CONTENTS

A New Microporous Surface Coating for Enhancement of Pool and Flow Boiling Heat Transfer

SEUNG M. YOU, KEVIN N. RAINEY AND
CURTT N. AMMERMAN

Robust Algorithms for Tomographic Reconstruction of Line-of-Sight Thermal Flow Field Images

KENNETH D. KIHM, KOJI OKAMOTO AND HAN SEO KO

Multiplicity and Stability of Convection in Curved Ducts: Review and Progress

LIQIU WANG AND TIANLIANG YANG

CONTRIBUTORS

Numbers in parentheses indicate the pages on which the author's contributions begin.

CURTT N. AMMERMAN (73), Los Alamos National Laboratory, P.O. Box 1663, MS H821, Los Alamos, NM 87545, USA

Y.I. CHO (1), Department of Mechanical Engineering and Mechanics, Drexel University, Philadelphia, PA 19104, USA

A.F. FRIDMAN (1), Department of Mechanical Engineering and Mechanics, Drexel University, Philadelphia, PA 19104, USA

KENNETH D. KIHM (143), Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843-3123, USA

W.T. KIM (1), Physical Water Treatment Center, Phoenixville, PA 19460, USA

HAN SEO KO (143), School of Mechanical Engineering, Sungkyunkwan University, Suwon Kyungyi-Do, Korea

S.H. LEE (1), Department of Mechanical Engineering and Mechanics, Drexel University, Philadelphia, PA 19104, USA

KOJI OKAMOTO (143), Department of Quantum Engineering & Systems Science, The University of Tokyo, Japan

KEVIN N. RAINEY (73), Los Alamos National Laboratory, P.O. Box 1663, MS-P940, Los Alamos, NM 87545, USA

LIQIU WANG (203), Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, P.R. of China

TIANLIANG YANG (203), Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, P.R. of China

SEUNG M. YOU (73), The University of Texas at Arlington, Department of Mechanical Aerospace Engineering, Arlington, TX 76019-0023, USA

PREFACE

For more than 40 years, the serial publication *Advances in Heat Transfer* has filled the information gap between regularly published journals and university-level textbooks. The series presents review articles on topics of current interest. Each contribution starts from widely understood principles and brings the reader up to the forefront of the topic being addressed. The favorable response by the international scientific and engineering community to the 38 volumes published to date is an indication of the success of our authors in fulfilling this purpose.

In recent years, the editors have published topical volumes dedicated to specific fields of endeavor. Examples of such volumes are volume 22 (Bioengineering Heat Transfer), volume 28 (Transport Phenomena in Materials Processing), and volume 29 (Heat Transfer in Nuclear Reactor Safety). The editors intend to continue publishing topical volumes as well as the traditional general volumes in the future. Volume 32, a cumulative author and subject index for the first 32 volumes, has become a valuable tool to search the series for contributions relevant to their current research interests.

The editorial board expresses its appreciation to the contributing authors of Volume 38 who have maintained the high standards associated with *Advances in Heat Transfer*. We also would like to welcome Professor Avram Bar-Cohen who joins the editorial board of *Advances in Heat Transfer* with the publication of this volume. Lastly, the editors would like to acknowledge the efforts of the staff at Academic Press and Elsevier Inc. who have maintained the attractive presentation of the volumes over the years.

Physical Water Treatment for Fouling Prevention in Heat Exchangers

Y.I. CHO[1], A.F. FRIDMAN[1], S.H. LEE[1], and W.T. KIM[2]

[1]*Department of Mechanical Engineering and Mechanics, Drexel University, Philadelphia, PA 19104, USA*
[2]*Physical Water Treatment Center, Phoenixville, PA 19460, USA*

I. Introduction

A. Background

Physical water treatment (PWT) is an attempt to treat hard water for the purpose of preventing or mitigating fouling using a physical means, without adding chemicals to water. PWT has a long history of applications. Early patents for PWT devices were issued almost 100 years ago [1]. Numerous field applications of various PWT devices have been made and resulted in both successful and unsuccessful case studies. After 100 years of field applications, scientists are still looking for the operating principle and mechanism of PWT, and there are several proposed mechanisms as summarized in several articles [2–11].

The goal of PWT technology is to prevent or mitigate scale build-up at heat exchangers. Figure 1 shows that scaling (or fouling) problems involve three different academic disciplines: physics, chemistry, and mechanical engineering (heat transfer). The fouling problem starts because of hard water being heated inside heat transfer equipment. The precipitation of dissolved mineral ions and subsequent scale deposit on the heat transfer surface critically depend on water chemistry. In addition to water chemistry, the deposit and removal rates of scale also depend on flow velocity, heat flux, and heat exchanger geometry. The PWT tools we attempt to use to solve the fouling problem are governed by various laws of physics. Hence, we need to include all three disciplines to study the PWT technology. In general, PWT (without a filtration system) is not a bona fide water-softener as it does not remove any mineral ions from water. Neither does PWT add anything to water. Given that PWT does not change anything in the structure of water molecules, one may not expect to detect any changes in

Advances in Heat Transfer
Volume 38 ISSN 0065-2717
DOI 10.1016/S0065-2717(04)38001-9

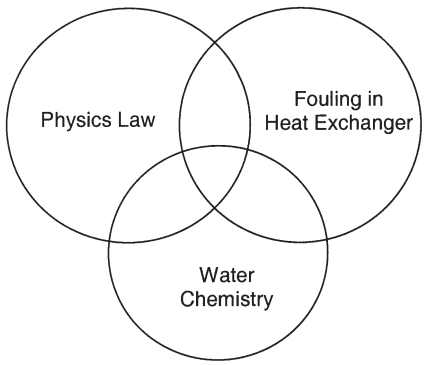


FIG. 1. Scaling problem involves three different disciplines.

water characteristics even with sophisticated analytical tools. If the benefit of PWT takes place at the heat transfer surface, one should carefully examine the fouling mechanism at heat exchangers.

The field of fouling is a unique area of research in mechanical engineering where a large number of papers and books have been published [12–18]. Billions of dollars are spent each year in the maintenance and replacement of heat transfer equipment that has been coated with various forms of scale through the fouling process [5]. A large amount of money and labor would be saved if these effects could be fully understood and controlled.

Let us briefly examine why the fouling problem exists. Due to the large amount of carbon dioxide released from various combustion sources, rain-water becomes acidic and absorbs mineral ions (mostly calcium) by dissolving limestone in subterranean aquifers [19]. These ions, when transported through piping in ordinary plumbing systems can form scaling. Scaling typically occurs when water with a high concentration of these ions is subjected to a change in temperature, which forces these ions out of solution. For every solute and solvent combination, there is a characteristic amount of solute that will dissolve under a given condition such as temperature, pressure, and pH. The solubility of a substance is the amount that will dissolve in a given solvent to produce a saturated solution [20]. Two forms of solubility are possible: normal and inverse. Some salts have greater solubility as the temperature is raised; these salts are called normal solubility salts, and examples include NaCl, $NaNO_3$, and silica. Other salts show less solubility as the temperature is raised, usually termed inverse solubility salts such as $CaCO_3$, $CaSO_4$, and $MgSiO_3$ [20]. Due to the heating (or cooling) process in thermal equipment, the temperature-dependent solubility leads to a supersaturated state. Calcium carbonate

has an ''inverse solubility'', which as a result, drives the ions out of solution at a high temperature. This often accounts for crystallization fouling in and around heat exchangers [13,15,18]. Thus, calcium carbonate scaling is one of the most common scale problems in many heat exchangers where water is used as a cooling medium.

The study on the effect of magnetic or electric fields on the fouling behavior of hard water is relatively new. Despite numerous positive results, however, there has been very little data that define the parameters that control the interactions between water and magnetic or electric fields. There is consequently much skepticism in the scientific community over the effects of these magnetic and/or electric devices, since a majority of the hard science has been replaced with claims based on inaccurate or misleading principles.

In order to understand the mechanism of PWT, we should find the passageway between the PWT and fouling mitigation. The key to the passage appears to be the phenomenon of bulk precipitation, a chemical process which may be affected by a number of parameters including the external applications of magnetic or electric fields.

B. Crystallization and Particulate Fouling

In order to explain the mechanism of PWT, we briefly review two types of fouling: crystallization and particulate fouling. Although PWT has been used for fouling mitigation in various heat exchangers, the highest consistency in successful applications of PWT have been reported in heat transfer equipment in cooling tower water loops, where the repetitive treatment of PWT in recirculating cooling systems may have contributed to success.

Circulating cooling tower water contains an excess amount of mineral ions such as calcium and magnesium due to the evaporation of water, making the circulating water hard even if the make-up water is relatively soft. When hard water is heated inside heat transfer equipment, the calcium and bicarbonate ions precipitate due to the changes in solubility, forming hard scale on heat-transfer surfaces and clogging heat exchanger pipes and manifolds. When any undesirable material deposits on a heat exchanger surface, it is called fouling [12–18].

When the circulating water is not treated, one often finds hardened scale depositing on the heat transfer surface (Fig. 2a). This is due to crystal formation on the heat-transfer surface, often known as crystallization fouling. Factors affecting nucleation and subsequent crystal formation are the concentration of fouling materials (foulants), temperature, pH, pressure, time, flow velocity, mechanical motions, radiation, and impurities.

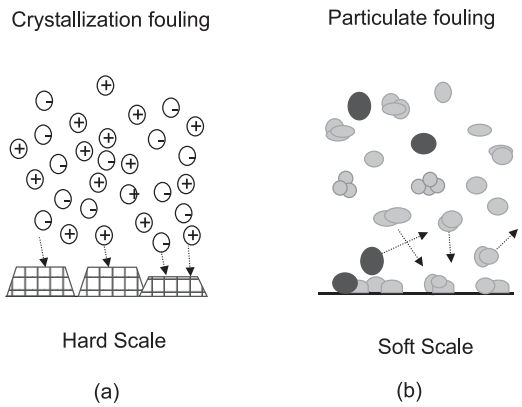


FIG. 2. Sketches of crystallization and particulate foulings.

Mechanical motions include shaking, mixing, and friction, which can be characterized as turbulence.

The initial crystals formed by a precipitation reaction may not be the most stable solids (i.e., thermodynamically stable phase) for reaction conditions. Over a period of time, however, the crystal structure changes to that of a stable phase. This change may be accompanied by additional precipitation from solution. A phenomenon, called ripening, may also take place whereby the crystal size of the precipitate increases [1]. Such hardened scale is common in heat transfer equipment using untreated water as a cooling medium. It is well known that such hardened scale cannot be removed by brush punching. Acid cleaning is often required to remove the hardened scale, which may cause premature equipment failure and replacement as well as producing chemical wastes for disposal.

Particulate fouling is a deposition process of particles carried by a flowing fluid as well as by matters generated in a solution. When compared with the scale produced in crystallization fouling, the scale produced in particulate fouling is much softer, as illustrated by Fig. 2b. The term ''particle'' is general and may refer to particulate matter, bacteria, corrosion products, and so on. The term ''particle'' also refers to particles that may be generated at the surface, such as the products of chemical reaction or crystallization [12]. The arrival of particles at a surface can take place by two mechanisms, i.e., gravitational settling or particle transport.

Beal [21] categorized the particulate fouling into three major processes: transport of the particles from the bulk fluid to the surface, the attachment of the particles to the surface, and re-entrainment of previously deposited particles from the surface back into the bulk fluid. Bott [12]

divided particle deposition into transport mechanism and agglomeration. First the particle has to be transported to the surface by one or a combination of mechanisms including Brownian motion, turbulent diffusion, or by virtue of the momentum possessed by the particle. Other factors that must be borne in mind in a consideration of the particle deposition include the agglomeration of particles in the bulk or at or near the surface, and the complex interactions of the forces near the wall.

It is of note that calcium bicarbonate is used to represent a group of mineral salts that deposit as foulants on the surface of a heat exchanger. Calcium carbonate is one of the most common scale types in cooling tower applications. The two structures of calcium carbonate crystal commonly found in nature are calcite and aragonite in morphology. They have the same chemical component, $CaCO_3$, but differ in many aspects. Calcite is formed at room temperature (i.e., below 30°C), easily removable with weak hydrochloric acid, is less adherent than aragonite, and has a hexagonal crystal shape with a specific gravity of 2.71 [1]. Aragonite is formed at high temperature (i.e., above 30°C) and is difficult to remove, having an orthorhombic crystal shape and a specific gravity of 2.94 [1]. Aragonite is a more troublesome form of calcium carbonate than calcite because it forms a harder and denser deposit than calcite in boilers and other heat transfer equipment [1]. It has been of interest to see whether calcium carbonate scales produced in water treated by a PWT device is calcite or aragonite.

C. Devices for Physical Water Treatment

In order to prevent or mitigate the fouling problem, various physical water treatment (PWT) methods have been applied using one of the following means: magnetic fields, electric fields, alteration of surface charges, mechanical disturbance such as ultrasound, vortex flows, and sudden pressure changes. Magnetic devices in particular have a long and controversial history in their efficacy and efficiency [1,2,10]. Numerous researchers investigated the feasibility of using a permanent magnet for reducing mineral fouling [22–31]. Some of them reported that magnetic treatment changed water properties and/or produced nucleation sites in a bulk solution, whereas others reported that it did not work.

Two most common types of permanent magnets are clamp-on magnets (see Fig. 3a) and magnets positioned at the center of a pipe (see Fig. 3b) such that water flows through an annulus gap between the pipe and magnets. Permanent magnets often have magnetic field strength ranges from 2000 to 6000 gauss (0.2–0.6 T). When charged molecules or ions pass

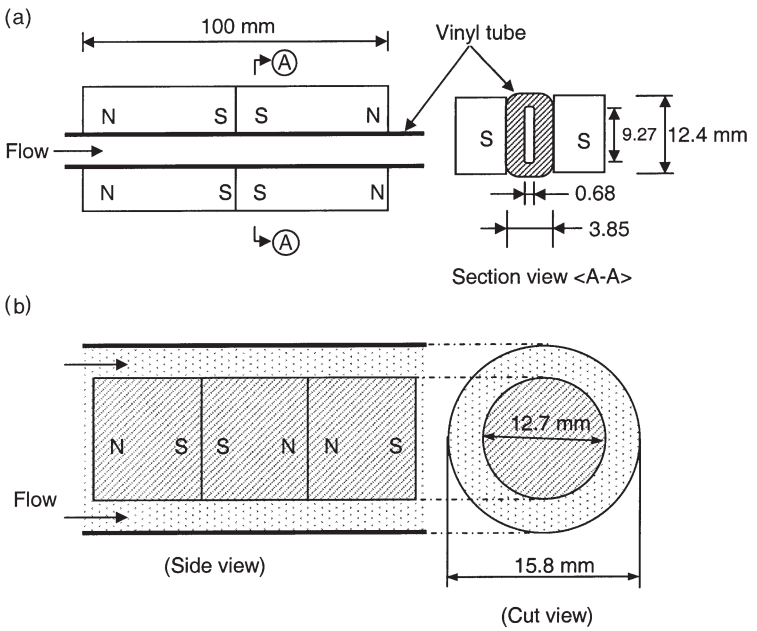


FIG. 3. Sketches of two common geometries of permanent magnets used for PWT. (a) PMDU (permanent magnet made at Drexel University), test results obtained with this magnet is identified as MWT(A) in Section 5. Total four permanent magnets were used, an arrangement that maximized the magnetic treatment on water. The maximum strength of the magnetic field was measured to be 0.16 T (or 1600 gauss) by using an Alphalab DC Magnetometer, from Salt Lake City, UT. (b) MWT(B), water moves through annulus gap between pipe wall and cylindrical magnets.

through a region under magnetic fields, electric fields are induced, which can be expressed as [32,33]:

$$E = V \times B \tag{1}$$

where E is an induced electric field [V/m], V is a flow velocity vector, and B is a magnetic field strength vector [Wb/m^2]. The magnitude of the electric field induced by permanent magnets at a water velocity of 3.0 m/s becomes (3 m/s) × (4000 gauss) = 1.2 V/m.

Figure 4 shows a sketch of a solenoid coil device, and a method to measure the induced electric field inside a tube. As shown in Fig. 4a, a solenoid coil was wrapped over a plastic tube, used only for forming the solenoid coil into a cylindrical shape; the plastic material has no affect on the electromagnetic field due to its non-ferrous composition.

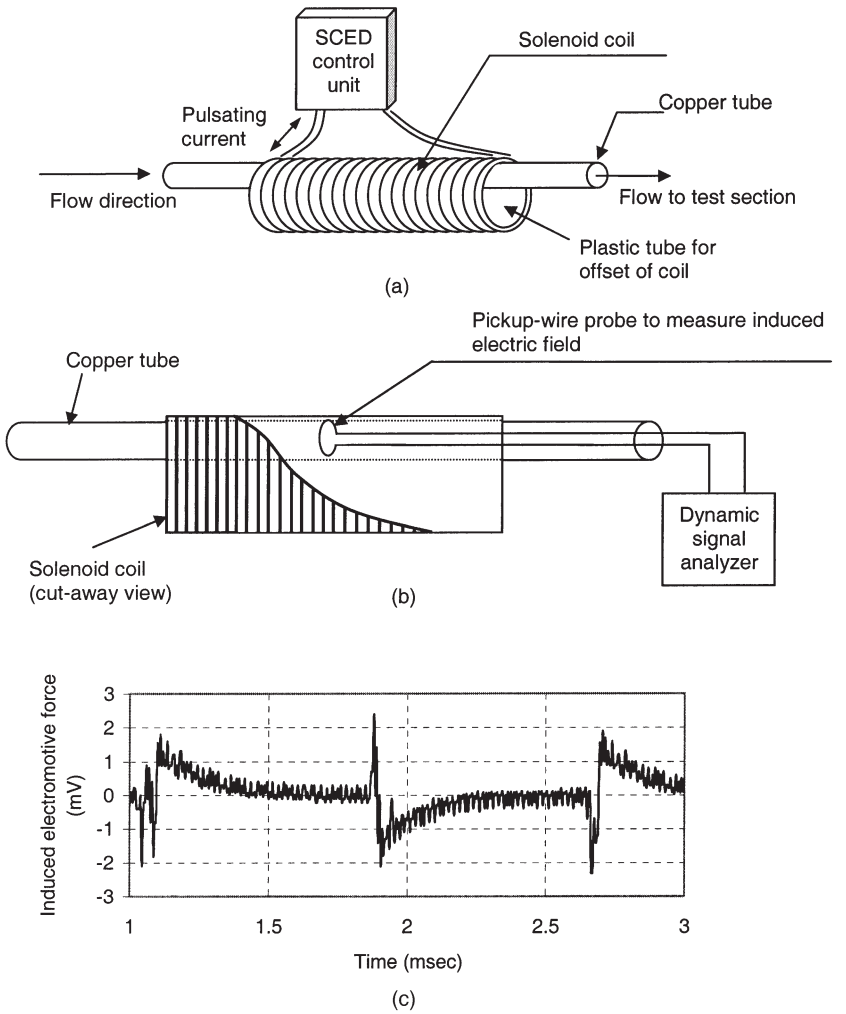


FIG. 4. (a) Installation of an EAF unit and a solenoid coil; The solenoid coil was wrapped over a plastic tube with an outside diameter (O.D.) of 50 mm. 14-Gauge wire was wound with 80 turns. Two ends of the solenoid coil were connected to a SCED-control unit. (b) Measurement of an induced electric field inside a copper tube; (c) Induced electromotive forces generated by a time-varying electric current.

Two ends of the solenoid coil were connected to an electronic control unit. The copper tube was located at the off-centered position relative to the solenoid coil since the strength of the induced electric field had a maximum value at the surface of the coil and a minimum value at the

center of the coil. The induced pulsating electric field was generated inside the pipe by Faraday's Law [32,33]:

$$\int E \cdot \mathrm{d}S = -\frac{\partial}{\partial t}\int B \cdot \mathrm{d}A \tag{2}$$

where E [V/m] is an induced electric field vector, S is a line vector along the circumferential direction, B [Wb/m^2] is a magnetic field strength vector, and A is the cross-sectional area of the solenoid coil.

Figure 4b shows a schematic diagram of a test set-up with a circular-pickup-wire probe to measure the induced electric field inside a copper tube. Figure 4c shows a typical output of the induced electric field measured at the axial center of the solenoid coil, showing the peak-to-peak voltage generated by a time-varying electric current inside the copper tube. The control unit was adjusted to produce a pulsing current at approximately 650 Hz. The magnitude of induced electric field by a solenoid coil via Faraday's Law was approximately 0.1–0.2 V/m.

Another type of PWT device uses a high-voltage electrode with a dielectric coating on the electrode surface positioned at the center of a pipe filled with water. The device works by transmitting an electric signal which varies with time. Ultrasonic waves were also used for the PWT as they influence crystallization by increasing the rates of both nucleation and growth [34]. The ultrasonic waves between 10 and 60 KHz are used to cause spontaneous nucleation, and nuclei are formed in the bulk solution, which form sludge of scale deposit removed from the system by blowdown.

D. The Hypothesis of Physical Water Treatment

Although the success may not be achieved consistently in all applications, one can make a statement that when a PWT device is properly used in a circulating cooling tower water application, the PWT produces bulk precipitation. The water treated by PWT devices tends to produce particles of $CaCO_3$ and other mineral salts in bulk water as external magnetic or electric fields precipitate mineral ions and form clusters (i.e., colloidal particles of submicron size [35]) in solution. The clusters grow in size as the solubility of the mineral ion drops due to increased temperature of the solution inside heat transfer equipment. These large particles adhere to the heat transfer surface as the adhesion probability increases with increasing particle size for particle diameter larger than 5 μm. In other words, PWT devices produce soft sludge coatings on the heat transfer surface. Subsequently, when the shear force generated by the

flow velocity in heat transfer equipment is sufficiently large to remove the soft sludge coating, then PWT device can be said to prevent new scale deposit or significantly mitigate the scale. In other words, when there is not sufficient shear force, PWT will not work or will be less effective. Thus, the question becomes how such a small energy or small electric fields of 0.1–1 V/m produced by a PWT device can cause the precipitation of calcium carbonate in bulk water. This is a key issue that will be addressed in the next section.

II. The Mechanism of Physical Water Treatment

A. Literature Reviews on the Mechanism of the PWT

The mechanism of physical water treatment (PWT) has been studied by a large number of researchers [2–6,9,34]. More recently, Parsons [10] summarized the most likely mechanism as (1) the effect of microcontaminants, (2) interfacial phenomena, or (3) surface/interfacial phenomenon. When the mechanism of PWT was discussed in the past, the effect of magnetic fields on $CaCO_3$ scale prevention was most often reported simply because permanent magnets were the oldest embodiments of PWT, and $CaCO_3$ is the most common form of scale in many heat transfer equipment using water as a cooling medium.

In addition to permanent magnets, there have been other types of PWT devices that mitigate scaling in a heat exchanger, including solenoid coils, electrostatic devices, catalytic surfaces, ultrasound and devices using sudden pressure changes, turbulence, and vortex flow. If one or more of these PWT devices, properly applied, were proven to prevent or mitigate fouling in some specific types of heat exchangers, there would also have to be clear physical and chemical reasoning behind the success. The mechanism explaining how the permanent magnet works should, in principle, be able to connect with the mechanisms of other non-magnetic types of PWT devices from a physicochemical standpoint. On the other hand, if the PWT does not work in certain types of heat exchangers, it would still be very useful to identify why the PWT does not work so that the proposed mechanism can be tested in those cases where PWT does not work.

Intuitively, one can consider that the magnetic field may not directly affect water or dissolved mineral ions in water for the purpose of mitigation of scale in heat exchangers. There is a possibility that the mechanism of magnetic water treatment may be different from that of non-magnetic devices. As an example, catalytic surfaces, which are similar to the materials

commonly used for heterogeneous catalysis, have been recently used as a PWT device for the purpose of fouling mitigation. This class of PWT devices does not use either magnetic or electric field to mitigate the scaling in heat exchangers. Thus, when one considers the mechanism of PWT, one should not limit to the PWT using magnetic and electric fields only. In addition, the fact that a PWT device does not work consistently is closely related to the lack of understanding of the fundamentals of PWT. Once the fundamentals are fully understood, the conditions under which the PWT device consistently works or does not work can be identified.

It is often hypothesized that the permanent magnet modifies solution properties, that is, the precipitation of mineral ions in a bulk solution and the alteration of water properties, when mineral ions and molecules are exposed to the alternating magnetic fields or induced electric fields produced by magnetic fields. It was suggested that magnetically treated water produced different types of scale formation due to the modification of solution properties [28,29]. Kronenberg [28] reported that calcium carbonate from magnetically treated water formed a soft sludge instead of hard lime scale clinging to the surface. Donaldson and Grimes [36] speculated that magnetic fields acted at the surface of crystallites, modifying the nature of the charges at the surface. Donaldson [37] explained the mechanism of magnetic water treatment in terms of modification of crystal nucleation, i.e., the magnetic field could modify the nature of charges at the particle surface and subsequently modify the growth of crystals. He attributed to the Lorentz force and the flow rate of fluid the production of energy to remove or prevent scale formation. Parsons *et al.* [3] reported that physical water treatment led to the formation of irregularly shaped crystal forms.

It has been suggested in general that both the magnetic and electric fields affect the characteristics for the nucleation of mineral ions and other electrically charged sub-micron particles [38]. The influence that an externally applied electro-magnetic field may exert on nucleation and particle behavior was studied from the point of view of fouling mitigation. One of the most common statements on the fouling mitigation by external electro-magnetic fields is a generation of nucleation in a bulk solution. Based on Russian scientists' research, Al-Qahtani [39] reported that the magnetic treatment processes accelerated the coagulation–flocculation of solid particles suspended in water and increased the crystal formation on the bulk solution instead of deposition on heat-transfer surfaces.

Söhnel and Mullin [40] estimated the homogeneous nucleation using the supersaturation ratio and found that for most PWT devices the rate of

nucleation was negligibly small. They concluded that the direct influence of a magnetic field on the process of crystallization of $CaCO_3$ was not very significant. They ruled out bulk precipitation as a result of direct influence of magnetic or electric fields, as well as hydrodynamic cavitation.

Based on the amount of scientific evidence available, the interaction between magnetic/electric fields and fluids warrants a closer look. There are a number of questions and applications within PWT technology that require clarification. There is also a definite need for a genuine scientific understanding of the mechanisms that lead to fouling or prevention of fouling in order to remove the inconsistencies and ineffectiveness from existing PWT applications.

The following questions remain to be answered:

- Can a magnetic/electric field directly influence a reaction leading to the precipitation of $CaCO_3$?
- Can a magnetic field rotate and/or orient an ion or molecule to stimulate chemical reactions, leading to the precipitation of $CaCO_3$?
- How is the effectiveness of PWT related to the velocity of the fluid?
- How long does the effect of PWT last?
- Can PWT alter the properties of water itself or does it just change water composition?

Generally, this section will address the first three questions, leaving the last two questions to be studied in greater depth once the mechanism is clearly defined. This section presents an introductory explanation of a few mechanisms under which charged particles in a fluid medium would be manipulated by electromagnetic forces. In this case, the fluid we are considering is water, and the charged particles are dissolved mineral ions in water that eventually cause the fouling in heat exchangers if left untreated.

B. Background: Water Chemistry of Precipitation of $CaCO_3$

The chemistry behind the fouling mechanism is described next. The main reaction in mineral fouling in cooling water applications is the precipitation of calcium carbonate [20]

$$Ca^{2+}_{(aq)} + CO^{2-}_{3\,(aq)} \Leftrightarrow CaCO_3 \tag{3}$$

and

$$CaCO_{3(s)} + H^+ \Leftrightarrow Ca^{2+}_{(aq)} + HCO^{-}_{3\,(aq)} \tag{4}$$

In a saturated condition, the forward reaction of the precipitation of $CaCO_3$ does not take place as both calcium and bicarbonate ions are surrounded (i.e., hydrated) by water molecules. When water is supersaturated (i.e., hard water) and there is sufficient energy added to water as manifested by increased temperature inside a heat exchanger, the water molecules are disturbed or become freed from these calcium and bicarbonate ions, resulting in the precipitation of $CaCO_3$, a phenomenon that occurs often inside a heat exchanger. Equation (4) shows the dissolution of solid calcium carbonate by acid, a process that takes place during acid cleaning of scaled heat exchangers.

Generally, we are concerned with the two reactions, Eqs. (3) and (4). In reality, precipitation and dissociation reactions are much more complicated. There are actually three reactions that control the rate at which dissolved calcium and carbonate ions recombine and crystallize. Reaction 1 relates the dissociation of bicarbonate ions into the hydroxide ions and carbon dioxide (in aqueous solution).

$$HCO^-_{3\,(aq)} \Leftrightarrow OH^-_{(aq)} + CO_{2(aq)} \quad \text{(Reaction 1)} \tag{5}$$

The forward reaction indicates the dissociation of the bicarbonate ions, which is the most critical step for PWT theory. Of note that the bicarbonate ions do not cause any harm in terms of scaling as long as they remain as bicarbonate ions. On the other hand, carbonate ions are a dangerous species insofar as they spontaneously form scaling on the surface of heat transfer equipment. Reaction 1 shows the first step in the conversion of bicarbonate to carbonate ions. The presence of hydroxide ions is best indicated by a local increase in pH, and carbon dioxide typically evolves from the water as a gas over time.

The pH value of cooling water available in open cooling towers is often greater than 8.0 as CO_2 escapes from water over time in a circulating open cooling tower system. Thus excess hydroxide ions go to the pipe wall, making pipe walls and heat transfer surfaces negatively charged. The positively charged calcium ions are attracted to negatively charged heat transfer surfaces where calcium ions form the ''electric double layer'' near the negatively charged surface. This layer is 10–30 Å thick (i.e., 1–3 nm) for a concentrated electrolyte solution and is about 10 nm for a dilute solution like water [41,42].

In Reaction 2 hydroxide ions produced from Reaction 1 further react with existing bicarbonate ions, producing carbonate ions and water.

$$OH^-_{(aq)} + HCO^-_{3\,(aq)} \Leftrightarrow CO^{2-}_{3\,(aq)} + H_2O \quad \text{(Reaction 2)} \tag{6}$$

Reaction 3 is the reaction between calcium and carbonate ions, resulting in the precipitation and crystallization of calcium carbonate.

$$Ca^{2+}_{(aq)} + CO^{2-}_{3\,(aq)} \Leftrightarrow CaCO_{3(s)} \quad \text{(Reaction 3)} \tag{7}$$

At this point, carbonate ions combine with calcium ions that have been present in solution throughout the process. Based on the Gibbs free energy value for this reaction (−47.7 kJ/mol), calcium carbonate will tend to form spontaneously once the initial activation energy is overcome [20]. Reaction 3 often occurs at the heat exchanger surface, since the positively charged calcium ions are located at the electric double layer at this point. In the case of PWT, this nucleation reaction takes place in the bulk instead of on the heat exchanger surface, a phenomenon which would result in bulk precipitation leading to particulate fouling [9,43,44]. Thus, it is essential that Reaction 1 occurs in a pipe prior to the heat exchanger such that carbonate ions become available in bulk water for Reaction 3 before water enters the heat exchanger.

Table I presents the thermochemistry of the above three reactions [20]. The values of ΔH and ΔG give some useful insights into the behavior of the system of reactions. The endothermic Reaction 1 is the most rate-limiting, since it needs a relatively large input of energy to continue in the forward direction based on the high enthalpy (ΔH) value. Since the Gibbs free energy (ΔG) is relatively high, this reaction will tend to form bicarbonate (HCO_3^-), unless this energy restriction is overcome.

Once the energy input for Reaction 1 is overcome, Reaction 2 is free to proceed. Reaction 3 will proceed when a smaller amount of energy is supplied, like that from a heat exchanger. The three chemical reactions above form the mechanism for calcium carbonate fouling involving water inside a heat exchanger. Keeping this mechanism in mind, three possible methods by which PWT can affect this fouling mechanism have been studied. The following is a description of each of the three models regarding

TABLE I
THERMOCHEMICAL DATA

	Reaction 1	Reaction 2	Reaction 3
ΔH	48.26 (endothermic)	−41.06 (exothermic)	12.36 (endothermic)
ΔG	43.6 (spontaneous backward)	−20.9 (spontaneous forward)	−47.7 (spontaneous forward)

All values are in kJ/mol [20].

the physical effects of PWT on the fouling process in water. The goal of the study was to investigate (1) whether or not PWT [i.e., electric field (E) or magnetic field (B)] can directly affect the three chemical reactions via the orientation of molecules, (2) whether or not PWT can produce the Hall Effect for the separation of ions, and (3) whether or not PWT can lead to heterogeneous catalysis at the surface to provide the local energy that is necessary to overcome the energy value that is required to initiate the dissociation of the bicarbonate, i.e., Reaction 1.

C. The Direct Effect of Electric and Magnetic Fields on the Potential Energy at Surface and Chemical Reaction Rate

Dissolved calcium and bicarbonate ions do not react at normal temperature as both ions are surrounded by water molecules, which can be disrupted by the addition of external energy for the reaction. Thus, if a magnetic or electric field can preferentially orient the water molecules surrounding the calcium and bicarbonate ions, PWT can precipitate them into calcium carbonate particles. We call this effect a direct effect of the magnetic or electric field.

Let us assume that a direct electromagnetic force could produce an influence on dissolved ions (such as bicarbonate, HCO_s^-), a process that might causes a slightly preferred orientation so that the reaction of two bicarbonate ions producing carbonate ions may take place. In other words two bicarbonate ions collide in a specific orientation such that the oxygen with an open end in one bicarbonate ion meets with the hydrogen from the other bicarbonate ion. This preferred orientation change was expected to lead to enhanced molecular interactions [20].

$$HCO_3^- + HCO_3^- \longrightarrow CO_3^{2-} + H_2O + CO_2(\text{gas}) \tag{8}$$

Based on a typical potential of 10 V [32,33,41,42], the energy necessary to orient an ion with the size of 10^{-10} m (i.e., 1 Å) would be 10 V/(10^{-10} m) = 10^{11} (V/m), which is much too large compared with the electric field (i.e., 0.1–1 V/m) produced by PWT. This large value tells us that a direct affect on this molecule by an electromagnetic field is not possible. There would have to be a field-strength considerably higher than we are seeing in order to produce any observable direct effect.

Laying the possibility of directly affecting a single molecule aside, we can look at the interactions of two or more molecules, in particular Reaction 1, which is rate-limiting. Reaction 1 is endothermic ($\Delta H = 48.26$ kJ/mol) and will need to be stimulated in order for it to proceed. The ability of a direct magnetic or electric field to produce the desired

energy for activation is investigated. In order for hydroxide ions to be produced as a result of the dissociation of a bicarbonate ion, the energy required for the reaction can be calculated as

$$\frac{48{,}260\ \text{J}}{\text{mol}}\frac{1\ \text{mol}}{6.02\times10^{23}\ \text{ions}} = 0.801\times10^{-19}\ \text{J} \quad (\text{or} \sim 0.5\ \text{eV}) \tag{9}$$

If the dissociation reaction occurred, the corresponding electric field necessary to induce the electron to move a distance of 1 Å [41,42] required by the reaction would have to be 0.5 V/(10^{-10} m) = 5×10^{9} (V/m). Energy requirements are much too large to induce changes directly by a magnetic or electric field in the bulk liquid. Based on this data, we conclude that the direct effect of the magnetic or electric field on fluids would only be capable of influencing a reaction at field strengths that are much higher than what is possible with current PWT technology.

Along this line of thought, the zeta potential has been mentioned to describe the mechanism of PWT devices. Zeta potential is defined as the potential between the charge surface and the electrolyte solution at a shear surface (i.e., much smaller than the Debye length) [41,42]. Several researchers [2,4,6,7,45] indicated that the magnetic fields could affect the zeta potentials or the distance of the electric double layer. Gehr *et al.* [4] reported an approximate 23% drop in zeta potential in magnetically treated water. They rotated a tube containing the water sample in a magnetic field produced by a MRI spectrometer (with a magnetic field strength of 4.75 T), which was significantly larger than most PWT devices using permanent magnets (i.e., 0.2 T). They suggested that the effect of magnetic treatment indirectly increased the coagulating tendency of the suspension. Gamayunov [6] suggested a possible action of the magnetic effect on the electric double layer. He proposed that Lorentz forces displaced co- and counter-ions, resulting in changes in the charge distribution within the electric double layer, resulting in a semi-permanent change in the charge and potential at the boundary of the Stern layer [41,42].

In order for a magnetic or electric field to directly affect the electric double layer near a negatively charged surface, one needs a magnetic or electric field that provides force, that is, as big as the force in the electric double layer. Typically, there are approximately two electron volts across an electric double layer, and the Debye length for a dilute solution like water is about 10 nm (i.e., 100 Å) [41,42]. Thus, the electric field required to directly affect the electric double layer becomes approximately 2 V/10 nm = 2×10^{8} (V/m). Thus, current PWT devices cannot directly affect or disrupt the electric double layer. If we use the zeta potential measured in a negatively charged surface of suspension in aqueous

solution, which is in a range of 10–20 mV [46], the required electric field to disrupt the electric double layer becomes 10 mV/10 nm $= 1 \times 10^6$ (V/m). Still the required electric field is by several orders of magnitude greater than what is produced by any PWT device (i.e., 0.1$\sim$1 V/m).

A number of researchers postulate that the magnetic or electric field somehow disrupts the electric double layer surrounding suspended sub-micron size colloidal particles which are negatively charged in a cooling tower water application. The underlying assumption is that circulating cooling water contains enough colloidal particles as seeds but these particles are surrounded by calcium ions, preventing them from acting as seeds for particulate fouling. Note that the colloidal particles are also negatively charged in water, and there is an electrical double layer at each particle–solution interface [42]. As one can see from the above estimation, the PWT device cannot produce enough force to directly disrupt the electric double layer. Thus, we need to find out a different mechanism for bulk precipitation.

Furthermore, it has been proposed that the magnetic or electric field generated by the PWT devices breaks the hydrogen bonds in water molecules, affecting the average cluster size in water. The energy required to break the hydrogen bonds can be estimated in a similar manner. In order to break the bond, we need energy of at least 0.4 eV over 2-Å distance (i.e., the distance between oxygen atom in one water molecule and hydrogen atom in adjacent water molecule). Note that the energy required to break the hydrogen bond is about 40 times stronger than the energy required for rotation or orientation of the molecule. Thus, we calculate an electric field required to break the hydrogen bond as 2×10^9 V/m. Thus, one can conclude that we cannot break the hydrogen bonds using PWT devices to make more free water molecules.

Note that the typical electric field involved in the rotation of a large molecule (i.e., dipole momentum) is about 0.01–0.03 eV over a distance of 10 Å [41,42]. This results in an electric field strength of 10^7 V/m or 100 KV/cm. That is why in the field of plasma science, the breakdown electric field of 30 KV/cm is used to generate electric discharge in air.

Direct influences of a magnetic or electric field do not appear to be possible. In general, if the direct influence is possible, we need the energy source of 1 eV over one angstrom, which gives 10^{10} V/m. In terms of the permanent magnet strength, one can briefly estimate the required magnetic field strength as follows:

$$\mathrm{eV}B = \mathrm{e}E \tag{10a}$$

$$B \text{ (required magnetic field strength)} = E/V \tag{10b}$$

If we use the water velocity of 1 m/s for V, we estimate the required permanent magnet strength of 10^{10} T. If we use an electron velocity of 10^6 m/s, we still get a magnetic field strength of 10^4 T. Considering that the present permanent magnet used for PWT generally has the value of 0.2–0.5 T, we need to conclude that there is no direct effect by the magnetic field.

D. Separation of Charges (The Hall Effect)

Another possible mechanism of PWT was a separation of charged particles that would produce high concentrations of calcium and bicarbonate ions locally, producing dissociation of bicarbonate ions (Reaction 1) and resulting in the precipitation of calcium carbonate. We examined whether or not PWT might be producing a sufficient drifting force that separates ions based on charges. This driving force is caused by a physical process known as the Hall Effect. The Hall Effect states that charged particles in an applied magnetic field drift in different directions based on the sign of their charges [32,33]. The positively charged particles would be driven in one direction, while the negatively charged particles would be driven in the opposite direction, resulting in a separation of charges.

This separation needs to be large enough to keep a number of oppositely charged particles at different areas inside the pipe. The number of charges in water is on the order of 10^{28} molecules per m^3. Assuming the calcium hardness of 200 ppm (i.e., 2×10^{-4}), the number of charges that are necessary to be affected by a magnetic or electric field should be on the order of 2×10^{24} molecules per m^3. This was compared with the maximum charge density that can be taken out by a reasonable magnetic field (0.3 T). Based on a typical set of conditions, the number of charges per unit volume (number density of charges; n) was calculated using Gauss' Law as

$$n = \frac{VB\varepsilon_0}{eL}, \tag{11}$$

where B is the magnetic field (0.3 T), ε_0 is the permittivity constant (8.85×10^{-12} F/m), L is the distance between two electrodes (~ 0.01 m), V is the velocity of the fluid (~ 1.5 m/s), and e is the charge of a single ion (1.6×10^{-19} C) [32,33]. Once the values were plugged in, it became apparent that the maximum number density affected by a magnetic field is approximately 2.5×10^9 per m^3, which is much less than the number density that we had determined was present (approximately 10^{24} per m^3). In other words, the permanent magnet cannot separate charged particles in water over the distance between the electrodes.

If one considers a separation over a short distance (i.e., Debye radius), for example, 10 interatomic distance (i.e., 10 $\mathring{A} = 10^{-9}$ m) [41,42], then the number density that can be separated by the magnetic field of 0.3 T becomes 3×10^{15} per m^3. This number is still many orders of magnitude smaller than 10^{24} per m^3.

In summary, it is not possible to keep the charges separated at any meaningful distance using the Hall Effect. Therefore, we conclude that the applied magnetic or electric field from PWT devices cannot separate calcium and bicarbonate ions in water and the Hall Effect is not the mechanism of PWT.

E. PWT Works at Surface via Heterogeneous Catalysis

The critical step for the bulk precipitation of calcium carbonate in water treated by PWT is the dissociation of the bicarbonate ions, Reaction 1. In this section we examine the possibility of applying the concept of heterogeneous catalysis to explain this dissociation reaction. The energy required to affect Reaction 1 will be compared with the energy that is attainable on a small scale as a result of electrostatic effects of surface irregularities (similar to the heterogeneous electrocatalysis). In heterogeneous catalysis, the reaction rate is increased without the catalyst being consumed by the reaction. In this case, we assume that the catalytic surface is the pipe wall. In this approach, ions are first diffused from the bulk water to the pipe wall, i.e., to the irregularities in the form of bumps at the wall of the pipe.

Consider a very small local bump, which serves as an electrode. The size of the local roughness ranges from 0.1 mm for a galvanized iron pipe to 0.0015 mm for a drawn tubing [47]. The electric field around these local bumps (E_L) can be much larger than the space-averaged field within the pipe (E_0), which will be explained in detail next. Figure 5 illustrates the geometric scales between the pipe radius and the radius of the small local bump.

The average electric field produced by electrostatic interaction of charged particles at the surface is described as the charge of a particle divided by its radius of curvature squared:

$$E_0 \propto \frac{q}{R^2} \tag{12}$$

In the case of an infinitesimal bump irregularity of radius r, the local electric field produced would be:

$$E_L \propto \frac{q}{Rr} \tag{13}$$

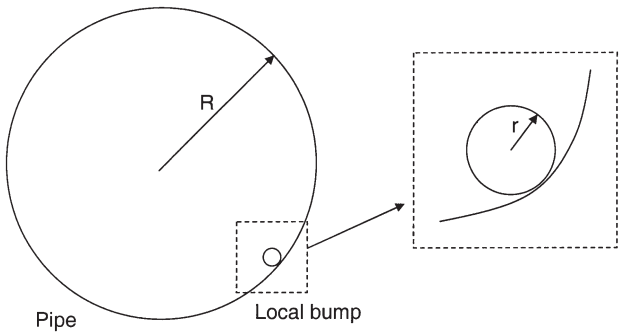


FIG. 5. Radius of curvature of the surface irregularity (r) vs. the curvature of the pipe surface (R).

Dividing Eq. (13) by Eq. (12), we get:

$$\frac{E_L}{E_0} = \frac{R}{r} \tag{14}$$

The surface radius of curvature (R) in a pipe is on the order of 10 cm, and the radius of curvature of the irregularity (r) is on the order of 10^{-7} m [47]. Thus, the electric field local to the bump irregularity is on the order of

$$\frac{E_L}{E_0} = \frac{10^{-1}\ \text{m}}{10^{-7}\ \text{m}} = 10^6 \tag{15}$$

The local electric field (at the surface) is 10^6 times greater than that in the bulk solution. Thus, despite the fact that the space-averaged electric field (E_0) typically produced by a PWT device in the bulk is much too small (~ 1 V/m), the corresponding strength of the electric field local to the surface irregularities becomes

$$E_L = E_0 \times 10^6 = 1 \times 10^6 \frac{\text{V}}{\text{m}} \tag{16}$$

This value is of the same order of magnitude necessary to drive these reactions in a region local to these surface irregularities.

Knowing the strength of an induced electric field local to the pipe wall ($E_L = 1 \times 10^6$ V/m), we can take a careful look at a single positive charge (i.e., calcium ion) located at the surface of the wall. The calcium ion acts as a charge center, accelerating the dissociation process of bicarbonate ions, producing hydroxide ions OH^- and CO_2(gas) as the

local electrostatic effect is strong enough to induce the rate-limiting (Reaction 1). The OH^- ions are collected on the surface first. When enough OH^- ions are collected on the surface, Reactions 2 and 3 initiate, allowing carbonate ions to be formed (Reaction 2) which proceed to induce the formation or precipitation of calcium carbonate (Reaction 3). The heterogeneous electrostatic catalysis process of PWT can generally be described as follows:

Step 1: Local roughness produces catalytic centers.
Step 2: Large local E field at surface.
Step 3: Reaction 1 is catalyzed, producing OH^- ions at surface.
Step 4: Reactions 2 and 3 follow in bulk as sufficient amounts of OH^- ions are collected, producing carbonate ions which subsequently combine with calcium ion.

Copper, which is in most common household piping, is used in industry to facilitate reactions through heterogeneous catalysis [48]. For example, it is used industrially as a catalyst in the synthesis of methanol [48]. These catalytic effects are closely related to the surface characteristics of the copper. It is therefore not unreasonable to suspect that a similar mechanism might be occurring at the copper pipe walls in heat transfer and water treatment application. Of note is that the above-mentioned heterogeneous catalytic effect of PWT can take place in both metal and non-metal surfaces.

There is another physical principle that can relate the use of magnetic field to stimulate heterogeneous catalysis and accelerate some chemical reactions in hard water. As the strength of the magnetic field is not strong enough to cause any direct energy-related effects on water properties or water chemistry, permanent magnets might produce charge centers for heterogeneous catalysis. In the case of the heterogeneous initiation reaction of the dissociation of bicarbonate ions, Reaction 1, there is a certain energy barrier that has to be overcome in order for the reaction to take place. In chemical processes, some transitions are permitted whereas other transitions are prohibited, a phenomena known as Pauli's Exclusion Principle [32,33,48]. The permanent magnet splits this energy barrier into several small steps, a phenomenon known in physics as the Zeeman Kinetic Effect [32,33,48]. This is the theoretical basis of the magnetic field effects on water. The Zeeman Kinetic Effect splits energy barriers, simplifies some quantum-mechanical selection rules, and multiplies chemical reaction paths. The Zeeman Kinetic Effect makes more permitted transitions available and thus can stimulate the process of bicarbonate dissociation (Fig. 6). In summary, the magnetic fields can influence the heterogeneous catalysis via the Zeeman Kinetic Effect. Subsequently, the heterogeneous

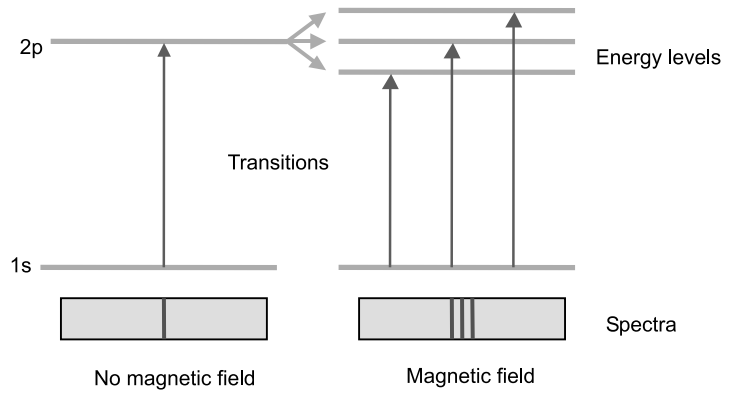


FIG. 6. Zeeman Effect.

catalysis accelerates some chemical reactions in hard water, which result in the precipitation of calcium carbonate as seed particles in bulk water as described earlier.

F. BULK PRECIPITATION

In this section we examined three possible mechanisms for physical treatment of water, which are direct affects of magnetism, charge separation from drift of charges in opposing directions, and heterogeneous catalysis based on surface effects. The first two models were proven to be ineffective since the magnetic field applied would have to be much stronger to see any effects. For the direct effect, the energy that can be provided by a PWT device simply cannot affect the structure or orientation of molecules, and for the mechanism of drifting charges, the field strength is not great enough to keep the charges apart. In order for a strong field to occur in a local space it is proposed that the reactions are the result of heterogeneous catalysis that is caused by irregularities at the surface of the pipe wall.

We propose that the precursor for PWT is the hydroxide ion, and the end result is the bulk precipitation of calcium carbonate, acting as seeds for the particulate fouling inside a heat exchanger. Subsequently, when the soft sludge coating generated by particulate fouling can be removed by the shear force of flow, then fouling can be prevented or mitigated by PWT. If there is not sufficient shear force inside heat transfer equipment, PWT will not work. Note that as a result of the dissociation of bicarbonate ions and the formation of hydroxide ions at the surface, one

can expect that the value of pH should increase locally at the surface of the pipe wall [8,30].

III. Effect of Particles in Water on Fouling Mitigation

The mechanism of PWT is postulated as bulk precipitation, which causes fouling mitigation via particulate fouling. The feasibility of adding small particles to water to prevent or mitigate fouling in a heat exchanger has been reported in the literature. This is a mechanical method of preventing mineral scaling, an alternative to chemical treatment.

Troup and Richardson [34] reviewed both chemical and physical methods of water treatment. They mentioned that the effect of solid material suspended in a supersaturated solution was to shorten the nucleation time. They noted that even dust particles from air could induce nucleation in water. The concept of adding solid particles for the prevention of scale was known in commercial heat exchanger operation, as ''seeding''. Badger *et al.* [49] and Charuit *et al.* [50] reported that fouling could be reduced by simply feeding fine $CaCO_3$ or $CaSO_4$ particles in industrial heat exchangers. Alternatively, Springer [51] and Aslan [52] reported that colloidal particles of $Al(OH)_3$ were introduced in water by passing an alternating current between two aluminum electrodes so that the $Al(OH)_3$ particles could act as seeds or secondary nucleation sites, thus reducing the scale formation on the heat exchanger. They reported that the precipitating effect of the $Al(OH)_3$ particles produced sludge of scaling deposit easily removable by blowdown. Similarly, Boute *et al.* [53] used the corrosion concept of mild steel by passing a small current between two mild steel electrodes to introduce colloidal particles of $Fe(OH)_2$ to water for the purpose of scale prevention. They also observed a sludge coating of scale which could be removed in the blowdown.

Middis *et al.* [54] reported that the addition of small amounts of fiber to a calcium sulfate solution caused a marked reduction in the rate of growth of the fouling deposition. They attributed this reduction to the fact that fibers reduced the rate of deposition by blocking the reactants from reaching the heat transfer surface and by continuous physical collision with the heat transfer surface. Later, Kazi *et al.* [55] studied the effect of wood pulp fibers that were added to supersaturated calcium sulfate solutions at various concentrations to mitigate fouling. Average length and diameter of fibers were 2.5 mm and 30 microns. Heat transfer enhancement above the solution alone was observed initially with the fibers and the onset of fouling delayed (see Fig. 7). When fouling eventually developed, the final asymptotic level was lower than the fiber-free case for the

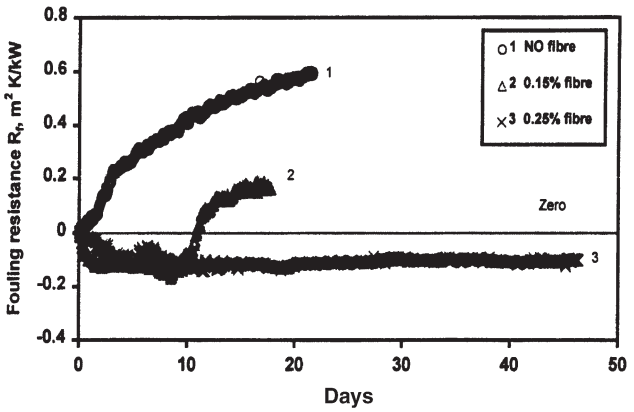


FIG. 7. Fouling resistance vs. time for the effect of the addition of particles in calcium sulfate solution on fouling mitigation [55].

experimental conditions specified. At a fiber concentration of 0.15%, heat transfer augmentation occurred for 11 days. At 0.25% fiber concentration, heat transfer augmentation (i.e., no fouling) was sustained over the experimental duration of 45 days. Thus, they concluded that the service-life cycle of a heat exchanger could be prolonged with the addition of asymmetric, flexible, natural fibers.

Kazi *et al.* [55] explained that fibers modified the onset of deposition by boundary layer scavenging and interacted with turbulent eddies to reduce the rate of mass transfer of the foulant to the heated surface. When scale formed, the crystalline structure of the scale was interrupted by the fibers, which appeared to roughen the heat transfer surface initially and increased the heat transfer coefficient. However, the scale deposit continued to build up very slowly, causing the thermal resistance to eventually override the turbulence augmented heat transfer effect of the fibers, as the thermal conductivity of the fibers was poor. However, their study indicated that any suspended particles could help mitigate crystallization fouling on the heat transfer surface.

Cowan and Weintritt [1,56] reported that the degree of supersaturation was an important factor in the formation of adherent or non-adherent scale. Note that the adherent and non-adherent scales meant crystallization and particulate fouling, respectively. For example, calcium sulfate, calcium carbonate, and barium sulfate would precipitate at a low degree of supersaturation but formed adherent large crystals on the heat transfer surface. At a high degree of supersaturation, all precipitated but formed non-adherent small crystals. For example, for calcium carbonate, which is the

most important scale composition for cooling water scaling, a calcium hardness of approximately 550 ppm (i.e., corresponding electric conductivity of 2000 μmhos/cm) fell within the range of the low degree of supersaturation. Thus, at this concentration, calcium carbonate scale became adherent on the heat transfer surface. When the concentration increased to 1100 ppm calcium hardness, the degree of supersaturation was high, and scale became non-adherent or less-adherent [1]. Cowan and Weintritt [1, p. 213] suggested that precipitates could be tolerated as long as they were non-adherent. However, in spite of this beneficial effect of the high degree of supersaturation, we often do not use it in practice because of the uncertainty involving the high degree of supersaturation.

Cowan and Weintritt [1, p. 387] also reported that numerous attempts have been made by seeding a slurry of seed crystals in water to provide a more extensive and favorable surface for deposition to occur on the heat transfer surface. Ideally, seeds of the same composition as the scaling-forming constituents could be used for the maximum benefit. We believe that this is exactly what is happening with PWT as calcium carbonate precipitates in bulk water prior to the treated water entering a heat exchanger.

The observation reported by Cowan and Weintritt [1] provides an insight in understanding the mechanism of PWT. We postulated earlier that PWT produced seed crystals in water in the form of fine calcium carbonate particles. These fine particles could be generated by PWT, assisted by the high degree of supersaturation. What is important is that there are fine particles in bulk water. As the supersaturated water is heated inside a heat exchanger, the solubility decreases according to the inverse solubility law of the calcium carbonate, forcing other dissolved calcium ions to come out of water. At this point, calcium ions adhere to the surface of the fine particles in bulk water as well as to the heat transfer surface, making the fine particles grow in size. As the particles grow to a certain size (i.e., greater than 10 microns), the probability of adhesion becomes large [12]. However, the deposit from the particulate fouling tends to be a soft-sludge coating, which can often be removed by the shear force produced by flow velocity. That is why water with a high degree of supersaturation produces precipitations that are non-adherent. More discussion on this issue will be made when the effect of concentration on PWT is presented later.

IV. Validation Methods for the Efficiency of PWT

This section addresses issues involved in the validation of PWT and review the past attempts to validate PWT. The successful results from

PWT are often reported in the field in terms of scale prevention. Depending on the type of field application, there are several key parameters affecting the scaling process and the performance of PWT, including flow velocity, heat flux, water hardness, surface condition of heat exchanger, blowdown schedule, etc. For example, flow velocity varies greatly from water-cooled condensers (i.e., 1–3 m/s) to evaporators used in various process plants (i.e., almost zero). As expected, PWT has seen a large number of successful case studies in condenser applications in recirculating cooling loops, whereas it has experienced limited success in distillation and evaporator applications for scale prevention or mitigation. Thus, it is essential to identify the conditions where PWT does and does not work, which should in turn help us understand the fundamental mechanism of PWT.

A. Criteria for Validation Methods

As the validation of PWT in the field becomes expensive and in many cases impractical, most previous validation attempts used laboratory studies, which often reported negative results as they focused on the changes in water properties by PWT instead of scaling prevention. As the efficiency of PWT critically depends on the test conditions at a laboratory, the outcome of the study could be significantly influenced by test conditions. In order to understand the science behind PWT, one needs to focus on test conditions which show the efficacy of the PWT first, then explore the cases where it does not work to develop an understanding of how it works.

For example, PWT does not work in a hot water heater where an electric heating element is used in a pool of water, e.g., a 50-gallon water heater commonly used at home in the U.S. We conducted a large number of fouling tests with domestic electric hot water heaters and found that PWT could not prevent scaling in these cases. The reason can be explained as follows: the PWT device causes bulk precipitation, leading to particulate fouling and a soft-sludge coating that must be removed by the shear force created by the flow. When there is no cleaning shear force in heat transfer equipment, the soft-sludge coating eventually becomes hardened with time, and the effect of PWT will be practically zero. Thus, if one chooses a hot water heater for the validation study of PWT, we cannot study the science behind PWT.

Wagner [57] described the DVGW (Deutsche Vereinigung des Gas- und Wasserfaches) Standard W 512, which was proposed as a validation method of PWT devices. However, since there is no shear force inside the boiler used in the DVGW Standard, one can almost predict that no PWT

device will pass the DVGW test. The only device that can pass such a test may be a water-softening device such as a reverse osmosis or ion exchanger which removes mineral ions from water. In order to properly evaluate the efficacy of a PWT device, one has to have a decent flow velocity of at least 1 m/s so that the soft-sludge coating, if formed, can be removed from the heat transfer surface by the shear force produced by the velocity [43].

B. Observation of $CaCO_3$ Crystals Using Microscope

Kronenberg [58] used a rather simple method to evaluate the effect of magnetic treatment. He placed a small water sample (i.e., 20–100 μL) on a glass slide and allowed it to evaporate. Water evaporated leaving a ring of several hundred prismatic $CaCO_3$ crystals along its perimeter. He defined the efficacy of magnetic treatment by the reduction of the number of prismatic crystals and counted them using a polarized light microscope. A efficacy of above 80% was achieved as the number of permanent magnets was increased from 2 to 8. At a given arrangement of magnets, efficacy varied with increasing flow velocity through the magnetic field, reaching peak efficacy of 90% at an optimum velocity (i.e., 2.5 m/s) and then gradually decreasing beyond the optimum velocity. As the axial spacing between two magnets doubled, the efficacy vs. flow velocity curve was shifted to the right, resulting in the optimum velocity at 4.5 m/s with a reduced efficacy of 75%.

Lee [59] also conducted similar tests and obtained photographs of the microscopic result of the water-drop-evaporation experiment with water samples treated by a PWT device using a solenoid coil (i.e., see Fig. 4). The photographs shown in Fig. 8 illustrate the final stage of the untreated and treated cases. Both cases had a similar giant crystal band around the perimeter of the water drop. The untreated case showed many tiny crystal particles, and the average size of each crystal was 10–15 μm, except for the outer crystal band. In contrast, huge crystals covered the whole surface of the water drop in the treated case. Several big dents were evident on the huge crystals. The photographs show a part of the perimeter of a drop because almost all the solidification of the mineral ions took place at or close to the perimeter of the drops. It is speculated that this was caused by vigorous radial convection inside the evaporating drop, which transported the heavier minerals toward the drop perimeter [58].

In the no-treatment case, no crystal was found in the water drop for the first 30 min. After 45 min, numerous tiny particles appeared around the boundary of the water drop. During the next 10 min, more crystals were made but the size of each crystal did not increase further. In comparison,

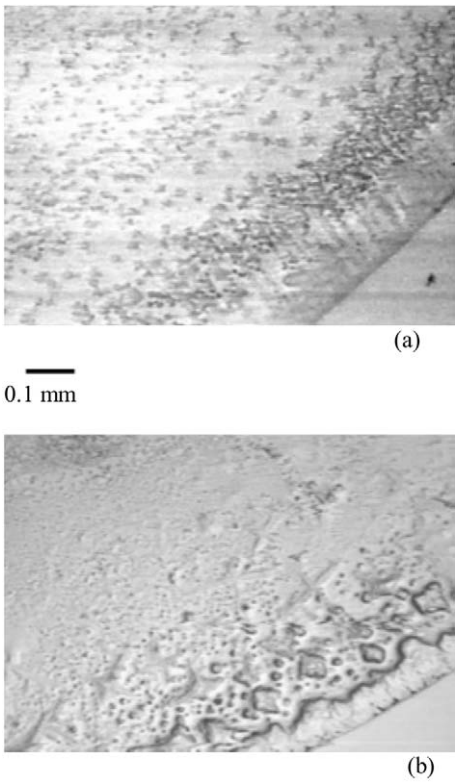


FIG. 8. Photographs of calcium carbonate particles from water-drop-evaporation test taken using a light microscope. (a) Untreated water sample and (b) Water sample treated by PWT.

far fewer but much larger crystals were produced in the treated-water case during the same time intervals. The diffusion of ions would be the dominant mechanism of crystallization in the untreated case. On the other hand, a number of seed crystals were produced by PWT in the treated water, and new calcium ions continued to adhere as pure water evaporated, making the remaining water on the glass slide extremely supersaturated and allowing seed crystals to grow into huge crystals. The results observed by Kronenberg [58] and Lee [59] indirectly support the theory of bulk precipitation.

C. Analytical Measurements

In order to validate PWT, many researchers conducted analytical measurements for water analyses and crystal structures using standard

laboratory test protocols, which include UV-absorbance, particle size measurement (i.e., Hiac/Royco VersaCount or Malvern ZetaMaster), SEM (scanning electron microscopy) analysis, X-ray diffraction method, turbidity (i.e., Hach Turbidimeter), zeta potential, and force–distance measurements (i.e., Topometres Explorer) [60,61].

They often used distilled water and added chemicals to simulate the circulating cooling water in the field. For example, water samples were often prepared in distilled water by adding $CaCl_2$ and Na_2CO_3, i.e., artificial hard water. In general, with the artificial hard water and commercial permanent magnets for PWT, the measurements of particle size, zeta potential, and AFM (atomic force microscopy) did not show any significant difference between untreated and magnetically treated water samples. Note that the above artificial hard water is different from the typical circulating cooling tower water made from "city water" in the characteristics of scale formation. The cooling tower water is supersaturated in calcium and has abundant bicarbonate ions with pH values of 8–10, whereas there are a large number of sodium and chloride ions in the artificial hard water, which may significantly affect the heterogeneous catalytic performance of PWT and alter the course of the precipitation of $CaCO_3$. As we found from previous fouling studies [62–65], PWT worked less effectively with artificial water, and one can expect that the results of the analytical measurement may not show difference for artificial water treated by PWT.

Gehr *et al.* [4] reported an approximate 23% drop in zeta potential in magnetically treated water, which was supersaturated (3 g/l at 21°C) and similar to water used in the field. A tube containing the water sample was rotated in a magnetic field produced by a MRI spectrometer (with a magnetic field strength of 4.75 T), which was significantly larger than most PWT devices using permanent magnets (i.e., 0.2 T). They suggested that the effect of magnetic treatment indirectly increased the coagulating tendency of the suspension.

As the hydroxide ion (OH^-) was considered as the precursor of the bulk precipitation caused by PWT devices, one can ask whether or not PWT may increase the value of local pH on the surface of a pipe where the PWT device is installed. Busch and Busch [8] and a former student [27,30] measured the changes in pH on the pipe surface as well as in bulk. They used deionized water that was made conductive by adding sodium chloride. They circulated the solution through a rectangular channel surrounded by a set of permanent magnets. They showed a plot of surface pH minus bulk pH as a function of time for magnetic treatment using a metallic housing and a PVC pipe. The results shown in Fig. 9 showed that the pH at the surface was greater than the bulk value, indicating that the metal surfaces for the magnetic treatment were more alkaline than the

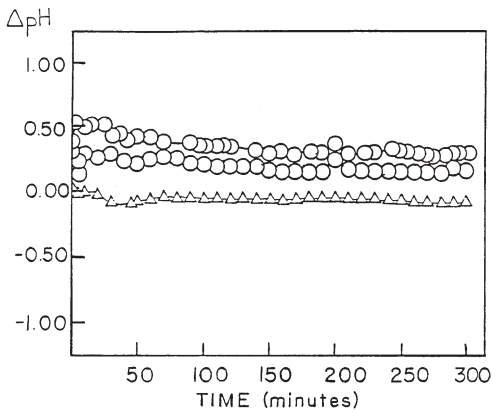


FIG. 9. pH measurements between surface and bulk for magnetically treated water [8,27,30].

bulk solution. As the test reported by Busch and Busch [8] was conducted with deionized water, it would be interesting to see if such findings can be replicated with natural hard water.

D. Particle Size Measurements

For circulating cooling tower water, it would be of interest to see whether or not PWT causes some changes in particle size between untreated and treated water samples. Lee [59] measured the distribution of particles in cooling water with a sub-micron laser particle sizer (Nicomp, Model 200, from Santa Barbara, CA). For the cases of make-up city water (from Philadelphia, PA) and untreated cooling tower water samples, the latter had slightly larger particles than the former in particle size distributions. Note that the cycles of concentration of the circulating water sample was approximately four, i.e., the electric conductivity of the make-up water was 510 μmhos/cm, whereas that of the no-treatment case was 1970 μmhos/cm. Hence, high concentration in the circulating water increased the particle size slightly as shown in Fig. 10.

Figure 11 represents the results of particles distribution for no-treatment and two cases with permanent magnets, identified as PWT-2.3 m/s and PWT-3.0 m/s. The average particle size of PWT-3.0 m/s case was about twice that of the no-treatment case. Note that the average particle size was determined by the software in the particle counter. For the case of PWT-2.3 m/s, the average size of particles for treated water sample was 34,200 nm, i.e., 18 times bigger than that of the no-treatment case which had an average size of 1900 nm. This result shows the existence of huge

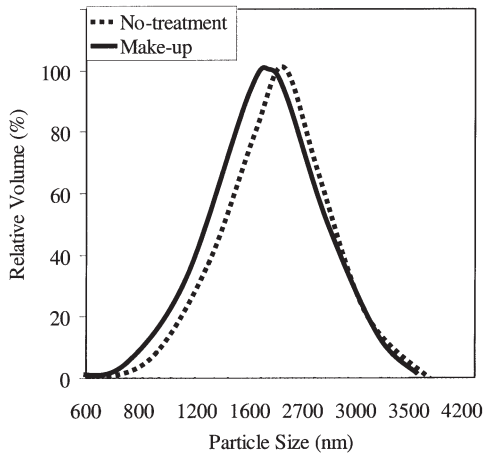


FIG. 10. Results of particle distribution measurements of cooling water for no-treatment and make-up water.

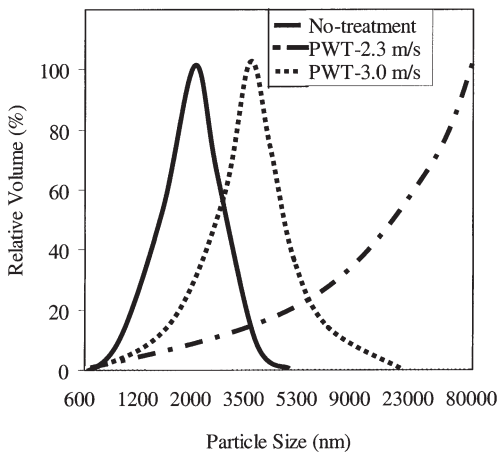


FIG. 11. Results of particle distribution measurements of cooling water for no-treatment, PWT-2.3 m/s and 3.0 m/s cases.

particles in the bulk cooling water when the cooling water was treated with the PWT device. The results of the particle-size measurement reported by Lee [59] provide some support on bulk precipitation and the existence of large colloidal particles suspended in the cooling water treated by PWT.

E. Validation Method Using Fouling Experiments

The validation method introduced in this section uses fouling experiment which was developed by a study carried out at Drexel University under the support and supervision of ASHRAE Technical Committee 3.6 (Water Treatment) during the period of September 2000 to May 2002 [66]. The objective of this study was to investigate the efficiency of physical water treatment (PWT) technologies in preventing and controlling calcium scale accumulation on heat transfer surfaces in re-circulating open cooling-tower water systems. This section describes the experimental facility and test procedure used to validate the PWT. The tests were limited to cooling-tower water applications, where water was repeatedly treated by a PWT device. All tests were conducted with a biocide, glutaraldehyde (Union Carbide, Houston, TX), to eliminate the effect of biofouling on this mineral fouling study.

Figure 12 shows a schematic diagram of the test facility, which consisted of a water-circulating loop, a cooling tower, a side-stream flow loop, a

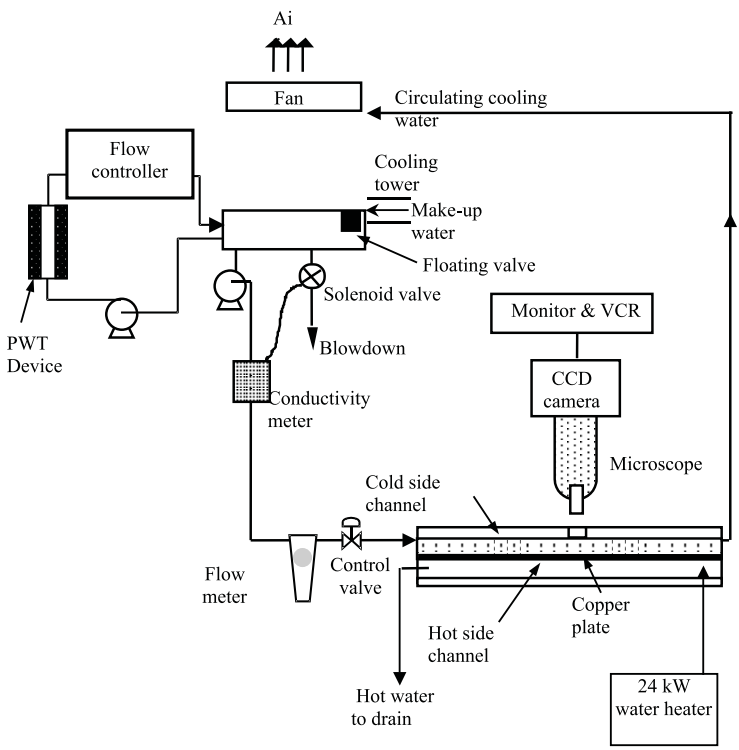


FIG. 12. Schematic diagram of an experimental system proposed for validation test using fouling experiments.

permanent magnetic device, a pump, a heat-transfer test section, an electric-conductivity meter, a floating-ball valve for automatic feeding of make-up water, and an automatic blowdown system controlled by a solenoid valve which was controlled in turn by an electric-conductivity meter. In order to change the flow velocity through permanent magnets without changing the flow velocity in the main heat-transfer test section, a side-stream flow loop was set up in a sump tank of the cooling tower.

Figure 13 shows a cross-sectional view of the heat-transfer test section that consisted of a copper plate as the heat-transfer surface, a cooling-water channel (i.e., the upper flow channel) and a hot-water channel (i.e., the lower flow channel). Cooling water moved in the opposite direction of the hot water, forming a counter-flow heat exchanger.

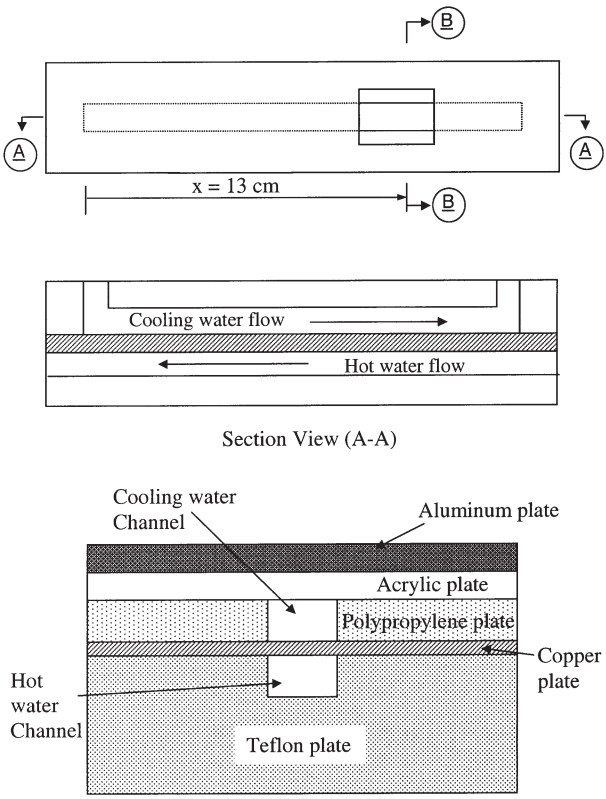


FIG. 13. Cross-sectional view of a rectangular heat transfer test section used for validation study using fouling experiment.

Since the solubility of calcium ions decreases with increasing temperature, scale was expected to deposit on the cooling-water side of the copper plate. At the end of the test, the fouled copper plate was removed for SEM analyses of scale, and a new copper plate was installed for the next test. For the cooling-water channel, the two sidewalls were made of polypropylene plates, whereas the top wall was made of a transparent acrylic plate due to its thermal insulation properties and for visual observation of crystal growth. The hot-water channel was machined out of a Teflon block, which had excellent thermal insulation properties. The dimensions of the cooling-water channel were 1.6 mm × 6.4 mm × 254 mm (height × width × length) and the dimensions of the hot-water channel were 13.7 mm × 6.4 mm × 254 mm (height × width × length).

In order to directly observe scale crystals on the heat-transfer surface as the fouling progressed, a fouling test facility with a microscopic-imaging system was developed. The microscopic-imaging system consisted of a microscope, an external illuminator, a Sony CCD (charged couple device)-IRIS camera, a 35-mm Nikon camera, a video recorder, and a monitor. The images of $CaCO_3$ scale crystals on the cooling-water side were observed using the microscope and continuously recorded on VHS video-tape via the CCD camera. Hard copies of the microscopic images of scale crystals were produced from the VHS tape using a software called Snappy. A magnification of 40 × was used for observation. In addition, still pictures were taken with the 35-mm Nikon-camera.

The inlet temperature of the circulating water prior to the heat-transfer test section was maintained at $20 \pm 1^{\circ}C$ by means of a cooling tower, whereas the inlet temperature of the hot water was maintained at 89 to 92°C throughout the entire experiment. Heat fluxes between the heating and cooling sides of copper plate were in a range of 380 and 485 kW/m^2. Note that the heat flux in typical water chillers is in a range of 20–30 kW/m^2. In the present study, a much higher heat flux was used to accelerate the fouling so that each run could be completed in one week. The use of such a high heat flux in the laboratory fouling study is not uncommon [67–69]. Of note is that Vermeiren [70] conducted research that suggested that PWT should not be used in cases where heat flux was greater than 35 kW/m^2.

In order to calculate the convective heat transfer coefficient over the copper plate during the fouling test, Chiranjivi and Rao's convective heat-transfer correlation was used, which was for flow in a bottom-wall-heated rectangular duct [71]:

$$h = 0.79\mathrm{Re}^{0.4}\mathrm{Pr}^{0.52}\frac{k_w}{D_h} \tag{17}$$

where h is the convective heat-transfer coefficient, k_w is the thermal conductivity of water, and D_h is the hydraulic diameter for the rectangular cooling channel. To calculate the average surface temperature, Newton's law of cooling was introduced as:

$$\dot{Q} = hA(T_s - T_m) \tag{18}$$

where A is the heat-transfer surface area, T_s is the surface temperature, and T_m is the mean temperature of the fluid. The average surface temperature of the copper plate, T_s, was estimated to be $68 \pm 2^{\circ}$C using Eqs. (17) and (18). The heat-transfer-surface temperature was applied in the calculation of Langelier saturation index (LSI).

After the cooling water passed through the heat-transfer test section, it was sent to the cooling tower, where the water was cooled by evaporation. In order to maintain a constant volume of water in the system, tap water was added as make-up water, a procedure that was controlled by an automatic floating-valve system. The tap water supplied by the City of Philadelphia was used for make-up water. As water quality varied with time, the average total hardness was 190 mg/L, and the average electric conductivity was 520 μmhos/cm (data reported by the Bureau of Laboratory Service, Philadelphia).

Figure 14 shows test procedure and variations of electric conductivities vs. time used for the cases identified as MWT(A), MWT(B), and PMDU. All fouling tests were performed while maintaining the same level of conductivity in the cooling tower. In the fouling study the electric

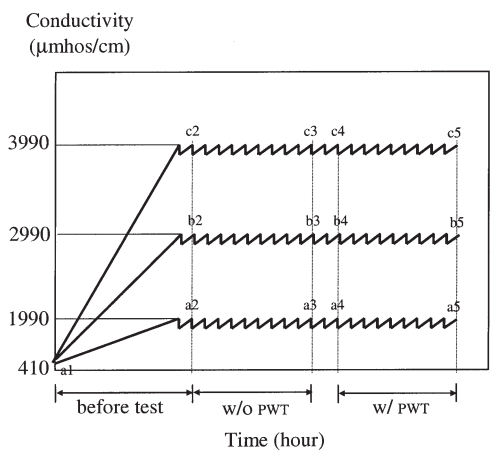


FIG. 14. Test procedure and variations of electric conductivity vs. time. Tap water was used to make a concentrated water by evaporating pure water using a laboratory cooling tower.

conductivity was varied because it was much more convenient to control the electric conductivity experimentally than the hardness of circulating water. Before starting each test, the water sample was prepared by circulating tap water through the cooling tower so that the desired electric conductivity was obtained through evaporation. Once the circulating water reached the desired conductivity, the fouling test began with hot water in the hot water channel. At the end of the test for the no-treatment case, the cooling-tower water was continuously circulated while treated by the PMDU, but without heat transfer for one day (i.e., see the period between a3 and a4). After one-day pretreatment, a brand new copper plate was installed in the heat-transfer test section, and the fouling test began. Each fouling experiment was repeated for the reliability of test results at the same experimental conditions.

The fouling resistance is often calculated using the following equation:

$$R_f = \frac{1}{U_f} - \frac{1}{U_i} \tag{19}$$

where U_f is an overall heat transfer coefficient from a heat exchanger experiencing fouling, and U_i is the overall heat transfer coefficient from a clean heat exchanger. Both overall heat transfer coefficients were calculated from the following equation:

$$U = \frac{\dot{Q}}{A\Delta T_{lm}} \tag{20}$$

where ΔT_{lm} is a log-mean-temperature difference, which can be described as

$$\Delta T_{lm} = \frac{(T_{h,i} - T_{c,o}) - (T_{h,o} - T_{c,i})}{\ln\left[(T_{h,i} - T_{c,o})/(T_{h,o} - T_{c,i})\right]} \tag{21}$$

The heat transfer rate, $\dot{Q}$, was estimated from both the heating and cooling channels as

$$\dot{Q} = [\dot{m}c_p(T_i - T_o)]_h = [\dot{m}c_p(T_o - T_i)]_c \tag{22}$$

The heat transfer rate at the hot side was compared with that at the cold side. The difference was less than 5% as the heat loss from the hot side to the surroundings was kept minimal by insulating the heat transfer test section. The heat transfer rate in the cooling-water side, $\dot{Q}_c$, was used to calculate the overall heat transfer coefficient in the fouling study. The results of fouling experiments were obtained by varying

parameters such as the arrangements of permanent magnets, flow velocity through magnetic fields, concentration of hard water, etc., and will be given in Section V.

V. Velocity Effect on Magnetic Water Treatment

This section reports experimental results of fouling tests, which were conducted to check whether or not PWT using a permanent magnet can prevent or mitigate mineral fouling. When one conducts a validation study using a permanent magnet, the flow velocity through the magnetic device becomes one of the most critical experimental parameters. If the overall performance of PWT using the permanent magnet significantly depends on the velocity, the outcome of the fouling test should also directly depend on the selection of flow velocity. Thus, in this section we review the issues involved in flow velocity in the use of permanent magnets.

A. Background

Numerous researchers investigated the feasibility of PWT using permanent magnets in reducing mineral fouling in a heat exchanger. Some of them [22–28] reported that PWT changed water properties and/or produced nucleation sites in bulk solution, whereas others [72,73] reported that PWT did not work. After reviewing the literature, it is clear that the effect of PWT on scale prevention was not consistent because the effectiveness of PWT critically depended on test conditions, including the flow velocity through permanent magnets, water hardness (i.e., levels of supersaturation), and recirculation, i.e., whether PWT was used in a recirculation mode or once-through system.

The effectiveness of the permanent magnet appears to be strongly related to the flow velocity passing through the magnetic fields [8,25]. Grutsch and McClintock [25] reported through a full-scale field test that scaling was completely prevented at a flow velocity of 6 m/s using a magnetic device (1700 gauss), where the cooling-tower water passed through an annulus-shaped gap between magnets installed at the center of the pipe and pipe wall. Szostak [26] reported that he could operate a cooling tower at 3000-ppm hardness without scale buildup by using a magnetic fluid conditioning system. He designed the internal structure of the magnetic device such that the cross-sectional area was reduced to increase the flow velocity to 6 m/s through the magnetic treatment device. In addition, Busch and Busch reported that the induced voltage by permanent magnets was linearly increasing with flow velocity [8]. They reported that a certain

minimum flow rate might be necessary to produce an anti-scaling effect with permanent magnets. Thus, we investigated the effect of flow velocity through permanent magnets on the mitigation of mineral fouling in a heat exchanger in an open-cooling-tower system.

Grutsch and McClintock [25] reported through a full-scale field test that scaling was completely prevented at a flow velocity through magnets of 6 m/s using a magnetic device (1700 gauss), where the cooling-tower water moved between the magnets at the center and pipe wall. When the flow velocity at the upstream pipe was 4.5 ft/s, they could obtain the maximum induced voltage. Note that the cross-sectional area at the upstream pipe was four times larger than the cross-sectional area of the pipe with permanent magnets so that the actual flow velocity through the permanent magnets was 18 ft/s (5.5 m/s).

B. Experimental Methods

The schematic diagram of the fouling test facility and heat transfer test section were shown in Figs. 12 and 13, respectively. Two different types of permanent magnets were tested in the study: one was a clamp-on type and the other was in in-line geometry where multiple permanent magnets were positioned at the center of the feed pipe. Figures 3a and 3b show two different permanent magnets identified as MWT(A) and MWT(B), respectively, as well as their arrangements and the cross-sectional dimensions of the channel where water was treated by these permanent magnet devices. The MWT(A) was made at the Fouling Laboratory at Drexel University, whereas the MWT(B) was a commercial product. The strength of the magnetic field produced by both devices was 0.16 T (or 1600 gauss) as measured by an AlphaLab DC Magnetometer. The dimensions of the flow channels made by the two permanent magnets in MWT(A) and MWT(B) are given in Figs. 3a and 3b, respectively. In MWT(B), cooling water passed through the annulus gap between cylindrical permanent magnets and a circular tube casing.

The objective of the study was to investigate whether or not there was an optimum flow velocity for each magnetic device which maximized the performance of magnetic water treatment. The flow velocity was varied in order to measure the performance over a range of flow velocities. For MWT(A) cases, the flow velocity was varied from 2 to 8.5 m/s through the PWT device in the side-stream flow loop, while the flow velocity at the heat-transfer test section was fixed at 1.2 m/s. The corresponding Reynolds numbers at the MWT(A) device were varied from 5600 to 24,000 based on the hydraulic diameter of the flow channel made by the magnets. For MWT(B) cases, the flow velocity was varied from

1.1 to 3.0 m/s through the PWT device in the side-stream flow loop, while the flow velocity at the heat-transfer test section was fixed at 1.5 m/s. The fouling resistance was calculated from the four temperature data and flow rate of the cooling water side using Eqs. (19)–(21).

C. Results and Discussion: Flow Velocity Effect

Since the conductivity of the make-up and circulating cooling tower waters for MWT(A) case were 520 and 2450 μmhos/cm [74], approximately 5 COC (cycles of concentration) were projected to be in effect, based on the conductivity. Using a mass balance, actual COC were found to be 6.17 on average (see Ref. [74] for detailed calculations using the mass balance method). The COC calculated from conductivity values were underestimated by approximately 20% because the method did not consider the precipitation of dissolved mineral ions in water. For MWT(B) case, the set point of conductivity was 2000 μmhos/cm so that approximately 4 COC were in effect based on conductivity (see Table II).

The Langelier saturation index (LSI) was calculated to examine the precipitation tendency of circulating water. The values of the LSI were in the range of 2.14 to 2.35 in MWT(A) tests and from 2.48 to 2.51 in MWT(B) tests, indicating that the circulating water in both cases had the capability to cause severe mineral fouling.

TABLE II
Water Quality Data Measured for Make-up and Circulating Water

	Make-up water	No treatment	Magnetic water treatment			
			$V = 1.1$ m/s	$V = 1.7$ m/s	$V = 2.3$ m/s	$V = 3.0$ m/s
Conductivity (μmhos/cm)	510	1970 (3.9)*	1980 (3.9)	1990 (3.9)	1975 (3.9)	1995 (3.9)
pH	7.4	8.3	8.3	8.4	8.4	8.4
Total hardness (mg/L)	176	760 (4.2)	748 (4.2)	744 (4.2)	755 (4.2)	745 (4.2)
Ca^+ hardness (mg/L)	131	440 (3.4)	436 (3.4)	448 (3.4)	448 (3.4)	456 (3.4)
Total alkalinity (mg/L)	68	228 (3.4)	236 (3.4)	232 (3.4)	230 (3.4)	228 (3.4)
Chloride (mg/L)	93	440 (4.7)	456 (4.7)	436 (4.7)	440 (4.7)	436 (4.7)
LSI**	−0.01	2.48	2.48	2.49	2.49	2.51

* ( ) represents the ratio of each treatment to the make-up water.
** LSI = Langelier Saturation Index at $T = 68°C$.

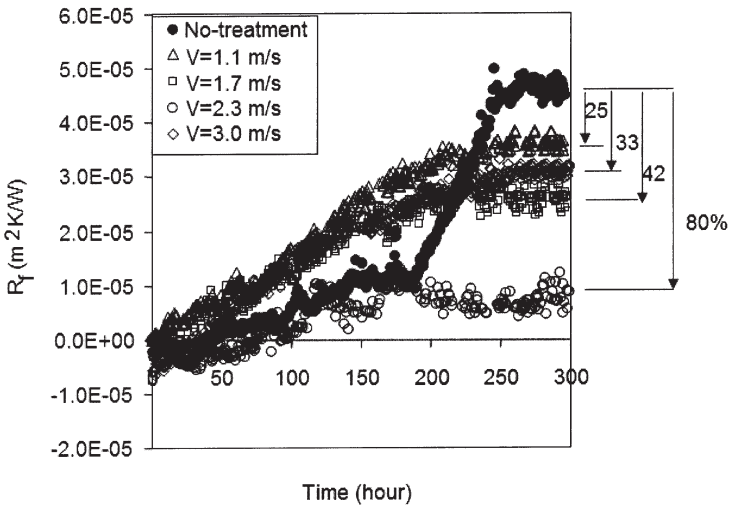


FIG. 15. Fouling resistances vs. time for five different cases with MWT(B).

Figure 15 represents fouling resistance values vs. time for five different cases with MWT(B). In this case, the best magnetic treatment was obtained at a flow velocity through the MWT(B) of 2.3 m/s, where the fouling resistance was reduced by 80% from the value of the no-treatment case. When the flow velocity through the MWT(B) was not at this optimum value, the treatment was less effective. Similar fouling-resistance curves vs. time were obtained for five different runs with MWT(A). For the case of 6.3 m/s, the rate of increase in the fouling resistance was the least among the five different velocities, thus resulting in the least fouling resistance at the end of the 160-h test. When the flow velocity through the MWT(A) device increased from 2.0 to 6.3 m/s, the magnetic treatment became more efficient, rendering the best performance at 6.3-m/s case, where the fouling resistance was reduced by 57%. However, when the flow velocity was further increased to 8.5 m/s, the effectiveness of the magnetic device significantly dropped.

Figure 16 shows the percentage reductions of the fouling resistance vs. flow velocity through permanent magnet for both MWT(A) and MWT(B). As can be seen in Fig. 16, the maximum effectiveness for MWT(A) and MWT(B) was obtained at a flow velocity of 6.0 and 2.3 m/s, respectively. Note that Grutsch and McClintock [25] reported that 6.0 m/s was the required velocity for the maximum performance of their magnetic water treatment device. Also note that the magnets (i.e., the ones used in the present study and that of Grutsch and McClintock) had an almost identical strength of the magnetic field, e.g., 1600–1700 gauss. The difference

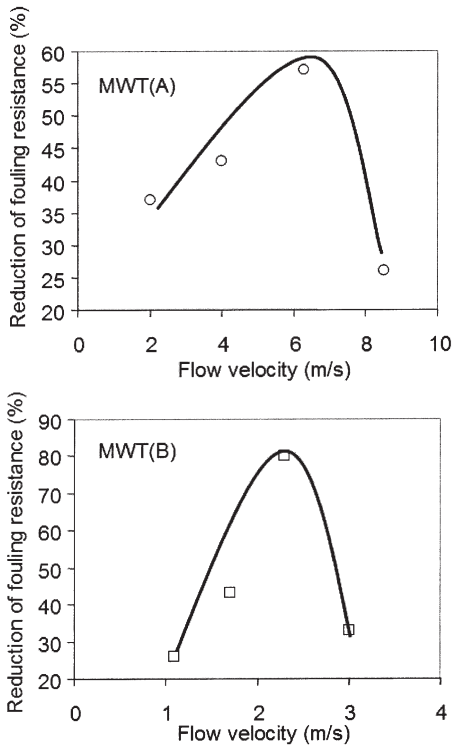


FIG. 16. Percentage reductions of the fouling resistance vs. flow velocity through permanent magnet for both MWT(A) (top figure) and MWT(B) (bottom figure).

in the optimum flow velocity observed in different magnetic devices may be attributed to different arrangements of the permanent magnets, i.e., the difference of pole positions and subsequently different profiles of the magnetic fields.

It is not clear why the performance of a permanent magnet device shows the optimum behavior at a specific velocity. One may speculate as follows: for the no-treatment case, the scale deposition was mainly due to the diffusion of dissolved mineral ions, which results in crystallization fouling and tenacious hardened scale. When the cooling water was treated by PWT at the optimum velocity through magnets, the mineral ions that were dissolved in the cooling water precipitated in bulk water, producing seed crystals for particulate fouling. The precipitated particles grew in size through recirculation and eventually were deposited on the heat exchanger surface in the form of particulate fouling. This may explain

why the fouling curves for the PWT cases show more rapid increase in the beginning of tests compared with the no-treatment case. However, these particles deposited on the heat transfer surface as a soft sludge coating, could easily be removed by the shear force created by the flow in the heat exchanger. On the other hand, for the no-treatment case the hardened scale deposit is produced mostly by the diffusion of ions, often known as precipitation fouling, and could not be removed easily by the same shear force. Thus, the fouling resistance for the no-treatment case was substantially greater at the end of fouling tests than the PWT cases.

When the flow velocity through the magnets was different from the optimum velocity, it is not clear why the performance of the magnetic treatment significantly decreased. When the flow velocity was greater than the optimum value, one can consider that the contact residence time between water and magnetic fields was too short for water to be effectively treated by the magnetic field. In other words, the precursor of the bulk precipitation is believed to be OH^- ions, which needs some time to dissociate the bicarbonate ions and to precipitate calcium carbonate particles in bulk. When the flow velocity was much smaller than the optimum value, the agitating force via Lorentz force may not be sufficient to cause bulk precipitation. The production of the hydroxide ions may not have been sufficient, and thus the magnetic treatment may have been less effective.

Figure 17 shows photographs of fouled surfaces for MWT(A) cases taken after the fouled heat transfer surfaces were completely dried. The case with a flow velocity of 6.3 m/s gave the least amount of scale deposit. No treatment case resulted in the most severe fouling according to the photographs. When the circulating water passed through the MWT(A) device at a velocity greater than the optimum velocity, i.e., at 8.5 m/s, the MWT(A) was not effective, as manifested by Fig. 17(d).

Figure 18 shows photographs of fouled heat-transfer surfaces for MWT(B) cases taken after the fouled heat transfer surfaces were completely dried. No-treatment cases showed much more severe fouling than PWT cases, regardless of flow velocity. The amount of fouling deposit decreased with increasing flow velocity through MWT(B) from 1.1 m/s to 2.3 m/s. When the flow velocity was increased beyond the optimum velocity of 2.3 m/s, i.e., to 3.0 m/s, the amount of scale deposition was substantially increased.

Figure 19 shows scanning electron microscope (SEM) photographs of scale samples obtained with MWT(A) at the end of 160-h tests with a magnification of $1000\times$. In the SEM photographs, only the top surfaces of the scale layer are seen. Figure 19(a) shows a SEM photograph for the no-treatment case after a 160-hr test. The crystals are close to rhombic

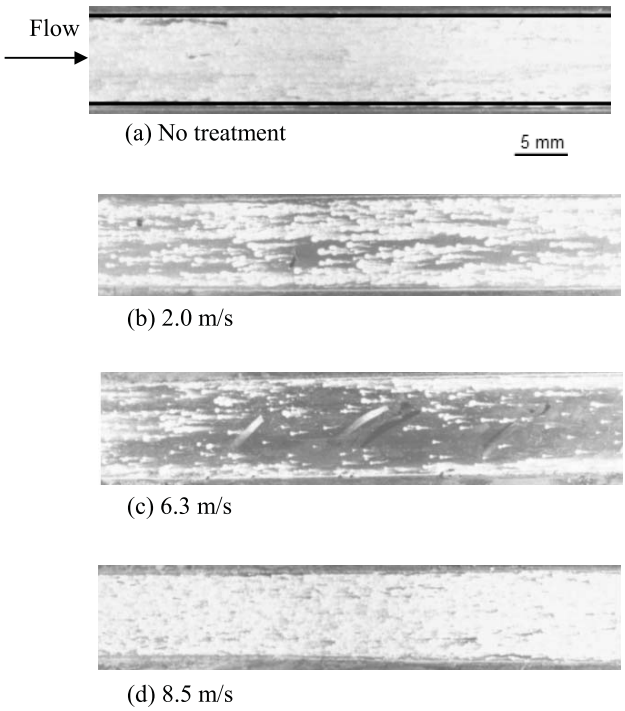


FIG. 17. Photographs of fouled heat transfer surfaces with MWT(A) .

in shape with round edges. The growing surface of each crystal (i.e., the interface between solution and solid phase) appears to be very smooth, indicating that the crystals might have grown through crystallization reaction assisted by ion diffusion. In other words, the scale layer for the no-treatment case might have been formed due to a mechanism of crystallization fouling on the surface.

Figure 19(b) shows a SEM photograph for MWT(A) with a flow velocity of 2 m/s. Compared to the pattern in Fig. 19(a), the growing surface of scale appears to be rough and irregular, suggesting that the mechanism of scale formation in a magnetically treated solution is quite different from the crystallization reaction. A similar pattern of scale was observed for MWT(A) with a flow velocity of 4 m/s. The top surface of each mineral scale in both cases with 2.0 and 4.0 m/s appears to be flat as shown in Figs. 19(b) and (c), whereas the cases with higher flow velocities of 6.3 and 8.5 m/s resulted in a very rough surface of scale layer (see Figs. 19(d) and (e)). It is speculated that the different patterns of scale seen

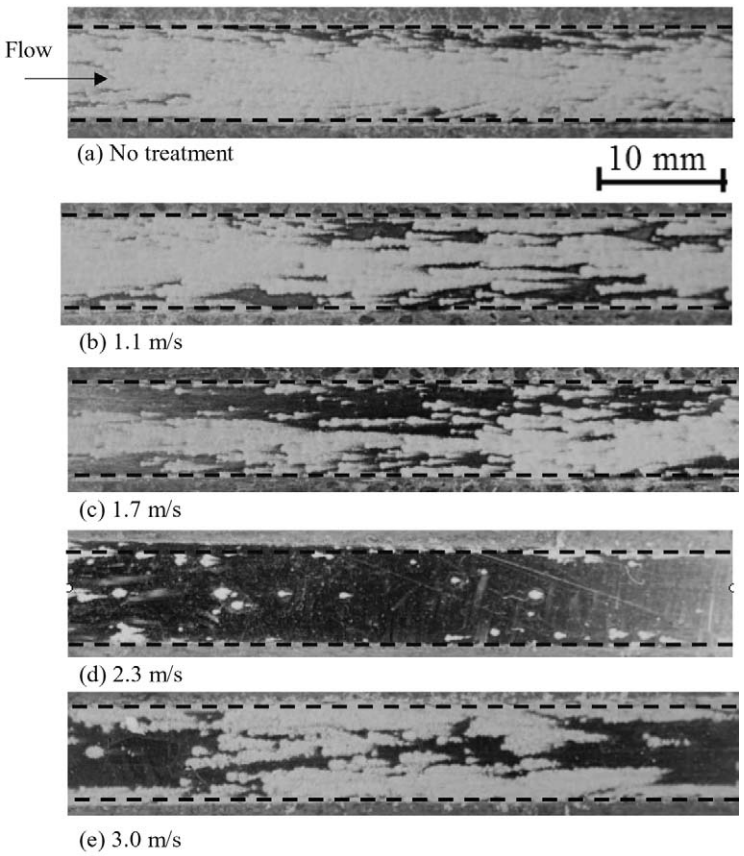
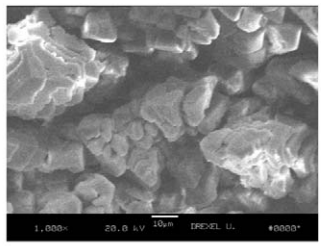


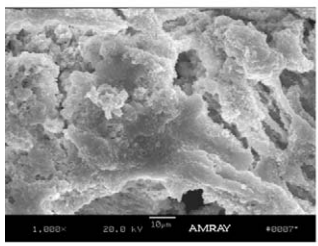
FIG. 18. Photographs of fouled heat transfer surfaces with MWT(B).

from the top surface resulted from the different fouling mechanisms. In the cases with higher flow velocities of 6.3 and 8.5 m/s, one may speculate that the scale layer appears to be formed by the accumulation of numerous small-sized crystallites suspended in circulating water instead of crystallization reaction. In short, PWT changed the nature of crystal formation, forming an irregular growing surface and sludge-like crystal formation, whereas the no-treatment case showed a regular pattern of crystal growth on the heat-transfer surface.

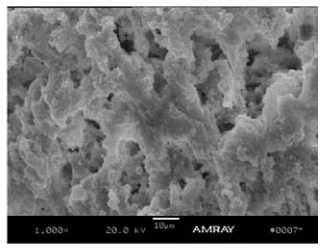
Figure 20 shows scanning electron microscope (SEM) photographs of scale samples obtained with MWT(B) at the end of 300-h tests with magnifications of 500 and 3000 $\times$. For the no-treatment case, numerous tiny crystals were agglomerated. In the case of MWT(B) with a flow velocity



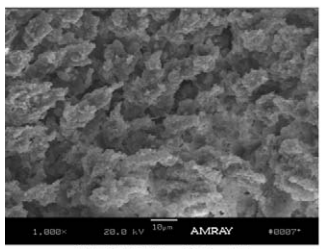
(a) No treatment X1000



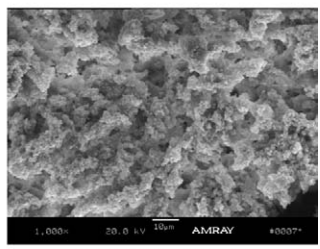
(b) MWT (A) – (2 m/s)



(c) MWT(A) – (4 m/s)



(d) MWT(A) – (6.3 m/s)



(e) MWT(A) – (8.5 m/s)

FIG. 19. Scanning electron microscopy photographs.

of 1.7 m/s, the average crystal size was much bigger than that of the no-treatment case. In the case of MWT(B) with a flow velocity of 2.3 m/s, it was not easy to determine the grain boundary of each crystal. Crystals often appear as a single large irregular crystal. Based on the SEM results, one may speculate that the MWT(B) produced $CaCO_3$ particles in water, which grew in size and eventually deposited on the heat-transfer surface in an amorphous colloidal form. The major difference in the morphologies of the scale layer between the untreated and PWT cases may shed some light on the mechanism of PWT (i.e., crystallization fouling vs. particulate fouling). More SEM photographs for the other test cases are available elsewhere [43,59,66,75].

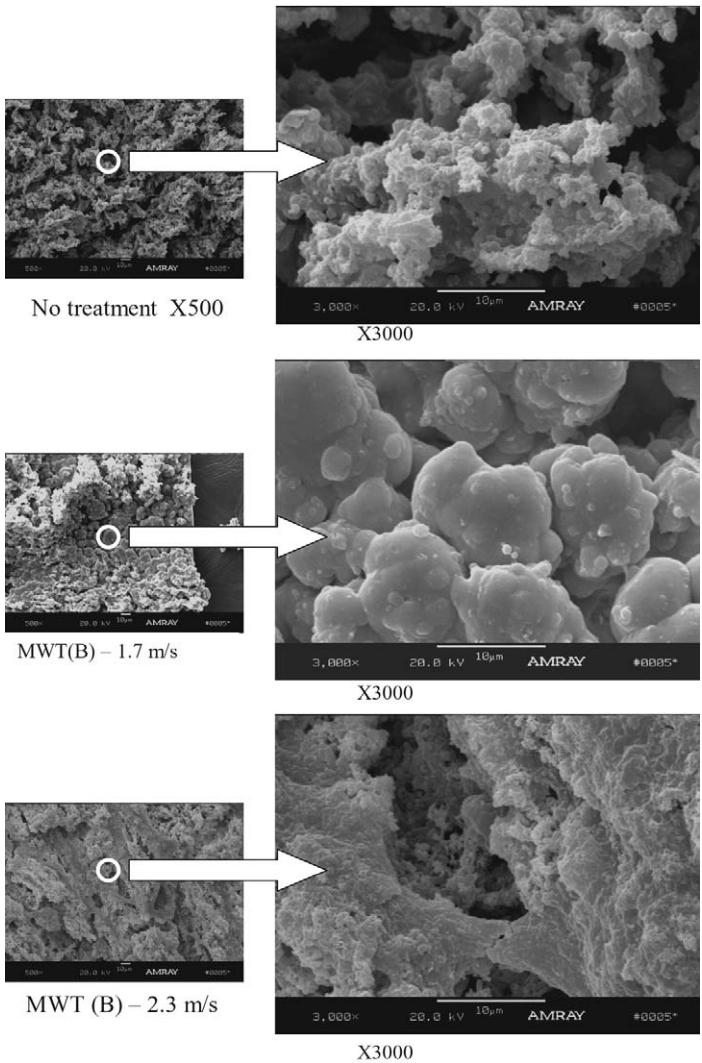


FIG. 20. Scanning electron microscopy photographs.

Figure 21 shows a series of time-dependent photographs taken using the microscopic imaging system for three MWT(A) cases. The photographs represent the process of scale-layer formation, revealing the population of nucleation sites, the growth of each scale crystal, secondary scale formation behind an individual scale crystal, and its ripening.

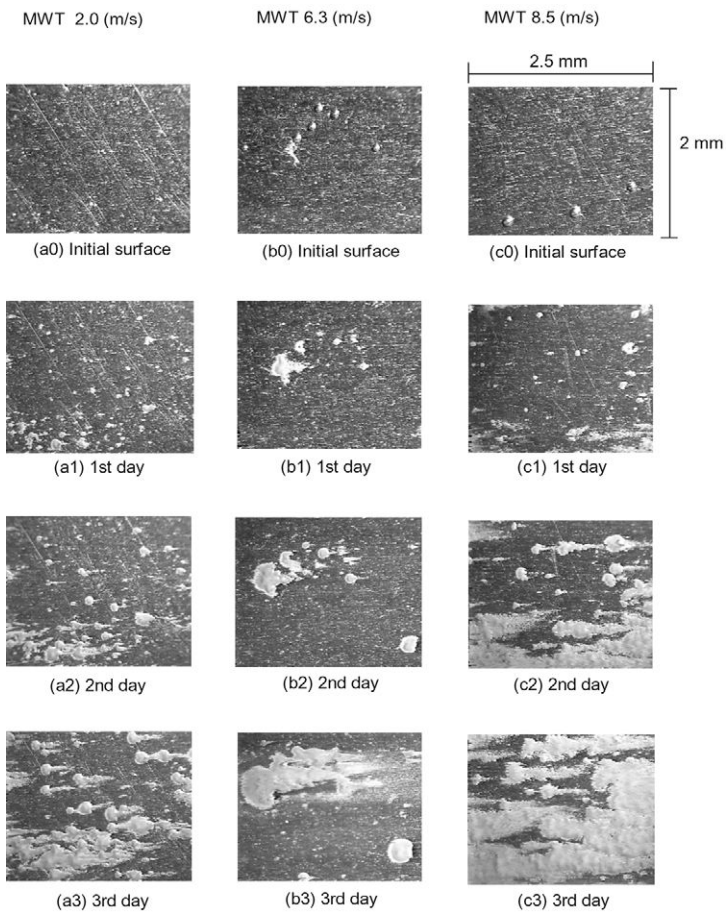


FIG. 21. (a) Time history photographs of heat transfer surface during fouling experiment.

The following scale-formation characteristics in PWT are summarized as follows:

- Scale sites were sparsely distributed.
- As scale crystals grew, the flow around the crystals produced secondary scale formation, forming a comet-like scale pattern.
- The scales formed early continued to grow without any new scale site nearby, keeping the comet-like shape.
- The area of each scale became larger and larger and eventually touched adjacent scales, thus being fused into one (i.e., scale layer).

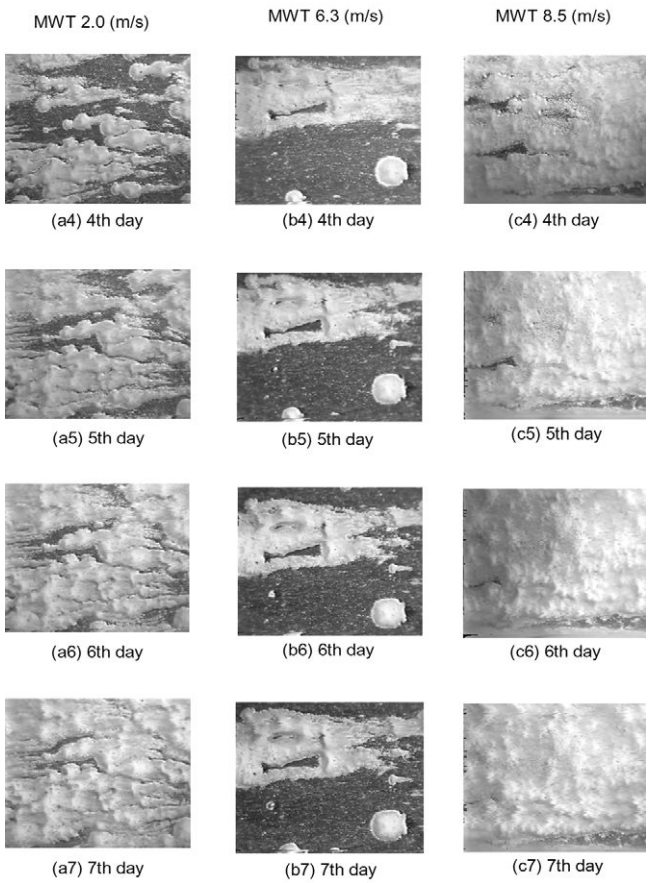


FIG. 21. Continued.

As shown in Fig. 21(a5), the heat-transfer surface was almost fully covered with scaling in about 5 days for the case of MWT(A) using a flow velocity of 2 m/s. After the fifth day, the scale layer continued to grow, covering the heat transfer surface almost completely, suggesting that the flow velocity of 2 m/s through MWT(A) was not sufficient to effectively mitigate fouling. Figure 21(c4) shows that the heat-transfer surface was fully coated in 4 days when the flow velocity through MWT(A) was 8.5 m/s. However, when circulating water was treated at a flow velocity of 6.3 m/s through MWT(A), the population of nucleation sites was not only much lower, but the scale-growth rate also decreased substantially. At this velocity, a good part of the surface remained scale-free for 7 days

as shown in Figs. 21(b0) through (b7), whereas in other cases the heat-transfer surfaces were eventually covered at the end of tests shown in Figs. 21(a7) and (c7).

The present study examined magnetic water treatment used for the purpose of the mitigation of mineral scale. In particular, the study investigated the effect of flow velocity through physical water treatment (PWT) devices on their fouling mitigation as judged by fouling curves. The study found that each PWT device had its optimum operating condition and that a PWT device could substantially reduce mineral fouling in a heat exchanger using cooling-tower water if this optimum condition was provided. However, the benefit of PWT was reduced when the optimum operating conditions were not provided.

In the case of no-treatment, crystallization fouling seemed to be dominant. With PWT, on the other hand, $CaCO_3$ particles were produced in water (i.e., bulk precipitation) and grew, eventually depositing on the heat-transfer surface in the form of particulate fouling. The scale deposit produced by the particulate fouling in the case of PWT was a soft sludge coating, a part of which was removed by the shear force generated by water flow. Microscope-assisted video imaging system and SEM photographs were obtained to render further support to the particulate fouling hypothesis.

D. Effect of Magnetic Treatment Under No Flow Conditions

Higashitani *et al.* [45,61] studied the characteristics of $CaCO_3$ crystals formed by measuring the variation of absorbance of suspensions of $CaCO_3$ particles, a phenomenon which was reproduced by Parsons [10]. Higashitani *et al.* generated static magnetic fields using an electromagnet, resulting in a uniform magnetic flux density in the space where the sample was placed. They prepared solutions by mixing $CaCl_2$ and Na_2CO_3 solutions which were exposed to magnetic fields (0.3–0.6 T with an exposure time greater than 10 min) while the solution was kept in a beaker. As this is one of the few studies reporting the effect of PWT conducted with zero flow velocity, it is important to review this study in understanding the mechanism of PWT using a magnetic field.

They found that the nucleation frequency of particles was suppressed but the growth of existing $CaCO_3$ particles was accelerated, and the magnetic treatment effect produced a memory effect even after 120 h after magnetic exposure. The positive results demonstrated by Higashitani *et al.* [45,61] and Parsons [10] under quiescent conditions suggest that the magnetic fields do not directly alter water molecules or dissolved calcium or bicarbonate ions.

The goal of any physical water treatment is to prevent or mitigate the scale deposit in a heat exchanger when natural water is used as a cooling medium. In general, PWT devices work much less efficiently with artificially hardened water, prepared by mixing $CaCl_2$ and Na_2CO_3 solutions [62,63]. Furthermore, it has been reported that PWT devices showed the best performance when applied in a circulating cooling tower system. Not only repeated treatments, but also the circulating cooling tower system itself, with relatively low temperatures and high water velocity at the condenser, provides optimum conditions for bulk precipitation and particulate fouling. Although the validation method used by Higashitani *et al.* [45,61] is useful in studying the mechanism of the PWT, the validity of PWT should be judged ultimately by a bona fide fouling test.

VI. Concentration Effect

The effectiveness of PWT at laboratory tests, in general, depends on the test conditions, including the flow velocity through permanent magnets, water hardness (i.e., levels of supersaturation or concentration), and recirculation (i.e., a recirculation mode or once-through system). In a cooling-tower application, the concentration of dissolved mineral ions continues to increase due to the evaporation of water. In order to maintain a consistent level of concentration of mineral ions, a portion of water is discharged, and the same amount of water is added in the form of make-up water. The consumption of fresh water, coupled with maintaining the maximum efficiency of the cooling system, is one of the major issues related to cooling water treatment. In other words, if one can use highly concentrated hard water without facing a serious fouling problem, it can be very helpful not only economically but also environmentally.

Furthermore, since we postulate bulk precipitation as the main mechanism of PWT, it is of interest to see whether or not a similar bulk precipitation can take place when the concentration of solution increases to a high level (i.e., beyond 1000 ppm of calcium hardness). If bulk precipitation takes place for the high concentration case, we may get less fouling in a high concentration case than in a low concentration case, shedding some insight on the (bulk precipitation) mechanism of PWT. If a high concentration can be used together with PWT and other mechanical means such as filtration, one can further increase the cycles of concentration in an open-cooling-tower system while holding mineral fouling in a heat exchanger to an acceptable level. Hence, the objective was to investigate the effect of concentration on the fouling resistance with and without PWT.

The schematic diagram of the test facility used in the study was shown in Fig. 12. In order to change the flow velocity through the permanent magnets, a side-stream flow loop was set up in a sump tank of the cooling tower, where a separate pump was used to vary the flow velocity through PWT device independent of the main flow velocity at the heat transfer section.

The flow velocity of water in the cooling-water channel of the main heat transfer test section was fixed to be 1.5 m/s for all runs in the study, and the corresponding Reynolds number was 4470 based on the hydraulic diameter. The flow velocity through a permanent magnet (identified as PMDU in this section) was 6.3 m/s and its corresponding Reynolds number was 17,000. The inlet temperature of the circulating water prior to the heat transfer test section was maintained at $20 \pm 1^{\circ}C$ by means of the evaporative cooling tower, whereas the inlet temperature of the hot water was maintained at 89 to 92°C throughout all the experiments. Heat flux between heating and cooling sides of the copper plate was 485 kW/m^2. The test procedure and variations of electric conductivities vs. time was essentially similar to those shown in Fig. 14. Fouling tests were conducted for three cases with three different water concentrations (i.e., 2000, 3000, and 4000 μmhos/cm) using the PMDU. In each case, the fouling data for the no-treatment case was obtained as a reference.

Table III shows the water analysis data for the make-up, and circulating cooling-tower water for both the no-treatment case and the PWT cases using PMDU. Water samples were taken from the circulating water at times marked by a_i, b_i, and c_i where i varies from 1 to 5 (see Figure 14). Large differences were not observed in concentrations of dissolved materials between no-treatment case and the case with PWT. The Langelier saturation index (LSI) was calculated to examine the precipitation tendency of circulating water. The values of the LSI were in the range of 2.1 to 2.7, indicating that the circulating water held the conditions to cause severe mineral fouling.

Figure 22 shows a series of real-time photographs of fouled heat transfer surfaces during fouling tests obtained using a microscopic imaging system for the no-treatment case and the 3000-μmhos/cm case with PWT (i.e., PMDU). Note that the cooling water was moving from left to right. After 30 h, numerous small particles deposited on a heat-exchanger surface in the case of the no-treatment, whereas several big particles were found for PWT treatment. The number of the small particles for the no-treatment case increased with time and eventually covered the whole heat-exchanger surface at $t = 150$ h. Thus, one can conclude that the crystallization fouling resulted from the mass diffusion of calcium ions was dominant in the no-treatment case.

TABLE III
WATER QUALITY DATA MEASURED FOR MAKE-UP AND CIRCULATING WATER

	Make-up	w/o PWT		w/ PWT	
	(a) Conductivity = 2000 μmhos/cm				
	a1	a2	a3	a4	a5
Conductivity (μmhos/cm)	410	1955	1910	1950	1900
pH	7.1	8.2	8.2	8.2	8.4
Total hardness (mg/L)	116	626	601	624	608
Ca^{++} hardness (mg/L)	84	448	427	444	436
Mg^{++} hardness (mg/L)	32	180	174	180	172
Total alkalinity (mg/L)	64	172	151	168	164
Chloride (mg/L)	80	436	363	432	360
LSI (at $T = 68°C$)	0.08	2.17	2.09	2.15	2.34
	(b) Conductivity = 3000 μmhos/cm				
	b1	b2	b3	b4	b5
Conductivity (μmhos/cm)	410	2950	2850	2950	2900
pH	7.1	8.2	8.0	8.2	8.3
Total hardness (mg/L)	116	982	880	980	972
Ca^{++} hardness (mg/L)	84	648	624	644	636
Mg^{++} hardness (mg/L)	32	334	256	336	336
Total alkalinity (mg/L)	64	232	220	232	228
Chloride (mg/L)	80	668	640	664	656
LSI (at $T = 68°C$)	0.08	2.40	2.17	2.40	2.49
	(c) Conductivity = 4000 μmhos/cm				
	c1	c2	c3	c4	c5
Conductivity (μmhos/cm)	420	3950	3780	3950	3800
PH	7.4	8.2	8.2	8.2	8.2
Total hardness (mg/L)	137	1496	1360	1488	1424
Ca^{++} hardness (mg/L)	114	1020	792	1012	820
Mg^{++} hardness (mg/L)	23	476	568	476	604
Total alkalinity (mg/L)	70	322	296	318	260
Chloride (mg/L)	73	864	860	864	872
LSI (at $T = 68°C$)	0.55	2.70	2.56	2.69	2.52

For the PWT case, the size of large particles was bigger, and these big particles developed tails, which were densely populated by scale crystals. This tail-shaped scale deposition could be attributed to a streamlined secondary flow behind a large particle, where flow velocity behind the particle was substantially smaller than that in the mainstream. When the cooling water was treated by PWT, the mineral ions that dissolved in the cooling water precipitated in the bulk water, and therefore, these

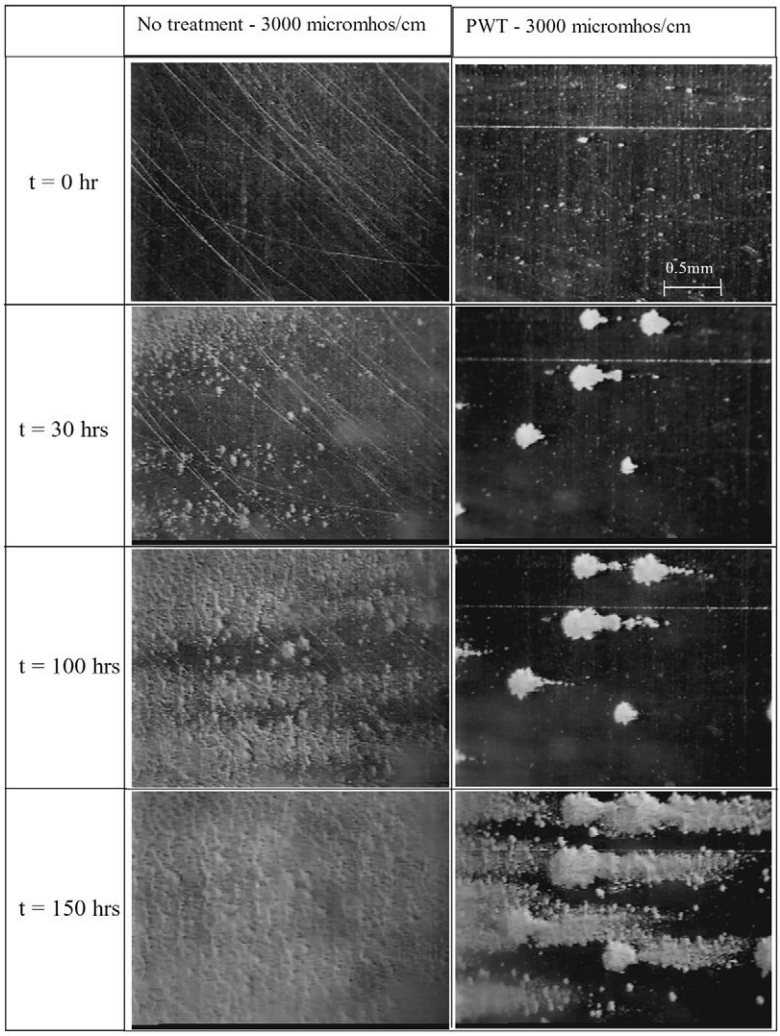


FIG. 22. Photographs of fouled surfaces obtained using a microscopic imaging system.

precipitated particles were deposited on the heat exchanger surface in the form of particulate fouling. It is speculated that particles were more difficult to be deposited on the heat transfer surface than dissolved ions, because large particles could be removed by shear force more easily than small ions. Thus, the number of the crystals in the no-treatment case was significantly greater than that in the PWT case, while the size of the

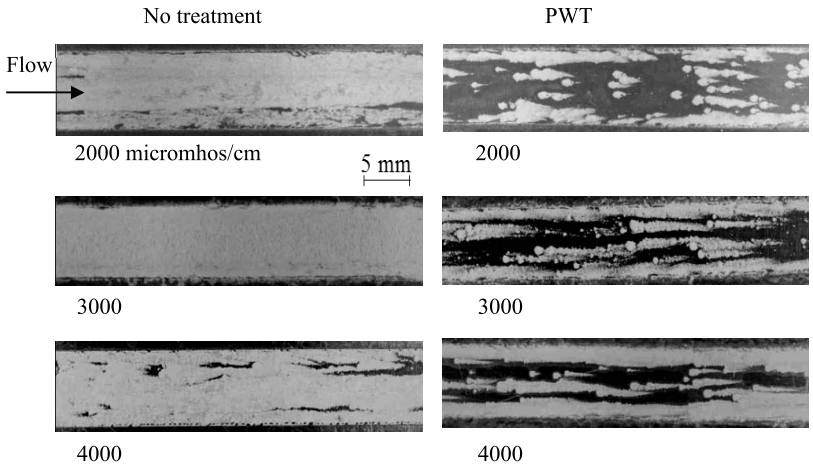


FIG. 23. Photographs of fouled surfaces taken after fouled heat-transfer surfaces were completely dried.

crystals in the no-treatment case was much smaller, due to the crystallization fouling than that in the PWT case. The present result is consistent to what Higashitani *et al.* reported earlier [45,61], who reported that magnetic treatment of water suppressed the nucleation frequency of $CaCO_3$ particles but accelerated the growth of the particles.

Figure 23 shows photographs of fouled surfaces taken after the fouled heat-transfer surfaces were completely dried. No-treatment cases resulted in much more severe fouling than PWT cases, regardless of water hardness. But in PWT cases, the amount of fouling deposit increased with increasing conductivity.

Figure 24 shows fouling resistance values vs. time for three different conductivity cases (i.e., 2000, 3000, and 4000 μmhos/cm). When one compares the three untreated cases, asymptotic fouling resistances slightly but consistently decreased with increasing concentration, a phenomenon that was not consistent with the diffusion theory of crystallization fouling. As manifested by water analyses data shown in Table IIIc (i.e., see Ca^{++} hardness), there was calcium precipitation in the case of 4000-μmhos/cm case for both untreated and PWT. One may ask why the fouling resistance for the 4000-μmhos/cm case was smaller than that for the 2000-μmhos/cm case. For the untreated 2000 μmhos/cm case, the scale deposition may be mainly due to the diffusion of dissolved mineral ions, i.e., crystallization fouling. However, for the untreated 4000 μmhos/cm case, the high level of supersaturation might have produced a large number of the seed particles in bulk, a phenomenon which could result in more particulate fouling and

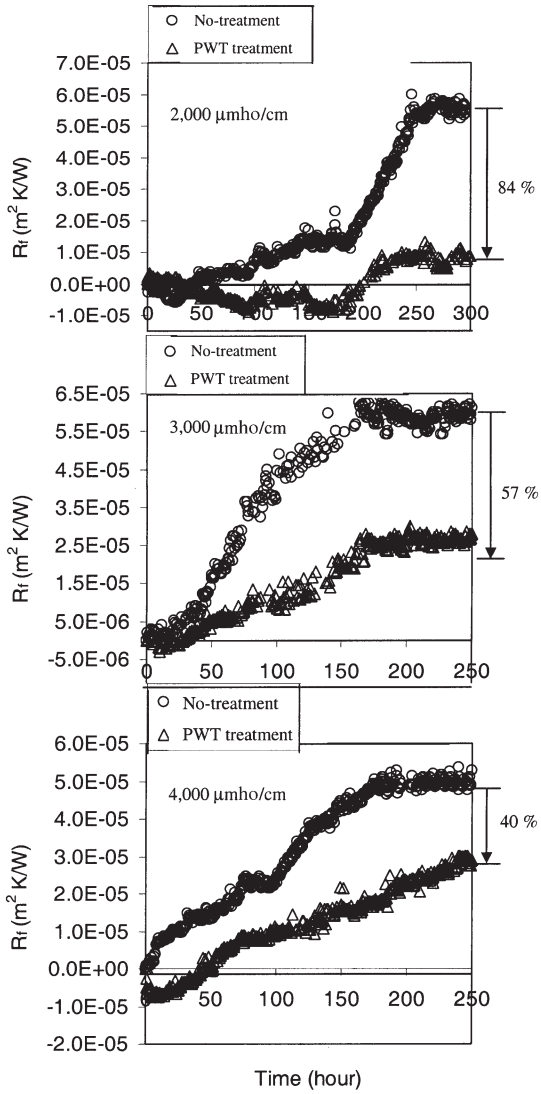


FIG. 24. Fouling resistance values vs. time for three different conductivity cases.

less crystallization fouling at the heat transfer surface. Subsequently, the soft sludge coating might have been removed by the shear produced by flow for the 4000-μmhos/cm case.

Turning to the PWT cases, the asymptotic value of the fouling resistance for the 2000-μmhos/cm case decreased by 84% from the value of the

no-treatment case, whereas for the 4000-μmhos/cm case, it decreased only by 40%. Thus, one may ask why PWT suddenly became less effective for the high concentration case. The fact that PWT became less effective at high-concentration sheds some light on the mechanism of PWT. For the 4000-μmhos/cm case, there were already a large number of seed particles produced from the high level of supersaturation. The additional production of seed particles by PWT made a less significant impact on overall fouling mitigation, rendering indirect support to the bulk precipitation hypothesis. Accordingly, one may ask what happens if the large particles produced by PWT and circulation in the cooling tower system can be removed continuously. We believe that the combined use of PWT and filtration could result in reducing the blowdown of cooling water, substantially saving water and costs. This hypothesis has to be validated in the future. Note that SEM and X-ray diffraction results obtained from the scale samples in the above cases are available elsewhere [43,66].

In summary, in the case of no-treatment, the crystallization fouling seemed to be dominant particularly at low concentration (i.e., 2000 μmhos/cm). With increasing concentration, bulk precipitation from high supersaturation produced seed particles in water, resulting in both crystallization and particulate foulings and reducing fouling resistance. PWT cases produced $CaCO_3$ particles (i.e., bulk precipitation), which grew inside the heat exchanger, eventually depositing on the heat-transfer surface in the form of particulate fouling. Real-time microscopic photographs were obtained to render further support on the hypothesis of PWT.

VII. Surface Tension Measurements

This section reports a study that was conducted to investigate whether or not PWT devices could reduce the surface tension of water. Furthermore, the study utilized visualization of the behavior of dye in water and related the dye behavior to the changes in the surface tension of water. Figure 3a shows the arrangement and dimensions of a PWT device based on permanent magnets fabricated at Drexel University (PMDU), and Fig. 4 shows a sketch of a PWT device using a solenoid coil (SCED). The SCED-control unit produced a pulsing current at a frequency of 600 Hz. Details of the operating principle of the SCED treatment including the precipitation mechanism by the induced pulsating electric field can be found elsewhere [9,44,63].

The repeated effect of PWT on the surface tension of water was studied with Philadelphia City tap water and natural hard water that was made through evaporation in a cooling-tower system. Tap water in the reservoir

tank was pumped up and passed through a flow meter and a PMDU or SCED system and finally collected in another tank at the ground level. The treatment level of PWT was varied from 0 to 30, i.e., 0 means no-treatment and 30 means the water sample passed through the PWT device 30 times. The flow velocities through the PMDU and SCED devices were 6.3 and 1.0 m/s, respectively. Note that the flow velocity of 6.3 m/s was the optimum treatment velocity for the PMDU.

After finishing each test run, the surface tension of the sample water was measured with precision glass capillary tubes, Corning Pyrex. Figure 25a shows a schematic diagram of the capillary-tube system used for surface tension measurements. A glass capillary tube was attached beside a ruler,

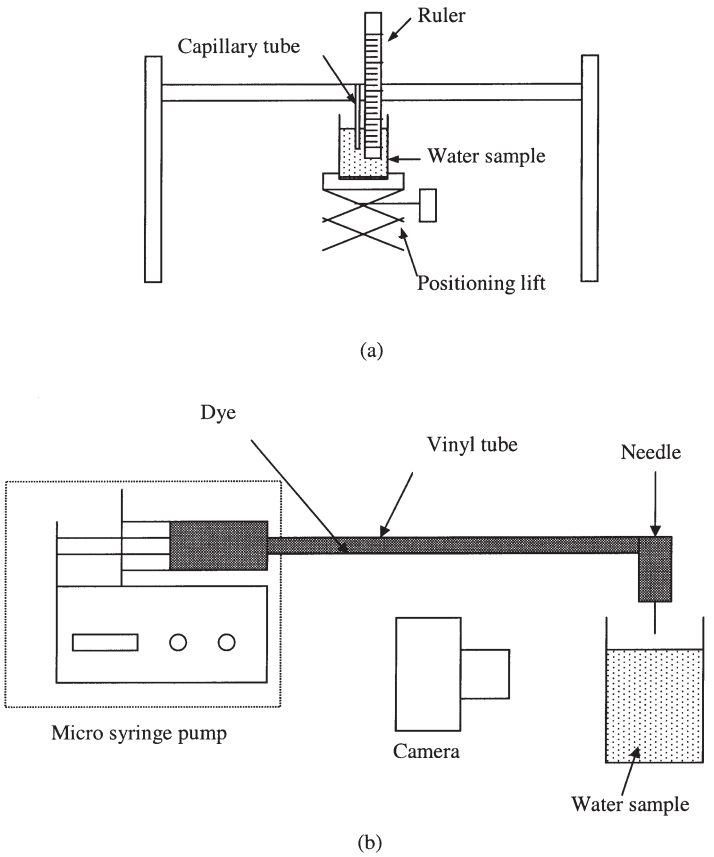


FIG. 25. (a) Schematic diagram of capillary tube system. (b) Schematic diagram of dye injection experiment system.

and then a 100-mL beaker was placed on an adjustable jack so that the capillary tube was positioned at the center of the sample water in the beaker. The capillary tube was first wetted with the sample water by raising the beaker, and then the beaker was lowered slowly. When the water level reached near the bottom of the capillary tube, a point exactly 5 mm from the bottom, the height of the water level inside the capillary tube was read. This step was repeated 10 times for each water sample and the average of the 10 measurements was used. The maximum reading error on the height of the water was estimated to be $\pm 1\%$. The dimension of the tube was 1.15 mm $\times$ 100 mm (I.D. $\times$ Length). Through the whole experiment, the temperature of the sample water was controlled at 25°C.

Figure 4b shows a schematic diagram of the system used for the present dye flow-visualization experiment, consisting of a micro-syringe pump, a needle with a vinyl tube, and a camera system. The dye flow-visualization test was performed to investigate the surface tension effect on the motion of dye. Eriochrome Black T (SG = 1.109, Aldrich Inc., from St. Louis, MO) was used for the test. A micro syringe pump (kd Scientific Co., from Holliston, MA) was used to deliver an exact amount of dye, 0.005 ± 0.0003 g, without manual interventions. The tip of the needle (21G1.5, Precision Glide) used for dye injection was always positioned at 3 mm above the surface of the sample water.

Table IV gives the result of water analysis that was performed with standard titration methods for both tap water and natural hard water that were made from a cooling-tower system. The natural hard water had 5–8 times higher total hardness than that of tap water. Before starting with a PWT device, the relationship between the effect of the surface tension and the experimental system itself was established as the baseline. After 30 passes, the surface tension dropped by 2% maximum compared to that of the initial state. This result may be attributed to the fact that pump

TABLE IV
WATER QUALITY DATA OF TAP WATER AND NATURAL HARD WATER

	Tap water	Hard water
Conductivity (μmhos/cm)	450–570	2990
pH	7.1–7.4	8.3
Total hardness (mg/L)	140–204	1190
Ca^{++} hardness (mg/L)	120–142	780
Total alkalinity (mg/L)	65–75	320
Chloride (mg/L)	75–95	640

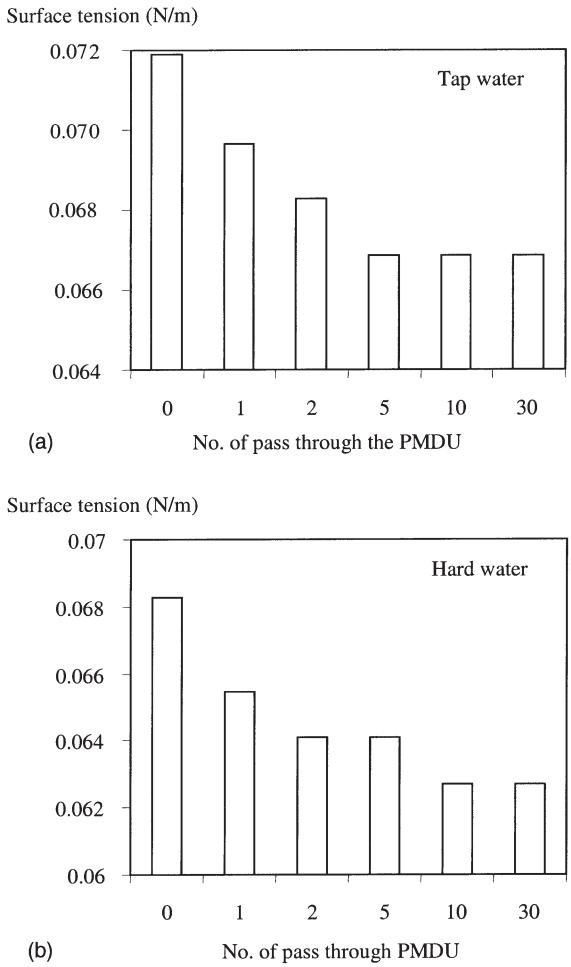


FIG. 26. (a) Results of the surface tension of tap water (570 μmhos/cm) with changing of the number of PMDU treatment and (b) for natural hard water (2990 μmhos/cm).

agitation itself might have reduced the surface tension through bulk precipitation although the amount was small.

Figure 26 shows the results of the surface tension of the sample water treated by the PMDU. In the case of the tap water (see Fig. 26(a)), the maximum reduction of the surface tension was 7.7%, and for the natural hard water case it was 8.2% (see Fig. 26(b)). This suggests that water hardness does not affect the reduction of the surface tension very significantly. A possible explanation could be as follows: when hard water

passes through the PMDU, mineral ions dissolved in the cooling-tower water collide with anionic ions such as bicarbonate and make colloidal particles in the bulk water. Therefore, as the number of passes through the PWT device increases, the number and/or the size of the colloidal particles increase, reducing the surface tension of the water.

Surface tension can be described as the surface energy per unit area. The surface energy of a liquid–liquid state is less than that of a solid–liquid state in water [76–78]. Hence, the surface energy at the interface of a water molecule and glass tube is much greater than that between two water molecules. But as the number of the colloidal particles increases in the water, the surface energy at the interface of water molecule and colloidal particle increases. In other words, the surface energy at the interface between the water molecule and glass tube decreases relatively. Therefore, the surface tension of the water will decrease as the number of colloidal particles increases in the water.

Figure 27 shows the results of the surface tension of the water sample as the water was treated repeatedly by the SCED. As the number of passes through the SCED was increased, the surface tension of the treated water sample decreased by 5.9 and 7.8% from those of the untreated tap water and natural hard water, respectively. The reduction amount of the surface tension in the natural hard water was slightly larger than that in the tap water, suggesting that the efficiency of SCED might be somehow proportional to the water hardness; in other words, the fraction of the collisions of ions that led to bulk precipitation might have increased in the hard water. Generally, the surface tension results gave positive correlations with the results of the fouling-resistance experiment [43].

Figure 28 shows photographs of the dye-injection experiment with three water samples (i.e., no treatment, 2 passes, and 10 passes through the SCED). The dye drop in the no-treatment case rapidly spread out along the radial direction; in other words, dye did not fall through the water but stayed on the top surface of the water indefinitely. But when the dye was introduced to the water sample that passed 10 times through the SCED, the dye drop quickly fell through the water as it was released from the needle and reached the bottom of the beaker in approximately 10 s. This result can be attributed to the reduction of the surface tension caused by PWT. Similar results were observed for cases with the PMDU and are available elsewhere [59].

In summary, the surface tension measurements and a dye-injection experiment were performed to examine whether or not the changes in the surface tension could be manifested qualitatively in the diffusion characteristics of the dye in water. As the PWT was repeated, the surface

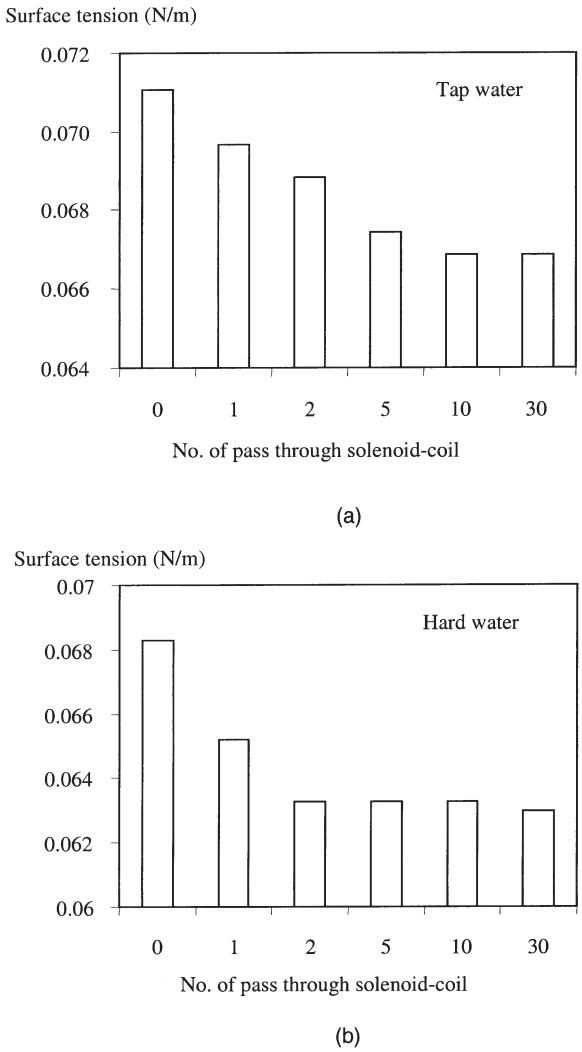


FIG. 27. (a) Results of the surface tension of tap water (450 μmhos/cm) with changing of the number of SCED treatment and (b) for natural hard water (3000 μmhos/cm).

tension of the treated water reduced by approximately 8% compared to that of the no-treatment case. The study suggests that the simple measurement of the surface tension and dye flow visualization can be used to qualitatively evaluate the efficiency of a physical water treatment method for the prevention of mineral fouling in heat exchangers.

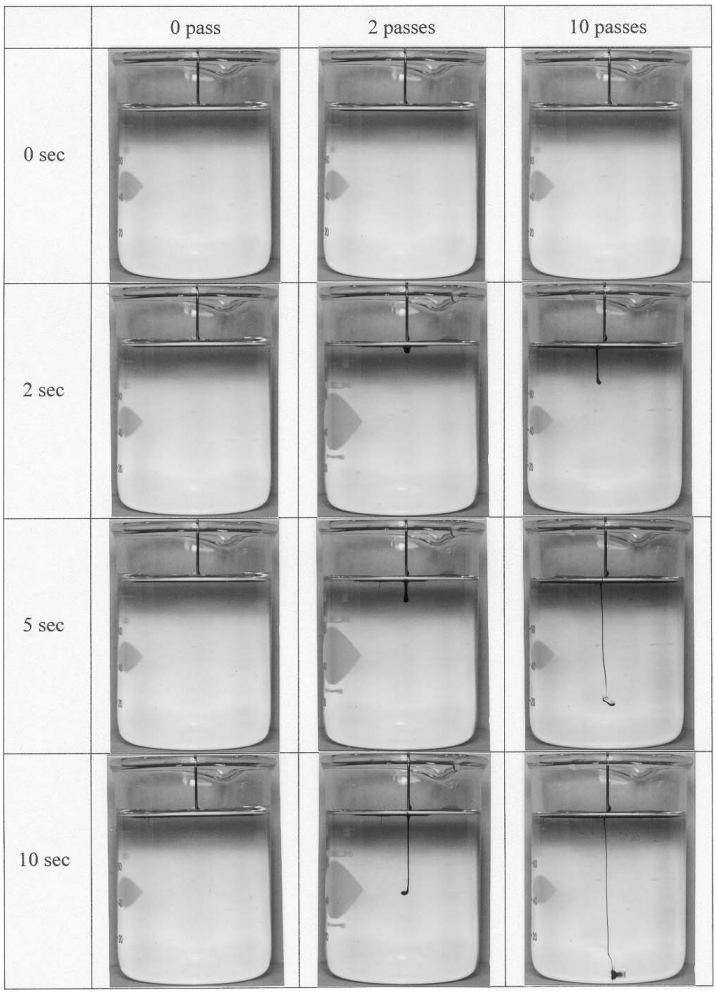


FIG. 28. Photographs of dye injection experiment with water samples treated by SCED (3000 μmhos/cm).

VIII. Needs for Multiple Magnets

There are a number of different types and designs of permanent magnets that are used for the purpose of physical water treatment. Due to vastly different designs, some permanent magnets work very well, while others do not work at all. The former represents a group of products which have been perfected through the years of trial-and-error and continuous

improvement. In order to understand the mechanism of magnetic water treatment, it is essential to understand what the critical condition is for the proper operation of the permanent magnets used. Kronenberg [28,58] reported that the efficacy of magnetic treatment of water increased as the number of magnets increased. Parsons reported at a Cranfield University symposium (UK) [10] that the most successful reports of magnetic water treatment applications occurred in continuous recirculating flow systems and under dynamic magnetic treatment, i.e., when the solution moved sufficiently rapidly through the predominantly orthogonal and alternating magnetic fields. Thus, the study investigated the effect of different arrangements of permanent magnets on the mitigation of fouling. More specifically, the study focused on finding whether one needed to have alternating magnetic fields for effective water treatment or not.

Figure 29 shows three different types of permanent magnets used in the study and the corresponding cross-sectional views of the flow channel where water was treated by permanent magnets. The strength of the magnetic field of each type was similar to each other at 0.10–0.12 T (or 1000–1200 gauss), which was measured by an AlphaLab DC Magnetometer. Except for Type I, the other four cases (i.e., Type II-A, B and Type III-A, B) used a vinyl tube of the same size (4.32 mm—I.D., 6.35 mm—O.D.) where permanent magnets were installed.

Type I represents a commercial unit, which is considered to be the optimum geometric arrangement of permanent magnets, a unit which was composed of three cylindrical magnets in the center of the device jacket with a small gap between any two magnets, see Fig. 3b. Magnetic poles are arranged as NS–SN–NS along the flow direction. Cooling water was treated by passing through the narrow annulus gap made by the permanent magnets and housing jacket. The cross-sectional dimension of the flow channel in Type I was ID = 15.7 mm and OD = 12.7 mm so that flow accelerated as it passed through the permanent magnets. Type II-A was a ring-type magnet installed around a vinyl tube filled with water. In Type II-B, two ring magnets were used, with different magnetic poles facing each other (i.e., NS–NS). Type III-A was composed of two rectangular magnets, each positioned at the top and bottom of a water-carrying vinyl tube. The two rectangular magnets had different poles. In Type III-B, two more sets of rectangular magnets were installed in a manner that magnetic fields alternate along the flow direction.

The flow velocity of water was maintained to be 1.5 m/s in the heat transfer test section, whereas the flow velocity at the permanent magnets was 2.5 m/s for all runs. The inlet temperature of the circulating water prior to the heat transfer test section was maintained at $20 \pm 1^{\circ}C$ by means of an evaporative cooling tower, whereas the inlet temperature of the hot

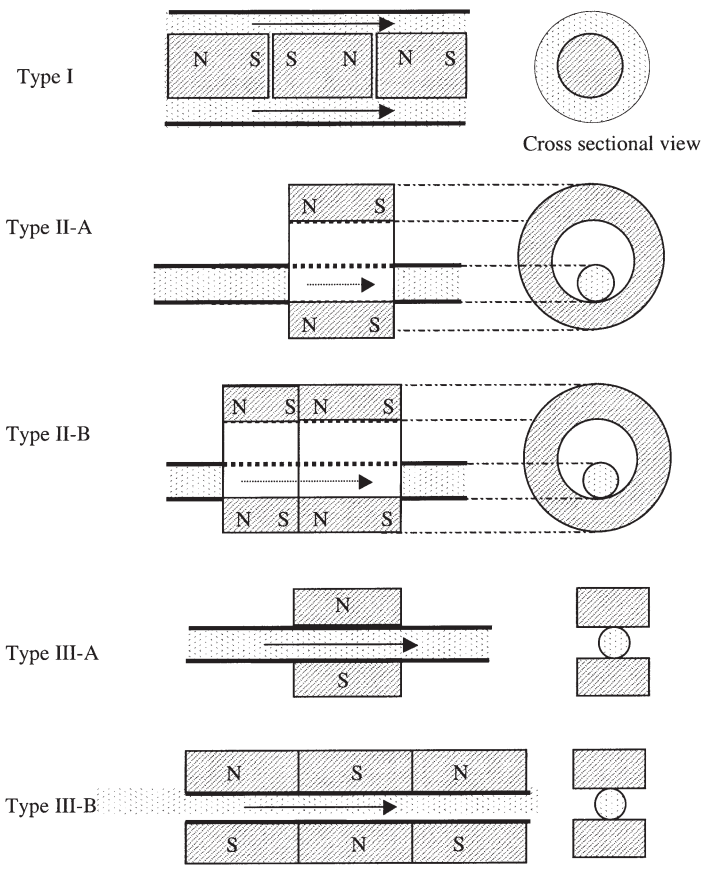


FIG. 29. Various arrangements of permanent magnets used for the purpose of the physical water treatment (PWT).

water was maintained at 89 to 92°C throughout for all the experiments. Heat flux between heating and cooling sides of the copper plate was 485 kW/m^2. As city tap water was not hard enough for fouling tests in Philadelphia, hard water was prepared by evaporating pure water using a cooling tower setting electric conductivity at 2000 μmhos/cm.

Table V shows water analysis data for the make-up, and circulating cooling-tower water for the four PWT treatment cases. There was no significant difference in water qualities between no-treatment case and the case with PWT because water quality was controlled by maintaining exactly the same electric conductivity in all tests. The Langelier saturation index (LSI) was calculated to examine the precipitation tendency of the

TABLE V

RESULTS OF WATER ANALYSES FOR MAKE-UP AND SAMPLES TAKEN FROM COOLING TOWER AT THE END OF EACH TEST

	Make-up	Type I	TypeII-A	Type II-B	Type III-A	Type III-B
pH	7.4	7.8	7.8	7.7	7.7	7.6
Sp Cond @ 25C (μmhos/cm)	599	2130	2120	2120	2130	2130
Alk, "M" as $CaCO_3$ (ppm)	65	196	196	196	196	196
Sulfur, as SO_4 (ppm)	59	172	171	173	172	172
Chloride, as Cl (ppm)	92	427	419	420	418	420
Hardness, Total, as $CaCO_3$ (ppm)	185	660	659	664	660	665
Calcium, Total, as $CaCO_3$ (ppm)	126	428	429	431	428	431
Magnesium, Total, as $CaCO_3$ (ppm)	59	231	230	232	230	234
Copper, Total, as Cu (ppm)	0.1	0.36	0.36	0.36	0.36	0.36
Iron, Total, as Fe (ppm)	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Sodium, as Na (ppm)	31	153	146	150	149	151
PO_4, Ortho, as PO_4	1.1	2.1	2.0	2.0	2.1	2.0
Silica, Total, as SiO_2 (ppm)	1.9	19.1	19.1	19.4	19.2	19.3

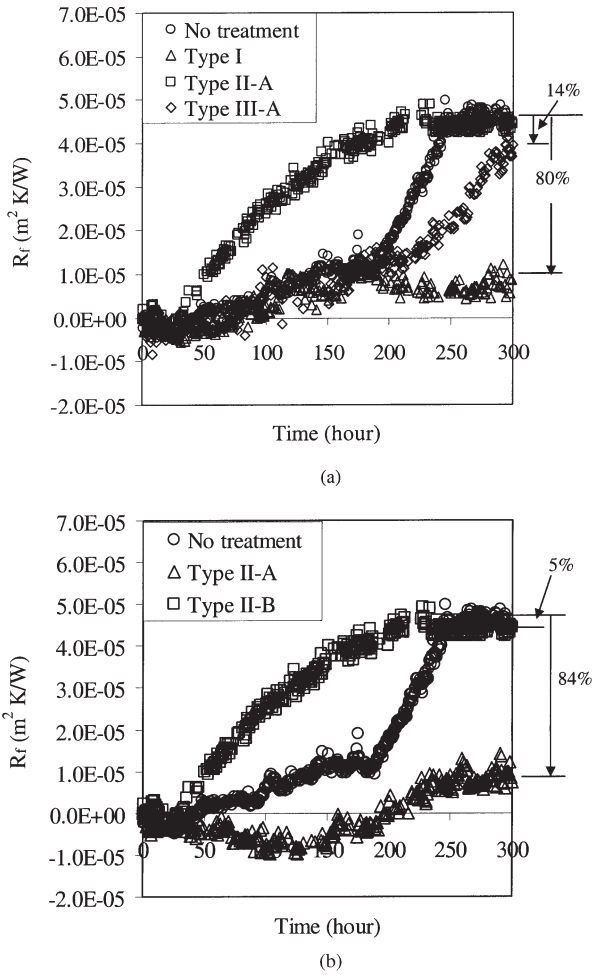
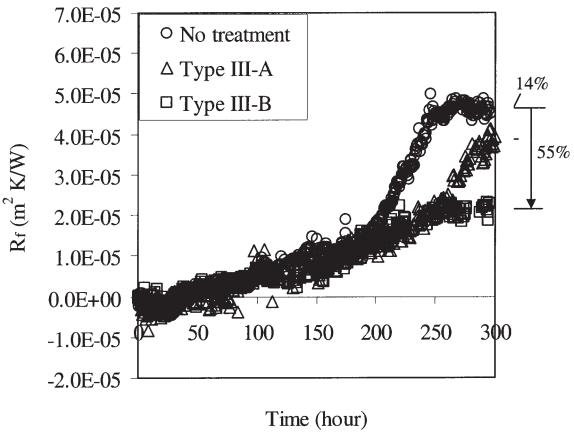


FIG. 30. Fouling resistances over time for three different permanent magnet arrangements: (a) Types I, II-A and III-A; (b)Types II-A and II-B; and (c) Types III-A and III-B.

circulating water. The values of the LSI were in the range of 1.6–1.9, indicating that the circulating water held the condition to cause severe mineral fouling.

Figure 30 represents fouling resistances vs. time for a no-treatment case and three other cases with different arrangements of permanent magnets (i.e., Type I, II-A and B, and III-A and B). With a Type I device, the fouling resistance decreased by 80% from the value of the no-treatment



(c)

FIG. 30. Continued.

case at the end of the test, see Fig. 30(a). On the contrary, the fouling resistance for the case of Type II-A increased almost at the level of no-treatment case (with negligible 5% reduction). For the case of Type III-A, the fouling resistance only reduced by 14% comparing to that from no-treatment case.

Clearly, the performance of permanent magnets critically depends on the arrangement of the permanent magnets. For the no-treatment case, scale deposition is believed to be mainly due to the diffusion and crystallization of dissolved mineral ions at the heat transfer surface, i.e., crystallization fouling. When cooling water was treated by the PWT device of Type I, mineral ions that dissolved in the cooling water were believed to be precipitated in the bulk water, and the precipitated particles grew in size inside the heat transfer test section and deposited on the surface in the form of particulate fouling. Such particulate fouling leads to soft-sludge type deposits, which can easily be removed by the shear force created by flow velocity [79–83].

Based on the fouling resistance results, one can conclude that the PWT device of Type I was effective in precipitating dissolved ions into colloidal particles in water. Hence, the reduction in the fouling resistance was great, i.e., 80%. On the other hand, the arrangements shown in the Types II-A and III-A cases were not very effective in preventing fouling, i.e., not effective in precipitating dissolved ions into colloidal particles.

Since both Types II-A and III-A did not have alternating magnetic fields along the flow direction and Type I did have them, we speculated that

the alternating magnetic fields might be necessary for the effective treatment of water. It is of note that the magnetic field and velocity vector are perpendicular in Types II and III, a phenomenon which should produce a strong induced electric field, whereas they are not perpendicular in the Type-I arrangement. Thus, we hypothesized that the performance of the Types II and III arrangements can be significantly improved if alternating magnetic fields can be provided.

To prove this hypothesis, Type II-B device was tested, which provided alternating magnetic fields. Figure 30(b) shows the results of the fouling resistance vs. time for Types II-A and B together with those of the no-treatment case for reference. Type II-B case produced a significant reduction of 84% in the fouling resistance compared with those from the no-treatment, whereas Type II-A showed only 5% reduction in the fouling resistance, indicating that the alternating magnetic fields may be required for the effective physical water treatment.

Type III-B is a modification made from Type-III-A device, where three pairs of rectangular magnets were used to provide alternating magnetic fields. Figure 30(c) shows the corresponding fouling resistance results, indicating that the Type III-B arrangement had substantially smaller fouling resistances (i.e., 55% drop) than those obtained in the Type III-A case (i.e., only 14% drop). The results also support the need of alternating magnetic fields for the effective treatment of water for the purpose of mineral fouling mitigation. The scanning electron microscopic (SEM) photographs of the scale layers for all six cases (i.e., no treatment, Type I, Type II-A, Type II-B, Type III-A, and Type III-B) are available elsewhere [66]. In general, the SEM photographs can be grouped into two broad categories from the surface morphology point of view: one includes the cases of no-treatment, Types II-A and III-A and the other includes Types I, II-B, and III-B. The Type III-B case is almost in the middle of two groups but is grouped to the latter. For the Type II-A case, the scale layer consisted of numerous tiny round shape crystals, which were agglomerated together. Each crystal size was estimated as approximately 1 μm. In the case of Type II-B, each crystal grain boundary was not clearly defined. The scale layer appeared to be a one huge irregular crystal, and this big piece of crystal was composed of several layers.

The present study investigated the effect of PWT using permanent magnets on the mitigation of mineral fouling. When the alternating magnetic fields (i.e., Types I, II-B, and III-B) were used, the asymptotic values of fouling resistance decreased by approximately 80% from the values of the no-treatment case. On the other hand, for the cases without alternating magnetic fields (i.e., Types II-A and III-A), the fouling resistance did not show any reduction compared with those of the no-treatment case.

In the cases of no-treatment and the cases without alternating magnetic fields, crystallization fouling seemed to be dominant as depicted by SEM photographs. Alternatively, the PWT cases using alternating magnetic fields produced scale crystals made of multiple layers of $CaCO_3$, which are believed to be due to particulate fouling. Real-time microscopic photographs and SEM photographs [66] were obtained to render further support to the particulate-fouling hypothesis.

IX. Summary

Physical water treatment is an attempt to treat hard water without adding chemicals to it. The purpose of PWT is to prevent or mitigate fouling so that heat exchangers can be operated without the performance degradation due to fouling. The central hypothesis of PWT was bulk precipitation, and the science behind bulk precipitation was examined. Heterogeneous catalytic reaction was proposed as the most likely mechanism of PWT, leading to bulk precipitation.

Bulk precipitation leads to particulate fouling, which produces a soft sludge coating on the heat transfer surface. Thus, for the successful operation of PWT, the ability to remove the soft sludge coating is essential. Various experimental studies were conducted to produce the supporting data for the validation of PWT.

It is recognized that a PWT device alone cannot solve all the problems associated with hard water in circulating cooling tower water applications. The PWT device (i.e., permanent magnets, solenoid coil, and others) should be used together with filtration so that the particles produced by PWT could be continuously removed by filtration. Furthermore, copper and silver ionization systems or other bio-control means could be used in conjunction with a PWT device to eliminate both mineral and bio foulings. Such a total physical water treatment method should provide a realistic alternative to chemical water treatment in the future.

References

1. Cowan, J. C. and Weintritt, D. J. (1976). "Water-Formed Scale Deposits". Gulf Publishing Company, Houston, TX.
2. Baker, J. S. and Judd, S. J. (1996). Magnetic amelioration of scale formation. *Water Res.* **30**(2), 247–260.
3. Parsons, S. A., Judd, S. J., Stephenson, T., Udol, S., and Wang, B.-L. (1997). Magnetically augmented water treatment. *Inst. Chem. Eng.* **75**(B2), 98–104.

4. Gehr, R., Zhai, Z. A., Finch, J. A., and Rao, S. R. (1995). Reduction of soluble mineral concentrations in $CaSO_4$ saturated water using a magnetic field. *Water Res.* **29**, 933–940.
5. Parsons, S. (2003). "Antiscale Magnetic Treatment". Cranfield University, www.cranfield.ac.uk/sims/water/magnets.htm.
6. Gamayunov, N. I. (1994). Action of a static magnetic field on moving solutions and suspensions. *Colloid J.* **56**, 234–241.
7. Herzog, R. E., Shi, Q., Patil, J. N., and Katz, J. L. (1989). Magnetic water treatment: the effect of iron on calcium carbonate nucleation and growth. *Langmuir* **5**(3), 861–867.
8. Busch, K. W. and Busch, M. A. (1997). Laboratory studies on physical water treatment and their relationship to a possible mechanism for scale reduction. *Desalination* **109**, 131–148; Also "Proceedings of the 45th International Water Conference". Pittsburgh, PA. October 23, 1984.
9. Cho, Y. I., *et al.* (1997). Theory of electronic anti-fouling technology to control precipitation fouling in heat exchangers. *Int. Comm. Heat Mass Transfer* **24**, 757–770.
10. Parsons, S. A. (1999). Overview of Recent Magnetic Treatment Research at Cranfield University, MAG3, Anti-scale magnetic Treatment and Physical Conditions, Cranfield University, UK.
11. Grutsch, J. F. and McClintock, J. W. (1982). Controlling scale by magnetic treatment. Report submitted to the American Petroleum Institute on Refinery Environmental Conservation.
12. Bott, T. R. (1995). "The Fouling of Heat Exchangers". Elsevier Science, New York.
13. Hasson, D., Avriel, M., Resnick, W., Rozenman, T., and Windreich, S. (1968). Mechanism of calcium carbonate scale deposition on heat-transfer surface. *Ind. Eng. Chem. Fund* **7**, 58–63.
14. Taborek, J., Aoki, T., Ritter, R. B., and Palen, J. W. (1972). Fouling: The major unresolved problem in heat transfer. *Chem. Eng. Prog.* **68**, 59–67.
15. Watkinson, A. P. and Martinez, O. (1975). Scaling of heat exchanger tubes by calcium carbonate. *Trans. ASME* **97**, 504–508.
16. Morse, R. W. and Knudsen, J. G. (1997). Effect of alkalinity on the scaling of simulated cooling tower water. *Can. J. Chem. Eng.* **55**, 272–278.
17. Somerscales, E. F. (1990). Fouling of heat transfer surfaces: an historical review. *Heat Trans. Eng.* **11**, 19–36.
18. Muller-Steinhagen, H. (1999). Cooling water fouling in heat exchangers. *In* "Advances in Heat Transfer", Vol. 33, pp. 415–496. Academic Press, New York.
19. Brezonik, P. (1994). "Chemical Kinetics and Process Dynamics in Aquatic Systems". Lewis Publishers, Chelsea, MI.
20. Snoeyink, V. L. and Jenkins, D. (1982). "Water Chemistry". Wiley, New York.
21. Beal, S. K. (1970). Deposition of particles in turbulent flow on channel or pipe walls. *Nucl. Sci. Eng.* **40**, 1–11.
22. Klassen, V. I. (Russian 1969). Magnetic water: between scylla and charybdis. *Khimiya I Zhizn* **9**, 24–27.
23. Diamant, R. M. E. (1970). Magnetic Treatment of Water. *Hospital Eng.*
24. Sandulyak, A. V. and Krivtsov, V. V. (1982). Heat and hydrodynamic conditions of magnetic treatment of water against scale formation. *Soviet Power Eng.* **5**, 362–367.
25. Grutsch, J. F. and McClintock, J. W. (1984). Corrosion and deposit control in alkaline cooling water using physical water treatment at AMOCO's Largest Refinery. *Corrosion 84* **330**, 1–26.
26. Szostak, R. J. (1985). Magnetic fluid conditioning system allows 3000 ppm hardness without cooling tower scale buildup. *Chem. Process.* 44–45.

27. Parker, D. H. (1985). ''An investigation of the role of physical water treatment devices in calcium carbonate scale formation''. MS Thesis, Baylor University, Texas.
28. Kronenberg, K. J. (1986). Physical water treatment de-mystified. *Magnets* 6–15.
29. Baker, J. S., Judd, S. J., and Parsons, S. A. (1997). Antiscale magnetic pretreatment of reverse osmosis feedwater. *Desalination* **110**, 151–166.
30. Busch, K. W., Busch, M. A., and Parker, D. H. (1986). An investigation of the role of physical water treatment devices in calcium carbonate scale formation, Baylor University, Waco, TX 1985; Laboratory studies involving physical water treatment devices, Corrosion 85, Paper No. 251, 1985; Corrosion-NACE, Vol. 42, pp. 211–221.
31. Bernardin, J. D. and Chan, S. H. (1991). Magnetic effects on simulated brine properties pertaining to magnetic water treatment, 28th National Heat Transfer Conference, HTD **164**, 109–117.
32. Serway, R. A. (1990). ''Physics for Scientists and Engineers'', 3rd edn., pp. 874–891. Saunders College Publishing, Philadelphia, PA.
33. Halliday, D., Resnick, R., and Walker, J. (1997). ''Fundamentals of Physics'', 5th edn. Wiley Interscience, New York, p. 711, 798.
34. Troup, D. H. and Richardson, J. A. (1978). Scale nucleation on a heat transfer surface and its prevention. *Chem. Eng. Commun.* **2**, 167–180.
35. Hiemenz, P. C. (1986). ''Principles of Colloid and Surface Chemistry'', 2nd edn. Marcel Dekker Inc., New York.
36. Donaldson, J. D. and Grimes, S. (1990). ''Lifting the scales from our pipes'', pp. 111–118. The City University, London, UK. *Newscientist; Tube International*, 1988.
37. Donaldson, J. D. (1988). Magnetic Treatment of Fluids-Preventing Scale. *Finishing* **12**, 22–32.
38. Kashchiev, D. (2000). ''Nucleation: Basic Theory With Applications'', 1st edn. Butterworth-Heinemann, Boston.
39. Al-Qahtani, H. (1996). Effect of magnetic treatment on gulf seawater. *Desalination* **107**, 75–81.
40. Söhnel, O. and Mullin, J. (1988). Some comments on the influence of a magnetic field on crystalline scale formation. *Chem. Ind.* 356–358.
41. Probstein, R. F. (1989). ''Physicochemical Hydrodynamics''. Butterworths, Boston, p. 186, 189.
42. Levine, I. N. (1978). ''Physical Chemistry''. McGraw-Hill, New York, p. 244.
43. Cho, Y. I., Lee, S. H., and Kim, W.T. (2003). *Physical Water Treatment for the Mitigation of Mineral Fouling in Cooling-Tower Water Applications*, ASHRAE Symposium, Chicago, IL, Jan. 2003, Paper Number CH-03–3-3 (RP-1155), 2003; ASHRAE Transactions, Vol. 109, Part 1.
44. Cho, Y. I. and Choi, B. G. (1999). Validation of an electronic anti-fouling technology in a single-tube heat exchanger. *Int. J. Heat Mass Transfer* **42**(8), 1491–1499.
45. Higashitani, K. and Oshitan, J. (1997). Measurements of magnetic effects on electrolyte by atomic force microscopy. *Trans. I ChemE (Part B)* **75**, 115–119.
46. Grutsch, J. F. (1978). Wastewater treatment: the electrical connection. *Environ. Sci. Technol.* **12**, 1022–1027.
47. Munson, B. R., Young, D. F., and Okiishi, T. H. (2002). ''Fundamentals of Fluid Mechanics'', 4th edn. Wiley, New York, p. 476.
48. Satterfield, C. N. (1991). ''Heterogeneous Catalysis in Industrial Practice'', 2nd edn. McGraw-Hill, New York.
49. Badger, W. L. and Associates (1959). *U.S. Office Saline Water Research Develp. Report* 26.
50. Charuit, P., Marchand, R., Viard, M., Courvoisier, P., and Vignet, P. (1970). 3rd Int. Symp. Fresh Water From Sea, Vol. 1, 417.

51. Springer, H. B. (1954). "Proc. 8th Cong. Int. Soc. Sugar Cane Technology". British West Indies, p. 754.
52. Aslan, G. L. (1958). *Rev. Alum* **260**, 1265.
53. Boute, C., Devienne, L., and Pourbaix, M. (1952). *Soc. Roy. Belge Ing. Ind. Mem.* **4**, 7.
54. Middis, J., Paul, S. T., Muller-Steinhagen, H. M., and Duffy, G. G. (1997). Reduction of heat transfer fouling by the addition of wood pulp fibers. *Heat Trans. Eng.* **19**, 1–9.
55. Kazi, S. N., Duffy, G. G., and Chen, X. D. (2002). Fiber-modified scaling in heat transfer fouling mitigation. *Chem. Eng. Commun.* **189**, 742–758.
56. Weintritt, D. J. and Cowan, J. C. (1967). Unique characteristics of barium sulfate scale. *J. Petroleum Technol.* **19**, 1381–1384.
57. Wagner, I. (1999). The DVGW Standard W 512—A tool for testing the efficiency of physical water conditioners, MAG3, Anti-scale magnetic treatment and physical conditions, Cranfield University, UK.
58. Kronenberg, K. J. (1985). Experimental evidence for effects of magnetic fields on moving water. *IEEE Trans. Mag., MAG-21* 2059–2061.
59. Lee, S. H. (2002). A study of physical water treatment technology to mitigate the mineral fouling in a heat exchanger, Ph.D. Thesis, Drexel University, Philadelphia, PA, USA.
60. Barrett, S. and Parsons, R. A. (1999). A review of some of the analytical methods used to access the influence of magnetic treatment, Mag3, Anti-scale magnetic treatment and physical conditions, Cranfield University, UK.
61. Higashitani, K., Kage, A., Katamura, S., Imai, K., and Hatade, S. (1993). Effects of magnetic field on formation of $CaCO_3$ particles. *J. Colloid Interface Sci.* **156**, 90–95.
62. Fan, C. (1997). A study of electronic descaling technology to control precipitation fouling, Ph.D. Thesis. Drexel University, Philadelphia, PA, USA.
63. Cho, Y. I., Fan, C., and Choi, B. G. (1998). Use of electronic anti-fouling technology with filtration to prevent fouling in a heat exchanger. *Int. J. Heat Mass Transfer* **41**, 2961–2966.
64. Cho, Y. I., Choi, B. G., and Drazner, B. J. (1998). Electronic anti-fouling technology to mitigate precipitation fouling in plate-and-frame heat exchangers. *Int. J. Heat Mass Transfer* **41**, 2565–2571.
65. Cho, Y. I. and Choi, B. G. (1999). Validation of an electronic anti-fouling technology in a single-tube heat exchanger. *Int. J. Heat Mass Transfer* **42**, 1491–1499.
66. Cho, Y. I. (2002). Final Report, ASHRAE (available from www.ashrae.org; Bookstore, keyword = magnetic water treatment).
67. Morse, R. W. and Knudsen, J. G. (1997). Effect of alkalinity on the scaling of simulated cooling tower water. *Can. J. Chem. Eng.* **55**, 272–278.
68. Kim, N. H. and Webb, R. L. (1991). Particulate fouling of water in tubes having a two-dimensional roughness geometry. *Int. J. Heat Mass Transfer* **34**, 2727–2738.
69. Helalizadeh, A., Muller-Steinhagen, H., and Jamialahmadi, M. (2000). Mixed salt crystallization fouling. *Chem. Eng. Process.* **39**, 29–43.
70. Vermeiren, T. (1958). Magnetic treatment of liquids for scale and corrosion prevention. *Corrosion Technol.* 215–219; also cited as Private Communications in Troup, D. H. and Richardson, J. A. (1978). Scale nucleation on a heat transfer surface and its prevention. *Chem. Eng. Commun.* **2**, 167–180.
71. Ebadian, M. A. and Dong, Z. F. (1998). "Handbook of Heat Transfer" (Rohsenow *et al.* ed.), 3rd edn., Chapter 5, pp. 5.106–5.107. McGraw-Hill, New York.
72. Hasson, D. and Bramson, D. (1985). Effectiveness of magnetic water treatment in suppressing calcium carbonate. *Ind. Eng. Chem. Proc. Des. Dev.* **24**, 588–592.
73. Alleman, J. E. (1985). Quantitative assessment of the effectiveness of permanent magnetic water conditioning devices, *Purdue Univ. Civil Eng. Res. Report*, Contract 0651–57–1284.

74. Kim, W. T. (2001). A study of physical water treatment methods for the mitigation of mineral fouling, Ph.D. Thesis. Drexel University, Philadelphia, PA, USA.
75. Lee, S. H. and Cho, Y. I. (2002). Velocity effect on electronic-antifouling technology to mitigate mineral fouling in enhanced-tube heat exchanger. *Int. J. Heat Mass Transfer* **45**, 4163–4174.
76. Lower, S. K. (1997). Solids in contact with natural waters. *A Chem1 Reference Text*, Simon Fraser University, 21–24.
77. Mysels, K. J. (1976). ''Introduction to Colloid Chemistry'', pp. 184–187. Interscience Publishers, Inc. NY.
78. Vold, R. D. and Vold, M. J. (1983). ''Colloid and Interface Chemistry'', pp. 116–148, 223–259. Addison-Wesley Publishing Company, Inc.
79. Liu, R. and Cho, Y. I. (1999). Combined use of an electronic antifouling technology and brush punching for scale removal in a water-cooled plain tube. *Exp. Heat Transfer* **12**, 203–213.
80. Cho, Y. I. and Liu, Rong (1999). Control of fouling in a spirally-ribbed water chilled tube with electronic anti-fouling technology. *Int. J. Heat Mass Transfer* **42**, 3037–3046.
81. Cho, Y. I. and Taylor, W. T. (1999). An innovative electronic descaling technology for scale prevention in a chiller, Ashrae Transaction, Symposia 1999, Vol. 105, Part 2, SE-99–3-1, pp. 581–586.
82. Cho, Y. I., Liu, R., McFarland, W. J., and Fusegni, L. (2000). Study of scale-removal methods in a double-pipe heat exchanger. *Heat Transfer Eng.* **21**, 50–57.
83. Kim, W. T., Cho, Y. I., and Bai, C. (2001). Effect of electronic anti-fouling treatment on fouling mitigation with circulating cooling tower water. *Int. Commun. Heat Mass Transfer* **28**, 671–680.

A New Microporous Surface Coating for Enhancement of Pool and Flow Boiling Heat Transfer

SEUNG M. YOU[1], KEVIN N. RAINEY[2], and CURTT N. AMMERMAN[3]

[1]*The University of Texas at Arlington, Department of Mechanical Aerospace Engineering, Arlington, TX 76019-0023, USA; Tel.: 817-272-5635; E-mail: you@uta.edu*
[2]*Los Alamos National Laboratory, P.O. Box 1663, MS-P940, Los Alamos, NM 87545, USA*
[3]*Los Alamos National Laboratory, P.O. Box 1663, MS-H821, Los Alamos, NM 87545, USA*

I. Introduction

The ever-increasing heat flux from electronic devices has driven the electronic cooling community toward ebullient cooling technologies. A prominent cooling scheme for accommodating high power dissipation is immersion cooling with dielectric fluids, such as 3M's Flourinert series and DuPont's Freons. Boiling, rather than single-phase heat transfer, is becoming necessary because of the high heat fluxes generated from advanced electronic devices. Boiling is attractive for this application because it produces much higher heat transfer coefficients and can yield a far more uniform temperature distribution across the surface of a device and/or an array of devices than typical single-phase cooling schemes. However, some of the relevant physical and thermal properties of most suitable dielectric fluids are generally poor from a boiling standpoint (highly wetting, low contact angle, low specific heat), therefore, boiling enhancement schemes are required in order to utilize these cooling techniques for electronic devices.

Surface microstructure enhancement is one method that can be used to enhance boiling heat transfer. One of the earliest methods used to produce an enhanced boiling surface microstructure was to roughen the surface using sandpaper or some other abrasive process. Surface roughening has been proven to significantly enhance the nucleate pool boiling heat transfer

Advances in Heat Transfer
Volume 38 ISSN 0065-2717
DOI 10.1016/S0065-2717(04)38002-0

in water, organic fluids, cryogenics, and refrigerants. Kurihara and Myers [1], as well as Benjamin and Balakrishnan [2], showed that the nucleate boiling enhancement was the result of increased active nucleation site density. With regard to critical heat flux (CHF), the effects of surface roughening are much less clear. Using data for various fluids, Ramilison *et al.* [3] showed that surface roughness influences CHF. The quenching data of Chowdhury and Winterton [4] appears to support these findings although the authors do not mention it. On the other hand, Berenson [5], Nishio and Chandratilleke [6], and Sakurai *et al.* [7] found no significant effect of surface roughening on CHF.

Under flow boiling conditions, Yu *et al.* [8] found that the heat transfer coefficient increased with increasing roughness. Alternatively, Kandlikar and Spiesman [9] found a correlation between surface cavity number and size distribution with increased nucleate boiling performance, but that it did not necessarily correspond to surface roughness (although their roughest surface performed better). Additionally, Groenveld [10] found that CHF was increased with roughness in subcooled flow boiling, which was attributed to increased mixing at the wall [11]. Due to severe aging effects, however, surface roughening has received little sustained interest. Chaudri and McDougall [12] measured the long-term nucleate pool boiling performance of an abrasive-treated tube and found only temporary benefits. In addition, Nishikawa *et al.* [13] observed that the effects of surface roughness diminish with increasing pressure.

Griffith and Wallis [14] showed that the geometry of a microcavity containing trapped vapor was directly related to the bubble nucleation process. Most notably, they found that re-entrant type cavities were stable, easily activated boiling sites. From their study (as well as others), many commercial enhanced surfaces utilizing re-entrant type cavities or grooves have been developed, such as Furukawa's ECR-40, Wieland's GEWA, Union Carbide's High-Flux, Hitachi's Thermoexcel, and Wolverine's Turbo-B. Experiments by many researchers have shown that these and other surfaces can enhance pool boiling heat transfer and CHF on various geometries and in various fluids, including highly wetting fluids. Thome [15,16], Bergles [17], Webb [18,19] and Webb and Chien [20] provide nice reviews of pool boiling on these surfaces. Thome [16] concluded that the primary enhancement mechanisms for re-entrant-type enhanced surfaces were enhanced nucleation from the larger embryonic bubbles, increased thin film evaporation due to the large internal surface area of the porous structure, and two-phase convection within the porous structure. Chien and Webb [21] performed visualization experiments on re-entrant cavities and found that for saturated boiling, an active tunnel was vapor-filled with liquid menisci in the corners, indicating that the primary

boiling mechanism within the cavity was evaporation, which supported the suction-evaporation boiling model described by Nakayama *et al*. [22].

Under flow boiling conditions, Czikk *et al*. [23] reported that the heat transfer coefficient for a High-Flux internally coated tube was typically an order of magnitude greater than that for smooth-tube data using liquid oxygen, ammonia, and R-22. Wadekar [24] also showed a significant enhancement using the High-Flux surface. Using a tube internally coated with 115-μm copper particles, Ikeuchi *et al*. [25] reported a factor of five enhancement of the heat transfer coefficient compared to a smooth tube in R-22. Similar enhancements have been reported for axial microfin/microgroove flow channels [26,27]. Celato and Cumo [28] provide a review for the enhancement of flow boiling CHF from these surfaces. In addition, Mizunuma *et al*. [29], Maddox and Mudawar [30], and Kandlikar and Howell [31] found that both nucleate boiling and CHF were enhanced for discrete surfaces in flow boiling when incorporating a microfinned surface.

Although the structured/porous enhancement surfaces described previously could significantly enhance nucleate boiling heat transfer and CHF, their application processes are generally not benign (e.g., sintering, flame-spraying, machining) and many of them still suffer from excessive incipience superheat in highly wetting fluids [17]. Tests by Tong *et al*. [32] have shown that the embryonic bubbles required to initiate boiling when using highly wetting dielectric fluids are very small (on the order of 0.1 μm). The feature sizes of most of the enhancements schemes listed previously are too large to effectively trap vapor for highly wetting dielectric fluids. In addition, these surface treatments may not be easily applied to an electronic chip because they could alter or destroy the chip surface. To address these problems, several researchers have studied smaller-scale enhancement techniques such as graphite-fiber/metal matrices, various etching techniques, chemical vapor deposition, microfins, dendritic structures, and microporous coatings. Yang *et al*. [33] tested a copper/graphite-fiber matrix surface in pool boiling of R-113. This type of surface was found to reduce the hysteresis at incipience and enhance nucleate boiling heat transfer over plain surfaces [34]. In addition, various micro-pin-fin geometries have been used successfully to enhance nucleate boiling heat transfer and CHF [35–37]. Honda and Wei [38] provide a good review of various small-scale-feature enhancement surfaces.

The focus of this chapter is to describe the development and subsequent analysis of another small-scale-feature surface—the microporous coating. Following the discoveries of You *et al*. [39], the microporous coating has developed into an enhancement technique that is benign enough to apply directly to an electronic chip surface. The microporous coating provides a significant enhancement of nucleate boiling heat transfer and CHF while

reducing incipient hysteresis. The coating is a surface treatment used to increase vapor/gas entrapment volume and active nucleation site density by forming a porous structure with cavities much smaller than conventional metallic porous coatings. In the following sections, the development, optimization, and performance of the microporous coating are reported.

II. Nucleate Boiling Heat Transfer and Highly Wetting Fluids

A. Nucleate Boiling Heat Transfer

Figure 1 shows a typical pool boiling curve. Four different regimes are identified: the natural convection regime (A–B), the nucleate boiling regime (C–D), the transitional boiling regime (D–E), and the film boiling regime (E–F). The end point of the natural convection regime (point B) is the incipient point, or onset of nucleate boiling (ONB), where the first bubbles form on the heated surface. The upper end point of the nucleate boiling regime (point D) is the critical heat flux (CHF), or the onset of film boiling. For a heat-flux-controlled boiling curve, the transitional regime is skipped and the boiling curve passes directly from point D to point F at CHF. The transitional boiling regime between points D and E can only be obtained when the heating surface is temperature-controlled.

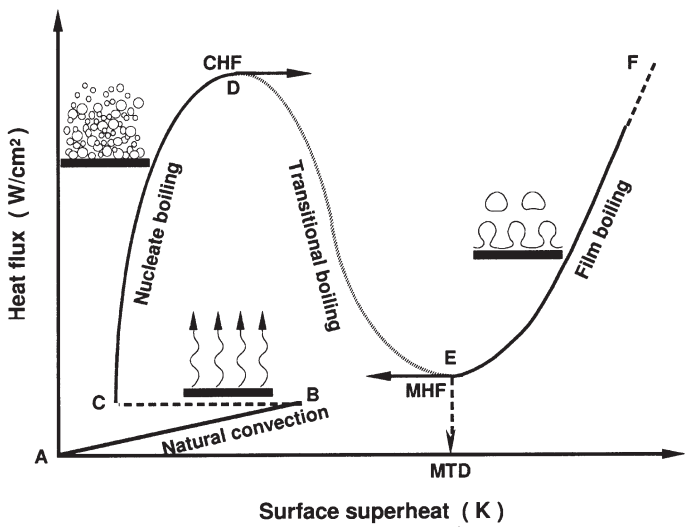


FIG. 1. Typical boiling curve.

1. *Incipience*

For most systems, boiling initiation occurs through the heterogeneous nucleation process. In this process, small amounts of air or vapor are trapped in tiny cavities (surface irregularities) on the heating surface by the advancing fluid front. These trapped gas pockets act as embryos in the formation of boiling bubbles. Figure 2 illustrates the gas entrapment process.

As heat is applied to the surface, a superheated liquid layer is formed. This superheated layer drives fluid vaporization at the embryonic bubble vapor/liquid interface. If enough heat is applied, the embryonic bubble will grow until it detaches and a new bubble will grow in its place. The amount of liquid superheat required to initiate bubble growth from an embryonic bubble depends on the fluid properties and the embryonic bubble size. Smaller embryonic bubbles require larger superheats to initiate growth due to increased bubble vapor pressure. The pressure inside a bubble can be estimated from

$$P_b - P_l = \frac{2\sigma}{r_b} \tag{1}$$

where P_b is the bubble pressure and is equal to the sum of the partial pressures of vapor and gas within the bubble ($P_{vb}+P_{gb}$); P_l is the liquid pressure and for saturation conditions is equal to the saturation pressure at the bulk liquid temperature, $P_{\text{sat}}(T_l)$; σ is the surface tension; and r_b is the bubble radius. The pressure inside a smaller embryonic bubble must be higher to achieve a balance with the surface tension force. A higher bubble pressure generally means that a higher superheat is needed to cause vaporization at the bubble vapor/liquid interface (higher pressure results in higher saturation temperature).

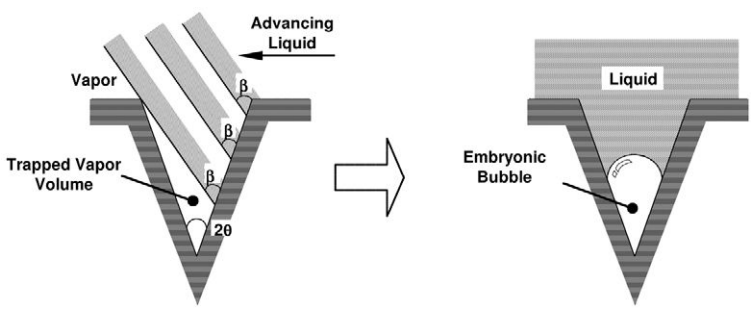


FIG. 2. Vapor/gas entrapment process.

2. *Nucleate Boiling*

The heat transfer mechanisms associated with pool boiling are liquid convection and latent heat transfer. Liquid convection could be further subdivided into natural convection (where nucleation does not occur), microconvection (local fluid motion induced by bubble departure), and Marangoni convection (surface-tension-driven flow over a growing bubble). In the nucleate boiling regime, the two dominant mechanisms are latent heat and microconvection. Marangoni convection becomes significant only when subcooled conditions are present, and the contribution of natural convection is usually very small during nucleate boiling.

Flow boiling refers to the situation where boiling occurs on a surface immersed in a flowing fluid. The main difference between pool and flow boiling is that natural convection is replaced with forced convection. Many researchers [40,41] have found that the nucleate flow boiling performance is insensitive to both flow velocity and liquid subcooling. Although forced convection significantly affects the single-phase convection regime, its effect generally diminishes in the nucleate boiling regimes as nucleate boiling is usually a dominant heat transfer mechanism.

3. *Critical Heat Flux*

CHF is the transition point from the highly efficient nucleate boiling regime to the very inefficient, and usually damaging, film boiling regime. When CHF is reached in a heat-flux-controlled boiling situation, vapor completely covers the heating surface. This vapor blanketing causes the surface temperature to rise dramatically due to the added thermal resistance of the vapor layer. Zuber [42] postulated that CHF was reached when the vapor leaving the heater surface became hydrodynamically unstable. He developed the following correlation to predict CHF under saturated pool boiling conditions:

$$q''_{\mathrm{CHF},Z} = \left(\frac{\pi}{24}\right) h_{lv} \rho_v^{1/2} [g\sigma(\rho_l - \rho_v)]^{1/4} \tag{2}$$

where h_{lv} is the enthalpy of vaporization, ρ_v and ρ_l are the vapor and liquid densities, and g is the acceleration due to gravity. Zuber's correlation has been found to predict CHF well for plain surfaces, however, it does not take into account surface microstructure (e.g., nucleation site density), which appears to significantly affect CHF [43]. (CHF is generally low in highly wetting fluids due to their poor thermal properties.)

Therefore, there are generally three important parameters that should be considered when evaluating boiling heat transfer performance: incipient

superheat, nucleate boiling efficiency, and CHF. Regardless of the application or fluid used, it is generally desirable to minimize the wall superheat required to initiate boiling to produce less thermal stress, to enhance the nucleate boiling heat transfer efficiency, and to increase CHF to give a larger heat flux operating range.

B. Boiling Enhancement

The enhancement of nucleate boiling has been studied extensively. Methods employed can be categorized in four general areas: surface enhancement (e.g., coatings, surface roughening), area enhancement (e.g., fins), modifying fluid conditions (e.g., subcooling, forced convection, nano-particles), and external manipulation (e.g., mechanical vibration, electrohydrodynamics (EHD)). Surface enhancement is generally used to artificially increase the active nucleation site density, which has been found to significantly improve heat transfer efficiency. This can be done by simply roughening a surface, mechanically structuring the surface, or applying a relatively thin porous coating. Area enhancement is used to increase the available heat transfer surface area without significantly increasing the overall size. Fins are the most common method used. Modifying fluid conditions can significantly affect incipience and CHF, which determine the usable heat flux operating range. Bubble manipulation methods such as vibration and EHD enhance heat transfer by improving the vapor removal mechanism.

1. *Pressure, Subcooling, and Dissolved Gas*

Increasing pressure and increasing liquid subcooling are two methods that can be used to enhance nucleate boiling heat transfer efficiency. In addition, most highly wetting fluids of interest tend to absorb large amounts of non-condensable gas under subcooled conditions, which can also enhance boiling efficiency.

It can be seen from Eq. (1) that the pressure inside the bubble must be higher than that of the surrounding liquid to balance the surface tension force. This is why a superheated liquid layer near the heating surface is required to initiate bubble growth regardless of the fluid conditions. Increasing the system pressure would, at first glance, seem to have no effect on incipience, however, increasing the pressure increases the saturation temperature. An increased saturation temperature generally causes a reduction in surface tension, which in turn reduces the bubble pressure required to maintain the force balance. A system pressure increase, therefore, has the effect of reducing the superheat required for incipience, as was observed by You *et al*. [44].

Prior to discussing the effects of liquid subcooling, some terms need to be defined. The term "pure subcooled" refers to a completely degassed, subcooled fluid condition. "Gas saturated" is used to describe the condition of the fluid as being saturated with non-condensable gas. "Gassy subcooled" is used to describe the fluid condition near a heater surface during boiling, which contains an indeterminable amount of dissolved gas—a condition that is somewhere between gas-saturated and pure-subcooled. The liquid subcooling level is defined as $T_{sat}(P_{sys})-T_{bulk}$. For pure subcooled conditions, P_{sys} equals the fluid vapor pressure, P_v, while for gas-saturated conditions, P_{sys} equals the sum of the fluid vapor and non-condensable gas partial pressures, or $P_{sys} = P_v + P_g$ [45].

With regard to pure subcooling, decreasing the bulk liquid temperature (T_l) apparently has little or no effect on boiling incipience [44]. Decreasing T_l (while keeping P_l constant) does not change the saturation temperature. Therefore, the bubble surface tension, and hence P_b, remains the same and the superheat required for boiling initiation remains unchanged.

The effects of dissolved non-condensable gas can also be seen from Eq. (1). Adding dissolved gas ($P_{gb} \neq 0$) while keeping everything else constant has the effect of reducing the vapor pressure within the bubble (P_{vb}), which reduces the wall superheat required to activate the bubble [46].

C. Highly Wetting Fluids and Re-Entrant Cavities

The term "highly wetting fluid" refers to a class of liquids that exhibit very low surface tension and near-zero static contact angle on most surfaces. Examples are 3M's Fluorinert liquids, Dupont's Freons, and Solvay Solexis' perfluoropolyether liquids. These liquids are generally inert, dielectric, non-flammable, and can have low boiling points making them desirable for use in direct immersion cooling schemes for electronics. Their poor thermal properties and highly wetting characteristics, however, interfere with boiling efficiency. Table I lists some physical and thermal properties of some of these liquids along with water for comparison. (Bar-Cohen [47] expands on nucleate pool boiling behavior in highly wetting fluids.)

As described in Section II.A.1, incipience superheat is directly related to the initial embryonic bubble size. The size of the embryonic bubble depends on the size and shape of the cavity as well as the dynamic contact angle at the vapor/liquid/solid interface [32]. Theoretically, a larger cavity will result in a larger embryonic bubble. Highly wetting fluids, however, which have small dynamic contact angles, will displace most of the trapped vapor from large cavities when wetting the surface (Fig. 2). With highly wetting fluids, therefore, the resulting small embryonic bubbles produce a pronounced incipient superheat excursion under increasing heat flux conditions. If the

TABLE I
SATURATED PROPERTY COMPARISON OF SELECTED HIGHLY WETTING FLUIDS AND WATER AT 1 atm

Property	FC-72	FC-87	R-123	Water
Chemical formula	C_6F_{14}	C_5F_{12}	$CHCl_2CF_3$	H_2O
Boiling point [K]	329.7	302.3	301	373.15
Density—liquid [kg/m^3]	1598.9	1571.6	1456	958.3
Density—vapor [kg/m^3]	13.37	12.38	6.46	0.597
Specific heat—liquid [J/kg · K]	1121	1103	982	4220
Latent heat [kJ/kg]	94.280	98.97	170	2256.7
Thermal conductivity—liquid [W/m · K]	0.0545	0.055	0.080	0.679
Viscosity—liquid [kg/m · s]	4.47E-4	6.69E-4	4.21E-4	2.78E-4
Surface tension [mN/m]	8.10	8.91	14.83	58.91

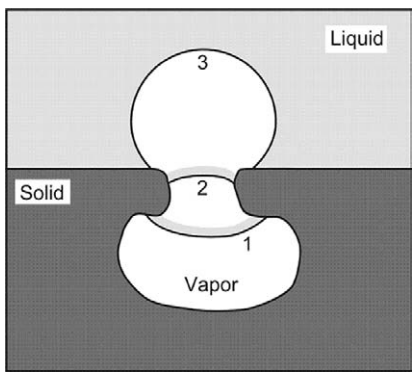


FIG. 3. Growth of bubble from a re-entrant-type cavity.

embryonic bubble size could be artificially increased, the superheat excursion could be significantly reduced.

A re-entrant cavity can be used to artificially increase the embryonic bubble size. Figure 3 shows a representation of a re-entrant cavity. The concept of a re-entrant cavity is to increase the cavity volume without increasing the cavity opening so as to maximize the amount of gas/vapor entrapped by the wetting front. The boiling performance of highly wetting fluids, therefore, can be greatly enhanced through the use of re-entrant cavities. Re-entrant cavities can be formed through mechanical enhancement processes as well as porous enhancement coatings [15].

III. Development of the Microporous Surface Coating

A. Particle Layering Technique

In 1992, You *et al.* [39] realized the need for a boiling enhancement technique that was both effective with highly wetting fluids and benign enough to be applied to bare semiconductor chips. They developed a method for depositing 0.3 and 0.3- to 3-μm alumina (Al_2O_3) particles on the surfaces of heating elements. The alumina particles were sprayed onto a 1-cm^2 alumina-coated, platinum heating element using a compressed-air nebulizer. The particles impinged on the surface via a $\sim$1-m/s air jet. The particles adhered to the surface by what was believed to be van der Waals forces. Figure 4 shows a scanning electron microscope (SEM) image of the 0.3- to 3-μm particles deposited on the heating element.

Subsequent pool boiling tests in saturated FC-72 showed that particle layering enhanced boiling heat transfer by increasing the number of nucleation sites and by the generation of re-entrant cavities. Pool boiling curves for the two different particle layered surfaces are shown in Fig. 5 compared with data from an untreated reference surface. The particle layering increased CHF values by 36% for the 0.3-μm particles and by 39% for the 0.3- to 3-μm particles. Wall superheat values in nucleate boiling decreased by approximately 37% and 46% for the 0.3-μm and 0.3- to 3-μm particle surfaces, respectively. In addition, the incipient superheat was reduced by

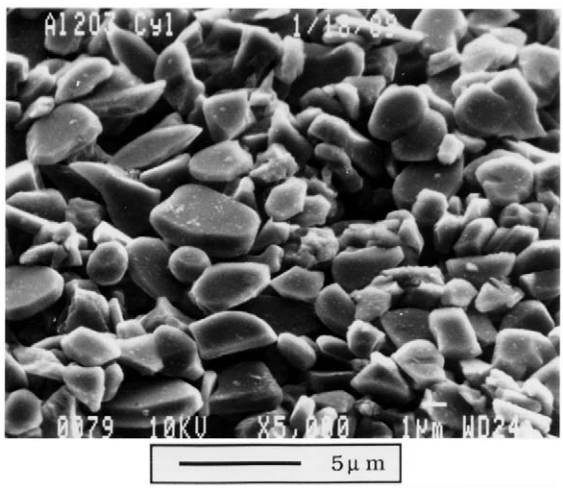


FIG. 4. Surface microgeometry taken by SEM; 0.3- to 3-μm-alumina-particle-layered surface.

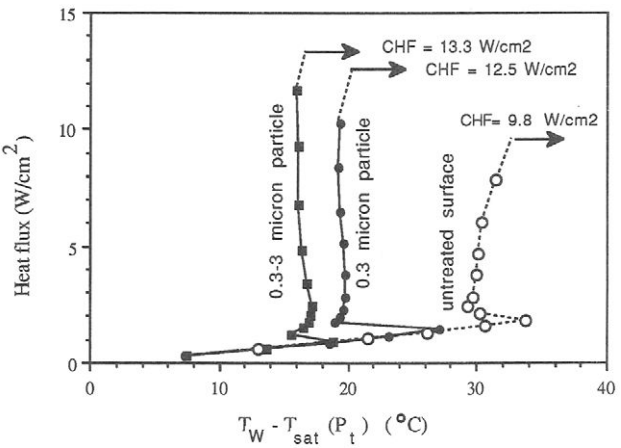


FIG. 5. Pool boiling curves, effect of particle layering.

about 5°C for the 0.3-μm particles and by about 15°C for the 0.3- to 3-μm particles relative to the reference surface.

Although pool boiling heat transfer was considerably improved by using the particle layering technique, stronger particle attachment mechanisms were needed for the technique to be widely applicable (the particles could be easily wiped off). To increase particle attachment strength, a coating of glue was applied to the heater surface prior to spraying the particles. After spraying the particles, the glue was cured with ultraviolet light creating a relatively permanent coating. Pool boiling tests were again performed showing that the 0.3- to 3-μm particle coating with glue did not perform as well as the coating without glue, but was still better than the reference surface. Comparing SEM photos of the coated surfaces with and without the glue revealed that the glue eliminated many of the small-scale structures. It could be deduced from these results, therefore, that larger particles were needed when using glue.

B. Silver Flake Paint

In 1995, O'Connor and You [48] presented the results of enhancement studies in which an epoxy was mixed with silver flake particles and applied to a heating surface. This new paint mixture both dramatically improved boiling enhancement and significantly increased the particle attachment strength. More specifically, O'Connor and You combined 3- to 10-μm silver flakes, epoxy, and isopropyl alcohol to form their paint mixture. The paint

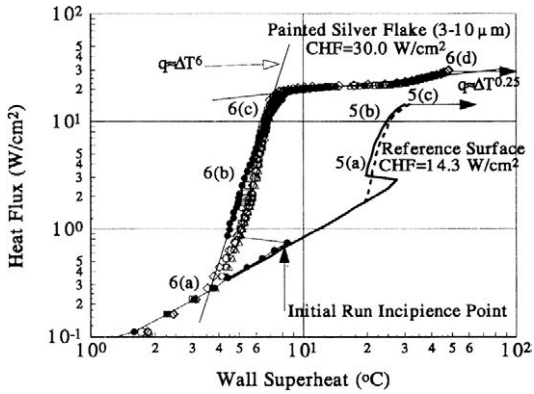


FIG. 6. Pool boiling curves, effect of silver flake paint.

was applied to a heating element using a "drip" application method resulting in a cured paint thickness of approximately 25 μm.

The resulting pool boiling curves for this coating tested in saturated FC-72 are shown in Fig. 6 compared with a plain reference surface. The coated surface significantly enhanced nucleate boiling along with incipient superheat and CHF. Photographic observations revealed that the painted structure consistently produced a larger number of smaller bubbles than the plain surface throughout the entire boiling curve, including conditions just prior to CHF. This phenomenon was unique and led to enhanced boiling heat transfer mechanisms. Incipient superheat values for the reference surface ranged from approximately 20 to 27°C, whereas for the enhanced surface they ranged from approximately 3 to 8.5°C. This difference in incipient superheat can be attributed to larger embryonic bubbles in the re-entrant cavities provided by the surface microstructure. At 4 W/cm^2, bubble departure diameters for the reference surface ranged from 0.6 to 1.0 mm, however for the painted surface, departure diameters ranged from 0.4 to 0.8 mm. The smaller bubble departure size along with an increased number of nucleation sites and bubble departure frequency resulted in a reduction in nucleate boiling superheats of approximately 70%. In addition, the enhanced-surface CHF was 109% higher than for the reference surface.

C. MICROPOROUS COATING

The microporous coating subsequently evolved into a three-component mixture of particles, binder, and carrier based on the silver flake paint by

TABLE II
BASELINE COMPOSITIONS FOR COATING DEVELOPMENT STUDY

Paint	Particles (g)	Particle size (μm)	Binder (ml)	Carrier (ml)
ABM	1.5	1–20	0.4	10
CBM	1.5	1–50	0.4	10
DBM	1.5	8–12	0.4	10
DOA	1.5	8–12	0.1	10
SOA	1.5	3–10	0.1	10

O'Connor and You [48]. The particles provide the structural material to create the required re-entrant cavities. The binder, or glue, binds the particles together and to the heating surface. Since the particles are solid and the binder is usually very viscous, a volatile carrier liquid is used to allow for adequate mixing of the particles and binder, and for application of the particles and binder to the surface. The carrier evaporates after application.

In 1997, Chang and You [49] performed a coating development study to investigate the impact of particle type, binder, and carrier on boiling heat transfer performance. Their baseline coating compositions are shown in Table II. The coatings are identified as "ABM", "CBM", "DBM", "DOA", and "SOA", named using letters of the three components of each paint (particles: A = aluminum particles, C = copper particles, D = diamond particles, S = silver flakes; binder: B = Devcon brushable ceramic epoxy, O = Omegabond 101 epoxy; carrier: A = isopropyl alcohol, M = methyl ethyl ketone). These coatings were painted on copper reference surfaces using the drip application method.

Figure 7 shows SEM images of top views of the different coatings described in Table II. For the diamond coatings (DOA and DBM), nearly identical SEM images were observed, therefore only one image is shown representing both coatings (Fig. 7(c)). Aluminum particle sizes in the ABM coating (1–20 μm) are smaller than those for the CBM coating (1–50 μm); however, the large copper particles within the CBM coating have additional cavity structures on their surfaces with sizes on the order of microns. The photograph of the DOA (or DBM) coating shows more uniform particle shapes and sizes of diamond (8–12 μm). The SOA coating has randomly oriented layers of silver flakes (3–10 μm). Since a thin layer of epoxy is evenly distributed over all particles, the particles are clearly visible from the SEM images. The coatings are multilayered,

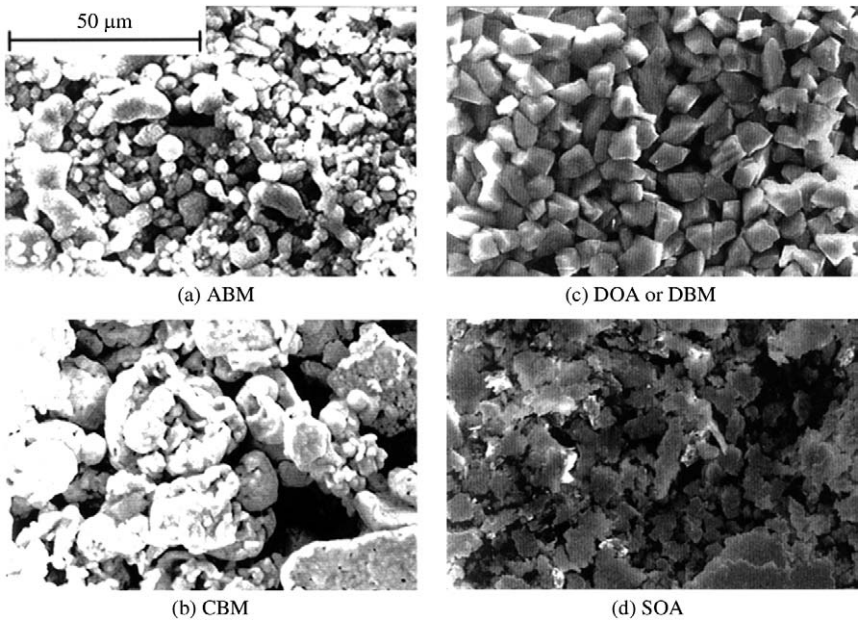


FIG. 7. SEM images of microporous surfaces.

microporous structures having thicknesses of approximately 100 μm for the CBM coating, 50 μm for the ABM, DOA, and DBM coatings, and 30 μm for the SOA coating. All of the coatings form numerous re-entrant cavity structures.

In Fig. 8, the boiling curves for surfaces using these five coatings are compared with that of an uncoated reference surface. The incipient superheats for the coated surfaces were about 80% lower than that of the reference surface. At incipience, all the microporous-enhanced surfaces produced nucleate boiling over the entire surface area, generating discrete tiny bubbles less than 0.2 mm in diameter. At similar heat fluxes, the uncoated reference surface experienced up to about 10 times higher wall superheats than the coated surfaces. Throughout the nucleate boiling regime, the coated surfaces consistently augmented heat transfer coefficients by approximately 330% over those of the reference surface. This enhancement was the result of dramatically increased active nucleation site density produced by the coatings' surface microstructures. The microporous enhancement coatings produced repeatable CHF values with $\approx 100\%$ enhancement over the reference surface.

The five coatings in Table II produced nearly identical pool boiling enhancement. Figure 8 indicates that the active cavity size and density

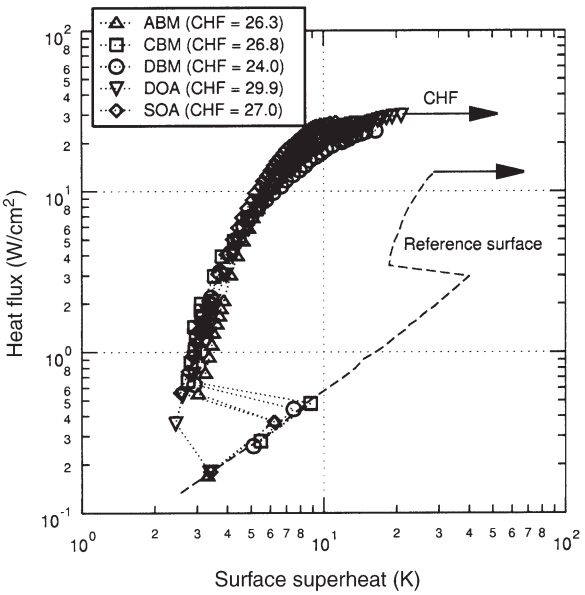


FIG. 8. Effects of boiling enhancement coatings.

generated by the coatings match each other well. The next section investigates performance optimization of the microporous coating by examining the effects of coating thickness, particle size, mixture component ratios, and application method.

IV. Optimization of the Microporous Coating

A. Coating Thickness

O'Connor *et al.* [50] studied the effect of the microporous coating thickness on nucleate boiling performance. Their results are summarized in Fig. 9. O'Connor *et al.* used a single paint mixture (1.5 g of 8- to 12-μm diamond particles, 0.1 ml of Omegabond 101 epoxy, and 10 ml of isopropyl alcohol) and varied the thickness by applying successive layers onto a reference surface (also shown in Fig. 9). The coating was painted onto the heater surface using the drip application method. Paint thicknesses of 40, 80, 120, and 170 μm were tested in a heat flux range between 5 and 19 W/cm^2 in a pool of saturated FC-72 at 1 atm.

As shown in Fig. 9, the wall superheat increases with increasing paint thickness, as expected. The superheats were subsequently corrected to take

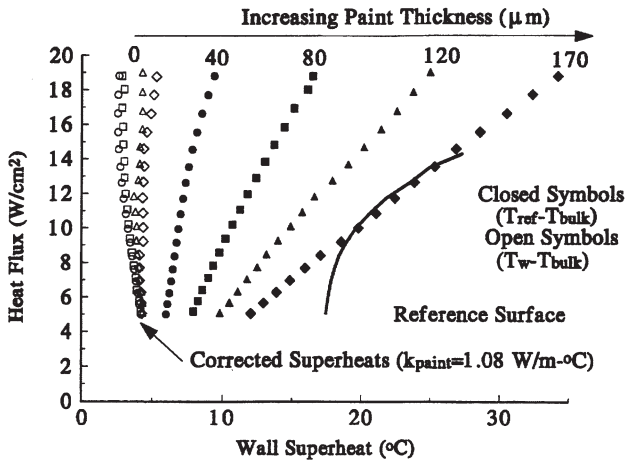


FIG. 9. Diamond (8–12 μm) paint thickness study.

into account the thermal conduction resistance of the coating (open symbols in Fig. 9). Since the corrected data collapsed together, it appeared that the increased heater surface temperatures were due to the thermal conduction resistance of the coating only and not to any degradation of nucleation enhancement. Based on these results, ≈ 40 μm was determined as the optimum paint thickness for this study. The data in Fig. 9 indicate a coating thickness less than 40 μm may be better, however, 40 μm was the practical lower limit using the drip method of application. Full heater surface coverage was too difficult to achieve with this method for thinner coatings.

B. PARTICLE SIZE

Chang and You [51] examined the effects of particle size on boiling performance of microporous enhanced surfaces using five different sizes of diamond particles in a DOA coating. Diamond particles were used for this study because they were available in a wide range of sizes with uniform size distributions. Pool boiling tests were conducted in saturated FC-72 at 1 atm.

Table III shows the average particle size (d_m), coating thickness (t_c), thickness-to-particle-size ratio (t_c/d_m), and powder porosity (ε) of each coating used in the study. Each coating thickness was maintained at around 5 times its mean particle diameter, except for the coatings with $d_m = 2$ and 70 μm where the resulting thickness ratio was $t_c/d_m \approx 15$ and 3.6, respectively. SEM images of top views of the five coatings are shown

TABLE III
PARTICLE SIZES USED FOR DOA COATINGS

Particle size (d_m)	Coating thickness (t_c)	t_c/d_m	Powder porosity (ε)
2 μm ± 1	30 μm ± 10	≈15	48%
10 μm ± 2	50 μm ± 20	≈5	41%
20 μm ± 3	100 μm ± 30	≈5	40%
45 μm ± 5	200 μm ± 50	≈4.4	40%
70 μm ± 10	250 μm ± 50	≈3.6	47%

in Fig. 10. All of the particles appear to have similar polygon shapes. The coatings were generated using the drip application method.

The porosity of a powder is defined as the ratio of porous volume to an effective powder volume (particle and porous volume combined). The porosity data tells how much of the volume is vacant within the powder. The porosity of these coatings was calculated using the following equation:

$$\varepsilon = \frac{V_p - (m_p/\rho_p)}{V_p} \tag{3}$$

where V_p, m_p, and ρ_p are the volume, mass, and density of the powder, respectively. All five cases shown in Table III have essentially the same porosity.

The boiling performance of these coatings was found to be dependent on the ratio of coating thickness to particle size. In addition, the use of epoxy to bond particles to surface and particles to each other created additional thermal resistance in the coating. O'Neill *et al.* [52] proposed a nucleate boiling model for a porous coated surface. This model was based upon the higher thermal conductivity value (≈400 W/m · K) of sintered copper particles employed in their High-Flux surface. In their model, a thin, superheated liquid film was assumed to exist on the surface of each particle stacked within the porous layer. They proposed that heat is transferred by conduction through the particle matrix and then by conduction across the thin liquid film where evaporation occurs. For the epoxy-based coatings of Chang and You, however, the coating layer is not a good thermal conductor. The low thermal conductivity value is thought to be due to the epoxy component within the coating (as shown in Section IV.A), which impedes conduction between the micron-size particles. Based on this low thermal conductivity, a superheated layer is assumed to develop from the heater base surface through the epoxy-based coating layer.

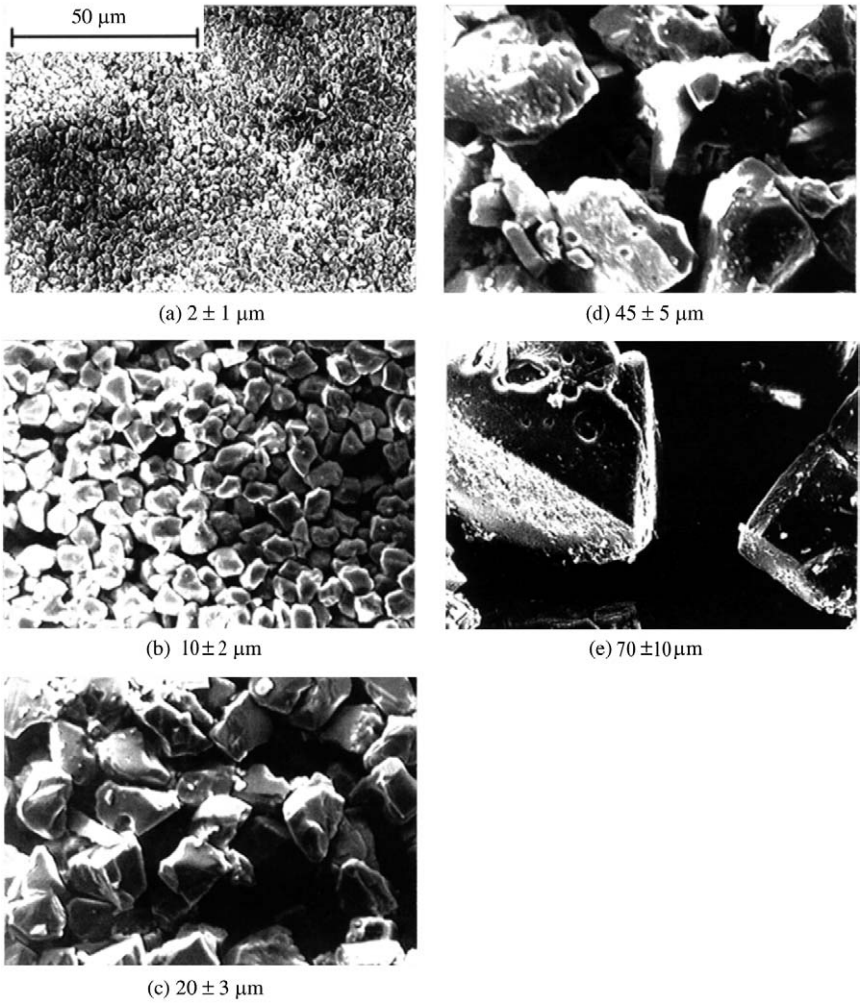


FIG. 10. SEM images of DOA-coated surfaces (top views).

A transient thermal boundary layer concept by Hsu [53] was used to determine activation of cavities during nucleation in the coatings in Table III. Based on Hsu's model, the superheated liquid layer thickness, δ_{99}, was estimated for the present test liquid (FC-72). A transient conduction solution [54] for a constant-heat-flux surface in an infinite medium was generated assuming the bubble departure frequency was 46 Hz [55]. From this analysis, the superheated liquid layer thickness was computed to be approximately 100 μm. This estimation of superheated liquid layer thickness

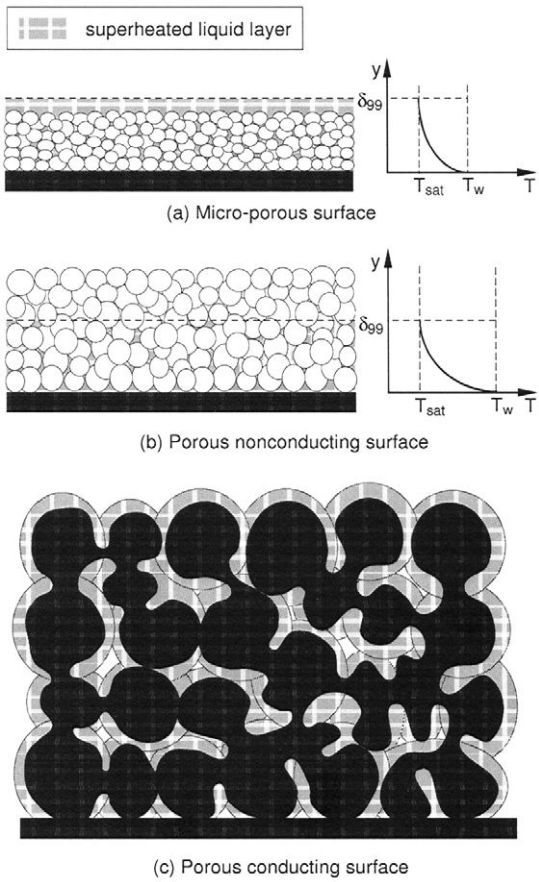


FIG. 11. Theoretical models of microporous and porous coated surfaces. (a) Microporous surface, (b) porous non-conducting surface, and (c) porous conducting surface.

indicated that for the coatings in Table III whose thickness was greater than 100 μm, only the lower portion of the coating layer is activated during nucleate boiling.

By comparing coating thicknesses with the estimated δ_{99}, the coatings in Table III were classified into two groups: "microporous" and "porous". Figures 11(a) and (b) show theoretical models of microporous and porous surfaces, respectively. The microporous coating thicknesses are less than or equal to δ_{99}, and the porous coating thicknesses are greater than δ_{99}. Chang and You called their epoxy-based porous surface a "porous non-conducting surface" in contrast to the highly conductive coating of O'Neill *et al.*, which was called a "porous conducting surface". Figure 11(c)

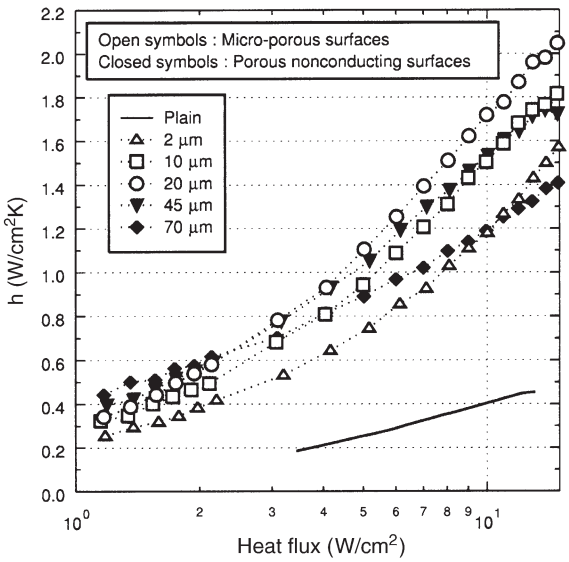


FIG. 12. Particle size effects on nucleate boiling enhancement (microporous, $t_c < \delta_{99}$ and porous non-conducting, $t_c > \delta_{99}$).

shows a theoretical model of the porous conducting surface where a superheated liquid layer develops over each particle in the porous structure.

Nucleate boiling performance characteristics of the microporous and the porous non-conducting coatings shown in Table III are compared in Fig. 12 for heat fluxes below 15 W/cm^2. A similar value of slope is observed for all the microporous surfaces, and decreased values of slope are observed for the porous non-conducting surfaces. Heat transfer is enhanced as particle size (and hence, coating thickness) is increased for the microporous structures. The microporous surfaces produce larger and more active nucleation sites as the coating thickness and pore size are increased, as long as the coating thicknesses are smaller than δ_{99}.

The porous non-conducting surfaces (with larger particle sizes and thicker coatings) produced higher heat transfer coefficients than the microporous surfaces below heat fluxes of ≈ 2.5 W/cm^2. In this lower heat flux range, discrete vapor bubbles were observed to depart from the coating surfaces (as opposed to larger merged vapor bubbles departing above ≈ 2.5 W/cm^2). These less-crowded vapor bubbles can escape the porous coating layer with little hydraulic resistance.

The increase in heat transfer coefficient (h) for the porous non-conducting surfaces over the discrete vapor bubble regime is thought to be

due to increased active nucleation site density. At heat fluxes higher than ≈ 2.5 W/cm^2, this trend was reversed: lower h was observed with larger particle size coatings. This decrease in nucleate boiling performance can be attributed to the penalty associated with the increased coating thickness in two ways. First, liquid supply to the innermost active portion ($< \delta_{99}$) of the porous non-conducting coating is hindered due to additional hydraulic resistance over liquid/vapor exchange channels, reducing cooling performance. Secondly, the increase of coating thickness produces additional thermal resistance due to the low thermal conductivity of the coating layer, thus increasing the temperature drop at higher heat fluxes.

The variations in incipient superheat for the five different particle sizes are shown in Fig. 13. Each data point in Fig. 13 is a compilation of five or more runs from each of three different heaters. As the particle size is increased in the microporous coating regime ($0 < d_m < 20$ μm), the incipient superheat decreases significantly. Above $d_m = 20$ μm, incipient superheats were relatively insensitive to particle size.

Bar-Cohen and Simon [56] discussed a correlation for incipient wall superheat required to initiate boiling in a highly wetting liquid using Eq. (1). Using vapor pressure values corresponding to the measured incipient superheats in Fig. 13, the embryonic bubble radii were estimated to have ranges of 0.047 to 0.13 μm, 0.14 to 0.61 μm, 0.47 to 1.1 μm, and 0.79 to

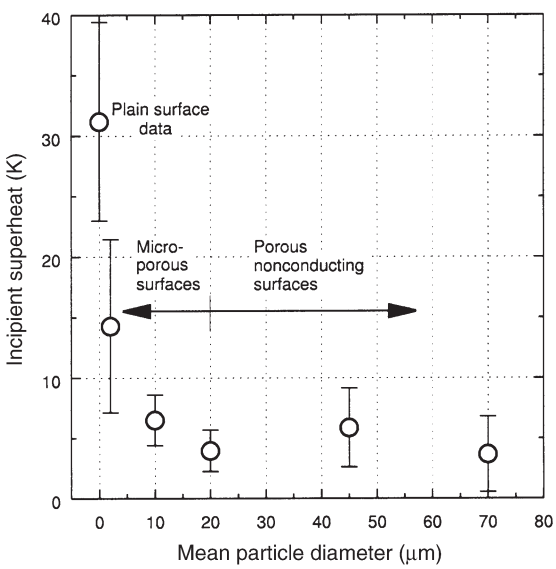


FIG. 13. Incipient superheat vs. particle diameter.

2.3 μm for $d_m = 0$, 2, 10, and 20 μm, respectively. These estimated embryonic bubble sizes were comparable to the cavity sizes observed from the surfaces in Fig. 10. As particle size was increased further, similar ranges of incipient superheat (< 10 K) were observed. As previously illustrated by Bankoff [57], vapor/gas entrapment from the advancing liquid film over surface cavities is strongly affected by the liquid contact angle and the cavity geometry. For the porous non-conducting surfaces, it can be assumed that a similar size of cavity geometry with the microporous surfaces was used to capture embryonic bubbles, whereas larger cavity geometries ($d_m > \approx 2$ μm) were flooded with the contact of FC-72 liquid film.

CHF variations for the five coatings are shown in Fig. 14. The coatings consistently show enhanced CHF values over the uncoated reference surface. An empirical approach was employed to correlate CHF data with the coating parameters. Since the porosities of the five coatings were similar (Table III), the CHF data in Fig. 14 are correlated solely as a function of mean particle diameter, d_m:

$$q''_{\mathrm{CHF}} = 20.62 + 3.11 \ln(d_m) = 0.25 [\ln(d_m)]^2 - 0.016 [\ln(d_m)]^3 \qquad (4)$$

For the fitted line in Fig. 14, CHF increases dramatically in the micro-porous surface regime, and then levels off in the porous non-conducting surface regime. Surface area enhancements were ruled out by observing no

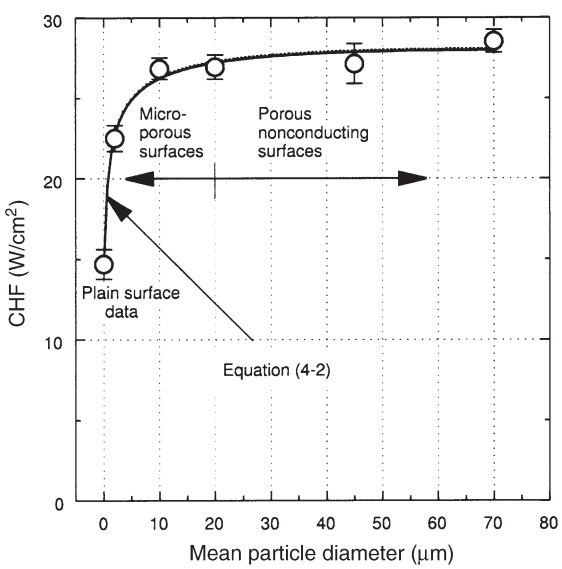


FIG. 14. CHF vs. particle diameter.

differences between natural convection data for the coated and uncoated (reference) surfaces. The increase in CHF with these coatings, therefore, is due to the enhancement of the active nucleation site density within the coating layer, which effectively delays the onset of film boiling. This enhancement mechanism seems to reach a limit within the porous non-conducting regime.

C. Mixing Ratio

The ABM and CBM coatings in Table II were investigated to study mixing ratio variations. Two parameters were used to describe the different mixing ratio variations: porosity and effective volume ratio. The effective volume ratio (EVR) is defined as the ratio of the effective powder volume to the binder (epoxy) volume used in manufacturing the coating. The EVR data show a volume mixing ratio between the powder and the binder and are inversely proportional to the amount of binder contained in the paint. The EVR of these coatings was calculated using the following equation:

$$\mathrm{EVR} = \frac{(m_p/\rho_p)/(1-\varepsilon)}{V_b} \tag{5}$$

where V_b is the binder volume.

Table IV shows the three-composition paints used in this mixture-ratio study. The same porosity of 70% was measured for both the aluminum powder (1–20 μm) and the copper powder (1–50 μm). To examine the effects of different EVR values, the weight of the particle content and the volumetric MEK content were fixed at 1.5 g and 10 ml, respectively, and

TABLE IV
ABM and CBM Coatings with Different Mixture Ratios

Product	Particle	Binder	Carrier	EVR	No. of tests
ABM	Aluminum:	0.3 ml	MEK (10 ml)	6.2	5
	1 to 20 μm (1.5 g)	0.4 ml		4.6	7
		0.5 ml		3.7	5
		0.7 ml		2.6	4
		0.9 ml		2.1	4
CBM	Copper:	0.1 ml	MEK (10 ml)	5.6	4
	1 to 50 μm (1.5 g)	0.3 ml		1.9	5
		0.4 ml		1.4	5
		0.5 ml		1.1	4

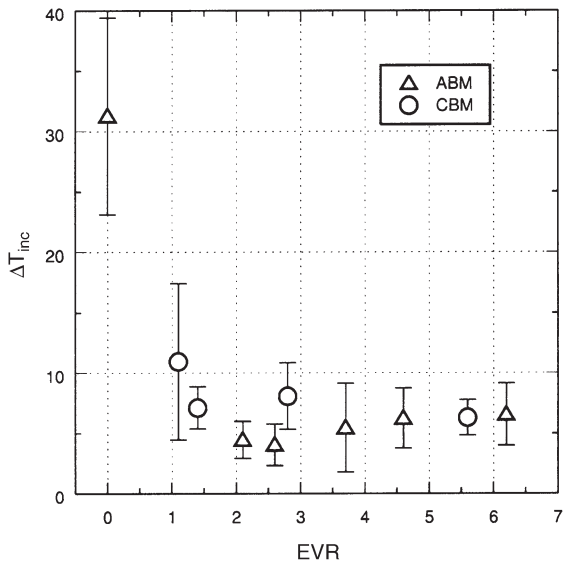


FIG. 15. Incipient superheat vs. EVR.

the volumetric content of the binder was varied. Five ABM coatings with five different EVR values (6.2, 4.6, 3.7, 2.6, and 2.1) and four CBM coatings with four different EVR values (5.6, 1.9, 1.4, and 1.1) were manufactured. Table IV shows the different mixture ratios and resulting EVR values that were studied. Between 4 and 7 different heater surfaces were used to test each respective EVR value. The measured coating thicknesses were ≈ 50 and ≈ 100 μm for the ABM and CBM coatings, respectively. All boiling tests were carried out in saturated FC-72 at 1 atm.

Figure 15 illustrates incipient superheat as a function of EVR. The ABM and CBM coatings show a similar trend of incipient superheat variation with EVR. The smaller the EVR value, the larger the amount of binder contained within the paint. The zero EVR value represents no particles included within the paint. The uncoated reference surface was used for the zero EVR case, which produced the highest superheat of 31.2 ± 8.2 K. The data in Fig. 15 indicate that the incipient superheats of the microporous-enhanced surfaces show a minimal impact from EVR at EVR values greater than about 1.1.

Heat transfer from the microporous surfaces in Table IV and the reference surface were compared under fully developed nucleate boiling in FC-72. Figure 16 shows enhancements of the heat transfer coefficients for heat fluxes of 5 and 10 W/cm^2. The microporous surfaces produced enhanced heat transfer coefficients that were 3 to 4.5 times larger than that

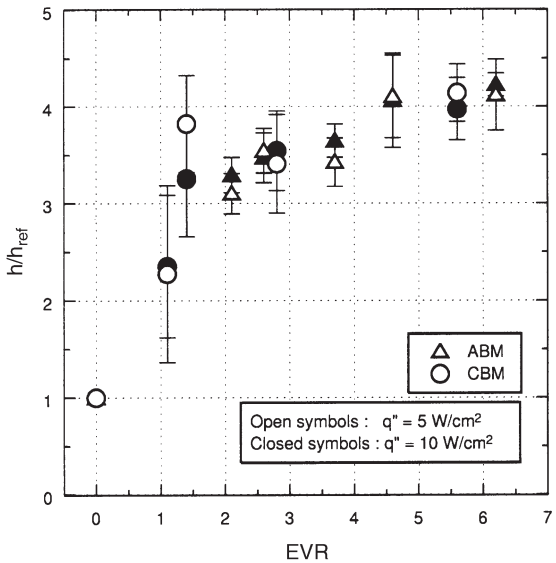


FIG. 16. Heat transfer coefficient enhancement vs. EVR.

of the uncoated surface at EVR conditions greater than 1.1. Pool boiling curves at EVR values of 6.2 (ABM) and 1.1 (CBM) are plotted in Figs. 17 and 18, respectively. These figures show reproducibility of nucleate boiling enhancement at EVR = 6.2 and large scatter in the boiling curves as EVR decreases. In Fig. 18, the large differences in nucleate boiling curves are a result of an increased number of clogged pores due to the increased epoxy content within the coated layer.

CHF enhancement data are plotted in Fig. 19. The microporous surfaces of Table IV show a 1.7- to 2-times increase in CHF over the uncoated reference surface across the range of EVR values tested. Near CHF, the vapor clouds resting over the coated surface inhibit the supply of fresh liquid into the heated area, therefore most of the vacant space within the microstructure is thought to be nearly dried out.

D. Application Method

Chang and You [49] sought a new coating method to make the boiling enhancement paint widely applicable to a variety of geometries. The dripping method could only be used for a flat, horizontal area upon which the paint solution could rest. Therefore, a spray method using compressed air was developed and tested. The paint solution was poured into a sprayer container, which was then filled with compressed air up to about 4 atm.

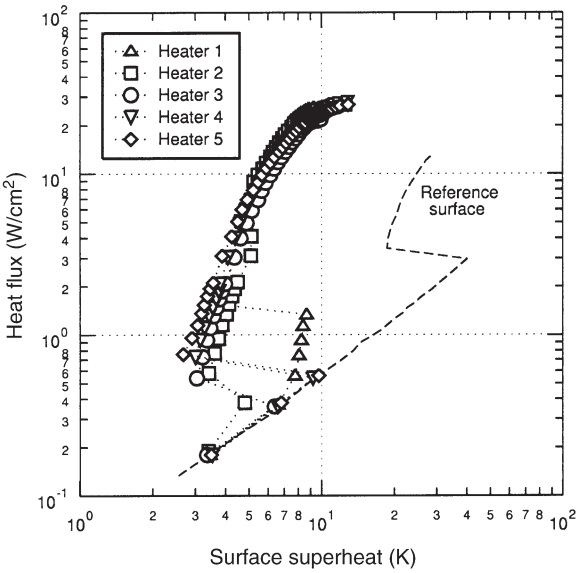


FIG. 17. Boiling from ABM surfaces (EVR = 6.2).

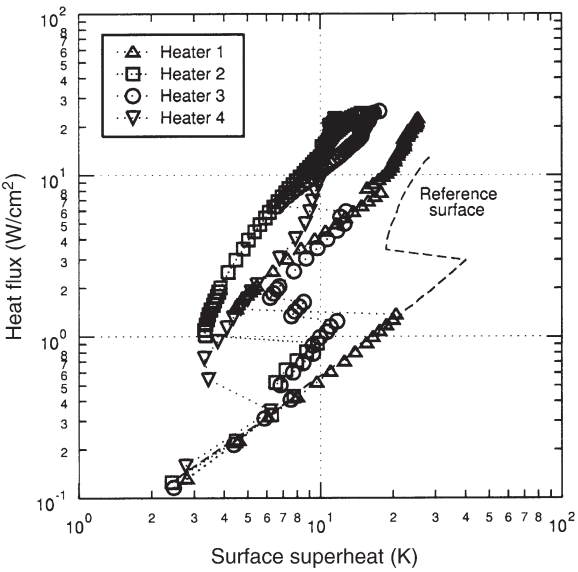


FIG. 18. Boiling from CBM surfaces (EVR = 1.1).

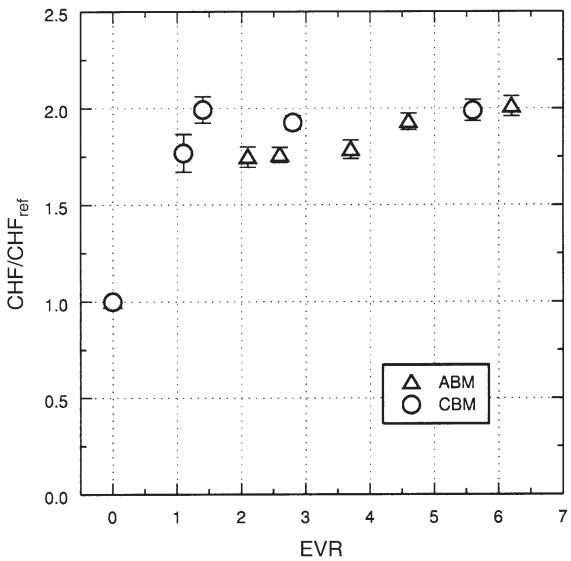


FIG. 19. CHF enhancement vs. EVR.

The spraying jet was controlled by a trigger mechanism on the sprayer. The paint solution, carried by air, was sprayed on the reference surface. The spraying time and the distance between the jet exit point and the target surface determined the attachment area and the thickness of the enhancement layer, however, a specific distance was required to produce a uniform coating layer. The ABM solution was more compatible with this spraying technique than the CBM coating. The 0.1-mm-diameter nozzle exit was easily clogged with the copper particle content due to the higher density and larger size of the copper particles.

Two ABM-coated surfaces with the same EVR value of 4.6 were made using two different coating methods: dripping and spraying. Pool boiling curves for these ABM surfaces are compared in Fig. 20. Both methods show nearly identical incipient superheat, nucleate boiling, and CHF values. These tests prove that the two coating methods can be used interchangeably. In addition, Rainey and You [43] found that a smaller, artist-type airbrush could also be used to apply the coating.

E. Microporous Coating Performance

1. *Pool Boiling Curve*

The performance of the microporous coating was compared with that of the well-known High-Flux commercial surface. Top and side SEM images

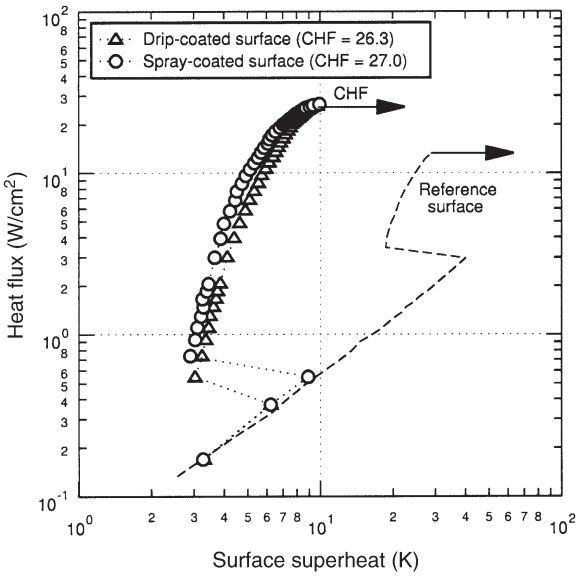
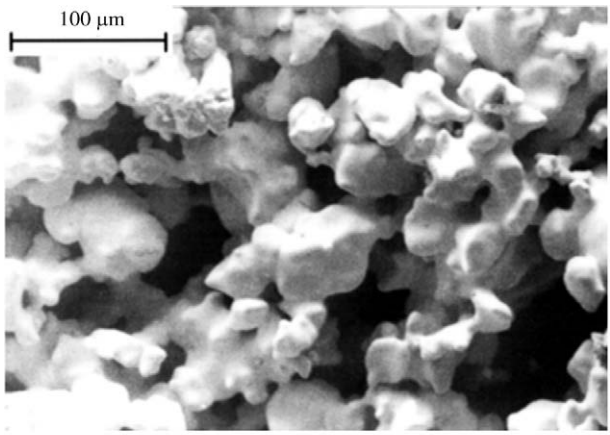


FIG. 20. Effects of coating methods (ABM, EVR = 4.6).

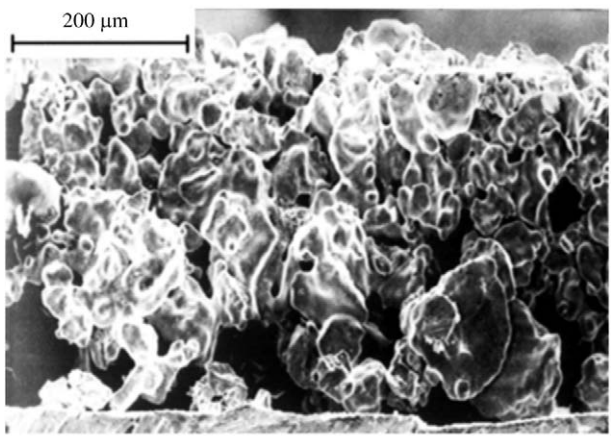
of the High-Flux surface are illustrated in Figs. 21(a) and (b), respectively. The High-Flux surface contains cavities with larger nominal sizes (10–100 μm) than any of the microporous surfaces used by Chang and You. The High-Flux surface, known to have porosity between 40% and 50%, appears to have an extended boiling surface area due to its large pores within the structure. The side view indicates a thickness of ≈400 μm and a structure containing large interconnected internal cavities.

Figure 22 illustrates the test data of the High-Flux surface compared with the spray-coated ABM surface (EVR = 4.6) in saturated FC-72 at 1 atm. Incipient superheat of the High-Flux surface was observed to be ≈5 K, which is comparable with that of the microporous surfaces. In the low heat-flux range (< 10 W/cm^2), the High-Flux surface shows higher heat transfer coefficients than the ABM surface by up to 2.6 times. At heat fluxes above 10 W/cm^2, the ABM surface shows comparable performance with the High-Flux surface. The High-Flux surface produced a higher CHF than the ABM surface: 33.5 vs. 27.0 W/cm^2.

Based on the observations in Figs. 21 and 22, the High-Flux surface (porous conducting surface) possesses geometry characteristics that are very beneficial to the enhancement of boiling performance at low heat fluxes and at CHF. In addition to the extended effective boiling surface area



(a) Top view



(b) Side view

FIG. 21. SEM images of High-Flux surface.

within the highly conductive layer, the large pores provide more efficient liquid/vapor exchange channels, which produce increased convection heat transfer. As heat flux was increased, the benefit of the large-size structure diminished due to increased vapor content within the porous structure, which produced comparable nucleate boiling performance with the ABM surface (microporous surface). Although the porous conducting surface showed higher boiling performance, the microporous coating has shown the possibility of broad application based on its simple and benign application process.

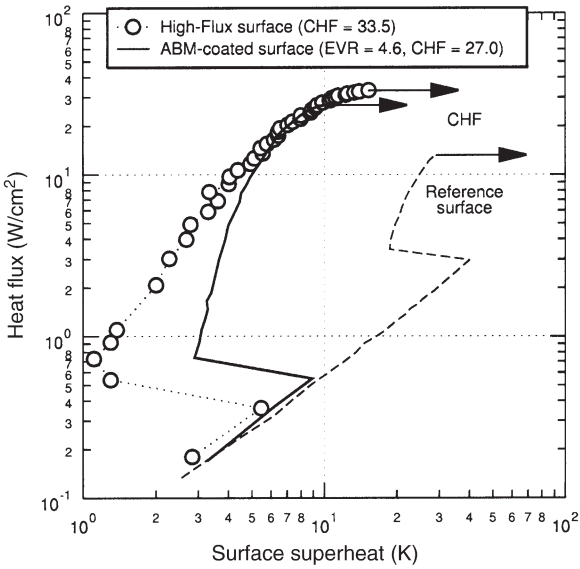


FIG. 22. Boiling from High-Flux surface.

2. *Adhesion, Durability, and Reliability*

Chang and You [49] investigated the long-term performance of the microporous coating. The ABM-coated surface with an EVR value of 4.6 was powered at 15 W/cm^2 and allowed to boil at that heat flux for 110 h. The applied heat flux, higher than the CHF of the reference surface, was chosen to provide a severe operating condition. During this test, slugs of bubbles covered the entire heater surface and rapidly growing and departing bubbles induced strong fluid motion. During the prolonged boiling, the wall temperature slowly increased by about 1°C over the 110-h period. However, when the test was restarted at the same heat flux condition, the initial wall temperature was observed to be identical with that of the previous run. The consistent performance during the durability tests displayed confidence that the attachment of the porous-coated layer was secure.

Initial concerns on the reliability of the coating layer focused on the loss of coating particles due to thermal and mechanical stresses generated during service. Thermal stresses arise mainly due to the difference in the thermal expansion between particle, resin glue, and substrate. The source of mechanical stress is interface tension at the sites where bubbles form. All these concerns were put into consideration and tested by subjecting the coated layer to ultrasonic turbulence. It is known that the site departure

frequency for the coating is on the order of 10 Hz. The ultrasonic bath employed produces mechanical agitation with ~ 4 kHz frequency. It is assumed that the mechanical stress that the coating layer experiences during bubble growth and departure is the same as the stress produced by ultrasonic turbulence. This assumption is believed to be conservative.

In order to count the fraction of particles that are detached in a given testing time, the coating layer is initially coated with conductive carbon, less than 500-Å thick. Carbon coating is used because the area that is not covered with carbon produces electric charge when placed in a scanning electron microscope. By comparing the coated layer before and after testing under the SEM, any failure in the coating layer can be easily detected because it produces unusually bright contrast.

Coated layers with two different substrate conditions were tested. The first substrate was prepared using 600-grit sandpaper, while the second was prepared using 100-grit sandpaper. Different substrate surface conditions change the locking surface area of the glue and are therefore expected to produce different reliability behavior. The ultrasonic testing was done under water for a maximum of 4 days. The surface condition was inspected every 24 h. Due to ultrasonic energy, the bath temperature quickly rose to 60°C during testing.

The test results indicate that the bond strength of the coating is stronger than anticipated. In all cases tested in this experiment, virtually no indication of particle loss was detected. It appears that the substrate prepared with 100-grit paper yields better bonding than the 600-grit surface, however, the difference was minimal. Occasionally, a failed area was detected at the location of pre-existing damage (a scratch) that was intentionally introduced for comparison. Even in this case, the extent of the coating detachment is far less than 1% of the total surface area. Under the assumed conditions, 1 day in the ultrasonic bath is equivalent to about 1 year of service because the agitation produced by the bath is about 400 times more than the actual service condition. Testing for 4 days should simulate approximately 4 years of continuous boiling. Since no damage was detected even for 4 days of vigorous testing, it was not possible to produce a correlation that predicts the kinetics of the damage process. It is safe to conclude, however, that the coating is fairly reliable–enough to survive up to 10 years of service without any noticeable failure.

V. Pool Boiling Parametric Effects with the Microporous Coating

Since the microporous coating was originally developed for electronics cooling applications, it is conceivable that the coating will need to perform

under varying geometrical and fluid condition constraints. Therefore, the effects of heater geometry and fluid conditions on the microporous coating's enhancement performance were determined.

A. Heater Size and Orientation Effects

Following Chang and You's [58] initial study of orientation effects using 1-cm^2 heaters with the microporous coating, Rainey and You [59] performed a comprehensive study of the effects of heater size and orientation on the performance of the microporous coating. More specifically, Rainey and You studied 1-, 4-, and 25-cm^2 flat, microporous-coated and plain copper heaters at orientations ranging from $\theta = 0^\circ$ (upward facing) to 180° (downward facing) under saturated pool boiling conditions in FC-72. Figure 23 shows the effects of heater size on the $\theta = 0^\circ$ nucleate boiling curves of the microporous-coated surfaces along with those of the plain reference surfaces. From Fig. 23, it appears that the nucleate boiling performance of the microporous coating is unaffected by heater size in this size range. Park and Bergles [60] observed similar behavior for their vertically oriented heaters.

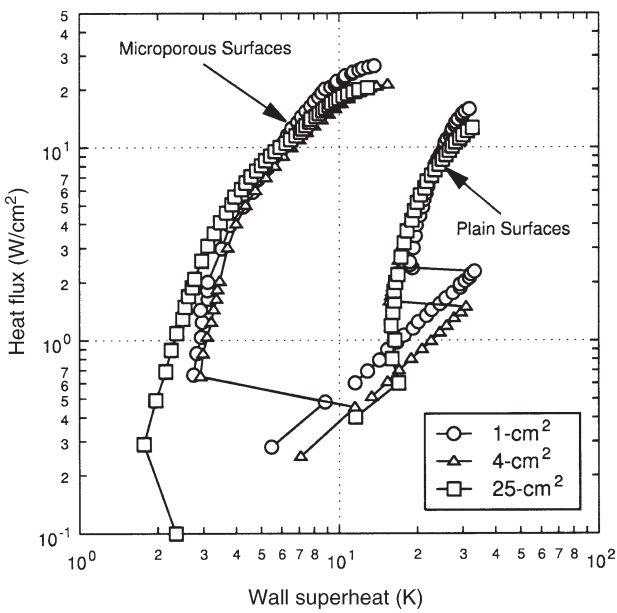


Fig. 23. Reference surfaces' pool boiling data ($\theta = 0^\circ$, increasing heat flux).

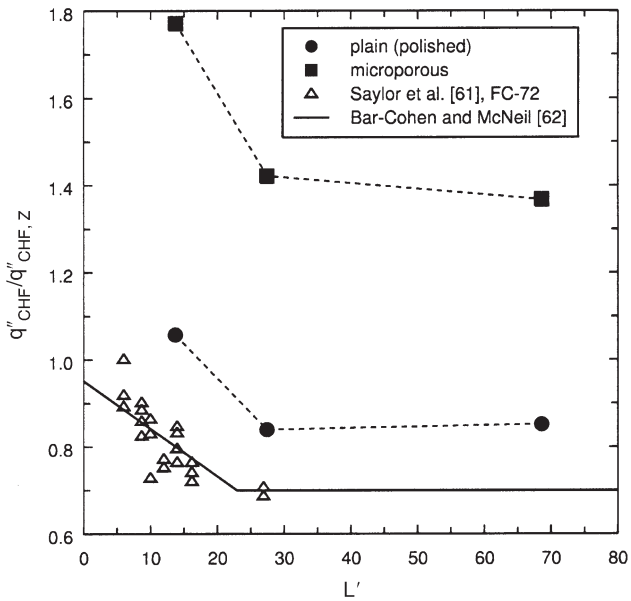


FIG. 24. Effect of heater size on CHF ($\theta = 0°$).

Figure 24 shows the normalized $\theta = 0°$ CHF values versus a dimensionless heater length scale, L', for these surfaces along with the data of Saylor *et al.* [61] (Saylor *et al.* tested flat, smooth surfaces in saturated FC-72 at 1 atm). The data are normalized with respect to Zuber's [42] CHF correlation [Eq. (2)]. The 1-cm^2 heater (equivalent to L' of ≈ 14) exhibited a significantly higher CHF value than the 4- and 25-cm^2 heaters. This trend was found to be similar to Saylor *et al.*'s data although shifted upward due to the microporous coating's enhancement of CHF.

The CHF values appear to be asymptotic for large surfaces but increase with decreasing heater size past a "transition point". Using Saylor *et al.*'s data, Bar-Cohen and McNeil [62] estimated this transition point as $L'_{\text{trans}} \approx 20$ (which corresponds to a 2.25-cm^2 heater). This behavior is explained in terms of fluid re-wetting. As the heat flux increases, the amount of vapor covering the surfaces increases, which in turn increases the resistance to fluid re-wetting the heater surface. For small surfaces, a significant portion of the re-wetting fluid is supplied from the sides rather than from above (as would be the case for an infinite flat plate situation). The re-wetting resistance should be a function of flow path length parallel to the heater surface. The larger 4- and 25-cm^2 surfaces have longer fluid paths to reach the center portion of the heater surface

as well as relatively small entrance flow areas (from the sides) compared to their heater surface areas. Because of this increased re-wetting resistance, the larger surfaces experience higher wall superheats near CHF and thus lower CHF values. Park and Bergles' [60] results also showed a trend of increasing CHF with decreasing heater size.

To investigate the effects of surface orientation on the microporous coating's performance, the heater surfaces were rotated from $\theta = 0^\circ$ to 180°. Figure 25 depicts the pool boiling curves at various angles for the 25-cm^2 microporous-coated and plain surfaces. Unlike the plain reference surface, the heat transfer coefficient in the nucleate boiling regime for the microporous surface is unaffected by surface inclination. This behavior is attributed to the higher number of active nucleation sites provided by the microporous structure, which significantly enhances heat removal mechanisms. This nucleation enhancement dominates any effect of changing the inclination angle.

Although the inclination angle had no noticeable effect on the nucleate boiling heat transfer of the microporous coating, the CHF values shown in Fig. 25 are significantly reduced with increased inclination angle. Figure 26 better illustrates this effect. The data in Fig. 26 were normalized with respect to the CHF value at $\theta = 0^\circ$ and plotted versus inclination angle.

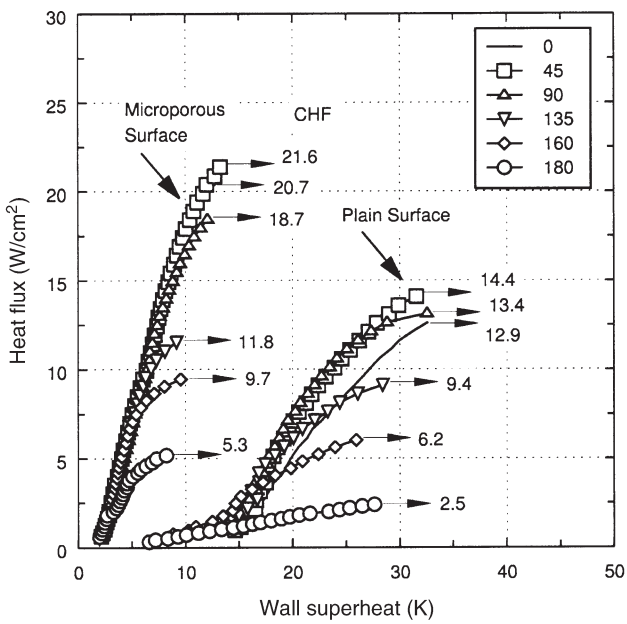


FIG. 25. Effects of inclination angle for 25-cm^2 surface (increasing heat flux).

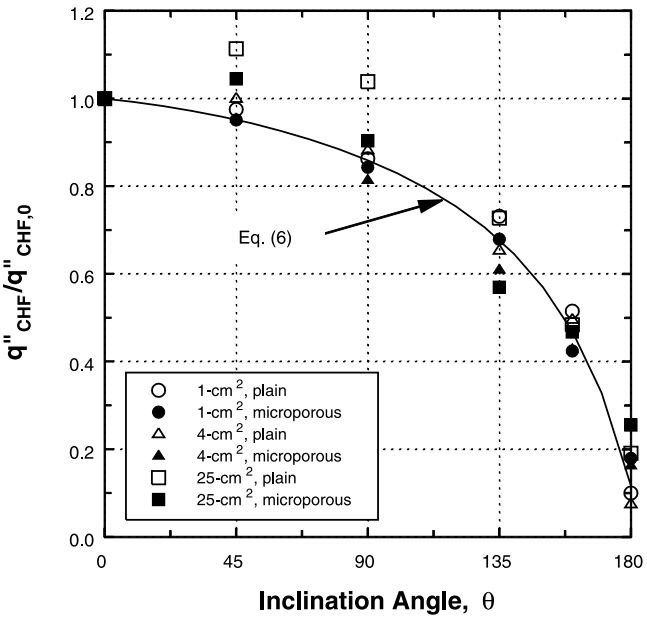


FIG. 26. Normalized CHF data versus inclination angle.

The CHF values decrease as inclination angle increases. In Fig. 26, the 25-cm^2 surfaces appear to show slight increases in CHF at 45° and 90°. This may be due to increased convection heat transfer from bubbles sweeping along the longer-length, 25-cm^2 surfaces. However, even though this behavior was repeatable, the increase in CHF values are nearly within experimental uncertainty, therefore, a larger test surface would be required to confirm this observation. Also plotted in Fig. 26 is the empirical correlation developed by Chang and You [58] given by:

$$\frac{q''_{\mathrm{CHF}}}{q''_{\mathrm{CHF},0^\circ}} = 1.0 - 0.00120 \cdot \theta \cdot \tan(0.414 \cdot \theta) - 0.122 \cdot \sin(0.318 \cdot \theta) \qquad (6)$$

The CHF data for all of the surfaces, including the plain surfaces, follow a similar trend with Eq. (6). The similarity in the normalized CHF data for each of the different surface conditions shows that the CHF mechanism is a strong function of surface orientation, even for enhanced surfaces.

Combining the nucleate boiling curves for all heater sizes and orientations dramatically illustrates the consistency of the nucleate boiling performance for the microporous coating. As can be seen in Fig. 27, the microporous surface boiling curves for all three heater sizes collapse to

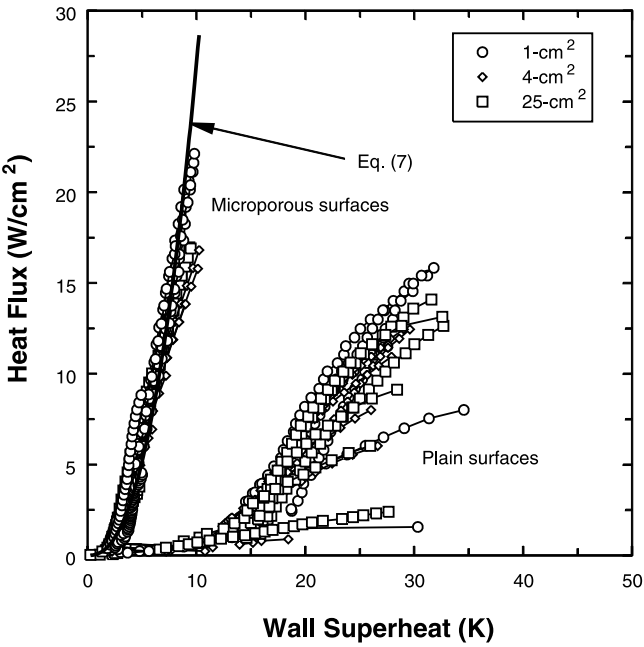


FIG. 27. Effects of size and inclination angle for all surfaces (increasing heat flux).

one line, which can be correlated with the following simple power-law fit:

$$q'' = 1100 \cdot \Delta T_{\text{sat}}^{2.4} \tag{7}$$

In comparison, the plain reference surfaces showed significant scatter along with poorer overall boiling performance.

B. FLUID CONDITIONS

Rainey *et al.* [63] studied the effects of pressure, subcooling, and dissolved air on the boiling performance of the microporous coating. Using 1-cm^2 flat microporous-coated and plain (machine-roughened using a slitting saw blade; similar in roughness to a 600-grit sanded surface) copper heaters under pool boiling conditions in FC-72, the pressure was varied from 30 to 150 kPa and the liquid subcooling from 0 to 50 K. Figure 28 shows the effects of pressure at saturation conditions for the microporous surface along with that of the plain reference surface. The increased pressure had the effect of reducing the incipient superheat, increasing the nucleate boiling heat transfer coefficient, and increasing CHF. These trends are consistent with other researchers [46,64,65]. Two important observations should be made regarding these effects. First, at low pressures, the incipient superheat

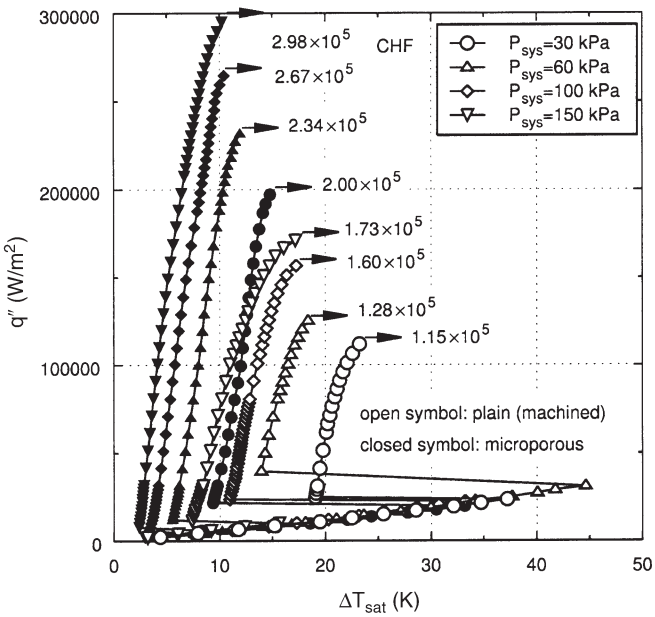


FIG. 28. Pressure effect for plain and microporous saturated boiling curves.

can become quite large, increasing the possibility that the heat source transitions directly to film boiling from natural convection [66]. Second, although the heat transfer coefficient and CHF both increase with increasing pressure, the absolute heater surface temperature also increases. This becomes important when there are maximum temperature constraints on the heat source (e.g., electronic chips).

Figure 29 shows the effect of pressure on CHF for the microporous and plain data, along with roughened (sand-blasted) surface data of Mudawar and Anderson [64]. Zuber's [42] CHF correlation [Eq. (2)] predicts the effects of pressure in the range of test conditions for the plain surface reasonably well. Zuber's correlation, however, does not take into account the effects of surface microstructure and therefore does not predict CHF well for enhanced surfaces. Zuber's correlation was modified by multiplying by an empirically determined constant to account for the effects of surface microstructure. The empirical constant for the microporous-coated surface is 1.78, resulting in a modified correlation given as:

$$\text{Plain:} \qquad q''_{\text{CHF,sat}} \approx q''_{\text{CHF},Z} \tag{8}$$

$$\text{Microporous:} \qquad q''_{\text{CHF,sat}} = 1.78 q''_{\text{CHF},Z} \tag{9}$$

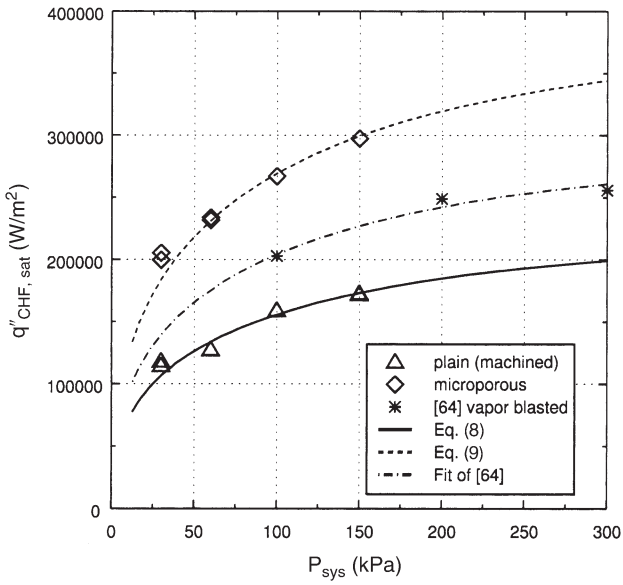


FIG. 29. Effect of pressure on saturated CHF.

For comparison, Zuber's correlation fits the data of Mudawar and Anderson with an empirical constant of 1.35. The empirically derived constant is most likely related to the active nucleation site density just prior to CHF, however, more study should be done in this area to better account for this behavior.

Tests were conducted with the microporous coating in both gas-saturated and pure-subcooled FC-72. As mentioned in Section II.B.1, the term "gas saturated" refers to the condition when the bulk fluid is saturated with dissolved non-condensable gas (in this case, air). The amount of the dissolved air depends on both the temperature and pressure of the fluid. For highly wetting fluids such as FC-72, this can amount to $\approx 40\%$ dissolved air by volume at standard conditions. The term "pure subcooled" refers to the condition when all non-condensables are removed from the fluid (degassed).

Figure 30 shows gas-saturated boiling curves at 100 kPa for subcooling levels of 0, 10, and 30 K. Regardless of the subcooling level, the nucleate boiling curves for the microporous surface collapsed to one line, which is consistent with the prevailing trend on gas-saturated effects in the literature [35,64,67]. Increased subcooling decreases the bubble departure diameters and frequencies [68], which reduces the amount of heat transferred through latent heat and microconvection. However, increased

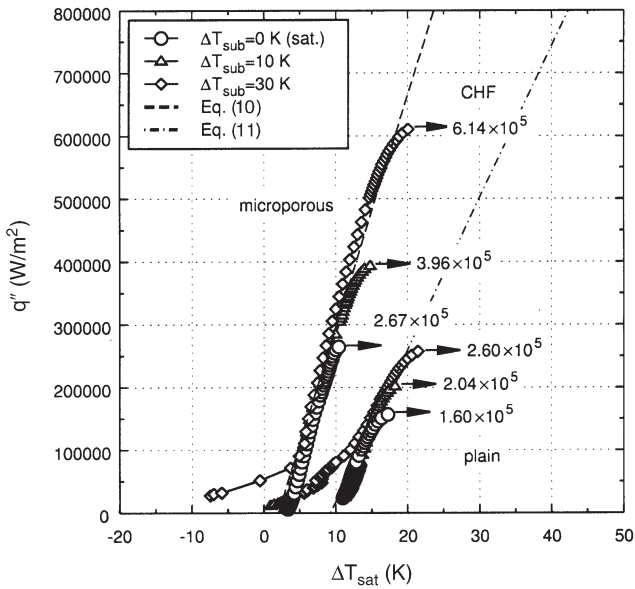


FIG. 30. Gas-saturated boiling curves at $P_{sys} = 100$ kPa.

subcooling also decreases the superheated liquid layer thickness [69], which increases natural and Marongoni convection heat transfer [70]. These combined effects appear to produce a relative insensitivity of the nucleate boiling curve to gas-saturated liquid subcooling.

The gas-saturated data for both microporous and plain surfaces were well correlated for the effects of both pressure and subcooling level with a linear fit given by:

$$\text{Plain:} \qquad q'' = 2.42 \cdot 10^4 \Delta T_{sat} - 3.11 \cdot 10^6 P_{sys} - 0.571 \qquad (10)$$

$$\text{Microporous:} \qquad q'' = 3.71 \cdot 10^4 \Delta T_{sat} - 3.22 \cdot 10^7 P_{sys} - 1.32 \qquad (11)$$

Although it is not clear why, these nucleate boiling data are decidedly linear in nature and do not follow the typical power-law fit of Rohsenow [71].

For many applications, it is not desirable to degas the working fluid of a cooling system prior to use. Therefore, the microporous coating's performance for both gas-saturated and pure-subcooled conditions were compared. Figure 31 shows the gas-saturated boiling curves from Fig. 30 at subcooling levels of 10 and 30 K along with pure-subcooled boiling curves at the same subcooling levels. As can be seen from Fig. 31, there is little difference in the boiling curves. The only significant effects of the

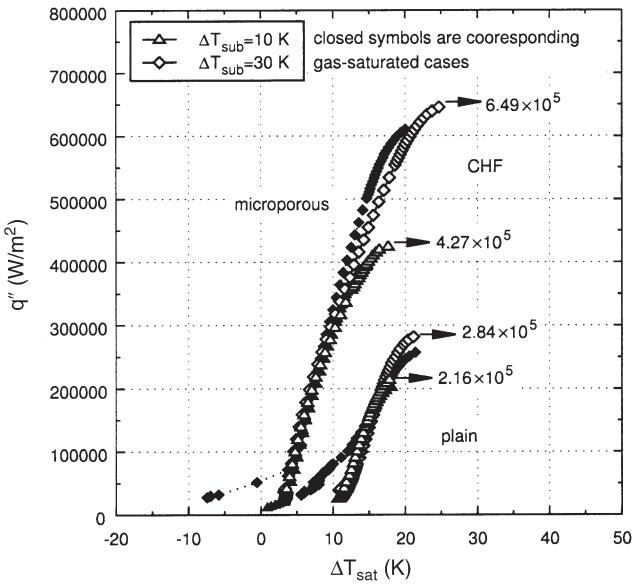


FIG. 31. Comparison of pure-subcooled and gas-saturated boiling curves at $P_{sys} = 100$ kPa.

dissolved air are seen as an enhancement of the heat transfer at lower heat fluxes. You *et al.* [46] attributed this enhancement of the heat transfer for gas-saturated cases to a larger number of active nucleation sites due to the gas partial pressure within the nucleating bubbles. For a given system pressure, an increase in gas partial pressure within the bubbles would reduce the bubble vapor pressure, and hence the wall temperature required to activate the nucleating bubbles (see Section II.B.1).

As heat flux is increased in Fig. 31, the gas-saturated and pure-subcooled boiling curves merged showing little effect of the dissolved air on the nucleate boiling performance. Hong *et al.* [45] explained this behavior as due to depletion of the dissolved air near the heater surface. For a given system pressure, depletion of air within the fluid near the heated surface after boiling long enough provides an identical nucleate boiling situation with the pure-subcooled case. Additionally, You *et al.* [46] observed that the dissolved gas enhancement at low heat fluxes increased with increasing gas content. As dissolved gas concentration increases, the gas-depleted liquid layer becomes gassy more efficiently. The data in Fig. 31 also showed this trend.

The CHF values for the gas-saturated cases are comparable to those of the pure-subcooled values (Fig. 31), also showing little effect of the

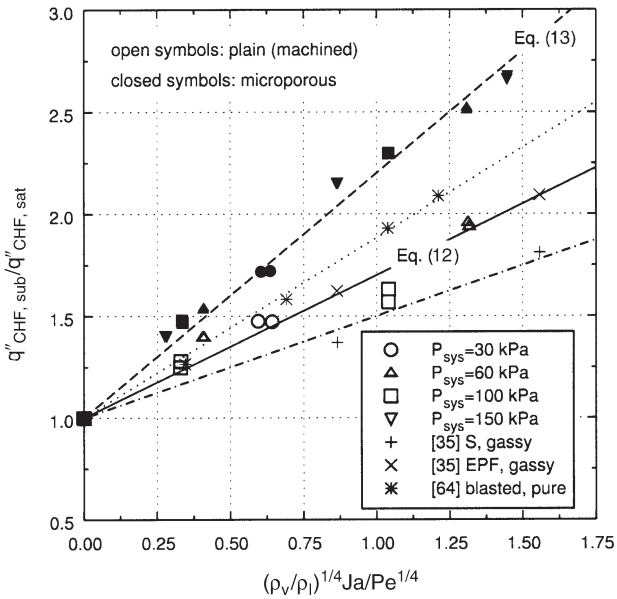


FIG. 32. Effect of subcooling on CHF (gas saturated).

dissolved air. In contrast, using small-diameter wire heaters, Hong *et al.* [72] observed that pure-subcooled CHF values were significantly higher than gas-saturated CHF values. This effect diminished, however, as the wire diameter increased. The depletion of dissolved gas near the heater surface is most likely greater for larger heaters (less effective dissolved gas transport back to the heater surface while boiling), which would produce CHF values closer to the pure-subcooled values as is observed for the boiling curves in Fig. 31.

Figure 32 illustrates the effects of subcooling on CHF for the plain and microporous surfaces at all tested pressures. For comparison, CHF data from Honda *et al.* [35] and Mudawar and Anderson [64] are also shown in Fig. 32. Regarding the effects of pure subcooling on CHF, Inoue *et al.*'s [73] correlation was used to fit the microporous and plain data, given by:

$$\text{Plain:} \qquad q''_{\text{CHF,sub}} = q''_{\text{CHF,sat}}(1 + 0.70(\rho_v/\rho_l)^{1/4}\text{Ja}/\text{Pe}^{1/4}) \qquad (12)$$

$$\text{Microporous:} \qquad q''_{\text{CHF,sub}} = q''_{\text{CHF,sat}}(1 + 1.20(\rho_v/\rho_l)^{1/4}\text{Ja}/\text{Pe}^{1/4}) \qquad (13)$$

Equations 12 and 13 correlate the data well. Interestingly, the final form of the equation is very similar to Elkassabgi and Lienhard's [74] low-subcooling-region correlation. Another interesting observation from Fig. 32

is that the microporous surface exhibits a greater enhancement from increased subcooling than the plain (machined) surface. This is believed to be due to the much higher active nucleation site density of the microporous surface compared to the plain surface, creating smaller and more numerous departing bubble diameters [48,49].

C. Double Enhancement

In an attempt to further increase boiling enhancement performance, Rainey and You [43] combined the microporous coating with pin fins. More specifically, they spray-coated copper heaters that had 1-mm, square-cross-section pin fins, a base dimension of 1 cm^2, and fin lengths of between $L = 0$ and 8 mm. This combination is called "double enhancement" because of the enhancement of area (pin fins) along with the enhancement due to the microporous coating.

The saturated pool boiling results for these double-enhanced surfaces are shown in Fig. 33 for FC-72 at 1 atm. The reference surface area used for heat flux calculations in Fig. 33 is the base area, excluding the area of the fins. The double-enhanced data collapse to one line regardless of fin

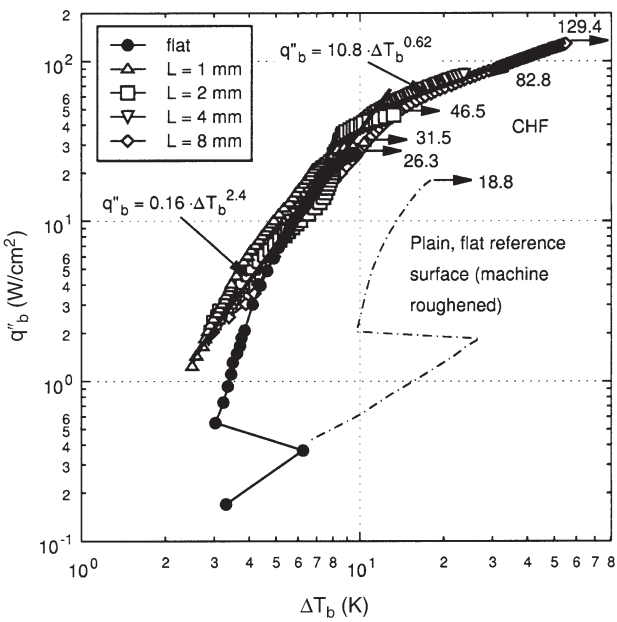


FIG. 33. Microporous-coated, finned-surface boiling curves based on base surface area.

length. This is reasonable because the heat transfer surface area does not affect nucleate boiling performance unless the number of active nucleation sites becomes limited and does not increase with increasing heat flux. The microporous coating removes this form of surface area dependency by providing a much higher number of active nucleation sites. The effect of changes in surface area only becomes evident at CHF.

The slope of the double-enhanced nucleate boiling curves in Fig. 33 changed significantly at about 50 W/cm^2. Below 50 W/cm^2, the boiling curves can be approximated by the power law fit of $q'' = 0.16\Delta T_b^{2.4}$. Above 50 W/cm^2, the boiling curves can be approximated by $q'' = 10.8\Delta T_b^{0.62}$. It is believed that this change in slope is caused by a localized dryout situation at the base of the fins.

Figure 34 shows the effect of fin length on CHF (again, the reference area is the base area). Zuber's [42] CHF correlation [Eq. (2)] is shown for comparison. This curve was generated by multiplying Zuber's CHF value by the total heated area (including the base and fins) divided by the base area. From Fig. 34 it can be seen that CHF for the microporous finned surfaces appears to increase linearly with fin length, following Zuber's correlation. The plain, finned surfaces follow Zuber's correlation up to a fin length of 4 mm and then degrade significantly at 8 mm. This degradation

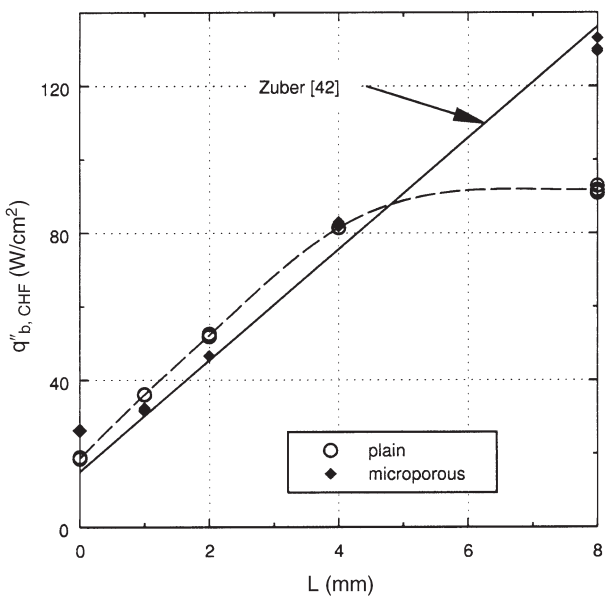


FIG. 34. CHF versus fin length.

is attributed to a non-boiling tip condition that the microporous surface did not have. Due to the microporous coating's nucleation enhancement characteristics, it is able to sustain nucleate boiling at lower wall superheats than a plain surface, however, it is speculated that the microporous surfaces would also experience this same degradation at longer fin lengths. The linear relationship with fin length indicates that the microporous finned surfaces are able to effectively utilize the entire fin surface in this range of fin lengths.

For fin lengths between 0 and 4 mm, the microporous-coated finned surfaces did not enhance CHF compared to the plain finned surfaces in Fig. 34, in contrast to the flat surface behavior. Figure 35 focuses on fin lengths between 0 and 2 mm to illustrate how the addition of fins affects CHF. The CHF data in Fig. 35 are based upon the total heated area and they are normalized with respect to Zuber's CHF value. CHF enhancement is observed at the 0 mm case (flat surface), however, the enhancement becomes negligible as the fin length increases to 2 mm. The finned-surface CHF is limited by hydrodynamic mechanisms rather than by the surface microstructure (number of active nucleation sites). Unlike flat surfaces, finned surfaces have adjacent and opposing walls, which increases the level of bubble crowding at a given heat flux. As the heat flux approaches CHF, the increased bubble crowding more effectively blocks

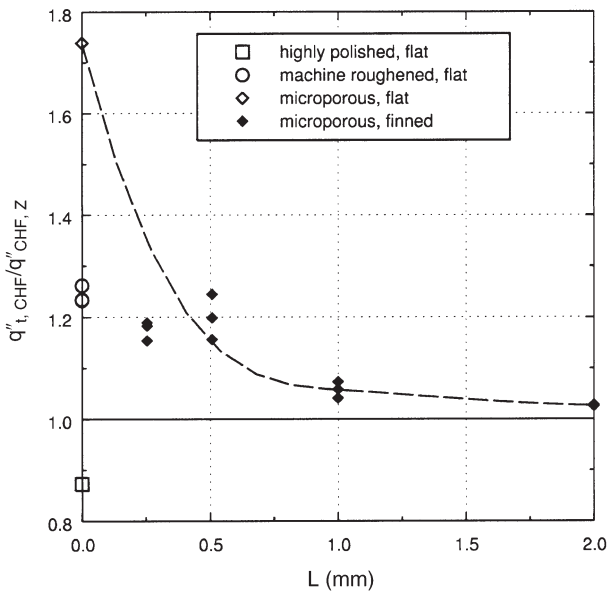


FIG. 35. Normalized CHF versus fin length in the short-fin region.

the re-wetting liquid flow and causes CHF to occur at lower heat fluxes than an equivalent flat surface.

D. Cylindrical Heaters

In an effort to study the use of the microporous coating in heat exchanger applications, Chang and You [75] tested horizontal microporous-coated cylindrical heaters and compared them to plain and commercially enhanced surfaces. Figure 36 compares plain and microporous-coated cylindrical heater pool boiling data in both saturated FC-87 and saturated R-123 at 1 atm. Although nucleate boiling heat transfer was enhanced with the microporous coating, the CHF values were not. This is different from the flat surface behavior where the microporous coating significantly increases CHF. This CHF behavior for the cylindrical surfaces was attributed to a "bubble blanketing" phenomenon in which the bubbles from the lower portion of the surface obscure the upper portion of the surface. The increased void fraction around the upper portion inhibits the liquid from re-wetting the surface. This bubble crowding causes the CHF to be limited by hydrodynamic mechanisms rather than by the surface microstructure, similar to the microporous finned surface situation described in the previous section.

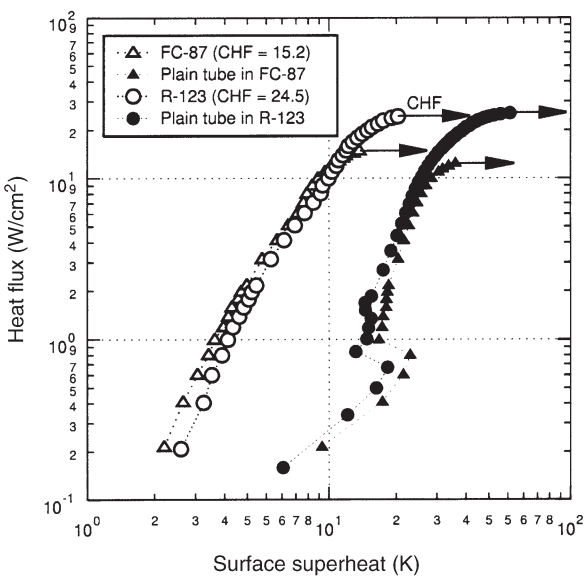


Fig. 36. Pool boiling data of plain and microporous-coated cylindrical surfaces in saturated FC-87 and R-123.

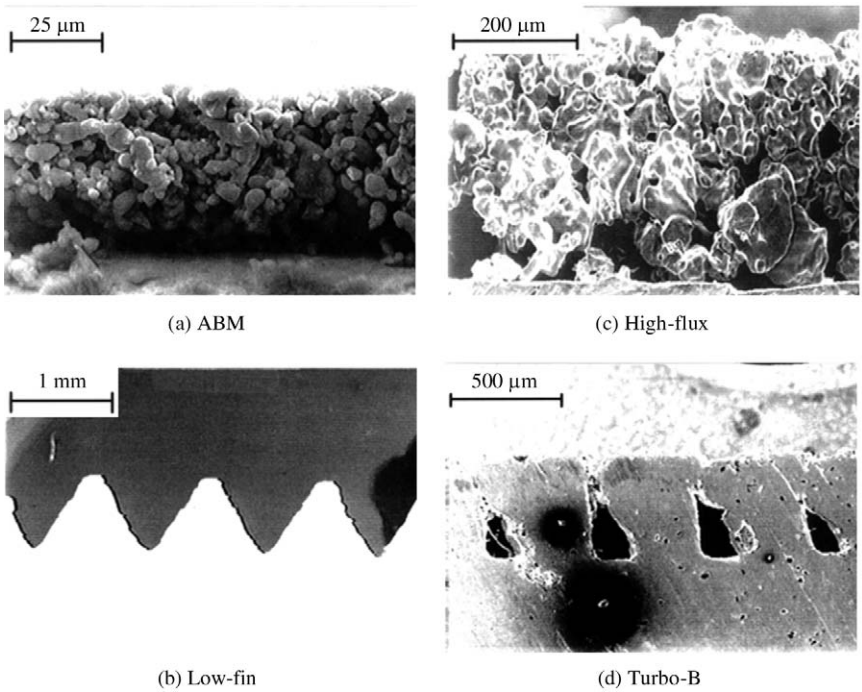


FIG. 37. SEM images of enhanced surfaces.

Next, plain/smooth, microporous/smooth, microporous/low-fin, High-Flux-coated, and Turbo-B cylindrical surfaces were compared. For reference, Fig. 37 shows SEM images of these different surfaces. The High-Flux and Turbo-B surfaces are commercially available boiling enhancement surfaces. Figure 38 shows the results of the pool boiling comparison in saturated R-123 at 1 atm. Although the High-Flux surface had better nucleate boiling performance, CHF was slightly higher for the Turbo-B and microporous/low-fin double enhancement surfaces. From Fig. 38, it can be seen that the microporous coating is close in performance to the two commercially available enhancement surfaces.

VI. Flow Boiling Heat Transfer with the Microporous Coating

A. Small Channel

Ammerman and You [76] studied the performance of the microporous coating in a small-channel flow boiling configuration. More specifically,

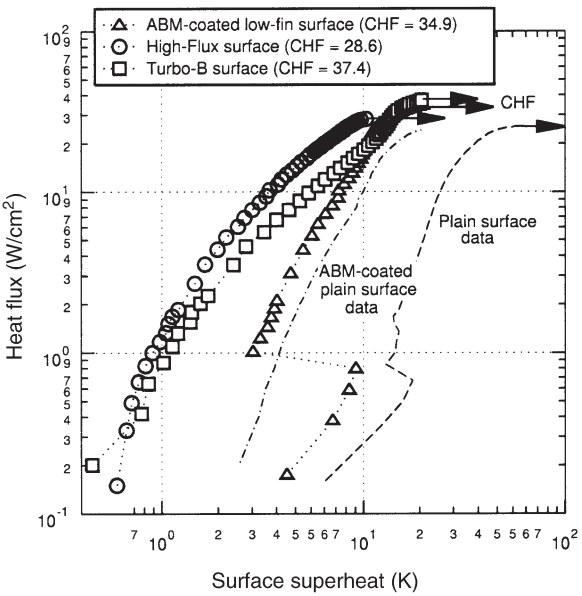


FIG. 38. Pool boiling data of enhanced surfaces in saturated R-123.

they tested a horizontally-oriented, square-cross-section flow channel with a side length of 2 mm and a heated length of 8 cm (only one side of the channel was heated). The heated surface of the channel was tested both with, and without, the microporous surface coating. Convective boiling curves were generated using FC-87 at single-phase inlet velocities ranging from 0.31 to 3.15 m/s and pure inlet subcooling levels from 2 to 31°C.

A comparison of boiling curves for the plain and microporous-coated channels are shown in Fig. 39 at a velocity of 0.31 m/s. CHF values in units of W/cm^2 are supplied in the legend in parentheses for each case on this and subsequent plots of boiling curves. Considerable enhancement can be seen in both heat transfer coefficient and CHF due to the application of the coating. At a heat flux of 20 W/cm^2, for example, the heat transfer coefficient is doubled. Widespread boiling for the coated channels occurs at much lower wall superheats (3 to 5°C) as compared with those of the plain channel (15 to 20°C). Pressure drop values for the 0.31-m/s data are not available since they were less than 5 kPa, which was below the measurement resolution of the test apparatus.

Boiling curves for the plain and coated channels are shown in Fig. 40 for a velocity of 1.25 m/s. The coated channel shows substantial heat transfer and CHF enhancement over the uncoated channel. At a heat flux

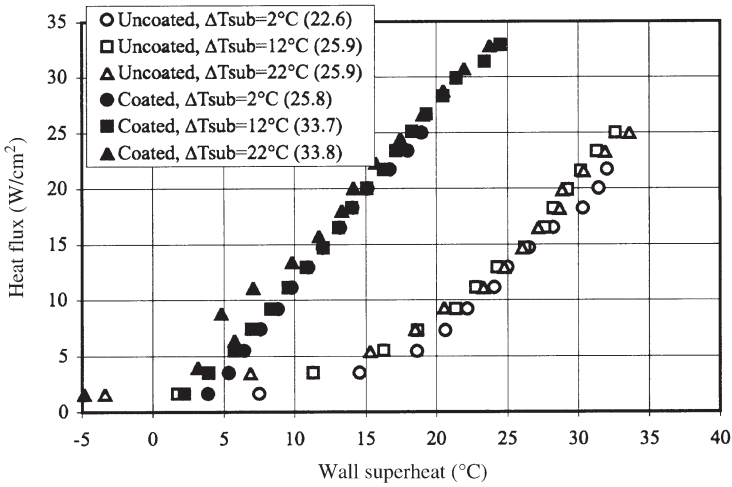


FIG. 39. Boiling curves showing coating effects at 0.31 m/s.

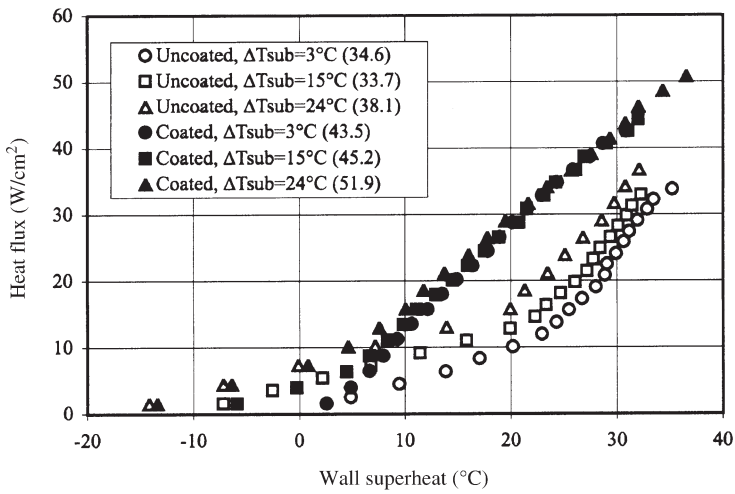


FIG. 40. Boiling curves showing coating effects at 1.25 m/s.

of 30 W/cm^2, for example, the coated-channel heat transfer coefficient is approximately 1.5 times that of the plain. As in the case of the 0.31-m/s data, widespread boiling occurs at much lower wall superheats for the coated channel ($\approx 5°C$ vs. $\approx 20°C$), even though the coated channel is primarily within the subcooled boiling regime throughout its length.

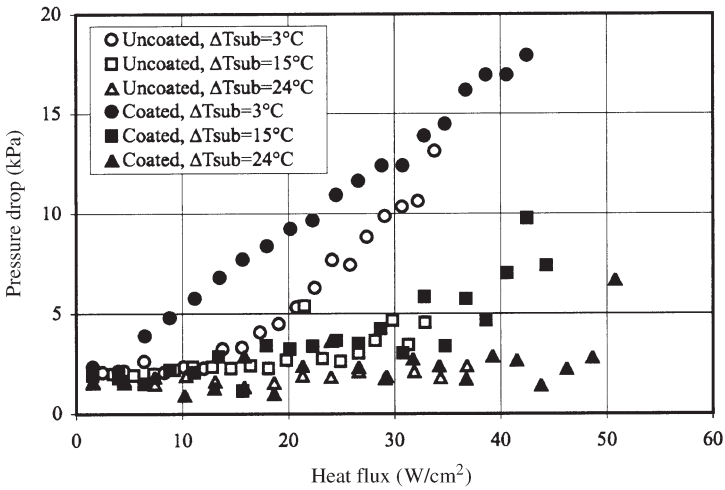


FIG. 41. Pressure drop for the uncoated and coated channels at 1.25 m/s.

Pressure drop data for the plain and coated channels at 1.25 m/s are shown in Fig. 41 versus heat flux. The effects of measurement uncertainty in pressure drop (± 3.5 kPa) are evident in this figure. The presence of the coating does not have much, if any, of an effect on pressure drop for subcooling levels of 15 and 24°C. For the 3°C subcooled case, however, the coated-channel pressure drop exceeds that of the plain channel in the lower heat flux range. Test results showed that the test section exit quality for this subcooling level was saturated. For the coated channel, boiling occurring at lower heat fluxes combined with saturated conditions within the channel results in a larger void fraction, and hence higher pressure drop. As boiling becomes more widespread within the plain channel, its pressure drop begins to approach that of the coated channel.

Based on a qualitative comparison of heat transfer and pressure drop for the 1.25 m/s case, application of the coating clearly provides an advantage. For higher subcooling levels, significant heat transfer enhancement is accompanied by little or no pressure drop penalty. The widespread subcooled boiling generated by the presence of the coating does not increase void fraction because boiling vapor is rapidly condensed. At 3°C subcooling, however, early initiation of boiling results in an immediate increase in pressure drop, tending to offset the benefit of the heat transfer enhancement.

Boiling curves for uncoated and plain channels are shown in Fig. 42 for a velocity of 3.15 m/s. The single-phase region for the coated-channel cases is followed by a gradual transition to subcooled boiling (indicated by

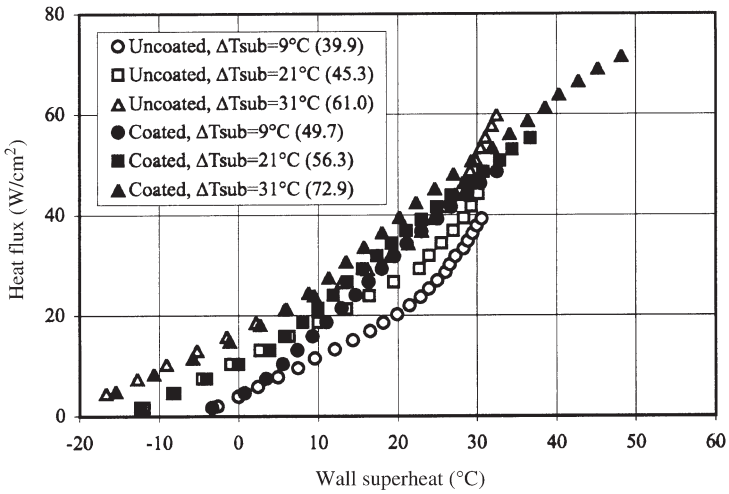


FIG. 42. Boiling curves showing coating effects at 3.15 m/s.

the minor change in slope), which provides modest enhancement in heat transfer over the plain channels. For the 9°C subcooled case at 30 W/cm^2, for example, the coated-channel heat transfer coefficient is approximately 1.4 times that of the plain channel. This enhancement is diminished relative to that seen for lower velocities. This reduction in enhancement is a result of the increase in efficiency of the plain channels with increasing velocity. The coated-channel data are relatively insensitive to velocity effects due to the strong boiling dominance made possible by the coating. The plain-channel data, however, show a steady increase in heat transfer rate with increasing velocity. Once subcooled boiling is established in the plain channels, their boiling-curve slopes exceed those of the coated-channel curves. At a wall superheat of approximately 30°C, this difference in slope causes the uncoated heat transfer rate for the 31°C subcooled case to exceed the coated heat transfer rate.

This transition in heat transfer superiority from the coated to the plain channel may be due to a combination of two effects. First, boiling dominance within the coated channels suggests that, for a given heat flux, wall superheat has a fixed value regardless of velocity and subcooling (within the velocity and subcooling ranges tested). Alternatively, convection dominance within the plain channels results in a heat transfer dependence on velocity and subcooling. For a given heat flux, therefore, wall superheat within the plain channel will continue to decrease as velocity increases, thus shifting the boiling curves to the left. Secondly, the coating layer becomes an added conductive resistance in the heat path (see Section IV.A).

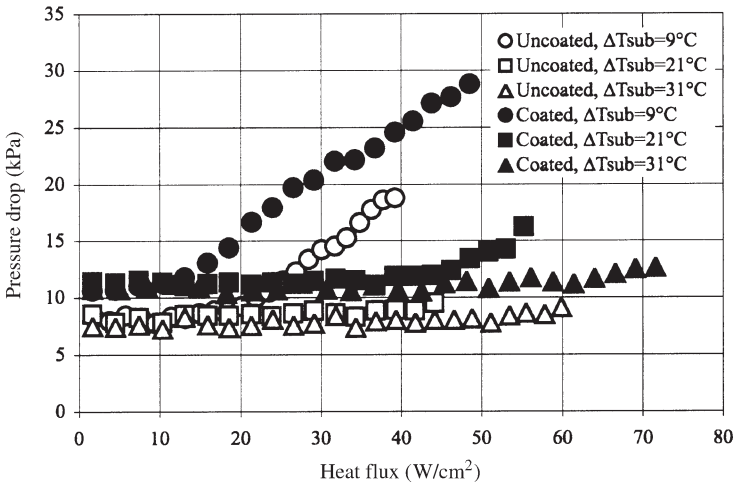


FIG. 43. Pressure drop for the uncoated and coated channels at 3.15 m/s.

This added resistance has the effect of decreasing the boiling curve slope. Minimizing the coating thickness, therefore, is desirable.

A comparison of uncoated- and plain-channel pressure drop for 3.15 m/s is shown in Fig. 43 versus heat flux. A general trend can be seen in which the coated-channel pressure drop departs from its single-phase value at lower heat fluxes than the uncoated channel due to early initiation of boiling. This was also observed in Fig. 42 for the 1.25 m/s case. There is a noticeable single-phase pressure drop difference between the uncoated and coated channels. It is possible that the application of the porous coating roughens the surface causing an increase in the turbulent friction factor.

A qualitative comparison of heat transfer and pressure drop for the 3.15-m/s case indicates that addition of the coating produces only minor benefits at this velocity. At heat fluxes below 20 W/cm^2, boiling is suppressed and no heat transfer benefit is seen for the coated channels. Between approximately 20 and 40 W/cm^2, the coated channels appear to have an advantage as long as increases in pressure drop do not occur. At heat fluxes above 40 W/cm^2, the benefit of the coating appears to come only in the form of increased CHF.

The microporous coating effect on CHF can be seen in the comparison plots shown in Figs. 39, 40, and 42. The increase in CHF due to application of the coating ranged from approximately 14 to 36%. This is somewhat less than the percentages achieved for flat heaters in pool boiling. This is not surprising, however, since a limited amount of liquid exists within the

small-diameter channel. The coated-channel CHF data were correlated with the following equation:

$$\frac{q''_{\mathrm{CHF}}}{Gh_{lv}} = 0.52\mathrm{We}_D^{-0.37}(-x_i)^{0.1}\left(\frac{\rho_v}{\rho l}\right)^{0.6} \tag{14}$$

where G is the mass flux, We_D is the Weber number based on channel hydraulic diameter, and x_i is the equilibrium inlet quality. This correlation performed well in illustrating the trend of the coated channel data. The coated-channel CHF data are plotted against this prediction in Fig. 44.

For this small-channel application, therefore, the microporous coating produced considerable heat transfer enhancement and showed a remarkable insensitivity to both velocity and subcooling effects (as seen when comparing Figs. 39, 40, and 42). Comparison of pressure drop between the plain and coated channels revealed that the benefits of the coating are optimum at higher subcooling levels where subcooled boiling prevails. Even though subcooled, the presence of the coating promotes widespread boiling at lower wall superheats. The vapor generated in this regime is quickly condensed, however, and leads to only minor impacts on pressure drop. For these subcooled boiling cases, the coating provides a significant, penalty-free heat transfer enhancement. This attribute of the coating makes it superior to current compact heat exchanger enhancement methods such as the addition of fins or ribs inside channels.

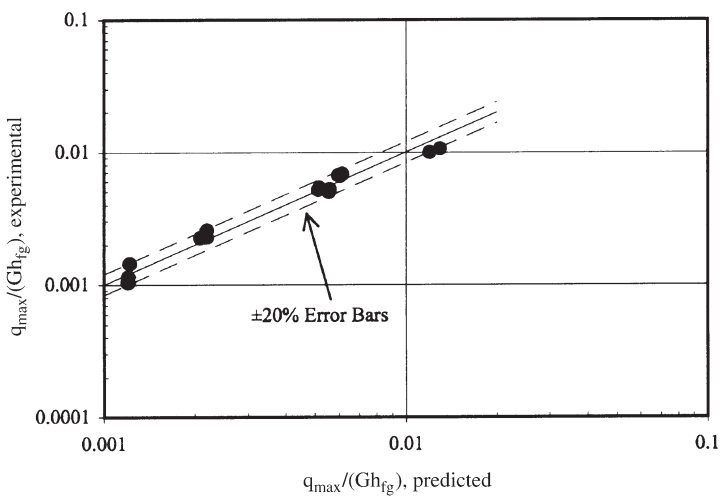


FIG. 44. Coated-channel CHF correlation (Eq. 14).

B. High Flow Rate, Discrete Surface

Rainey *et al*. [77] studied the effects of applying the microporous coating to a small flat surface in flow boiling conditions. Specifically, they placed a 1-cm^2 surface in the side of a 1.27-cm square-cross-section flow channel and tested it in pure-subcooled flow boiling conditions in FC-72. Subcooling levels were varied from 4 to 20°C and velocities ranged from 0.5 to 4 m/s. Figures 45 and 46 show the boiling curves (heat flux and heat transfer coefficient, respectively) of the microporous-coated surface along with the plain reference surface (the single-phase convection portions have been omitted). The saturated pool boiling curves for each surface are also shown.

The results in Fig. 45 illustrate that the fully-developed nucleate boiling heat transfer of the microporous surface was insensitive to both velocity and subcooling level, similar to Ammerman and You's [76] observations. The primary slope of each of the plain-surface boiling curves also showed insensitivity to velocity and subcooling levels. The boiling curves in Fig. 45

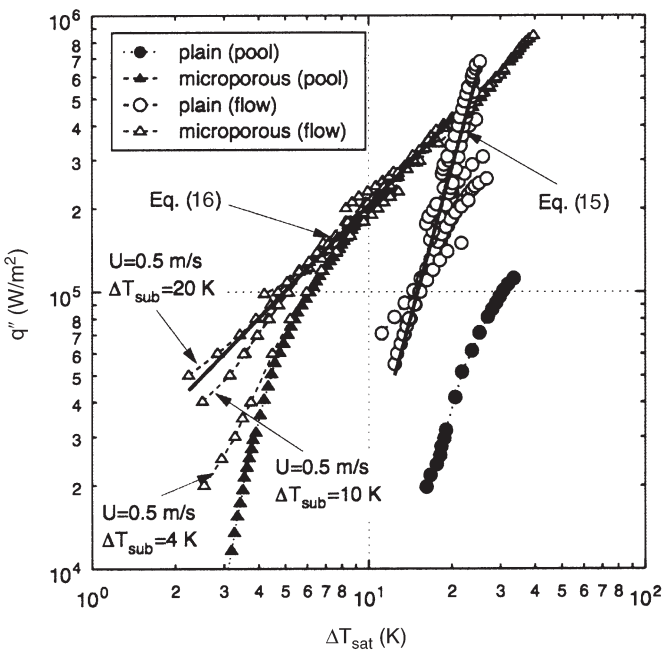


Fig. 45. Correlation of nucleate boiling curves.

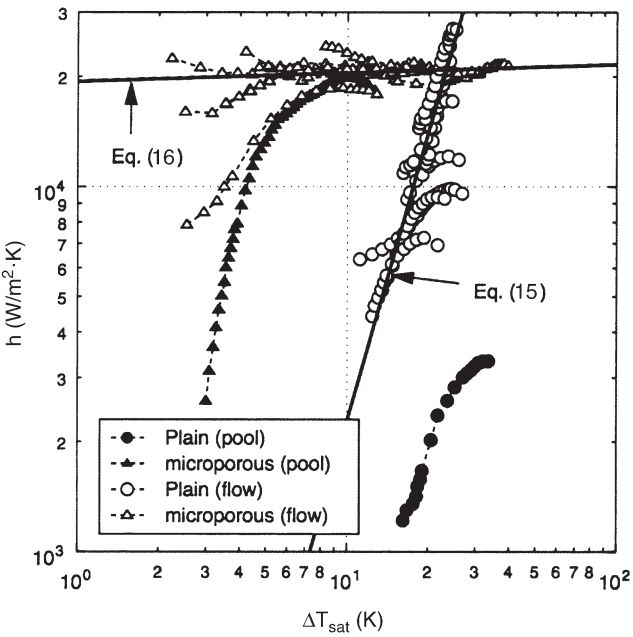


FIG. 46. Average nucleate boiling heat transfer coefficients.

were correlated with the following equations:

$$\text{Plain:} \qquad q'' = 5.39\Delta T_{\text{sat}}^{3.63}. \tag{15}$$

$$\text{Microporous:} \qquad q'' = 1.94 \times 10^4 \Delta T_{sat}^{1.02}. \tag{16}$$

From Fig. 45 it can be seen that the nucleate boiling performance of the microporous coating is much better than the plain surface at low heat fluxes, however, at higher heat fluxes the plain surface actually provides better boiling performance (higher slope than the microporous surface). This behavior is similar to that observed for the small channel at higher velocities. The coating generates widespread subcooled nucleate boiling (boiling dominance) at lower wall superheats than the plain surface, thus providing significant enhancement. At higher wall superheats, however, the added thermal conduction resistance of the coating tends to diminish the enhancement. The total thermal resistance for the plain heater is dependent only on the thermal resistance of flow boiling, while the total thermal resistance for the microporous heater is dependent on the thermal resistance of flow boiling plus the thermal resistance of conduction for the thin

microporous coating layer. Referring to Fig. 46, the nearly constant heat transfer coefficient ($\approx 2 \times 10^4$ W/m$^2 \cdot$K) for the microporous surface means that the thermal resistance of flow boiling is much less than the thermal resistance of conduction for the microporous coating layer, therefore, the total thermal resistance is approximately equal to the thermal resistance of conduction for the microporous coating layer only. From this realization, the effective thermal conductivity of the microporous coating layer was estimated as 1 W/m$\cdot$K (coating thickness ≈ 50 μm), which compares well with the previously estimated values of 1.08 W/m$\cdot$K by O'Connor *et al.* [50] for the DOA microporous coating and 0.95 W/m$\cdot$K by O'Connor and You [48] for the silver flake microporous coating.

The effects of fluid velocity and subcooling level on microporous surface CHF are shown in Fig. 47. As is shown, CHF increases with increasing velocity and subcooling. The subcooling effect is due to a reduction, via condensation, of vapor near the surface, which decreases the resistance to liquid re-wetting the surface, thus delaying CHF. With respect to fluid velocity, the CHF data show a change in slope near 2 m/s similar to that observed by Mudawar and Maddox [78]. Using Haramura and Katto's [79] macrolayer dryout model, Mudawar and Maddox explained that the low-velocity CHF was caused by dryout of the liquid sublayer beneath a large

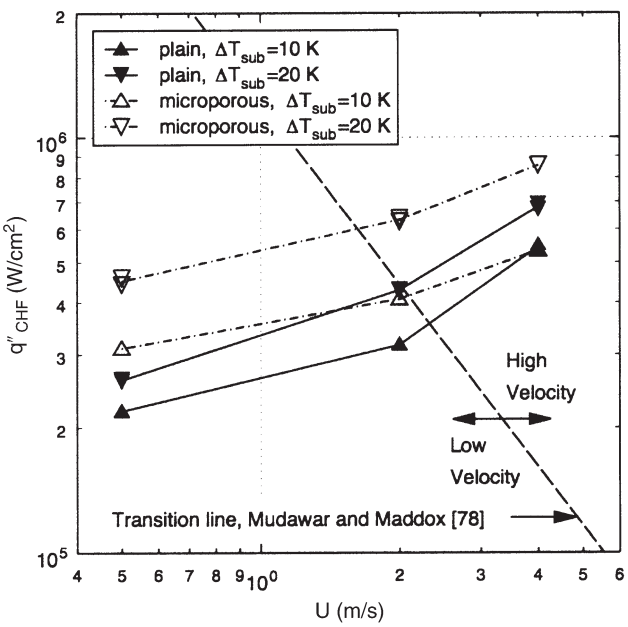


FIG. 47. Effect of velocity and subcooling on CHF.

continuous vapor blanket near the downstream edge of the heater. In the high-velocity CHF regime, they observed that the thin vapor layer covering the surface was broken into continuous vapor blankets much smaller than the heater surface. This most likely decreases the resistance to the fluid re-wetting the liquid sublayer, providing an additional enhancement to CHF and subsequent increase in slope as seen in Fig. 47. The dashed line in Fig. 47 represents the transition line proposed by Mudawar and Maddox to separate low- and high-velocity behavior.

VII. Physics of Boiling Heat Transfer Enhancement

Although it was known that the microporous coating provided significant enhancement to nucleate boiling heat transfer, it was not known by what mechanism the heat transfer was enhanced. To help answer this question, Ammerman and You [80] developed a method to measure the vapor volume flow rate departing from a boiling wire called the "Consecutive Photo Method". The method involves photographing departing bubbles from a wire heater with a high-speed video camera. Consecutive video images are then analyzed with image processing software to measure bubble size and frequency, which can be used to calculate the vapor volume flow rate leaving the wire. The latent heat transfer contribution to the total heat flux is then calculated from:

$$q''_{\mathrm{LH}} = \frac{\left(\rho_v \dot{V}_g h_{lv}\right)}{\pi D_w L_w} \tag{17}$$

where $\dot{V}_g$ is the volume flow rate of vapor leaving the wire and D_w and L_w are the diameter and length of the heated wire, respectively.

Using this method, Kim *et al.* [81] studied the nucleate pool boiling and CHF enhancement mechanisms of the microporous coating. As mentioned in Section II.A.2, the heat transfer mechanisms associated with pool boiling are natural convection, Marangoni convection, microconvection, and latent heat transfer. To simplify the analysis, only fully developed nucleate boiling in saturated liquid conditions (FC-72) was investigated. These test conditions allowed both natural and Marangoni convection to be neglected, leaving microconvection and latent heat as the only two significant contributors to heat transfer. It would have been preferable to study the same flat square heater geometry used in the previous microporous studies, however, the consecutive-photo method can only be used in situations where departing bubbles can be clearly seen in the photos, such as with the quasi-2D, 390-μm-diameter wire heaters used in this study.

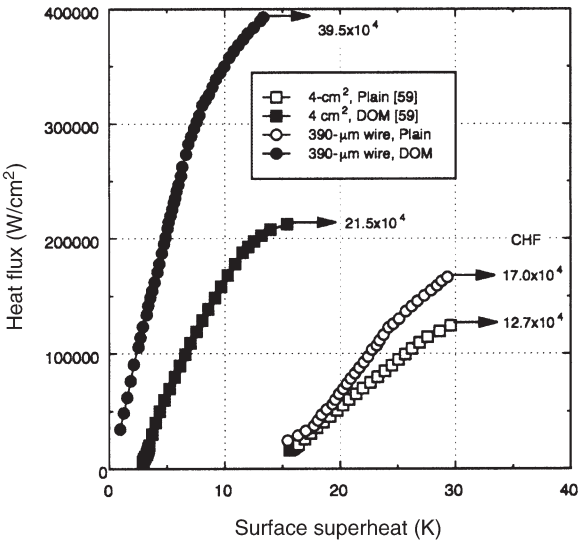


FIG. 48. Reference pool boiling curves.

Figure 48 shows the boiling curves for the microporous-coated and plain, 390-μm-diameter wire heater. Also shown for comparison are microporous-coated and plain boiling curves for a 4-cm^2 flat heater from Rainey and You [59]. The microporous-coated wire showed significant enhancement over the plain wire. At heat fluxes of 6 and 16 W/cm^2, the microporous heat transfer coefficient is approximately 900% and 650% greater, respectively, than that for the plain wire. In addition, CHF for the microporous wire is 130% greater than that for the plain wire, which is similar but higher than the enhancement of about 70% observed for the 4-cm^2 surfaces. It has already been shown that below a certain heater size (length-scale), the nucleate boiling heat transfer coefficient and CHF will increase with decreasing heater size [62,72]. In addition, the wire surfaces show higher levels of enhancement of nucleate boiling and CHF with the addition of the microporous coating. This may be due to increased sensitivity to changes in convection heat transfer from the length-scale effects just mentioned. However, the important observation to make from Fig. 48 is that the 390-μm wire surface behaves in a similar manner to the larger 4-cm^2 flat surface with regard to microporous coating enhancement, therefore, the conclusions drawn from the heat transfer mechanism analysis of 390-μm wire heaters can be applied to relatively larger, flat surfaces.

Figure 49 shows photos of both the plain and microporous-coated wires at 6 and 16 W/cm^2. (The line located 2 mm above the plain wire in

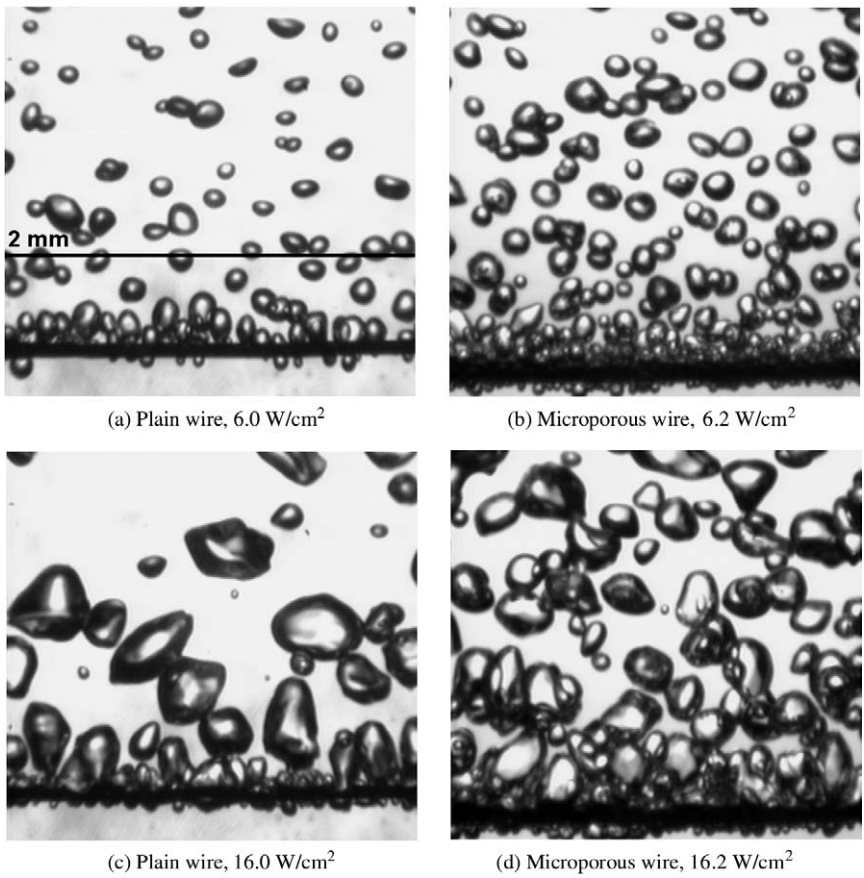


FIG. 49. Photographs of boiling behavior from plain and microporous wires.

Fig. 49(a) indicates where bubble measurements were taken.) At 6 W/cm^2, the photos show a much higher active nucleation site density for the microporous wire compared to the plain wire, along with what appears to be a higher bubble departure frequency. Increasing the heat flux to 16 W/cm^2 produced more active nucleation sites for the plain surface, however, the microporous wire nucleation site density appears unchanged. It is possible that the microporous active nucleation site density is maximized at 6 W/cm^2. Another observation made from Fig. 49 is that although the bubbles for both wires are similarly sized at 6 W/cm^2, the bubble diameters from the plain wire are noticeably larger than for the microporous wire at 16 W/cm^2.

In addition to 6 and 16 W/cm^2, bubble measurements were taken at 11 and 13 W/cm^2. Figure 50 compares the average bubble diameter and

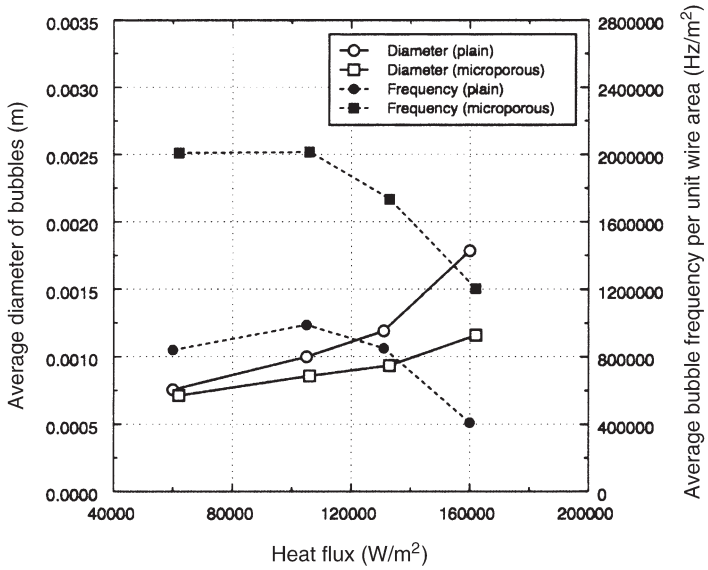


FIG. 50. Average diameter and frequency of bubbles departing from the wire.

frequency for the plain and microporous wires. The microporous wire shows significantly higher bubble frequencies than the plain wire, which is evidence of the higher nucleation site density. As the heat flux is increased, the frequency drops due to increased bubble merging, which was confirmed visually. In addition, the average bubble diameter increases more rapidly with increased heat flux for the plain wire compared to the microporous wire. This behavior seems counter-intuitive because of the much higher bubble frequency (active nucleation site density) of the microporous wire producing more bubble crowding. It is expected that the average bubble diameter would increase with increased heat flux due to a higher vapor volume flow rate, however, this does not explain why the microporous wire exhibited a slower increase in average diameter compared to the plain wire. Three possible mechanisms are proposed to explain this behavior [81]:

(1) First, consider that the bubble will detach once it reaches the "critical" diameter when the buoyancy and/or inertia force are greater than the surface tension force holding the bubble to the surface. When two bubbles of the same size merge, the resulting bubble is twice the volume but the diameter only changes by 26%. Since buoyancy is related to the bubble volume and the surface tension force is related to bubble perimeter in contact with heater

surface (bubble diameter), merging of bubbles will significantly accelerate the departure process. Therefore, when neighboring bubbles merge while still attached to the wire, it is expected that the new, larger bubble would become greater than the critical diameter and immediately depart. The lower active site density of the plain wire would allow an individual bubble to grow to a larger diameter prior to merging with neighboring bubbles and thus have a larger diameter upon departure than the microporous wire.

(2) A second possible explanation would be that increased bubble interaction on the microporous wire from its higher active site density might cause the bubbles to depart earlier. In other words, the effective critical bubble diameter would become smaller due to more turbulence caused by more frequent and crowded bubble departures. A smaller critical diameter due to a higher active site density for the microporous wire may partially offset the effect of increased merging at higher heat fluxes.

(3) A third possible explanation concerns the vapor behavior within the microporous coating. The majority of initial nucleation processes is expected to occur inside the coating near the heating surface. The rest of the coating would act to distribute the vapor to all of the openings on the top surface (cavities within the coating are interconnected). This would provide many more apparent nucleation sites than even a roughened plain surface and would cause a bubble to be formed by a relatively high velocity jet of vapor normal to the surface. The addition of internal vapor inertia on the bubbles could cause premature departure. Evidence of this effect can be seen when comparing the photos of Fig. 49. The photos of the microporous wire show many bubbles protruding from the bottom of the wire while the plain surface bubbles appear to slide upward before becoming large. The addition of vapor inertia would cause many bubbles to grow downward for the microporous wire and also may delay the bubble from sliding upward until the bubble is larger.

Regardless of the actual explanation(s) for the difference in average bubble size behavior between the plain and microporous wire, the effect appears to be a direct result of differences in the active or apparent nucleation site density.

Based on vapor volume flow rate measurements, the relative contributions of convective and latent heat transfer were determined. Figure 51 shows the latent heat flux versus the total heat flux for both the plain and microporous-coated wires. The plain wire showed a steady increase in

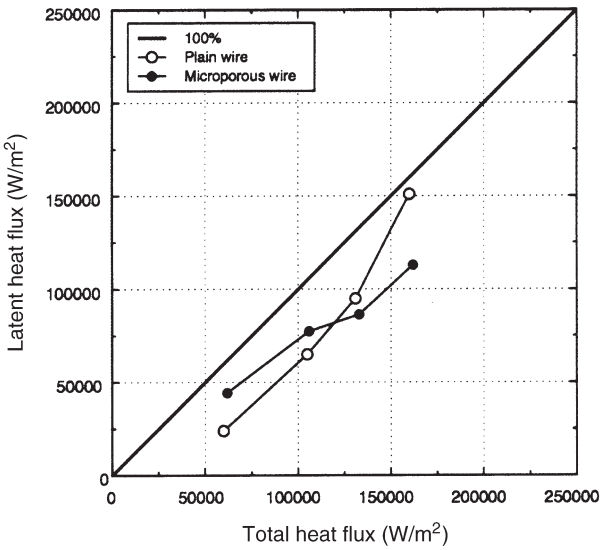


FIG. 51. Latent heat flux contribution for plain and microporous wires.

latent heat flux from about 40% at 6 W/cm^2 to about 94% at 16 W/cm^2. This means that as the heat flux is increased towards CHF (17.0 W/cm^2 for the plain wire), more and more vapor is generated. This is consistent with the observations of Paul and Abdel-Khalik [82]. In contrast, the microporous wire showed a relatively constant latent heat transfer contribution of about 70% between 6 and 16 W/cm^2. (Unfortunately, measurements of vapor volume flow rate any closer to its higher CHF of 39.5 W/cm^2 were not possible due to excessive bubble crowding in the images.)

Figures 50 and 51 reveal that at lower heat fluxes (below about 12 W/cm^2), the microporous coating has a higher latent heat transfer due to its significantly higher bubble departure frequency (active nucleation site density). At higher heat fluxes, the microporous coating has a higher microconvection component, again due to higher departure frequencies (along with smaller departing bubbles).

The enhancement of nucleate boiling performance from the microporous wire over the plain wire is surprisingly similar to the enhancement from the addition of surfactant to water observed by Ammerman and You [83]. They found that by decreasing the surface tension of water, the average bubble diameter was reduced while the average bubble departure frequency increased. From their boiling photos, the active nucleation site density appears to be significantly increased as well. Ammerman

and You further found that these effects on bubble dynamics caused the enhancement of nucleate boiling from increased latent heat at low heat fluxes and from increased convection contribution at high heat fluxes, which is strikingly similar to the results in Fig. 51. They attributed the increased latent heat at low heat fluxes to the activation of normally dormant nucleation sites (increased active nucleation site density). At high heat fluxes, the reduction in surface tension decreased the amount of bubble agglomeration at the surface (decreased diameter/increased frequency). The increased frequency was thought to reduce the superheated liquid layer thickness at the surface. Although the cause for the increased bubble frequencies and reduced bubble diameters is different between the application of the microporous coating and the addition of surfactant to water, the resulting effect on the heat transfer mechanisms appears to be the same.

The consistent contributing factor to nucleate boiling enhancement using the microporous coating is the increase in bubble departure frequency. The rise in departure frequency is probably due to an increase in both single-site frequency and nucleation site density. The occurrence of either of these phenomena would have the effect of lowering the heater surface temperature (increasing the heat transfer coefficient) by reducing the average temperature within the superheated liquid layer that surrounds the heated surface. A notional example of this concept is shown in Fig. 52. If a significant increase in single-site frequency occurred (1 to 2 orders of magnitude), a reduction in waiting time would result during the bubble growth and departure cycle. A reduction in waiting time would lead directly to lower wall superheat. If a rise in the number of active nucleation sites occurred, the combined sum of the individual bubble influence areas (the heater surface areas affected by microconvection due

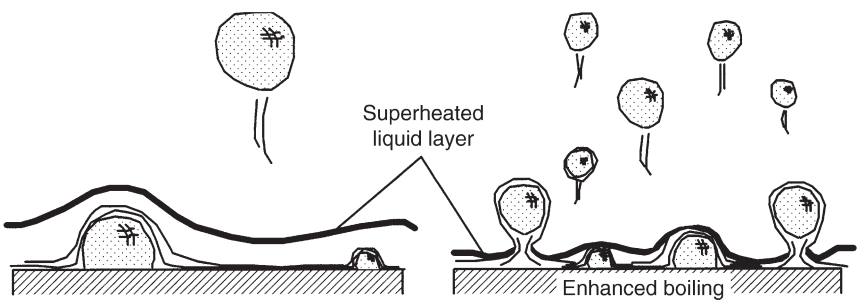


FIG. 52. Example of the effect of increased bubble departure frequency.

to departing bubbles) would increase, even though the resulting bubble diameters are smaller. An increase in area of influence along with an increase in frequency enables the removal of larger amounts of superheated liquid and inhibits the growth of the superheated liquid layer, thus reducing wall superheat.

The CHF enhancement gained through the use of the microporous coating is also believed to be a result of increased bubble departure frequency, combined with smaller departure diameters. Figure 51 shows that just prior to CHF, the latent heat component for the plain wire is approaching 100%. Zuber's [42] CHF correlation assumes hydrodynamic instability at CHF and is based on the assumption that the latent heat contribution is 100% at CHF. For the microporous wire, Fig. 51 shows that the latent heat contribution of 70% is safely away from this region of hydrodynamic instability. In addition, this reduced latent heat contribution is generated with a higher frequency of smaller diameter bubbles, i.e., less agglomeration, which also contributes to a more hydrodynamically stable environment near the heater surface. The higher microconvection component may also inhibit vapor generation at higher heat fluxes, thus further delaying the occurrence of CHF.

VIII. Summary and Conclusions

The preceding sections have analyzed the development and the performance enhancement associated with the microporous coating in nucleate pool and flow boiling of highly wetting fluids. Highly wetting fluids are inert, dielectric, and can have low boiling points making them desirable for electronics cooling applications, however, their poor thermophysical properties require the use of heat transfer enhancement methods to provide a useful operating range. The microporous coating was developed specifically to enhance nucleate boiling performance in highly wetting fluids and to be benign enough for direct application to sensitive electronic chips. The microporous coating is a mixture of small particles (1- to 20-μm-diameter) and epoxy binder that creates a thin ($\approx$ 50-μm thick), porous structure that contains re-entrant cavities. The goal of the coating development was to improve incipience using re-entrant cavities and to enhance nucleate boiling heat transfer and CHF by significantly increasing the number of active nucleation sites.

Under pool boiling conditions, the microporous coating has been shown to significantly enhance all relevant aspects of nucleate boiling in highly wetting fluids. The incipient superheat is reduced, the nucleate boiling heat

transfer coefficient is increased, and the CHF point is extended. In addition, the performance of the coating has shown to be very repeatable and predictable. A full set of correlations for predicting nucleate pool boiling heat transfer and CHF for flat and square pin-finned surfaces has been developed that takes into account the effects of pressure, subcooling, dissolved air, and heater orientation. Adhesion and durability testing indicate that the coating has excellent attachment strength and is reliable enough to survive up to 10 years of continuous service without any noticeable failure.

Under flow boiling conditions, the enhanced heat transfer benefit of the microporous coating depends on the boiling regime. In the boiling-dominated, subcooled convective boiling regime, the coating enables widespread subcooled boiling, thus providing significant heat transfer enhancements. In addition, since the small boiling bubbles generated by the coating are rapidly condensed due to subcooling, this boiling enhancement is obtained with little or no impact on internal flow pressure drop. As the subcooling level is reduced and the fluid condition approaches saturation, the coating produces a significant void fraction, which can result in higher pressure drop penalties. In the convection-dominated, high-velocity regime, the benefit of the coating is diminished. As the single-phase convection heat transfer contribution increases, the poor thermal conductivity of the coating layer can present a significant thermal resistance.

An assessment of the physics behind the enhancement provided by the microporous coating was performed. The reduction in incipience superheat is attributed to the creation of re-entrant cavities, which trap larger embryonic bubbles, thus significantly reducing the superheat required to initiate bubble growth. The consistent contributing factor to nucleate boiling enhancement using the microporous coating is the increase in bubble departure frequency. The rise in departure frequency has the effect of lowering the heater surface temperature (increasing the heat transfer coefficient) by reducing the average temperature within the superheated liquid layer that surrounds the heated surface. The CHF enhancement gained through the use of the microporous coating is also a result of increased bubble departure frequency, combined with smaller departure diameters. This combination results in less bubble agglomeration, which contributes to a hydrodynamically more stable environment near the heater surface. The high departure frequency of smaller bubbles also results in a larger microconvection contribution to the total heat removal at higher heat fluxes. This higher microconvection component may inhibit vapor generation at higher heat fluxes, thus further delaying the occurrence of CHF.

Nomenclature

c_p	specific heat capacity [J/kg · K]
d_m	average particle size [m]
D	channel hydraulic diameter [m]
D_w	wire diameter [m]
g	gravitational acceleration [m/s^2]
G	mass flux [kg/m^2 · s]
h	heat transfer coefficient [W/m^2 · K]
h_{ref}	heat transfer coefficient of reference surface [W/m^2 · K]
h_{lv}, h_{fg}	latent heat of vaporization [J/kg]
Ja	volumetric Jacob number, $\text{Ja} = (\rho_l c_p \Delta T_{\text{sub}})/(\rho_v h_{lv})$
L	fin length or heater side length [m]
L_w	wire length [m]
L'	dimensionless heater length scale; $L' = L\sqrt{g(\rho_l - \rho_v)/\sigma}$, [–]
L'_{trans}	dimensionless heater length at transition point [–]
m_p	mass of powder [kg]
P_b	bubble pressure [kPa]
P_l	liquid pressure [kPa]
P_{sys}, P_t	system pressure [kPa]
P_v	partial pressure of vapor [kPa]
P_g	partial pressure of non-condensible gas [kPa]
P_{vb}	partial pressure of vapor in bubble [kPa]
P_{gb}	partial pressure of non-condensible gas in bubble [kPa]
Pe	effective Péclet number, $\text{Pe} = \sigma^{3/4}/(\alpha\rho_v{}^{1/2}(g(\rho_l - \rho_v))^{1/4})$ [–]
q''	heat flux [W/m^2]
q''_{LH}	heat flux contribution due to latent heat [W/m^2]
q''_b	heat flux based on base surface area [W/m^2]
q''_t	heat flux based on total surface area [W/m^2]
q''_{CHF}	critical heat flux [W/m^2]
$q''_{b,\text{CHF}}$	critical heat flux based on base surface area [W/m^2]
$q''_{t,\text{CHF}}$	critical heat flux based on total surface area [W/m^2]
$q''_{\text{CHF,sat}}$	critical heat flux at saturation conditions [W/m^2]
$q''_{\text{CHF,sub}}$	critical heat flux at subcooled conditions [W/m^2]
$q''_{\text{CHF},0^\circ}$	critical heat flux at $\theta = 0^\circ$ [W/m^2]
$q''_{\text{CHF},Z}$	critical heat flux from Zuber's [42] correlation [W/m^2]
r_b	bubble radius [m]
t_c	coating thickness [m]
T_b	wall temperature at base of finned surface [K]
T_w	wall temperature [K]
T_{sat}	saturation temperature [K]
T_l	liquid temperature [K]
T_{bulk}	bulk fluid temperature [K]
U	fluid velocity [m/s]
V_b	volume of binder (epoxy) [m^3]
$\dot{V}_g$	volume flow rate of gas/vapor [m^3/s]
V_p	volume of powder [m^3]
We_D	Weber number based on hydraulic diameter; $\text{We}_D = G^2 D/\rho_l\sigma$ [–]
x_i	equilibrium inlet quality [–]

Greek Symbols

α	thermal diffusivity [m^2/s]
β	dynamic contact angle [deg]
δ_{99}	superheated liquid layer thickness [m]
ΔT_b	wall superheat at base of finned surface; $T_b - T_{\text{sat}}(P_{\text{sys}})$ [K]
ΔT_{bulk}	wall superheat; $T_w - T_{\text{bulk}}$ or $T_b - T_{\text{bulk}}$ [K]
ΔT_{inc}	wall superheat at incipience [K]
ΔT_{sat}	wall superheat; $T_w - T_{\text{sat}}(P_{\text{sys}})$ [K]
ΔT_{sub}	liquid subcooling (gassy or pure); $T_{\text{sat}}(P_{\text{sys}}) - T_{\text{bulk}}$ [K]
ε	powder porosity [–]
θ	inclination or cavity angle [deg]
ρ_v	saturated vapor density [kg/m^3]
ρ_l	saturated liquid density [kg/m^3]
ρ_p	powder density [kg/m^3]
σ	surface tension [N/m]

References

1. Kurihara, H. M. and Myers, J. E. (1960). The effects of superheat and surface roughness on boiling coefficients. *AIChE J.* **6**(1), 83–91.
2. Benjamin, R. J. and Balakrishnan, A. R. (1997). Nucleation site density in pool boiling of saturated pure liquids: effect of surface microroughness and surface and liquid physical properties. *Exp. Ther. Fluid Sci.* **15**(1), 32–42.
3. Ramilison, J. M., Sadasivan, P., and Lienhard, J. H. (1992). Surface factors influencing burnout on flat heaters. *J. Heat Transfer* **114**(1), 287–290.
4. Chowdhury, S. K. R. and Winterton, R. H. S. (1985). Surface effects in pool boiling. *Int. J. Heat Mass Transfer* **28**(10), 1881–1889.
5. Berenson, P. J. (1962). Experiments on pool-boiling heat transfer. *Int. J. Heat Mass Transfer* **5**, 985–999.
6. Nishio, S. and Chandratilleke, G. R. (1989). Steady-state pool boiling heat transfer to saturated liquid helium at atmospheric pressure. *JSME Int. J. Series II* **32**(4), 639–645.
7. Sakurai, A., Shiotsu, M., Hata, K., and Fukuda, K. (1997). Effect of surface conditions on pool boiling critical heat fluxes in saturated and subcooled liquid nitrogen at various pressures. *In* ''Advances in Cryogenic Engineering'' (P. Kittel, ed.), Vol. 43, pp. 1305–1312. Plenum Press, New York.
8. Yu, J., Momoki, S., and Koyama, S. (1999). Experimental study of surface effect on flow boiling heat transfer in horizontal smooth tubes. *Int. J. Heat Mass Transfer* **42**, 1909–1918.
9. Kandlikar, S. G. and Spiesman, P. H. (1998). Effect of surface finish on flow boiling heat transfer. *ASME Publ.* HTD-Vol. 361-1, 157–163.
10. Groenveld, D. C. (1981). ''Heat Transfer Phenomena Related to the Boiling Crisis''. AECL-7239, Chalk River Nuclear Laboratory.
11. Chernobai, V. A. (1980). Model for heat transfer crisis for water boiling in pipes. *High Temperature—USSR* **18**, 797–801.
12. Chaudri, I. H. and McDougall, I. R. (1969). Aging studies in nucleate pool boiling of isopropyl acetate and perchlorethylene. *Int. J. Heat Mass Transfer* **12**, 681–688.
13. Nishikawa, K., Fujita, Y., Ohta, H., and Hidaka, S. (1982). Effect of the surface roughness on the nucleate boiling heat transfer over the wide range of pressure. *In* ''Proceedings of the 7th International Heat Transfer Conference'', pp. 61–66. Munich, Germany.
14. Griffith, P. and Wallis, J. D. (1960). The role of surface conditions in nucleate boiling. *Chem. Eng. Prog. Symp. Ser No. 49* **56**, 49–63.
15. Thome, J. R. (1990). ''Enhanced Boiling Heat Transfer'', pp. 64–141. Hemisphere Publishing, Washington, D.C.
16. Thome, J. R. (1992). Mechanisms of enhanced nucleate pool boiling. *In* ''Proceedings of the Engineering Foundation Conference on Pool and External Flow Boiling'', pp. 337–343. ASME, New York.
17. Bergles, A. E. (1997). Enhancement of pool boiling. *Int. J. Refrig.* **20**(8), 545–551.
18. Webb, R. L. (1994). ''Principles of Enhanced Heat Transfer'', Chapter 11. pp. 311–372. Wiley Interscience, New York.
19. Webb, R. L. (1983). Nucleate boiling on porous coated surfaces. *Heat Transfer Eng.* **4**(3–4), 71–82.
20. Webb, R. L. and Chien, L. H. (1999). Boiling on structured surfaces. *In* ''Heat Transfer Enhancement of Heat Exchangers'' (S. Kakaç *et al.*, eds.), pp. 249–284. Kluwer Academic Publishers, Netherlands.

21. Chien, L. H. and Webb, R. L. (1998). Visualization of pool boiling on enhanced surfaces. *Exp. Ther. Fluid Sci.* **16**(4), 332–341.
22. Nakayama, W., Daikoku, T., and Nakajima, T. (1982). Effects of pore diameters and system pressure on saturated pool nucleate boiling heat transfer from porous surfaces. *J. Heat Transfer* **104**, 286–291.
23. Czikk, A. M., O'Neill, P. S., and Gottzmann, C. F. (1981). Nucleate pool boiling from porous metal films: effect of primary variables. *ASME Publ.* HTD-Vol. 18, 109–122.
24. Wadekar, V. V. (1998). A comparative study of in-tube boiling on plain and high flux coated surfaces. *Enhanced Heat Transfer* **5**, 257–264.
25. Ikeuchi, M., Yumikura, T., Fujii, M., and Yamanaka, G. (1984). Heat-transfer characteristics of an internal micro-porous tube with refrigerant-22 under evaporating conditions. *ASHRAE Trans.* **90**(Part 1A), 196–211.
26. Lallemand, M., Branescu, C., and Haberschill, P. (2001). Local heat transfer coefficients during boiling of R22 and R407C in horizontal smooth and microfin tubes. *Int. J. Refrig.* **24**, 57–72.
27. Cheng, L. and Chen, T. (2001). Study of flow boiling heat transfer in a tube with axial microgrooves. *Exp. Heat Transfer* **14**, 59–73.
28. Celato, G. P. and Cumo, M. (1992). Burnout at very high heat fluxes. *Wärme- und Stoffübertragung* **27**, 233–244.
29. Mizunuma, H., Behnia, M., and Nakayama, W. (2002). Forced convective boiling of a fluorocarbon liquid in reduced size channels—an experimental study. *J. Enhanced Heat Transfer* **9**, 69–76.
30. Maddox, D. E. and Mudawar, I. (1989). Single- and two-phase convective heat transfer from smooth and enhanced microelectronic heat sources in a rectangular channel. *J. Heat Transfer* **111**, 1045–1052.
31. Kandlikar, S. G. and Howell, M. J. (1996). Investigation of nucleation and heat transfer for subcooled flow boiling on microfin surfaces. *2nd European Thermal-Sciences and 14th UIT National Heat Transfer* **1**, 241–248.
32. Tong, W., Bar-Cohen, A., Simon, T. W., and You, S. M. (1990). Contact angle effects on boiling incipience of highly-wetting liquids. *Int. J. Heat Mass Transfer* **33**(1), 91–103.
33. Yang, W. J., Takizawa, H., and Vrable, D. L. (1991). Augmented boiling on copper-graphite composite surface. *Int. J. Heat Mass Transfer* **34**(11), 2751–2758.
34. Liang, H. S. and Yang, W. J. (1998). Nucleate pool boiling heat transfer in a highly-wetting liquid on micro-graphite-fiber composite surfaces. *Int. J. Heat Mass Transfer* **41**(13), 1993–2001.
35. Honda, H., Takamastu, H., and Wei, J. J. (2002). Enhanced boiling of FC-72 on silicon chips with micro-pin-fins and submicron-scale roughness. *J. Heat Transfer* **124**(2), 383–390.
36. Mudawar, I. and Anderson, T. M. (1989). High flux electronics cooling by means of pool boiling—Part II: Optimization of enhanced surface geometry. *ASME Publ.* HTD-Vol. 111, 35–49.
37. Guglielmini, G., Misale, M., and Schenone, C. (1996). Experiments on pool boiling of a dielectric fluid on extended surfaces. *Int. Commun. Heat Mass Transfer* **23**(4), 451–462.
38. Honda, H. and Wei, J. J. (2003). Advances in enhanced boiling heat transfer from electronic components. Presented at the 6th ASME/JSME Thermal Engineering Joint Conference, March 16–20, Hawaii Island, Hawaii, USA.
39. You, S. M., Simon, T. W., and Bar-Cohen, A. (1992). A technique for enhancing boiling heat transfer with application to cooling of electronic equipment. *IEEE Trans. Comp., Hybrids, Manuf. Technol.* **15**, 823–831.
40. McAdams, W. H. (1954). "Heat Transmission", pp. 378–408. McGraw-Hill, New York.

41. Willingham, T. C. and Mudawar, I. (1992). Channel height effects on forced-convection boiling and critical heat flux from a linear array of discrete heat sources. *Int. J. Heat Mass Transfer* **35**(8), 1865–1880.
42. Zuber, N., (1959). "Hydrodynamic Aspects of Boiling Heat Transfer", AEC Report No. AECU-4439, Physics and Mathematics.
43. Rainey, K. N. and You, S. M. (2000). Pool boiling heat transfer from plain and microporous, square pin-finned surfaces in saturated FC-72. *J. Heat Transfer* **122**(3), 509–516.
44. You, S. M., Simon, T. W., and Bar-Cohen, A. (1990). Experiments on boiling incipience with a highly-wetting dielectric fluid: effects of pressure, subcooling and dissolved gas content. *Proc. 9th Int. Heat Transfer Conf.* **2**, 337–342.
45. Hong, Y. S., Ammerman, C. N., and You, S. M. (1997). Boiling characteristics of cylindrical heaters in saturated, gas saturated, and pure-subcooled FC-72. *J. Heat Transfer* **119**, 313–318.
46. You, S. M., Simon, T. W., and Bar-Cohen, A. (1990). Experiments on nucleate boiling heat transfer with a highly-wetting dielectric fluid: effects of pressure, subcooling and dissolved gas content. *ASME Publ.* HTD-Vol. 131, 45–52.
47. Bar-Cohen, A. (1994). Fundamentals of nucleate pool boiling of highly-wetting dielectric liquids. *In* "Cooling of Electronic Systems" (S. Kakaç *et al.*, eds.), pp. 415–455. Kluwer Academic Publishers, Netherlands.
48. O'Connor, J. P. and You, S. M. (1995). A painting technique to enhance pool boiling heat transfer in saturated FC-72. *J. Heat Transfer* **117**, 387–393.
49. Chang, J. Y. and You, S. M. (1997). Enhanced boiling heat transfer from microporous surfaces: effects of a coating composition and method. *Int. J. Heat Mass Transfer* **40**, 4449–4460.
50. O'Connor, J. P., You, S. M., and Price, D. C. (1995). A dielectric surface coating technique to enhance boiling heat transfer from high power microelectronics. *IEEE Transactions on Components, Hybrids and Manufacturing Technology Part A* **18**(3), 656–663.
51. Chang, J. Y. and You, S. M. (1997). Boiling heat transfer phenomena from microporous and porous surfaces in saturated FC-72. *Int. J. Heat Mass Transfer* **40**, 4437–4447.
52. O'Neill, P. S., Gottzmann, C. F., and Terbot, J. W. (1972). Novel heat exchanger increases cascade cycle efficiency for natural gas liquefaction. *Adv. Cryogenic Eng.* **17**, 420–437.
53. Hsu, Y. Y. (1962). On the size range of active nucleation cavities on a heating surface. *J. Heat Transfer* **84**, 207–216.
54. Carslaw, H. S. and Jaeger, J. C. (1986). "Conduction of Heat in Solids", 2nd edn., pp. 75–77. Clarendon Press, Oxford.
55. Ammerman, C. N., You, S. M., and Hong, Y. S. (1996). Identification of pool boiling heat transfer mechanisms from a wire immersed in saturated FC-72 using a single-photo/LDA method. *J. Heat Transfer* **118**, 117–123.
56. Bar-Cohen, A. and Simon, T. W. (1988). Wall superheat excursions in the boiling incipience of dielectric fluids. *Heat Transfer Eng.* **9**(3), 19–31.
57. Bankoff, S. G. (1958). Entrapment of gas in the spreading of a liquid over a rough surface. *AIChE J.* **4**(1), 24–26.
58. Chang, J. Y. and You, S. M. (1996). Heater orientation effects on pool boiling of microporous-enhanced surfaces in saturated FC-72. *J. Heat Transfer* **118**(4), 937–943.
59. Rainey, K. N. and You, S. M. (2001). Effects of heater orientation on pool boiling heat transfer from microporous coated surfaces. *Int. J. Heat Mass Transfer* **44**(14), 2589–2599.
60. Park, K. A. and Bergles, A. E. (1988). Effects of size of simulated microelectronic chips on boiling and critical heat flux. *J. Heat Transfer* **110**, 728–734.

61. Saylor, J. R., Simon, T. W., and Bar-Cohen, A. (1989). The effect of a dimensionless length scale on the critical heat flux in saturated, pool boiling. *ASME Publ.* HTD-Vol. 108, 71–80.
62. Bar-Cohen, A. and McNeil, A. (1992). Parametric effects of pool boiling critical heat flux in dielectric liquids. *In* "Proceedings of the Engineering Foundation Conference on Pool and External Flow Boiling", Santa Barbara, CA, pp. 171–175.
63. Rainey, K. N., You, S. M., and Lee, S. (1992). Effect of pressure, subcooling, and dissolved gas on pool boiling heat transfer from microporous surfaces in FC-72. *J. Heat Transfer* **125**(1), 75–83.
64. Mudawar, I. and Anderson, T. M. (1990). Parametric investigation into the effects of pressure, subcooling, surface augmentation and choice of coolant on pool boiling in the design of cooling systems for high-power-density electronic chips. *J. Electronic Packaging* **112**, 375–381.
65. Watwe, A. A., Bar-Cohen, A., and McNeil, A. (1997). Combined pressure and subcooling effects on pool boiling from a PPGA chip package. *J. Electronic Packaging* **119**, 95–105.
66. Chang, J. Y., You, S. M., and Haji-Sheikh, A. (1998). Film boiling incipience at the departure from natural convection on flat, smooth surfaces. *J. Heat Transfer* **120**(2), 402–409.
67. Carey, V. P. (1992). "Liquid-Vapor Phase-Change Phenomena", pp. 299–310. Taylor and Francis, Bristol, PA.
68. Yin, C. P., Yan, Y. Y., Lin, T. F., and Yang, B. C. (2000). Subcooled flow boiling heat transfer of R-134a and bubble characteristics in a horizontal annular duct. *Int. J. Heat Mass Transfer* **43**(11), 1885–1896.
69. Wiebe, J. R. and Judd, R. L. (1971). Superheat layer thickness measurements in saturated and subcooled nucleate boiling. *J. Heat Transfer* **93**(4), 455–461.
70. Marek, R. and Straub, J. (2001). The origin of thermocapillary convection in subcooled nucleate pool boiling. *Int. J. Heat Mass Transfer* **44**(3), 619–632.
71. Rohsenow, W. M. (1962). A method of correlating heat transfer data for surface boiling of liquids. *J. Heat Transfer* **74**, 969–975.
72. Hong, Y. S., Ammerman, C. N., and You, S. M. (1998). Effects of length scale, subcooling, and dissolved gas content on the pool boiling critical heat flux of cylindrical heaters. *In* "Proceedings of the 11th International Heat Transfer Conference", pp. 389–394. Kyongju, Korea.
73. Inoue, T., Kawae, N., and Monde, M. (1998). Effect of subcooling on critical heat flux during pool boiling on a horizontal wire. *Heat Mass Transfer* **33**(56), 481–488.
74. Elkassabgi, Y. and Lienhard, J. H. (1988). Influences of subcooling on burnout of horizontal cylindrical heaters. *J. Heat Transfer* **110**, 479–486.
75. Chang, J. Y. and You, S. M. (1997). Enhanced boiling heat transfer from micro-porous cylindrical surfaces in saturated FC-87 and R-123. *J. Heat Transfer* **119**, 319–325.
76. Ammerman, C. N. and You, S. M. (2001). Enhancing small-channel convective boiling performance using a microporous surface coating. *J. Heat Transfer* **123**, 976–983.
77. Rainey, K. N., Li, G., and You, S. M. (2001). Flow boiling heat transfer from plain and microporous coated surfaces in subcooled FC-72. *J. Heat Transfer* **123**(5), 918–925.
78. Mudawar, I. and Maddox, D. E. (1989). Critical heat flux in subcooled flow boiling of fluorocarbon liquid on a simulated electronic chip in a vertical rectangular channel. *Int. J. Heat Mass Transfer* **32**(2), 379–394.
79. Haramura, Y. and Katto, Y. (1983). A new hydrodynamic model of critical heat flux, applicable widely to both pool and forced convection boiling on submerged bodies in saturated liquids. *Int. J. Heat Mass Transfer* **26**(3), 389–399.

80. Ammerman, C. N. and You, S. M. (1998). Consecutive-photo method to measure volume flow rate during boiling from a wire immersed in saturated liquid. *J. Heat Transfer* **120**(3), 561–567.
81. Kim, J. H., Rainey, K. N., You, S. M., and Pak, J. Y. (2002). Mechanism of nucleate boiling heat transfer enhancement from microporous surfaces in saturated FC-72. *J. Heat Transfer* **124**(3), 500–506.
82. Paul, D. D. and Abdel-Khalik, S. I. (1983). A statistical analysis of saturated nucleate boiling along a heated wire. *Int. J. Heat Mass Transfer* **26**, 509–519.
83. Ammerman, C. N. and You, S. M. (1996). Determination of the boiling enhancement caused by surfactant addition to water. *J. Heat Transfer* **118**, 561–567.

Robust Algorithms for Tomographic Reconstruction of Line-of-Sight Thermal Flow Field Images

KENNETH D. KIHM[1], KOJI OKAMOTO[2], and HAN SEO KO[3]

[1]*Department of Mechanical Engineering, Texas A&M University, College Station, Texas, USA*
[2]*Department of Quantum Engineering & Systems Science, The University of Tokyo, Japan*
[3]*School of Mechanical Engineering, Sungkyunkwan University, Korea*

I. Introduction

Once properly formulated, tomographic image reconstruction is analogous to a function optimization for the reconstructing field under a constraint specified by a type of projection. A new tomographic image reconstruction technique is presented that uses Genetic Algorithm (GA), a robust and combinatorial function optimization based on the genetic principle, i.e., natural selection and survival of the fittest from the pool of solution candidates. The parallel optimization feature of GA allows simultaneous optimization of multiple parameters, while most conventional tomographic techniques reconstruct images for a single-parameter, such as in algebraic reconstruction technique (ART). For example, the GA-based tomography has inherent advantages for multi-phase flows that often require simultaneous reconstruction of bubble locations, sizes, and shapes. The GA requires the fitness values of the solution candidates as the only feedback information for the solution searching and evolution, not requiring an inversion of the projection data in order to update the reconstructing field. Thus, the GA-based tomography can readily reconstruct non-linear projections such as laser speckle photography, which requires additional manipulations for a linearization before the inversion. Also, a hybrid idea of using a Concurrent Simplex (CS) with GA is presented to expedite the evolving calculation steps of the GA-based tomography computations.

Most optical visualization techniques measure line-of-sight projected images that integrate optical information along the beam path occupied by a test field. Unless the optical information is uniformly distributed along

Advances in Heat Transfer
Volume 38 ISSN 0065-2717
DOI 10.1016/S0065-2717(04)38003-2

the beam path, a sophisticated reconstruction process is necessary to recover the cross-sectional field information from the projected images. Such procedures are generally called *optical tomography* or tomographic image reconstruction.

An axisymmetric test field requires only a single projection at any arbitrary projection angle to reconstruct the cross-sectional information that varies only along the radius [1]. For asymmetric test fields, however, a number of projections from many different angles must be simultaneously interrogated to optimally determine the cross-sectional information. Stacking up the reconstructed two-dimensional information will construct a full three-dimensional field. Existing tomographic techniques belong to either of two primary categories: a direct mathematical calculation of the cross-sectional field from projected images [2,3] or a successive iteration of the field by regression optimization to match the projection images [4,5].

The former category, referred to as Fourier Convolutions (FC) based on the Fourier slicing theorem [6], usually requires a large number of projections to ensure acceptable accuracy for one-step mathematical inversion. Spatial and other constraints often limit the number and interval of projections. The reconstruction accuracy of the FC technique can be noticeably degraded for the case of limited projections [7]. Nevertheless, the FC technique is widely used for medical imaging systems such as MRI (magnetic resonance imaging) or CT (computer tomography), where the field (patient's body) remains steady during the period of multiple image recordings and allows hundreds of equally angled projections. On the other hand, other problems involving time-dependent phenomena, typically such as in engineering measurements, require a simultaneous recording with a sufficiently short temporal resolution.

The latter category of regression optimization, represented by the algebraic reconstruction technique (ART), usually needs fewer projections since its regression feature allows iterative solution search [8]. However, the ART is a monotonic optimization searching for the solution field gradient and it carries a danger of entrapment by a local peak [9]. Also, when the number of iterations exceeds a certain optimum the ill-posedness in ART can increase artifacts and degrade the reconstruction [10]. In addition, the iteration of ART relies on the feedback information based on the difference between the measured projection of the actual field and the virtual projection from the guessed field. The discrepancy between the two projections is back projected to update the previously guessed field and the iterative cycle repeats until a convergence is met. When the forward projection is a linear summation of the field values along the light ray, the back projection of the feedback information is easily obtained.

Fortunately, many optical projections devise a linear summation of the field (e.g., various imaging techniques based on light attenuation, such as X-ray imaging in proportion to the medium density, MRI based on the isotope concentration, and interferometry where the fringe shift is a direct integration of the field density). The simple and straight features of ART techniques are widely accepted for linear projection problems [11]. However, for non-linear imaging such as laser speckle photography [12], the use of ART needs an additional manipulation for linearization of the projected data to acquire proper feedback information [13].

Genetic algorithm or GA [14], the combinatorial and robust function optimization scheme, can contribute to establishing a new tomographic method to alleviate the afore mentioned problems associated with existing techniques. The genetic principle implies natural selection and survival of the fittest based on parallelism, elitism, and scaled randomness. Computer scientists and mathematicians have long recognized the superiority of GA in finding solutions over other optimization schemes. The robust feature of GA is expected to overcome the local-peak-trap problem with progressive optimization.

Literature on thermal engineering applications of GA is scarce. Tsukiji *et al*. [15] proposed a new GA particle tracking method to improve the effectiveness of identifying particle pairs between sequential flow images for PIV (particle image velocimetry). They showed that GA was advantageous in that the particle image pairs were identified faster and more accurately compared with the conventional correlation method. The pioneering work by Tsuru *et al*. [16] is the first published attempt to use GA for tomographic flow visualization. Their results showed a viability of the GA-based tomography.

This chapter describes the principle of GA and its application for tomography. Section II describes the line-of-sight nature of prevalent optical imaging techniques for engineering and medical applications, and Section III presents a brief review of two conventional tomographic reconstruction techniques, namely Fourier Transformation Technique (FTT) and Algebraic Reconstruction Technique (ART), and their application examples. Section IV introduces the genetic algorithm and its relevance to tomographic reconstruction. To enhance understanding of GA, the section also presents example calculations using computer synthesized phantom fields and examines the reconstruction accuracy. Section V describes a hybrid GA tomographic technique with a Concurrent Simplex (CS) optimization. The section presents illustrative example calculations to examine the reconstruction efficiency of the hybridized GA tomographic algorithm. Section VI presents a comprehensive three-dimensional reconstruction technique and its application.

Section VII presents a concluding remark with recommended future research direction.

II. Line-of-Sight Imaging

The line-of-sight imaging is widely used as described in the Introduction. Usually, the beam is used as the detecting materials, including photon beam (visible light, X-ray, gamma-ray, etc.), neutron beam, electron beam, ultrasound beam, and so on. These beams run straight from the beam generator to a detector. The detected value is usually the function of line integral along the pass-line or the line-of-sight, ℓ (Fig. 1).

$$p = \int_{\ell} f(x)ds \tag{1}$$

$$\frac{dp}{ds} = f(x) \tag{2}$$

where p is the detected value, typically beam intensity and phase information. $f(x)$ is the local property at location x, which may be a function of temperature, pressure, density, material, beam intensity (p), and so on. The detected value p, has the whole integrated information of the local properties. So, if $f(x)$ is reconstructed from the integrated value, p, we should know proper and sufficient clue on the local property distributions, which is the representation of temperature, pressure and so on.

For a linear projection case, the function $f(x)$ can be expressed as either $f(x) = g(x)$ or $f(x) = g(x)p$ and Eqs. (1) and (2) are expressed as,

$$p = \int_{\ell} g(x)ds \tag{3}$$

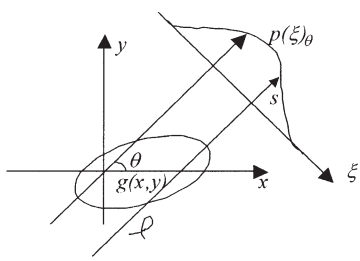


FIG. 1. Schematic of the line-of-sight projection data.

$$p = p_0 \exp\left(\int_\ell g(x)ds\right) \tag{4}$$

The first equation is usually used in the phase detection, e.g., refraction index integral. The second one is the typical equation for the intensity detection, e.g., X-ray image. In the phase detection case, the property $g(x)$ is a function of local refraction index, which is a function of density, temperature, and so on. If the variation along the pass-line, ℓ, is negligible, Eq. (3) is reduced as, $p = g\ell$. Then, if we can detect the phase, p, it directly converted to local property, g, i.e., temperature, density, and so on. This case is the simplest expression of the line-of-sight imaging.

For the intensity detection case, Eq. (4), the function $g(x)$ usually has a negative sign, which represent the absorption or attenuation of the beam by a medium material. The absorption is a function of material, e.g., air, water, or organs. A well-known example will be the X-ray imaging of the human body for medical uses. It displays the organs, bones, and so on, depending on the differentiated levels of X-ray attenuation in different medium materials and the X-ray line-of-sight path lengths.

In practice, the function $g(x)$ is not uniform along the pass-line, ℓ. Therefore, the distribution function $g(x)$ cannot be reconstructed directly from the single projection data, p. To properly reconstruct the data, we must have sufficiently additional information. Theoretically, we can accurately reconstruct the distribution, $g(x)$, only with infinitely many projections at different angles, but practically, we can only have a finite number of projections, which are often severely limited because of the geometrical as well as economical constraints. This chapter presents a concise discussion on a new idea for robust reconstruction schemes under limited number of projections.

III. Conventional Tomographic Reconstruction Techniques

This chapter briefly reviews two conventional tomographic reconstruction schemes, namely, A. Fourier Transformation Technique (FTT) and B. Algebraic Reconstruction Technique (ART) and Multiplicative Algebraic Reconstruction Technique (MART).

A. Fourier Transformation Technique (FTT)

Let the problem be three-dimensional in space, Eq. (1) is re-written as,

$$p(\xi, \eta)_{\theta,\varphi} = \int_{\ell(\theta,\varphi)} f(x, y, z)d\ell \tag{5}$$

where p represents the projection image, and ξ, η are the local coordinates on the projection image. The projection angles are θ, ϕ, and the global coordinates are x, y, z. We may mathematically and sufficiently reconstruct the distribution $f(x, y, z)$ from infinite projection images $p(\xi, \eta)_{\theta,\phi}$. For simplicity, in most cases, the cross-sectional distribution $f(x, y)$, instead of the full three-dimensional distribution $f(x, y, z)$, is reconstructed from the projection data $p(\xi)_\theta$. Then, the cross-section data will be stacked in z-direction to conform to the three-dimensional data $f(x, y, z)$.

For the case of cross-sectional reconstruction, Eq. (5) is rewritten as,

$$p(\xi)_\theta = \int_{\ell(\theta)} f(x, y) d\ell = \int\int_A f(x, y)\, \delta(x \cos\theta + y \sin\theta - \xi) dx dy \tag{6}$$

Fourier Transformation of Eq. (5) is expressed as (Fig. 2),

$$F(\alpha, \beta) = \int\int f(x, y) e^{(\alpha x + \beta y)i} dx dy \tag{7}$$

$$F(\xi \cos\theta, \xi \sin\theta) = \int\int p(\xi)_\theta e^{-(\alpha x + \beta y)i} dx dy \tag{8}$$

Then, the Fourier Transformed data F will be a data along the line on the Fourier domain as schematically shown in Fig. 2. By varying the projection angle θ, the Fourier domain will be filled with many data lines. This shows that the full field of the Fourier domain $F(\alpha, \beta)$, or equivalently the reconstructed data field, can be determined by sufficiently many projection angles θ. After the complete Fourier domain data are calculated from many projection data $p(\xi)_\theta$, the inverse Fourier transformation of $F(\alpha, \beta)$

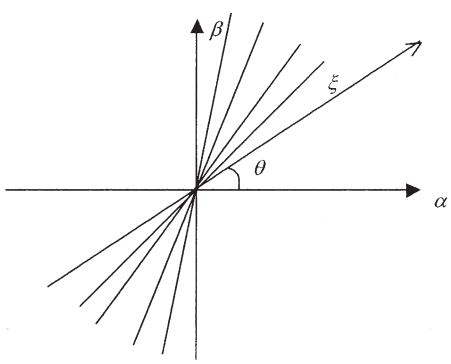


FIG. 2. Fourier Domain $F(\alpha, \beta)$.

gives the actual distribution data $f(x, y)$. To fill-out the Fourier domain, we need sufficiently many sets of projection data at different projection angles, preferably at equal divisions.

B. Algebraic Reconstruction Technique (ART) and Multiplicative Algebraic Reconstruction Technique (MART)

For a cross-sectional density field, one can represent the field as a series of basis functions allowing their parameters to be optimally determined. The tomography undertakes the optimization task for the linear case where each basis function is defined by a single parameter (usually its unknown height with a fixed spread). The location of each basis function is given as

$$\hat{f}(x, y) = \sum_{n=1}^{ij} O_n b(x - x_n, y - y_n) \tag{9}$$

where $\hat{f}$ is an object function that represents the field to be reconstructed, b is a general form of the basis function located at (x_n, y_n), and O_n is the height coefficient of the nth basis function centered at a fixed location of (x_n, y_n). The (x_n, y_n) positions form a rectangular array of i equally spaced points in the x direction and j in the y direction as schematically illustrated in Fig. 3. Thus, ij is the total number of coefficients to be estimated by the reconstruction algorithm. The use of a smooth basis function such as the cubic B-spline function can accurately represent a relatively smooth object field with far fewer coefficients (unknowns) than with the square-pixel basis function. An optimized set of these unknowns must be found to minimize the deviations between the virtual projection $\hat{p}$ of an intermediate object function $\hat{f}$ and the measured projection p of the actual field f.

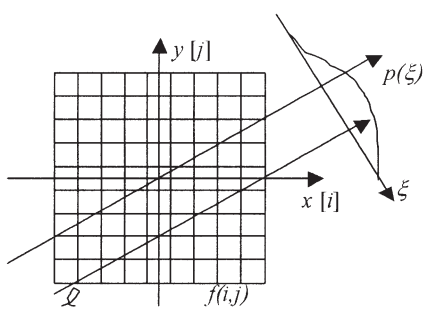


Fig. 3. Concept of Algebraic Reconstruction Technique (ART).

The projected ray sum p_m in Fig. 3 can be obtained as follows:

$$p_m = \sum_{n=1}^{ij} O_n \int_l b(x - x_n, y - y_n)dl \tag{10}$$

where $m = 1, 2, \ldots, PR$ is the total number of ray sums for the number of equally angled projections P, and the number of ray sums per each projection R. Equation (10) can be expressed in matrix form as follows:

$$\boldsymbol{p} = W\boldsymbol{O} \tag{11}$$

where $\boldsymbol{p}$ is the measurement ray sum vector, W is the projection matrix, and $\boldsymbol{O}$ is the object vector to be reconstructed. Solving this set of PR linear algebraic equations (one equation for each measured ray sum) is the goal of the series-expansion technique such as the ART or MART in ij unknowns.

The ART algorithm is the basically iterative solver of systems of linear equations [Eq. (11)] adapted to the particular problem of tomographic reconstruction for an object vector $\boldsymbol{O}$. The ART uses the feedback information on the deviation of the virtual projection from the measured projection and iteratively optimizes the object coefficient vector $\boldsymbol{O}$ by an algebraic updating as follows:

$$O^{q+1} = O^q + \frac{p_m - \langle w_m, O^q \rangle}{\left(\sum_{n=1}^{ij} w_{m,n}\right)^2} w_m, \qquad \sum_{n=1}^{ij} w_{m,n} \neq 0 \tag{12}$$

where $\langle x, y \rangle$ denotes the inner product of vectors x and y, q indicates the iteration number, w_m is the mth row of the projection matrix, and p_m is the corresponding measured ray sum. If $\sum_{n=1}^{ij} w_{m,n} = 0$, O is left unchanged. This algorithm performs the iteration on a ray-by-ray basis until convergence is reached.

Equation (11) can be expressed by the use of the matrix equation as follows:

$$\langle w_m, O \rangle = \begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \\ \vdots \\ \hat{p}_m \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & & & \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{bmatrix} \begin{bmatrix} O_1 \\ O_2 \\ \vdots \\ O_n \end{bmatrix} \tag{13}$$

where the first matrix denotes the virtual projection from calculation, the second matrix is termed the projection matrix, and the third is the object matrix to be reconstructed iteratively. Thus, the numerator of Eq. (12), the deviation between the virtual projection $\hat{p}_m$ and the measured projection p_m of the actual field, can also be written from Eq. (13):

$$p_m - \langle w_m, O^q \rangle = p_m - \hat{p}_m \tag{14}$$

where $m = 1, 2, \ldots, PR$. The coefficients of the projection matrix for the case of the square-pixel basis function are all zero and one. If the ray passes through a square-pixel, then the value of the coefficient of the square-pixel is one. Otherwise, it is zero. However, the coefficients of the projection matrix for the cubic B-spline basis function have all different values depending on the distance from the fixed location (x_n, y_n). Thus, the cubic B-spline basis function is advantageous for the relatively complex object.

Multiplicative algebraic reconstruction technique (MART) differs from the ART only in the way the deviation between the virtual projection $\hat{p}_m$ and the measured projection p_m is distributed among the object coefficients. The MART uses an element C_n of the multiplicative correction vector $\boldsymbol{C}$ as follows:

$$\begin{aligned} O_n^{q+1} &= C_n^q O_n^q \\ C_n^q &= \begin{cases} 1 - 0.5 W_{m,n}\left(1 - \dfrac{p_m}{\hat{p}_m}\right), & \hat{p}_m \neq 0 \\ 1, & \text{otherwise} \end{cases} \end{aligned} \tag{15}$$

where q denotes the qth iteration and the normalized weighting factor $W_{m,n}$ is equal to $w_{m,n}/w_{\max}$ where $w_{\max}$ is the largest element of the projection matrix W. One advantage of using MART is to ensure a non-negative object field in reconstructing non-negative scalar.

For a three-dimensional reconstruction, x, y, and z directions are used to express the location of each basis function as follows:

$$\hat{f}(x, y, z) = \sum_{n=1}^{ijk} O_n b(x - x_n, y - y_n, z - z_n) \tag{16}$$

where O_n is the height coefficient of the nth basis function centered at a fixed location of (x_n, y_n, z_n) which can be shown by a rectangular array of i equally spaced points in the x direction, j in the y direction, and k in the z direction. Thus, ijk is the total number of coefficients to be calculated by the three-dimensional reconstruction algorithm.

The number of the projected ray sum p_m also changes for the three-dimensional reconstruction as follows:

$$p_m = \sum_{n=1}^{ijk} O_n \int_l b(x - x_n, y - y_n, z - z_n)dl \tag{17}$$

where $m = 1, 2, \ldots,$ PZR is the total number of ray sums for the number of equally angled projections in one plane P, the number of planes in z-direction Z, and the number of ray sums per each projection R. The simultaneous three-dimensional tomography can be performed using the projected data with various heights from multiple projection angles of view instead of piling up the two-dimensional reconstructions. Thus, the multiple projection angles in one plane, multiple planes in z-direction, and multiple rays in each projection angle are required for the simultaneous three-dimensional tomography.

C. Applications of Fourier Transformation Technique (FTT) and Algebraic Reconstruction Technique (ART)

The FTT and the ART for tomographic reconstruction are computationally investigated using two different computer synthesized phantoms. The two phantoms: Cosine phantom and a CosGauss phantom have been previously suggested as reference fields in literature [11]. The Cosine phantom is a smooth and asymmetric function whose values and gradients at the edges of its domain are zero [13]:

$$\begin{aligned} f(x, y) &= \frac{1}{4}\left[1-\cos\left\{2\pi(x+0.5)^{4/5}\right\}\right]\left[1-\cos\left\{2\pi(y+0.5)^{2/3}\right\}\right], \quad |x|<0.5 \text{ and } |y|<0.5 \\ &= 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise} \end{aligned} \tag{18}$$

This phantom constructs an asymmetric single peak as shown in Fig. 4(a). The CosGauss phantom is a superposition of the Cosine phantom and the two Gaussian peaks centered at ($x_1 = 0.20$, $y_1 = 0.10$) and ($x_2 = -0.20$, $y_2 = -0.35$) as shown in Fig. 4(b) [13]:

$$\begin{aligned} f(x, y) = {} & 0.327\text{Cosine}(x, y) \\ & + 0.872 \exp\left\{-[9(x - 0.2)]^2 - [6(y - 0.1)]^2\right\} \\ & + 0.872 \exp\left\{-[8(x + 0.2)]^2 - [30(y + 0.35)]^2\right\} \end{aligned} \tag{19}$$

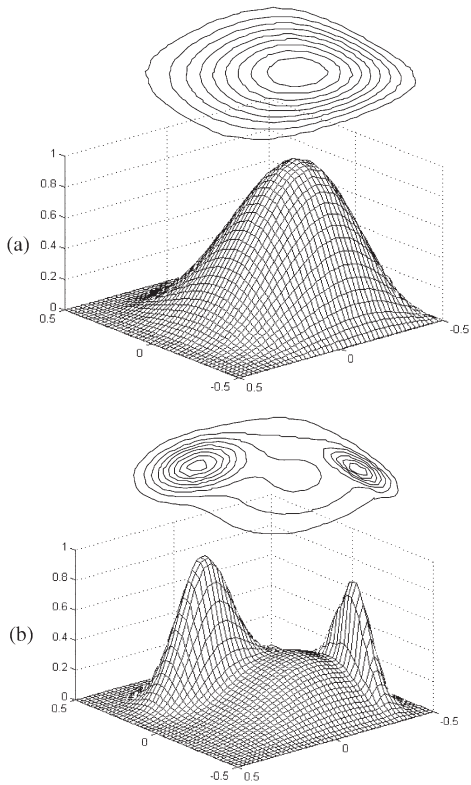


FIG. 4. Reference phantom. (a) Cosine function and (b) CosGauss function.

Because the Cosine and the CosGauss phantoms are asymmetric, 5 and 3 projection angles are used to reconstruct the density field. These projection angles are shown over the contours of the phantoms in Fig. 5.

Three different error measurements are used in this research [13]. The first is the average error of the reconstructed object function $\hat{f}$ and the reference phantom function f:

$$\Phi_{\text{avg}} = \frac{\sum_{n=1}^{ij} |f(x_n, y_n) - \hat{f}(x_n, y_n)|}{ij} \tag{20}$$

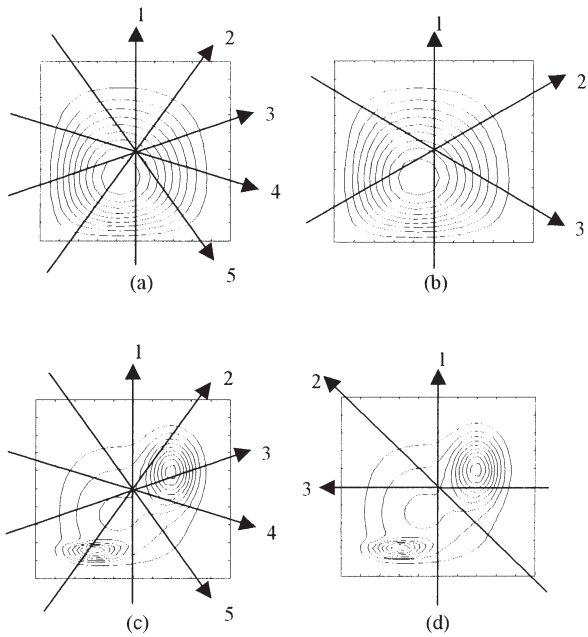


FIG. 5. Projection angles of reference phantom fields. (a) 5 projections of Cosine function, (b) 3 projections of Cosine function, (c) 5 projections of CosGauss function, and (d) 3 projections of CosGauss function.

where ij is the total number of the basis functions used to conform to the reconstructing object functions. The second is a normalized rms error:

$$\Phi_{\text{rms}} = \sqrt{\frac{\sum_{n=1}^{ij}\left[f(x_n, y_n) - \hat{f}(x_n, y_n)\right]^2}{\sum_{n=1}^{ij}\left[f(x_n, y_n) - \bar{f}\right]^2}} \tag{21}$$

where $\bar{f}$ is the average value of the phantom field f. The normalized rms error is large if there are a few large errors in the reconstruction. The third one is a normalized absolute error:

$$\Phi_{\text{abs}} = \frac{\sum_{n=1}^{ij} |f(x_n, y_n) - \hat{f}(x_n, y_n)|}{\sum_{n=1}^{ij} |f(x_n, y_n)|} \tag{22}$$

which emphasizes the effect of many small errors.

Note that these three errors measure the reconstruction quality based on the comparison between the reconstructed field (object function) and

the true field (phantom function). In a real experiment, however, the true field is unknown and the quality of reconstruction is only measured by comparing the virtual projection against the measured projection.

The two reconstruction algorithms, FTT [Eq. (8)] and ART, were used for reconstruction of the two phantom fields under the speckle photography and the interferometry projections. The speckle photographic projection of the phantoms was simulated by analytically differentiating the functional representations of Eqs. (18) and (19) in the normal direction to the ray. For the Cosine phantom each projection consists of 45 rays ($R = 45$); the object field is described by 15×15 discrete pixels for the FTT reconstruction, and by 15×15 cubic B-spline basis functions for the ART ($ij = 225$). For the case of the CosGauss phantom each projection consists of 63 rays ($R = 63$) and the number of pixels or basis functions is $ij = 25 \times 25 = 625$ for both the FTT and the ART. When the ART iteration exceeds the optimum iteration steps, artifacts may violate the convergence resulting in gradually increasing errors [10]. Thus, ART calculations are ceased when the minimum is reached for the discrepancy between the reconstructed field and the phantom field. The locally continuous cubic B-spline basis function permits interpolated evaluations of field values at any locations between nodal points. This permits the reconstructed results to be plotted with the increased resolutions of 45×45 grid points for all of the results of the ART. The FTT reconstructions are also plotted with 45×45 grid points by the linear interpolation.

The reconstruction calculations were performed with 5 and 3 equally angled projections ($P = 5$ and 3). The FTT reconstructs the Cosine phantom within an acceptable accuracy (Fig. 6(a)) under 5 projections, primarily because of the simplicity of the slightly off-centered single peak. However, the reconstruction result of the double-peak phantom of the CosGauss function (Fig. 7(a)) under 5 projections shows noisy and erroneous data. In addition to the noise arising near the field boundary, four or more misrepresenting peaks are present. Under 3 projections, which may be regarded as the minimally required projections to perceive an asymmetric object no matter how simple the shape, the FTT result shows excessively distorted contours beyond recognition, except for a glimpse of an approximate location of the peak (Fig. 8(a)). The degraded accuracy of the FTT under limited projections is more pronounced for the double-peak phantom of the CosGauss function (Fig. 9(a)). The results are very noisy and give hardly any clue regarding the phantom field. The FTT reconstructions of speckle photographic projections suffer severely from limited projections.

The ART reconstructs the Cosine phantom with a better accuracy (Fig. 6(b)) than the FTT. The ART reconstruction of the CosGauss

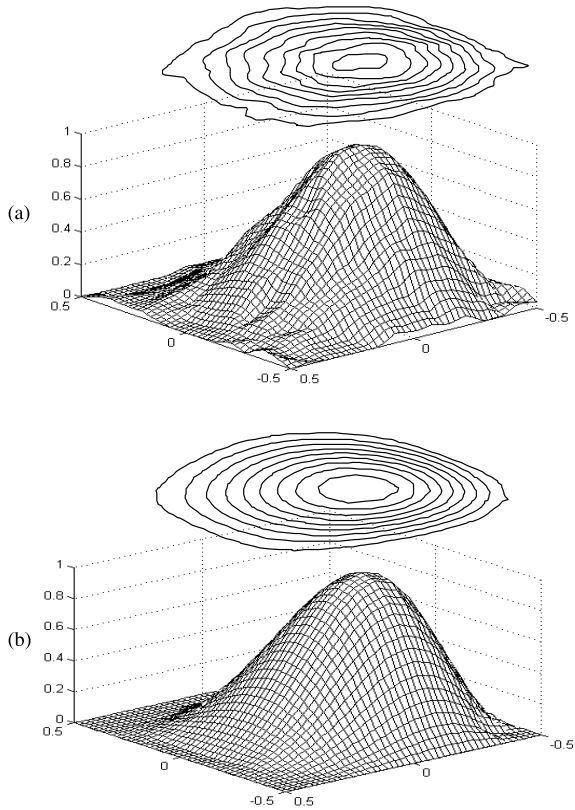


FIG. 6. Reconstructed fields of Cosine function under 5 projections and 45 rays for speckle photography using (a) FTT and (b) ART [13].

phantom (Fig. 7(b)) shows significant improvement over the FTT reconstruction result of Fig. 7(a). The two peaks are clearly pronounced and the peak locations are fairly accurately reconstructed. The peak heights, however, are 15 to 20% underpredicted compared to the phantom heights. It is believed that this discrepancy can be attributed primarily to the numerical errors due to the discrete representation of the phantom field using the basis function, and numerical integration of the projected data, carried out to convert to interferometric data for the ART calculation. Under 3 projections, the ART is still able to reconstruct the Cosine phantom field within an acceptable accuracy (Fig. 8(b)). However, the ART reconstruction of the CosGauss phantom under 3 projections (Fig. 9(b)) is hardly acceptable because of the excessive artifacts of several satellite peaks. Note that some artifacts also appear along the edge

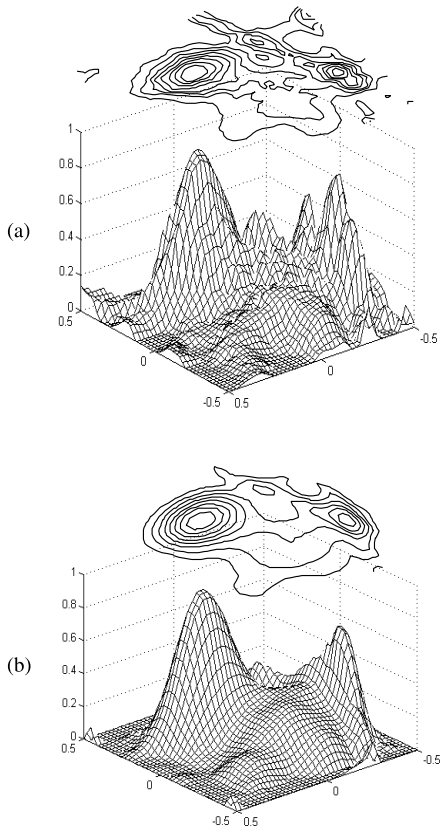


FIG. 7. Reconstructed fields of CosGauss function under 5 projections and 63 rays for speckle photography using (a) FTT and (b) ART [13].

(Figs. 7(b) and 9(b)) because the beam deflection angle data have been numerically integrated to algebraic interferometric fringe numbers to make using the ART method possible. Nevertheless, the ART method shows a noticeably improved reconstruction performance in comparison with the FTT method.

Table I presents a comparison of three errors [Eqs. (20) to (21)] for the reconstruction results. The average, rms, and absolute errors are various measures of the reconstruction accuracy for the reference phantom field. These errors have more significance when comparing tomographic algorithms using a *known* reference field. For all the tested cases, the reconstruction errors of the ART are significantly lower than those of

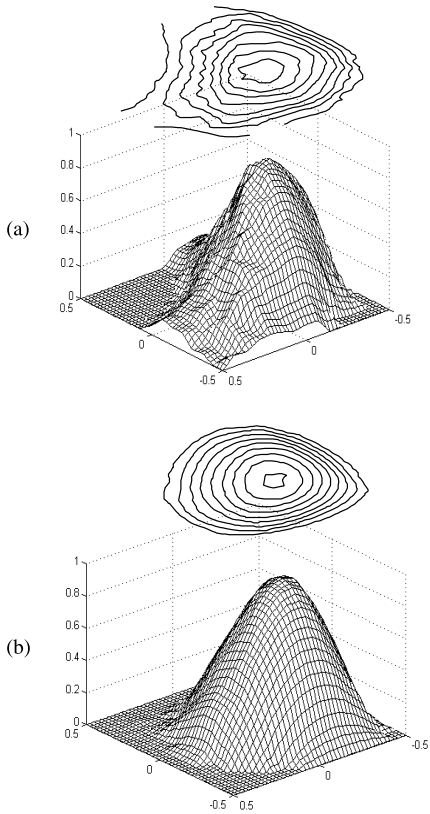


FIG. 8. Reconstructed fields of Cosine function under 3 projections and 45 rays for speckle photography using (a) FTT and (b) ART [13].

the FTT method. For both phantoms for the speckle photography, the differences of the reconstruction errors between the FTT and ART under 5 projections are smaller than those under 3 projections. It means that the ART method has more pronounced improvement over the FTT method under fewer projections.

The asymmetric helium jet from a half-blocked nozzle (Fig. 10) has been tested for the FTT and the ART reconstructions from the speckle photographic and the interferometric projections. The orifice was installed at the end of 12.7 mm ID copper tube. For detailed positioning, the entire jet unit is attached to a rotation stage and it sits on a vertically adjustable stand. Line-of-sight speckle photographic and interferometric images were taken at 4 equally-angled projections with 45 data realizations per

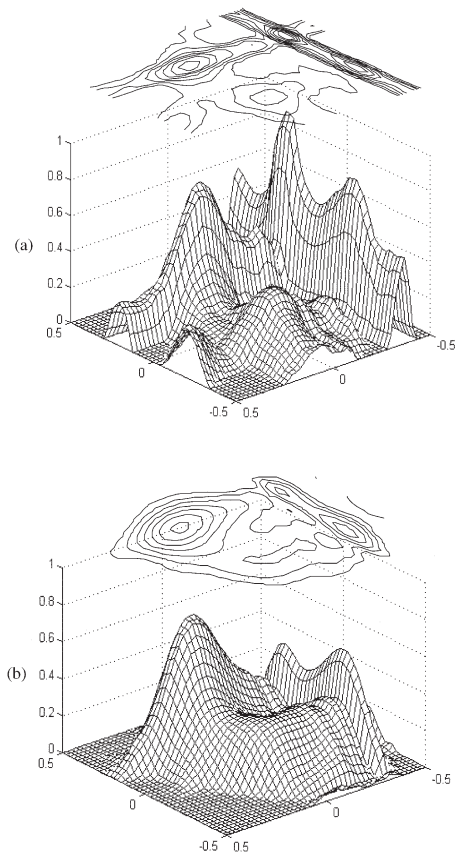


FIG. 9. Reconstructed fields of CosGauss function under 3 projections and 63 rays for speckle photography using (a) FTT and (b) ART [13].

projection for the half-blocked nozzle. 15×15 Cubic B-spline basis functions were used to constitute the cross-sectional object field of the half-blocked nozzle for the ART, and 15×15 discrete pixel values were used for the FTT. The reconstructions for all the present cases were carried out for the cross-section located 9.0 mm above the orifice. A pitot tube measured flow velocity profiles at the tube exit, which converted to the spatially averaged velocity. The helium jet was maintained laminar at $\mathrm{Re} = 990$ for the half-blocked orifice, based on the tube inner diameter $= 12.7$ mm and average velocity at exit $= 15.4$ m/s.

Figure 11 shows a hybrid system for the speckle photography and the interferometry at the Microscale Fluidics and Heat Transport Laboratory of Texas A&M University [http://go.to/microlab]. A Spectra-Physics

TABLE I
COMPARISON OF RECONSTRUCTION ERRORS BETWEEN FTT AND ART METHODS FOR COMPUTER SYNTHESIZED PHANTOMS

	# of Projection	Phantom field	Reconst. tech.	# of Iteration	Φ_{avg}	Φ_{rms}	Φ_{abs}
	5	Cosine	FTT	—	2.85	12.97	12.15
			ART	23	0.93	5.07	3.99
Speckle		CosGauss	FTT	—	5.16	43.34	37.26
Photography			ART	526	2.50	22.75	18.06
	3	Cosine	FTT	—	8.98	40.58	38.33
			ART	3000	2.98	14.70	12.73
		CosGauss	FTT	—	10.84	93.41	78.30
			ART	3000	4.65	38.58	33.59

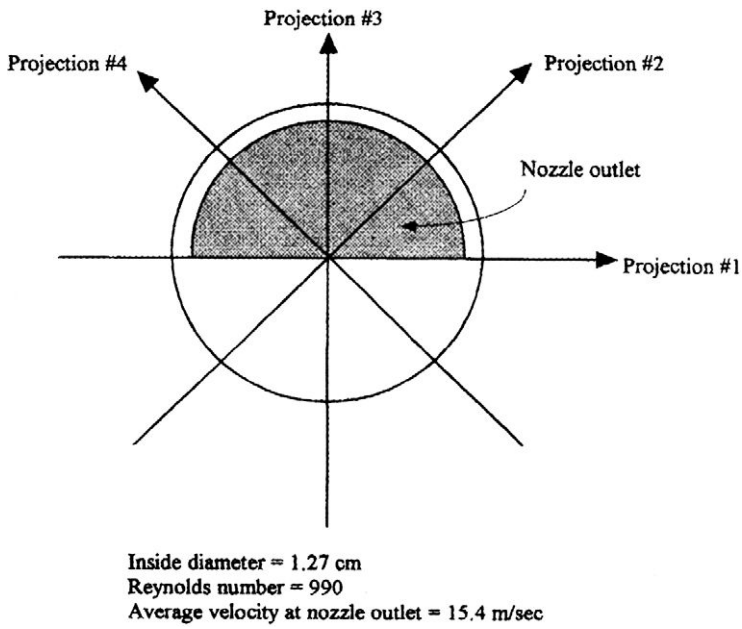


FIG. 10. Schematic of half-blocked nozzle [13].

35 mW He–Ne laser is used as the light source for both of the systems. The flat mirror in front of the laser, as shown in Fig. 11, allows the He–Ne laser to be used for both speckle photography and interferometry light sources. If the flat mirror is placed in front of the laser, then the light will be reflected and directed to the beam expander for the speckle photography.

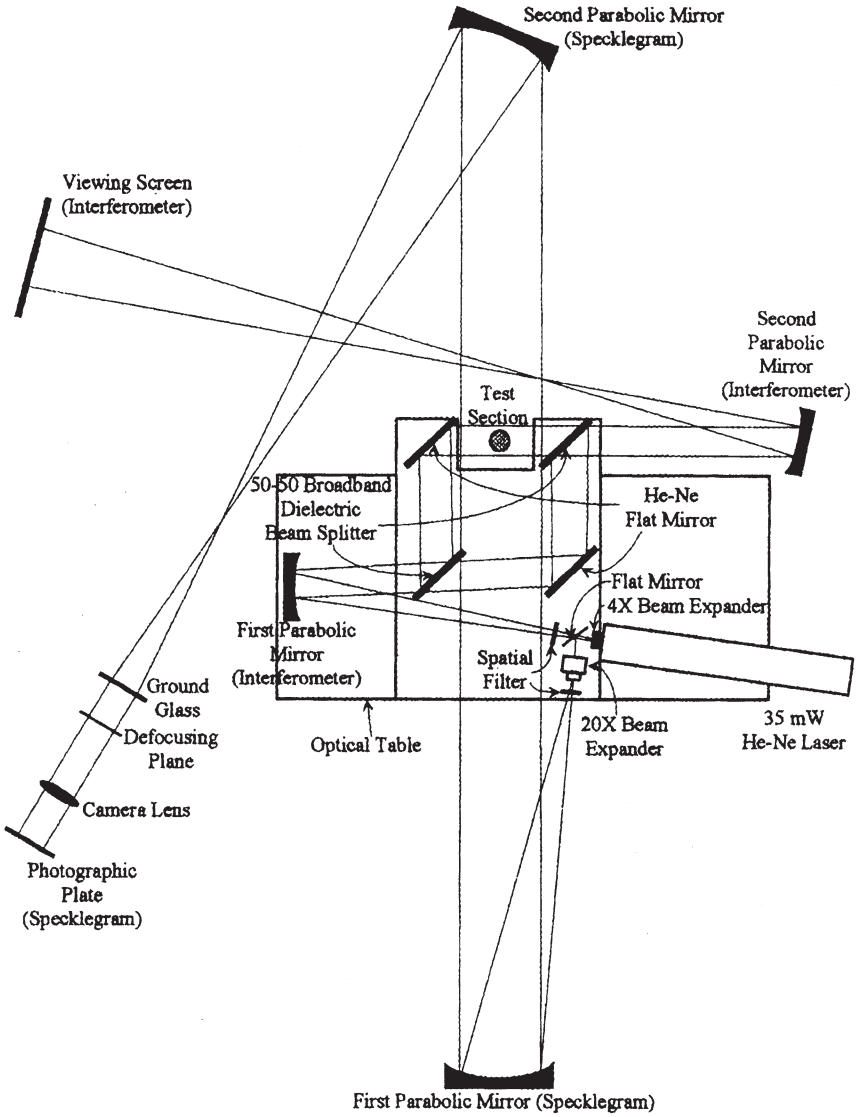


FIG. 11. Hybrid system of speckle photography and interferometry [http://go.to/microlab].

If the flat mirror is removed, the interferometry will be established. The laser and other optics, except for the second parabolic mirror, are mounted on the optical table with a sealed hole table top to minimize disturbances with four high performance pendulum isolates to prevent vibrations from the floor. The two beam splitters and two flat mirrors are mounted

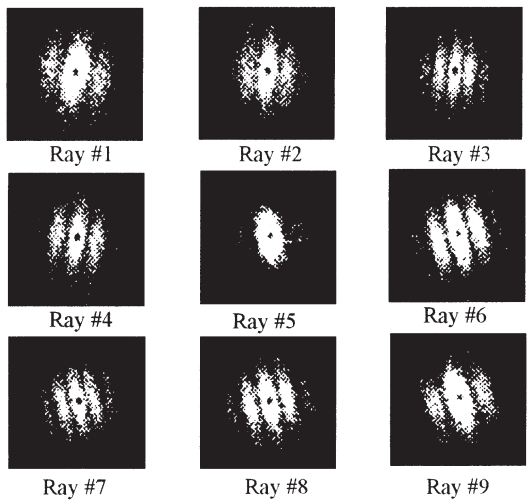


FIG. 12. Young's fringes generated from laser specklegram recorded for helium jet of half-blocked nozzle (projection #3).

on the optical breadboard which measures $3' \times 2'$ and is made of anodized aluminum plates. The breadboard is directly attached to the optical table.

Specklegrams of helium jets from the half-blocked nozzle were taken with the afore mentioned 4×5 format camera at 4 different projection angles using a 4 ms shutter speed. Figure 12 shows the resulting Young's fringe patterns when the specklegram of projection #3 was interrogated in the traverse direction located 9.0 mm above the nozzle. Each ray is separated by 2 mm in the traversing direction. At the center of the nozzle (ray #5), a zero density gradient condition dictated by symmetry provided no speckle dislocation and hence no formation of Young's fringes. Other locations provided measurable density gradients, as an inverse relation exists between fringe spacing and angular refraction. That is, narrowly spaced fringes denote regions of high density gradients, while widely spaced fringes denote lower gradients. Further, as one approaches the ambient condition near the edge of the phenomenon, a zero density gradient again exists and no speckle dislocation is observed. The direction of the Young's fringes has been changed after ray #5, the center of the nozzle for projection #3. It means that the density gradient becomes negative and the mixture density itself increases from the center to the ambient air density.

For each projection, nine Young's fringes were recorded by the video camera and they were digitized into the computer by the video capture board to analyze the fringe spacings. From these nine Young's fringe patterns for each projection, nine distances between Young's fringes have been measured by the computer. Then, the deflection angles for the speckle photographic projection have been calculated to reconstruct the density fields by the FTT and the ART. From 9 ray sums per projection, the interpolation has been performed to increase the discrete projection data (45 data points per projection). These 45 deflection angle data points per projection have been used to reconstruct the density field by the FTT for the speckle photography. For the successive ART the regression technique fitted the discrete projection data (45 per projection) to the ninth order polynomials and the polynomials were analytically integrated to the interferometric data.

Although the interpolation and the integration have been performed for the ART, the results in Fig. 13 show the better reconstruction capability of the ART over FTT for the speckle photography. The ART reconstruction result (Fig. 13(b)) shows more accurate density field contours and lower noise than the FTT result (Fig. 13(a)). The ART result contains virtually no directional dependence and no ambiguities outside the jet region. The use of cubic B-spline basis functions helps to reconstruct smooth and more comprehensive fields. The FTT result, however, shows noisy and immature reconstruction compared with the ART result. The square-pixel construction of the objective field in the FTT produces the discrete results showing bumps in the field. Non-zero values outside the jet are believed to be the mathematical artifacts due to the only 4 projection data sets that are insufficient for the one-time analytical inversion of FTT calculation.

Figure 14 shows the Mach–Zehnder interferograms of helium jets for the half-blocked nozzle. The interferograms were recorded by the video camera and they were transferred to the computer to overlap with the interferograms for no flow case. Then, the fringe shifts were obtained from the superposition of the interferograms for finite fringe mode to reconstruct the density field by the FTT and the ART. The advantage of the interferometry is that one can qualitatively predict the flow profiles from the interferograms.

For each projection of the interferograms, nine fringe shifts have been obtained from the digitally overlapped interferograms. Then, the regression of the ninth order polynomial has been performed for these nine fringe shift values to calculate 45 data points per projection for the interferometric projection of the ART. The 45 projected data of the

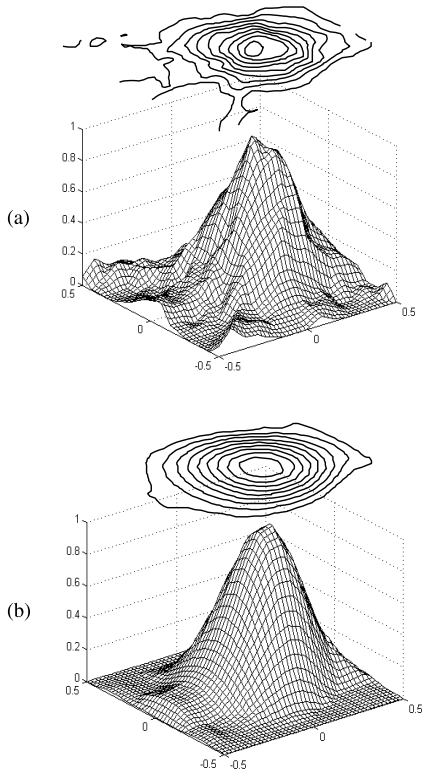


FIG. 13. Reconstructed fields of half-blocked nozzle under 4 projections and 45 rays for speckle photography using (a) FTT and (b) ART [13].

interferometry have been differentiated to directly use the FTT method derived for the speckle photographic projections because the reconstructions of the interferometry can be compared with the results of the speckle photography by the same FTT method. Also, the relations between the interferometric and the speckle photographic projections can be confirmed.

Figure 15 also shows the better result of the ART than that of the FTT. In this case of the interferometry, the ART reconstruction does not require the integration of the interferometric projections. Thus, the noise of the reconstructed field (Fig. 15(b)) is reduced as compared with the result of the speckle photography (Fig. 13(b)). The FTT reconstruction (Fig. 15(a)) also shows a strong directional dependence with enlarged noise along the projection direction because of the numerical differentiation of the interferometric projections and the limited projection angles.

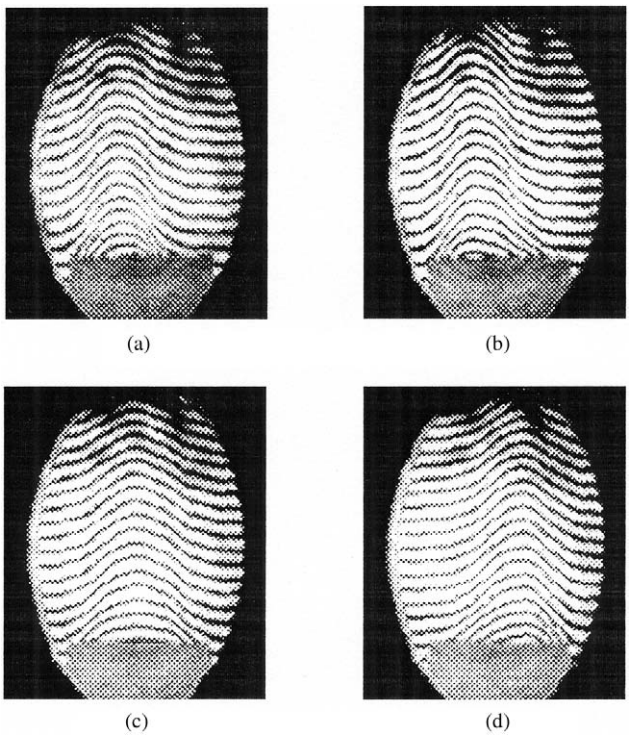


FIG. 14. Mach–Zehnder interferograms of helium jet for half-blocked nozzle. (a) Projection #1, (b) projection #2, (c) projection #3, and (d) projection #4 (the projections are numbered as shown in Fig. 10).

The reconstructed density fields of helium–air mixture were also compared with the direct measurements of helium concentration using an oxygen analyzing probe. The oxygen analyzer measures the density of air/helium mixture directly by detecting the oxygen percentage of sampled gas [30]. The oxygen analyzer uses a small probe fixed on a micropositioning x–y translation stage to sample small amount of flow and to measure the oxygen percentage of the sampled gas (Fig. 16). The principle of the analyzer is based on a Zirconia sensor that measures the quantity of oxygen ions at an elevated temperature. The oxygen percentage is readily converted to the gas density ρ as follows:

$$\rho = H * \rho_{\text{he}} + (1 - H) * \rho_{\text{air}} \quad \text{with} \quad H = 1 - R_m/R_o \tag{23}$$

where H is the helium mass fraction, R_m is the measured oxygen percentage, R_o is the reference oxygen percentage (20.6% at the laboratory),

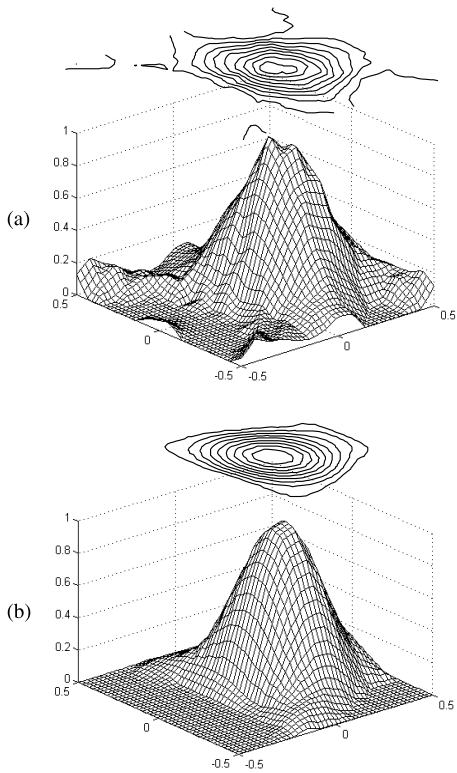


FIG. 15. Reconstructed fields of half-blocked nozzle under 4 projections and 45 rays for interferometry using (a) FTT and (b) ART [13].

ρ_{he} is the pure helium density, and ρ_{air} is the pure air density. The measured gas density is readily converted to normalized helium concentration as follows:

$$\rho^* = \frac{\rho - \rho_{air}}{\rho_{he} - \rho_{air}} \tag{24}$$

where ρ^* ranges from 0 for pure air to 1.0 for pure helium. This normalized helium concentration, obtained from the oxygen analysis, is used for all the comparison.

Figure 17 shows the reconstructed results along the symmetric lines, projection #3 shown in Fig. 10, in comparison with the measured helium concentration. The solid circles represent the average values of fluctuating readings from the oxygen sensor unit and the fluctuating ranges are

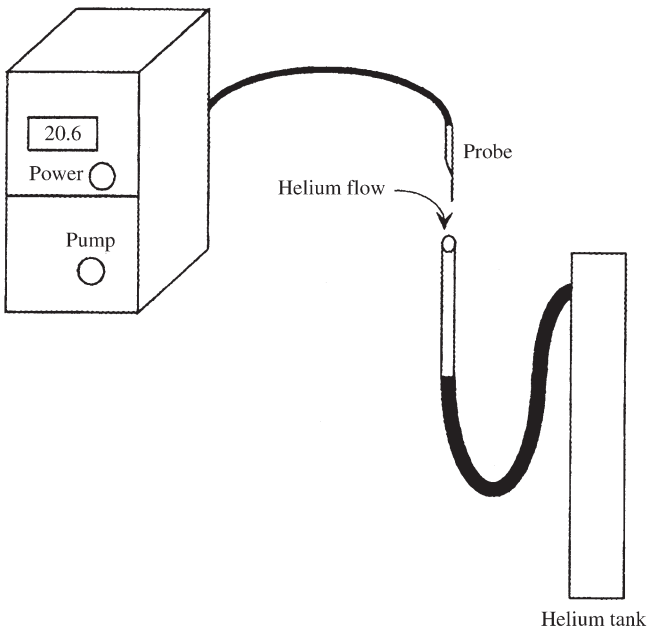


FIG. 16. Schematic of oxygen analyzer.

marked with the uncertainty bars for the peak reading case. Drastic reduction was not possible in the relatively large fluctuations in the oxygen sensor reading, instead an average value of the fluctuating range was taken once the fluctuations remained within a fixed range. The ART result (solid curve) shows more accurate identification of the peak and agree more quantitatively with the oxygen analyzer results as shown in Fig. 17(a). The FTT results (dashed curve) do not accurately predict the peak location and slope. The ART reconstructions of the speckle photography and the interferometry under 4 projection angles are also compared with the results of the oxygen analyzer (Fig. 17(b)). From this graph, the ART reconstruction of the interferometry and the oxygen analyzer results agree well in the blocked region (from -0.5 to -0.1 of the position). The ART reconstruction of the speckle photography includes small noise near the edge because of the integration of the projected data. However, for the location of the peak, the ART reconstruction of the speckle photography shows better results than that of the interferometry as shown by Fig. 17(b). In most of the regions, the ART reconstructions of the speckle photography and the interferometry are well identified with the oxygen analyzer results.

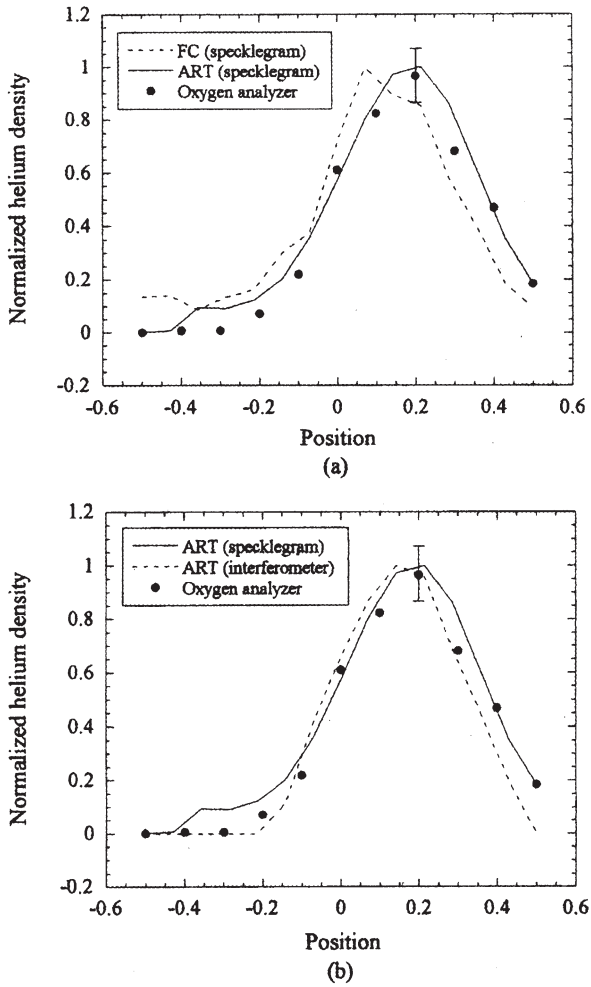


FIG. 17. Comparison of helium density distribution with oxygen analyzer along centerline of half-blocked nozzle. (a) FTT (FC) and ART for speckle photography and (b) ART for speckle photography and interferometry [13].

IV. Genetic Algorithm (GA)-Based Tomography

Particularly under limited number and angle of projections, a new idea of reconstruction technique is proposed to utilize a robust function optimization, namely a Genetic Algorithm (GA). Section IV.A explains the essential feature of the GA function optimization. Section IV.B presents an

illustrative example for the GA function optimization. Section IV.C discusses the relevance of GA solution search under a given constraint, to tomographic image reconstruction from measured line-of-sight optical projections. Section IV.D presents an example for the GA-based optical tomography.

A. What is GA?

The Genetic Algorithm (GA) is a robust, intelligent, and combinatorial function optimization algorithm for finding unknowns from a set of knowns or constraints based on the mechanics of the genetic principle [17]. Given a set of unknowns $(a, b, c, \ldots)$ to be searched under a specified constraint $f(a, b, c, \ldots) = A$, the solution search is considered to use the mechanics of natural selection and natural genetics, adopted by genetic algorithm, among a large pool of randomly generated solution candidates for the unknowns. Although the goal is to find a single optimized set of unknowns (*the fittest individual*), the GA initiates and proceeds with multiple sets of solution candidates (*the population pool*). This feature of handling multiple solution candidates is called *implicit parallelism*. The key concept of the GA solution search is based on the fittest survival and their dominating evolution, for either the survival of living creatures in nature or the solution search from candidates under a given constraint. The assigned values of $(a, b, c, \ldots)$ for each solution candidate are analogous to the set of *genes* of an individual.

As an illustrated example (shown in Table II), consider N-individuals $(I_i,\ i = 1, 2, 3, \ldots, N)$ of randomly generated solution candidates with A_i being the constraint imposed for the ith individual. The fitness *ranking* of individuals is established by their fitness values defined as $|A_i - A|$, with the lowest value referring to the highest ranking and fitness. The current *generation* is now ready to evolve to the next generation.

TABLE II
Construction and Evaluation of a Population

Population	Genes	Constraint	Fitness	Ranking
I_1	$a_1, b_1, c_1, \ldots$	A_1	$\|A_1 - \mathrm{A}\|$	
I_2	$a_2, b_2, c_2, \ldots$	A_2	$\|A_2 - \mathrm{A}\|$	
I_3	$a_3, b_3, c_3, \ldots$	A_3	$\|A_3 - \mathrm{A}\|$	
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
I_N	$a_N, b_N, c_N, \ldots$	A_N	$\|A_N - \mathrm{A}\|$	

According to the ranking, each individual is assigned a *mating probability*. The selection of a mating couple, from the mating pool, is based on a scaled random or biased random procedure, i.e., an individual with a higher mating probability is assured to have a higher chance of being selected while an individual with a lower mating probability is occasionally and randomly selected for mating. The idea is to save high-quality genes hidden in lower-ranked individuals. The two individuals selected from the mating pool randomly swap their genes or *crossover* for two new children. Then only two individuals with the least fitness values survive the four of the two parents and two children. This is called *elitism* to ensure the fitness is always improved or at least stays at the present level.

Each gene carries a specified *mutation* probability and genes are subject to mutate randomly to create a mutant individual. Only the individual with the least fitness value survives among the original and the mutant. The primary role of the mutation of genes is to prevent individuals from settling at a local peak, and to provide a robust feature in searching for the ultimate solution peak. Any twin or triplet individuals, if ever occurring, will count as one to avoid the potential of "*dominance by common*". A new generation evolves from the set of procedures of these mating selections, crossover, mutation, and generation evolution is repeated until the fitness value of the fittest individual enters a specified tolerance from the solution constraint. The genes carried by this optimized individual determine the solution.

B. Illustrative Example of Function Optimization Using GA

For the purpose of discussion, a simple density field is considered as a reference field (Fig. 18), whose functional representation is given as [18]:

$$\rho(r) = K_a e^{-K_b r^2} + K_c r e^{-K_d r^2} \quad (K_a = 0.5,\ K_b = 10,\ K_c = 5, \text{ and } K_d = 10) \tag{25}$$

The reference field is assumed to be conformed by a summation of basis functions f:

$$\rho^*(r) = \sum_{i=1}^{M} f(r - r_i;\ a_i,\ t_i) \tag{26}$$

where f represents a two-parameter basis function, such as in a Gaussian function, located at $r = r_i$, with its height a_i and shape factor t_i, and M denotes the number of basis functions to be used. The unknowns a_i and t_i are analogous to *genes*. The superscript * refers to the solution candidate

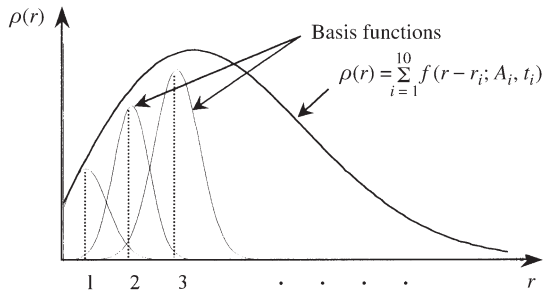


FIG. 18. A reference field conformed by the summation of basis functions [18].

field and the solution constraint is thereby given as $A = |\rho(r) - \rho^*(r)|$, which needs to be minimized.

The use of continuous basis functions enhances the spatial resolution of the solution without significantly increasing the computational expenses. In principle, basis function can be any type of continuous function such as Gaussian, Rayleigh, sine or cosine, or log-normal. Proper discretion in selecting a basis function type to best represent the tested field may expedite the reconstruction optimization. For example, an air jet diffused into a helium environment may select a mass diffusion solution for basis function, since the diffusion equation describes the physical nature of the jet [19]. On the other hand, the basis function representation requires optimizing multiple sets of parameters simultaneously while the discrete field representation requires optimizing just a single set of parameters for nodal values.

The evolution process of the GA can effectively optimize the finding of multiple sets of unknowns, as schematically illustrated in Fig. 19. The GA-based optimization starts with a large group (*initial population pool*) of solution candidates (*individuals*). Each individual carries randomly generated magnification and shape factors, respectively, a_{ij} and t_{ij} for the ith basis function of the jth solution candidate. The solution evolves by an internal manipulation of the genes within the population group under specified rules imposed by the genetic principles.

The GA optimization procedure, as described in Section IV.A, is extended to search for the best possible solutions for a_{ij} and t_{ij}. The Selection operator selects individuals based on a probability associated with their fitness. This ensures the "environmental" pressure forcing survival of the fittest. Once a couple is selected, they have a pre-specified chance of creating two children through randomly exchanging mutual genes by the crossover operator. During each gene transfer, there is a pre-specified chance of a gene being altered by the mutation operator. The mutation

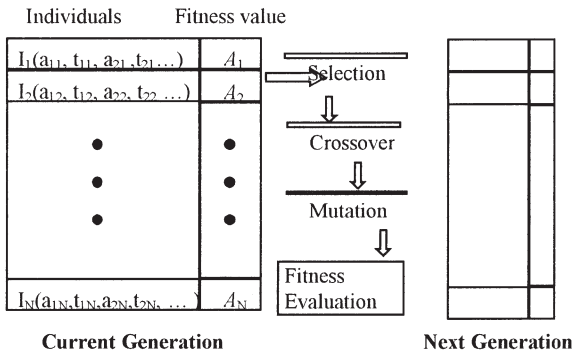


FIG. 19. Schematic evolution of the GA function optimization.

operator is responsible for exploring new points in the search space and is the primary source of robustness. These operators operate on each population and the next generation evolves the same size as the current generation. The new generation is then operated on in a similar manner to evolve further generations. This process ceases when the best individual of the present generation conforms to the actual field to a desired degree of accuracy, i.e., $A_{\text{the fittest}} < \varepsilon$ or a specified number of generations is reached.

The reference density field given by Eq. (25) is assumed to be represented by the superposition of a total of 10 ($M = 10$) Gaussian basis functions, i.e.,

$$\rho^*(r) = \sum_{i=1}^{10} a_i \exp\left[-(r - r_i)^2 / t_i\right] \tag{27}$$

The initial population pool adopts 100 randomly generated solution candidates of a_{ij} and t_{ij} with $i = 1, 2, \ldots, 10$ and $j = 1, 2, \ldots, 100$. Strictly speaking, they are not entirely randomly generated, as the random variations for a_{ij} and t_{ij} are set to be bounded to $(-1.0, 1.0)$ and $(0, 10)$, respectively, judging from the magnitude of the reference function. The ranges of selection of unknown parameters can be pre-determined by using available information on them and this effort to expedite the optimization is called an *intellectual random*.

The computer program for the GA-tomographic reconstruction calculations has been coded with the C and C^{++} languages. The available code library GALIB [http://lancet.mit.edu/ga/], also written in C^{++}, has been downloaded and modified to use for our calculations for GA optimization. The Concurrent Simplex optimization needed for the hybrid GA optimization (Section IV.C) uses a C^{++} language program.

Figure 20 shows the evolution of the reconstructed field given by the best individual for different generations. The best individual of the

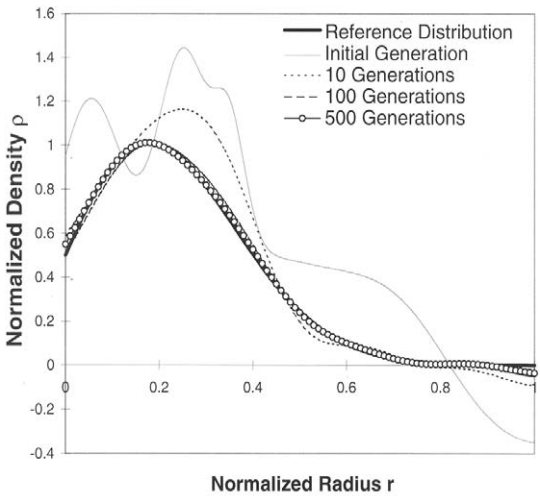


FIG. 20. Evolution of the reconstructed field by the best individuals with progressing generations [18].

randomly selected initial population of 100 shows large deviations from the reference field as expected. However, the discrepancies have quickly decreased as the generations have evolved and the reconstructed density field from the 100th generation has already approached the reference field in very close proximity.

Figure 21 shows the history of convergence along the evolving generation. The convergence of both the average and the best individual has reached the asymptotic level near the 400th generation. Although the average fitness fluctuates temporarily, the fitness of the best individual at least maintains its current level or improves with generations. This is due to the afore mentioned elitism imposed in the GA in that the current best maintains the title until a better individual with a smaller fitness value captures it.

Figure 22 shows the 10 basis functions that constitute the best individual at the 1000th generation. Whereas it might be expected that the individual basis functions would have particular characteristics in certain areas of the density field, i.e., relatively narrow and tall near the left end and short and broad toward the right end, it must be remembered that it is the sum of the basis functions that determines the overall density field. Also, the negative basis functions compensate for the excessive contributions from the positive basis functions, which facilitates a more effective optimization process.

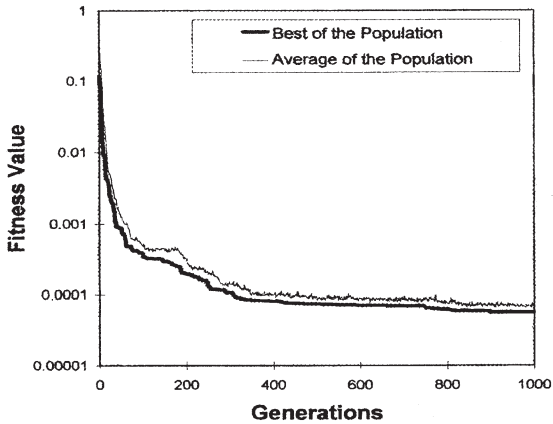


FIG. 21. History of fitness value with evolving generations [18].

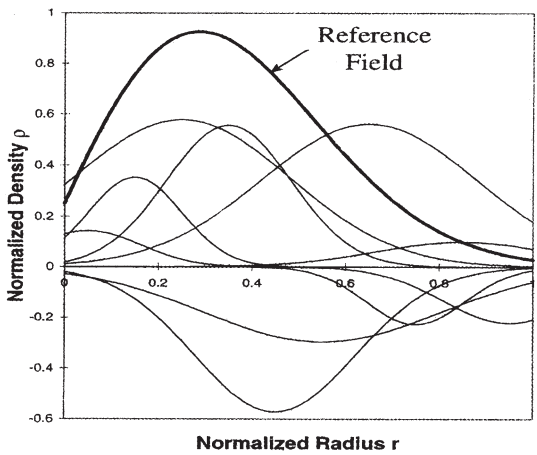


FIG. 22. Optimized basis functions conforming to the reconstructed field at the 1000th generation [18].

C. RELEVANCE OF GA TO OPTICAL TOMOGRAPHY

Given a density field $\rho(x, y)$, a line-of-sight projection, such as in interferometry, is constructed by an integration of the density field along each ray (Fig. 23). Thus, the measured projection image is given as:

$$\psi_i = \psi(t_i; \theta_i) = \text{Const} \int_{s_i} \rho(x, y) ds_i \tag{28}$$

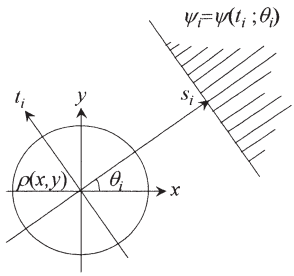


FIG. 23. Line-of-sight projection of a density field.

where s_i and t_i denote the coordinates parallel and perpendicular to the ray, respectively, θ_i denotes the projection angle of the ith projection, and the Const is a function of incident laser wave length. Thus, Eq. (28) is considered as the constraint for the determination of the solution $\rho(x, y)$. Tomography comprises a numerical programming to reconstruct an unknown field $\rho(x, y)$ from its measured projections ψ_i for $i = 1, 2, 3, \ldots, N_p$, with N_p denoting the number of projections.

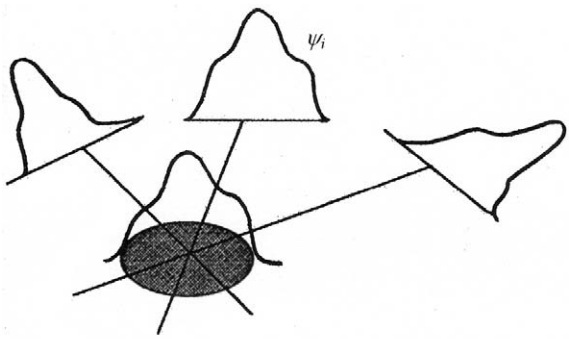
Substituting the basis function representation, Eq. (26), into Eq. (28) leads to a system of non-linear equations

$$\Psi = W(O) \tag{29}$$

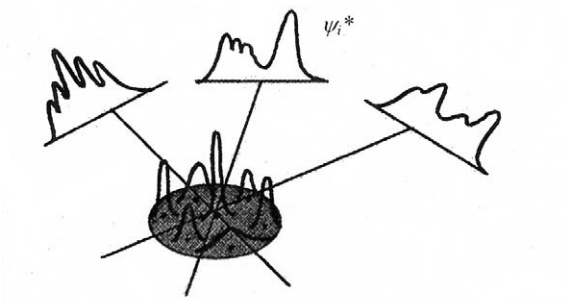
where $\Psi = (\psi_1, \psi_2, \psi_3, \ldots)$ is known as the measurement vector whose components represent the projected values, W is the projection function specified depending on the type of projection method (interferometry, speckle photography [20], X-rays etc.), and $O = (a_1, t_1, a_2, t_2, a_3, t_3, \ldots)$ is the object (solution) vector, whose components represent heights and spreads of all the basis functions. Note that the selection of a two-parameter basis function makes the inversion of Eq. (29) a multi-parallel process that requires a simultaneous optimization of the two sets of parameters. The tomographic image reconstruction in the current context is now reduced to simultaneously finding optimized amplitudes and shape factors for all basis functions.

The deviation of each individual's projection from the true projection of the actual field is measured by a fitness value defined as:

$$A_j = \sum_{i=1}^{N_p} \int_{t_i} (\psi(t_i; \theta_i) - \psi^*(t_i; \theta_i))^2 dt_i \tag{30}$$



(a) Measured projections from the actual field



(b) Virtual projections from the combined EDFs

FIG. 24. Measured projections and virtual projections [19].

where $\psi(t_i;\theta_i)$ represents the measured projection of the true field and $\psi^*(t_i;\theta_i)$ represents a virtual projection of an individual solution candidate.

Figures 24(a) and (b) schematically illustrate the measured and virtual projections. The Elementary Density Function (EDF) in Fig. 24(b) is equivalent to a basis function. The GA-based evolution optimizes the parameters of the basis functions under the constraint in that the virtual projections approach the measured projections. The tomographic image reconstruction is equivalent to a function optimization for a_i and t_i under the constraint of minimum of Eq. (30).

D. Illustrative Examples with Interferometric Projections of a Density Field

Fermat's principle [21] states that the wavelength of light becomes shorter in a medium with refractive index n compared with the wave length

in vacuum ($n_o = 1.0$). When light rays travel through the test section with refractive index distribution $n(x, y)$ in the projection direction s (Fig. 20), the number of waves increases within the test section compared with that for a vacuum, and the interferometric projection is expressed as:

$$\psi_i = \psi(t_i; \theta_i) = \frac{1}{\lambda_o} \int_{s_i} [n(x, y) - n_o] ds_i \tag{31}$$

For finite-fringe interferometry, ψ_i is identical to the number of crossing fringes recorded on the ith projection. The refractive index can be converted to the medium density using the Gladstone–Dale relationship [22]:

$$n(x, y) = 1 + K \cdot \rho(x, y) \tag{32}$$

where the Gladstone–Dale constant K is a function of wavelength. Combining Eqs. (31) and (32), the interferometric projection function is derived.

$$\psi(t_i; \theta_i) = \frac{K}{\lambda_o} \int_{s_1}^{s_2} [\rho(x, y) - \rho_o] ds_i \tag{33}$$

Similarly, when $s_1 = s_1(x_1, y_1)$ and $s_2 = s_2(x_2, y_2)$,

$$\psi(t_i; \theta_i) = \frac{K}{\lambda_o} \int_{y_1}^{y_2} \int_{x_1}^{x_2} [\rho(x, y) - \rho_o] \delta(x \cos \theta_i + y \sin \theta_i - t_i) dx dy \tag{34}$$

where the density distribution $[\rho(x, y) - \rho_o]$ can also be expressed in terms of basis functions as in Eq. (26). $\delta(\cdot)$ represents a Dirac-delta function that is equal to 1 only on a line of $x \cos \theta_i + y \sin \theta_i - t_i = 0$, and 0 otherwise.

The viability of the GA-based tomographic scheme has been examined using an asymmetric single peak reference field (Fig. 25(a)). This reference field is synthesized by computer as a linear combination of diffusive concentration profiles that are defined by two parameters specifying its height and spread [18]. The peak location is skewed from the center of the test field and the peak density extends above 1.4 in an arbitrary unit. The projection data of the fringe shift numbers were calculated from the reference phantom field using the interferometric projection function shown in Eq. (12), at six equal angular intervals, $\theta = 0$, 30, 60, 90, 120, and 150°. Figure 11-b shows the reconstruction performed using the GA-based tomography. Under the identical projection conditions, a comparative study has been made with reconstruction images by ART (Fig. 25(c)) and FTT (Fig. 25(d)).

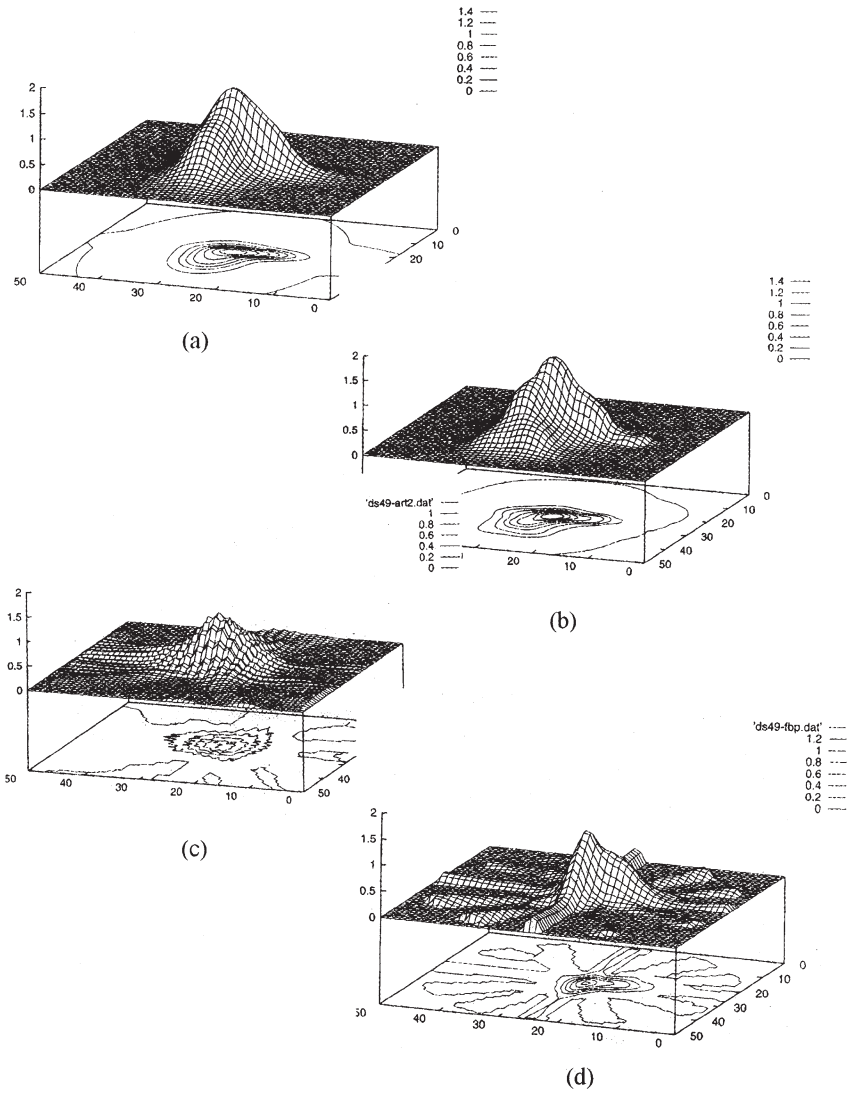


FIG. 25. Computer synthesized reference phantom (a) and reconstructed images by Genetic Algorithm (GA) (b). Algebraic Reconstruction Technique (ART) (c) and Fourier transformation technique (FTT) (d) [19].

The number of GA generations was expanded up to 1000. The basis function is selected as the same type used to synthesize the reference field, but with arbitrarily different height and spread parameters. The computation was carried out on a Pentium 90 MHz PC machine

in 1995. The GA conversion used a population of 100 with 20 children in each generation. The number of computational nodal points was 19 with 19 basis functions for the GA and was 10,000 (100×100) for the Fourier transformation technique (FTT) and algebraic reconstruction technique (ART). The number of ART iterations was set to 1000. FTT, which is a direct mathematical inversion of projected image [1,2], does not require iterations. Note that the ART update is possible only for a single set of unknowns and such simplicity makes the ART iteration extremely fast.

The GA-based conversion (Fig. 25(b)) successfully identifies the locations and shapes of the ridges, accurately predicts the peak density height of approximately 1.4, and shows close proximity in the density contour map with its counterpart of the reference field. While all three results identify the existence and approximate location of the peak, ART (Fig. 25(c)) and FTT (Fig. 25(d)) results show underestimation of the peak density as noticed on the topographic contour maps. In addition, the discrete nature of the ART iteration was unable to show the detailed shape of the peak. The ART will need an increased number of nodal points for a smoother-looking conversion. The insufficient number of six projections for the FTT conversions makes the directional noise grow to an unacceptable level and the number of projections must be substantially increased to enhance the accuracy. These defects of the ART and FC conversions are also apparent in the contour maps.

The Pentium machine required approximately 2400 s for the entire GA evolutions up to 1000. The CPU time needed for the GA is noticeably longer than that of the FC calculation (7 s) and of the ART conversion (25 s). However, the accuracy of conversion must not be sacrificed for the sake of CPU time saving. Although not shown, the ART conversions with 10,000 iteration steps (approximately 250 s CPU time) and 50,000 steps (1250 s) did not show any significant improvement and the results were still not as accurate as the GA result. The FC calculations that were repeated with 1000×1000 nodal points using the same six projections (about 350 s CPU time), did not show any major improvement either. The noticeably longer time required for GA is because of its massive and parallel computation for multiple and simultaneous optimization processes. A substantial effort in an attempt to reduce the computational time by introducing a hybridized GA idea will be discussed in Section IV.C.

The band of symbols in Fig. 26 shows the evolving history of the normalized fitness values [Eq. (5)] for the population of 100 individuals. Fitness values quickly decrease during the beginning 100 generations or so, and gradually approaches the minimal level. The best (or smallest) fitness

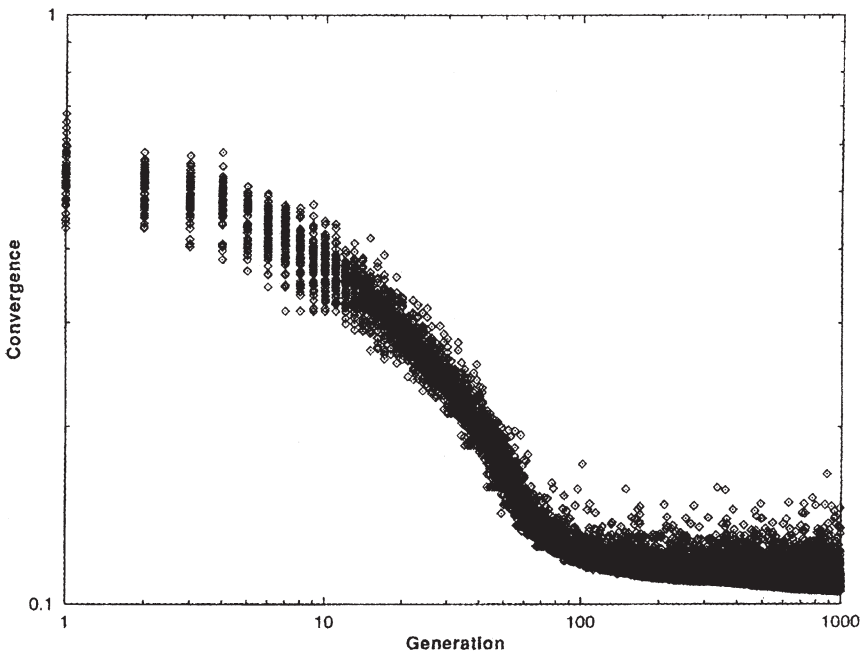


FIG. 26. Convergence values of the whole population of 100 individuals versus generation evolution for the GA-based reconstruction shown in Fig. 22(b) [19].

value decreases consistently with evolving generations, while the range of fitness values in each generation scatters. This means that the worst or largest fitness value does not necessarily improve with generations while the top individual consistently improves. This explains the "elitism" imposed in the GA optimization, which focuses on improving the elite members rather than the "also-runs", since it is the best individual that leads the optimization. Along the same strategy, the best individual keeps its title until a new challenger wins the title.

Figure 27(a) shows a phantom composed of 90 identical mass diffusion solutions and one taller solution of the same type. The GA-based tomography uses 91 basis functions of the same type but with randomly assigned heights and spreads. The reconstruction, the optimization of the heights and spreads, was carried out from six interferometric projections at six different angles with equal intervals. Each generation held a population of 40 with 20 children and up to 20 allowable mutations. The calculations were carried out on the CRAY-YMP super computer platform in 1996. Figure 27(b) shows the reconstructed field by the top individual of the

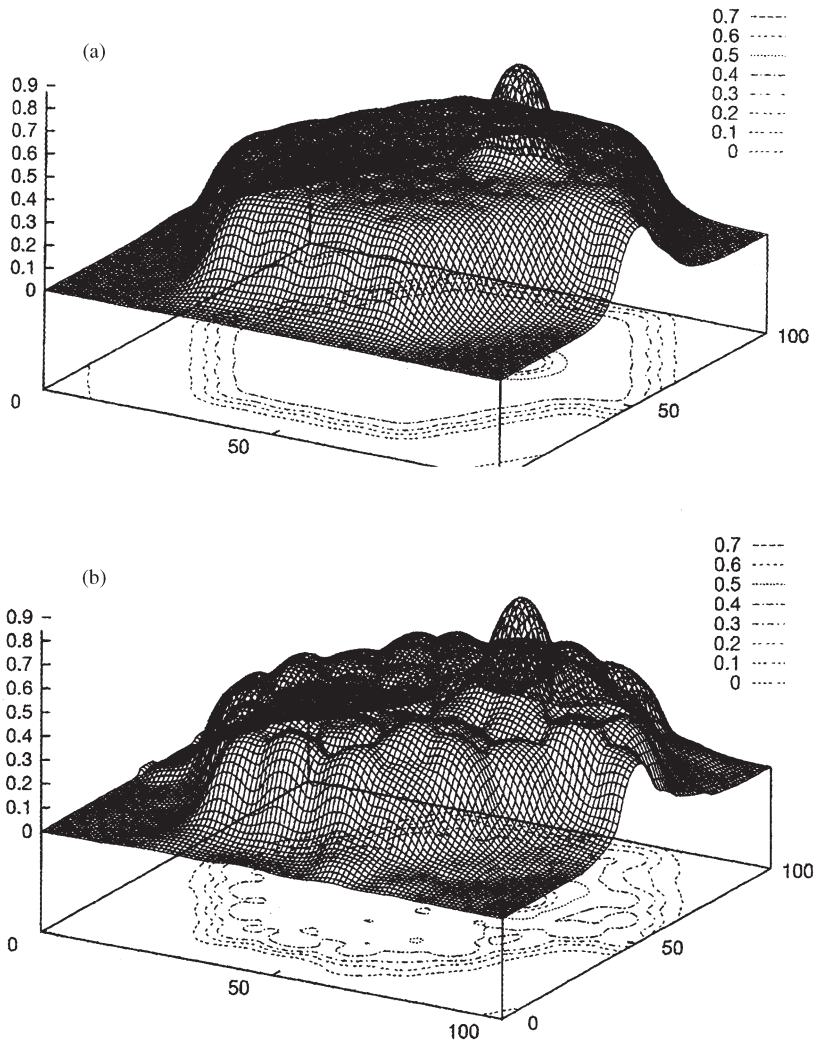


FIG. 27. Computer synthesized reference field (a) and reconstructed field (b) after 1000 generations of the GA-based tomography [19].

1000th generation. The GA-based tomography reconstructs reasonably well the shape and location of the peak and the density contours.

Figure 28(a) shows a computer-generated distribution of sharp twin peaks to accommodate a more critical comparison of the GA method with the conventional ART method [23]. The insufficient number of projections

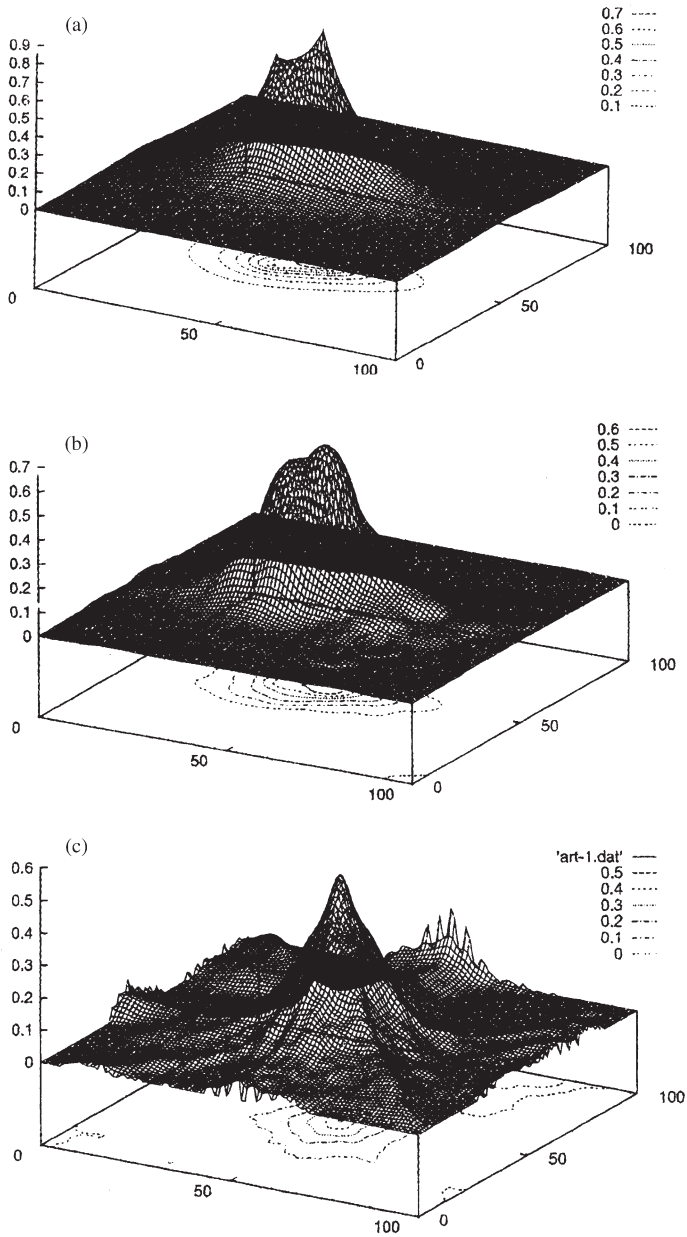


FIG. 28. Computer synthesized twin-peak (a), reconstructed image after 1000 generations of the GA-based tomography (b), and 1000 iterations of the ART with six equally-angled interferometric projections (c) [19].

($N_p = 6$) did not allow the FC method to construct any sensible conversion results although not presented. The GA method uses 91 basis functions of the axisymmetric mass diffusion solution distributed evenly in the rectangular domain and each generation held a population of 40 with 20 children and up to 20 mutations allowed. The ART calculations used 100×100 computational nodal points on the CRAY-YMP super computer in 1995. The required computational times including the vectorization procedure were 2 h for 1000 generations of the GA and approximately 4 min for 1000 iterations of the ART.

The GA result (Fig. 28(b)) shows a fairly acceptable distribution identifying the accurate location of the twin peaks although the GA method is unable to reconstruct the exact shape of the sharp peaks. This deficiency is attributed to the fact that the selected basis functions of bell-shaped diffusion solution are not able to conform to the sharp peaks. Nevertheless, the contour map of the GA results shows a fairly good reconstruction, apparently evidencing the twin peaks. The ART result (Fig. 28(c)), however, failed to show the existence of the twin peaks and merely identified the approximate area with a merged single peak. Also, the maximum density has been underestimated from over 0.7 (in an arbitrary unit) of the original distribution to about 0.5 of the ART reconstructed data. The contour map of the ART results is inaccurate with faulty contours and noises.

The point to note is that the present comparison has been made only with the very basic ART and classical FTT methods just to demonstrate the idea of a new tomographic method using the genetic algorithm. More rigorous and extensive comparisons with many up-to-date ART and FTT methods will be essential to carry out for future study of the GA-based tomography.

V. Hybridized GA with Concurrent Simplex Optimization

Despite its robustness and more accurate reconstruction potential, the GA-based tomography requires an excessive computational time because of its massive combinatorial handling of multiple solution candidates. In an effort to accelerate the calculation procedure without sacrificing the advantages of the GA-based tomography, a hybridization is proposed for GA evolution with a downhill Simplex optimization method [24]. Section V.A explains the basic idea of the Simplex optimization, and Section V.B introduces the hybridized GA optimization with the Concurrent Simplex (CS). Section V.C presents an illustrative example of hybridized GA function optimization for the same reference field that was

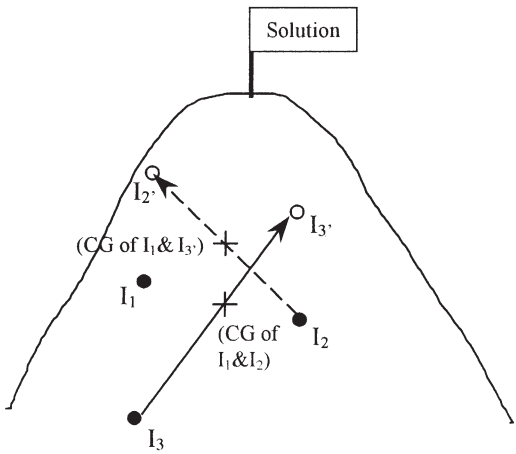


FIG. 29. Schematic illustration of Simplex optimization [26].

used in Section IV.B. Finally, Section V.D presents the performance of the hybridized GA-based tomography applied to reconstruct discrete fields such as bubbly two-phase flows.

A. What is a Simplex Optimization?

Figure 29 shows a schematic of the basic Downhill Simplex where a population of three individuals is optimized to approach the peak. Individual I_3 is the worst in terms of fitness in the first generation and I_3 is then reflected with respect to the center of gravity (CG) of better individuals, I_1 and I_2. In the next generation consisting of I_1, I_2 and $I_{3'}$, the newly created $I_{3'}$ becomes the best and I_2 becomes the worst. Now, I_2 is reflected with respect to the CG of I_1 and $I_{3'}$ to achieve the better individual $I_{2'}$ for the following generation. The reflection procedure is repeated until the best individual approaches within a specified range from the peak or the number of generations reaches a specified number.

The variant of the downhill simplex used herein is called the Concurrent Simplex [25], where multiple worst individuals, instead of a single worst individual, are selected and they are reflected with respect to the CG or centroid of all the remaining (better) individuals.

B. Hybridized GA Optimization

In the hybrid GA (Fig. 30), each new generation is created partially by the preservation of the top ranking elite group (N_E), partially by the

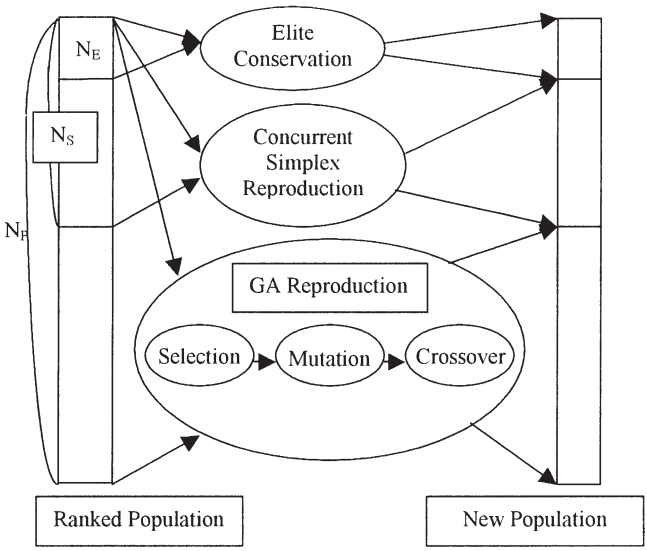


FIG. 30. Evolution protocol of the elite-based hybrid Simplex-GA [26].

Concurrent Simplex (CS) optimization ($N_S - N_E$), and partially evolved by the GA operators ($N_P - N_S$). The preservation of the elite group ensures that the prestigious individuals survive the generation without being deteriorated by the lower ranked individuals, which would be the case for the substitution of the GA optimization for the entire population. Note that the elite group is also subjected to the concurrent simplex and to the GA optimization with the remaining population. This multiple and parallel optimization protocol is intended to evolve the best possible individual from the given gene pool. The CS exploits the continuous real-valued nature of the search space providing, hopefully, a good direction toward the optimum solution.

C. Illustrative Example of Hybridized GA Optimization

Calculations of the hybridized GA optimization has been made for the reference field given in Eq. (1) and Fig. 1 using 10 basis functions of the Gaussian type [26]. In the present calculations, the specified probabilities of crossover and mutation were given as 50 and 1%, respectively. A series of computations were carried out to examine five different hybridization ratios (N_S/N_P) of 0% (pure GA), 30, 50, 70, and 100% (pure CS with no GA). The ratio of the elite group, N_E/N_P, is set to 1/10 of the hybridization ratio. A single initial population group was used for all the

calculations to provide that the runs for all degrees of hybridization began with the same initial condition. For each percentile hybridization, 10 individual calculations using different seed number series were performed to average the different random number sequence of each calculation. The 100% CS method has no random processes involved and only one run for the given initial population was necessary.

Convergence history of fitness values is shown in Fig. 31. As expected, the 100% CS does not converge well, exhibiting locking after less than 100 generations. This locking presumably exhibits a local peak trap. Any level of hybridization accelerates convergence over the pure GA case, with 70% hybridization being the best for the present case. There are several parameters that affect the performance of the hybrid GA optimization. These parameters are indeed stochastic and randomly assigned during the iterative process. The performance also depends on the nature of the reconstructing field. At present, it does not seem straightforward to conjecture precisely the reason why 70% is better than the other cases.

More importantly, the results by the hybrid GA reach asymptotic value around the 150th generation, while the conventional GA does not reach close to an asymptotic value until around the 400th generation, in addition to its poor convergence. This shows a noticeable reduction in the computational time providing a better convergence.

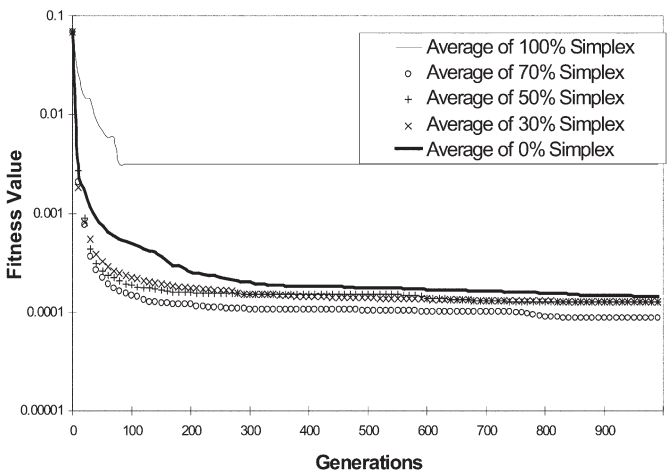


FIG. 31. Convergence of fitness values for different levels of hybridization with the Concurrent Simplex [26].

D. Hybridized GA-Tomography for Discrete Fields

Two-phase flows, such as bubbly liquid flows or solid-laden liquid flows, are popular in many industrial processes, and the non-intrusive determination of the shapes, number, locations, and sizes of the bubbles is increasingly demanded in order to monitor the component fractions and their distributions. A general expression for optical projections is given from considering Eqs. (3) and (10) and Fig. 23:

$$\Psi(t;\theta) = \int\int I(x, y)\delta(t - x\cos\theta - y\sin\theta)dxdy \tag{35}$$

where the wave impedance $I(x, y)$ is defined to specify the type of optical projection, for example, $I(x, y) = (K/\lambda)\cdot[\rho(x, y) - \rho_o]$ for interferometric projection shown in Eq. (10). The impedance or density field I must be reconstructed from the measured projection data Ψ by inverting the integral of Eq. (35). Table III presents the projection Ψ and the impedance I for different optical techniques [27].

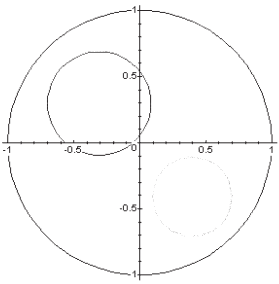
The use of the conventional algebraic reconstruction algorithm (ART) finds limited usefulness in quickly estimating the number of bubbles and their approximate sizes and locations. This approximate information can be used as intelligent initial information for more sophisticated and time-consuming hybrid GA-tomography.

The procedure involved for the execution of the Hybrid-GA is as follows: Considering a computer synthesized reference phantom field consisting of two circular bubbles (Fig. 32(a)), the ART program is executed to get an approximate size, shape, and number of bubbles. A value of "1" is assigned for the pixels located inside the bubbles and "0" otherwise. During the iterative update, pixel points with values larger than 0.5 are considered inside the bubble, and grid points of less than 0.5 are considered outside the bubble. The number of bubbles, an approximate estimate of center locations

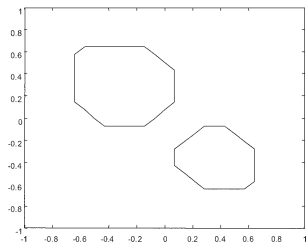
TABLE III

Examples of the Line-of-Sight Projection ψ and the Impedance I for Different Optical Techniques

Optical technique	Projection (ψ)	Impedance (I)
Interferometry	Fringe pattern	Density field
Speckle photography	Beam deflection angles	Lateral gradient of density field
X-ray (CT) or γ-ray	Light attenuation	Density field
Ultrasonic (non-optical)	Pressure wave transit time	Density function (speed of sound)

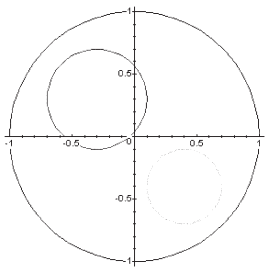


(a) Reference Phantom field

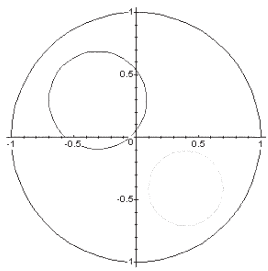


(b) ART Reconstruction

Circle 1: (0.4, -0.4, 0.3)
Circle 2: (-0.3, 0.3, 0.4)



(c) GA Reconstruction



(d) CS-GA Reco

FIG. 32. Reconstruction of a phantom field (a), using ART (b), GA (c), and Concurrent Simplex-GA (d) [28].

(x, y), and radii (r) of the bubbles are estimated from the ART output. This result is inputted as the first individual of the initial population for the GA-based reconstruction, which uses two circular basis functions and simultaneously optimizes the basis function parameters, (x_i, y_i, r_i), $i = 1, 2, \ldots, N_B$ (number of bubbles). The remaining individuals of the initial population are randomly generated to constitute a pool of solution candidates.

The ART calculations use 15 by 15 square-pixel grid points, 15 projections at equal angular intervals, and 15 data points per projection. The ART results after 1000 iterations provides an estimation of $(x_i, y_i, r_i) = (0.4475, -0.4475, 0.2775)$ and $(-0.2775, 0.2775, 0.3625)$ for each of the two approximate bubbles (Fig. 32(b)). The same is inputted for

the first individual of the population for the GA run and the remaining populations are randomly selected. The final reconstruction of the GA and the Concurrent Simplex-GA are shown in Fig. 32(c) and Fig. 32(d), respectively. The GA calculations are taken for 1000 generations with 15 projections and 15 rays per projections. For both GA and CS-GA use 167 individuals for the evolving population, 50% for the crossover probability, and 10% for the mutation probability. The CS-GA uses 40% Simplex ratio and the calculations were carried out in 1999 by a Pentium Pro 200 MHz PC and Sun-Sparc Unix machine [28].

In order to understand the influence of the various parameters on the fitness values and the reconstruction accuracy, the ratio of simplex is varied from 0 to 80% and the results are shown in Fig. 33. After 15,000 trials, equivalent to the number of generations, the CS-GA gives a virtually perfect reconstruction of the phantom, irrespective of the ratio of CS, except for the case of 0% Simplex. Table IV shows the fitness values of the best individual after 15,000 trials and the required computational time on the Unix platform. For the given example and the GA parametric conditions the Optimum CS ratio varies from 40 to 60%. A higher Simplex ratio, though converging faster initially, loses its ability to search a wider solution surface once it approaches a peak. On the other hand, a lower CS ratio takes more time to reach a particular value of fitness as it lacks the advantage of expedited convergence of Simplex.

For the case of the 40% Simplex ratio, a comparison of the fitness values with and without imposing the preliminary information, estimated from the ART result, is shown in Fig. 34. The inclusion of the imposed individual by ART expedites the convergence and also improves the

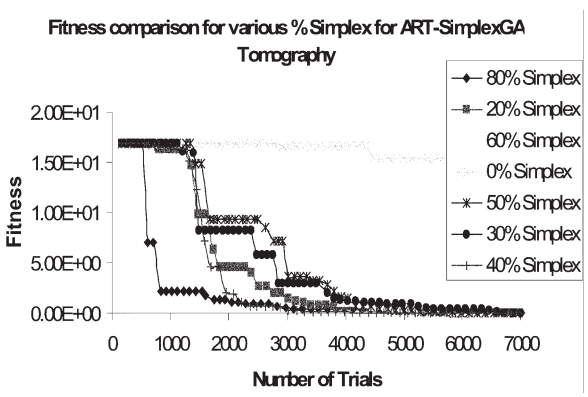


FIG. 33. Fitness comparison for various Concurrent Simplex ratios [28].

TABLE IV
FINAL FITNESS VALUES FOR 15,000 TRIALS FOR DIFFERENT CONCURRENT SIMPLEX (CS) RATIOS

Ratio of simplex	Computational time (s)	Final fitness values
0	253	7.42
20%	245	1.00×10^{-9}
30%	252	3.76×10^{-8}
40%	241	2.96×10^{-9}
50%	242	2.62×10^{-8}
60%	252	7.50×10^{-8}
80%	242	1.01×10^{-7}

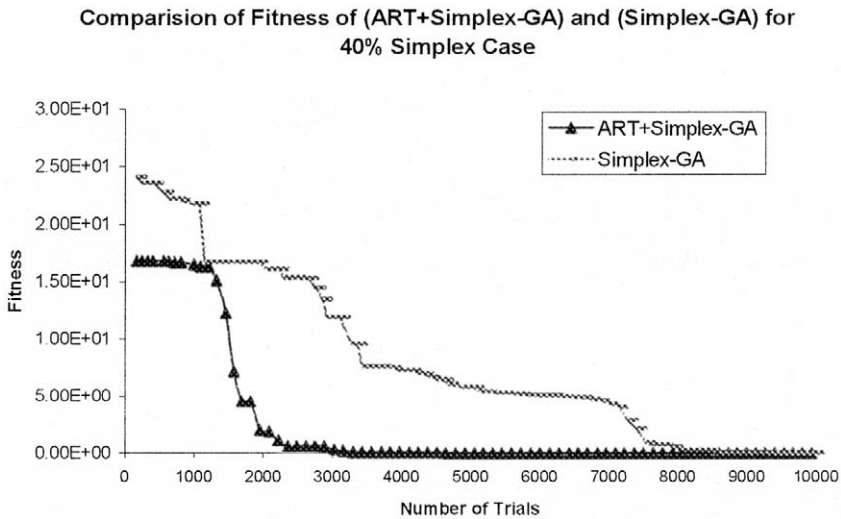


FIG. 34. Fitness value comparison between with and without ART imposed result for 40% Concurrent Simplex ratio [28].

asymptotic convergence values as early as 3000 trials vs. 10,000 trials without ART.

While the ART method converges to a nearly unique solution regardless of its initial guess of solution field, the GA-based method is a statistical method that can result in different solutions depending on the calculation procedures. After several trials, the best solution (individual) must be selected to give the highest fitness or most accurate reconstruction. The way to ensure the repeatability of the GA-method is to specify a fixed random

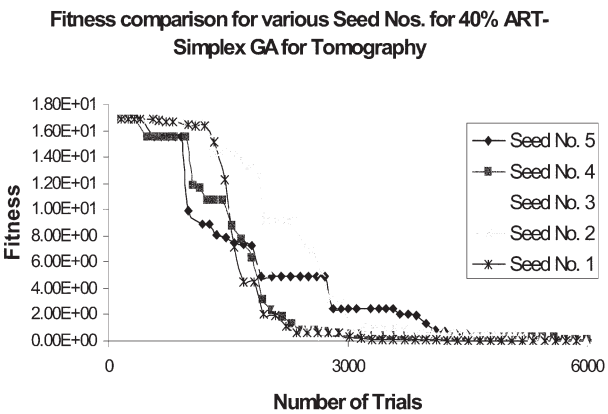


FIG. 35. Fitness comparison for various Seed Numbers for 40% Concurrent Simplex case of the Hybrid CS-GA [28].

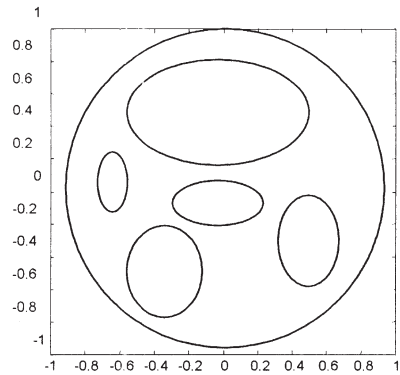
number seeder, a lengthy series of random numbers, which controls each and every manipulation step. Several different random number seeders must be examined to select the best seeder that gives the highest reconstruction accuracy. Though this additional effort of the GA in determining the available best solution seems cumbersome and detrimental, such a trial-and-error search is commonly required in most statistical iterative methods.

To examine the statistical nature of the Hybrid CS-GA reconstruction calculations, additional analysis has been made by using different Seed numbers, and the result is shown in Fig. 35. For all five seeds, the fitness reaches a sufficiently common low range beyond a certain number of trials.

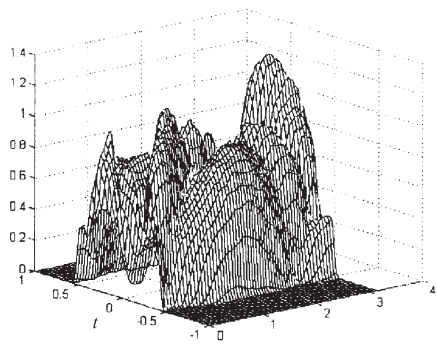
For a more elaborate example, a phantom density field is chosen as a summation of five elliptic bubbles (Fig. 36(a)). Thus, the jth bubble is described as an elliptic function g_j as [29]:

$$g_j(x, y; \mathrm{A}_j) = \begin{cases} 1 & \left(\dfrac{x - x_j}{a_j}\right)^2 + \left(\dfrac{y - y_j}{b_j}\right)^2 < 1 \\ 0 & \text{otherwise} \end{cases} \tag{36}$$

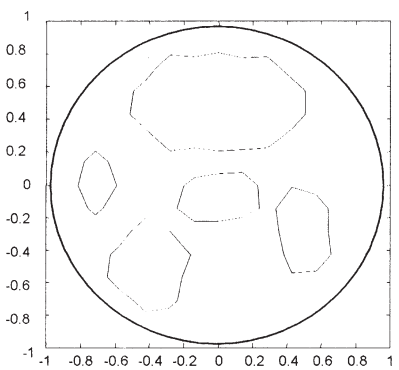
where $\mathbf{A}_j = (x_j, y_j, a_j, b_j)$, whose components correspond to the center location and the major and minor axes of the ellipse, respectively. Binary basis function values are given to represent normalized field impedance (density) of 1 for the region inside the bubble and 0 for liquid or solid outside the bubble.



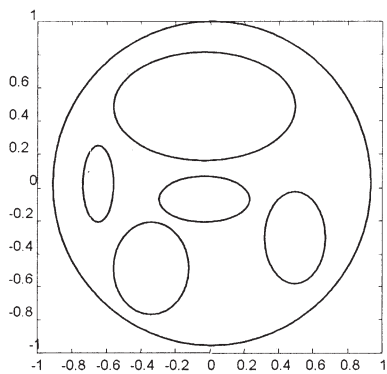
(a) Phantom field



(b) Radon transformation



(c) ART reconstruction



(d) GA-based reconstruction using the ART reconstruction results (c) as preliminary information

FIG. 36. Computer synthesized two-phase field (a), its line-of-sight (Radon transform) projection (b), reconstruction image using ART (c), and reconstruction image by the combined use of ART and GA (d) [27].

The line-of-sight projection, Eq. (25), is also known as the Radon transformation [1] of function $I(x, y)$, and the Radon transformation of the reference phantom field is shown as a function of the normalized projection plane coordinate t and the normalized projection angle θ in Fig. 36(b). If the inverse Radon transformation is known for all $\Psi(t, \theta)$ and sufficient numbers of projections are available, a direct mathematical reconstruction would be possible from Fig. 36(b) to Fig. 36(a). The inverse Radon transformations of most asymmetric impedance functions

are not available and only a finite number of projections are allowed in practice.

As discussed earlier, ART gives an approximate reconstruction field that can be used as input for a more rigorous hybrid-GA reconstruction algorithm. Figure 36(c) shows the reconstructed field using ART where the field is described by 15×15 square pixels with 225 unknown pixel values, p_i $(i = 1, 2, 3, \ldots, 225)$. The calculation used 15 equally angled projections ($N_p = 15$) and 15 rays ($N_R = 15$) per projection. The calculation was stopped at approximately 650 iterations, beyond which the convergence is nearly saturated. Although the calculation demands very low CPU, the reconstruction deficiency is apparent from the use of finite square pixel representation of the field. The reconstructed bubble images roughly identify elliptic shapes. On the other hand, ART provides the number of bubbles in most cases, and their approximate sizes and locations. This insinuates that the ART reconstruction results can be used as intelligent initial information for more sophisticated and time-consuming GA-based reconstruction.

The adoption of genetic algorithm (GA) optimization for image reconstruction shows good potential for improvement, particularly when multiple sets of unknowns have to be simultaneously optimized. One distinctive advantage of GA-based reconstruction is to allow the use of any arbitrary types of basis functions whose summation conforms to the density field to be reconstructed. In addition, the basis function can have as many parameters as necessary since the GA-based tomography can simultaneously handle multiple sets of unknowns. The most obvious choice of the present example, reflected from the preliminary ART solution, will be five elliptic-type basis functions.

The GA-based reconstruction result of Fig. 36(d) was obtained by using five elliptic basis functions conforming to the field with 20 unknown parameters $\mathbf{A}_j = (x_j, y_j, a_j, b_j)$, $j = 1, 2, \ldots, 5$. The algorithm uses a population of 100 individuals evolving to 1000 generations with 50% crossover probability and 10% mutation probability. The number of projections and the number of rays per projection are 15 for both, as in the ART calculations in Fig. 36(c). The result shows a noticeable improvement over the ART results, and no apparent distinction is discerned compared with the phantom given in Fig. 36(a).

As expected, the GA's simultaneous handling of multiple sets of parameters requires a massive computation amount and demands more CPU time by two to three orders of magnitude compared with ART. At the cost of computation, the GA-based tomography reconstructs far more accurate shapes and locations of the bubbles. If the GA were used without the help of ART in estimating the bubble parameters, its CPU time could be

TABLE V
COMPARISON OF FEATURES OF ALGEBRAIC RECONSTRUCTION TECHNIQUE (ART) AND GENETIC ALGORITHM (GA)-BASED RECONSTRUCTION

ART	GA-based tomography
The number of bubbles and approximate bubble shapes/locations are quickly available.	Accurate reconstruction of bubble shapes and sizes is possible by using an appropriate type of basis functions.
Linear optimization restricted to a single set of parameters: the algorithm is simple and fast.	Nonlinear optimization allowing multiple sets of parameters: robust, but needs lengthy computation.
Large number of total unknowns identical to the total number of pixels: $15 \times 15 = 225$.	Fewer number of total unknowns: 5 (elliptic bubbles) $\times$ 4 (x and y locations of the bubble center and its major/minor axes) $= 20$.

significantly longer. The CPU time of GA increases geometrically with the number of unknown parameter sets. Table V summarizes the individual ART and GA techniques in terms of their comparison, emphasizing the potential merit of their combined use.

VI. Three-Dimensional Reconstructions

As stated earlier in Section III, three-dimensional distribution functions $f(x, y, z)$ can be reconstructed by stacking up a number of two-dimensional cross-section data $f(x, y)$, individually reconstructed at various z locations. Each cross-sectional data $f(x, y)$ can be reconstructed from a set of multiple projection data $p(\xi, z)_\theta$ taken at different projection angles. In this case, there will be no correlation between two different cross-sectional data. When the number of projections is limited, in particular, the individually cross-sectioned reconstruction may well amplify errors in simply stacking up the plane. Therein, a direct three-dimensional reconstruction may be required to account for the field continuity and correlation in the z-direction.

The concept of the elementary distribution function (EDF) has been extended to three-dimensional domain to represent $f(x, y, z)$ as the summation of $f_i^*(x, y, z)$, as schematically shown in Fig. 37.

$$f^*(x, y, z) = \sum f_i^*(x, y, z) \tag{37}$$

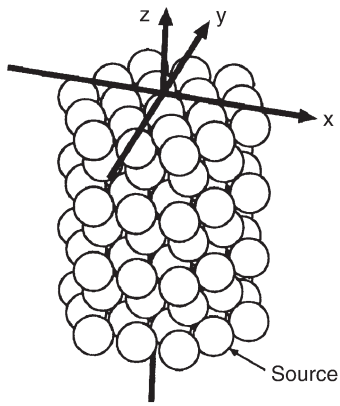


FIG. 37. Sample location of 3-D EDF distributions [30].

where $f_i^*(x, y, z)$ is an assumed density distribution that is shown as each sphere in Fig. 37. From assumed EDFs, $f_i^*(x, y, z)$ for $i = 1, 2, \ldots, N$, the projection data, $p^*(\xi, \eta)_{\theta,\phi}$, are calculated from the specified projection function or physical relationship between the field and the projection. If the calculated projection data $p^*(\xi, \eta)_{\theta,\phi}$ falls within a small convergence from the measured projection $p(\xi, \eta)_{\theta,\phi}$, the assumed density distribution $f^*(x, y, z)$ is accepted as the solution for the reconstructed field.

Any arbitrary function can be used as the form of EDFs, for example, a diffusion function with two parameters [16]. For the present three-dimensional reconstruction, the Gaussian function is used for EDF, i.e.,

$$f_i^*(x, y, z) = a_i \exp\left(-\frac{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}{\sigma_i^2}\right) \tag{38}$$

where a_i is the magnification factor of the EDF, σ_i is the spatial distribution parameters, and (x_i, y_i, z_i) denotes the center of the Gaussian function, which is also called the source location. In the spatial domain, the source locations are placed in the form of a close-packed hexagon, as shown in Fig. 37. The closed form of the Gaussian EDFs allowed mathematical evaluation of the projection data $p^*(\xi, \eta)_{\theta,\phi}$. The optimizing parameters, a_i, σ_i, x_i, y_i, z_i that respectively determine the EDF amplitude, its spread and three-dimensional locations, are optimized by the GA-based tomography.

As an illustrative example, a steady state asymmetric flow field is examined by the Mach–Zehnder interferometric projections [30]. Figure 38

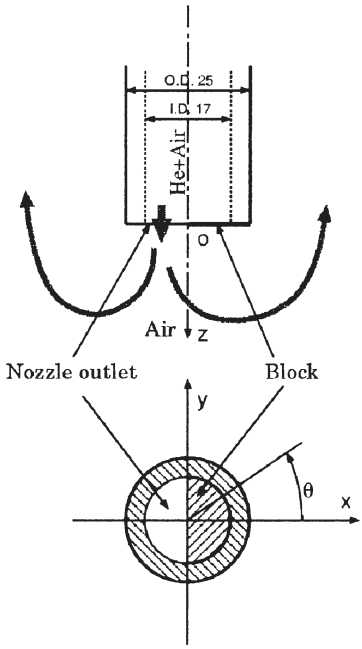


FIG. 38. Schematic test section [30].

shows the experimental apparatus schematically, i.e., stratified flow field by a downward jet of 20% helium–80% air mixture gas into ambient air. The nozzle outlet is a half circle causing the flow field to be asymmetric. The flow rate of the mixture gas is set to 2 L/min and the average flow velocity of 29.3 cm/s at the exit. Figure 39 shows the example of the interferogram images taken at two different θ-angles. The number of ''crossing'' fringes and their directions represent the line-of-sight optical path length differentials because of the density variation in the field. Figure 40 shows a comparison of the reconstructed density distributions on a selected vertical cross section, (a) the FTT reconstruction from 12 projection images, and (b) the GA reconstruction from 3 projection images. In the FTT result, despite its four-times larger number of projections, the edge area shows substantial amount of errors, primarily because of the truncated projection data. The GA result shows fairly comparable reconstruction of the three-dimensional density distribution even from just 3 projection images. The GA technique proves useful to achieve a direct three-dimensional reconstruction from a substantially fewer number of projection images.

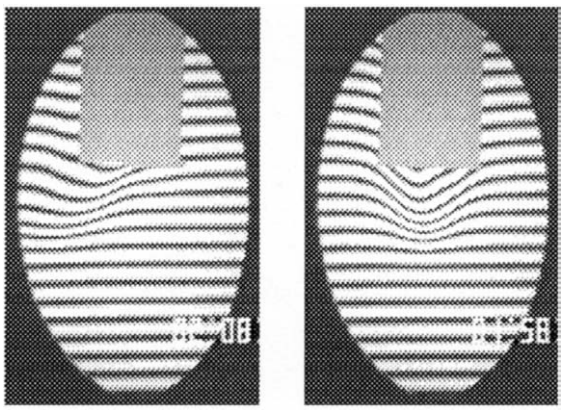


FIG. 39. Line-of-sight interferometric fringes [30].

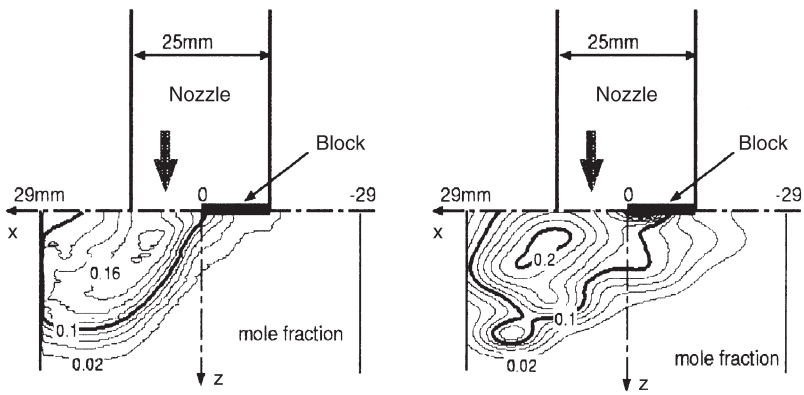


FIG. 40. Three-dimensionally reconstructed results [30].

VII. Concluding Remarks

Focusing solely on the fitness of individual solution candidates, the ability of the GA to proceed toward the optimum solution without any preconceived ideas of what a good solution should be, gives the genetic algorithm superiority over other optimization techniques. The GA-based tomography shows generally more accurate image reconstruction compared with the conventional Fourier Transformation Technique (FTT) or Algebraic Reconstruction Technique (ART), at the cost of longer computer CPU times that are required to carry out the massive calculations for its combinatorial and multi-parallel optimization feature. In order

to expedite the evolutionary solution search of the GA, a hybridization idea has been proposed to use Concurrent Simplex (CS) in conjunction with the GA. Also, a tandem use of ART to provide intelligent initial information for the more rigorous GA evolution shows a significant expedition in the solution search. The hybridized tomographic reconstruction algorithm (ART + CS + GA) shows highly accurate image reconstruction for both continuous and discrete impedance fields.

Suggestions for future development of the GA-based tomographic algorithm can include:

(1) More elaborate selection for the basis function type should be further studied to optimally represent the reconstructing field to the extent that the preliminary knowledge on the field is available.

(2) The GA operation has a number of adjustable parameters including the number and selection of initial population, the mating probability, the number of genes to be crossed, the mutation probability, and the random number seed controlling the individual trials. In the present study, the selections of these parameters have been made based on trial-and-error to assure we have the best reconstruction results. This has been possible only because the reference field as known is synthesized by computer. In practice, the reconstruction accuracy may not be easily assessed since the field is unknown. More rigorous investigation on the optimization of those GA parameters could include bringing advanced intelligence from computer engineering.

(3) Since the GA-based tomography requires excessive computational times, the hybridizing idea for expedition of the calculation process needs to be developed more. For example, more advanced ART techniques may be used to reconstruct the field to the extent possible, and then, the GA-CS tomography can be used to improve the reconstruction accuracy and to ensure avoiding a local-peak trap. The optimal ratio of CS should also be further investigated.

Nomenclature

A	a constraint specified by a type of projection	f	actual field
		$\hat{f}$	guessed or intermediate objective function to be optimized
a	height of a basis function		
$a, b, c, \ldots$	genes or unknowns to be determined	g	elliptic basis function
		H	helium mass fraction
b	basis function	I	individual solution candidate, or impedance
C	multiplicative correction vector		
		K	Gladstone–Dale constant

l	coordinate parallel to the ray direction
M	the total number of basis functions
N	the total number of individuals
N_p	the total number of projections
N_R	the total number of data realization per projection
n	index of refraction
O	object (solution) vector
p	pixel value
R	oxygen percentage
(r, θ)	polar coordinates
t	spread of a basis function
(s, t)	projection coordinate
W	projection function
$W_{m,n}$	weighting factor of MART
w_m	mth row of projection matrix
(x, y)	Cartesian coordinates

Greek Letters

δ	Dirac-delta function
λ	wave length of light
ρ	density field
ψ	projection image

Subscripts

E	elite group
S	simplex group
i	the ith individuals
ij	the ith basis function and the jth solution candidate
o	vacuum

Superscripts

*	refers to a solution candidate

Abbreviations

ART	Algebraic Reconstruction Technique
CG	Center of Gravity
CS	Concurrent Simplex
CT	Computer Tomography
EDF	Elementary Density Function, equivalent to a basis function
FC	Fourier Convolution
FTT	Fourier Transformation Technique
GA	Genetic Algorithm
MART	Multiplicative Algebraic Reconstruction Technique
MRI	Magnetic Resonance Imaging
PIV	Particle Image Velocimetry

Glossary

Crossover	Randomly swapping genes between the two mating solution candidates
Elitism	Survival of the solution candidates with the highest fitness—a device to ensure the fitness improvement with evolution
Fitness	Proximity of a solution candidate to a true solution
Generation	Iterative evolution of GA
Genes	Unknowns
Implicit Parallelism	Feature of handling multiple sets of unknowns simultaneously
Individual	A solution candidate
Mating	Coupling of two solution candidates
Mutation	Random choice of gene(s) of an individual and random change of the gene value—a device to prevent a local-peak trapping
Population	A pool of solution candidates
Ranking	Order of fitness of solution candidates

References

1. Deans, A. R. (1983). "The Radon Transformation and Some of Its Applications". Wiley, New York.
2. Kak, A. C. and Slaney, M. (1987). "Principles of Computerized Tomographic Imaging", pp. 275–296. IEEE Press, New York.

3. Liu, T. C., Merzkirch, W., and Oberste-Lehn, K. (1989). Optical tomography applied to speckle photographic measurement of asymmetric flows with variable density. *Exp. Fluids* **7**, 157–163.
4. Sweeny, D. W. and Vest, C. M. (1973). Reconstruction of three-dimensional refractive index fields from multidirectional interferometric data. *App. Opt.* **12**, 2649–2664.
5. Rangayyan, D. W., Dhawan, A. P., and Gordon, R. (1985). Algorithms for limited view computed tomography: an annotated bibliography and a challenge. *App. Opt.* **24**, 4000–4012.
6. Brigham, E. O. (1988). ''The Fast Fourier Transform and Its Applications''. Prentice-Hall, Englewood Cliffs.
7. Goulard, R. and Emmerman, P. J. (1980). Combustion diagnostics by multiangular absorption. *In* ''Inverse Scattering Problems in Optics'', pp. 215–235. Springer Verlag, Berlin and New York.
8. Gordon, R. (1974). A tutorial on ART. *IEEE Trans.* **NS-21**, 78–92.
9. Swindell, W. and Barrett, H. H. (December, 1977). Computerized tomography: taking sectional X-rays. *Phys. Today* 32–41.
10. Decker, A. J. (1993). Tomographic methods in flow diagnostics. *NASA Report* **106330**.
11. Verhoeven, D. (1993). Limited-data computed tomography algorithms for the physical sciences. *App. Opt.* **32**, 3736–3754.
12. Merzkirch, W. (1987). ''Flow Visualization''. Chapter 3. Academic Press, Orlando.
13. Ko, H. S. and Kihm, K. D. (1999). An extended algebraic reconstruction technique (ART) for density-gradient projections: laser speckle photographic tomography. *Exp. Fluids* **27**, 542–550.
14. Goldberg, D. E. (1989). ''Genetic Algorithms in Search, Optimization and Machine Language''. Addison-Wesley, Reading, Chapters 1 and 2.
15. Tsukiji, T., Ohyama, R., and Kaneko, K. (1994). Investigation of particle tracking method using genetic algorithm. *In* ''Proceedings 3rd Asian Symposium on Visualization'', pp. 663–668. Chiba, Japan.
16. Tsuru, D., Okamoto, K., Madarame, H., and Fumizawa, M. (1994). Measurement of density distribution using optical tomography with genetic algorithm. *In* ''Proceedings 4th Fluid Control Measurement and Visualization (FLUCOME 94)'', pp. 971–976. Toulouse, France.
17. Holland, J. (1975). ''Adaptation in Natural and Artificial Systems''. U. Michigan Press, Ann Arbor.
18. Kihm, K. D. and Lyons, D. P. (1996). Optical tomography using a genetic algorithm (GA). *Optics Lett.* **21**(17), 1327–1329.
19. Kihm, K. D., Okamoto, K., Tsuru, D., and Ko, H. S. (1996). Adoption of a genetic algorithm (GA) for tomographic reconstruction of line-of-sight pptical images. *Exp. Fluids* **22**, 137–143.
20. Kihm, K. D. (1997). Laser speckle photography technique applied for heat and mass transfer problems. *In* ''Advances in Heat Transfer'', pp. 255–311. Academic Press, San Diego.
21. Born, M. and Wolf, E. (1959). ''Principle of Optics'', pp. 128–130. Pergamon Press, Oxford.
22. Vest, C. M. (1979). ''Holographic Interferometry'', pp. 344–377. John-Wiley, New York.
23. Kihm, K. D., Okamoto, K., and Tsuru, D. (July, 1996). Adoption of genetic algorithm (GA) for accurate tomographic image reconstruction under limited projections. *In* ''Proc. 1996 ASME Fluid Engineering Division Summer Meeting'', Vol. FED-Vol. 239-IV, pp. 459–464. San Diego, California.
24. Nelder, J. A. and Mead, R. *Computer* **7**, 308.

25. Randolph, D. (1994). ''A Hybrid of the Genetic Algorithm and Concurrent Simplex'', M.S. Thesis, Texas A&M University: College Station.
26. Lyons, D. P. and Kihm, K. D. (1997). Tomographic image reconstruction using a hybrid genetic algorithm. *Opt. Lett.* **22**(12), 847–849.
27. Kihm, K. D., Ko, H. S., and Lyons, D. P. (1998). Tomographic identification of gas bubbles in two-phase flows with the combined use of the algebraic reconstruction technique (ART) and the genetic algorithm (GA). *Opt. Lett.* **23**(9), 658–660.
28. Rajesh, A., Ko, H. S., and Kihm, K. D. (1998). Hybrid approach to tomographic reconstruction of bubbles in two-phase flows using genetic algorithm, *Proc. ASME-IMECE*, HTD-Vol. 361-5, November, 1998, pp. 515–521. Anaheim, California.
29. Rajesh, A. and Kihm, K. D. (May 1999). Tomographic reconstruction of bubbles—a hybrid approach using ART, genetic algorithms and simplex method. *In* ''Proc. Second Pacific Symposium on Flow Visualization and Image Processing'' (PSFVIP-2), Paper No. PF111. Honolulu, Hawaii.
30. Okamoto, K., Tsuru, D., Feng, J. and Fumizawa, M. (June 25–28, 1996). Reconstruction of three-dimensional distribution with genetic algorithm. *In* ''The Ninth International Symposium on Transport Phenomena in Thermal-Fluids Engineering Proceeding'', pp. 1030–1035. Singapore.

Multiplicity and Stability of Convection in Curved Ducts: Review and Progress

LIQIU WANG and TIANLIANG YANG

Department of Mechanical Engineering, The University of Hong Kong, Hong Kong, P.R. of China

Abstract

The present review starts with a literature review of multiplicity and stability of convection in curved ducts, a phenomenon widely observed from piping systems for various fluids to blood flows in the human arterial system. As a result of this review, several key unresolved issues are identified: (i) multiple steady-state solutions up to a high value of Dean number; (ii) conclusive determination of linear stability of multiple solutions and their dynamic responses to finite random disturbances; (iii) gain and loss of flow stability along solution branches without passing limit/bifurcation points; (iv) development of temporal oscillations in the range of dynamic parameters where no stable steady fully developed flows exist; and (v) phenomena related to the transition to the turbulence at high Dean numbers such as temporal oscillation, period doubling, intermittency, and chaotic oscillation.

An attempt to address these issues has led to our recent numerical study on the fully-developed bifurcation structure and stability of forced convection in a curved duct of square cross-section. In addition to the extension of three known solution branches to the high Dean number region, three new asymmetric solution branches are found from three symmetry-breaking bifurcation points on the isolated symmetric branch. The flows on these new branches are either an asymmetric 2-cell state or an asymmetric 7-cell structure. The linear stability of multiple solutions is *conclusively* determined by solving the eigenvalue system for all eigenvalues. Only 2-cell flows on the primary symmetric branch and on the part of isolated symmetric branch are linearly stable. The symmetric 6-cell flow is also linearly unstable under asymmetric disturbances although it was ascertained to be

Advances in Heat Transfer
Volume 38 ISSN 0065-2717
DOI 10.1016/S0065-2717(04)38004-4

stable under symmetric disturbances in the literature. The linear stability is observed to change along some solution branches even without passing any bifurcation or limit points. Furthermore, dynamic responses of the multiple solutions to finite random disturbances are also examined by the direct transient computation. It is found that physically realizable fully-developed flows possibly evolve, as the Dean number increases, from a stable steady 2-cell state at a lower Dean number to a temporal periodic oscillation state, another stable steady 2-cell state, a temporal oscillation with the intermittency, and a chaotic temporal oscillation. Among them, three temporal oscillation states have not been reported in the literature. A temporal periodic oscillation between symmetric/asymmetric 2-cell flows and symmetric/asymmetric 4-cell flows are found in the range where there are no stable steady fully-developed solutions. The symmetry-breaking point on the primary solution branch is determined to be a subcritical Hopf point by the transient computation.

I. Introduction

The present review is mainly on the fully-developed bifurcation structure and stability of the forced convection in a stationary curved duct of square cross-section. The flow geometry is illustrated in Fig. 1 with (R, Z, ϕ) as the radial (normal), spanwise, and streamwise directions, respectively. A viscous fluid is driven by a streamwise pressure gradient to flow through a square duct with a streamwise curvature and a uniform wall heat flux. Such flows and transport phenomena have been and still are the object of intense investigations due to its intrinsic interest and its relevance to a host of areas

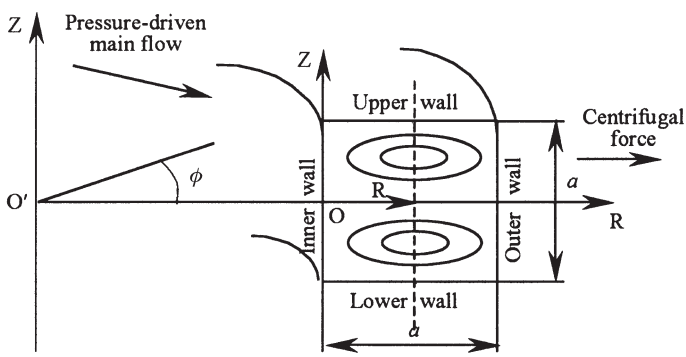


FIG. 1. Physical problem and coordinate system.

involving curved passages and surfaces. For instance, all piping systems have curved sections. It is important to know the pressure drop in order to calculate the pumping power needed to overcome curvature-induced pressure losses. Because secondary flows can enhance heat and mass transfer, knowledge of the magnitude of this effect in different ranges of operating parameters is important in designing and operating heat exchangers, chemical mixers, blood oxygenators, and dialysis machines. To avoid or reduce the flow-induced vibration and noise in piping systems, heat exchangers, or chemical mixers, we need to know when temporal oscillation appears. Blood flow in the human arterial system has been of particular interest. The largest vessel in this system, the aorta, is highly curved. Of particular interest concerning the flow in this vessel is the location of sites of maximum and minimum wall shear stress, because such sites may be the sites of cholesterol buildup on vessel walls, and thus may play a major role in atherogenesis [1].

The present review surveys the literature and presents the results from our recent study on the multiplicity and stability of convection in curved ducts. The readers are referred to Refs. [2–5] for more comprehensive reviews on general aspects of flow and heat transfer in curved ducts/pipes. For our purposes, we focus our literature review mainly on flow bifurcation and stability of fully-developed flows in channels with an infinite span and in ducts with a square cross-section (Section II). The review uncovers several key unresolved issues of academic and practical significance and stimulates our recent efforts for a relatively comprehensive study [6]. We detail our theoretical model in Section III, the numerical method in Section IV, and the main findings in Section V.

II. Literature Review

A. Channels with an Infinite Span

Using analytic means for small gap channels of an infinite span, Dean [7] first demonstrated that the curved channel flow is unstable to small amplitude disturbances for values of $Re\sqrt{\sigma}$ greater than a critical value (here Re is the Reynolds number, and σ is the curvature ratio defined by a/R_c, the ratio of channel radial dimension a over the radius of curvature R_c, Fig. 1). This parameter is now named as the Dean number De, the ratio of the square root of the product of inertia and centrifugal forces to the viscous force. In terms of Dean number versus spanwise wavenumber at a particular curvature ratio, the neutral stability curve separates unstable flow at higher Dean numbers from *curved channel Poiseulle flow* (CCPF) at

lower Dean numbers. With CCPF, the velocity in the channel is purely streamwise with a profile similar to the parabola of plane channel flow. The main difference is evident in the maximum velocity, which is shifted slightly toward the outer wall. The focus of points along the neutral stability curve defines the critical Dean number De_c and critical spanwise wavenumber α_c. Dean [7] found that $De_c = 35.94$ and $\alpha_c = 3.96$. The confirmation of Dean's results was provided by independent theoretical analyses of Yih and Sangster [8], Reid [9], and Hammerlin [10] using different methods and by experiments of Brewster *et al.* [11] who find that $De_c = 36.5 \pm 1.1$ and $\alpha_c = 4.9 \pm 0.8$.

It is of interest to compare the Reynolds number for the onset of instability given by Dean with the one at which Tollmien–Schlichting waves [12] may set in for *strictly* straight plane channel flows. It is known that in the latter case, the onset of instability occurs when $Re_c \doteq 7070$ [13,14]. Clearly, the criterion given by Dean must cease to be valid when

$$\sigma < 2.6 \times 10^{-5},$$

which is very small. Therefore, we typically expect to observe streamwise oriented vortices instead of TS-waves in curved channel flow when the Reynolds number is increased from zero.

Kelleher *et al.* [15] and Ligrani and Niver [16] made experiments of flow visualization in curved channels with a high spanwise aspect ratio. They observed mushroom shaped smoke patterns, which are believed to be indicative of Dean vortices. As shown by hot-wire measurements, the vortices give rise to large spanwise variations of the streamwise velocity [15].

The secondary instability of Dean-vortices was first examined numerically by Finlay *et al.* [17] for a curved channel of the infinite span. Due to the instability of Dean-vortices subjected to streamwise wavy disturbances, two traveling vortex flows, namely, the undulating vortex flow and the twisting vortex flow, are developed on top of the primary instability (Dean instability). These two kinds of vortex flows were confirmed experimentally by Ligrani and Niver [16] and Ligrani *et al.* [18]. The linear stability analysis of spanwise and radial profiles of the streamwise velocity shows that the spanwise inflectional profile is associated with the secondary twisting motion while the undulating motion is related to the radial profile [19].

The secondary instability of Dean-vortices with respect to 2D spanwise-periodic perturbations (i.e., Eckhaus stability) was investigated numerically by Guo and Finlay [20] for an infinite-span cross section. Because of this instability, merging and splitting of vortex pairs appear. Experimental confirmation of the splitting and merging of vortex pairs was made by

Ligrani and Niver [16], Alfredsson and Persson [21], and Matsson and Alfredsson [22]. Further developments on the Eckhaus instability problem have been made by Matsson and Alfredsson [23] and Bottaro [24].

Three-dimensional spatially developing vortices were studied by Guo and Finlay [25]. They found that the selection of the spanwise wave number depends on the energy level of vortices. Different spanwise wave lengths could appear simultaneously without interaction when the energy level is low. However, merging and splitting of vortices occur if the vortices have reached a certain non-linear level.

Measurements on the streamwise development of the Dean instability [23] show that the merging and splitting of vortex pairs are not time-dependent as in the case by Ligrani and Niver [16]. Numerical simulations by Bottaro [24], however, show that the Eckhaus instability is time-dependent in his temporal model (streamwise periodic boundary conditions), and steady in his spatial model (inlet–outlet conditions). For the experiments in Ref. [23] and the spatial model in Ref. [24], the vortices are locked into place by zero-frequency wall disturbances and thus the merging/splitting phenomena appear as stationary, spatially developing events. For the experiments in Ref. [16] and the temporal model in Ref. [24], however, the vortices are free to meander along the span. Thus the merging/splitting events are very receptive and responsive to free-stream disturbances and appear time-dependent.

B. Ducts with a Square Cross-Section

Existence of upper and lower walls induces pressure gradients along the walls, which manifest themselves as one pair of counter-rotating vortices on the cross plane [26,27]. These vortices are named Ekman vortices in the literature and are present for any non-zero value of a Dean number. As a consequence, the fully developed flow in a duct with upper and lower walls is a symmetric 2-cell state even at lower Dean numbers, rather than the purely streamwise CCPF. As the Dean number increases, this 2-cell flow loses its stability due to a centrifugal instability [28] (also called Dean instability in the literature). As explained by Cheng *et al*. [29], in the region near the central outer wall, the pressure gradient across the duct in the radial direction is positive but the centrifugal force decreases from a maximum value to zero at the outer wall. The instability due to the imbalance between the inward pressure gradient and the outward centrifugal force generates an unstable region. Cheng and Akiyama [30] first reported, by their numerical study of loosely coiled ducts of rectangular cross-section, that this instability manifests itself in the form of a secondary pair of counter-rotating vortices near the central outer wall. The fully developed

flow in a square duct is thus a 4-cell state in which the secondary pair of vortices near the central outer wall are named Dean vortices. This 4-cell flow was further detailed numerically by Cheng *et al.* [29,31] and Joseph *et al.* [32], and was confirmed experimentally by Cheng *et al.* [33]. The critical Dean number for the switch from the 2-cell pattern to the 4-cell pattern was found to be about 118 by Cheng *et al.* [29,33], while Joseph *et al.* [32] observed that the critical value is about 100 for square ducts. The difference regarding this critical value indicates the existence of a range of Dean numbers where 2- and 4-cell solutions co-exist.

Such co-existence of 2- and 4-cell solutions was described in Akiyama [34]. Masliyah [35] established this co-existence both numerically and experimentally for the fully developed flow in a curved duct of semicircular cross-section with a flat outer wall. Numerical studies by Winter and Brindley [36], Shanthini and Nandakumar [37], Winters [38], and Daskopoulos and Lenhoff [39] also found dual solutions for fully developed flow in a curved duct of square cross-section.

Detailed bifurcation structure and linear stability of solutions for steady, fully developed flows in a curved duct of square cross-section was determined numerically by Winters [38]. Three solution branches were found (Fig. 2). For a loosely coiled duct with $\sigma = 0.04$, the primary branch S_1 starts at a low Dean number as a symmetric (w.r.t. the horizontal central plane) 2-cell flow state (two large counter-rotating Ekman vortices) and develops,

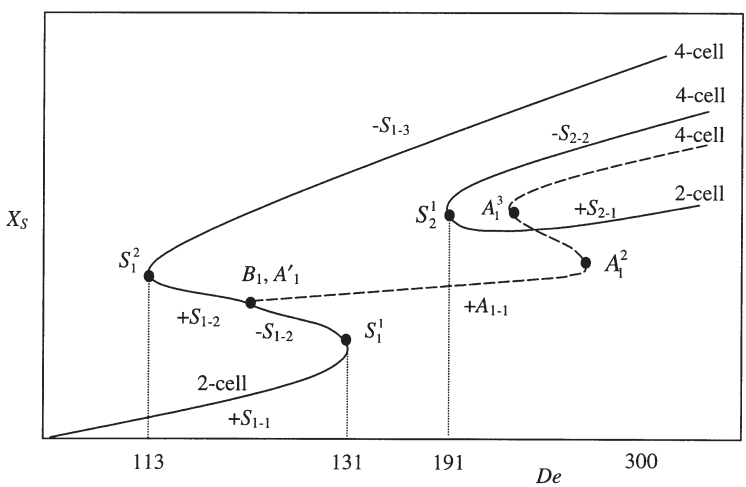


FIG. 2. Schematic state diagram of fully developed flows in a curved duct of square section. X_s is a measure of the solution. The sign of each subbranch is $(-1)^n$, where n is the number of negative eigenvalues of the Jacobian matrix. After Winters [38].

as the Dean number increases, to a symmetric 4-cell flow (two large Ekman vortices and two small Dean vortices near the center of the outer wall) through two limit points S_1^1 at $De = 131$ and S_1^2 at $De = 113$. Between these two limit points, 2-cell and 4-cell solutions co-exist. The linear stability analysis by Winters [38] determined that the 2-cell flow is stable and the 4-cell flow is unstable with respect to asymmetric disturbances. Bara *et al.* [40] experimentally confirmed, using a curved duct of 270° and constant curvature ratio of 0.0662, the solution structure up to $De = 150$, including the dual solution region. They observed the symmetric 4-cell flow although it is unstable because the asymmetric disturbances in their apparatus were small or the flow had not reached the fully developed state in their 270° curved duct.

The second branch found by Winters [38] is a pair of asymmetric solutions A_{1-1} and A_{1-2}. This branch has two pairs of limit points and arises from a symmetry-breaking bifurcation point B_1 at $De = 130$ on the primary branch S_1. The flow, consisting of two asymmetric Ekman vortices, is unstable. The third branch is an isolated branch S_2 of 2-cell and 4-cell flows above $De = 191$ (S_2^1). Winters [38] determined that the isolated symmetric 4-cell sub-branch is unstable while the isolated 2-cell sub-branch is stable. The location of limit and bifurcation points does not change much for curvature ratios less than 0.2, but at higher curvature ratios (tighter coil), they move to higher Dean numbers.

Therefore, no stable steady fully developed solutions exist between $De = 131$ and $De = 191$. This raises the question of what the physically realizable solutions are in this region. Clearly, the flows must develop steady spatial (streamwise) oscillations or fully developed temporal oscillations, or even some combination of the two. Experimentally, both oscillations were observed between symmetric/asymmetric 2-cell flows and symmetric/asymmetric 4-cell flows for a Dean number up to 269 [40–45]. The co-existence of both kinds of oscillations could be the result of development of different disturbances in the experimental apparatus. Sankar *et al.* [46] numerically predicted steady spatial oscillations at Dean numbers above 128 by using a parabolized version of steady, three-dimensional Navier–Stokes equations. However, there is no known study on prediction of fully developed temporal oscillations and how such oscillations relate to the solution structure outlined above.

It is noted that the upper limit of oscillating flow region is higher than predicted by Winters [38] (at $De = 191$). This leads to the question of whether the isolated 2-cell branch is stable near the limit point S_2^1 of $De = 191$. The stability indicator used by Winters [38] was the sign of Jacobian determinant which equals $(-1)^n$, where n is the number of negative eigenvalues of the Jacobian matrix. A negative value of this

indicator indicates instability, but a positive value is *insufficient* to indicate stability [38]. Therefore, the linear stability of solutions with a positive Jacobian determinant, including the isolated 2-cell branch, has not been conclusively determined. Furthermore, the branch stability is often determined by the stability of one point on the branch. This is partly due to the fact that the computation of the complete eigenvalue spectrum along solution branches is a computationally expensive process and partly due to the assumption prevalent in the literature that the linear stability of solutions along a solution branch is unchanged without passing limit/bifurcation points. However, based on the bifurcation and stability theory, such a change in stability is possible [47]. Therefore, a more detailed and careful linear stability analysis is desirable to *conclusively* determine the linear stability of multiple solutions.

Daskopoulos and Lenhoff [39] extended the work by Winters [38] to a higher Dean number. While no new solution branches were detected, four limit points on the isolated 2-cell branch were found as the Dean number increased. Some of these points lead to solutions with six or eight vortices. Note that Daskopoulos and Lenhoff [39] have imposed a symmetry condition on the horizontal central plane. Thus questions concerning asymmetric solutions, their stability, and symmetry-breaking bifurcations have been left unanswered.

While the linear stability analysis is efficient in terms of the computation efforts required, it suffers two fundamental defects. First, it is not applicable to a finite disturbance. With a finite disturbance, a stable solution based on the linear stability may not be always stable. Second, it provides no answer to questions related to the dynamic behavior of the solutions, including how flows approach a stable solution after a disturbance, what happens to an unstable solution after a disturbance, whether all unstable solutions at a given set of parameters respond to disturbances in the same way, and whether the disturbances turn an unstable solution into a stable one at the same parameter value. Clearly, a fully transient computation is necessary to examine dynamic responses of multiple solutions to finite random disturbances.

The bifurcation and stability of the multiple solutions available in the literature are limited to low Dean numbers (less than 250). However, solutions and their stability at high Dean numbers are more relevant to practical applications. Upon increasing the Dean number, a richer bifurcation structure and a more intricate stability feature are expected because of a stronger non-linearity of the problem. New limit/bifurcation points and solution branches are expected. Because of the lack of solution structures at high Dean numbers, there is a long-standing controversy over such solutions obtained by different methods that aimed at obtaining a solution

without considering the multiplicity [48]. To resolve this controversy, we need to find the bifurcation structure at high Dean numbers. Furthermore, it is also reasonable on the physical ground to expect phenomena related to the transition to the turbulence at high Dean numbers such as oscillation solutions, periodic doubling, intermittency, and chaotic oscillation.

It is well known that both the period-doubling bifurcation and the intermittency serve as the route to chaos in closed flow systems such as the Taylor–Couette flow and the Rayleigh–Benard problem [49–51]. It is interesting to know whether this may be changed by the pressure-driven flows in open flow systems, which require the flow to be introduced and removed from the observational domain. The Dean problem represents a typical open flow system and can serve as a candidate to address this question by examining the dynamic responses of multiple solutions.

To address these issues, we have made a relatively comprehensive study on the bifurcation structure and stability of multiple solutions for the laminar forced convection in a curved duct of square cross-section (Fig. 1) [6]. The governing differential equations in primitive variables and in stream function and vorticity are solved for detailed bifurcation structure by a finite-volume/Euler–Newton continuation method with the help of the bifurcation test function, the branch switching technique and the parameterization of arc-length or local variable. Transient calculation is made to examine in detail the response of every solution family to finite random disturbances. Eigenvalue systems are solved for all eigenvalues by QR decomposition method to conclusively determine the linear stability of the multiple solutions. The Hopf bifurcation can also be detected by the transient computation and by the appearance of a pair of pure imaginary conjugated eigenvalues in the linear stability analysis. The power spectra are constructed by the Fourier transformation of temporal oscillation solutions to confirm the chaotic flow at large Dean numbers. We restrict ourselves to the hydrodynamically and thermally fully-developed region and two-dimensional disturbances. So far, a detailed 3D numerical computation of flow bifurcation and stability is still too costly to conduct. A 2D model is still useful for the fundamental understanding of curved duct flows. However, our assumption of fully developed flow limits our analysis to the one preserving the streamwise symmetry. There may be further bifurcation to flows that break this symmetry and that cannot be found by our 2D model.

In addition to the three solution branches in [38,39], three new asymmetric branches are found from three symmetry-breaking bifurcation points on the isolated symmetric branch S_2. Only symmetric 2-cell flows on the primary branch and on some part of isolated 2-cell branch are linearly stable. The linear stability of the isolated 2-cell branch changes along

the branch. In the range where there is no stable steady fully developed solution, flows develop into a temporal periodic oscillation between symmetric/asymmetric 2-cell flows and symmetric/asymmetric 4-cell flows. As the Dean number increases, physically realizable fully-developed flows will evolute from a stable steady 2-cell at a lower Dean number to a temporal periodic oscillation, another stable steady 2-cell state, an intermittent oscillation, and a chaotic state. Among them, three states of temporal oscillation have not been reported in the literature. Furthermore, the intermittency is found to lead curved channel flows to chaos. The period-doubling bifurcation is, on the other hand, suppressed by the pressure-driven flows in the Dean problem, a typical open flow system. The temporal periodic oscillation serves as the transition from one stable steady flow to another.

Since rotation can be viewed as a perturbation for the non-rotating cases and mixed convection as a perturbation for the forced convection, the flow and temperature fields presented here also serve as starting points of our continuation algorithms for the multiplicity and stability of mixed convection in rotating curved ducts [6].

III. Governing Equations

Consideration is given to a hydrodynamically and thermally fully developed laminar flow of viscous fluid in a curved square duct under the thermal boundary conditions of uniform wall heat flux and peripherally uniform wall temperature at any streamwise position (Fig. 1). The geometry is toroidal and hence the finite pitch effect is not considered. Properties of the fluid, density in particular, are taken to be constant. Therefore, a gravity potential can be introduced for the purely hydrostatic effect of gravity [52].

Consider a toroidal coordinate system (R, Z, ϕ) as shown in Fig. 1. Let U, V, and W be velocity components in directions of R, Z, and ϕ respectively, t the time, and T, T_w the temperatures of the fluid and the wall. Continuity, Navier–Stokes and energy equations governing the fully developed laminar flow and heat transfer are given, in terms of dimensionless variables, as [30,53–55]:

Continuity equation:

$$\frac{\partial}{\partial r}\{[1+\sigma(r-0.5)]u\}+\frac{\partial}{\partial z}\{[1+\sigma(r-0.5)]v\}=0. \tag{1}$$

Momentum equations:

$$\frac{\partial u}{\partial \tau}+u\frac{\partial u}{\partial r}+v\frac{\partial u}{\partial z}-\frac{16w^2Dk^2}{\sigma[1+\sigma(r-0.5)]}$$
$$=-\frac{\partial p}{\partial r}+\left\{\frac{\partial^2 u}{\partial r^2}+\frac{\partial^2 u}{\partial z^2}+\frac{\sigma}{1+\sigma(r-0.5)}\frac{\partial u}{\partial r}-\frac{\sigma^2 u}{[1+\sigma(r-0.5)]^2}\right\}, \quad (2)$$

$$\frac{\partial v}{\partial \tau}+u\frac{\partial v}{\partial r}+v\frac{\partial v}{\partial z}=-\frac{\partial p}{\partial z}+\left[\frac{\partial^2 v}{\partial r^2}+\frac{\partial^2 v}{\partial z^2}+\frac{\sigma}{1+\sigma(r-0.5)}\frac{\partial v}{\partial r}\right], \quad (3)$$

$$\frac{\partial w}{\partial \tau}+u\frac{\partial w}{\partial r}+v\frac{\partial w}{\partial z}+\frac{\sigma uw}{1+\sigma(r-0.5)}$$
$$=\frac{1}{1+\sigma(r-0.5)}+\left\{\frac{\partial^2 w}{\partial r^2}+\frac{\partial^2 w}{\partial z^2}+\frac{\sigma}{1+\sigma(r-0.5)}\frac{\partial w}{\partial r}\right.$$
$$\left.-\frac{\sigma^2 w}{[1+\sigma(r-0.5)]^2}\right\}. \quad (4)$$

Energy equation:

$$\frac{\partial \theta}{\partial \tau}+u\frac{\partial \theta}{\partial r}+v\frac{\partial \theta}{\partial z}-\frac{4wDk}{\sigma Pr[1+\sigma(r-0.5)]}$$
$$=\frac{1}{Pr}\left[\frac{\partial^2 \theta}{\partial r^2}+\frac{\partial^2 \theta}{\partial z^2}+\frac{\sigma}{1+\sigma(r-0.5)}\frac{\partial \theta}{\partial r}\right]. \quad (5)$$

The dimensionless variables are defined as:

$$r=\frac{R}{a}, \quad z=\frac{Z}{a}, \quad \tau=\frac{t}{\nu/a^2},$$
$$u=\frac{aU}{\nu}, \quad v=\frac{aV}{\nu}, \quad w=\frac{W}{W_1},$$
$$p=\frac{P}{\rho(\nu/a)^2}, \quad \theta=\frac{T_w-T}{\Delta T},$$

where ν and ρ are the kinematic viscosity and the density of the fluid; a is the duct radial dimension, P is a pseudo-pressure, a combination of fluid pressure and the gravity potential; W_1 and ΔT are the representative streamwise velocity and temperature difference, respectively, which are

defined as

$$W_1 = \frac{a^2 c_1}{\mu}; \quad \Delta T = a c_2 Pr.$$

Here μ is the viscosity of the fluid; Pr the Prandtl number; c_1 the streamwise pressure gradient which is a positive constant for hydrodynamically fully developed flow ($c_1 = -\partial P / R_c \partial \phi$ with R_c as the curvature radius) [55]; c_2 the streamwise temperature gradient which is a constant for the thermally fully developed flow, but can be positive or negative depending on the heating or cooling of the fluid ($c_2 = \partial T / R_c \partial \phi$) [55,56].

It is customary to use the mean streamwise velocity W_m and the difference between the wall temperature and the bulk mean temperature ($T_w - T_b$) for the non-dimensionalization of streamwise velocity and temperature, respectively. However, the employment of these quantities results, unavoidably, in the appearance of two unknown dimensional parameters in the governing equations, which comprise the unknowns W_m and T_b, respectively. Consequently, the iterative procedure should be applied, assuming some initial estimated values to them. It requires an additional computation time. In order to avoid this additional increase in computation time, we follow [53–55] in using W_1 and ΔT as the representative streamwise velocity and the representative temperature difference, respectively. They involve the streamwise pressure gradient c_1 and the streamwise temperature gradient c_2, which are usually given as design parameters so that it does not induce the difficulty in using computation results for design. The velocity W_1 is proportional to the pressure drop in the streamwise direction. For the flow in a stationary straight circular tube, the mean streamwise velocity W_m is related to W_1 as $W_m = W_1/8$ [54,55]. The temperature difference ΔT is, on the other hand, proportional to the fluid temperature difference between the inlet and outlet of the duct.

Three dimensionless parameters are defined as:

$$\sigma = \frac{a}{R_c}, \quad Pr = \frac{\nu}{\alpha}, \quad Dk = \frac{\sigma a W_1}{4\nu},$$

with α as the thermal diffusivity. The dimensionless groups adopted here are those in Refs. [53–55]. The curvature ratio σ is a geometry parameter, representing the degree of curvature. The Prandtl number Pr, a thermophysical property parameter, represents the ratio of momentum diffusion rate to that of thermal diffusion. Dk is a pseudo Dean number with W_1 as the characteristic velocity [53–55].

Boundary conditions (non-slip, impermeability, and uniform peripheral temperature) may be written, in terms of dimensionless variables, as

$$u = v = w = \theta = 0 \quad \text{at } r = 0, 1 \quad \text{for } -0.5 \le z \le 0.5, \tag{6}$$

$$u = v = w = \theta = 0 \quad \text{at } z = -0.5, 0.5 \quad \text{for } 0 \le r \le 1. \tag{7}$$

Within the scope of the present study, Eqs. (1)–(5) under the boundary conditions (6)–(7) constitute the mathematical model of the problem under consideration in forms of primitive variables u, v, w, p, and θ. We use them to examine the steady bifurcation structure by removing all time-dependent terms. We also use them to study stability properties of multiple solutions through a direct transient computation examining dynamic responses of every steady solution family to finite random disturbances. To verify the bifurcation structure obtained and save some computer memory demand by the linear stability analysis, we also use a stream function/vorticity (ψ/ω) version of (1)–(7) to obtain the steady bifurcation structure and to perform the linear stability analysis. By introducing the stream function ψ of secondary flow satisfying the continuity equation (1)

$$u = \frac{1}{1 + \sigma(r - 0.5)} \frac{\partial \psi}{\partial z}, \quad v = -\frac{1}{1 + \sigma(r - 0.5)} \frac{\partial \psi}{\partial r}, \tag{8}$$

and the vorticity ω

$$\omega = \frac{\partial u}{\partial z} - \frac{\partial v}{\partial r} = \frac{1}{1 + \sigma(r - 0.5)} \left(\frac{\partial^2 \psi}{\partial z^2} + \frac{\partial^2 \psi}{\partial r^2} - \frac{\sigma}{1 + \sigma(r - 0.5)} \frac{\partial \psi}{\partial r} \right), \tag{9}$$

we obtain ψ/ω version of (1)–(7)

$$\begin{aligned} &\frac{\partial \omega}{\partial \tau} + \frac{1}{1 + \sigma(r - 0.5)} \frac{\partial(\psi, \omega)}{\partial(z, r)} \\ &\quad = \frac{\partial^2 \omega}{\partial r^2} + \frac{\partial^2 \omega}{\partial z^2} + \frac{\sigma}{1 + \sigma(r - 0.5)} \frac{\partial \omega}{\partial r} - \frac{\sigma^2}{[1 + \sigma(r - 0.5)]^2} \omega \\ &\qquad + \frac{\sigma}{[1 + \sigma(r - 0.5)]^2} \omega \frac{\partial \psi}{\partial z} + \frac{32 w D k^2}{\sigma[1 + \sigma(r - 0.5)]} \frac{\partial w}{\partial z}, \end{aligned} \tag{10}$$

$$\frac{\partial w}{\partial \tau}+\frac{1}{1+\sigma(r-0.5)}\frac{\partial(\psi, w)}{\partial(z, r)}$$
$$=\frac{\partial^2 w}{\partial r^2}+\frac{\partial^2 w}{\partial z^2}+\frac{\sigma}{1+\sigma(r-0.5)}\frac{\partial w}{\partial r}+\frac{1}{1+\sigma(r-0.5)}$$
$$-\frac{\sigma^2}{[1+\sigma(r-0.5)]^2}w-\frac{\sigma}{[1+\sigma(r-0.5)]^2}w\frac{\partial \psi}{\partial z}, \tag{11}$$

$$\frac{\partial \theta}{\partial \tau}+\frac{1}{1+\sigma(r-0.5)}\frac{\partial(\psi, \theta)}{\partial(z, r)}$$
$$=\frac{1}{Pr}\left(\frac{\partial^2 \theta}{\partial r^2}+\frac{\partial^2 \theta}{\partial z^2}\right)+\frac{1}{Pr}\left(\frac{\sigma}{1+\sigma(r-0.5)}\frac{\partial \theta}{\partial r}+\frac{4wDk}{\sigma[1+\sigma(r-0.5)]}\right), \tag{12}$$

$$\psi=\frac{\partial \psi}{\partial r}=\frac{\partial \psi}{\partial z}=w=\theta=0 \quad \text{at } r=0, 1 \quad \text{for } -0.5\le z\le 0.5, \tag{13}$$

$$\psi=\frac{\partial \psi}{\partial r}=\frac{\partial \psi}{\partial z}=w=\theta=0 \quad \text{at } z=-0.5, 0.5 \quad \text{for } 0\le r\le 1. \tag{14}$$

After the velocity and temperature fields are obtained, the computation of the local friction factor and Nusselt number is of practical interest. Following the usual definitions, the expression for the product of friction factor and Reynolds number fRe and Nusselt number Nu can be written based on the local streamwise velocity gradient and temperature gradient at the wall as [53],

$$fRe=\frac{2}{w_m}\left(\frac{\partial w}{\partial n}\right)_{\text{wall}}, \tag{15}$$

$$Nu=\frac{1}{\theta_b}\left(\frac{\partial \theta}{\partial n}\right)_{\text{wall}}, \tag{16}$$

where w_m and θ_b are the mean streamwise velocity and the bulk mean temperature, respectively. The mean friction factor and Nusselt number can be obtained by peripherally averaging their local values.

IV. Numerical Algorithm

A. Bifurcation Structure

For steady bifurcation structure, we remove the time-dependent terms in Eqs. (2)–(5). The governing equations (1)–(5) are then discretized under

the boundary conditions (6) and (7) by the finite volume method which is an adaptation of that in Refs. [53,57,58]. Main features of this method include a staggered mesh system, a power-law formulation for the combined effect of convection and diffusion terms, and central difference scheme for source terms. After this discretization in the flow domain, we obtain a set of non-linear algebraic equations, the so-called discretization equations, which approximate the governing differential equations (1)–(5) under the boundary conditions (6) and (7) and can be written symbolically as

$$\mathbf{f}(\mathbf{y}, \alpha) = 0, \tag{17}$$

where $\alpha = [\sigma, Pr, Dk]$ (the parameter vector), $\mathbf{y}$ is the dependent variable vector of size N with N as the total number of equations, and $\mathbf{f}$ is a vector-valued function of size N.

Among the three parameters, the curvature ratio σ is a more detailed measure of the effect of geometry and the extent to which the centrifugal force varies on the cross section. Winters [38] found that it does not contribute to the structural change of bifurcation while higher curvature ratios (larger than 0.2) shift limit and bifurcation points to higher Dean numbers. This is consistent with Dean's theoretical finding that the Dean problem is governed by the Dean number alone for the loosely coiled ducts [2,5]. Therefore, we fix its value at 0.02, typically used in cooling systems of rotor drums and conductors of electrical generators. Readers are referred to [59,60] for the multiplicity and stability of forced convection in tightly coiled ducts of square cross-section ($\sigma = 0.3$). While the Prandtl number Pr affects the temperature field, it also has no effect on flows because the flow field is decoupled with the temperature field for the forced convection. In the present work, we set the Pr at 0.7, a typical value for air.

Equation (17) has a solution $\mathbf{y} = 0$ when $Dk = 0$. This is physically obvious and can be confirmed by a dimensional version of (1)–(7) [55]. This point serves as the starting point of our continuation schemes in tracing solution branches as the control parameter λ varies. We use Dk, as the control parameter, for the regular portion of solution branches, u at $r = 0.9$ and $z = 0.14$ (Dk as one unknown variable) or the arc-length s for turning limit points [47,61]. The u at (0.9, 0.14) is selected as the control parameter λ in turning limit points because it varies significantly as Dk changes. For the case of using s as the control parameter λ in turning limit points in the arc-length continuation, we need an additional algebraic equation:

$$\sum_{j=1}^{N} \left.\frac{\partial y_j}{\partial s}\right|_{s_0} \left[y_j(s) - y_j(s_0)\right] + \left.\frac{\partial Dk}{\partial s}\right|_{s_0} \left[Dk(s) - Dk(s_0)\right] - (s - s_0) = 0, \tag{18}$$

in addition to Eq. (17) which now takes the form of

$$\mathbf{f}(\mathbf{y}(s), D_k(s)) = 0. \tag{19}$$

Here, $s = \|(\mathbf{y}, D_k)^T\|_2$, $s - s_0$ is the step size along the branch and y_j is a component of $\mathbf{y}$. Equation (18) is obtained by differentiating the defining equation of s

$$\sum_{j=1}^{N} \left[y_j(s) - y_j(s_0)\right]^2 + [Dk(s) - Dk(s_0)]^2 = (s - s_0)^2$$

with respect to s.

Equation (17) [or Eqs. (18) and (19)] is solved for the whole branch solution by Euler-Newton method for all unknown dependent variables simultaneously. Let $\mathbf{y}(\lambda_i)$ be the solution at $\lambda = \lambda_i$ which has been obtained, the Euler predictor for $\mathbf{y}(\lambda_i + \Delta\lambda)$ is, where $\Delta\lambda$ is the step size,

$$\mathbf{y}^0(\lambda_i + \Delta\lambda) = \mathbf{y}(\lambda_i) + \left.\frac{\partial \mathbf{y}}{\partial \lambda}\right|_{\lambda_i} \Delta\lambda, \tag{20}$$

where $(\partial \mathbf{y}/\partial\lambda)|_{\lambda_i}$ is the solution of the system of linear algebraic equations

$$\mathbf{J}|_{\mathbf{y}(\lambda_i),\lambda_i} \left.\frac{\partial \mathbf{y}}{\partial \lambda}\right|_{\lambda_i} = -\left.\frac{\partial \mathbf{f}}{\partial \lambda}\right|_{\lambda_i}. \tag{21}$$

Here $\mathbf{J}$ with $J_{ij} = (\partial f_i/\partial y_j)$ as components is the Jacobian matrix, and is evaluated numerically by the forward difference. Linear equation (21) is then solved by the SPARSPAK (the Waterloo sparse matrix package), which is a collection of FORTRAN subroutines for solving large sparse systems of linear equations [62]. The basic approach used in the SPARSPAK is to compute a triangular factorization of the coefficient matrix using Gaussian elimination without any row interchanges.

Using $\mathbf{y}^0(\lambda_i + \Delta\lambda)$ from the Euler-predictor [Eq. (20)] as the initial guess of $\mathbf{y}(\lambda_i + \Delta\lambda)$, the damped Newton iteration of Eq. (17) [or Eqs. (18) and (19)], in which the Jacobian matrix $\mathbf{J}$ is updated once for every four iterations, leads to:

$$\boldsymbol{J}\left[\mathbf{y}^{k+1}\left(\lambda_i + \lambda\right) - \mathbf{y}^k\left(\lambda_i + \lambda\right)\right] = \mathbf{f}^k\left[\mathbf{y}^k\left(\lambda_i + \lambda\right), \lambda_i + \lambda\right], \quad k = 1, 2, \ldots, \tag{22}$$

where k denotes the kth iteration. Equation (22) is solved for $\mathbf{y}^{k+1}(\lambda_i+\Delta\lambda)$ by the SPARSPAK. The solution was assumed to be convergent in a numerical sense if the Euclidean norm of error in each of the primitive variables (i.e., velocity components, temperature, and pressure) over all grid points is less than 10^{-10} between successive iterations.

In order to detect bifurcation points, a test function τ_{ij} defined by [47,63]:

$$\tau_{\mathrm{ij}} = \mathbf{e}_i^T \mathbf{J}\mathbf{h} \tag{23}$$

is evaluated at every continuation step after the Newton iteration is convergent. Here, $\mathbf{e}_i$ is a unit column vector of size N and the column vector $\mathbf{h}$ of size N satisfies

$$\mathbf{J}_{ij}\mathbf{h} = \mathbf{e}_i, \tag{24}$$

where the matrix $\mathbf{J}_{ij}$ is a reduced $\mathbf{J}$ by replacing its ith row by a unit vector, i.e.,

$$\mathbf{J}_{ij} = \left[\left(\mathbf{I} - \mathbf{e}_i\mathbf{e}_i^T\right)\boldsymbol{J} + \mathbf{e}_i\mathbf{e}_j^{\boldsymbol{T}}\right]. \tag{25}$$

Equation (24) guarantees that $\mathbf{h}$ is a non-trial solution and is solved by the SPARSPAK. When the system is at the bifurcation point (i.e., singular $\mathbf{J}$), τ_{ij} vanishes. The values of i and j can be any integer less than N. For our problem, we take the grid indexes of point (0.9, 0.14) as their values. This choice works well as the flow at this point changes sensitively with Dk.

Branch switching refers to the calculation of at least one solution on the emanating branch arising at the bifurcation point. This first solution could serve as the starting point for a subsequent tracing of the entire branch. Note that at least two solutions $[\mathbf{y}(\lambda_1), \lambda_1]$ and $[\mathbf{y}(\lambda_2), \lambda_2]$ with opposite sign of the test function τ_{ij} are available close to the bifurcation in practice. We use the branch switching technique developed by Seydel [47,64] which approximates the difference between the branches by

$$\delta_0 \max\{1, |y_k(\lambda^*)|\}\mathbf{h}.$$

Here, δ_0 is an empirical variable. Seydel [47] found that $\delta_0 = 0.02$ is successful for a wide range of practical problems. For our problem, a value in the range 0.01 to 0.04 enables us to switch branches successfully. y_k is the kth component of $\mathbf{y}$ with the index k from 1 up to N. λ^* is either λ_1 or λ_2. $\mathbf{h}$ is a column vector of size N defined and determined by Eq. (24) with

$\mathbf{J}_{ij}$ evaluated at λ^* and $\mathbf{y}(\lambda^*)$. Using

$$\mathbf{y}^0(\lambda^*) = \mathbf{y}(\lambda^*) \pm \delta_0 \max\{1, |y_k(\lambda^*)|\}\mathbf{h} \tag{26}$$

as the initial guess, the damped Newton iteration of Eq. (17) [or Eqs. (18) and (19)]

$$\mathbf{J}|_{\mathbf{y}^k(\lambda^*),\lambda}\left[\mathbf{y}^{k+1}(\lambda^*) - \mathbf{y}^k(\lambda^*)\right] = \mathbf{f}^k\left[\mathbf{y}^k(\lambda^*), \lambda^*\right], \quad k = 1, 2, \ldots, \tag{27}$$

could lead to a *first solution* on the emanating branch.

By using the ψ/ω-version of Eqs. (10)–(14) to compute the bifurcation structure, we remove time-dependent terms in Eqs. (10)–(12). The governing equations (10)–(12) are then discretized under the boundary conditions (13)–(14) by the finite volume method. The boundary conditions for the vorticity are obtained by the boundary wall formula [65]. All terms are approximated with the central difference scheme. The power-law formulation and the staggered mesh system are found unnecessary here because of the elliptic nature of ψ and ω and the disappearance of pressure terms in Eqs. (10)–(14) [57]. After this discretization in the flow domain, we obtain a set of discretization equations which can also be written symbolically as Eq. (17). The algorithm in seeking its solution for whole bifurcation structure of solutions is the same as that for the primitive variables.

B. Linear Stability

To save the computer-memory demand, we use the ψ/ω-version of governing equations for the linear stability analysis. After discretizing all terms except the time-dependent ones by the finite volume method mentioned above, we obtain a set of differential/algebraic equations which can be symbolically written as

$$\mathbf{M}(\mathbf{y})\frac{d\mathbf{y}}{d\tau} = \mathbf{f}(\mathbf{y}, \alpha), \tag{28}$$

where $\mathbf{M}(\mathbf{y})$ is the mass matrix of size N. For any steady solution $\mathbf{y}_s$ found by the Newton method during continuation, the linear stability analysis can be made by substituting $\mathbf{y} = \mathbf{y}_s + \boldsymbol{\varepsilon}\mathbf{e}^{\boldsymbol{\beta}\tau}$ into the above equation, where $\boldsymbol{\varepsilon}$ is a small disturbance vector of size N and $\boldsymbol{\beta}$ the eigenvalue vector of size N. After the linear terms of ε are collected, a generalized eigenvalue problem is obtained:

$$\mathbf{M}^{-1}(\mathbf{y}_s)\mathbf{J}|_{\mathbf{y}_s} - \boldsymbol{\beta}\mathbf{I} = 0, \tag{29}$$

where $\mathbf{I}$ is the identity matrix. Apparently, the Jacobian $\mathbf{J}$ obtained after Newton iterations can be used here directly. $\mathbf{M}^{-1}(\mathbf{y}_s)$ is obtained by the Gauss–Jordan method. After $\mathbf{M}^{-1}(\mathbf{y}_s)\mathbf{J}|_{\mathbf{y}_s}$ is transformed into its Hessenberg form by the elementary similarity transformation, we solve Eq. (29) for all eigenvalues by the QR decomposition method.

The linear stability of $\mathbf{y}_s$ can then be ascertained by the sign of real parts of all eigenvalues. If all real parts are negative, $\mathbf{y}_s$ is linearly stable. If there is any positive real part, the steady solution $\mathbf{y}_s$ is not stable to an infinitesimal disturbance. Hopf bifurcation can also be detected by the appearance of a pair of pure imaginary conjugated eigenvalues.

C. Dynamic Responses to Finite Random Disturbances

For transient computation aiming for the response of multiple steady solutions to the finite two-dimensional random disturbances, we obtain the discretization equations by integrating the governing equations with the time-dependent terms over every control volume *and* over the time period from τ to $\tau+\Delta\tau$ (the finite volume method). The fully implicit method is used because of its superior numerical stability. The system of discretization equations is then solved by the Euler–Damped Newton method described above by viewing time τ as the continuation parameter. The initial condition at $\tau=0$, which also serves as the starting point of the continuation scheme, is formed by the steady solution $\mathbf{y}_s(Dk)$ plus a finite random disturbance. Here, the subscript s denotes the steady solution. The random disturbance is generated by $d^{(k)}\boldsymbol{\chi}^{(k)}\mathbf{y}_s(Dk)$. Here d is the maximum percentage of disturbing value over the steady value $\mathbf{y}_s$. The superscript k represents the ordinal of the disturbance. $\boldsymbol{\chi}$ is a vector of size N whose components take random values from -1 to 1 and are generated by the computer. To examine dynamic responses of a steady solution to different finite random disturbances, we normally generate three sets of disturbances denoted by $k=1$, 2, and 3, with $d=4\%$, 10%, and 15% respectively.

D. Grid-Dependence Check

With the recognition that at a point in the parameter space, several different solutions with quite different flow structures can coexist, a uniform grid that is fine enough to resolve all the different flow structures appears proper. In the present computations, the grid dependence was checked by three pairs of grid sizes uniformly distributed in the flow domain. They are 40×40, 50×50, and 60×60. The pair of numbers $(L\times K)$ represent the number of grid points used in r- and z-directions, respectively. The bifurcation diagrams obtained by using these three grid sizes are shown in

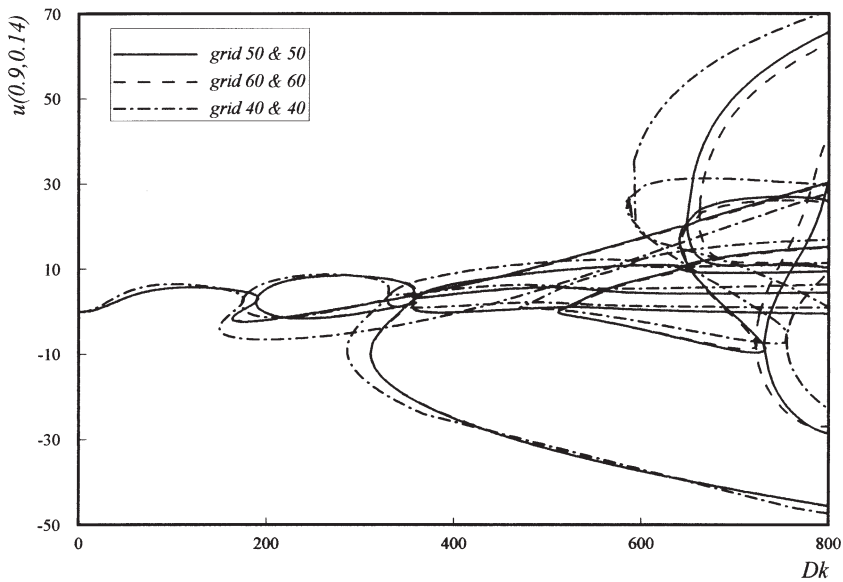


FIG. 3. Effect of grid size on bifurcation structure ($\sigma = 0.02$ and $Pr = 0.7$).

Fig. 3 with the u velocity component at (0.9, 0.14) as the state variable and Dk as the parameter at $\sigma = 0.02$ and $Pr = 0.7$. It shows that the bifurcation diagram has no qualitative change for all three pairs of grid sizes. The quantitative change is also very small for most of interested region of Dk from 50×50 to 60×60. The numerical data are shown in Table I for eight cases at $Dk = 20$, 100, 300, 450, 550, 700, and 800 (all at the primary solution branch S_1), respectively, with $\sigma = 0.02$, and $Pr = 0.7$. Four representative properties, namely, Dean number ($De = Re\sqrt{\sigma}$, $Re = W_m a/\nu$ with W_m as the streamwise mean velocity), maximum of absolute values of secondary flow stream function ($|\psi|_{\max}$), maximum streamwise velocity ($w_{\max}$) and maximum temperature ($\theta_{\max}$), as well as the CPU time for one continuation step along S_1, are listed in Table I for comparison. The computations were carried out on the DIGITAL Personal Workstation AU 600. The general trend of these results as the grid size is decreased tends to indicate that the solutions for the case of (50×50) grids are accurate to within 1% tolerance. We also checked the detailed variations of flow and temperature fields on various solution branches for different grid sizes, and found that 50×50 is indeed a reasonably accurate choice for the grid size for square ducts, a result being in agreement with that in Ref. [53]. It is worth noting that the CPU time increases rapidly as the grid spacing

TABLE I

VARIATIONS OF De, $|\psi|_{\max}$, $w_{\max}$, $\theta_{\max}$ AND CPU TIME IN SECONDS WITH DIFFERENT GRIDS ON PRIMARY SOLUTION BRANCH S_1

Dk	Grids	$\sigma = 0.02$ $Pr = 0.7$				CPU (s)
		De	$\lvert\psi\rvert_{\max}$	$w_{\max}$	$\theta_{\max}$	
20	40×40	19.5	0.880	0.07150	15.5	73.5
	50×50	19.5	0.885	0.07153	15.5	155.1
	60×60	19.5	0.887	0.07154	15.5	301.3
100	40×40	76.5	5.488	0.04938	39.6	72.2
	50×50	76.5	5.528	0.04939	39.7	157.1
	60×60	76.5	5.546	0.04945	39.7	296.4
300	40×40	173.3	8.332	0.03405	59.8	73.1
	50×50	173.6	8.386	0.03413	60.2	156.3
	60×60	173.8	8.426	0.03419	60.3	302.0
450	40×40	234.4	10.733	0.03024	71.5	71.7
	50×50	235.0	10.951	0.03037	72.1	154.7
	60×60	235.4	10.909	0.03045	72.4	301.1
550	40×40	272.8	12.095	0.02867	78.7	72.4
	50×50	273.8	12.203	0.02884	79.4	158.2
	60×60	274.3	12.246	0.02891	79.9	310.7
700	40×40	327.8	13.747	0.02693	88.4	72.9
	50×50	329.2	13.836	0.02710	89.3	155.9
	60×60	330.1	13.942	0.02718	89.9	306.4
800	40×40	362.8	14.622	0.02597	94.1	73.1
	50×50	364.6	14.787	0.02615	95.2	156.4
	60×60	365.6	14.857	0.02625	95.8	291.5

decreases. In order to have a balance between the cost of the computer time and the accuracy of the solution, we carried out all the computations with a 50×50 uniform mesh for square ducts.

E. ACCURACY CHECK

To verify the code, five representative properties obtained by the present work are shown in Table II together with those in Ref. [53] at $Dk = 100$, $\sigma = 0.02$ and $Pr = 0.7$, where there is only one solution. The results are in good agreement, with a very small difference (less than 2%) being due to the different numerical methods used in the two studies. We also compared, in Table III, the locations (in terms of their De values) of three limit points S_1^1, S_1^2, S_2^1 and one bifurcation point B_1 with those available in Ref. [38] (see next section for the notation of limit points). The results of

TABLE II

COMPARISON OF FIVE REPRESENTATIVE PROPERTIES AT $Dk = 100$, $\sigma = 0.02$, $Pr = 0.7$ WITH THOSE IN REF. [53]

Sources	Re	De	$\lvert\psi\rvert_{\max}$	$w_{\max}$	$\theta_{\max}$
Present work	541	76.5	5.528	0.0494	39.7
Wang and Cheng (1996) [53]	542	76.6	5.641	0.0496	39.9

TABLE III

LOCATIONS OF S_1^1, S_1^2, S_2^1 AND B_1: A COMPARISON WITH THOSE IN REF. [38]

	S_1^1	S_1^2	S_2^1	B_1
Present work	129.55	112.82	187.91	128.22
Winters (1987) [38]	131.13	113.35	190.77	129.71

the present analysis are also in good agreement with those in Ref. [38]. The very small difference (less than 1.5%) is believed to be due to the different numerical methods used in the two studies.

To further verify the code and accuracy, we re-obtained the bifurcation structure by using our ψ/ω-version of the governing equation and a 50×50 uniform grid. The bifurcation structure by these two methods were found to be in excellent agreement for most of the range ($Dk \leq 700$), and have very slightly quantitative difference when $Dk > 700$. The bifurcation structure presented here is that from the formulation of primitive variables.

V. Results and Discussion

A. Solution Structure

The bifurcation structure is shown in Fig. 4 for Dk values from 0 up to 800 at $\sigma = 0.02$ and $Pr = 0.7$. In Fig. 4, the u velocity at (0.9, 0.14) (where the flow is sensitively dependent on Dk) is used as the state variable, enabling the most clear visualization of all solution branches. Six solution branches, labeled as S_1, S_2, A_1, A_2, A_3, and A_4 respectively, are found. Here, S stands for symmetric solutions with respect to the horizontal central plane $z = 0$ and A for asymmetric solutions. Branch A_1 is bifurcated from S_1 at the symmetry-breaking bifurcation point B_1. Branches A_2, A_3, and A_4

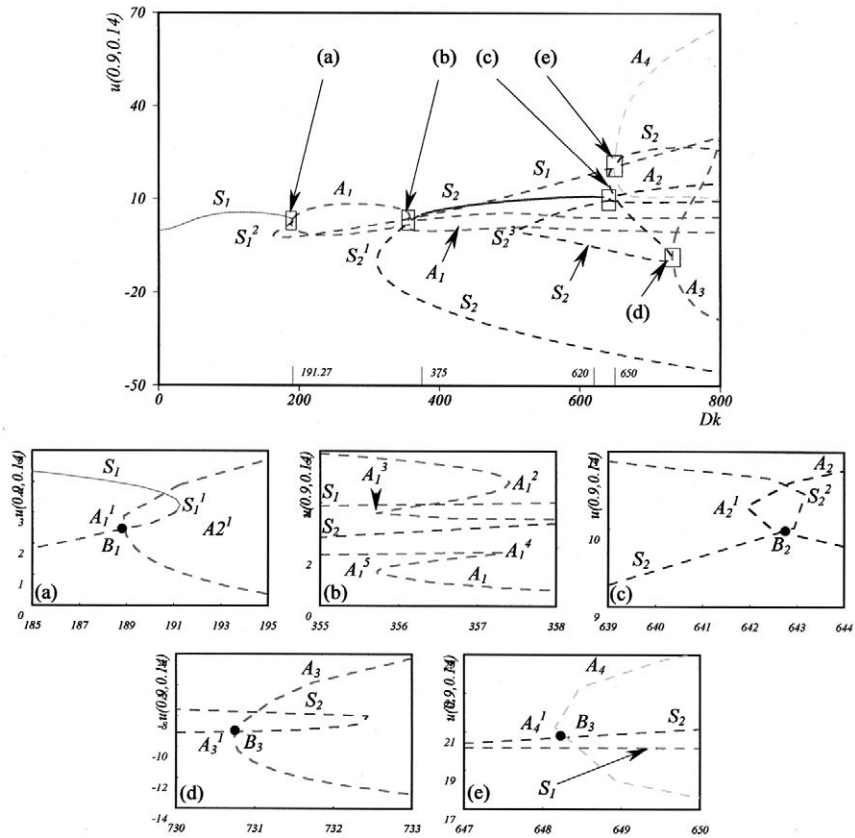


FIG. 4. Solution branches and their connectivity ($\sigma = 0.02$ and $Pr = 0.7$). (a) Connectivity between S_1 and A_1. (b) Four limit points on A_1. (c) Connectivity between S_2 and A_2. (d) Connectivity between S_2 and A_3. (e) Connectivity between S_2 and A_4.

are bifurcated from S_2 at three symmetry-breaking bifurcation points B_2, B_3, and B_4, respectively. Table IV lists Dk and De values of four symmetry-breaking bifurcation points B_1–B_4 and 15 limit points labeled according to their branch symbols with a superscript number. For example, S_1^2 represents the second limit point on the solution branch S_1. To visualize the details of branch connectivity and some limit points, the locally enlarged state diagrams are also shown in Fig. 4. As Fig. 4 is only a 1D projection of N dimensional solution branches, all intersecting points except four bifurcation points should not be interpreted as connection points of branches. In terms of Dk, the work by Winters [38] is limited to the range up to about 400. It is interesting to note that there is always an odd number of

TABLE IV

LOCATIONS OF ALL LIMIT POINTS AND BIFURCATION POINTS UP TO $Dk = 800$ AT $\sigma = 0.02$ AND $Pr = 0.7$

Points	Dk	De	Points	Dk	De
S_1^1	191.27	129.55	S_2^4	732.43	350.41
S_1^2	163.43	112.82	S_2^5	640.59	319.34
A_1^1	188.67	127.51	A_2^1	641.87	324.89
A_1^2	357.41	209.92	A_3^1	730.83	349.84
A_1^3	355.72	209.25	A_4^1	648.16	322.27
A_1^4	357.41	209.92	B_1	188.80	128.22
A_1^5	355.72	209.25	B_2	642.51	325.14
S_2^1	310.71	187.91	B_3	731.22	350.01
S_2^2	643.30	325.83	B_4	648.18	322.27
S_2^3	511.04	274.80			

co-existing steady solutions at each Dk from 0 to 800. Note also that each pair of singular points, (B_1, A_1^1), (B_2, A_2^1), (B_3, A_3^1), and (B_4, A_4^1) are very close. Whether each pair represents just a single point of higher nullity in the continuous problem and whether their slight separation is an artifact of the numerical discretization cannot be ascertained at present.

Among the six solution branches, S_1, S_2, and A_1 were first reported by Winters [38]. We here reconfirm and extend their findings up to $Dk = 800$. While no new limit and bifurcation points are found along S_1 and A_1, four additional limit points $S_2^2 - S_2^5$ (first reported in Ref. [39]) and three new symmetry-breaking bifurcation points B_2, B_3, and B_4 are detected along S_2. In particular, three symmetry-breaking bifurcation points B_2, B_3, and B_4 lead to three pairs of asymmetric solution branches A_2, A_3, and A_4.

1. *S_1 branch*

The primary branch S_1 is a symmetric solution branch and has two limit points S_1^1 and S_1^2. The two limit points divide the branch into three parts, S_{1-1}, S_{1-2} and S_{1-3} [Fig. 5(a)], and generate a range ($163.43 < Dk < 191.27$) where three steady solutions co-exist for a fixed value of Dk [Fig. 5(c)–(e)]. The secondary flows for these three sub-branches are the 2-cell state [two Ekman vortices, Fig. 5(b) and (c)], the weak 4-cell state with two Ekman vortices and two weak Dean vortices [Fig. 5(d)], and the 4-cell state with two Ekman vortices and two Dean vortices [Fig. 5(e)–(j)], respectively. Figures 5(b) to (j) detail the evolution of flow and temperature fields along this branch as Dk varies. In the figures, the stream function, streamwise velocity and temperature are normalized by their corresponding maximum absolute values $|\psi|_{\max}$, $w_{\max}$, and $\theta_{\max}$. A vortex with a positive (negative)

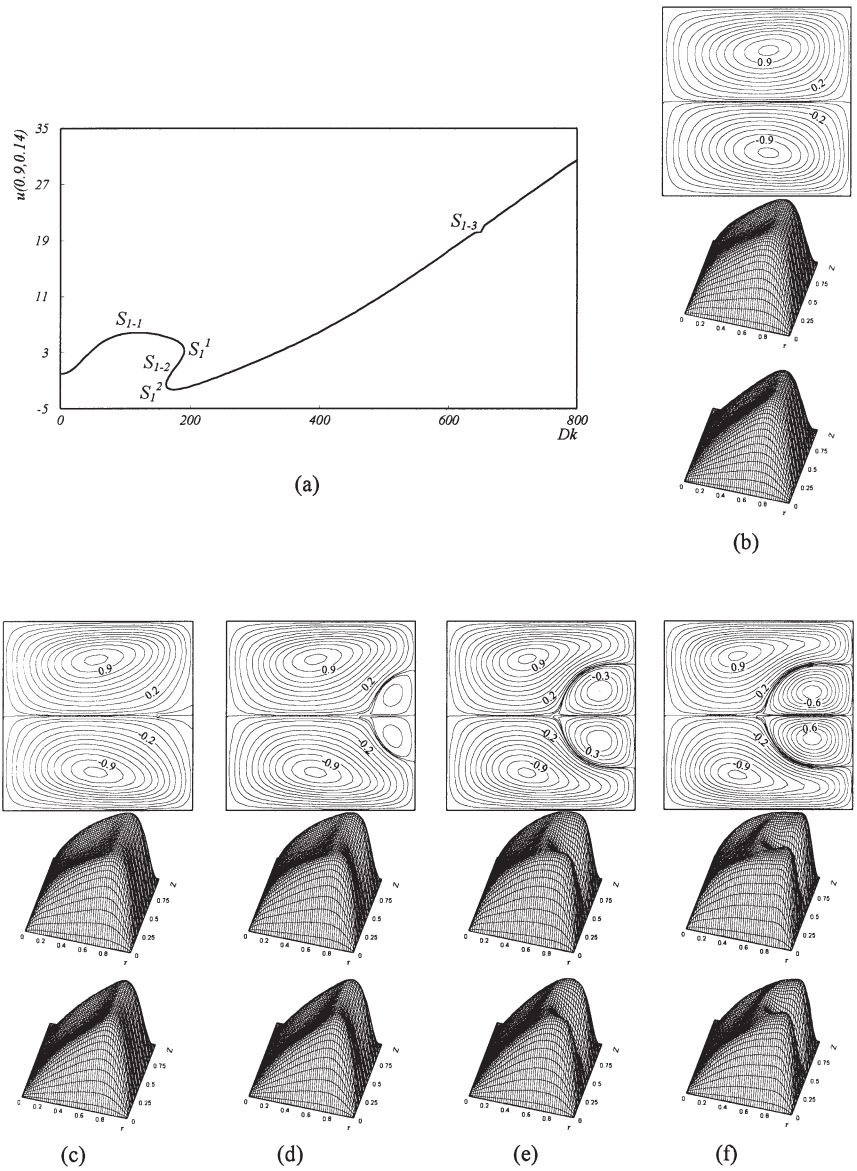


FIG. 5. Primary branch S_1: its limit points, flow and temperature fields ($\sigma = 0.02$ and $Pr = 0.7$). (a) Primary branch S_1 and its limit points. (b) $Dk = 100$ on S_{1-1}; $De = 76.5$, $|\psi|_{\max} = 5.53$, $w_{\max} = 0.04939$, $\theta_{\max} = 39.7$. (c) $Dk = 180$ on S_{1-1}; $De = 123.4$, $|\psi|_{\max} = 7.44$, $w_{\max} = 0.04288$, $\theta_{\max} = 53.8$. (d) $Dk = 180$ on S_{1-2}; $De = 123.3$, $|\psi|_{\max} = 7.06$, $w_{\max} = 0.04288$, $\theta_{\max} = 54.0$. (e) $Dk = 180$ on S_{1-3}; $De = 119.7$, $|\psi|_{\max} = 6.20$, $w_{\max} = 0.04048$, $\theta_{\max} = 50.0$. (f) $Dk = 300$ on S_{1-3}; $De = 173.6$, $|\psi|_{\max} = 8.39$, $w_{\max} = 0.03413$, $\theta_{\max} = 60.2$.

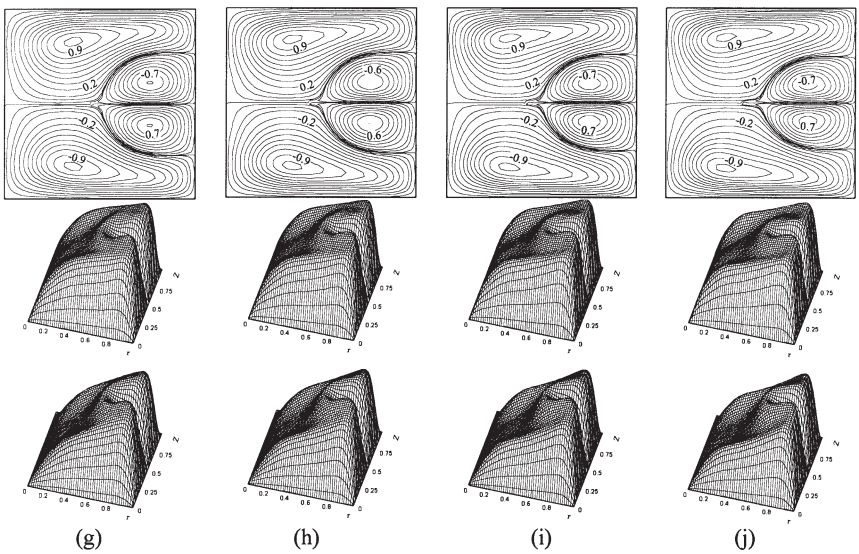


FIG. 5. (Continued) (g) $Dk = 450$ on S_{1-3}; $De = 235.0$, $|\psi|_{\max} = 10.86$, $w_{\max} = 0.03037$, $\theta_{\max} = 72.1$. (h) $Dk = 550$ on S_{1-3}; $De = 273.8$, $|\psi|_{\max} = 12.20$, $w_{\max} = 0.02884$, $\theta_{\max} = 79.4$. (i) $Dk = 700$ on S_{1-3}; $|\psi|_{\max} = 13.84$, $De = 329.2$, $w_{\max} = 0.02710$, $\theta_{\max} = 89.3$. (j) $Dk = 800$ on S_{1-3}; $De = 364.6$, $|\psi|_{\max} = 14.79$, $w_{\max} = 0.02615$, $\theta_{\max} = 95.2$. [Figures (b) to (j) are arranged as follows: Top—secondary flow; Middle—streamwise velocity; Bottom—temperature.]

value of the secondary flow stream function indicates a counter-clockwise (clockwise) circulation. The readers are referred to Ref. [53] for a detailed discussion of 2- and 4-cell flow structures in general, their relations with physical mechanisms and driving forces and their effects on the flow resistance and heat transfer in particular.

2. *A_1 branch*

In addition to the two limit points S_1^1 and S_1^2, the primary branch S_1 also has a symmetry-breaking bifurcation point B_1 at $Dk \doteq 188.8$ ($De \doteq 128.22$), originating an asymmetric solution branch A_1. A_1 has five limit points $A_1^1 - A_1^5$ [Fig. 6(a)]. The limit point A_1^1 divides the branch into upper sub-branch A_{1-1} and lower sub-branch A_{1-2}. The solutions on A_{1-2} can be formed by mirror images of corresponding solutions on A_{1-1} at the same Dk. While the sub-branch A_{1-1} contributes, through the two limit points A_1^2 and A_1^3 [Fig. 4(b)], three solutions for any value of Dk in a very small range $355.72 < Dk < 357.41$, the difference among these three solutions is negligibly small. Figures 6(b) to (f) illustrate the secondary flow pattern, streamwise velocity and temperature profile for several representative values of Dk on the sub-branch A_{1-1}. It is observed that flows on A_{1-1} are

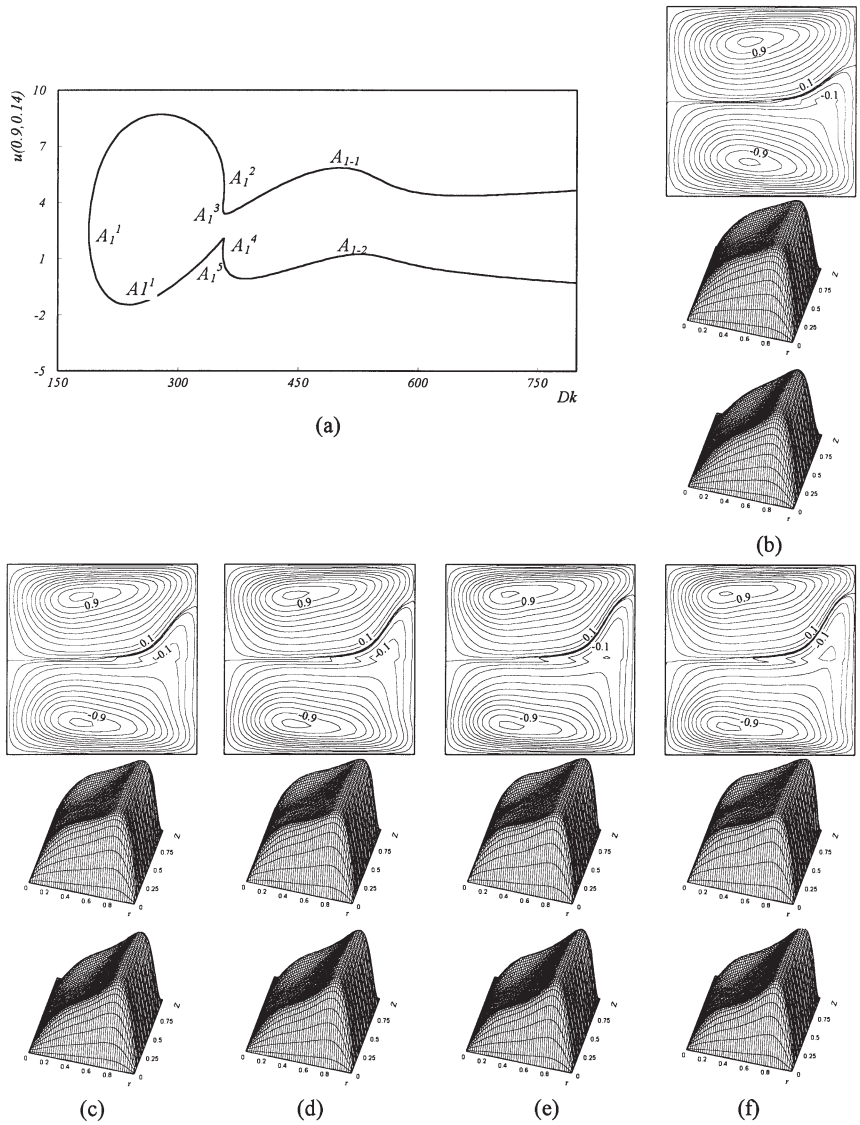
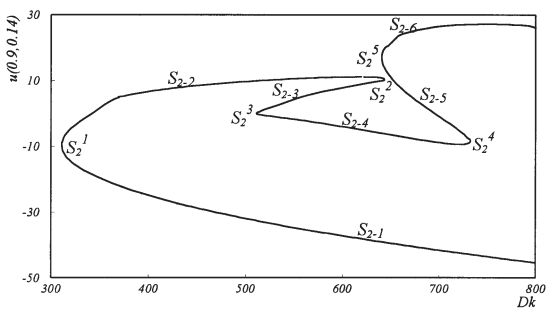


FIG. 6. Asymmetric branch A_1: its limit points, flow and temperature fields ($\sigma = 0.02$ and $Pr = 0.7$). (a) Asymmetric branch A_1 and its limit points. (b) $Dk = 300$ on A_{1-1}; $De = 183.3$, $|\psi|_{\max} = 9.60$, $w_{\max} = 0.03767$, $\theta_{\max} = 69.4$. (c) $Dk = 450$ on A_{1-1}; $De = 249.4$, $|\psi|_{\max} = 11.39$, $w_{\max} = 0.03382$, $\theta_{\max} = 83.8$. (d) $Dk = 550$ on A_{1-1}; $De = 290.0$, $|\psi|_{\max} = 12.34$, $w_{\max} = 0.03203$, $\theta_{\max} = 91.8$. (e) $Dk = 700$ on A_{1-1}; $De = 346.0$, $|\psi|_{\max} = 13.67$, $w_{\max} = 0.02998$, $\theta_{\max} = 101.9$. (f) $Dk = 800$ on A_{1-1}; $De = 381.3$, $|\psi|_{\max} = 14.47$, $w_{\max} = 0.02886$, $\theta_{\max} = 107.7$. [Figures (b) to (f) are arranged as follows: Top—secondary flow; Middle—streamwise velocity; Bottom—temperature.]

essentially an asymmetric 2-cell state. The vortex at the lower part of the duct stretches to the upper part near the outer wall. The streamwise velocity and temperature profiles are also asymmetric with respect to the horizontal central plane $z=0$. With an increase of Dk, both the asymmetry and the strength of secondary flow become stronger.

3. S_2 branch

The solution branch S_2 is an isolated symmetric branch. It is divided into six sub-branches, S_{2-1}–S_{2-6}, by five limit points, $S_2^1 - S_2^5$ [Fig. 7(a)]. To illustrate the flow structure and its evolution on each sub-branch, Figs. 7(b) to (m) show all solutions on this branch at $Dk=450$, 550, 700, and 800 in forms of secondary flow pattern, streamwise velocity, and temperature profile. The flow on S_{2-1} is a four-cell state with a stronger secondary flow at a higher value of Dk [Fig. 7(b), (d), (h), (l)]. However, this 4-cell structure differs from the one on S_{1-3} [Figs. 5(e) to (j)] mainly in the shape and size of Dean vortices. Two Dean vortices here stretch along span direction rather than the radial direction. The spanwise distance between



(a)

FIG. 7. Symmetric branch S_2: its limit points, flow and temperature fields ($\sigma=0.02$ and $Pr=0.7$). (a) Symmetric branch S_2 and its limit points. (b) $Dk=450$ on S_{2-1}; $De=241.9$, $|\psi|_{\max}=10.35$, $w_{\max}=0.03242$, $\theta_{\max}=77.9$. (c) $Dk=450$ on S_{2-2}; $De=249.1$, $|\psi|_{\max}=11.70$, $w_{\max}=0.03375$, $\theta_{\max}=83.4$. (d) $Dk=550$ on S_{2-1}; $De=279.1$, $|\psi|_{\max}=11.34$, $w_{\max}=0.03052$, $\theta_{\max}=84.2$. (e) $Dk=550$ on S_{2-2}; $De=289.2$, $|\psi|_{\max}=12.75$, $w_{\max}=0.03187$, $\theta_{\max}=91.1$. (f) $Dk=550$ on S_{2-3}; $De=290.8$, $|\psi|_{\max}=12.30$, $w_{\max}=0.03194$, $\theta_{\max}=91.7$. (g) $Dk=550$ on S_{2-4}; $De=287.8$, $|\psi|_{\max}=11.53$, $w_{\max}=0.03162$, $\theta_{\max}=90.9$. (h) $Dk=700$ on S_{2-1}; $De=331.6$, $|\psi|_{\max}=12.67$, $w_{\max}=0.02840$, $\theta_{\max}=92.3$. (i) $Dk=700$ on S_{2-4}; $De=339.6$, $|\psi|_{\max}=12.67$, $w_{\max}=0.02921$, $\theta_{\max}=98.5$. (j) $Dk=700$ on S_{2-5}; $De=339.4$, $|\psi|_{\max}=12.68$, $w_{\max}=0.02895$, $\theta_{\max}=98.2$. (k) $Dk=700$ on S_{2-6}; $De=340.8$, $|\psi|_{\max}=13.00$, $w_{\max}=0.02855$, $\theta_{\max}=98.6$. (l) $Dk=800$ on S_{2-1}; $De=364.7$, $|\psi|_{\max}=13.45$, $w_{\max}=0.02728$, $\theta_{\max}=97.2$. (m) $Dk=800$ on S_{2-6}; $De=374.9$, $|\psi|_{\max}=13.96$, $w_{\max}=0.02737$, $\theta_{\max}=103.4$. [Figures (b) to (m) are arranged as follows: Top—secondary flow; Middle—streamwise velocity; Bottom—temperature.]

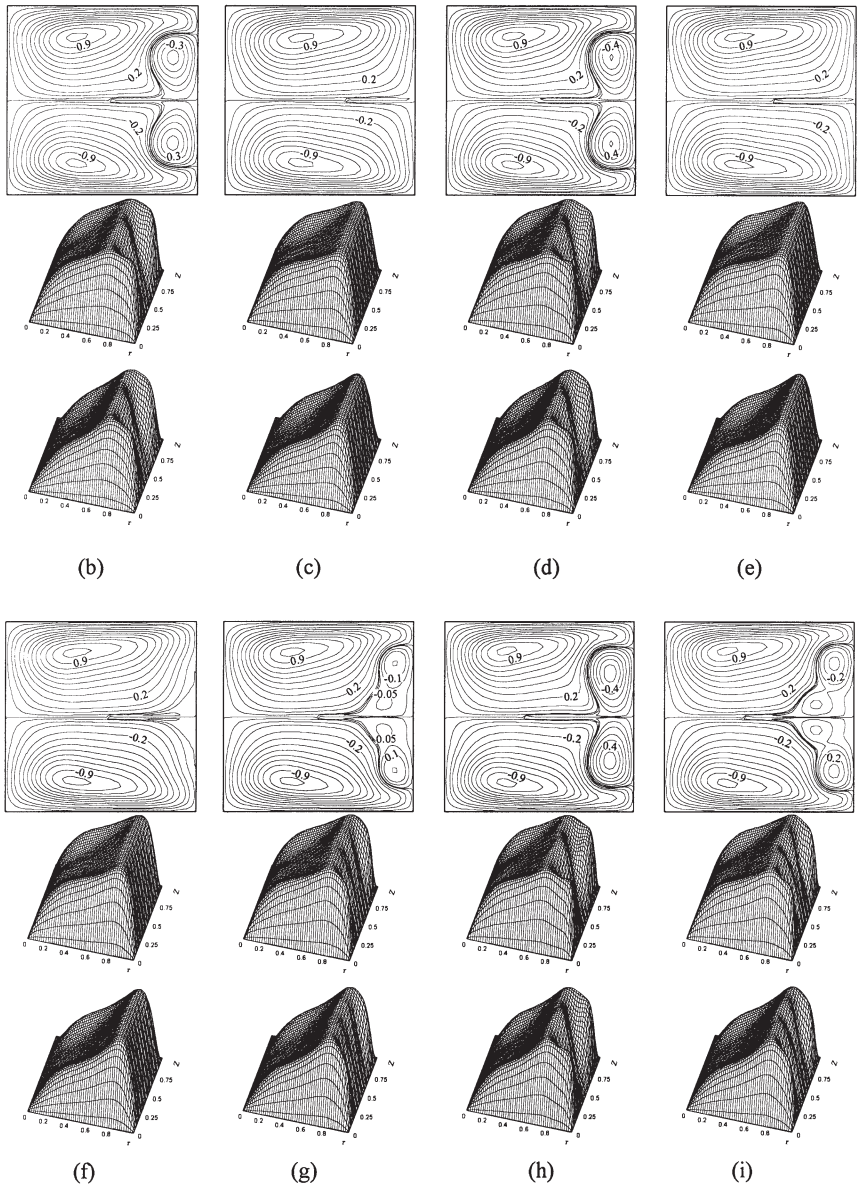


FIG. 7. Continued.

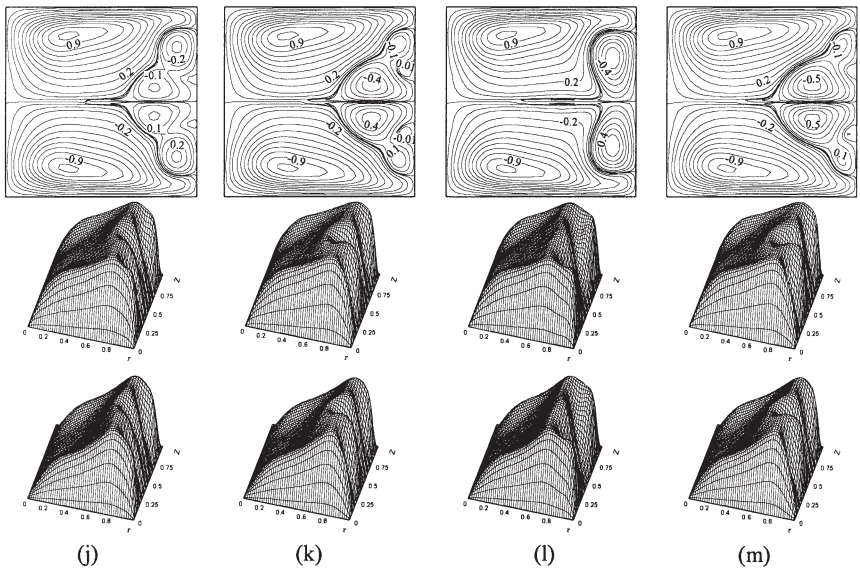


FIG. 7. Continued.

centers of two Dean vortices is noticeably larger than that of the 4-cell flow on S_{1-3}, resulting in another streamwise velocity/temperature peak between two Dean vortices in addition to the two peaks located between the Dean vortices and upper/lower walls, respectively. The flow on S_{2-2} is a 2-cell state, which is qualitatively similar to that on S_{1-1}, but with a stronger secondary flow and a further shift of streamwise velocity and temperature peaks to the outer wall [Figs. 5(b), (c) and Figs. 7(c), (e)]. The flow on S_{2-3} is a weak 4-cell state with a pair of very weak Dean vortices [Fig. 7(f)].

The limit point S_2^3 leads the weak 4-cell flow on S_{2-3} to a 6-cell state on S_{2-4} with two pairs of Dean vortices along the outer wall and a four-peak streamwise velocity and temperature [Figs. 7(g), (i)]. The second pair appears because of the splitting of the original pair. It could be interesting to study, in the future, the relation between this vortex splitting and the one due to the Eckhaus instability [20]. The secondary flow on this sub-branch also evolves to a stronger one through the growth of a second pair of Dean vortices as Dk increases [Figs. 7(g), (i)]. Flows on S_{2-5} and S_{2-6} are an 8-cell state with three pairs of Dean vortices [Figs. 7(j), (k), (m)]. The third pair is formed from the outer wall. This differs from the mechanism responsible for the appearance of the second pair. As usual, the secondary flows on S_{2-5} and S_{2-6} become stronger as Dk increases. For the same value of Dk, the 8-cell structure on S_{2-6} (the third pair of Dean vortices in particular) is stronger than that on S_{2-5}.

4. *A_2, A_3, and A_4 branches*

In addition to its five limit points, the solution branch S_2 also has three symmetry-breaking bifurcation points B_2, B_3, and B_4, originating three asymmetric solution branches A_2, A_3, and A_4, respectively [Fig. 8(a)]. Each of these branches has one limit point dividing the branch into an upper sub-branch (A_{2-1}, A_{3-1}, or A_{4-1}) and a lower sub-branch (A_{2-2}, A_{3-2}, or A_{4-2}). Solutions on the lower sub-branches are the mirror images of the corresponding solutions on the upper sub-branches at the same Dk. Figures 8(b)–(f) detail flow structures and temperature profiles on these three branches. It is observed that the strength of the secondary flow increases as Dk increases along each of these branches. The flow on the A_{2-1} is an asymmetric 2-cell state [Figs. 8(b), (d), qualitatively similar to that on A_{1-1}] while it is an asymmetric 7-cell structure on both A_{3-1} and A_{4-1} [Figs. 8(c), (e), (f)].

B. Stability of Multiple Solutions

Recognizing that there is no study on dynamic responses of multiple solutions to finite random disturbances in the literature, a relatively comprehensive transient computation is made to examine the dynamic behavior and stability of 56 typical steady solutions with respect to three sets of finite random disturbances with $d = 4\%$, 10%, and 15% respectively. It is found that the final dynamic evolution after a short transient temporal period is independent of the initial disturbances for all solutions in the region $0 \le Dk \le 650$. The results presented here are those obtained from the disturbance with $d = 10\%$ unless otherwise stated. At any fixed value of Dk in the range $0 \le Dk \le 620$, all steady solutions develop, after initial finite random disturbances, to the same final state, i.e., there is no co-existence of two or more stable states in this range within the scope of the present study. The stability of solutions on the sub-branch S_{2-2} changes as Dk changes even without passing any bifurcation or limit point. In particular, the sub-branch is unstable in the range $310.71 \le Dk \le 375$, stable in the range $375 < Dk \le 620$, and unstable again in the range $620 < Dk \le 643.3$.

Five sub-ranges are identified with each having distinct dynamic responses to the finite random disturbances. The first ranges from $Dk = 0$ to 191.27 (S_1^1), where the finite random disturbances lead all steady solutions at any fixed Dk to a 2-cell steady state on S_{1-1} with the same Dk. The second covers the range $191.27 < Dk \le 375$ where all steady solutions evolve to a temporal periodic solution. In the third sub-range, $375 < Dk \le 620$, the finite random disturbances lead all solutions to a 2-cell steady state on S_{2-2} with the same Dk. The fourth sub-range is from

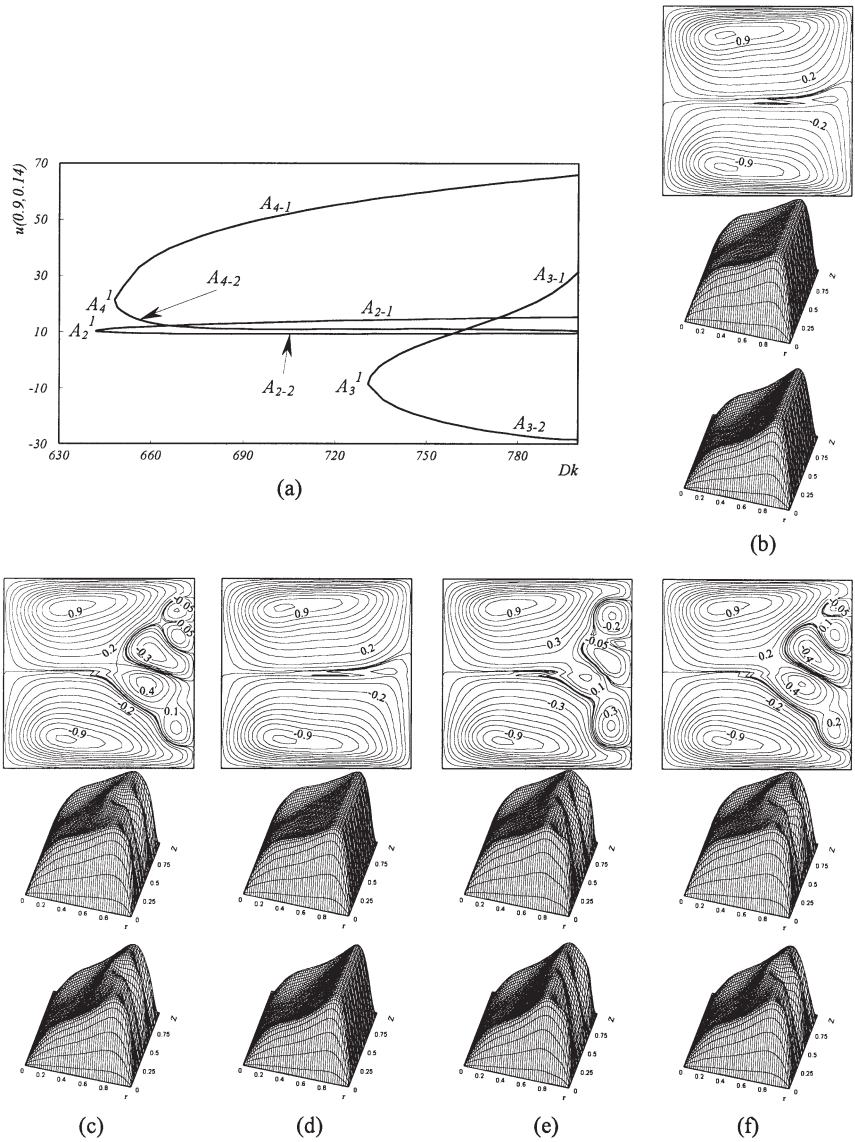


FIG. 8. Asymmetric branch A_2, A_3 and A_4: their limit points, flow and temperature fields ($\sigma = 0.02$ and $Pr = 0.7$). (a) Asymmetric branch A_2, A_3 and A_4 and their limit points. (b) $Dk = 700$ on A_{2-1}; $De = 346.2$, $|\psi|_{\max} = 14.02$, $w_{\max} = 0.02981$, $\theta_{\max} = 100.9$. (c) $Dk = 700$ on A_{4-1}; $De = 339.2$, $|\psi|_{\max} = 13.02$, $w_{\max} = 0.02852$, $\theta_{\max} = 98.4$. (d) $Dk = 800$ on A_{2-1}; $De = 381.8$, $|\psi|_{\max} = 14.82$, $w_{\max} = 0.02870$, $\theta_{\max} = 106.9$. (e) $Dk = 800$ on A_{3-1}; $De = 372.4$, $|\psi|_{\max} = 13.48$, $w_{\max} = 0.02799$, $\theta_{\max} = 103.4$. (f) $Dk = 800$ on A_{4-1}; $De = 371.6$, $|\psi|_{\max} = 13.97$, $w_{\max} = 0.02712$, $\theta_{\max} = 102.2$. [Figures (b) to (f) are arranged as follows: Top—secondary flow; Middle—streamwise velocity; Bottom—temperature.]

$Dk = 620$ to 650 where the solutions response to the finite random disturbances in the form of temporal oscillation with intermittency, a forecasting signal of chaotic flows. In the last sub-range $Dk > 650$, any finite random disturbance will deviate the solutions to chaotic oscillating states.

1. $0 \leq Dk \leq 191.27$

Winters [38] found that the Jacobian determinant is negative on S_{1-3} and the part of S_{1-2} between B_1 and S_1^1, and positive on S_{1-1} and the part of S_{1-2} between S_1^2 and B_1. The conclusion drawn in [38] is that the solutions on S_{1-3} and the part of S_{1-2} between B_1 and S_1^1 are unstable, but the solutions on S_{1-1} and the part of S_{1-2} between S_1^2 and B_1 *might* be stable. Our linear stability analysis confirms Winters' finding of the linear instability of S_{1-3} and the part of S_{1-2} between B_1 and S_1^1. For example, we found one unstable mode with $\beta = 0.3929 \times 10^{-3}$ at $Dk = 180$ on S_{1-3} and one unstable mode with $\beta = 0.863 \times 10^{-4}$ at $Dk = 190$ on S_{1-2}. Furthermore, our linear stability analysis, which is capable to ascertain the linear stability conclusively by finding all eigenvalues, shows that the solution on S_{1-1} is linearly stable, but the solution on the part of S_{1-2} between S_1^2 and B_1 is linearly unstable. All the eigenvalues at $Dk = 100$ on S_{1-1} are, for example, with a negative real part (the maximum one is -0.224×10^{-3}). At $Dk = 180$ on S_{1-2}, we found a linearly unstable mode with $\beta = 0.392 \times 10^{-3}$.

Figures 9(a) and (b) typify the responses of solutions on S_{1-1} to the finite random disturbances. In the figures, the deviation of velocity components from their initial steady values is plotted against the time τ at $(0.9, 0.14)$, $(0.94, 0.1)$, and $(0.96, 0.06)$ for $Dk = 100$ [Fig. 9(a)] and $Dk = 180$ [Fig. 9(b)], respectively. We plot both u- and v-velocity components for the first point $(0.9, 0.14)$ while only u-velocity component is shown for the last two points. To facilitate the comparison, we use these four velocity components (either velocity itself or derivation velocity from its initial steady value) in all figures illustrating dynamic responses of the multiple solutions to the finite random disturbances. It is observed that all deviation velocities vanish after a short period of time. The flows and temperature profiles return to their initial steady ones shown in Figs. 5(b) and (c). Therefore, the solutions on S_{1-1} are also stable with respect to the finite random disturbances in addition to their linear stability.

Figures 9(c) and (d) illustrate the typical responses of solutions on S_{1-2} and S_{1-3} to the finite random disturbances. It shows that the finite random disturbances lead eventually the solutions on these two sub-branches to the stable one on S_{1-1} at the same Dk. This was also confirmed by our detailed check of flow and temperature fields. Therefore, the solutions on S_{1-2} and S_{1-3} are unstable to the finite random disturbances and response the disturbances by evolving to the stable solution on S_{1-1} at the same Dk.

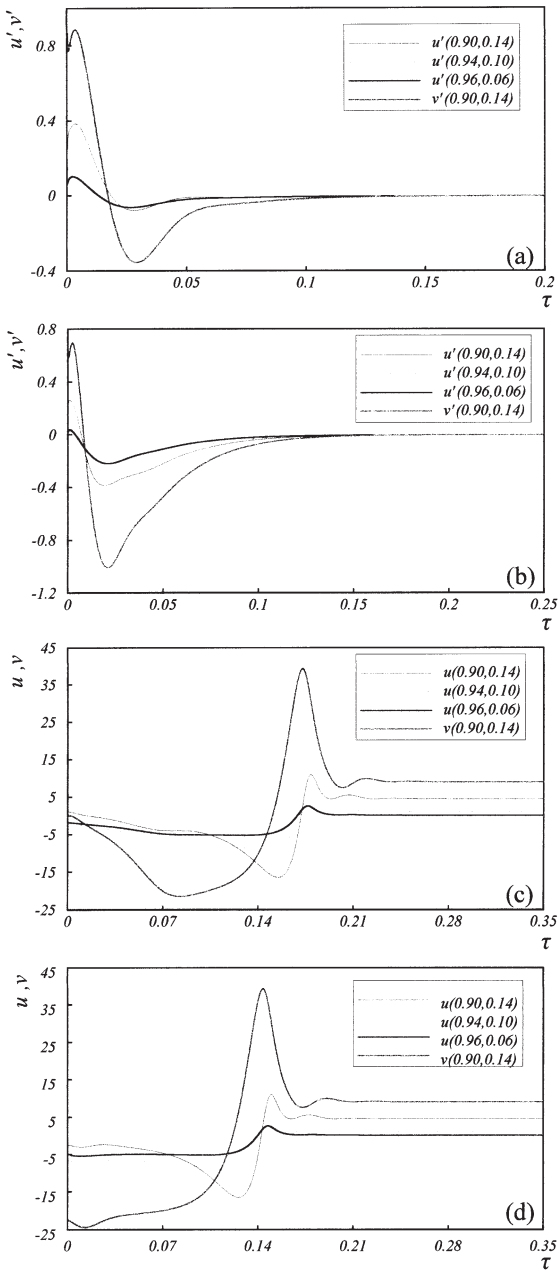


FIG. 9. Dynamic response of solutions on S_{1-1}, S_{1-2} and S_{1-3} to finite random disturbances: evolution to stable steady 2-cell state on S_{1-1} ($\sigma = 0.02$ and $Pr = 0.7$). (a) $Dk = 100$ on S_{1-1}; the final state is shown in Fig. 5(b). (b) $Dk = 180$ on S_{1-1}; the final state is shown in Fig. 5(c). (c) $Dk = 180$ on S_{1-2}; the final state is shown in Fig. 5(c). (d) $Dk = 180$ on S_{1-3}; the final state is shown in Fig. 5(c).

2. *191.27 < Dk ≤ 375*

Winters [38] found that the Jacobian determinant is positive on A_1. Therefore, the linear stability of this branch is not conclusive. Our linear stability analysis shows that this branch is linearly unstable. For example, we found one pair of unstable modes with $\beta = 0.278 \times 10^{-4} \pm 0.427 \times 10^{-3} i$ at $Dk = 200$ on A_{1-1}. The linear stability of S_{1-3} and S_{2-1} examined by Winters [38] is reconfirmed in this range by our linear stability analysis. We detected, for instance, one linearly unstable mode with $\beta = 0.614 \times 10^{-3}$ at $Dk = 300$ on S_{1-3}, and two pairs of unstable modes with $\beta = 0.498 \times 10^{-3} \pm 0.865 \times 10^{-3} i$ and $\beta = 0.423 \times 10^{-3} \pm 0.115 \times 10^{-3} i$ at $Dk = 350$ on S_{2-1}. In contrast to that in Ref. [38], however, we found that S_{2-2} is linearly unstable in this range. There exist, for example, one pair of linearly unstable modes with $\beta = 0.498 \times 10^{-3} \pm 0.865 \times 10^{-3} i$ at $Dk = 350$.

The dynamic response of the solution at $Dk = 300$ on A_{1-1} is shown in Fig. 10. The finite random disturbances here lead the solution to a temporal periodic state with a period of 0.159. Some typical secondary flow patterns are detailed in Fig. 11 within one period of τ. We clearly observe the temporal oscillations between symmetric/asymmetric 2-cell flows and symmetric/asymmetric 4-cell flows. With the spanwise direction vertical and the radial direction horizontal, the corresponding smoke patterns are also shown in Fig. 11 for the comparison. The smoke patterns come from our flow visualization by using experimental apparatus and the technique detailed in [41,42,45,66,67]. The numerical results agree very well with the flow visualization. Figure 12(a) illustrates the dynamic response of the solution at $Dk = 300$ on S_{1-3}. A comparison with that in Fig. 10 tends to show that the finite random disturbances finally lead to the same periodic state with the period of 0.159. We also made the detailed comparison of

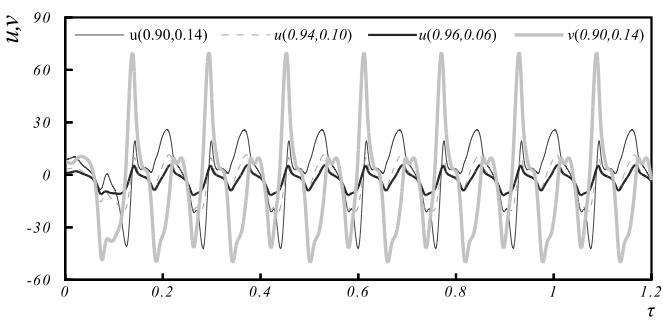


FIG. 10. Dynamic responses of solutions at $Dk = 300$ on A_{1-1} to finite random disturbances: periodic oscillation (period = 0.159, $\sigma = 0.02$ and $Pr = 0.7$).

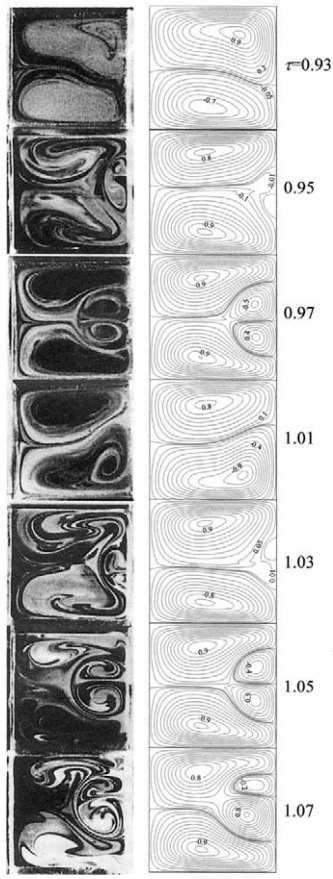


FIG. 11. Typical secondary flow patterns in one period of temporal periodic oscillation from the solution at $Dk = 300$ on A_{1-1} (Left: flow visualization; Right: numerical simulation).

flow and temperature fields within one period, and found that the finite random disturbances indeed lead all the solutions at the same Dk to the same periodic oscillation.

Figures 12(b) and (c) show the dynamic evolution of solutions at $Dk = 200$ on A_{1-1} and at $Dk = 350$ on S_{2-2}, respectively, to further detail some features of the temporally periodic oscillation. A similar pattern of peaks and valleys exists for all three cases in Figs. 10 and 12(b), (c) with different values of Dk. This signal shows the similarity of flow and temperature fields within one period among the periodic states for different values of Dk in the range $191.27 < Dk \leq 375$. Our detailed examination of

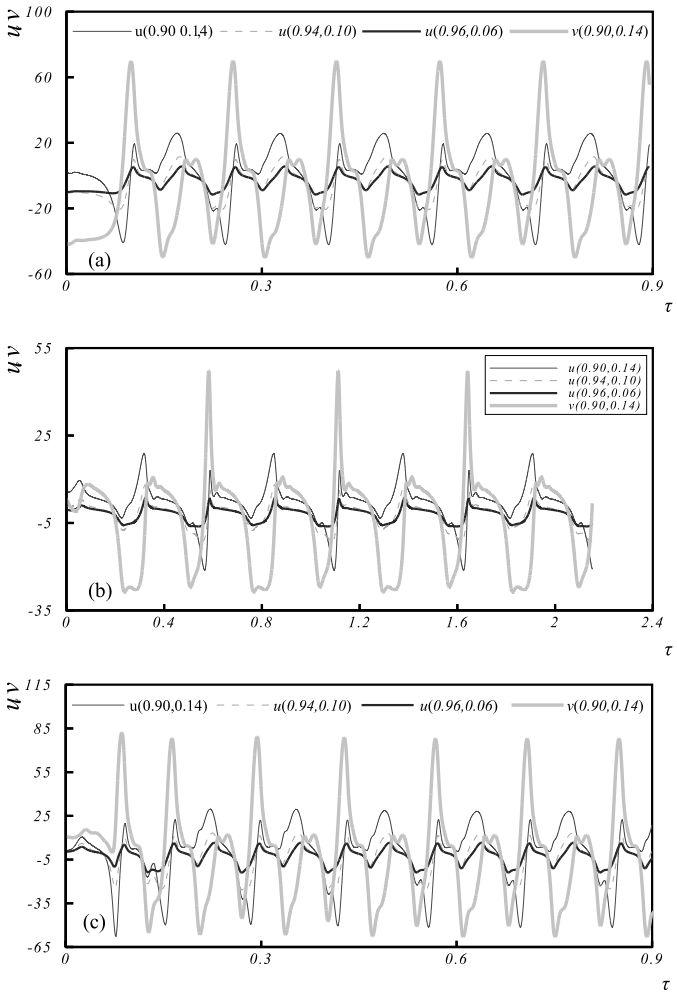


FIG. 12. Dynamic response of solution to finite random disturbances: periodic oscillation ($\sigma = 0.02$ and $Pr = 0.7$). (a) $Dk = 300$ on S_{1-3} (period = 0.159). (b) $Dk = 200$ on A_{1-1} (period = 0.530). (c) $Dk = 350$ on S_{2-2} (period = 0.140).

flow and temperature fields has confirmed this, and shown that the flow structures in Fig. 11 are typical for all Dk.

The period is 0.53 for the periodic flow in Fig. 12(b), and 0.14 for the periodic flow in Fig. 12(c). The variation of the period with Dk is shown and compared with that from the flow visualization in Fig. 13. It is observed that the period first increases with Dk, and then decreases. But, it increases once

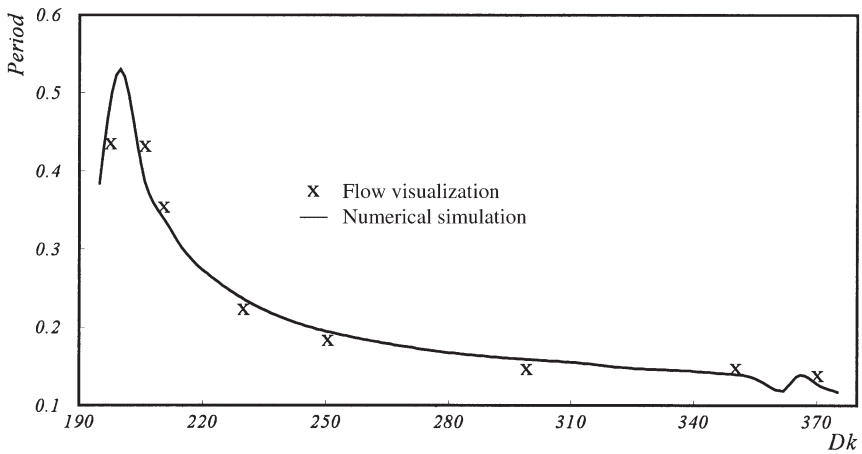


FIG. 13. Variation of period with Dk ($\sigma = 0.02$ and $Pr = 0.7$).

again slightly around $Dk = 362$ and then decreases as Dk increases. There is no period-doubling bifurcation. While it plays an important role of generating chaos in closed flow systems such as the Taylor–Couette flow and the Rayleigh–Benard problem [49–51], the period-doubling bifurcation seems to be suppressed by the pressure-driven flows in the parameter range under consideration for the flow in curved ducts, a typical open flow system which requires the flow to be introduced and removed from the observational domain. Here, the temporal periodic oscillation serves as the transition from one stable steady flow in $0 \leq Dk \leq 191.27$ to another in $375 < Dk \leq 620$ as discussed in the next subsection.

The dynamic responses of solutions on A_1 from $Dk = 188.67$ to 375 tend to show that B_1 is a subcritical Hopf bifurcation point.

3. *$375 < Dk \leq 620$*

In this range, our linear stability analysis shows that S_{2-1} is linearly unstable, but S_{2-2} is linearly stable. For instance, we found three unstable modes with $\beta = 0.721$, $0.203 \times 10^{-2} \pm 0.146 \times 10^{-2} i$ at $Dk = 450$ on S_{2-1}, but no unstable mode at $Dk = 450$ on S_{2-2} (the eigenvalues with the maximum real part are $\beta = -0.887 \times 10^{-3} \pm 0.616 \times 10^{-2} i$). It is of special interest that the linear stability of S_{2-2} changes from unstable in the last range to stable in this range. Daskopoulos and Lenhoff [39] found that S_{2-4} is linearly stable to symmetric disturbances. However, our linear stability analysis shows that it is linearly unstable to the asymmetric disturbances. We detected, for example, five unstable modes at $Dk = 550$ on S_{2-4} with $\beta = 0.01505$, $0.011277 \pm 0.003738i$, $0.002507 \pm 0.003373i$, respectively.

The 2-cell state on S_{2-2} is not only linearly stable but also stable to the finite random disturbances in this sub-range. This can be referred by the typical responses of solutions at $Dk = 450$, 600, 610, and 620 on this sub-branch to the disturbances in Fig. 14. Another striking feature can be obtained by comparing the dynamic process in Fig. 14 with that in Fig. 9. The transient solutions approach their stable steady states asymptotically for S_{1-1}, but in oscillation for S_{2-2}. The oscillation in Fig. 14 may be reviewed as the over-damped oscillation with the damping effect being weaker at higher Dk, and thus a longer oscillating time. This difference is in agreement with findings of our linear stability analysis: the eigenvalues β with maximum real part are real-valued for S_{1-1}, but complex-valued for S_{2-2}.

In this sub-range of the parameter space, the finite random disturbances will lead all the other solutions to the 2-cell steady state on S_{2-2} at the same Dk. Figure 15 typifies this process by the dynamic response of solutions at $Dk = 380$ on A_{1-1} and at $Dk = 550$ on S_{2-1} to the disturbances.

4. $620 < Dk \leq 650$

In this sub-range, all solution branches are linearly unstable. The S_{2-2}, in particular, loses its linear stability gained in the last sub-range. While three sets of finite random disturbances lead each solution to the same final temporal oscillation state, the solutions at different solution branches at the same Dk respond to the disturbances differently in the sense that the final oscillation is different. However, for a fixed value of Dk, all the oscillations are around the steady solutions on S_{2-2} before $Dk = 643.30$ (point S_2^2) and around the steady solutions on A_2 after $Dk = 643.30$. Whether it is A_{2-1} or A_{2-2} is difficult to distinguish because they are very close. Therefore, S_{2-2} and A_2 differ from the other solution branches in this sense.

Figure 16 details the dynamic responses to the disturbances of the solution at $Dk = 630$ on S_{2-1} [Fig. 16(a)], at $Dk = 630$ on S_{2-2} [Fig. 16(b)], at $Dk = 638$ on S_{2-2} [Fig. 16(c)], at $Dk = 640$ on S_{2-1} [Fig. 16(d)], at $Dk = 642$ on A_{2-1} [Fig. 16(e)] and at $Dk = 650$ on A_{4-1} [Fig. 16(f)], respectively. It is observed that there exist intermittent aperiodic bursts in the oscillation, one of routes to chaos [47,51,68]. The oscillation is quasi-periodic between two bursts with the periodicity being degraded as Dk increases. The frequency of the appearance of bursts increases as Dk increases.

Figure 17 details the secondary flow during quasi-periodic oscillations [Fig. 17(a), (e), (f)] and bursts [Fig. 17(b), (c), (d)] for the case shown in Fig. 16(a). It is observed that the flow oscillates among symmetric/asymmetric 2-cell patterns during quasi-periodic oscillations but among symmetric/asymmetric 4-cell structures during bursts.

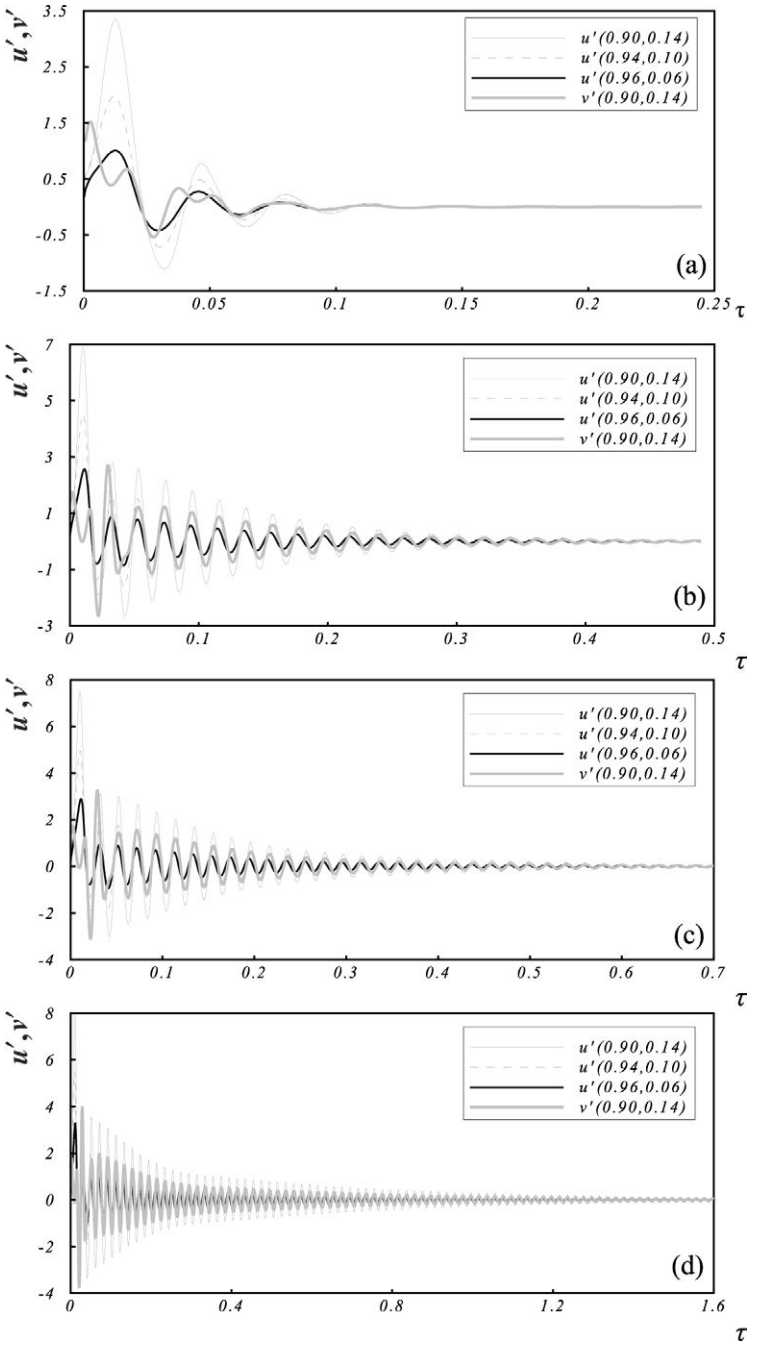


FIG. 14. Dynamic response of solutions on S_{2-2} to finite random disturbances: stable steady 2-cell state on S_{2-2} ($\sigma = 0.02$ and $Pr = 0.7$). (a) $Dk = 450$; the final state is shown in Fig. 7(c). (b) $Dk = 600$. (c) $Dk = 610$. (d) $Dk = 620$.

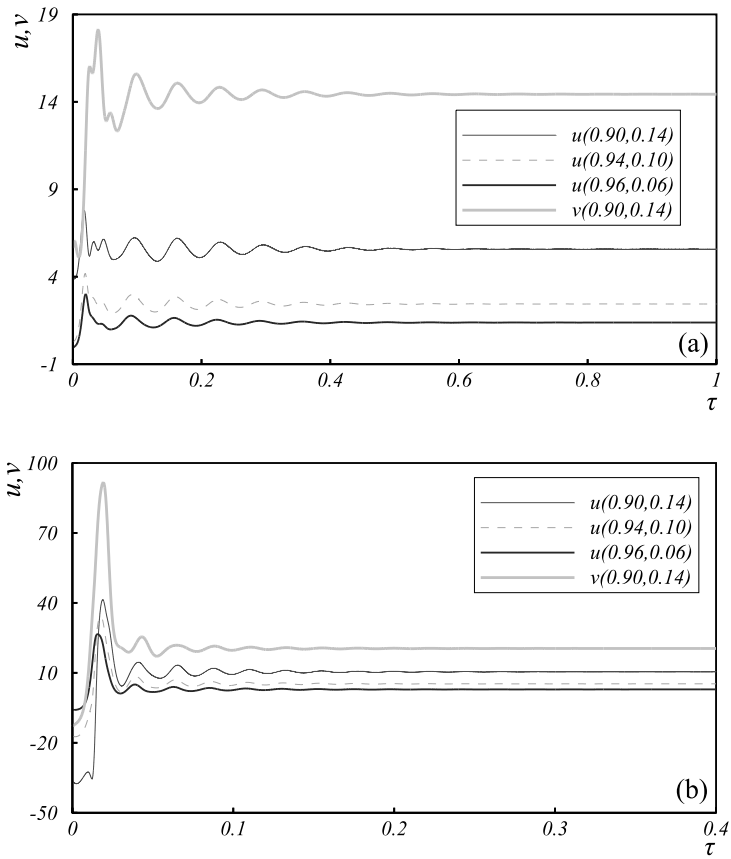


FIG. 15. Dynamic responses of solutions to finite random disturbances: evolution to stable steady 2-cell state on S_{2-2} ($\sigma = 0.02$ and $Pr = 0.7$). (a) $Dk = 380$ on A_{1-1}. (b) $Dk = 550$ on S_{2-1}.

5. $650 < Dk \leq 800$

In this sub-range, all solution branches are also linearly unstable. For example, we found five unstable modes at $Dk = 700$ on A_{4-2} with $\beta =$ 0.0106, 0.00927, 0.00389, $0.0035 \pm 0.00396i$, respectively.

Figure 18(a) shows the dynamic response of the solution at $Dk = 800$ on A_{3-1} to the finite random disturbance with $d = 10\%$. The bursts are still observed, but with a high appearance frequency and a generally small amplitude. The oscillation between two bursts, however, cannot be viewed as quasi-periodic any more. The power spectra of four velocity temporal series in Fig. 18(a) are constructed by the Fourier transformation and are

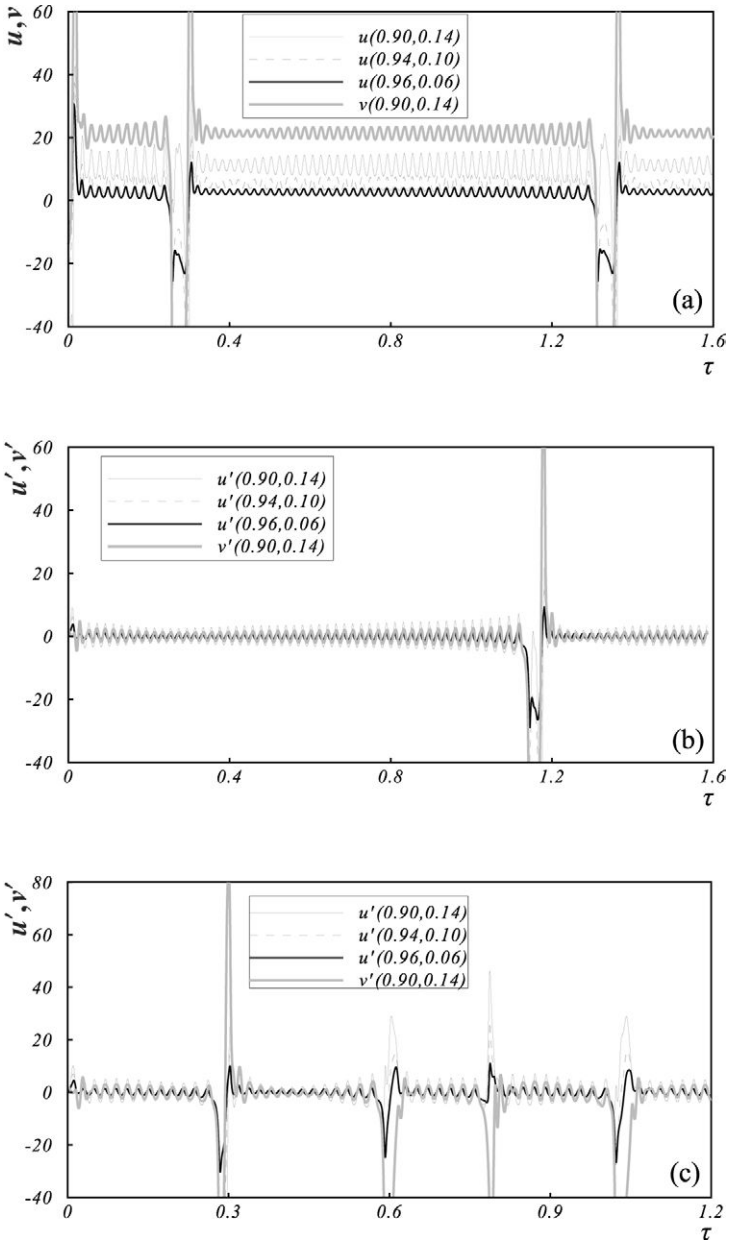


FIG. 16. Dynamic responses of solutions to finite random disturbances: intermittency ($\sigma = 0.02$ and $Pr = 0.7$). (a) $Dk = 630$ on S_{2-1}. (b) $Dk = 630$ on S_{2-2}. (c) $Dk = 638$ on S_{2-2}. (d) $Dk = 640$ on S_{2-1}. (e) $Dk = 642$ on A_{2-1}. (f) $Dk = 650$ on S_{2-5}.

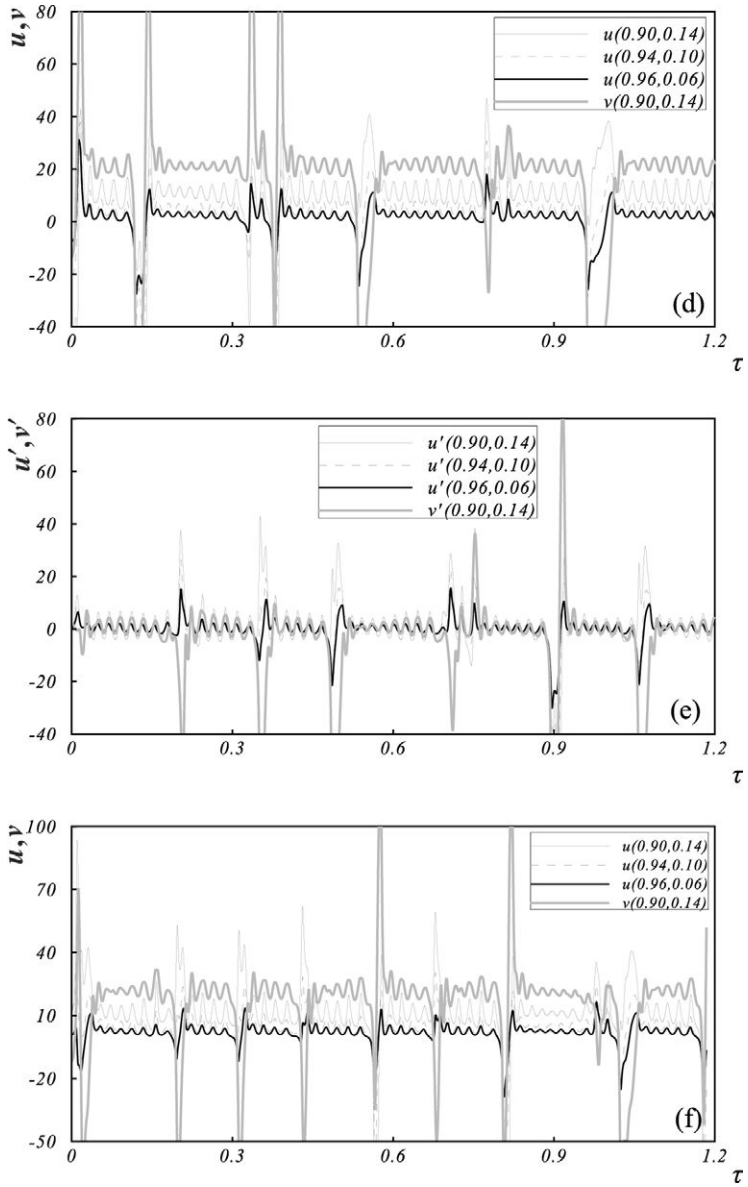
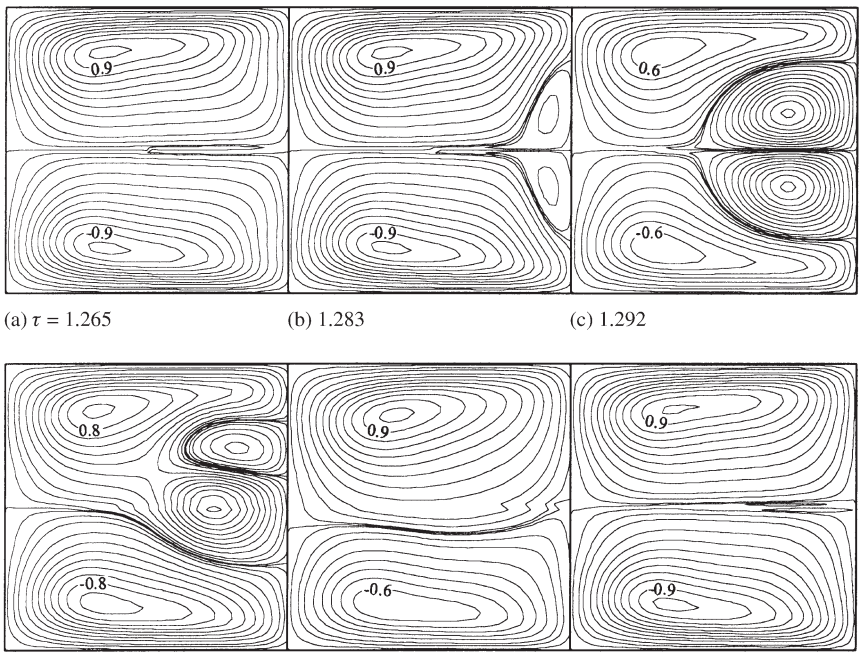


FIG. 16. Continued.



(a) $\tau = 1.265$ (b) 1.283 (c) 1.292

(d) 1.326 (e) 1.349 (f) 1.385

FIG. 17. Typical secondary flow patterns of intermittent flow from the solution at $Dk = 630$ on S_{2-1}.

shown in Fig. 18(b). They contain the broad-band noise, indicating the flow being chaotic [47]. The sensitivity to the initial conditions serves as another criterion of chaos [69]. Figure 18(c) shows the dynamic response of the solution to the disturbance with $d = 15\%$. The oscillation in Fig. 18(c) is observed to be different from that in Fig. 18(a). This further confirms that the oscillation in Fig. 18(a) is chaotic.

In spite of their chaotic feature, all oscillations from different solution branches and different disturbances at the same Dk appear oscillating around a common equilibrium state, the steady solution on A_2. As A_{2-1} and A_{2-2} are very close, it is difficult to distinguish whether it is A_{2-1} or A_{2-2}. Figure 19 shows dynamic responses of nine solutions at $Dk = 700$ on different solution branches to the disturbances by the u-velocity at (0.9, 0.14). It is observed that all u's evolve into A_2 at $Dk = 700$ in a short time period and then oscillate chaotically around it. This has also been confirmed by our detailed check of flow patterns and temperature profiles. Therefore, the solution branch A_2 differs from the others in its stability.

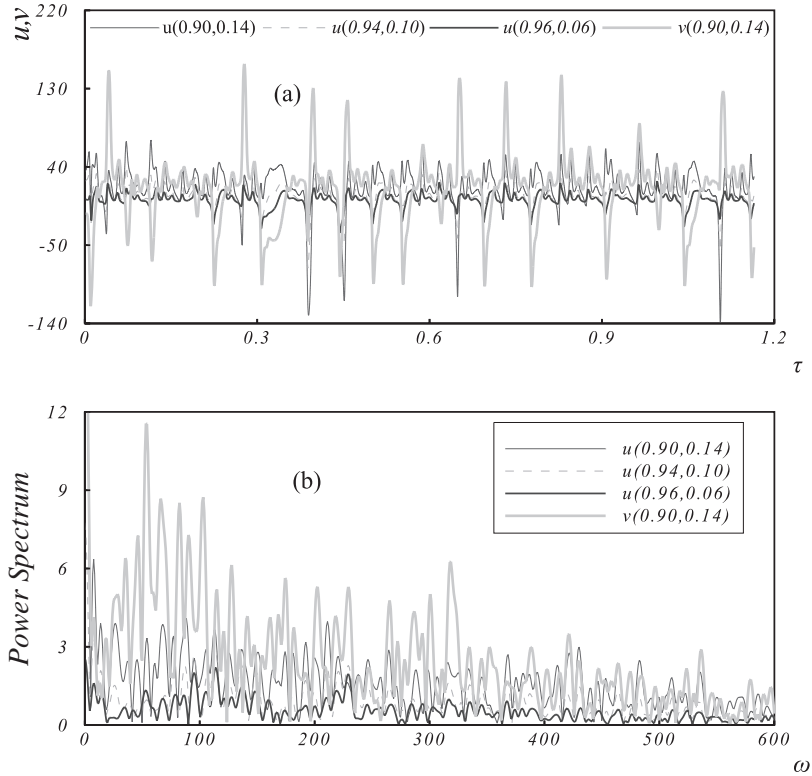


FIG. 18. Dynamic response of the solution at $Dk = 800$ on A_{3-1} to finite random disturbances: chaotic oscillation ($\sigma = 0.02$ and $Pr = 0.7$). (a) Dynamic response to the finite random disturbance with $d = 10\%$. (b) Power spectrum of the chaotic oscillation in (a). (c) Dynamic response to the finite random disturbance with $d = 15\%$.

Figure 20 details some typical secondary flow patterns for the temporal chaotic flow shown in Fig. 18(a). It is observed that the flow oscillates among 4-cell patterns during bursts [Fig. 20(c), (d)] but among 2-cell and 3-cell structures during the other period [Fig. 20(a), (b)].

The transition from the intermittent oscillation to the chaotic oscillation is believed to be a smooth process characterized by the increase of appearance frequency and the decrease of the amplitude of bursts as the increase of Dk. Figure 21 shows the dynamic response of the solution at $Dk = 660$ on A_{4-1} to the disturbance. The oscillation is observed to be chaotic. Therefore, it is not unreasonable to conjecture that the transition takes place gradually within $Dk = 650$–660.

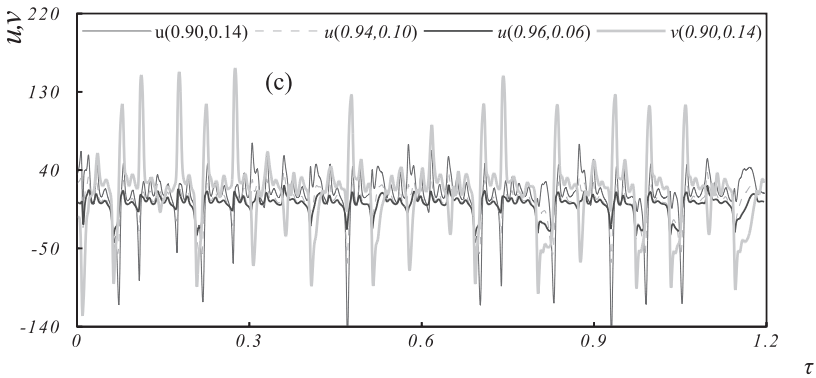


FIG. 18. Continued.

C. Friction Factor and Nusselt Number

For engineering applications, the most important results are friction factor and Nusselt number. Since the friction factor and Nusselt number depend on flow and temperature fields, the multiplicity and stability will strongly affect both their distributions and their mean values.

The mean friction factor and Nusselt number on various solution branches are shown in Figs. 22 and 23, respectively. They are shown on the basis of their values for a straight duct [$(fRe)_0 = 14.23$; $Nu_0 = 3.608$] to illustrate the curvature effect. Even for the same value of Dk and in terms of their mean values, both fRe and Nu are different for different solution branches. The temporal oscillation of flow and temperature fields in $191.22 < Dk \leq 375$ and $Dk > 620$ would also lead to a temporal oscillation of fRe and Nu. Such an oscillation could, in turn, generate the oscillation of pumping system and thermal stress in practice, which can cause the failure of pumping system and equipments. Therefore, the multiplicity and stability study is not only of theoretical importance, but also practical relevance.

VI. Concluding Remarks

The literature review identifies several unresolved issues regarding the multiplicity and stability of convection in curved ducts. The attempt to address these issues leads to a comprehensive numerical study. The governing differential equations from the conservation laws are discretized by the finite volume method to obtain discretization equations, a set of

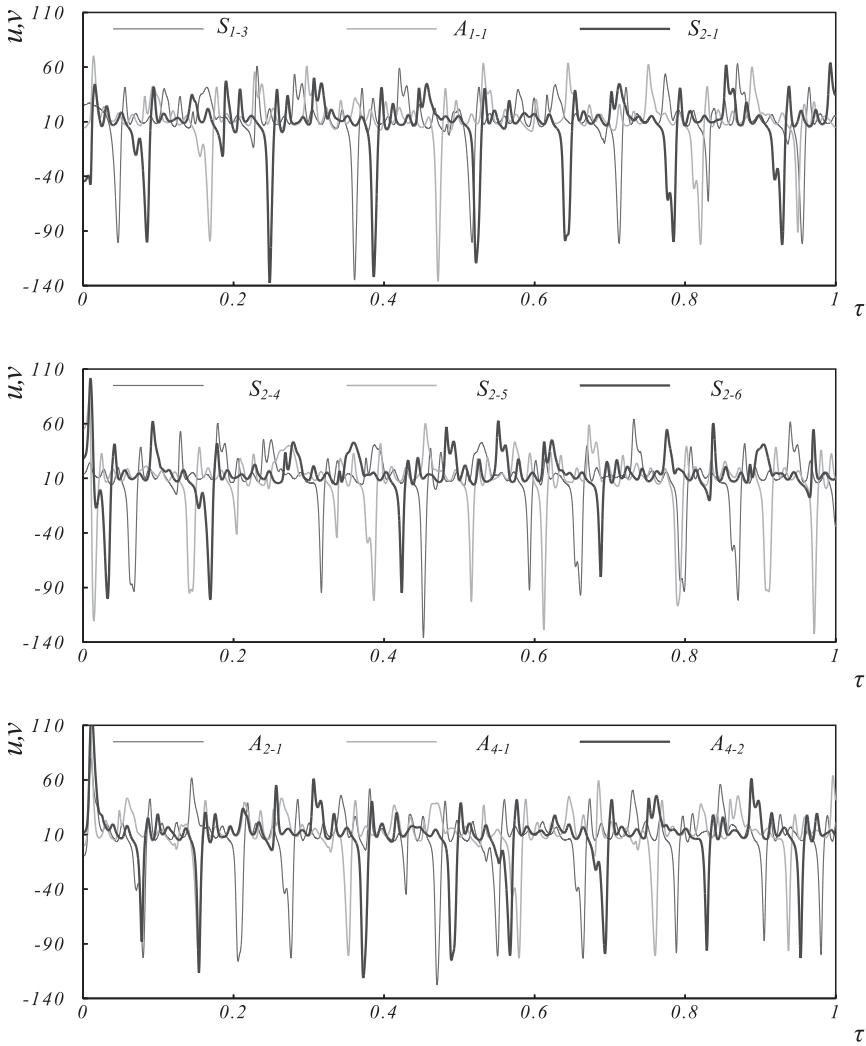
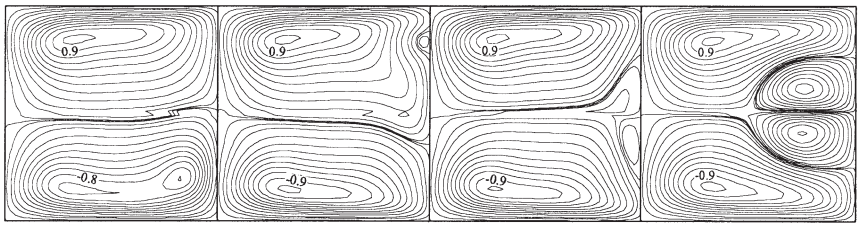


FIG. 19. Dynamic response of u at (0.9, 0.14) to finite random disturbances for nine solutions at $Dk = 700$ ($\sigma = 0.02$ and $Pr = 0.7$).

non-linear algebraic equations. The discretization equations are solved for parameter-dependence of flow and temperature fields by the Euler–Newton continuation with the solution branches parameterized by the pseudo Dean number Dk, the arclength, or the local variable. The bifurcation points are detected by the test function τ_{ij}. The Hopf bifurcation point is determined by the transient computation. The branch switching is made by a scheme



(a) $\tau = 0.86$ (b) 0.87 (c) 0.90 (d) 0.91

FIG. 20. Typical secondary flow patterns of chaotic flow from the solution at $Dk = 800$ on A_{3-1}.

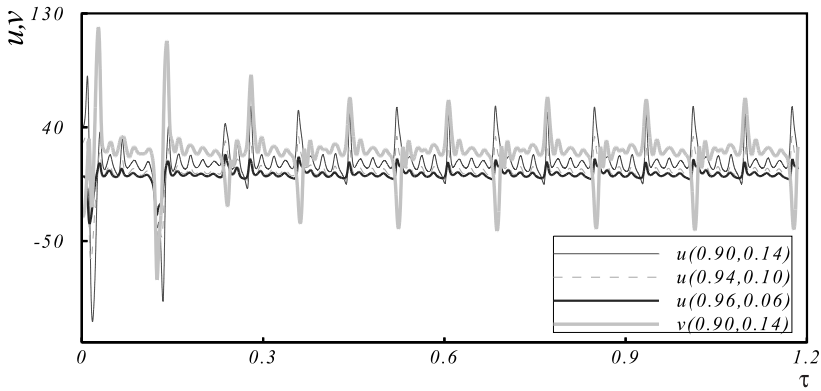


FIG. 21. Dynamic response of solution at $Dk = 660$ on A_{4-1} to finite random disturbances: chaotic oscillation ($\sigma = 0.02$ and $Pr = 0.7$).

approximating the difference between branches. Two symmetric and four asymmetric solution branches are found. Among them, three new asymmetric branches are with either a 2-cell or a 7-cell flow structure, and arise from three new symmetry-breaking bifurcation points on the symmetric branch S_2. The bifurcation point B_1 is found to be a subcritical Hopf point.

The linear stability of multiple solutions is made by solving the eigenvalue system for all eigenvalues. This can lead to a conclusive result regarding the linear stability of solutions. S_{1-1} ($0 \leq Dk \leq 191.27$) and the part of S_{2-2} between $Dk = 375$ and $Dk = 620$ are linearly stable. Both are with a 2-cell flow structure. All the other sub-branches and branches are linearly unstable. In particular, S_{2-4} is linearly unstable to the asymmetric disturbances although it is stable to the symmetric disturbances. S_{2-2} can gain and lose its linear stability without passing any bifurcation or limit points as Dk changes.

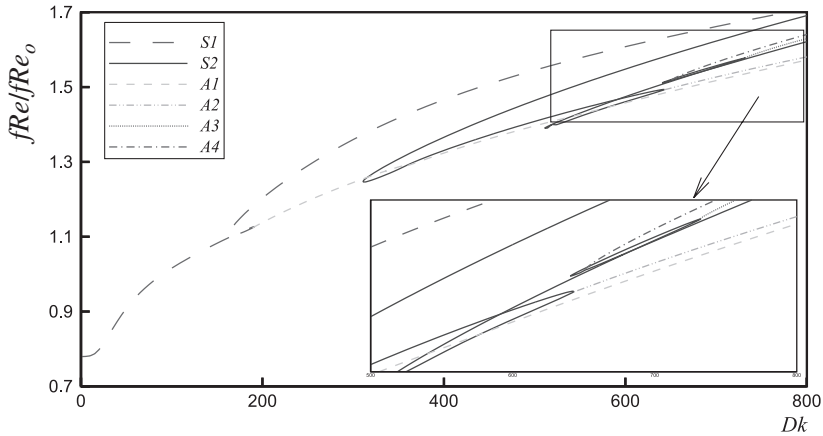


FIG. 22. Mean friction factor on various solution branches ($\sigma = 0.02$ and $Pr = 0.7$).

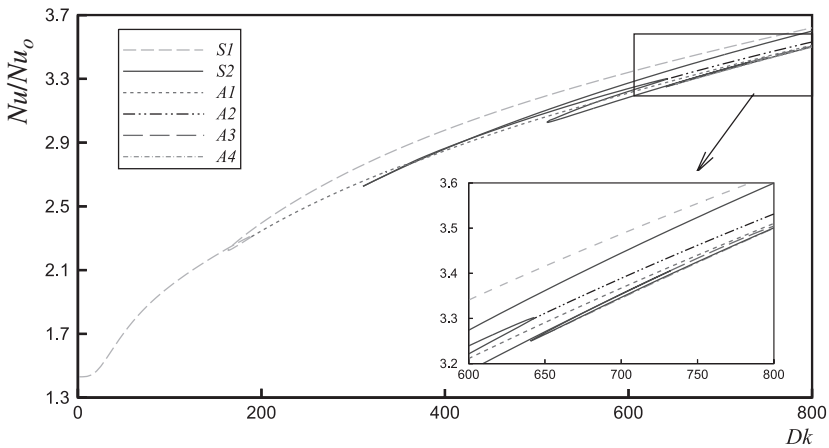


FIG. 23. Mean Nusselt number on various solution branches ($\sigma = 0.02$ and $Pr = 0.7$).

The dynamic responses of multiple solutions to the 2D finite random disturbances are examined by the direct transient computation. The solutions are found to respond to three sets of finite random disturbances in the same manner in the range of $Dk \leq 650$. At any fixed value of Dk in the range $0 \leq Dk \leq 620$, all steady solutions develop, after the initial finite random disturbances, to the same final state. The finite random

disturbances are found to lead the steady solutions to a stable steady 2-cell state on S_{1-1} in $0 \leq Dk \leq 191.27$, a temporal periodic oscillation between symmetric/asymmetric 2-cell flows and symmetric/asymmetric 4-cell flows in $191.27 < Dk \leq 375$, another stable steady 2-cell state on S_{2-2} in $375 < Dk \leq 620$, a temporal oscillation with the intermittency around the steady solution on S_{2-2} before the limit point S_2^2 and on A_2 after S_2^2 in $620 < Dk \leq 650$, and a chaotic oscillation around the steady solution on A_2 in $650 < Dk \leq 800$. The intermittent flow is characterized by the flow oscillation among the symmetric/asymmetric 2-cell patterns during the period between two bursts and among the symmetric/asymmetric 4-cell structures during the burst. The chaotic flow is, on the other hand, featured by the flow oscillation among the 4-cell patterns during the burst, and the 2-cell and 3-cell structures during the other period.

The period-doubling bifurcation is suppressed by the pressure-driven flows for flows in curved ducts in the parameter range considered in the present work. The temporal periodic oscillation serves as the transition from one stable steady flow to another. However, the intermittency leads the curved duct flow to chaos.

Acknowledgements

The financial support from the CRCG and the Research Grants Council of the Hong Kong Special Administrative Region (Project No. HKU7086/00E) to LW is gratefully acknowledged.

References

1. Pedley, T. J. (1980). "The Fluid Mechanics of Large Blood Vessels". Cambridge Univ. Press, Cambridge.
2. Berger, S. A., Talbot, L., and Yao, L. S. (1983). Flow in curved pipes. *Annu. Rev. Fluid Mech.* **15**, 461–512.
3. Nandakumar, K. and Masliyah, J. H. (1986). Swirling flow and heat transfer in coiled and twisted pipes. *Adv. Transport Process.* **4**, 49–112.
4. Itō, H. (1987). Flow in curved pipes. *JSME Int. J.* **30**, 543–552.
5. Berger, S. A. (1991). Flow and heat transfer in curved pipes and tubes. AIAA 91-0030, 1–19.
6. Yang, T. L. (2001). Multiplicity and stability of flow and heat transfer in rotating curved ducts, PhD Thesis, Department of Mechanical Engineering, The University of Hong Kong.
7. Dean, W. R. (1928). Fluid motion in a curved channel. *Proc. Roy. Soc. A* **121**, 402–420.
8. Yih, C. S. and Sanyster, W. M. (1957). Stability of laminar flows in curved channels. *Phil. Mag.* **2**, 305–310.

9. Reid, W. H. (1958). On the stability of viscous flow in a curved channel. *Proc. Roy. Soc. A* **244**, 186–198.
10. Hammerlin, G. (1958). Die stabilität der strömung in einem gekrümmten kanal. *Arch. Rat. Mech. and Anal.* **1**, 212–224.
11. Brewster, D. B., Grosberg, P., and Nissan, A. H. (1959). The stability of viscous flow between horizontal concentric cylinders. *Proc. Roy. Soc. A* **251**, 76–91.
12. Chandrasekhar, S. (1961). "Hydrodynamic and Hydromagnetic Stability". Oxford University Press, London.
13. Lin, C.C. (1955). "The Theory of Hydrodynamic Stability". Cambridge, England.
14. Gibson, R. D. and Cook, A. E. (1974). The stability of curved channel flow. *Q. J. Mech. Appl. Maths.* **27**, 149–160.
15. Kelleher, M. D., Flentie, D. L., and Mckee, R. J. (1980). An experimental study of the secondary flow in a curved rectangular channel. *J. Fluids Eng.* **102**, 92–96.
16. Ligrani, P. M. and Niver, R. D. (1988). Flow visualization of Dean vortices in a curved channel with 40 to 1 aspect ratio. *Phys. Fluids* **31**, 3605–3618.
17. Finlay, W. H., Keller, J. B., and Ferziger, J. H. (1988). Instabilities and transition in curved channel flow. *J. Fluid Mech.* **194**, 417–456.
18. Ligrani, P. M., Finlay, W. H., Fields, W. A., Fuqua, S. T., and Subramaman, C. S. (1992). Features of wavy vortices in a curved channel from experimental and numerical studies. *Phys. Fluids A* **4**, 695–709.
19. Le Cunff, C. and Bottaro, A. (1993). Linear stability of shear profiles and relation to the secondary instability of the Dean flow. *Phys. Fluids A* **5**, 2161–2171.
20. Guo, Y. and Finlay, W. H. (1991). Splitting, merging and wavelength selection of vortices in curved and/or rotating channel flow due to Eckhaus instability. *J. Fluid Mech.* **228**, 661–691.
21. Alfredsson, P. A. and Persson, H. (1989). Instabilities in channel flow with system rotation. *J. Fluid Mech.* **202**, 543–557.
22. Matsson, O. J. E. and Alfredsson, P. H. (1990). Curvature- and rotation-induced instabilities in channel flow. *J. Fluid Mech.* **210**, 537–563.
23. Matsson, O. J. E. and Alfredsson, P. H. (1992). Experiments on instabilities in curved channel flow. *Phys. Fluids A* **4**, 1666–1676.
24. Bottaro, A. (1993). On longitudinal vortices in curved channel flow. *J. Fluid Mech.* **251**, 627–660.
25. Guo, Y. and Finlay, W. H. (1994). Wavenumber selection and irregularity of spatially developing nonlinear Dean and Görtler vortices. *J. Fluid Mech.* **264**, 1–40.
26. Itō, H. (1951). Theory on laminar flow through curved pipes of elliptic and rectangular cross-sections. *Rep. Inst. High Speed Mech.* **1**, 1–16.
27. Cuming, H. G. (1952). The secondary flow in curved pipes. *Aeronaut. Res. Counc. Rep. Mem.* No. 2880.
28. Stuart, J. T. (1963). Hydrodynamic stability. *In* "Laminar Boundary Layers" (L. Rosenhead, ed.), p. 505. Oxford University Press.
29. Cheng, K. C., Lin, R. C., and Ou, J. W. (1976). Fully developed laminar flow in curved rectangular channels. *J. Fluids Eng.* **98**, 41–48.
30. Cheng, K. C. and Akiyama, M. (1970). Laminar forced convection heat transfer in curved rectangular channels. *Int. J. Heat Mass Transfer* **13**, 471–490.
31. Cheng, K. C., Lin, R. C., and Ou, J. W. (1975). Graetz problem in curved square channels. *J. Heat Transfer* **97**, 244–248.
32. Joseph, B., Smith, E. P., and Adler, R. J. (1975). Numerical treatment of laminar flow in helically coiled tubes of square cross section. *AIChE J.* **21**, 965–974.

33. Cheng, K. C., Nakayama, J., and Akiyama, M. (1970). Effect of finite and infinite aspect ratios on flow patterns in curved rectangular channels. *In* ''Flow Visualization''. pp. 181–186. Hemisphere.
34. Akiyama, M. (1969). Laminar forced convection heat transfer in curved rectangular channels, M.Sc. thesis, Mechanical Engineering Department, University of Alberta.
35. Masliyah, J. H. (1980). On laminar flow in curved semicircular ducts. *J. Fluid Mech.* **99**, 469–479.
36. Winters, K. H. and Brindley, R. C. G. (1984). Multiple solutions for laminar flow in helically-coiled tubes. *AERE Rep.* 11373, AERE Harwell, UK.
37. Shanthini, W. and Nandakumar, K. (1986). Bifurcation phenomena of generalized Newtonian fluids in curved rectangular ducts. *J. Non-Newtonian Fluid Mech.* **22**, 35–60.
38. Winters, K. H. (1987). A bifurcation study of laminar flow in a curved tube of rectangular cross-section. *J. Fluid Mech.* **180**, 342–369.
39. Daskopoulos, P. and Lenhoff, A. M. (1989). Flow in curved ducts: bifurcation structure for stationary ducts. *J. Fluid Mech.* **203**, 125–148.
40. Bara, B., Nandakuma, K., and Masliyah, J. H. (1992). An experimental and numerical study of the Dean problem: flow development towards two-dimensional multiple solutions. *J. Fluid Mech.* **244**, 339–376.
41. Cheng, K. C. and Wang, L. Q. (1993). The effects of centrifugal and Coriolis forces on secondary flow phenomena in rotating curved square channels: visualization for positive and negative rotations. *ASME: Experimental and Numerical Flow Visualization* **FED-Vol. 172**, 359–367.
42. Cheng, K. C. and Wang, L. Q. (1994). Visualization of secondary flow phenomena in rotating curved square channels with positive and negative rotations. *Proc. Fourth Triennial International Symposium on Fluid Control, Fluid Measurement, Fluid Mechanics, Visualization and Fluids* **1**, 561–569.
43. Mees, P. A. J., Nandakumar, K., and Masliyah, J. H. (1996). Secondary instability of flow in a curved duct of square-cross section. *J. Fluid Mech.* **323**, 387–409.
44. Mees, P. A. J., Nandakumar, K., and Masliyah, J. H. (1996). Steady spatial oscillations in a curved duct of square cross-section. *Phys. Fluids* **8**, 3264–3270.
45. Wang, L. Q. and Cheng, K. C. (1997). Visualization of flows in curved channels with a moderate or high rotation speed. *Int. J. Rotating Machinery* **3**, 215–231.
46. Sankar, S. R., Nandakumar, K., and Masliyah, J. H. (1988). Oscillatory flows in coiled square ducts. *Phys. Fluids* **31**, 1348–1358.
47. Seydel, R. (1994). ''Practical Bifurcation and Stability Analysis: From Equilibrium to Chaos''. Springer-Verlag, New York.
48. Van Dyke, M. (1978). Extended Stokes series: laminar flow through a loosely coiled pipe. *J. Fluid Mech.* **86**, 129–145.
49. Cvitanović, P. (1984). ''Universality in Chaos''. Adam Hilger, Bristol.
50. Sparrow, C. (1982). ''The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors''. Springer-Verlay, New York.
51. Bergè, P., Manneville, P., and Pomeau, Y. (1980). Intermittency in Rayleigh-Benard Convection. *J. Phys. Lett.* **41**, L341–L345.
52. Morris, W. D. (1981). ''Heat Transfer and Fluid Flow in Rotating Coolant Channels''. Wiley, New York.
53. Wang, L. Q. and Cheng, K. C. (1996). Flow transitions and combined free and forced convective heat transfer in rotating curved channels: the case of positive rotation. *Phys. Fluids* **6**, 1553–1573.
54. Miyazaki, H. (1971). Combined free and forced convective heat transfer and fluid flow in a rotating curved circular tube. *Int. J. Heat Mass Transfer* **14**, 1295–1309.

55. Miyazaki, H. (1973). Combined free and forced convective heat transfer and fluid flow in a rotating curved rectangular tubes. *J. Heat Transfer* **95**, 64–71.
56. Hong, S. W., Morcos, S. M., and Bergles, A. E. (1974). Analytical and experimental results for combined forced and free convection in horizontal tubes. *Proceedings of the Fifth International Heat Transfer Conference* **3**, 154–158.
57. Patankar, S. V. (1980). "Numerical Heat Transfer and Fluid Flow". Hemisphere Publishing Corporation, New York.
58. Wang, L. Q. (1997). Buoyancy-force-driven transitions in flow structures and their effects on heat transfer in a rotating curved channel. *Int. J. Heat Mass Transfer* **40**, 223–235.
59. Wang, L. Q. and Pang, O. (2004). Bifurcation and stability of forced convection in tightly coiled ducts: multiplicity. *Int. J. Heat Mass Transfer* (submitted).
60. Wang, L. Q. and Pang, O. (2004). Bifurcation and stability of forced convection in tightly coiled ducts: stability. *Int. J. Heat Mass Transfer* (submitted).
61. Keller, H. B. (1977). Numerical solution of bifurcation and nonlinear eigenvalue problems. *In* "Applications of Bifurcation Theory" (P. H. Rabinowitz, ed.). Academic, New York.
62. George, A. and Ng, E. (1984). SPARSPAK: Waterloo sparse matrix package user's guide for SPARSPAK-B, Department of Computer Science, University of Waterloo, Research Report CS-84-37.
63. Seydel, R. (1979). Numerical computation of branch point in nonlinear equations. *Numerical Mathematics* **33**, 339–352.
64. Seydel, R. (1983). Branch switching in bifurcation problems for ordinary differential equations. *Numerical Mathematics* **41**, 91–116.
65. Atkin, D. J., Maskell, S. J., and Patrick, M. A. (1980). Numerical prediction of separated flows. *J. Numer. Methods Eng.* **15**, 129–144.
66. Wang, L. Q. and Cheng, K. C. (1995). Flow in curved channels with a low negative rotating speed. *Phys. Rev. E* **51**, 1155–1161.
67. Wang, L. Q. and Cheng, K. C. (2001). Visualization of flows in channels with curvature and rotation. *In* "Processing by Centrifugation" (L. L. Regel and W. R. Wilcox, eds.), pp. 339–360. Kluwer Academic/Plenum Publishers, New York.
68. Pomeau, Y. and Manneville, P. (1980). Intermittent transition to turbulence in dissipative dynamical systems. *Commun. Math. Phys.* **74**, 189–197.
69. Peitgen, H. O., Jürgens, H., and Saupe, D. (1992). "Chaos and Fractals, New Frontiers of Science". Springer-Verlag, New York.

AUTHOR INDEX

Numerals in parentheses following the page numbers refer to reference numbers cited in the text

D

E

F

G

H

N

O

P

R

Y

Z

SUBJECT INDEX

Note: page numbers in *italic* type indicate figures and tables

D

E

F

G

H

I

L

M

Q

R

S