

Foundations and Trends® in
Stochastic Systems
1:2 (2006)

Controlled Markov Chains, Graphs and Hamiltonicity

Jerzy A. Filar

now

the essence of knowledge

Controlled Markov Chains, Graphs, and Hamiltonicity

Controlled Markov Chains, Graphs, and Hamiltonicity

Jerzy A. Filar

*School of Mathematics and Statistics
University of South Australia
Mawson Lakes
SA 5095
Australia
j.filar@unisa.edu.au*

now

the essence of **knowledge**

Boston – Delft

Foundations and Trends[®] in Stochastic Systems

Published, sold and distributed by:

now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:

now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is J. A. Filar, Controlled Markov Chains, Graphs, and Hamiltonicity, *Foundations and Trends[®] in Stochastic Systems*, vol 1, no 2, pp 77–162, 2006

ISBN: 978-1-60198-088-5

© 2007 J. A. Filar

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, mechanical, photocopying, recording or otherwise, without prior written permission of the publishers.

Photocopying. In the USA: This journal is registered at the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by now Publishers Inc, for users registered with the Copyright Clearance Center (CCC). The 'services' for users can be found on the internet at: www.copyright.com

For those organizations that have been granted a photocopy license, a separate system of payment has been arranged. Authorization does not extend to other kinds of copying, such as that for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. In the rest of the world: Permission to photocopy must be obtained from the copyright owner. Please apply to now Publishers Inc., PO Box 1024, Hanover, MA 02339, USA; Tel. +1-781-871-0245; www.nowpublishers.com; sales@nowpublishers.com

now Publishers Inc. has an exclusive license to publish this material worldwide. Permission to use this content must be obtained from the copyright license holder. Please apply to now Publishers, PO Box 179, 2600 AD Delft, The Netherlands, www.nowpublishers.com; e-mail: sales@nowpublishers.com

**Foundations and Trends[®] in
Stochastic Systems**
Volume 1 Issue 2, 2006
Editorial Board

Editor-in-Chief:

David Yao

IEOR Department

Columbia University

New York, NY 10027-6699

USA

yao@columbia.edu

Editors

Venkat Ananthamram (UC Berkeley)

Soren Asmussen (Lund University/
Aarhus University)

Dimitris Bertsimas (MIT)

Jim Dai (GATech)

Paul Dupuis (Brown University)

Paul Glasserman (Columbia

University)

Peter Glynn (Stanford University)

Debasis Mitra (Bell Laboratories)

Richard Weber (Cambridge

University)

Editorial Scope

Foundations and Trends[®] in Stochastic Systems will publish survey and tutorial articles in the following topics:

- Brownian Models
- Logistics and Supply Chain Applications
- Markov Decision Processes
- Mathematical Finance and Insurance
- Point Processes
- Queuing Theory
- Reliability Theory
- Risk Analysis
- Simulation
- Stochastic Control
- Stochastic Networks
- Stochastic Optimization
- Stochastic Programming
- Telecom and Internet Applications

Information for Librarians

Foundations and Trends[®] in Stochastic Systems, 2006, Volume 1, 4 issues. ISSN paper version 1551-3106. ISSN online version 1551-3092. Also available as a combined paper and online subscription.

Controlled Markov Chains, Graphs, and Hamiltonicity

Jerzy A. Filar

*School of Mathematics and Statistics, University of South Australia,
Mawson Lakes, SA 5095, Australia, j.filar@unisa.edu.au*

Abstract

This manuscript summarizes a line of research that maps certain classical problems of discrete mathematics — such as the Hamiltonian Cycle and the Traveling Salesman Problems — into convex domains where continuum analysis can be carried out. Arguably, the inherent difficulty of these, now classical, problems stems precisely from the discrete nature of domains in which these problems are posed. The convexification of domains underpinning the reported results is achieved by assigning probabilistic interpretation to key elements of the original deterministic problems.

In particular, approaches summarized here build on a technique that embeds Hamiltonian Cycle and Traveling Salesman Problems in a structured singularly perturbed Markov Decision Process. The unifying idea is to interpret subgraphs traced out by deterministic policies (including Hamiltonian Cycles, if any) as extreme points of a convex polyhedron in a space filled with randomized policies.

The topic has now evolved to the point where there are many, both theoretical and algorithmic, results that exploit the nexus between

graph theoretic structures and both probabilistic and algebraic entities of related Markov chains. The latter include moments of first return times, limiting frequencies of visits to nodes, or the spectra of certain matrices traditionally associated with the analysis of Markov chains. Numerous open questions and problems are described in the presentation.

Contents

1	Embedding of a Graph in a Markov Decision Process	1
1.1	Introduction	1
1.2	A Graph and A Markov Decision Process	5
1.3	Perturbed Embedding Γ_ε	10
1.4	Background from Markov Chains and MDPs	14
2	Analysis in the Policy Space	19
2.1	HCP and Minimal Variance of Hitting Times	19
2.2	Determinant of the Inverse of Fundamental Matrix	34
3	Analysis in the Frequency Space	39
3.1	HCP in the Frequency Space of the Long-Run Average MDP	39
3.2	An Interior-Point Method in Frequency Space	47
3.3	HCP in the Frequency Space of the Discounted MDP	56
3.4	Branch and Fix Methods in the Frequency Space	60
3.5	A Fast Heuristic	64
4	Spectral Properties, Spin-offs, and Speculation	67
4.1	Introduction	67

4.2	Fractal-like Structure of Regular Graphs	69
4.3	Conclusions and Some Speculative Comments	74
	Acknowledgments	79
	References	81

1

Embedding of a Graph in a Markov Decision Process

1.1 Introduction

Arguably, the inherent difficulty of many problems of discrete mathematics and combinatorial optimization stems, precisely, from the discrete nature of the domains in which these problems are posed. This manuscript is devoted to a line of research that maps such problems into convex domains where continuum analysis can be easily carried out. This convexification of domains is achieved by assigning probabilistic interpretation to the key elements of the original problems even though these problems are deterministic.

While there are probably other instances of similar ideas being exploited elsewhere, our approach builds on the innovation introduced in Filar and Krass [35] where the Hamiltonian Cycle and the Traveling Salesman Problems were embedded in a structured singularly perturbed Markov Decision Process (MDP, for short). The unifying idea of [35] was to interpret subgraphs traced out by deterministic policies (including Hamiltonian Cycles, if any) as extreme points of a convex polyhedron in a space filled with randomized policies.

This approach was continued by Chen and Filar [20]¹ and, independently, by Feinberg [32, 33]. Further results were obtained by Filar and Liu [37], Andramonov et al. [7], Filar and Lasserre [36], Ejov et al. [24, 25, 26, 27, 28, 29, 30, 31] and in Borkar et al. [16, 17]. Thus there is now an active group of mathematicians in a number of countries interested in this approach to discrete problems. Majority of these contributions focused on the classical Hamiltonian Cycle Problem but, in principle, many of the techniques used could be adapted to other problems of discrete mathematics (as, indeed, was done by Feinberg in [33]).

The essence of the Hamiltonian Cycle Problem (HCP, for short) is contained in the following — deceptively simple — single sentence statement: *given a graph on N nodes, find a simple cycle that contains all vertices of the graph (Hamiltonian Cycle (HC)) or prove that HC does not exist.* The HCP is known to be NP-hard and has become a challenge that attracts mathematical minds both in its own right and because of its close relationship to the famous *Traveling Salesman Problem* (TSP). An efficient solution of the latter would have an enormous impact in operations research, optimization, and computer science. However, from a mathematical perspective the underlying difficulty of the TSP is, perhaps, hidden in the Hamiltonian Cycle Problem. Hence we focus on the latter.

Just to indicate the flavor of the results reported in this survey, consider a key observation that led to the recent results presented in Borkar et al. [16, 17]. Namely, that: *the “correct” convex domain where Hamiltonian Cycles should be sought, is the set \mathcal{DS} of doubly stochastic matrices² induced by a given graph.*

The above observation is nearly obvious, once we recall the famous (and nonobvious) Birkhoff-von Neumann Theorem which states that the set of all $N \times N$ doubly stochastic matrices is the convex hull of permutation matrices. Of course, in searching for a Hamiltonian Cycle of a given graph we need to restrict ourselves to the convex hull of only those permutation matrices that correspond to subgraphs of that

¹Despite the fact that [20] appeared before [35], the latter preceded [20].

²A square nonnegative matrix is doubly stochastic if both its row-sums and column-sums are equal to 1.

graph. Results in [16, 17] imply that after a suitable perturbation and defining the random variable

$\tau_1 :=$ the first hitting time of the home node 1 (after time 0),

the Hamiltonian Cycle Problem essentially reduces to “merely” minimizing the variance-like functional

$$E[(\tau_1 - N)^2]$$

over the space \mathcal{DS} . This probabilistic, almost statistical, interpretation should permit us to bring to bear a wide range of both analytical and algorithmic tools on the HCP.

Thus the theoretical aim of this line of research is to explain the essential difficulty of the Hamiltonian Cycle Problem — that is, its NP-hardness — in analytic terms such as a measure of variability, or the size of a gap between certain optimization problems, or by the nature of certain singularities.

The algorithmic aim of these studies is to construct a general purpose heuristic algorithm for the HCP and is based on the belief that some classical “static” optimization problems can be analyzed by embedding them in suitably constructed Markov Decision Processes.

In our setting, the theoretical and algorithmic aims are not separate. Indeed, results on one of these aims seem to influence progress on the other. For instance, the heuristic algorithm in Ejev et al. [24] follows directly from [35] and [20] but has identified difficulties that some of the theoretical developments in [16] are trying to resolve.

The general approach constitutes one of the few instances where probabilistic, continuous optimization and dynamic control methods are combined to deal with a hard problem of discrete mathematics. Arguably, simulated annealing could be seen as a precursor of this approach. However, it should be mentioned that relationships between Markov chains and graphs are also of recent interest to other researchers; notably Aldous and Fill [4] and Hunter [44].

Many of the successful classical approaches of discrete optimization to the HCP and TSP focus on solving a linear programming “relaxation” followed by heuristics that prevent the formation of sub-cycles

(e.g., see Lawler et al. [52]). In the present approach, we embedded a given graph in a singularly perturbed MDP in such a way that we can identify Hamiltonian Cycles with irreducible Markov chains and sub-cycles with non-exhaustive ergodic classes. This permitted a search for a Hamiltonian Cycle in the frequency space of an MDP that is a polytope with a nonempty interior, thereby converting the original discrete problem to a continuous one.

Next we shall, briefly, differentiate between our approach and some of the best known, well established, approaches to the HCP.

We first note that the present line of research is essentially different from that adopted in the study of *random graphs* where an underlying random mechanism is used to generate a graph (eg., see Karp’s seminal paper [47]). In our approach, the graph that is to be studied is given and fixed but a *controller* can choose arcs according to a probability distribution and with a small probability (due to a perturbation) an arc may take you to a node other than its “head.” Of course, random graphs played an important role in the study of Hamiltonicity, a striking result to quote is that of Robinson and Wormald [62] who showed that *with high probability* k -regular graphs³ are Hamiltonian for $k \geq 3$.

Typical general purpose heuristic algorithms can, perhaps, be classified (we cite only few representative papers) as *rotational transformation* algorithms Posa [60], *cycle extension* Bollobas et al. [14], *long path* algorithms [50], *low degree vertices* algorithms Broder et al. [18], Brunacci [19], *multipath search* Kocay and Li [50], and *pruning* algorithms Christofides [21]. Of course, much research has been done on algorithms for finding a Hamiltonian Cycle on various restricted graph classes (e.g., see Parberry [58]). Clearly, algorithms designed for particular classes of graphs tend to outperform the best general purpose algorithms when applied to graphs from these classes.

Finally, the reported results open up many natural directions for further investigation. The recently implemented heuristic interior-point algorithm (see Ejoy et al. [24] and Section 3.2) is based on the cited

³Namely, graphs where the in-degree and the out-degree at every node is equal to k .

stochastic embedding and is performing competitively with alternative — general purpose — algorithms on various test problems including the “Knight’s tour” problem on chessboards of the size up to 32×32 . See also the “Branch and Fix” heuristics of Sections 3.4 and 3.5.

1.2 A Graph and A Markov Decision Process

Consider a directed graph G with the node set S and the arc set \mathcal{A} . We can associate a Markov Decision Process Γ with the graph G as follows:

- The set of N nodes is the finite state space $S = \{1, 2, \dots, N\}$ and the set of arcs in G is the total action space $\mathcal{A} = \{(i, j), i, j \in S\}$ where, for each state (node) i , the action space is the set of arcs (i, j) emanating from this node and will be denoted by $\mathcal{A}(i)$.
- $\{p(j|i, a) = \delta_{aj} | a = (i, j) \in \mathcal{A}(i), i, j \in S\}$, where δ_{aj} the Kronecker delta, is the set of (one-step) transition probabilities. Note that, we are adopting the convention that a equals to both arc (i, j) and its “head” j , whenever there is no possibility of confusion as to the “tail” i .

A *stationary policy* f in Γ is a set of N probability vectors $f(i) = (f(i, 1), f(i, 2), \dots, f(i, N))$, where $f(i, k)$ denotes the probability of choosing an action k (arc emanating from i to k) whenever state (node) i is visited. Of course, $\sum_{k=1}^N f(i, k) = 1$ and if the arc $(i, k) \notin \mathcal{A}(i)$, then $f(i, k) = 0$. Equivalently, it will be sometimes convenient to represent a policy f as an $N \times N$ matrix whose (i, k) th entry is $f(i, k)$. The set of all stationary policies will be denoted by \mathcal{F} .

A *deterministic policy* f is simply a stationary policy that selects a single action with probability 1 in every state (hence, all other available actions are selected with probability 0). That is, $f(i, k) = 1$ for some $(i, k) \in \mathcal{A}(i)$. For convenience, we will write $f(i) = k$ in this case. The set of all deterministic policies will be denoted by \mathcal{D} .

It is easy to check that any stationary policy $f \in \mathcal{F}$ induces a probability transition matrix

$$P(f) = [p(j|i, f)], \quad i, j = 1, \dots, N,$$

where for all $i, j \in S$

$$p(j|i, f) = \sum_{a=1}^N p(j|i, a) f(i, a).$$

In the above summation, we assume that $p(j|i, a) = 0$ if the arc $(i, a) \notin \mathcal{A}$.

A *doubly stochastic policy* $f \in \mathcal{F}$ is one which induces a probability transition matrix $P(f)$ that is doubly stochastic; namely all of its rows and columns sum to unity. The set of all doubly stochastic policies will be denoted by \mathcal{DS} . It should be clear from the construction that

$$\mathcal{DS} \subseteq \mathcal{F}.$$

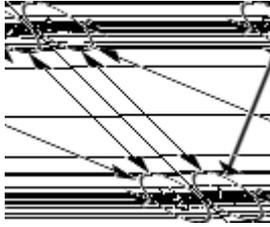
Assume now that 1 is the initial state (home node). We shall say that a deterministic policy f in Γ is a Hamiltonian Cycle (HC) (or simply “is Hamiltonian”) in G if the sub-graph G_f with the set of arcs $\{(1, f(1)), (2, f(2)), \dots, (N, f(N))\}$ is a HC in G .

If an analogous sub-graph G_f induced by a deterministic policy f contains cycles of length less than N , say m , we say that f has an *m-sub-cycle*.

However, such a straightforward identification of G with Γ leads to an inevitable difficulty of confronting multiple ergodic classes induced by various deterministic policies.

Note that if $f \in \mathcal{DS} \cap \mathcal{D}$, then the Markov chain induced by f corresponds to either a Hamiltonian Cycle or to a policy tracing out a union of disjoint sub-cycles in the graph G .

Example 1.1. All of the above can be illustrated on a complete graph G_4 on 4 nodes (without self-loops) in Figure 1.1. A policy f_1 such that $f_1(1) = 2$, $f_1(2) = 3$, $f_1(3) = 4$, and $f_1(4) = 1$ induces a sub-graph $G_{f_1} = \{(1, 2), (2, 3), (3, 4), (4, 1)\}$ that is a Hamiltonian Cycle. Policy f_1

Fig. 1.1 Complete graph G_4 .

also induces a Markov chain with the probability transition matrix

$$P(f_1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

that has a single ergodic class containing all four states. A policy f_2 such that $f_2(1) = 2$, $f_2(2) = 1$, $f_2(3) = 4$, and $f_2(4) = 3$ induces a sub-graph $G_{f_2} = \{(1, 2), (2, 1), (3, 4), (4, 3)\}$ which contains two 2-sub-cycles (see Figure 1.2). Policy f_2 also induces a Markov chain with the probability transition matrix

$$P(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

which has two ergodic classes corresponding to the sub-cycles of G_{f_2} .

Fig. 1.2 Sub-graph G_{f_2} .

Of course, randomized stationary policies can be regarded as convex combinations of deterministic policies. For instance, if in this example we take a policy f_3 that is a set of 4 probability vectors

$$\begin{aligned} f_3(1) &= (f_3(1,2), f_3(1,3), f_3(1,4)) = (1, 0, 0), \\ f_3(2) &= (f_3(2,1), f_3(2,3), f_3(2,4)) = (0.8, 0.2, 0), \\ f_3(3) &= (f_3(3,1), f_3(3,2), f_3(3,4)) = (0, 0, 1), \text{ and} \\ f_3(4) &= (f_3(4,1), f_3(4,2), f_3(4,3)) = (0.2, 0, 0.8), \end{aligned}$$

then it is clear that $f_3 = 0.2f_1 + 0.8f_2$ which induces the Markov chain probability transition matrix

$$P(f_3) = 0.2P(f_1) + 0.8P(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.8 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 1 \\ 0.2 & 0 & 0.8 & 0 \end{bmatrix}.$$

1.2.1 Classification of Deterministic Policies

We shall now describe a useful partition of \mathcal{D} that is based on the graphs “traced out” in G by deterministic policies. As above, with each $f \in \mathcal{D}$ we associate a sub-graph G_f of G defined by

$$(i, j) \in G_f \iff f(i) = j.$$

We shall also denote a simple cycle of length m and beginning at 1 by a set of arcs

$$c_m^1 = \{(i_1 = 1, i_2), (i_2, i_3), \dots, (i_m, i_{m+1} = 1)\}, \quad m = 2, 3, \dots, N.$$

Note that c_N^1 is a HC. If G_f contains a cycle c_m^1 , we write $G_f \supset c_m^1$. Let

$$C_m^1 := \{f \in \mathcal{D} | G_f \supset c_m^1\},$$

namely, the set of deterministic policies that trace out a simple cycle of length m , beginning at node 1, for each $m = 2, 3, \dots, N$. Of course, C_N^1 is the (possibly empty) set of policies that correspond to HCs and any single C_m^1 can be empty depending on the structure of the original



Fig. 1.3 A typical policy f in C_m^1 .

graph G . Thus, a typical policy $f \in C_m^1$ traces out a graph G_f in G that might look as Figure 1.2 where the dots indicate the “immaterial” remainder of G_f in the sense that it corresponds to states/nodes that will never be observed if the process begins at node 1 and the policy f is adhered to. The broken arrows indicate a sequence of one or more arcs (Figure 1.3). We now introduce the partition of the deterministic policies of the form:

$$\mathcal{D} = \left[\bigcup_{m=2}^N C_m^1 \right] \cup \mathcal{N}_c,$$

where \mathcal{N}_c contains all deterministic policies that are not in any of the C_m^1 s. A typical policy $f \in \mathcal{N}_c$ traces out a sub-graph G_f in G as in Figure 1.4, where the dots again denote the immaterial part of G_f . We shall call policies in \mathcal{N}_c *noose cycles*. For many operations related to Markov chains induced by deterministic policies properties of interest to us will be invariant under permutations of states/nodes that leave the home node unchanged. Thus unless stated otherwise, and without loss of generality, it is sufficient to consider only $f_m \in C_m^1$ tracing out the graph in Figure 1.5 as the representative of the whole class C_m^1 and also, $f_m^k \in \mathcal{N}_c$ that traces out Figure 1.6 as the representative of the entire class \mathcal{N}_c .



Fig. 1.4 A typical policy f in \mathcal{N}_c .

Fig. 1.5 A representative f_m of the whole class C_m^1 .Fig. 1.6 A representative f_m^k of the whole class \mathcal{N}_c .

1.3 Perturbed Embedding Γ_ε

We have seen that the direct embedding of G in Γ , in general, induces a multi-chain ergodic structure. This and some other technical difficulties would vanish if we force the MDP to be *unichain*. The latter is a Markov Decision Process in which every stationary policy induces a Markov chain containing only a single ergodic class plus a (possibly empty) class of transient states. A *completely ergodic* MDP is a unichain MDP in which the class of transient states is empty no matter which stationary policy induces the Markov chain.

There are many possible ways of perturbing the MDP Γ discussed in the preceding section to obtain a parameterized family of perturbed unichain or completely ergodic MDP's Γ_ε , where $\varepsilon \in (0, 1)$ will be called the *perturbation parameter*. However, all these perturbations share the characteristic of altering the ergodic structure of Markov chains induced by various stationary policies. Hence, they are so-called *singular perturbations*.

The question of what constitutes the “best” perturbation is potentially very interesting if the notion of optimality for such perturbations were formalized. In the results reported here three (generic) perturbations were considered.

1.3.1 The Symmetric Linear Perturbation

This is achieved by passing to a singularly perturbed MDP Γ_ε^s , that is obtained from Γ by introducing perturbed transition probabilities

$\{p_\varepsilon(j|i, a) \mid (i, j) \in \mathcal{A}, i, j \in S\}$, where for any $\varepsilon \in (0, \frac{1}{N-1})$

$$p_\varepsilon(j|i, a) := \begin{cases} 1 - (N - 1)\varepsilon & \text{if } a = j, \\ \varepsilon & \text{if } a \neq j. \end{cases}$$

Note that this perturbation ensures that every $f \in \mathcal{D}$ induces a Markov chain with a completely ergodic transition probability matrix $P_\varepsilon(f)$ whose dominant terms coincide with the 1-entries of the corresponding unperturbed probability transition matrix $P(f)$ that the same policy f induces in Γ .

For instance, if in the example of the preceding section we consider the policy f_2 that traces out the two sub-cycles depicted in Figure 1.2, it is clear that in Γ_ε^s the same policy induces the probability transition matrix

$$P_\varepsilon(f_2) = \begin{bmatrix} \varepsilon & 1 - 3\varepsilon & \varepsilon & \varepsilon \\ 1 - 3\varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 1 - 3\varepsilon \\ \varepsilon & \varepsilon & 1 - 3\varepsilon & \varepsilon \end{bmatrix},$$

where, for instance, the second entry in the first row can be formally derived as

$$p_\varepsilon(2|1, f_2) = \sum_{a \in \mathcal{A}(1)} p_\varepsilon(2|1, a) f_2(1, a) = p_\varepsilon(2|1, 2) f_2(1, 2) = 1 - 3\varepsilon.$$

Of course, $P_\varepsilon(f_2)$ now has only a single ergodic class and no transient states. The latter is a desirable property but comes at a price of replacing a sparse probability transition matrix with one that is “dense,” in the sense of not having any zero entries.

The above symmetric linear perturbation has been used in [16] and [17] where a perturbation was needed that also preserved double-stochasticity of a probability transition matrix.

1.3.2 The Asymmetric Linear Perturbation

In the sequence of papers that launched this topic (e.g., [20], [35], and [37]) an asymmetric linear perturbation was used. Its goal was not only to eliminate multiple ergodic classes but also to differentiate

the home node from all other nodes and to maintain the sparsity of probability transition matrices induced by deterministic policies. This was achieved by passing to a singularly perturbed MDP Γ_ε^a , that is obtained from Γ by introducing perturbed transition probabilities $\{p_\varepsilon(j|i,a) \mid (i,j) \in \mathcal{A}, i,j \in S\}$, where for any $\varepsilon \in (0,1)$

$$p_\varepsilon(j|i,a) := \begin{cases} 1 & \text{if } i = 1 \text{ and } a = j, \\ 0 & \text{if } i = 1 \text{ and } a \neq j, \\ 1 & \text{if } i > 1 \text{ and } a = j = 1, \\ \varepsilon & \text{if } i > 1, a \neq j \text{ and } j = 1, \\ 1 - \varepsilon & \text{if } i > 1, a = j \text{ and } j > 1, \\ 0 & \text{if } i > 1, a \neq j \text{ and } j > 1. \end{cases}$$

Note that 1 denotes the home node. For each pair of nodes i,j (not equal to 1) corresponding to a (deterministic) arc (i,j) , our perturbation replaces that arc by a pair of “stochastic arcs” $(i,1)$ and (i,j) (see Figure 1.7) with weights ε and $1 - \varepsilon$, respectively. This stochastic perturbation has the interpretation that a decision to move along arc (i,j) results in movement along (i,j) only with probability of $(1 - \varepsilon)$ and with probability ε it results in a return to the home node 1. We emphasize that the perturbation is chosen to ensure that the Markov chain defined by $P_\varepsilon(f)$ contains only a single ergodic class. On the other hand, the ε -perturbed process Γ_ε clearly “tends” to Γ as $\varepsilon \rightarrow 0$, in the sense that $P_\varepsilon(f) \rightarrow P_0(f)$ for every stationary policy f .

For instance, if in the example of the preceding section we consider the policy f_2 that traces out the two sub-cycles depicted in Figure 1.2,



Fig. 1.7 Perturbation of a deterministic action (arc).

it is clear that in Γ_ε^a the same policy induces the probability transition matrix

$$P_\varepsilon(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \varepsilon & 0 & 0 & 1 - \varepsilon \\ \varepsilon & 0 & 1 - \varepsilon & 0 \end{bmatrix}$$

that still has only a single ergodic class, but now has a nonempty class $\{3, 4\}$ containing the two transient states created by this perturbation.

1.3.3 The Asymmetric Quadratic Perturbation

Of course, there are situations where the presence of transient states is undesirable. For that reason, in [29] and [24] the asymmetric perturbation was modified slightly by introducing a quadratic (in ε) term. This was achieved by passing to a singularly perturbed MDP Γ_ε^q , that is obtained from Γ by introducing perturbed transition probabilities $\{p_\varepsilon(j|i, a) \mid (i, j) \in \mathcal{A}, i, j \in S\}$, where for any $\varepsilon \in (0, \frac{1}{\sqrt{N-2}})$

$$p_\varepsilon(j|i, a) := \begin{cases} 1 - (N - 2)\varepsilon^2 & \text{if } i = 1 \text{ and } a = j, \\ \varepsilon^2 & \text{if } i = 1 \text{ and } a \neq j > 1, \\ 1 & \text{if } i > 1 \text{ and } a = j = 1, \\ \varepsilon & \text{if } i > 1, a \neq j \text{ and } j = 1, \\ 1 - \varepsilon & \text{if } i > 1, a = j \text{ and } j > 1, \\ 0 & \text{if } i > 1, a \neq j \text{ and } j > 1. \end{cases}$$

In this instance, in the example of the preceding section we consider the policy f_2 that traces out the two sub-cycles depicted in Figure 1.2, it is clear that in Γ_ε^q the same policy induces the probability transition matrix

$$P_\varepsilon(f_2) = \begin{bmatrix} 0 & 1 - 2\varepsilon^2 & \varepsilon^2 & \varepsilon^2 \\ 1 & 0 & 0 & 0 \\ \varepsilon & 0 & 0 & 1 - \varepsilon \\ \varepsilon & 0 & 1 - \varepsilon & 0 \end{bmatrix},$$

which now has a single ergodic class and no transient states. The fact that this perturbation preserves much of the sparsity of $P(f)$ is easier to see in examples of higher dimensions.

Remark 1.1. It should be clear that the above perturbations link this topic with the well developed field of analytic perturbations of operators and Markov chains. The treatise by Kato [48] is the seminal reference for the latter topic. However, some of the techniques used here are, perhaps, somewhat more in the spirit of the recent developments such as those reported in Avrachenkov et al. [9, 10]. More particularly, in the context of perturbation and sensitivity analysis of MDP's the papers by Schweitzer [65, 66] and Veinott [68] are quite relevant to the results reported here.

1.4 Background from Markov Chains and MDPs

As before, let $f \in \mathcal{F}$ be a stationary policy and $P(f)$ be the corresponding probability transition matrix. By $P^*(f)$ we denote its *stationary distribution matrix*, that is defined as the *limit Cesaro-sum matrix*

$$P^*(f) := \lim_{T \rightarrow \infty} \frac{1}{T+1} \sum_{t=0}^T P^t(f), \quad P^0(f) = I,$$

where I is an $N \times N$ identity matrix. It is well known (e.g., see [13]) that the above limit exists and satisfies the identity

$$P(f)P^*(f) = P^*(f)P(f) = P^*(f)P^*(f) = P^*(f). \quad (1.1)$$

An important special case arises when the Markov chain corresponding to $P(f)$ contains only a single ergodic class. In this case, $P^*(f)$ consists of identical rows, each of which will be denoted by $\pi(f) = (\pi_1(f), \pi_2(f), \dots, \pi_N(f))$, where $\sum_{i=1}^N \pi_i(f) = 1$. Hence, $\pi(f)$ constitutes a probability vector that is often called the *stationary or invariant distribution* of such a Markov chain. It follows from the preceding identity that $\pi(f)$ is a solution of the linear system of equations:

$$\pi(f)P(f) = \pi(f); \quad \pi(f)\mathbf{1} = 1, \quad (1.2)$$

where $\mathbf{1} := (1, 1, \dots, 1)^T$. Indeed, $\pi(f)$ is the unique solution of (1.2).

An even more special but also extremely important case is that of an *irreducible* Markov chain, where $P(f)$ contains only a single ergodic class and no transient states. In this case, the invariant distribution vector $\pi(f)$ is still the unique solution of (1.2) and possesses the additional useful property that $\pi_i(f) > 0 \forall i = 1, 2, \dots, N$.

Another two very important matrices associated with the probability transition matrix $P(f)$ of a Markov chain induced by a policy $f \in \mathcal{F}$ are the *fundamental matrix* $G(f)$ that is defined by

$$G(f) := (I - P(f) + P^*(f))^{-1} = \lim_{\beta \rightarrow 1^-} \sum_{t=0}^{\infty} \beta^t (P(f) - P^*(f))^t, \quad (1.3)$$

and the closely related *deviation matrix*

$$D(f) := G(f) - P^*(f). \quad (1.4)$$

The following identities are well known (e.g., see [13, 45, 49, 61])

$$D(f)P^*(f) = P^*(f)D(f) = 0, \quad (1.5)$$

where the 0 on the right-hand side above is an $N \times N$ matrix with 0s in all entries, and

$$\begin{aligned} (I - P(f))D(f) + P^*(f) - I &= D(f)(I - P(f)) + P^*(f) - I \\ &= P^*(f), \end{aligned} \quad (1.6)$$

$$D(f)\mathbf{1} = \mathbf{0} \quad \& \quad G(f)\mathbf{1} = \mathbf{1}, \quad (1.7)$$

where $\mathbf{1}$ and $\mathbf{0}$ are N -vectors consisting entirely of 1s and 0s, respectively.

Another matrix, also induced by any policy $f \in \mathcal{F}$, that plays an important role in the theory of Markov Decision Processes is the resolvent-like matrix

$$[I - \beta P(f)]^{-1} = \sum_{t=0}^{\infty} \beta^t P^t(f), \quad (1.8)$$

where the parameter $\beta \in [0, 1)$, is frequently called the *discount factor*. Note that this choice of the domain for β ensures that the spectral

radius of $\beta P(f)$ is strictly less than 1, thereby guaranteeing the existence of the above inverse and the power series expansion.

In a traditional, Markov Decision Process setting, there is also a *reward or a cost*, denoted by $r(i, a)$, associated with each state i and action a . The interpretation is that this is the reward/cost associated with action a if that action is selected in state i . However, if actions are being selected in accordance with a stationary policy $f \in \mathcal{F}$, then the “lottery” on the actions available in state i is prescribed by f and hence the expected reward in that state is given by

$$r(i, f) := \sum_{a=1}^N r(i, a)f(i, a), \quad i \in \mathcal{S}. \quad (1.9)$$

This immediately defines the *expected reward/cost vector*, $r(f)$, induced by f the transpose of which is defined by

$$r^T(f) := (r(1, f), r(2, f), \dots, r(N, f)). \quad (1.10)$$

There are now two well-known MDP’s that have been extensively studied in the literature (e.g., see [61] and [38]). They are differentiated by the manner of aggregating the infinite stream of expected rewards/costs induced by a policy⁴ f . The first of these is the so-called *limiting (or long-run) average process (AMD, for short)* where the performance of the policy f is defined by the *value vector*

$$v(f) := P^*(f)r(f), \quad (1.11)$$

whose entries $v(i, f) := [P^*(f)r(f)]_i$ for each $i \in \mathcal{S}$ are simply the long-run average expected rewards induced by f when the process begins in state i .

Analogously, the second process is the so-called *discounted Markov Decision Process (DMD, for short)* where the performance of the policy

⁴Note that in the theory of MDP’s policies can be more general than the stationary policies introduced here. For instance, they may depend on past histories of states and actions. However, for our purposes stationary policies suffice.

f is defined by the *value vector*

$$v^\beta(f) := [I - \beta P(f)]^{-1} r(f), \quad (1.12)$$

whose entries $v^\beta(i, f) := \{[I - \beta P(f)]^{-1} r(f)\}_i$ for each $i \in \mathcal{S}$ are simply the discounted expected rewards induced by f when the process begins in state i .

The optimization problems normally associated with the AMD and DMD processes, respectively, are

$$\max_{f \in \mathcal{F}} v(f), \quad \text{and} \quad \max_{f \in \mathcal{F}} v^\beta(f), \quad (1.13)$$

where the maximization is taken componentwise in the above expressions.

The above optimization problems are well understood and, for most purposes, completely solved. In particular, it is a remarkable fact that in each case there exist deterministic policies $f^0, f_\beta^0 \in \mathcal{D}$ that, respectively, attain the maxima in (1.13), componentwise. Furthermore, if β is sufficiently near 1, there exists a deterministic policy that is simultaneously optimal for both the AMD and DMD processes. While many outstanding researchers contributed to this topic Blackwell's 1962 paper (see [13]) is, perhaps, the authoritative reference. There are also many treatments of this problem in text books (e.g., see [23, 38, 61]).

Remark 1.2. In our embedding of the Hamiltonian Cycle Problem in Markov Decision Processes, in most instances, we use rewards/costs only to differentiate the home node 1 from the other nodes. For that purpose it is sufficient to assume that $r(i, a) \equiv 0$ for all actions/arcs emanating from nodes other than the home node, and that $r(1, a) \equiv 1$ for all actions/arcs emanating from the home node 1. Hence, unless explicitly stated otherwise, we shall assume that

$$r^T(f) = e_1^T = (1, 0, \dots, 0), \quad \forall f \in \mathcal{F}.$$

Remark 1.3. The above notation was developed for the unperturbed MDP Γ , however, whenever we use one of the previously defined

ε -perturbed MDP's, we shall simply add a subscript ε to the relevant quantity. For instance, the probability transition, stationary distribution and fundamental matrices $P(f)$, $P^*(f)$, $G(f)$ are replaced by $P_\varepsilon(f)$, $P_\varepsilon^*(f)$, $G_\varepsilon(f)$, where $\varepsilon > 0$ is the perturbation parameter. Similar indexing by ε will also apply to other quantities. When the choice of the type of perturbation used is clear from context, the perturbed MDP will be denoted simply by Γ_ε .

2

Analysis in the Policy Space

2.1 HCP and Minimal Variance of Hitting Times

Throughout this section (based on results derived in Borkar et al. [16, 17]) we use the, previously introduced, symmetric linear perturbation of the Markov Decision Process; namely Γ_ε^s . Any given policy $f \in \mathcal{F}$ determines the probability distribution of the random variable τ_1 denoting the first hitting time of node 1. Consequently, $E_1^f(\tau_1)$ is the mean recurrence time for node 1, that is, the expected number of steps required to return to node 1 for the first time. Similarly, $E_i^f(\tau_1)$ will denote the mean first passage time from node i to node 1, that is, the expected number of steps to reach node 1 from node i , $i = 2, \dots, N$. Finally, $\text{Var}_i^f(\tau_1)$ will denote the variance of τ_1 , the first hitting time of node 1 from node i , under the probability distribution induced by the policy f .

Since the symmetric linear perturbation applied to $f \in \mathcal{DS}$ preserves double stochasticity and ensures irreducibility of $P_\varepsilon(f)$, the corresponding stationary distribution matrix $P_\varepsilon^*(f) = \frac{1}{N}J$, where J is an $N \times N$ matrix with 1 in every entry. Of course, this implies that $E_1^f(\tau_1) = N$.

Most of this section is devoted to demonstrating that — in the case of symmetric linear perturbation — the Hamiltonian Cycle Problem is

equivalent to the variance minimization problem

$$\min_{f \in \mathcal{DS}} \text{Var}_1^f(\tau_1), \quad (2.1)$$

provided that the perturbation parameter $\varepsilon > 0$ and is sufficiently small.

Before presenting a more rigorous argument, we note that such a result is intuitively reasonable when we consider the possible variability of τ_1 — as captured by its variance $\text{Var}_1^f(\tau_1) = E_1^f[(\tau_1 - N)^2]$, for $f \in \mathcal{DS}$ — for both Hamiltonian and non-Hamiltonian policies. In the former case, it is clear that this variance can be made nearly zero by following a Hamiltonian Cycle because that policy would yield a variance actually equal to zero were it not for the (small) perturbation ε . However, if the policy f is non-Hamiltonian, it is “likely” to be either tracing out short cycles (of length less than N) before returning to node 1, or getting trapped for a long time in some ergodic class not containing node 1. Both of these effects would increase the variability of τ_1 .

Of course, if the interpretation of the HCP as a variance minimization problem is to be useful a formula for evaluating this variance is needed. The next result (see also [16]) shows that, for a doubly stochastic policy, such a variance is closely related to the top left hand corner element of the fundamental matrix induced by that policy. It also supplies a more direct, but also more complex, expression for calculating this variance.

Proposition 2.1. Consider $f \in \mathcal{DS}$ and the induced perturbed probability transition matrix $P_\varepsilon(f)$ as well as the associated fundamental matrix $G_\varepsilon(f)$. The top left element $g_{11}(f)$ of $G_\varepsilon(f)$ can be evaluated as

$$g_{11}(f) = \frac{1}{N^2} \sum_i^N E_i^f[\tau_1] \quad (2.2)$$

$$= \frac{N+1}{2N} + \frac{1}{2N^2} E_1^f[(\tau_1 - N)^2] \quad (2.3)$$

$$= \frac{N+1}{2N} + \frac{1}{2N^2} \left\{ [(I + \hat{P}_\varepsilon(f))(I - \hat{P}_\varepsilon(f))^{-3} P_\varepsilon(f)]_{11} - N^2 \right\}, \quad (2.4)$$

where $\hat{P}_\varepsilon(f)$ is $P_\varepsilon(f)$ with the elements in the first column replaced by zeroes.

Outline of Proof. For notational simplicity we suppress the dependence on the policy f and the permutation parameter ε of most of the quantities of interest; for instance g_{ij} will denote the (ij) th entry of the fundamental matrix $G_\varepsilon(f)$. By Theorem 11.15 of the Grinstead and Snell book [41], we have the mean recurrence time $E_1(\tau_1) = 1/\pi_1 = 1/(1/N) = N$, and by Theorem 11.16 in page 459 of the same book, we have the mean first passage time $E_i(\tau_1) = (g_{11} - g_{i1})/\pi_1 = N(g_{11} - g_{i1})$. Therefore,

$$\begin{aligned} \sum_i^N E_i(\tau_1) &= N + \sum_{i=2}^N N(g_{11} - g_{i1}) = N + N(N-1)g_{11} - N \sum_{i=2}^N g_{i1} \\ &= N + N(N-1)g_{11} - N(1 - g_{11}) = N^2 g_{11}, \end{aligned}$$

with the second last equality obtained by the fact that every row/column of a fundamental matrix sums to unity (e.g., see (1.5)–(1.7)). Hence (2.2) holds.

Let X_m be a random variable representing the current node after m steps,

$$\frac{1}{N^2} \sum_{i=1}^N E_i[\tau_1] = \frac{1}{N^2} E_1 \left[\sum_{m=0}^{\tau_1-1} E_{X_m}[\tau_1] \right] = \frac{1}{2N^2} E_1[\tau_1(\tau_1 + 1)]$$

using Theorem 5.3.4 of [15] and the occupation measure identity in [59] for the first and second equalities, respectively.

Hence,

$$\begin{aligned} g_{11} &= \frac{1}{2N^2} E_1[\tau_1(\tau_1 + 1)] \\ &= \frac{1}{2N^2} \{E_1[(\tau_1 - E_1(\tau_1))^2] + E_1^2(\tau_1) + E_1(\tau_1)\} \\ &= \frac{1}{2} \frac{N+1}{N} + \frac{1}{2N^2} E_1[(\tau_1 - N)^2], \end{aligned}$$

and we obtain (2.3).

For (2.4), we need to show that $E_1(\tau_1^2) = [(I + \hat{P})(I - \hat{P})^{-3}P]_{11}$. First, we have

$$\begin{aligned}
P(\tau_1 = k | X_0 = 1) &= P(X_k = 1, X_i \neq 1, i = 1, \dots, k-1 | X_0 = 1) \\
&= \sum_{j=2}^N P(X_k = 1, X_{k-1} = j, X_i \neq 1, i = 1, \dots, k-2 | X_0 = 1) \\
&= \sum_{j=2}^N P(X_k = 1 | X_{k-1} = j, X_i \neq 1, i = 1, \dots, k-2, X_0 = 1) \\
&\quad \times P(X_{k-1} = j, X_i \neq 1, i = 1, \dots, k-2 | X_0 = 1) \\
&= \sum_{j=2}^N [\hat{P}^{k-1}]_{1j} P_{j1} = [\hat{P}^{k-1}P]_{11}.
\end{aligned}$$

The reader is invited to check that this leads to

$$\begin{aligned}
E_1(\tau_1^2) &= \sum_{k=1}^{\infty} k^2 P(\tau_1 = k | X_0 = 1) = \sum_{k=1}^{\infty} k^2 [\hat{P}^{k-1}P]_{11} \\
&= [(I + \hat{P})(I - \hat{P})^{-3}P]_{11}
\end{aligned}$$

since the spectral radius $\rho(\hat{P}) < 1$, and (2.4) follows. \square

In view of the discussion leading up to (2.1) and Equation (2.3) above it is now clear that the problem of minimizing the variance of τ_1 is also equivalent to minimizing the top left-hand entry, $g_{11}^\varepsilon(f) := [G_\varepsilon(f)]_{11}$, of the fundamental matrix induced by a doubly stochastic policy. That is, the Hamiltonian Cycle Problem is essentially the same as

$$\min_{f \in \mathcal{DS}} [g_{11}^\varepsilon(f)] = \min_{f \in \mathcal{DS}} \left[\frac{1}{2} \frac{N+1}{N} + \frac{1}{2N^2} E_1^f[(\tau_1 - N)^2] \right], \quad (2.5)$$

provided that the perturbation parameter $\varepsilon > 0$ and is sufficiently small.

It is now possible to explicitly calculate the top left-hand corner element of the fundamental matrix induced by a Hamiltonian policy

f_h . This was done in [16], by a direct computation, and is summarized in the next lemma.

Lemma 2.2. For a Hamiltonian policy f_h the value $g_{11}^\varepsilon(f_h) := [G_\varepsilon(f_h)]_{11}$ does not depend on the particular Hamiltonian Cycle represented by f_h and equals

$$g_{11}^\varepsilon(f_h) = [G_\varepsilon(f_h)]_{11} = \frac{1}{N} + \frac{N^2\varepsilon - 1 + (1 - N\varepsilon)^N}{N^2\varepsilon(1 - (1 - N\varepsilon)^N)}.$$

Moreover, the initial terms of the Taylor expansion of $g_{11}^\varepsilon(f_h)$ are

$$g_{11}^\varepsilon(f_h) = \frac{1}{2} \left(\frac{1 + N}{N} \right) + \frac{1}{12} (N^2 - 1)\varepsilon + \frac{1}{24} (N(N^2 - 1))\varepsilon^2 + \mathcal{O}(\varepsilon^3).$$

It follows that for any Hamiltonian policy

$$\text{Var}_1^{f_h}(\tau_1) = 2N^2 \left(g_{11}^\varepsilon(f_h) - \frac{1}{2} \left(\frac{1 + N}{N} \right) \right) = \mathcal{O}(\varepsilon).$$

The above formulae can be easily verified in the case of the following example.

Example 2.1. Let f_h be the standard HC on four nodes: $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$. This Hamiltonian Cycle induces the perturbed probability transition matrix $P_\varepsilon(f_h)$ of the form:

$$P_\varepsilon(f_h) = \begin{bmatrix} \varepsilon & 1 - 3\varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 1 - 3\varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 1 - 3\varepsilon \\ 1 - 3\varepsilon & \varepsilon & \varepsilon & \varepsilon \end{bmatrix}.$$

Since $P_\varepsilon^*(f_h) = 1/4J$, the matrix needed to calculate the fundamental matrix induced by f_h is

$$I - P_\varepsilon(f_h) + \frac{1}{4}J = \begin{bmatrix} 5/4 - \varepsilon & -3/4 + 3\varepsilon & -\varepsilon + 1/4 & -\varepsilon + 1/4 \\ -\varepsilon + 1/4 & 5/4 - \varepsilon & -3/4 + 3\varepsilon & -\varepsilon + 1/4 \\ -\varepsilon + 1/4 & -\varepsilon + 1/4 & 5/4 - \varepsilon & -3/4 + 3\varepsilon \\ -3/4 + 3\varepsilon & -\varepsilon + 1/4 & -\varepsilon + 1/4 & 5/4 - \varepsilon \end{bmatrix}.$$

And, hence, the fundamental matrix is obtained as

$$G_\varepsilon(f_h) = \left(I - P_\varepsilon(f_h) + \frac{1}{4}J \right)^{-1}$$

$$= \begin{bmatrix} \frac{1}{8} \frac{-5+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-3+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-1+12\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-1-4\varepsilon-8\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} \\ \frac{1}{8} \frac{1-4\varepsilon-8\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-5+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-3+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-1+12\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} \\ \frac{1}{8} \frac{-1+12\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{1-4\varepsilon-8\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-5+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-3+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} \\ \frac{1}{8} \frac{-3+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-1+12\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{1-4\varepsilon-8\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} & \frac{1}{8} \frac{-5+20\varepsilon-40\varepsilon^2+32\varepsilon^3}{6\varepsilon-1+16\varepsilon^3-16\varepsilon^2} \end{bmatrix}.$$

The left-hand corner value of the fundamental matrix

$$\frac{1}{8} \frac{-5 + 20\varepsilon - 40\varepsilon^2 + 32\varepsilon^3}{6\varepsilon - 1 + 16\varepsilon^3 - 16\varepsilon^2} = \frac{5}{8} + \mathcal{O}(\varepsilon)$$

agrees with the formula in Lemma 2.2.

The preceding discussion makes it plausible that whenever the underlying graph G is Hamiltonian, the minimum of $g_{11}^\varepsilon(f)$ over \mathcal{DS} is given by

$$g_{11}^\varepsilon(f_h) = \min_{f \in \mathcal{DS}} [g_{11}^\varepsilon(f)] = \frac{1}{2} \frac{N+1}{N} + \mathcal{O}(\varepsilon), \quad (2.6)$$

where f_h is a policy defining any Hamiltonian Cycle in the graph. A proof that this is, indeed, the case can be found in [17].

The remainder of this section is devoted to an argument explaining an equally interesting fact that, for $\varepsilon > 0$ and sufficiently small, there exists $\Delta(N) > 0$ such that whenever the graph G is non-Hamiltonian

$$\min_{f \in \mathcal{DS}} [g_{11}^\varepsilon(f)] - \frac{1}{2} \frac{N+1}{N} = \min_{f \in \mathcal{DS}} \frac{1}{2N^2} \text{Var}_1^f(\tau_1) \geq \Delta(N) - \mathcal{O}(\varepsilon). \quad (2.7)$$

We name the quantity $\Delta(N)$ the *lower bound on Hamiltonicity gap* (of order N) because it distinguishes all non-Hamiltonian graphs with N nodes from all Hamiltonian graphs with the same number of nodes.

2.1.1 A Lower Bound for the Non-Hamiltonian Case

Let f be an arbitrary doubly stochastic policy in a non-Hamiltonian graph. The key step in what follows is the derivation of an upper bound on the probability $P_f(\{\tau_1 = N\})$, that the process returns to node 1, for the first time, after exactly N steps.

Lemma 2.3. Suppose that the graph G is non-Hamiltonian and let f be an arbitrary doubly stochastic policy.

(a) If $\varepsilon = 0$, then

$$P_f(\{\tau_1 = N\}) \leq \frac{1}{4}$$

(b) If $\varepsilon > 0$ and small, then

$$P_f(\{\tau_1 = N\}) \leq \frac{1}{4} + \mathcal{O}(\varepsilon).$$

Proof. First consider the case $\varepsilon = 0$. Let f be an arbitrary doubly stochastic policy and let $\{X_t\}_0^\infty$ be the Markov chain induced by f and the starting state 1. Let $\gamma_1 = (X_0, X_1, \dots, X_N)$ be a path of N steps through the graph and let $\chi_1 = \{\gamma_1 | X_0 = X_N = 1, X_k \neq 1; k = 1, \dots, N-1\}$. That is, the event that the first return to 1 occurs after N steps is simply $\{\tau_1 = N\} = \{\chi_1\}$ and hence

$$P_f(\{\tau_1 = N\}) = \sum_{\gamma_1 \in \chi_1} p_{\gamma_1},$$

where p_{γ_1} denotes the probability (under f) of observing the path γ_1 . However, because the graph is assumed to be non-Hamiltonian, all the paths in χ_1 that receive a positive probability have the structure

$$\gamma_1 = \gamma'_1 \oplus \bar{\gamma}_1,$$

where \oplus denotes a “composition” of two path components: γ'_1 that consists of a non-self-intersecting “reduced path” from 1 to itself of length $m \leq N-2$ adjoined at some node(s) other than 1 by one or more loops of total length $N-m$, that together constitute the second component $\bar{\gamma}_1$. One can think of γ'_1 and $\bar{\gamma}_1$ as the first and second parts of a figure comprising of basic loop with one or more side-lobes attached to it, each of which is either a loop or a connected union of loops. The simplest instance of this is a figure of eight, with two loops of length m and $N-m$, respectively, attached at a node other than 1.

Let p_{γ_1} denote the probability of the original path and $p_{\gamma'_1}$ that of the reduced path. Let $q := p_{\gamma_1}/p_{\gamma'_1} < 1$, which is the contribution to

p_{γ_1} coming from the loops comprising $\bar{\gamma}_1$. More generally, define

$$\gamma_0 = \gamma'_1, \gamma_1 = \gamma'_1 \oplus \bar{\gamma}_1, \gamma_2 = \gamma'_1 \oplus \bar{\gamma}_1 \oplus \bar{\gamma}_1, \gamma_3 = \gamma'_1 \oplus \bar{\gamma}_1 \oplus \bar{\gamma}_1 \oplus \bar{\gamma}_1, \dots$$

The paths γ_n ($n \geq 2$) from 1 to itself all begin with the same reduced path γ'_1 but may repeat exactly the same loop traversals $\bar{\gamma}_1$ path for $n \geq 2$ times all contribute to the event $\{\tau_1 \neq N\}$, as does $\gamma_0 = \gamma'_1$.

The paths γ_n , $n \geq 2$ have probabilities $p_{\gamma_1} q^{n-1}$. The total probability that these paths and $\gamma_0 = \gamma'_1$ (but excluding the original γ_1) contribute to $\{\tau_1 \neq N\}$ is as follows:

$$\begin{aligned} \frac{p_{\gamma_1}}{q} + \sum_{n \geq 2} p_{\gamma_1} q^{n-1} &= p_{\gamma_1} \left(\frac{1}{q} + \frac{q}{1-q} \right) \\ &= p_{\gamma_1} \left(-1 + \frac{1}{q(1-q)} \right) \\ &\geq 3p_{\gamma_1}. \end{aligned}$$

From the above it follows that

$$P_f(\{\tau_1 \neq N\}) \geq \sum_{\gamma_1 \in \chi_1} 3p_{\gamma_1} = 3P_f(\{\tau_1 = N\}).$$

Hence,

$$1 = P_f(\tau_1 < \infty) = P_f(\tau_1 = N) + P_f(\tau_1 \neq N) \geq 4P_f(\tau_1 = N),$$

implying $P_f(\tau_1 = N) \leq \frac{1}{4}$, or, $P_f(\tau_1 \neq N) \geq \frac{3}{4}$.

Returning to the case $\varepsilon > 0$ and small, we note that in the Markov chain induced by f there are now two types of transitions: *strong* that correspond to f assigning a positive probability to arcs that are actually in the graph and *weak* that are strictly the result of our perturbation. The latter are of order ε . Thus the only impact that the perturbation makes on the argument presented above is to introduce an adjustment of order ε . This completes the proof's outline. \square

It is now relatively easy to show that the Hamiltonicity gap $\Delta(N) = \frac{3}{8N^2}$, as is proved in [17]. The key step in the argument is the inequality

$$\begin{aligned} E_1^f[(\tau_1 - N)^2] &= \sum_{k \geq 1} (k - N)^2 P_f(\tau_1 = k) \\ &\geq \sum_{k \geq 1, k \neq N} P_f(\tau_1 = k) = P_f(\tau_1 \neq N). \end{aligned}$$

Hence by Lemma 2.3(b) we obtain

$$E_1^f[(\tau_1 - N)^2] \geq \frac{3}{4} - \mathcal{O}(\varepsilon). \quad (2.8)$$

It now follows from Proposition 2.1 that

$$g_{11}(f) \geq \frac{1}{2} \frac{N+1}{N} + \frac{1}{2N^2} \left(\frac{3}{4} - \mathcal{O}(\varepsilon) \right) = \frac{N+1}{2N} + \Delta(N) - \mathcal{O}(\varepsilon). \quad (2.9)$$

The above demonstrates a lower bound for the non-Hamiltonian case, that — for $\varepsilon > 0$ and sufficiently small — is strictly greater than $g_{11}(f_h) = \frac{1}{2} \frac{N+1}{N} + \mathcal{O}(\varepsilon)$.

2.1.2 Non-Convex Minimization Problem

In view of the preceding discussion, it is clear that the optimization problem (2.6) and hence also the Hamiltonian Cycle Problem are equivalent to the following mathematical programming problem

$$\min[g_{11}^\varepsilon(f)] = \min \left\{ \left[I - P_\varepsilon(f) + \frac{1}{N} J \right]_{1,1}^{-1} \right\} \quad (2.10)$$

subject to

- (i) $\sum_a f(i, a) = 1, \quad i = 1, \dots, N, \quad f(i, a) \geq 0 \quad \forall i, a,$
- (ii) $\mathbf{1}^T P_\varepsilon(f) = \mathbf{1}^T.$

Of course, constraints (i) in the above ensure that f is a proper stationary policy, while constraints (ii) ensure that $P_\varepsilon(f)$ is a doubly stochastic probability transition matrix.

The following theorem summarizes some of the results discussed above (see also Proposition 2.1, Lemmas 2.2 and 2.3).

Theorem 2.4. The problems of minimizing the variance of τ_1 and of $g_{11}^\varepsilon(f)$ over \mathcal{DS} are equivalent in the sense that:

$$\arg \min_{f \in \mathcal{DS}} [g_{11}^\varepsilon(f)] = \arg \min_{f \in \mathcal{DS}} [\text{Var}^f(\tau_1)]. \quad (2.11)$$

Furthermore, for a Hamiltonian graph and $\varepsilon > 0$ and sufficiently small, any solution f^* of the above is a Hamiltonian policy.

Since, in some contexts, variance minimization is a convex optimization problem it is important to stress that this is not the case here, as will be demonstrated below.

Let $P_0 := P_\varepsilon(f_0)$, $P_1 := P_\varepsilon(f_1)$ be two doubly stochastic matrices corresponding to a pair of policies f_0, f_1 in \mathcal{DS} . Also let \tilde{P}_0, \tilde{P}_1 denote the corresponding matrices same with the first row and column removed. Now, define the straight line segment $[f_0, f_1] \in \mathcal{DS}$ by

$$f_\lambda := \lambda f_1 + (1 - \lambda)f_0, \quad \lambda \in [0, 1],$$

and corresponding segments of transition matrices

$$P(\lambda) := \lambda P_1 + (1 - \lambda)P_0, \quad \tilde{P}(\lambda) = \lambda \tilde{P}_1 + (1 - \lambda)\tilde{P}_0$$

for $\lambda \in [0, 1]$. Now set

$$J(\lambda) := g_{11}^\varepsilon(f_\lambda) = \lambda g_{11}^\varepsilon(f_1) + (1 - \lambda)g_{11}^\varepsilon(f_0), \quad \lambda \in [0, 1], \quad (2.12)$$

that is, $J(\lambda)$ is the short-hand notation for the objective function of (2.10) on the line segment $[f_0, f_1] \in \mathcal{DS}$. In [17] the following useful expression was derived for the m th derivative of $J(\lambda)$ on $\lambda \in (0, 1)$.

Lemma 2.5. For $m \geq 1$,

$$\begin{aligned} J^{(m)}(\lambda) &= \frac{d^m J(\lambda)}{d\lambda^m} = \left(\frac{m!}{N^2}\right) \mathbf{1}^T (I - \tilde{P}(\lambda))^{-1} ((\tilde{P}_1 - \tilde{P}_0) \\ &\quad \times (I - \tilde{P}(\lambda))^{-1})^m \mathbf{1}. \end{aligned} \quad (2.13)$$

Now, the expressions for the first and second derivative of $J(\lambda)$ were used in the calculations of the following example illustrating nonconvexity of $g_{11}^\varepsilon(f)$ over \mathcal{DS} .

The *adjacency matrix* A of an N node graph is simply the $N \times N$ matrix with 1's in all ij -entries that correspond to arcs (i, j) present in G and with 0's in all other entries.

Example 2.2. We consider the non-Hamiltonian “Petersen graph” on 10 nodes that has the adjacency matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}. \quad (2.14)$$

Next, we consider a deterministic policy f_0 consisting of two sub-cycles of length 5. In particular, f_0 is equivalent to the map

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \rightarrow \{2, 3, 4, 5, 1, 8, 9, 10, 6, 7\}.$$

Furthermore, let f_1 be equivalent to the map

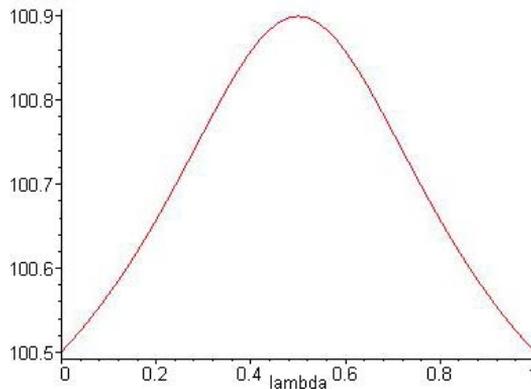
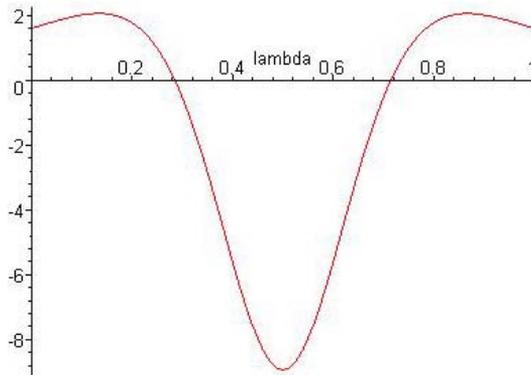
$$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \rightarrow \{5, 1, 2, 3, 4, 9, 10, 6, 7, 8\},$$

namely, f_1 consists of the “reverse sub-cycles” of f_0 .

Clearly, we can evaluate $J(\lambda)$ and its second derivative on the interval $\lambda \in (0, 1)$ in accordance with (2.12) and (2.13), respectively. The plots of these functions are displayed in Figures 2.1 and 2.2. It is clear from the figures that $J(\lambda) = g_{11}^\varepsilon(f_\lambda)$ is not convex.

Despite the above, there is an interesting convex subset of *symmetric doubly stochastic policies* such that our objective function over that subset is strictly convex.

Define $\mathcal{DSS} \stackrel{\text{def}}{=} \{f \in \mathcal{DS} | P_\varepsilon(f) = P_\varepsilon^T(f)\}$, namely, the set of doubly stochastic policies inducing symmetric probability transition matrices.

Fig. 2.1 Function $J(\lambda)$.Fig. 2.2 Function $J''(\lambda)$.

Note that for every $f \in \mathcal{DS}$ the transpose of the corresponding probability transition matrix $P_\varepsilon(f)$ is also doubly stochastic and if the “reverse arc” policy f^r is also in the graph G (as is always the case for undirected graphs), then $P_\varepsilon(f^r) = P_\varepsilon^T(f)$. In such a case $\frac{1}{2}(P_\varepsilon(f) + P_\varepsilon^T(f)) \in \mathcal{DSS}$. Clearly \mathcal{DSS} is a convex compact polytope in \mathcal{DS} .

In [17] it was shown that $g_{11}^\varepsilon(f)$ is strictly convex over \mathcal{DSS} . However, up to this stage, it is not clear whether the problem of

minimization (or maximization) of $g_{11}^\varepsilon(f)$ over that set is related to the original Hamiltonicity problem.

We conclude this section with some remarks concerning the extreme points of the feasible region of the mathematical program (2.10; (i)–(ii)). Recall the partition of all deterministic policies introduced in Section 1.2

$$\mathcal{D} = \left[\bigcup_{m=2}^N C_m^1 \right] \cup \mathcal{N}_c.$$

Since by Birkhoff-von Neumann Theorem, extreme points of the set of all $N \times N$ doubly stochastic matrices are permutation matrices, it follows that the extreme points of \mathcal{DS} are those deterministic policies that lie in $\bigcup_{m=2}^N C_m^1$. Note that these must lie either in C_N^1 , in which case they trace out Hamiltonian Cycles in G , or they lie in $\bigcup_{m=2}^{N-1} C_m^1$, in which case they trace out unions of two or more disjoint cycles in G . The latter we shall refer to as *short cycle policies* the name indicating that the cycle containing node 1 that they trace is of some length k that is less than N .

One important difference between Hamiltonian and short cycle deterministic policies arises in the asymptotic behavior of our objective function $g_{11}^\varepsilon(f)$ as $\varepsilon \rightarrow 0$. We have seen in Lemma 2.2 that, for a Hamiltonian policy f_h , $g_{11}^\varepsilon(f_h)$ has a Taylor's series expansion in ε and that $g_{11}^\varepsilon(f_h) \rightarrow \frac{1+N}{2N}$ as $\varepsilon \rightarrow 0$.

However, in the case of any short cycle doubly stochastic policy f , it was shown in [16] that $g_{11}^\varepsilon(f) \rightarrow \infty$ as $\varepsilon \rightarrow 0$. In particular, the following result holds.

Lemma 2.6. Let $f \in \mathcal{DS}$ be any short cycle deterministic policy; and let $k < N$ be the length of the cycle containing the home node 1. The following properties hold.

- (a) The value $g_{11}^\varepsilon(f)$ depends only on ε and k and equals

$$\begin{aligned} g_{11}^\varepsilon(f) &= \frac{1}{N} + \left(\frac{1}{N} \right) \frac{N - k}{1 - (1 - N\varepsilon)^k} \\ &\quad + \left(\frac{1}{N^2\varepsilon} \right) \frac{kN\varepsilon - 1 + (1 - N\varepsilon)^k}{1 - (1 - N\varepsilon)^k}. \end{aligned}$$

- (b) The functional $g_{11}^\varepsilon(f)$ has a pole of order 1 at $\varepsilon = 0$, and the initial terms of its Laurent expansion are:

$$g_{11}^\varepsilon(f) = \frac{N-k}{N^2k} \varepsilon^{-1} + \frac{2k+kN-N}{2kN} + \frac{N(k-1)(1+k)}{12k} \varepsilon + \frac{N^2(k-1)(1+k)}{24k} \varepsilon^2 + \mathcal{O}(\varepsilon^3).$$

Example 2.3. Let $N = 4$ and f be a short cycle doubly stochastic deterministic policy inducing the probability transition matrix $P_\varepsilon(f)$ that contains two cycles of length 2 (implying that $k = 2$ in Lemma 2.6) and has the form:

$$P_\varepsilon(f) = \begin{bmatrix} \varepsilon & 1-3\varepsilon & \varepsilon & \varepsilon \\ 1-3\varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 1-3\varepsilon \\ \varepsilon & \varepsilon & 1-3\varepsilon & \varepsilon \end{bmatrix}.$$

The corresponding matrix

$$I - P_\varepsilon(f) + \frac{1}{4}J = \begin{bmatrix} 5/4 - \varepsilon & -3/4 + 3\varepsilon & -\varepsilon + 1/4 & -\varepsilon + 1/4 \\ -3/4 + 3\varepsilon & 5/4 - \varepsilon & -\varepsilon + 1/4 & -\varepsilon + 1/4 \\ -\varepsilon + 1/4 & -\varepsilon + 1/4 & 5/4 - \varepsilon & -3/4 + 3\varepsilon \\ -\varepsilon + 1/4 & -\varepsilon + 1/4 & -3/4 + 3\varepsilon & 5/4 - \varepsilon \end{bmatrix}.$$

So, the fundamental matrix

$$\begin{aligned} & \left(I - P_\varepsilon(f) + \frac{1}{4}J \right)^{-1} \\ &= \begin{bmatrix} \frac{1}{16} \frac{8\varepsilon^2-6\varepsilon-1}{(-1+2\varepsilon)\varepsilon} & \frac{1}{16} \frac{8\varepsilon^2+2\varepsilon-1}{(-1+2\varepsilon)\varepsilon} & \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} & \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} \\ \frac{1}{16} \frac{8\varepsilon^2+2\varepsilon-1}{(-1+2\varepsilon)\varepsilon} & \frac{1}{16} \frac{8\varepsilon^2-6\varepsilon-1}{(-1+2\varepsilon)\varepsilon} & \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} & \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} \\ \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} & \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} & \frac{1}{16} \frac{8\varepsilon^2-6\varepsilon-1}{(-1+2\varepsilon)\varepsilon} & \frac{1}{16} \frac{8\varepsilon^2+2\varepsilon-1}{(-1+2\varepsilon)\varepsilon} \\ \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} & \frac{1}{16} \frac{4\varepsilon-1}{\varepsilon} & \frac{1}{16} \frac{8\varepsilon^2+2\varepsilon-1}{(-1+2\varepsilon)\varepsilon} & \frac{1}{16} \frac{8\varepsilon^2-6\varepsilon-1}{(-1+2\varepsilon)\varepsilon} \end{bmatrix}. \end{aligned}$$

The left-hand corner element, $\frac{8\varepsilon^2-6\varepsilon-1}{16\varepsilon(-1+2\varepsilon)}$ that can be checked to agree with the value given by the formula in Lemma 2.6.

Open Problems. In view of the preceding, the following mathematical programming deserves deeper analysis.

$$\min[g_{11}^\varepsilon(f)]$$

(RP)

subject to

$$\begin{aligned} \text{(i)} \quad & \sum_{a \in \mathcal{A}(i)} f(i, a) = 1 \quad \forall i \text{ and } f(i, a) \geq 0 \quad \forall i, a, \\ \text{(ii)} \quad & \sum_{i=1}^N \sum_{a \in \mathcal{A}(i)} p_\varepsilon(j|i, a) f(i, a) = 1 \quad \forall j. \end{aligned}$$

Double stochasticity of policies is ensured by (ii). The objective $g_{11}^\varepsilon(f)$, being an element of the fundamental matrix $G_\varepsilon(f)$, is clearly a rational function of the elements $f(i, a)$ of f . The extreme points of the feasible region are in one-to-one correspondence with those deterministic controls in Γ_ε that trace out sub-graphs in G corresponding to permutation matrices. The latter can be either unions of disjoint cycles (multiple ergodic classes, when $\varepsilon = 0$, with no transient states), or Hamiltonian Cycles.

Next, we present some evidence that the rational program (RP) offers interesting opportunities for both theoretical and algorithmic advances even though its objective functional is nonconvex in the underlying decision variables $f(i, a)$.

It is clear that for every $f \in \mathcal{DS}$ the objective function can be expressed as a Laurent series in ε . In fact, we know that such a series can, at worst, have a simple pole at $\varepsilon = 0$. That is,

$$g_{11}^\varepsilon(f) = \frac{1}{\varepsilon} c_{-1}(f) + c_0(f) + c_1(f)\varepsilon + c_2(f)\varepsilon^2 + \dots$$

However, from the earlier discussion it follows that whenever f is a Hamiltonian policy $c_{-1}(f) = 0$ and $c_0(f) = (1 + N)/2N$. Similarly, $c_{-1}(f) = (N - k)/kN^2$ and $c_0(f) = (2k + kN - N)/2kN$, whenever f corresponds to a permutation matrix whose cycle containing 1 is of length $k < N$. Thus, among the extreme points of \mathcal{DS} , Hamiltonian Cycles are the only ones that do not have a pole at $\varepsilon = 0$.

A number of natural research questions now arise. We list some of the most obvious ones.

- (Q1) Is it possible that, for ε small, local minima of (RP) occur only in special classes of graphs? If so, what is special about these classes of graphs?
- (Q2) Can a special starting point, that is a special doubly stochastic transition matrix, be identified from which there is always a descent path to a Hamiltonian Cycle, if one exists?
- (Q3) Can the probabilistic and MDP insights be used to develop an efficient heuristic for (RP) ?
- (Q4) Is it possible to approximate (RP) by a mathematical program (of lower theoretical complexity) whose objective function value at minimum is so close to that of (RP) to differentiate between a Hamiltonian and a non-Hamiltonian graph?
- (Q5) The previously discussed lower bound on the Hamiltonicity gap, $\Delta(N)$, quantifies the distinction between Hamiltonian and non-Hamiltonian graphs. However, it is a relatively crude estimate that applies to all graphs. Better lower bounds — applying to specific classes of graphs — might form measures of the difficulty of “NP-hardness” of the HCP in these cases. Classes of graphs with a large lower bound (ie., $\Delta(N) \rightarrow 0$ slowly with $N \rightarrow \infty$) could, perhaps, be classified as “ $\Delta(N)$ -easy” and those where the gap tends to 0 rapidly with N could be called “ $\Delta(N)$ -hard.” Such notions are yet to be rigorously formalized.

2.2 Determinant of the Inverse of Fundamental Matrix

In this section, we describe some interesting results that, in principle, could be considered as purely algebraic properties of certain matrices and their eigenvalues and determinants. However, it would have been difficult to discover these results without the insight of preceding sections highlighting the relationship of elements of fundamental matrices of Markov chains to Hamiltonicity of a graph. Consequently, the results

reported below are presented in a manner in which they were first discovered.

The first observation to be made is that for $f \in \mathcal{DS}$ the objective function $g_{11}^\varepsilon(f)$ of the minimization problem (2.6) is, of course, a ratio of two determinants

$$g_{11}^\varepsilon(f) = \frac{\text{cof}([A_\varepsilon(f)]_{11})}{|A_\varepsilon(f)|}, \quad (2.15)$$

where $A_\varepsilon(f) = (I - P_\varepsilon(f) + \frac{1}{N}J) = [G_\varepsilon(f)]^{-1}$, $\text{cof}([A_\varepsilon(f)]_{11})$ is the cofactor of the top-left entry of $A_\varepsilon(f)$ and $|A_\varepsilon(f)|$ is the determinant of that matrix.

Now, since Hamiltonian policies are the minimizers of $g_{11}^\varepsilon(f)$ over \mathcal{DS} , it is not unreasonable to conjecture that they may also be the maximizers of the determinant $|A_\varepsilon(f)|$ over that same space. This and an even stronger results are proved in [26].

In the remainder of this section no perturbation is assumed, equivalently $\varepsilon := 0$ in any of the previous perturbations, and we adopt the notation that for every $f \in \mathcal{F}$

$$A(f) := \left(I - P(f) + \frac{1}{N}J \right) = A_0(f) = \left(I - P_0(f) + \frac{1}{N}J \right).$$

Note that if $P(f)$ is no longer the transition probability matrix of an irreducible doubly stochastic policy, $A(f)$ need no longer equal to $[G_0(f)]^{-1}$, the inverse of the fundamental matrix of the Markov chain induced by f . Despite this, the determinant of $A(f)$ still possesses many interesting properties related to Hamiltonicity. These properties are first illustrated with the help of a simple example.

Example 2.4. As in Section 1.2 consider a complete graph G_4 on 4 nodes (without self-loops) in Figure 1.1. Take three deterministic policies: f_1 , f_2 and f_3 inducing, respectively, Markov chains with the probability transition matrices

$$P(f_1) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$P(f_2) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

and

$$P(f_3) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Now it is easy to verify that

$$A(f_1) = \left(I - P(f_1) + \frac{1}{4}J \right) = \frac{1}{4} \begin{bmatrix} 5 & -3 & 1 & 1 \\ 1 & 5 & -3 & 1 \\ 1 & 1 & 5 & -3 \\ -3 & 1 & 1 & 5 \end{bmatrix},$$

$$A(f_2) = \left(I - P(f_2) + \frac{1}{4}J \right) = \frac{1}{4} \begin{bmatrix} 5 & -3 & 1 & 1 \\ -3 & 5 & 1 & 1 \\ 1 & 1 & 5 & -3 \\ 1 & 1 & -3 & 5 \end{bmatrix},$$

and

$$A(f_3) = \left(I - P(f_3) + \frac{1}{4}J \right) = \frac{1}{4} \begin{bmatrix} 5 & -3 & 1 & 1 \\ 1 & 5 & -3 & 1 \\ -3 & 1 & 5 & 1 \\ 1 & -3 & 1 & 5 \end{bmatrix}.$$

It is now a simple matter to calculate the determinants of the above matrices. The reader is invited to check that

$$|A(f_1)| = 4, \quad |A(f_2)| = 0, \quad \text{and} \quad |A(f_3)| = 3.$$

It is not a coincidence that $P(f_1)$ corresponds to a Hamiltonian policy and the determinant of $A(f_1)$ is exactly equal to the number of nodes in that longest of all possible simple cycles. Analogously, $P(f_3)$ corresponds to a policy that traces out a 3-cycle and a transient state

and the determinant of $A(f_2)$ is again equal to the number of nodes in the cycle containing the home node. While, $P(f_2)$ may appear to be a “counter-example” to the emerging pattern, it can be excluded because it contains more than one ergodic class. Incidentally, the easiest way to verify that the determinant of $A(f_2)$ vanishes is to multiply it by the column vector $(1, 1, -1, -1)^T$.

As is often the case, a simple example illustrates the salient features of the general situation. Indeed, the following theorem proved in [26] summarizes these features.

Theorem 2.7. Consider an arbitrary graph G and the policy spaces \mathcal{F} and \mathcal{DS} of the associated, unperturbed, Markov Decision Process Γ . Then

- (a) if k is the length of the longest simple cycle in G we have that

$$\max_{f \in \mathcal{F}} |A(f)| \leq k$$

- (b) if G is Hamiltonian, then for any Hamiltonian policy f_h , it follows that

$$|A(f_h)| = \max_{f \in \mathcal{F}} |A(f)| = \max_{f \in \mathcal{DS}} |A(f)| = N,$$

- (c) if G is non-Hamiltonian,

$$\max_{f \in \mathcal{F}} |A(f)| \leq N - 1.$$

Corollary 2.1. Let $\lambda_1, \dots, \lambda_{N-1}$ be the eigenvalues of $P(f)$ with eigenvalue 1 of multiplicity 1 excluded.

- (i) For $f \in \mathcal{DS}$,

$$\prod_{i=1}^{N-1} (1 - \lambda_i) \leq N.$$

(ii) For a non-Hamiltonian graph G ,

$$\prod_{i=1}^{N-1} (1 - \lambda_i) \leq N - 1, \quad \text{for all } f \in \mathcal{DS}.$$

(iii) For a Hamiltonian graph G ,

$$\max_{f \in \mathcal{DS}} \prod_{i=1}^{N-1} (1 - \lambda_i) = N.$$

Remark 2.1. In fact, slightly more general results are proved in [26]. In particular, if for a parameter $\alpha > 0$ we define $A_\alpha(f) := (I - P(f) + \alpha J)$, then the main result of Theorem 2.7 extends to $\max_{f \in \mathcal{F}} |A(f)| \leq k\alpha N$. Hence the version stated above is merely the case of $\alpha = \frac{1}{N}$. Of course, for an arbitrary nonnegative value of α , the connection with the inverse of the fundamental matrix of $P(f)$ disappears completely.

Open Problem. It should be clear from the preceding discussion that there are a number of functionals — stemming from connections with Markov chains — that can be used as an objective function in optimization problems that identify Hamiltonian Cycles (if any) of a given graph. These include: $g_{11}^\varepsilon(f)$, $\text{Var}^f(\tau_1)$, $|A(f)|$, and $\prod_{i=1}^{N-1} (1 - \lambda_i)$. A recent report by Litvak and Ejov [53] shows that the trace $\text{tr}[G(f)]$ of the fundamental matrix can be added to that list. Hence a natural, but still unexplored, question is: *What are the comparative advantages or disadvantages of these alternative functionals?* This question should be investigated both from the algorithmic and theoretical perspectives.

3

Analysis in the Frequency Space

3.1 HCP in the Frequency Space of the Long-Run Average MDP

We now return to the asymmetric linear perturbation (see Section 1.3) and the MDP Γ_ε into which our original graph G has been embedded.

Recall that if $P(f)$ is the transition probability matrix of a Markov chain induced by f that is a Hamiltonian Cycle, then $P(f)$ is irreducible. However, if f traces out sub-cycles in G , then $P(f)$ contains multiple ergodic classes which complicates the analysis of the Markov Decision Process Γ . As mentioned earlier, a class of limiting average Markov Decision Processes that retains most of the desirable properties of the irreducible processes is the so-called “unchained” class.

Recall also that for the asymmetric linear perturbation, for each pair of nodes i, j (not equal to 1) corresponding to a “deterministic arc” (i, j) our perturbation replaces that arc by a pair of “stochastic arcs” $(i, 1)$ and (i, j) , with weights ε and $(1 - \varepsilon)$, respectively, ($\varepsilon \in (0, 1)$). Also, the perturbation has made the home node/state 1 rather special. In particular, the home state always belongs to the single ergodic class of $P_\varepsilon(f)$ for any $f \in \mathcal{F}$. Of course, some other states

could be transient. Finally, note that this perturbation changes Γ to an ε -perturbed Markov Decision Process Γ_ε .

3.1.1 Frequency Space of Γ_ε

We shall now undertake the analysis of the Hamiltonian Cycle Problem in the “frequency space” of the perturbed process Γ_ε . Consider the polyhedral set $X(\varepsilon)$ defined by the constraints

$$\begin{aligned} \text{(i)} \quad & \sum_{i=1}^N \sum_{a \in \mathcal{A}(i)} [\delta_{ij} - p_\varepsilon(j|i, a)] x_{ia} = 0; \quad j \in \mathcal{S}. \\ \text{(ii)} \quad & \sum_{i=1}^N \sum_{a \in \mathcal{A}(i)} x_{ia} = 1. \\ \text{(iii)} \quad & x_{ia} \geq 0; \quad a \in \mathcal{A}(i), \quad i \in \mathcal{S}. \end{aligned}$$

It is well known (e.g., see [38]) that with every $f \in \mathcal{F}$ we can associate the long-run frequency vector $x(f) \in X(\varepsilon)$. This is achieved by defining a map $M : \mathcal{F} \rightarrow X(\varepsilon)$ by

$$x_{ia}(f) = \pi_i(f) f(i, a); \quad f \in \mathcal{F}$$

for each $i \in \mathcal{S}$ and $a \in \mathcal{A}(i)$, where $\pi_i(f) = \pi_i^\varepsilon(f)$ is the i th element of the stationary distribution vector of the perturbed Markov chain transition matrix $P_\varepsilon(f)$, and $f(i, a)$ is the probability of choosing action a in state i . The quantity $x_{ia}(f)$ is interpreted as the *long-run frequency of the state-action pair* (i, a) induced by f which is consistent with the interpretation of $x_i(f) := \sum_{a \in \mathcal{A}(i)} x_{ia}(f) = \pi_i(f)$ as the *long-run frequency of visits to the state/node* i .

Next define a map $\hat{M} : X(\varepsilon) \rightarrow \mathcal{F}$ by

$$f_x(i, a) = \begin{cases} \frac{x_{ia}}{x_i}; & \text{if } x_i = \sum_{a \in \mathcal{A}(i)} x_{ia} > 0 \\ 1; & \text{if } x_i = 0 \text{ and } a = a_1 \\ 0; & \text{if } x_i = 0 \text{ and } a \neq a_1, \end{cases}$$

for every $a \in \mathcal{A}(i)$, $i \in \mathcal{S}$ where a_1 denotes the first available action in a given state according to some ordering. The following result can be found in [35], [38] and [45].

Lemma 3.1.

- (i) The set $X(\varepsilon) = \{x(f) | f \in \mathcal{F}\}$ and will henceforth be called the (long-run) “frequency space” of Γ_ε .
- (ii) For every $x \in X(\varepsilon)$,

$$M(\hat{M}(x)) = x$$

but the inverse of M need not exist.

- (iii) If x is an extreme point of $X(\varepsilon)$, then

$$f_x = \hat{M}(x) \in \mathcal{D}.$$

- (iv) If $f \in \mathcal{D}$ is a Hamiltonian Cycle, then $x(f)$ is an extreme point of $X(\varepsilon)$.

It is, perhaps, interesting to observe that all strategies in a given member of the partition $\mathcal{D} = [\bigcup_{m=2}^N C_m^1] \cup \mathcal{N}_c$ induce the same long-run frequency $x_1(f)$ of visits to the home node 1. This observation is captured in the following proposition which can be found in [20] and [38].

Proposition 3.2. Let $\varepsilon \in (0, 1)$, $f \in \mathcal{D}$, and $x(f)$ be its long-run frequency vector (that is, $x(f) = M(f)$). The long-run frequency of visits to the home state 1 is given by

$$x_1(f) = \sum_{a \in \mathcal{A}(1)} x_{1a}(f) = \begin{cases} \frac{1}{d_m(\varepsilon)}; & \text{if } f \in C_m^1, m = 2, 3, \dots, N \\ \frac{\varepsilon}{1 + \varepsilon}; & \text{if } f \in \mathcal{N}_c, \end{cases}$$

where $d_m(\varepsilon) = 1 + \sum_{i=2}^m (1 - \varepsilon)^{i-2}$ for $m = 2, 3, \dots, N$.

Outline of Proof. This result can be obtained by a direct calculation and the observation that the order of the nodes in a cycle, of length m , starting at the home node is immaterial. Below, we show the logical steps for the generic cycle of length m and $f \in C_m^1$. The case of $f \in \mathcal{N}_c$ is not very different.

and by recalling the definition of $d_m(\varepsilon)$ we have

$$x_1(f) = \pi_1(f) = \frac{1}{d_m(\varepsilon)} \quad \text{for } m = 2, 3, \dots, N.$$

□

The above proposition leads the following characterizations of the Hamiltonian Cycles of a directed graph.

Theorem 3.3.

- (i) Let $f \in \mathcal{D}$ be a Hamiltonian Cycle in the graph G . Then $G_f = c_N^1$, $x(f)$ is an extreme point of $X(\varepsilon)$ and $x_1(f) = \frac{1}{d_N(\varepsilon)}$.
- (ii) Conversely, suppose that x is an extreme point of $X(\varepsilon)$ and that $x_1 = \sum_{a \in \mathcal{A}(1)} x_{1a} = \frac{1}{d_N(\varepsilon)}$, then $f = \hat{M}(x)$ is a Hamiltonian Cycle in G .
- (iii) Hamiltonian Cycles of the graph G are in 1 : 1 correspondence with those points of $X(\varepsilon)$ which satisfy
 - (a) $x_1 = \sum_{a \in \mathcal{A}(1)} x_{1a} = \frac{1}{d_N(\varepsilon)}$.
 - (b) For every $i \in \mathcal{S}$, $x_i = \sum_{a \in \mathcal{A}(i)} x_{ia} > 0$ and $\frac{x_{ia}}{x_i} \in \{0, 1\}$ for each $a \in \mathcal{A}(i)$, $i \in \mathcal{S}$.

Outline of Proof.

- (i) Since f is a Hamiltonian Cycle, by definition, $G_f = c_N^1$. By Lemma 3.1, $x(f)$ is an extreme point of $X(\varepsilon)$, and by Proposition 3.2

$$x_1(f) = \frac{1}{d_N(\varepsilon)}.$$

- (ii) Again from Lemma 3.1, $f_x \in F_D$ and $x_1(f_x) = x_1 = \frac{1}{d_N(\varepsilon)}$ since $x = M(\hat{M}(x))$. Now, by Proposition 3.2, f_x is a Hamiltonian Cycle. □

Remark 3.1. It is, perhaps, significant to note that for all $\varepsilon \in (0, 1)$, $m = 2, 3, \dots, N - 1$

$$\frac{1}{d_m(\varepsilon)} > \frac{1}{d_{m+1}(\varepsilon)} > \frac{\varepsilon}{1 + \varepsilon}.$$

Thus Theorem 3.3 demonstrates that the extreme points x of $X(\varepsilon)$ can be “ranked” according to their values of the linear function $l(x) = \sum_{a \in \mathcal{A}(1)} x_{1a}$. Unfortunately, the Hamiltonian Cycles (if they exist) may attain only the “second lowest” value of $l(x)$, namely, $\frac{1}{d_N(\varepsilon)}$.

Remark 3.2. In view of the above, the challenge presented by our AMD formulation is to find a point in $X(\varepsilon)$ that simultaneously achieves the conditions $x_1 = \frac{1}{d_N(\varepsilon)}$ and $\frac{x_{ia}}{x_i} \in \{0, 1\}$ for each $a \in \mathcal{A}(i)$, $i \in \mathcal{S}$. Once such a point is found $f = \hat{M}(x)$ will trace out a Hamiltonian Cycle in G .

There are a number of ways by tackling the challenge stated in Remark 3.2. Below, we briefly mention two approaches: one that exploits a suitably constructed quadratic function, and one that introduces integer variables.

Let $D = \text{diag}(D_1, D_2, \dots, D_N)$ be a block-diagonal matrix with its i th block equal to D_i for $i = 1, 2, \dots, N$. Suppose that D_i is an $m(i) \times m(i)$ matrix with all diagonal elements equal to 0 and all off-diagonal elements equal to 1 (where $m(i)$ is the cardinality of $\mathcal{A}(i)$), for each $i \in \mathcal{S}$. Of course, D_i equals a 1×1 zero matrix if $m(i) = 1$. Consider the following (indefinite) quadratic program:

$$\min x^T D x$$

subject to

(QP)

$$\begin{aligned} \text{(i)} \quad & x \in X(\varepsilon) \\ \text{(ii)} \quad & x_1 = \sum_{a \in \mathcal{A}(1)} x_{1a} = \frac{1}{d_N(\varepsilon)}. \end{aligned}$$

The potential usefulness of this optimization problem stems from the results summarized below.

Theorem 3.4.

- (i) Let $f \in \mathcal{D}$ trace out a Hamiltonian Cycle in G . Then $x(f)$ is a global minimum of (QP) and $x^T D x = 0$.
- (ii) Conversely, let x^* be a global minimum of (QP) such that $(x^*)^T D x^* = 0$. Then $f_{x^*} = \hat{M}(x^*)$ is a deterministic strategy that traces out a Hamiltonian Cycle in G .

Outline of Proof.

- (i) Since $x \geq 0$ and D is a nonnegative matrix, we have that $x^T D x \geq 0$. The reader is invited to verify that the condition $\frac{x_{ia}}{x_i} \in \{0, 1\}$ for each $a \in \mathcal{A}(i)$, $i \in \mathcal{S}$, implies that $x^T D x = 0$.
- (ii) Once again it is not hard to check that $(x^*)^T D x^* = 0$ implies that

$$\frac{x_{ia}^*}{\sum_{a \in \mathcal{A}(i)} x_{ia}^*} \in \{0, 1\},$$

for all $a \in \mathcal{A}(i)$, $i \in \mathcal{S}$. The preceding results now support the validity of the statements of this theorem. \square

Remark 3.3. The reader is invited to verify that each block D_i of dimension $m(i) \geq 2$ has only two eigenvalues: -1 and $(m(i) - 1)$. Thus the optimization problem (QP) is, indeed, an indefinite quadratic program.

Example 3.1. Let G_4 be a complete graph on four nodes with no self-loops, it is easy to check that

$$D = \text{diag}(D_1, D_2, D_3, D_4),$$

where

$$D_i = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

for each $i = 1, 2, 3, 4$.

Further, the quadratic program (QP) can be written in the generic form:

$$\min x^T D x$$

subject to

$$A x = b$$

$$x \geq 0.$$

The vector $b^T = (0, 0, 0, 0, 1, 0.2695)$, and with ε fixed at 0.1. The coefficient matrix A , with the same epsilon value, is

$$\begin{pmatrix} 1 & 1 & 1 & -1 & -0.1 & -0.1 & -1 & -0.1 & -0.1 & -1 & -0.1 & -0.1 \\ -1 & 0 & 0 & 1 & 1 & 1 & 0 & -0.9 & 0 & 0 & -0.9 & 0 \\ 0 & -1 & 0 & 0 & -0.9 & 0 & 1 & 1 & 1 & 0 & 0 & -0.9 \\ 0 & 0 & -1 & 0 & 0 & -0.9 & 0 & 0 & -0.9 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

It is now easy to check that one global optimum of the above quadratic program is attained at

$$\hat{x}^T = (0.2695, 0, 0, 0, 0.2695, 0, 0, 0, 0.2425, 0.2183, 0, 0),$$

which induces the Hamiltonian Cycle via the transformation \hat{M} .

Taking a somewhat different approach, it can be checked that the most awkward requirement $x_{ia}/x_i \in \{0, 1\}$ for all $i \in \mathcal{S}, a \in \mathcal{A}(i)$ is, in this setting, equivalent to $\min\{x_{ia}, x_{ib}\} = 0$ for all $i \in \mathcal{S}, a, b \in \mathcal{A}(i)$ and $a \neq b$. This observation immediately leads to the following mixed integer programming formulation of the HCP:

$$\min \sum_i \sum_a c_{ia} x_{ia}$$

subject to:

(MIP)

$$\begin{aligned} x &\in X(\varepsilon) \\ x_1 &= 1/d_N(\varepsilon) \\ x_{ia} &\leq M y_{ia} : & i \in \mathcal{S}, a \in \mathcal{A}(i) \\ y_{ia} + y_{ib} &\leq 1; & i \in \mathcal{S}, a, b \in \mathcal{A}(i), a \neq b \\ y_{ia} &\in \{0, 1\}; & i \in \mathcal{S}, a \in \mathcal{A}(i). \end{aligned}$$

The above approach was a basis of an algorithm and some numerical experimentation that are reported in [7]. Clearly, in numerical experiments, the above parameter $M \geq 1/d_N(\varepsilon)$ and coefficients c_{ia} of a linear objective function can be experimented with. In preliminary numerical experiments reported in [7], randomly generated problems with up to 100 nodes and 300 arcs were solved in less than 150 cpu seconds on a Sun Workstation.

Remark 3.4. Solving the TSP

It is, perhaps, important to note that if the linear objective function in (MIP) is replaced by

$$\min \sum_i \sum_a c_{ia} y_{ia},$$

where c_{ia} is precisely the cost of arc (i, a) in the Traveling Salesman Problem based on the given graph G , then the solution of the (MIP) , so-modified, also provides a solution of the TSP in its y_{ia} -variables. This is because

$$y_{ia} = 1 \Leftrightarrow x_{ia} > 0$$

and hence the new objective function is precisely the cost of the TSP tour corresponding to the HC identified by the x_{ia} -variables. Since, we are minimizing the objective, this is also the minimal cost tour.

3.2 An Interior-Point Method in Frequency Space

The indefinite quadratic program (QP) introduced in the preceding section presents a challenge in its own right for specialists in mathematical programming. The challenge is to exploit some of the insights gained in the process of embedding the graph G in a Markov Decision Process so as to design an algorithm or a heuristic that might perform better than generic global optimization methods for nonconvex problems.

In this section, we outline a method, developed in [24] that is based on a well known interior-point method.

In order to ensure that our MDP is *irreducible* (rather than just unichain) we replace the asymmetric linear perturbation of the

preceding section with the asymmetric quadratic perturbation defined in Section 1.3 thereby inducing an MDP Γ_ε^q .

Recall, that with the above perturbation, for each pair of nodes i, a (not equal to 1) corresponding to a “deterministic arc” (i, a) , our perturbation replaces that arc by a pair of “stochastic arcs” $(i, 1)$ and (i, a) with weights ε and $(1 - \varepsilon)$, respectively. This stochastic perturbation has the following interpretation: a decision to move along arc (i, a) results in movement along (i, a) only with probability $(1 - \varepsilon)$ and with probability ε the process returns to the home node 1; a decision to move along arc $(1, a)$ results in movement along $(1, a)$ with probability $1 - (N - 2)\varepsilon^2$ and with probabilities ε^2 along the remaining $N - 2$ arcs $(1, a')$ such that $a' \neq a$. The quadratic in ε part of the perturbation ensures that there will arise only *irreducible* Markov chains defined by *stationary strategies* through the above perturbation.

Now, consider the Markov chain $P_\varepsilon(f)$ determined by an arbitrary stationary policy f . Thanks to the irreducibility of all such chains, the map \hat{M} defined in Section 3.3 simplifies somewhat and becomes the inverse of the map M . In particular, as before, $M : \mathcal{F} \rightarrow X(\varepsilon)$ is defined by

$$x_{ia}(f) := \pi_i(f)f(i, a) \quad \forall i \in S, \quad a \in \mathcal{A}, \quad (3.1)$$

where $\pi_i(f)$ is the i th entry of the stationary distribution $\pi(f)$ of that Markov chain.

Furthermore, the quadratic in ε part of the perturbation ensures that $x_i(f) = \sum_{a \in \mathcal{A}(i)} x_{ia}(f) > 0$ for each i . Therefore, M is invertible and its inverse $M^{-1} : X(\varepsilon) \rightarrow \mathcal{F}$ is defined by

$$M^{-1}(x)(i, a) = f_x(i, a) := \frac{x_{ia}}{x_i} \quad \forall i \in S, \quad a \in \mathcal{A}. \quad (3.2)$$

Recall that constraints (i)–(iii) defining the frequency space $X(\varepsilon)$ determine a polyhedron in \mathbb{R}^m (where m is the total number of arcs/actions) that can be represented, in matrix notation, as

$$X(\varepsilon) = \{x | W_\varepsilon x = \mathbf{0}, \mathbf{1}^T x = 1, x \geq \mathbf{0}\},$$

where $\mathbf{0}, \mathbf{1}$ are both m -vectors (with 0, respectively, 1 in every entry), and W_ε is an $N \times m$ matrix with rows numbered by vertices of the

graph G and columns by arcs \mathcal{A} whose $(j, (i, a))$ th entry is

$$w_\varepsilon[j, (i, a)] := \delta(i, j) - p_\varepsilon(j; i, a).$$

Remark 3.5. Construction of $X(\varepsilon)$ provides a connection between W_ε and the incidence $N \times m$ matrix $I(G)$ of the graph defined as

$$I(G)[j, (i, a)] = \begin{cases} 1: & i = j \\ -1: & i = a \\ 0: & \text{otherwise} \end{cases}$$

In particular, W_ε is a $(\varepsilon$ -quadratically) perturbed incidence matrix $I(G)$.

The following result, established in [24], is quite analogous to Theorem 3.3 of the preceding section.

Proposition 3.5. Let $\varepsilon \in \left(0, \frac{1}{\sqrt{N-1}}\right)$. Hamiltonian Cycles of the graph G are in 1 : 1 correspondence with those points of $X(\varepsilon)$ which satisfy

- (i) $x_1 = \sum_{a \in \mathcal{A}(1)} x_{1a} = \frac{1}{\tilde{d}_N(\varepsilon)}$, where

$$\tilde{d}_N(\varepsilon) = \frac{1 - (1 - \varepsilon)^{N-1}}{\varepsilon} + \varepsilon + (1 + (N - 2)\varepsilon)(1 - \varepsilon)^{N-1} = N + O(\varepsilon).$$
 - (ii) For every $i \in S$, $x_i = \sum_{a \in \mathcal{A}(i)} x_{ia} > 0$ and $\frac{x_{ia}}{x_i} \in \{0, 1\}$ for each $a \in \mathcal{A}(i)$, $i \in S$.
-

We now return to the optimization problem (QP) introduced in the preceding section. Recall that the objective function of (QP) was of the form:

$$Q(x) := x^T D x,$$

where $D = \text{diag}(D_1, D_2, \dots, D_N)$ is a block-diagonal matrix whose i th block D_i is an $m(i) \times m(i)$ matrix with 0's on the diagonal and 1's elsewhere (here $m(i)$ is the cardinality of $\mathcal{A}(i)$). Let

$$S(x) := \sum_i x_i^2 = \sum_i \left(\sum_{a \in \mathcal{A}(i)} x_{ia} \right)^2 = x^T J(G)x,$$

where $J(G)$ is the block-diagonal matrix consisting of N blocks with its i th block being an $m(i) \times m(i)$ matrix full of units, and

$$s(x) := \sum_i \sum_{a \in \mathcal{A}(i)} x_{i,a}^2 = x^T I_m x,$$

where I_m is the identity matrix of the size $m = \sum_{i=1}^N m(i)$. Clearly,

$$\begin{aligned} Q(x) &= x^T D x = S(x) - s(x) \\ &= \sum_i \left(\left(\sum_{a \in \mathcal{A}(i)} x_{ia} \right)^2 - \sum_{a \in \mathcal{A}(i)} x_{ia}^2 \right) \\ &= \sum_i \sum_{a, b \in \mathcal{A}(i), a \neq b} x_{i,a} x_{i,b}, \end{aligned}$$

is a difference of two convex functions, and (for nonnegative x) takes on value 0 if and only if at most one of variables x_{ia} is nonzero, for each node $i \in S$. In other words, such an element $x \in X(\varepsilon)$ “chooses” at most one neighbor of node i in graph G . However, because our perturbation ensures that $\sum_{a \in \mathcal{A}(i)} x_{ia} > 0$, $\forall i \in S$, any $x \in X(\varepsilon)$ “chooses” exactly one neighbor of node i .

As in the preceding section, the HCP now reduces to solving the following indefinite quadratic programming problem (QP):

$$\min Q(x) = \min x^T D x$$

subject to

$$\begin{aligned} \text{(i)} \quad & x \in X(\varepsilon) \\ \text{(ii)} \quad & x_1 = \sum_{a \in \mathcal{A}(1)} x_{1a} = \frac{1}{d_N(\varepsilon)}. \end{aligned} \tag{3.3}$$

Again, properties analogous to those of Theorem 3.3 hold, namely, (see [24] for details).

Proposition 3.6.

- (i) A policy $f \in \mathcal{F}$ is a *HC* in G if and only if $x(f) = M(f)$ is a global minimum of (QP) and $x^T(f) D x(f) = 0$.

- (ii) Conversely, if x^* be a global minimum of (QP) such that $(x^*)^T Dx^* = 0$, then $f_{x^*} = M^{-1}(x^*)$ is a deterministic strategy that traces out a *HC* in G .

We note that all equality constraints in (3.3) can be expressed as a single linear equation

$$Ax = b. \tag{3.4}$$

For our heuristic the single objective function $Q(x)$ is insufficient. Our algorithm requires a 1-parameter family of objective functions that (if the graph G contains a *HC*) achieve the (global) minimum at a *HC* or in a small neighborhood of a *HC* so that the *HC* can be recognized by the location of such global minima in $X(\varepsilon)$. It appears that the convex combination

$$f_\alpha(x) := S(x) - \alpha s(x) = \alpha Q(x) + (1 - \alpha)S(x), \quad 0 < \alpha \leq 1$$

provides such a family. The notation $\arg \min(f_\alpha(x))$ will be used to denote the location of a global minimum of $f_\alpha(x)$ in $X(\varepsilon)$.

Lemma 3.7. If G contains a *HC* then for every $\delta > 0$ there exists $\varepsilon_0(\delta, \alpha) > 0$ such that for every ε from $0 < \varepsilon \leq \varepsilon_0(\delta, \alpha)$ a global minimum $\{\arg \min(f_\alpha(x))\}$ lies in the δ -neighborhood of a *HC* in G .

3.2.1 An Interior Point Method

The theory [69] and the implementation [6] of interior point methods for optimization are well understood. Interior point methods have been used extensively in studying combinatorial optimization problems. These methods offer a number of advantages, specially when applied to very large problems. In [24] a specialized interior point algorithm was developed to solve the nonconvex quadratic problem (3.3) with the help of HOPDM solver [5, 39]. Below we outline the main issues of this application.

An interior point algorithm for quadratic programming implemented in HOPDM is the primal-dual method with multiple centrality

correctors. The algorithm is applied to the primal-dual formulation of the quadratic program

$$\begin{array}{ll}
 \text{Primal} & \text{Dual} \\
 \min & c^T x + \frac{1}{2} x^T D x \\
 \text{s.t.} & A x = b, \\
 & x \geq 0; \\
 & \\
 & \max \quad b^T y - \frac{1}{2} x^T D x \\
 & \text{s.t.} \quad A^T y + s - D x = c, \\
 & \quad y \text{ free, } x, s \geq 0,
 \end{array}$$

where $A \in \mathcal{R}^{m \times n}$, $D \in \mathcal{R}^{n \times n}$, $x, s, c \in \mathcal{R}^n$ and $y, b \in \mathcal{R}^m$. The main computational effort of this algorithm consists in the computation of the primal-dual Newton direction. This requires solving the following linear system:

$$\begin{bmatrix} A & 0 & 0 \\ -D & A^T & I \\ \mathcal{S} & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} \xi_p \\ \xi_d \\ \xi_\mu \end{bmatrix}, \quad (3.5)$$

where

$$\begin{aligned}
 \xi_p &= b - A x, \\
 \xi_d &= c - A^T y - s + D x, \\
 \xi_\mu &= \mu \mathbf{1} - X \mathcal{S} \mathbf{1},
 \end{aligned}$$

and X and \mathcal{S} denote $n \times n$ diagonal matrices in which vectors $x, s \in \mathcal{R}^n$ are spread across the diagonals, respectively, and μ is a parameter. After an elimination of

$$\Delta s = X^{-1} \xi_\mu - X^{-1} \mathcal{S} \Delta x,$$

the Newton system is reduced to

$$\begin{bmatrix} -D - X^{-1} \mathcal{S} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \xi_d - X^{-1} \xi_\mu \\ \xi_p \end{bmatrix}. \quad (3.6)$$

The matrix involved in (3.6) is symmetric but indefinite (even for a convex problem when D is positive definite). For the sake of efficiency, in HOPDM implementation [5], the matrix in the reduced Newton system is regularized with diagonal terms R_p and R_d

$$H = \begin{bmatrix} -D - X^{-1} \mathcal{S} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -R_p & 0 \\ 0 & R_d \end{bmatrix}, \quad (3.7)$$

to obtain a quasidefinite one. This allows the use of Cholesky-like factorization in which an $L\Lambda L^T$ decomposition is found with diagonal matrix Λ containing both positive and negative elements.

An approach equivalent to a partial convexification of the objective is utilized in [24]. In particular, the following parameterized objective was used, with $\alpha \in (0, 1)$

$$f_\alpha(x) := S(x) - \alpha s(x).$$

This function is still nonconvex. However, its negative curvature decreases when α is reduced. Lemma 3.7 provides a rationale for the use of $f_\alpha(x)$.

3.2.2 Heuristics Based on the QP Solution

For an arbitrary (nontrivial) problem there may be little chance to solve the nonconvex QP and obtain the global minimum, that is, attain the zero objective value in (3.3). If this happens, we immediately obtain a Hamiltonian Cycle. What is more likely to happen is that one of the numerous local minima of (3.3) is found. Such a solution has at least one set of frequencies for which there are two or more nonzero elements in x_{ia} , $a \in \mathcal{A}(i)$. Typically, when (3.3) corresponding to a large graph is solved the local minimum obtained has many nodes i with such a property. Consequently, we cannot translate the solution into a HC. However, this local solution provides us with a lot of useful information and allows the use of heuristics to find a HC. The heuristics rely on the interpretation of the solution x_{ia} as frequencies of traversing an arc (i, a) in the graph.

3.2.3 Arc Elimination

For a local minimum of the quadratic problem we compute

$$f(i, a) = \frac{x_{ia}}{x_i}, \quad a \in \mathcal{A}(i).$$

These variables can be interpreted as relative frequencies of leaving node i by appropriate arcs originating from i . If $f(i, a)$ is negligible, that is, if $f(i, a) < \delta$ for some prescribed tolerance δ , then the arc (i, a)

is eliminated from the graph as “unlikely” to be a part of a HC. After arc elimination the new quadratic program for the reduced graph is solved and the analysis is repeated. After a number of such reductions and repeated QP local solutions no more arcs satisfy the elimination criteria. If the last solution corresponds to the case that for each node i only one variable $f(i, a)$ is equal to 1, that is, out of all possible outgoing arcs only one is used, then a Hamiltonian Cycle is found. However, if two or more of variables $f(i, a)$ are bounded away from zero, then the method starts branching on these variables.

3.2.4 Branching

Branching is a technique widely used in integer programming. The heuristic in [24] analyzes the solution of the current quadratic program \mathcal{P}_k (corresponding to the reduced graph G_k) and if this solution has any node i with two or more variables which satisfy $f(i, a) \geq \delta$, then it replaces the problem with a set of $|\mathcal{A}(i)|$ new problems. Each of these corresponds to a different reduced graph in which (i, a) is the only arc leaving node i and all remaining arcs which originated from i have been removed.

This way branching forces the use of one particular arc leaving node i . By replacement of the original problem with a tree of sub-problems, branching inevitably increases the computational complexity. In [24] the branching strategy is combined with an arc elimination technique which usually results in a fast reduction of the size of the graph. However, search trees may occasionally grow to very large depths.

The above approach was implemented using HOPDM interior point solver (<http://www.maths.ed.ac.uk/~gondzio/software/hopdm.html>). It was run on the 200 MHz Pentium III PC with Linux operating system.

The authors of [24] tested this heuristic on classes of problems that included:

- (i) randomly generated graphs,
- (ii) Knight’s tour problems.

The first class of problems needs little introduction. In Table 3.1, the sizes of graphs (number of nodes and arcs) and the overall CPU times

Table 3.1 Solution times for randomly generated graphs.

Problem	Nodes	Arcs	Time (CPU secs)
Rand1	25	59	1.48
Rand2	30	72	0.44
Rand3	40	100	3.92
Rand4	50	150	7.92
Rand5	100	293	107.15
Rand6	110	323	12.94
Rand7	120	353	67.23
Rand8	130	392	19.11
Rand9	140	402	147.53
Rand10	150	420	1267.07

in seconds needed to find a HC are reported. The analysis of results collected in Table 3.1 indicates the potential of the approach but also indicates the fast growth of the solution CPU time when the size of the graph increases. The Knight's tour problem consists in finding a tour of the knight to visit each square of the $k \times k$ chessboard exactly once. This problem has received a lot of attention from the research community and a variety of algorithms have been developed for its solutions, see for example [58]. The problem has a solution only for even $k \geq 6$. Table 3.2 gives the solution for the 6×6 chessboard: the numbers in the fields provide the order in which the knight visits them. In Table 3.3 the sizes of graphs and the corresponding solution times for the Knight's tour problems are reported. Thus, it can thus be seen that the approach proposed [24] finds solutions to HC problems of moderate size in acceptable computation time. However, the solution time grows rapidly for a larger graph corresponding to the 32×32 chessboard.

Open Problem. From numerical experiments it appears that, typically, (QP) possesses many interior local minima and that, naturally, makes the goal of finding a global minimum difficult. One, poten-

Table 3.2 Solution for the 6×6 chessboard problem.

4	15	34	27	6	17
35	26	5	16	33	28
12	3	14	29	18	7
25	36	11	32	21	30
10	13	2	23	8	19
1	24	9	20	31	22

Table 3.3 Solution times for the Knight’s tour problem.

Problem	Nodes	Arcs	Time (CPU secs)
Chess6	36	160	1.25
Chess8	64	336	3.35
Chess10	100	576	29.77
Chess12	144	880	33.58
Chess14	196	1248	194.45
Chess20	400	2736	819.10
Chess32	1024	7440	35697.00

tial, remedy might be to exploit the fact that the objective function $Q(x) = S(x) - s(x)$ is a difference of two convex functions, indeed, an “ellipsoid minus a ball.” Hence, in Ejev et al. [29], in lieu of (QP), the convex maximization problem (CM) was considered:

$$\max[s(x)] \quad \text{s.t. } Ax = b \ \& \ S(\mathbf{x}) \leq \theta(\varepsilon), \quad (CM)$$

where $\theta(\varepsilon) := S(x_h)$ and $x_h = M(f_h)$ for any f_h that traces out a Hamiltonian Cycle. This way the HC’s (if any) remain in the feasible region. In [29] it was shown that, with ε sufficiently small, a graph is Hamiltonian if and only if the maximum of $s(x)$ equals $\theta(\varepsilon)$.

Geometrically speaking, at Hamiltonian Cycles, the surface of the ball $s(x) \leq \theta(\varepsilon)$ just touches the boundary of the ellipsoid $S(x) \leq \theta(\varepsilon)$.

A potential advantage of working with (CM) in place of (QP) may be due to the fact that (CM) has no interior local minima. Hence, an opportunity exists to adapt the interior-point heuristic from [24] to (CM).

3.3 HCP in the Frequency Space of the Discounted MDP

In an interesting, related, development Feinberg [32] considered the embedding of a Hamiltonian Cycle Problem in a discounted Markov Decision Process. In that formulation the perturbation parameter ε is not necessary but, instead, the discount factor $\beta \in [0,1)$ plays a crucial role.

In particular Feinberg’s embedding can be obtained either directly as at the beginning of Section 1.2, or from the previously considered perturbed MDP’s by merely setting $\varepsilon = 0$ in $p_\varepsilon(j|i, a)$ as defined earlier

and, as usual, by setting

$$r(i, a) = \begin{cases} 1 & \text{if } i = 1, a \in \mathcal{A}(1) \\ 0 & \text{otherwise.} \end{cases}$$

For any $f \in \mathcal{F}$ the expected discounted reward $v^\beta(i, f)$ is now defined as in Section 1.4. The analysis in [32] is based on the following observation. Let i_m denote the state/node visited at stage m , then an alternative probabilistic expression for the discounted reward starting from node 1 is

$$v^\beta(1, f) = \sum_{m=0}^{\infty} \beta^m P_1^f(i_m = 1), \quad (3.8)$$

where $P_1^f(\cdot)$ denotes the probability measure induced by f and the initial state $i_0 = 1$, and

$$P_1^f(i_m = 1) = \frac{1}{m!} \left[\frac{\partial^m}{\partial \beta^m} (v^\beta(1, f)) \right]_{\beta=0}. \quad (3.9)$$

It now follows immediately from (3.8) that if a policy f traces out a Hamiltonian Cycle, then the home node is visited periodically after N steps and this results in a deterministic sequence of discounted rewards

$$1, \beta^N, \beta^{2N}, \dots, \beta^{mN}, \dots$$

The above observations lead to novel characterizations of Hamiltonian Cycles that are summarized below.

Theorem 3.8. With the embedding in Γ_β described above the following statements are equivalent:

- (i) A policy f is deterministic and a Hamiltonian Cycle in G .
 - (ii) A policy f is stationary and a Hamiltonian Cycle in G .
 - (iii) A policy f is deterministic and $v^\beta(1, f) = (1 - \beta^N)^{-1}$ for at least one $\beta \in (0, 1)$.
 - (iv) A policy f is stationary and $v^\beta(1, f) = (1 - \beta^N)^{-1}$ for $2N - 1$ distinct discount factors $\beta_k \in (0, 1)$; $k = 1, 2, \dots, 2N - 1$.
-

The above characterization naturally leads to a number of mathematical programming formulations of both HCP and TSP that are

described in [32]. There is clearly a need to explore the algorithmic potential of these formulations. However, the remainder of the section is devoted to a discussion that complements the developments of the preceding sections.

3.3.1 Frequency Space of Γ_β

We shall undertake the analysis of the Hamiltonian Cycle Problem in the “frequency space” of the perturbed process Γ_β .

The *discounted frequency space*, $X_\beta := \{x(f) | f \in \mathcal{F}\}$ (induced by stationary policies) consists of vectors $x(f)$ whose entries are *discounted frequencies of the state-action pairs* $(i, a) \in \mathcal{A}(i)$ defined by

$$x_{ia}(f) := \{\gamma[(I - \beta P(f))^{-1}]\}_i f(i, a), \quad (3.10)$$

where the probability vector γ denotes the initial state distribution vector. Since, we want the latter to either correspond to starting in the home node with certainty, or near certainty, we shall consider a specially structured initial distribution. Namely, for $\mu \in (0, \frac{1}{N})$ we define

$$\gamma_i = \begin{cases} 1 - (N - 1)\mu: & \text{if } i = 1 \\ \mu: & \text{otherwise.} \end{cases}$$

The *discounted frequency of the state* i is defined as the aggregate

$$x_i(f) := \sum_{a \in \mathcal{A}(i)} x_{ia}(f).$$

The construction of x in (3.10) defines a map M of the strategy space \mathcal{F}_S into \mathbb{R}^m by

$$M(f) := x(f).$$

It is well known (e.g., see [38]) that for $\mu > 0$ map M is invertible and its inverse M^{-1} is defined by

$$M^{-1}(x)[i, a] = f_x(i, a) := \frac{x_{ia}}{x_i}.$$

Now, part (iii) of Theorem 3.8 can be generalized somewhat for the case of an initial state distribution γ with $\mu > 0$ and “sufficiently small.”

Namely, when $f \in \mathcal{F}$ is Hamiltonian then

$$v^\beta(1, f) = \kappa(\beta, \mu, N) := \frac{(1 - (N - 1)\mu)(1 - \beta) + \mu(\beta - \beta^N)}{(1 - \beta)(1 - \beta^N)}.$$

Note that when $\mu = 0$, the above reduces to $v^\beta(1, f) = (1 - \beta^N)^{-1}$.

This proposition together with the standard conversion of a discounted MDP to a linear program (e.g., see [43, 42, 45, 46] and [38]) suggests Hamiltonian Cycles can be sought among the extreme points of the following subset of the (*discounted*) *frequency space* $X(\beta)$ that is defined by the linear constraints:

$$\sum_{i=1}^N \sum_{a \in A(i)} (\delta_{ij} - \beta p(j|i, a)) x_{ia} = \gamma_j, \quad \forall j \in S. \quad (3.11)$$

$$x_{ia} \geq 0, \quad \forall i \in S, a \in A(i). \quad (3.12)$$

Consider also the additional constraint

$$\sum_{a \in A(1)} x_{1a} = \kappa(\beta, \mu, N). \quad (3.13)$$

Let $\mathcal{F}_\beta = \{x \mid x \text{ satisfies constraints (3.11)–(3.13)}\}$. It is well known that (when $\mu > 0$) every extreme point of $X(\beta)$ corresponds to a deterministic policy via the transformations M and M^{-1} introduced above. Hence, these extreme points must contain exactly N positive entries (one for each node). However, the additional constraint (3.12) in \mathcal{F}_β could introduce one more positive entry in its extreme points.

Let x_e be an extreme point of \mathcal{F}_β . It turns out that if it contains exactly N positive entries, then $f = M^{-1}(x_e)$ is a Hamiltonian policy. However, if x_e contains $N + 1$ positive entries, then $f = M^{-1}(x_e)$ is a randomized policy that is “nearly” a deterministic policy in the sense that a randomization occurs only in one state/node and on only two actions/arcs.

Since finding an extreme point of \mathcal{F}_β that corresponds to a deterministic policy solves the HCP it is important to understand the structure of these “nearly” deterministic extreme points as they are the ones preventing a trivial solution to the HCP.

The next result, proved in Ejev et al. [25], shows that these nearly-deterministic policies must always be convex combinations of short cycle and noose cycle policies.

Theorem 3.9. Let an extreme point x_e of \mathcal{F}_β induce a randomized policy f_α via the transformation $f_\alpha = M^{-1}(x_e)$. Then, the policy f_α is a convex combination of a short cycle policy and a noose cycle policy. That is, $f_\alpha = \alpha f_1 + (1 - \alpha) f_2$, where $f_1 \in \bigcup_{k=2}^{N-1} C_k^1$ and $f_2 \in \mathcal{N}_c$. Furthermore, f_1 and f_2 select the same arc at each node, other than at the single node where the randomization occurs.

Remark 3.6. The above result raises hope that LP-simplex type methods could be developed to find a Hamiltonian Cycle by finding a nonrandomized extreme point of the polytope \mathcal{F}_β . This is because it is now evident that nearly-deterministic extreme points of \mathcal{F}_β must be of a rather specialized type to satisfy the additional constraint (3.13).

3.4 Branch and Fix Methods in the Frequency Space

Theorems 3.8 and 3.9 and properties of extreme points of $X(\beta)$ suggest that one way to search for Hamiltonian Cycles of a graph G is to prevent randomizations from occurring on certain arcs. The latter can be achieved by *branching* at certain nodes in such a way that pairs of arcs that could have been used as “bridges” for randomizations are replaced by single-arc sub-graphs branching from the corresponding nodes. This was first observed in Filar and Lasserre [36] and also used as part of a heuristic in the interior-point method discussed in Section 3.2.

Note that, since our ultimate goal is to find an extreme point $x_e \in \mathcal{F}_\beta$ such that

$$f = M^{-1}(x_e) \in \mathcal{D}$$

we have a number of degrees of freedom in designing an algorithm. In particular, different linear objective functions can be chosen at each

stage of the algorithm, the parameter $\beta \in (0, 1)$ can be adjusted and $\mu \in (0, 1/N)$ can be chosen small but not so small as to cause numerical difficulties.

The logical structure of the “Branch and Fix” (B&F, for short) algorithm presented in Ejev et al. [25] is as follows. A sequence of linear programs is solved — two at each node of the B&F “tree” — with the generic structure:

$$\min L(x) \tag{†}$$

subject to

$$x \in \mathcal{F}_\beta,$$

additional constraints, if any, on arcs fixed earlier.

Step 1 — Initiation. Solve the original LP (†) without any additional constraints and with some choice of the objective function $L(x)$. Let x_0 be an optimal basic feasible solution obtained. Find $f_0 := M^{-1}(x_0)$. If $f_0 \in \mathcal{D}$; stop, the policy f_0 identifies a Hamiltonian Cycle.

Step 2 — Branching. Use the nearly-deterministic policy f_0 to identify a node i and two arcs (i, j_1) and (i, j_2) corresponding to the single randomization in f_0 . If there are d arcs $\{(i, a_1), \dots, (i, a_d)\}$ emanating from node i , construct subgraphs: G_1, G_2, \dots, G_d , where in G_k the arc (i, a_k) is the only arc emanating from node i . These graphs are identical to the “parent” graph G at all other nodes. Note that this process “fixes” an arc in each G_k .

Step 3 — Fixing. It turns out that in many graphs fixing one arc implies that other arcs can also be fixed.¹ A number of simple checks for determining additional arcs that can be fixed are described later in this section.

Step 4 — Iteration. Repeat Step 1 with the LP (†) constructed for the graph at the current node. Note that this node may correspond to G_1, G_2, \dots, G_d , or to a sub-graph constructed from one of these with the help of additional arc fixing.² Infeasibility, or failure to attain certain

¹For instance, this frequently happens in the case of cubic graphs that supplied many of our test examples.

²As is typical with B&B methods, the decisions guiding which branch to select first are important and open to alternative heuristics.

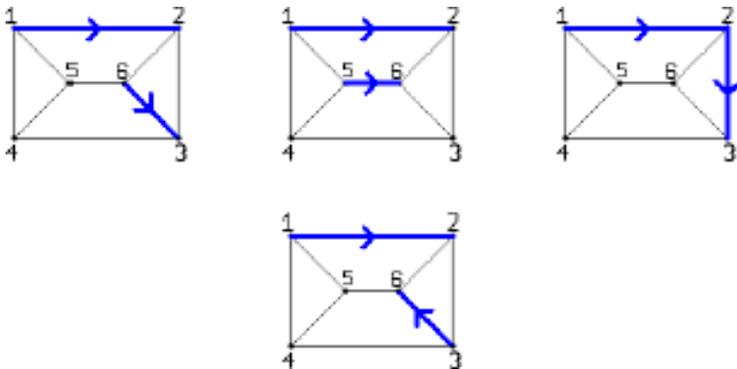
bounds, implies that this branch can be *fathomed* and thus discarded in subsequent iterations.

Remark 3.7. From the more detailed description of the algorithm presented in [25] it can be seen that adaptive bounds are derived that are also used to fathom branches of the B&F “tree.” However, it appears that infeasibility due to arc fixing contributes more to the “pruning” of this tree than failure to attain these bounds. Hence the name “Branch and Fix” seems more appropriate than the more standard “Branch and Bound.”

3.4.1 Arc Fixing Checks

In [25] it was found that in many test problems there are a number of logical checks that enable us to fix additional arcs once a decision is taken to fix one particular arc. This is a consequence of the fact that when there are nodes with low in/out-degrees, fixing one arc may have immediate implications on what arcs go in (or out of some other nodes). This is best illustrated with the help of an example.

Consider the simple 6-node cubic “envelope” graph. The figure below shows the kind of logical additional arc fixing that can arise.



Check 1: Consider the first figure. So far, the only fixed arcs are (1,2) and (6,3). As the only arcs that can go to node 5 are (1,5), (4,5), and (6,5), we must fix arc (4,5) as nodes 1 and 6 already have fixed arcs

going elsewhere. In general, if there is only one arc still available to go to a node, it must be fixed.

Check 2: Consider the second figure. So far, the only fixed arcs are $(1,2)$ and $(5,6)$. The only arcs left going to node 5 are $(4,5)$ and $(6,5)$, but we cannot choose $(6,5)$ as this will create a short cycle, so we must fix arc $(4,5)$. In general, if there are only two arcs available to go to a node, and one will create a loop, we must fix the other one.

Check 3: Consider the third figure. So far, the only fixed arcs are $(1,2)$ and $(2,3)$. As the only arcs that can come from node 6 are $(6,2)$, $(6,3)$, and $(6,5)$, we must fix arc $(6,5)$ as nodes 2 and 3 already have arcs going into them. In general, there is only one arc still available to come from a node, it must be fixed.

Check 4: Consider the fourth figure. So far, the only fixed arcs are $(1,2)$ and $(3,6)$. The only arcs left coming from node 6 are $(6,3)$ and $(6,5)$, but we cannot choose $(6,3)$ as this will create a small cycle, so we must fix arc $(6,5)$. In general, if there are only two arcs available to come from a node, and one will create a loop, the other one must be fixed.

3.4.2 Numerical Results

The above algorithm was implemented in [25] in MATLAB (version 7.0.1.24704) and used CPLEX (version 9.0.0) to solve all the linear programming sub-problems. The algorithm was tested on a range of relatively small graphs. The results reported were encouraging. The number of branches required to solve each of these problems was only a tiny fraction of the maximum possible number of branches.

For example, a 24-node cubic graph has $3^{24} \approx 2 \times 10^{11}$ possible choices for deterministic policies, but typically, the above algorithm finds a HC by examining only some 400 branches. Hamiltonian graphs perform better than non-Hamiltonian graphs. However, even in non-Hamiltonian graphs the algorithm performs rather well. For instance, a 28-node non-Hamiltonian graph has $3^{28} \approx 2 \times 10^{13}$ possible choices for deterministic policies, but the algorithm terminated after only 41268 branches.

Table 3.4 Preliminary results for the refined branch-and-bound method.

Graph	Branches	Upper bound	Time
Hamiltonian $N = 24$	394	2.8243×10^{11}	0:04
Dodecahedron: Hamiltonian $N = 38$	65	1.3509×10^{18}	0:01
8×8 Chessboard problem: Hamiltonian $N = 64$	1097	9.1654×10^{43}	0:38
Petersen: non-Hamiltonian $N = 10$	154	5.9049×10^4	0:01
Coxeter: non-Hamiltonian $N = 28$	41268	2.2877×10^{13}	8:23

Nonetheless, in general, non-Hamiltonian graphs tended to require more branches to solve than Hamiltonian graphs of the same size. This could be explained by the fact that a Hamiltonian graph will end branching immediately when a HC is found. The latter does not happen in non-Hamiltonian graphs.

Some results reported in [25] can be seen in Table 3.4. These include a comparison between the number of branches used by the algorithm and the maximum possible number of branches (upper bound) and the running time in seconds.

3.5 A Fast Heuristic

The results described in the preceding section stimulated a search for a faster heuristic based on the embedding in a discounted MDP. In particular, the “cut” frequency space \mathcal{F}_β can be constrained further, to define a “boxed-cut-frequency space” \mathcal{F}_β^b defined by the linear constraints as shown in Table 3.5.

$$\sum_{i=1}^N \sum_{a \in A(i)} (\delta_{ij} - \beta p(j|i, a)) x_{ia} = \delta_{1j} (1 - \beta^N), \quad \forall j \in S, \quad (3.14)$$

$$\sum_{a \in A(1)} x_{1a} = 1, \quad (3.15)$$

$$\sum_{a \in A(i)} x_{ia} \geq \beta^{N-1} \quad \forall i \neq 1, \quad (3.16)$$

$$\sum_{a \in A(i)} x_{ia} \leq \beta \quad \forall i \neq 1, \quad (3.17)$$

$$x_{ia} \geq 0, \quad \forall i \in S, a \in A(i). \quad (3.18)$$

The reader will note that in the above constraints the new x_{ia} ’s are merely the old x_{ia} ’s, from \mathcal{F}_β multiplied by a scaling factor of $(1 - \beta^N)$.

Table 3.5 Preliminary results for the boxed-cut-frequency heuristic.

Graph	Time
8×8 Knight's Tour problem (64 nodes, 336 arcs)	1 second
Perturbed Horton Graph (94 nodes, 282 arcs)	2 seconds
20×20 Knight's Tour Problem (400 nodes, 2736 arcs)	11 minutes
1000-node Heidelberg Graph (1000 nodes, 3996 arcs)	24 minutes
2000-node Heidelberg Graph (2000 nodes, 7992 arcs)	46 hours

Furthermore, the parameter μ in \mathcal{F}_β has been set to 0, as the constraints (3.15) and (3.16) already ensure that there will at least one arc, at each node, corresponding to a positive x_{ia} , so that the map M^{-1} is well defined. Finally, the “box constraints” (3.16) and (3.17) can be made extremely narrow by choosing β sufficiently near 1.

Of course, as usual, the additional “complementarity constraint”

$$x_{ia}x_{ib} = 0; \quad \forall i \in S, \quad a \neq b \in \mathcal{A}(i) \quad (3.19)$$

must be satisfied by a feasible point $x \in \mathcal{F}_\beta^b$ for $f_x = M^{-1}(x)$ to be Hamiltonian.

One heuristic that was recently experimented with, by the present author and his associates, simply searches for a feasible solution of (3.14)–(3.18) and (3.19) and uses a sophisticated simplex-based platform such as ILOG's OPL-CPLEX. Effectively, this delegates the branching needed to satisfy (3.19) to OPL-CPLEX but the user retains the flexibility of choosing β and a linear objective function. In particular, a choice of β seems to strongly influence the numerical performance.

Preliminary results outperformed all previous algorithmic approaches based on the MDP embedding in the sense that bigger graphs were solved faster. In particular, two well known “difficult” graphs, listed on the University of Heidelberg's web site³ were successfully solved even though the larger, 2000 node problem, took approximately 46 hours to find a Hamiltonian Cycle. On the other hand, the smaller, 1000 node problem, was solved in only 24 minutes even though both the interior-point heuristic of Section 3.2 and the B&F algorithm of Section 3.4 both failed to solve this example.

³See: www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/.

Open Problems. More generally, there are whole classes of recent algorithmic developments in the field of global optimization that are ripe for application to this challenging problem. Prominent among these are the algorithms for global optimization and discrete problems based on methods for local optimization that were recently pioneered by Murray and Ng (e.g., see [56] and [57]). Conceptually, some of these continuation methods seem ideally suited to our problem because they also embed a discrete optimization problem in a continuous one. Furthermore, the “cutting-angle” algorithms demand further exploration. Indeed, preliminary experimentation with these approaches yielded promising results that are described in Andramanov et al. [7]. Even though these results were not competitive with the heuristics in [24] or [25], it is likely that the latest versions of the cutting angle methods would perform better, especially if the starting points were well chosen. We refer the reader to some of the works of A. Rubinov, and L. Batten and their respective collaborators (e.g., see [8], [11] and [12]). Finally, we note that all the mathematical programming methods experimented with so far were essentially “primal methods.” The corresponding dual problems and “dual methods,” to the best of this author’s knowledge, remain unexplored.

4

Spectral Properties, Spin-offs, and Speculation

4.1 Introduction

One benefit of investigating a difficult problem such as the HCP is that these investigations lead to deeper understanding of related phenomena that may have intrinsic interest in their own right. Hence it is not surprising that, the line of research described in this survey has had some spin-offs that are, perhaps, worthy of an explicit mention.

The most natural of these is related to the relationship of a graph G with the spectra of matrices that contain all the pertinent information about that graph. In particular, recall that the *adjacency matrix* A of an N node graph is simply the $N \times N$ matrix with 1's in all ij -entries that correspond to arcs (i, j) present in G and with 0's in all other entries.

It is clear that A contains all the information about the graph G and since the spectrum of a matrix contains a lot of essential information about that matrix, it is natural to ask whether it is possible to differentiate between Hamiltonian and non-Hamiltonian graphs on the basis of the spectra of their respective adjacency matrices?

Regrettably, the answer to the above question is negative. For instance, in Filar et al. [34] an example is given of two, cubic, connected, co-spectral graphs where one is Hamiltonian and the other is not.

Nonetheless, there is a strong temptation to exploit spectral properties of adjacency matrices because of the following, now classical, theorem of Fröbenius (which is described in a more general form in Varga [67]). It concerns cyclic connected graphs. More precisely, we shall say that a connected graph is called *cyclic of index k* if its adjacency matrix has $k > 1$ eigenvalues λ_j that satisfy $|\lambda_j| = \max_i |\lambda_i|$ for $j = 0, 1, \dots, k - 1$.

Theorem 4.1. Consider a connected cyclic graph of index k and let $\lambda_0, \lambda_1, \dots, \lambda_{k-1}$ be those eigenvalues of its adjacency matrix that satisfy $|\lambda_j| = \max_i |\lambda_i|$. These k eigenvalues are the roots of

$$\lambda^k - \left(\max_i |\lambda_i|\right)^k = 0.$$

Specializing the above to our particular problem we note that a Hamiltonian Cycle is a subgraph of a given graph with all the original N vertices and only N selected arcs. Since a Hamiltonian Cycle will enter and leave each vertex exactly once, each row and column of the adjacency matrix for a Hamiltonian Cycle will contain exactly one entry of “1,” and all the rest are zeroes. The following result, that can be seen as a corollary of Fröbenius’s theorem, is proved in Ejov et al. [30] but can also be found in other sources.

Corollary 4.1. An $N \times N$ permutation matrix is the adjacency matrix of some Hamiltonian cyclic graph on N vertices if and only if its characteristic polynomial is $\lambda^N - 1$.

In Ejov et al. [30] Corollary 4.1 was exploited as a basis of an algebraic method to determine whether a given graph is Hamiltonian. To do so, in place of an adjacency matrix a probability transition matrix $P(f)$ for $f \in \mathcal{F}$ was used. Thus instead of 0–1 an entry $f(i, j)$ signifies a probability of choosing an arc $(i, j) \in \mathcal{A}(i)$ at node i , under a policy f , as has been done throughout this survey.

The problem, as always, is to choose values of $f(i, j)$ s to form a Hamiltonian Cycle within the graph, if at all possible. Thus, it is

now reasonable to believe (and was, indeed, established in [30]) that a Hamiltonian Cycle is equivalent to the solution of the system of polynomial equations

$$\left\{ \begin{array}{ll} f(i, j)(1 - f(i, j)) = 0, & \text{for all } (i, j) \in \mathcal{A}, \\ \sum_j f(i, j) - 1 = 0, & \text{for all } i, \\ \sum_i f(i, j) - 1 = 0, & \text{for all } j, \\ \det(\lambda I - P(f)) - \lambda^N + 1 = 0. \end{array} \right. , \quad (4.1)$$

where the last equation is interpreted as a system of equations obtained by equating the coefficients of all powers of λ to 0. If this system of polynomial equations has no solution, then the graph does not have a Hamiltonian Cycle.

A powerful method for solving systems of polynomial equations uses the technique of Gröbner bases, as applied using Buchberger's algorithm (e.g., see [3]). The technique transforms a system of polynomial equations into a "canonical" form which is much easier to solve. One important detail is related to Hilbert's Nullstellensatz, which states that a system of polynomial equations has no solution if its Gröbner basis is $\{1\}$. Thus the Gröbner basis method provides an algebraic check on whether a graph is Hamiltonian or not.

However, we should note that in general the computational complexity of solving a system of polynomial equations via Gröbner bases is at least exponential (e.g., see [22, 51]).

Open Problem. The question of whether Buchberger's algorithm can be speeded up in the particular case of equations (4.1) is one that, perhaps, deserves further investigations.

Finally, we mention that the perturbation methods discussed in Parts I–III of this survey influenced some developments in general theories of perturbations of Markov Decision Processes and linear operators; see for instance Abbad et al. [1, 2] and Avrachenkov et al. [9, 10].

4.2 Fractal-like Structure of Regular Graphs

In view of the above and the discussion in Section 2.2 it is still not clear whether spectral properties can make quantum difference to the

analysis of the Hamiltonian Cycle Problem. However, recent results reported in Ejov et al. [31] reveal a spectral structure of regular graphs that is interesting in its own right and may lead to further insights concerning what is special about non-Hamiltonian regular graphs.¹ This section describes these latest results.

For the sake of simplicity we will pay attention mainly to cubic graphs (or, in other words, to regular graphs of degree $d = 3$). This assumption is not restrictive since all our considerations remain valid for regular graphs of degree $d > 3$ with obvious minor modifications. Moreover, in a certain sense cubic graphs are the generic regular graphs (e.g., see Greenlaw and Petreschi [40]).

So let us consider the set of all regular cubic graphs with n vertices. They can be conveniently enumerated using the GENREG program of Meringer [54]. Each graph is completely determined by its adjacency matrix, which is symmetric.

Motivated by our interest in Markov chains, we replace the adjacency matrix A by the related doubly stochastic matrix $\frac{1}{3}A$. The theory of Markov chains then states that the probability of being at the j th vertex after a walk of length i in the graph with each edge equally likely to be chosen is the j th element of the vector $u^T (\frac{1}{3}A)^i$, where the k th element of the vector u is the probability of starting at the k th vertex.

The spectrum (the set of eigenvalues) of $\frac{1}{3}A$ is real and lies on the segment $[-1, 1]$. We take the exponential of the eigenvalues before finding their sample mean and variance.

Summarizing, the following procedure is applied. For a fixed even n adjacency matrices of all regular cubic graphs on n vertices are found. In each case, divide the adjacency matrix by three, find its eigenvalues, take their exponential, and then find their sample mean and variance. Each cubic graph is then represented by a single dot on a plot of mean versus variance. Thus, we can think of such a dot as a Cartesian “address” of a cubic graph.

Figure 4.1 shows the results of applying this procedure with $n = 8, 10, 12, 14, 16, 18$, where the number of regular cubic graphs in each case is 5, 19, 85, 509, 4060, 41301, respectively.

¹Recall that regular graphs are those with a constant in(out)-degree at each node.

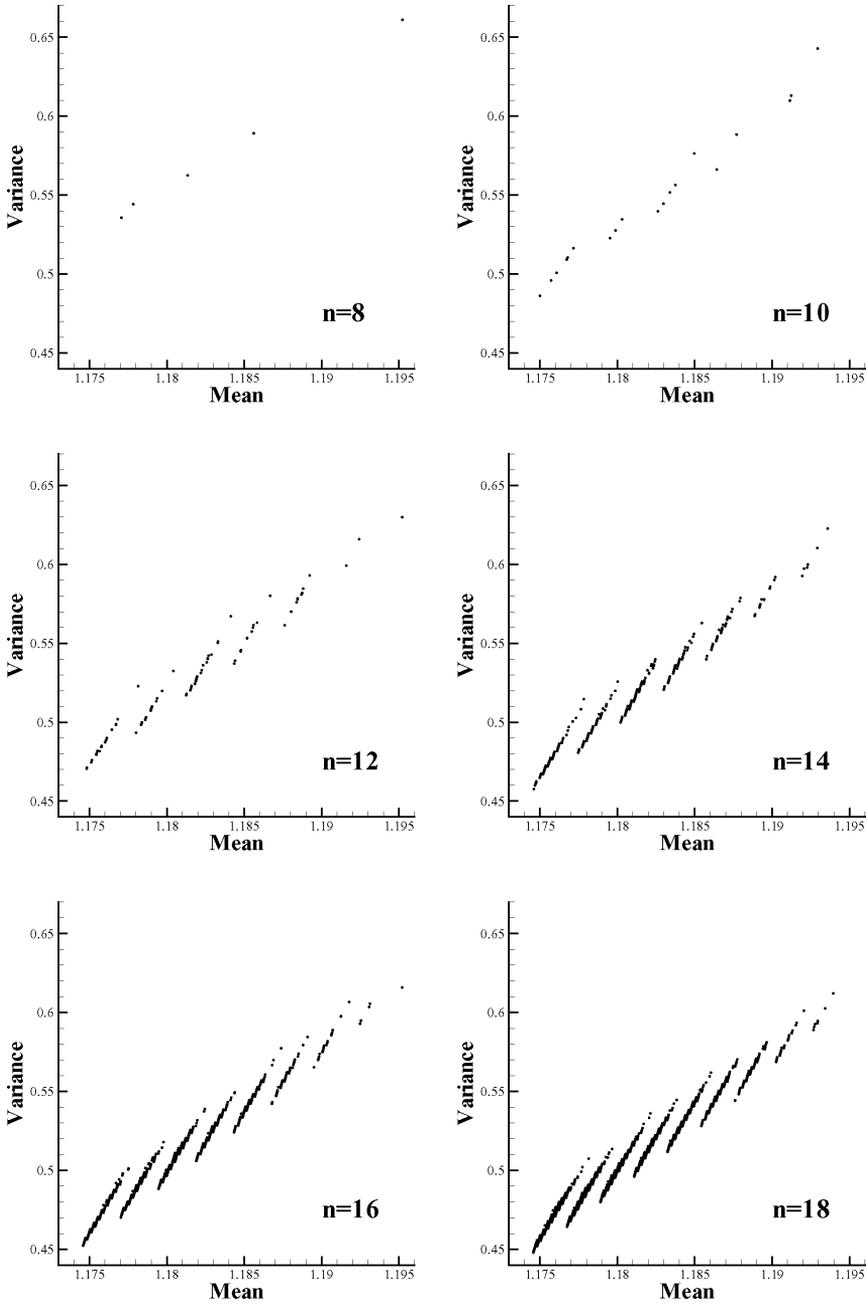


Fig. 4.1 Plots of mean versus variance for the exponential of the eigenvalues of the doubly stochastic matrices associated with all regular cubic graphs with various numbers of vertices.

There appears to be a very definite structure in these plots. In each case the data appear in distinct clusters that at this scale look somewhat like straight line segments with roughly the same slope and distance separating them. An adjective *multifilar* defined as “having threads across the field of view” seems appropriate to describe this structure.

An even greater level of structure exists within each thread. In Ejoy et al. [31] it is demonstrated that zooming in on each of the thread-like objects reveals that it too is made up of similar, smaller, sub-threads of approximately straight line segments, all roughly parallel and the same distance apart, with a steeper slope than the original one.

Since a *fractal* is defined as a self-similar image, where the same structure is evident when magnifying one part of the image, we see that cubic graphs — or, more precisely, their spectra — enjoy a fractal structure, when expressed in this manner. The larger the number of vertices, the more levels of magnification can be undertaken before the number of data points becomes small enough for the self-similar structure to be lost.

In [31] it is also noted that this behavior is not limited to cubic graphs. Plots for quartic graphs (every vertex of degree four) show the same structure. Remarkably, the famous Ihara-Selberg trace formula justifies the presence of such a fractal structure for regular graphs of arbitrary degree d .

In particular, one very explicit version of the Ihara-Selberg trace formula (due to Mnëv [55], Equation (30)) leads to the following interpretation. For any regular graph G of degree $d = q + 1$ on n vertices we have

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n e^{t\lambda_i} &= \frac{q+1}{2\pi} \int_{-2\sqrt{q}}^{2\sqrt{q}} e^{st} \frac{\sqrt{4q-s^2}}{(q+1)^2-s^2} ds \\ &+ \frac{1}{n} \sum_{\gamma} \sum_{k=1}^{\infty} \frac{\ell(\gamma)}{2^{k\ell(\gamma)/2}} I_{k\ell(\gamma)}(2\sqrt{q}t). \end{aligned} \quad (4.2)$$

In the above, $\{\lambda_1, \dots, \lambda_n\}$ is the spectrum of the adjacency matrix of G , γ runs over the set of all primitive closed geodesics. A *geodesic* is a

closed path of minimal length in its free homotopy class. A geodesic is called *primitive* if it is not a multiple of shorter closed geodesics in G . Also, $\ell(\gamma)$ is the length of γ , and $I_m(z)$ is the standard notation for the Bessel function of the first kind:

$$I_m(z) = \sum_{r=0}^{\infty} \frac{(z/2)^{n+2r}}{r!(n+r)!}.$$

All lengths $\ell(\gamma)$ are integers greater or equal to 3. Let us denote by n_ℓ the number of primitive closed geodesics of length ℓ in the graph G . The numbers n_ℓ are called the *multiplicities* of the length spectrum of the graph G . Thus the set $\{n_3, n_4, \dots\}$ provides a convenient way of describing the *length spectrum* of the graph G , that is, the set of lengths of primitive closed geodesics counted with multiplicities.

Now we rewrite Equation (4.2) in terms of the multiplicities n_ℓ (since we consider in detail only the case of cubic graphs, we also put $q = 2$):

$$\frac{1}{n} \sum_{i=1}^n e^{t\lambda_i} = J(t) + \frac{1}{n} \sum_{\ell=3}^{\infty} \ell n_\ell F_\ell(t), \quad (4.3)$$

where

$$J(t) = \frac{3}{2\pi} \int_{-2\sqrt{2}}^{2\sqrt{2}} e^{st} \frac{\sqrt{8-s^2}}{9-s^2} ds,$$

and

$$F_\ell(t) = \sum_{k=1}^{\infty} \frac{I_{k\ell}(2\sqrt{2}t)}{2^{k\ell/2}}.$$

The closed form expressions for the sample mean μ and variance σ can now be easily extracted from (4.3). Precisely, we have

$$\mu = \frac{1}{n} \sum_{i=1}^n e^{\frac{\lambda_i}{3}} = J(1/3) + \frac{1}{n} \sum_{\ell=3}^{\infty} \ell n_\ell F_\ell(1/3), \quad (4.4)$$

and²

$$\begin{aligned}\sigma &= \frac{1}{n} \sum_{i=1}^n \left(e^{\frac{\lambda_i}{3}} - \mu \right)^2 = \frac{1}{n} \sum_{i=1}^n \left(e^{\frac{2\lambda_i}{3}} \right) - \mu^2 \\ &= J\left(\frac{2}{3}\right) + \frac{1}{n} \sum_{\ell=3}^n \ell n_\ell F_\ell\left(\frac{2}{3}\right) - \mu^2.\end{aligned}\quad (4.5)$$

Substituting (4.4) into the last formula and neglecting quadratic terms in F_ℓ that are small, we obtain

$$\sigma \approx \left(J\left(\frac{2}{3}\right) - J\left(\frac{1}{3}\right)^2 \right) + \frac{2}{n} \sum_{\ell=1}^n \ell n_\ell \left(F_\ell\left(\frac{2}{3}\right) - 2J\left(\frac{1}{3}\right) F_\ell\left(\frac{1}{3}\right) \right).\quad (4.6)$$

Open Problem. Since every cubic (indeed, every regular) graph now has a 2-coordinate “address” in the multifilar structure presented above a natural question to pose is: *Is there something special about these addresses, in the case of, the less common, non-Hamiltonian graphs?* A preliminary investigation by the present author indicated that a “great majority” of non-Hamiltonian graphs reside at the tips of the threads in the multifilar structure. However, there were some exceptions where non-Hamiltonian graphs reside in the midst of many Hamiltonian graphs. What, if anything, is special about such “camouflaged” non-Hamiltonian graphs?

4.3 Conclusions and Some Speculative Comments

We have seen that the technique of embedding the Hamiltonian Cycle Problem in a Markov Decision Process has opened up many analytical approaches to this discrete mathematics problem. These included

- (a) Convexification of the domain into either a policy (transition matrix) space, or a frequency space,

²It should be mentioned that the plots in Figure 4.1 are constructed using the *unbiased* variance $sf_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (e^{\lambda_i/3} - \mu)^2$. In order to make formulas simpler, we consider here the variance $s_n^2 = \frac{1}{n} \sum_{i=1}^n (e^{\lambda_i/3} - \mu)^2$. The difference is insignificant, especially for large enough n .

- (b) Exploitation of properties of optimization problems formulated in these spaces including interior point methods,
- (c) Exploitation of distributional properties of random variables (such as first return times) related to stochastic processes induced by the embedding,
- (d) Spectral analysis techniques,
- (e) Exploitation of perturbation/parametric analysis of related problems.

It may be worthwhile to make a few additional, albeit speculative, comments about the last point above. In the majority of the approaches discussed above, the dependence of the problem on the perturbation parameter $\varepsilon > 0$, or the discount factor $\beta \in (0, 1)$ played an important role. Crucially, we note that the asymptotic behavior as $\varepsilon \downarrow 0$, or as $\beta \uparrow 1$ exhibited the presence of singularities.

It is the present author's belief that the underlying difficulty of the HCP can be understood to a much greater degree by a deeper understanding of the asymptotic behavior of the key matrices of Markov Decision Process and of their eigenvalues in the neighborhood of these singularities. Arguably, even the essence of the famous $P \neq NP$ conjecture may be related to this asymptotic behavior.

Furthermore, it is probably not a coincidence, that there exist certain connections between these two asymptotic behaviors: as $\varepsilon \downarrow 0$ and as $\beta \uparrow 1$. In particular, recall the important role played by the fundamental matrix $G(f)$ induced by a policy $f \in \mathcal{F}$. A rarely used property of fundamental matrices (e.g., see Blackwell [13]) implies that

$$G_\varepsilon(f) = \lim_{\beta \rightarrow 1} [I - \beta(P_\varepsilon(f) - P_\varepsilon^*(f))]^{-1} \quad (4.7)$$

$$= \lim_{\beta \rightarrow 1} \left[(I - \beta P_\varepsilon(f))^{-1} - \frac{\beta}{1 - \beta} P_\varepsilon^*(f) \right]. \quad (4.8)$$

Thus all results in Part II of this survey that were proved for “ ε sufficiently small” and concerned an objective function derived from elements of $G(f)$ were really statements concerning the iterated limit

$$\lim_{\varepsilon \rightarrow 0} \lim_{\beta \rightarrow 1} [U(\varepsilon, \beta, f)] = \lim_{\varepsilon \rightarrow 0} \lim_{\beta \rightarrow 1} [I - \beta(P_\varepsilon(f) - P_\varepsilon^*(f))]^{-1}, \quad (4.9)$$

where

$$\begin{aligned} U(\varepsilon, \beta, f) &:= [I - \beta(P_\varepsilon(f) - P_\varepsilon^*(f))]^{-1} \\ &= \left[(I - \beta P_\varepsilon(f))^{-1} - \frac{\beta}{1 - \beta} P_\varepsilon^*(f) \right]. \end{aligned} \quad (4.10)$$

The matrix $U(\varepsilon, \beta, f)$ is especially interesting because of the second equality in (4.10). That equality shows that $U(\varepsilon, \beta, f)$ is determined by a linear combination of the resolvent-type matrix $(I - \beta P_\varepsilon(f))^{-1}$ that plays a key role in the discounted Markov Decision Process and the stationary distribution matrix $P_\varepsilon^*(f)$ that plays a similar role in the long-run average Markov Decision Process. Furthermore, these two matrices are connected by a Tauberian type result (e.g., see Blackwell [13]) that states that for each fixed $\varepsilon > 0$

$$\lim_{\beta \rightarrow 1} (1 - \beta)(I - \beta P_\varepsilon(f))^{-1} = P_\varepsilon^*(f). \quad (4.11)$$

In Ejov et al. [28] the rather strong irreducibility conditions were shown to be necessary and sufficient for the interchange of the iterated limits of $U(\varepsilon, \beta, f)$, namely for the equality

$$\lim_{\beta \rightarrow 1} \lim_{\varepsilon \rightarrow 0} [U(\varepsilon, \beta, f)] = \lim_{\varepsilon \rightarrow 0} \lim_{\beta \rightarrow 1} [U(\varepsilon, \beta, f)]. \quad (4.12)$$

However, from the point of view of understanding Hamiltonicity of a graph, perhaps, it is more important to find an appropriate way for $\beta \rightarrow 1$ as a function of $\varepsilon \rightarrow 0$ to ensure that $U(\varepsilon, \beta, f)$ can be used as a basis of a suitable approximation to the fundamental matrix $G(f)$ for all f either in \mathcal{F} or in \mathcal{DS} .

Open Problems

- (a) Is it possible to use $U(\varepsilon, \beta, f)$ to approximate (RP) of Section 2.1 by a mathematical program (of lower theoretical complexity) whose objective function value at minimum is so close to that of (RP) as to differentiate between a Hamiltonian and a non-Hamiltonian graph? This may be possible over the space \mathcal{DS} . Recall that for any $\varepsilon > 0$ the stationary distribution matrix is simply $P_\varepsilon^*(f) = (1/N)J$, where J is a matrix with a 1 in every entry.

- (b) Another speculative question is whether advantage can be taken of results from the rich fields of statistics and information theory to bring to bear on these stochastic embeddings of discrete mathematics problems? After all, once we become interested in problems such as minimizing a variance of a random variable, sampling schemes ought to be relevant. Furthermore, the recurrent problem of finding a deterministic policy with certain properties refers to “least randomized” policies where the degree of randomization can be captured with entropy/information type functionals. We recall here some conceptual similarities with the emerging topic of “cross-entropy” methods [63, 64].

Acknowledgments

This survey summarizes a line of research initiated in 1985 when the present author posed this approach as a thesis topic for his former PhD student, Dmitry Krass. Since then investigations continued, albeit with numerous interruptions, but with important contributions from many collaborators, colleagues, and research students. The author gratefully acknowledges invaluable assistance and insights gained from collaboration or merely from discussions with Michael Andramonov, Ming Chen, Eric Denardo, Eugene Feinberg, Jacek Gondzio, Arie Hordijk, Lodewijk Kallenberg, Jean-Bernard Lasserre, Ke Liu, Nelly Litvak, Stephen Lucas, Wayne Lobb, Kien Ming Ng, Jessica Nelson, Minh Nguyen, Ulrich Rieder, Walter Murray, Panos Pardalos, Alex Rubinov, and Peter Zograf. Fellow collaborators Vivek Borkar and Vladimir Ejoy deserve a special thanks and an apology. Thanks, because they played a central role in the development of many key results, and an apology for infecting them with the “Hamiltonicity virus” that makes it hard to let go of this problem. The author’s current PhD students; Michael Haythorpe, Ali Eshragh Jahromi, and Giang Nguyen have helped enormously through collaboration, running programs and correcting ever changing versions of the manuscript, not to mention their youthful

energy and enthusiasm. Few, if any, typographical errors could have escaped Ali Eshragh's exceptional proof-reading eye, so if there are any such errors remaining they were almost certainly introduced by the author since Ali's last reading. Last, but not least, much of this research has been partially supported by grants from the Australian Research Council, most recently, by the Discovery grant DP0666632.

References

- [1] M. Abbad, T. Bielecki, and J. A. Filar, “Algorithms for singularly perturbed limiting average Markov control problems,” *IEEE Transactions on Automatic Control*, vol. 37, pp. 1421–1425, 1992.
- [2] M. Abbad and J. A. Filar, “Perturbations and stability theory for Markov control problems,” *IEEE Transactions on Automatic Control*, vol. 37, pp. 1415–1420, 1992.
- [3] W. Adams and P. Lostaunau, *An Introduction to Gröbner Bases*. Graduate Studies in Mathematics, American Mathematical Society, Vol. 3, 1994.
- [4] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*. Book in progress, available at <http://www.stat.berkeley.edu/users/aldous/RWG/book.html>.
- [5] A. Altman and J. Gondzio, “Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization,” *Optimization Methods and Software*, vol. 11–12, pp. 275–302, 1999.
- [6] E. D. Andersen, J. Gondzio, C. Mészáros, and X. Xu, “Implementation of interior point methods for large scale linear programming,” in *Interior Point Methods in Mathematical Programming*, (T. Terlaky, ed.), pp. 189–252, Kluwer Academic Publishers, 1996.
- [7] M. Andramonov, J. A. Filar, A. Rubinov, and P. Pardalos, “Hamiltonian cycle problem via Markov chains and min-type approaches,” in *Approximation and Complexity in Numerical Optimization: Continuous and Discrete Problems*, (P. Pardalos, ed.), pp. 31–47, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2000.
- [8] M. Andramonov, A. Rubinov, and B. M. Glover, “Cutting angle methods in global optimization,” *Applied Mathematics Letters*, vol. 12, pp. 95–100, 1999.

- [9] K. E. Avrachenkov, J. A. Filar, and M. Haviv, “Singular perturbations of Markov chains and Decision Processes,” in *Handbook of Markov Decision Processes*, (E. A. Feinberg and A. Shwartz, eds.), pp. 113–152, Boston: Kluwer, 2002.
- [10] K. E. Avrachenkov, M. Haviv, and P. G. Howlett, “Inversion of analytic matrix functions that are singular at the origin,” *SIAM Journal on Matrix Analysis and Applications*, vol. 22, pp. 1175–1189, 2001.
- [11] A. M. Bagirov, A. M. Rubinov, and J. Zhang, “Local optimization method with global multidimensional search,” *Journal of Global Optimization*, vol. 32, no. 2, DOI 10.1007/s10898-004-2700-0, June 2005.
- [12] L. Batten and G. Beliakov, “Fast algorithm for the cutting angle method of global optimization,” *Journal of Global Optimization*, vol. 24, pp. 149–161, 2003.
- [13] D. Blackwell, “Discrete dynamic programming,” *Annals of Mathematical Statistics*, vol. 33, pp. 719–726, 1962.
- [14] B. Bollobas, T. Fenner, and A. Frieze, “An algorithm for finding Hamiltonian paths and cycles in random graphs,” *Combinatorica*, vol. 7, no. 4, pp. 327–341, 1987.
- [15] V. S. Borkar, *Probability Theory: An Advanced Course*. New York: Springer-Verlag, 1995.
- [16] V. S. Borkar, V. Ejov, and J. A. Filar, “Directed graphs, Hamiltonicity and doubly stochastic matrices,” *Random Structures and Algorithms*, vol. 25, pp. 376–395, 2004.
- [17] V. S. Borkar, V. Ejov, and J. A. Filar, “On the Hamiltonicity gap and doubly stochastic matrices,” submitted to *Random Structures and Algorithms*, 2006.
- [18] A. Broder, A. Frieze, and E. Shamir, “Finding hidden Hamiltonian cycles,” *Random Structures and Algorithms*, vol. 5, no. 3, pp. 395–410, 1994.
- [19] F. Brunacci, “Two useful tools for constructing Hamiltonian circuits,” *EJOR*, vol. 34, pp. 231–236, 1988.
- [20] M. Chen and J. A. Filar, “Hamiltonian cycles, quadratic programming and ranking of extreme points,” in *Global Optimization*, (C. Floudas and P. Pardalos, eds.), pp. 32–49, Princeton University Press, 1992.
- [21] N. Christofides, *Graph Theory: An Algorithmic Approach*. New York: Academic Press, 1975.
- [22] D. Cox, J. Little, and D. O’Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. New York: Springer, Second ed., 1997.
- [23] C. Derman, *Finite State Markovian Decision Processes*. New York: Academic Press, 1970.
- [24] V. Ejov, J. A. Filar, and J. Gondzio, “An interior point heuristic for the Hamiltonian cycle problem via Markov decision processes,” *Journal of Global Optimization*, vol. 29, pp. 315–334, 2003.
- [25] V. Ejov, J. A. Filar, M. Haythorpe, and G. T. Nguyen, “Refined MDP-based branch-and-bound algorithm for the Hamiltonian cycles problem,” submitted to *Mathematics of Operations Research*, 2007.
- [26] V. Ejov, J. A. Filar, W. Murray, and G. T. Nguyen, “Determinants and longest cycles of graphs,” submitted to *SIAM Journal on Discrete Mathematics*, 2007.

- [27] V. Ejov, J. A. Filar, and M. Nguyen, “Hamiltonian cycles and singularly perturbed Markov chains,” *Mathematics of Operations Research*, vol. 29, no. 1, pp. 114–131, 2004.
- [28] V. Ejov, J. A. Filar, and F. Spieksma, “On regularly perturbed fundamental matrices,” *Journal of Mathematical Analysis and Applications*, vol. 336, pp. 18–30, 2007.
- [29] V. Ejov, J. A. Filar, and J. Thredgold, “Geometric interpretation of the Hamiltonian cycle problem via singularly perturbed Markov decision processes,” *Optimization*, vol. 52, pp. 441–458, 2003.
- [30] V. Ejov, J. A. Flar, S. K. Lucas, and J. Nelson, “Solving the Hamiltonian cycle problem using symbolic determinants,” *Taiwanese Journal of Mathematics*, vol. 10, pp. 327–337, 2006.
- [31] V. Ejov, J. A. Flar, S. K. Lucas, and P. G. Zograf, “Clustering of spectra and fractals of regular graphs,” *Journal of Mathematical Analysis and Applications*, vol. 333, pp. 236–246, 2007.
- [32] E. Feinberg, “Constrained discounted Markov decision process and Hamiltonian cycles,” *Mathematics of Operations Research*, vol. 25, pp. 130–140, 2000.
- [33] E. Feinberg and M. T. Curry, “Generalized pinwheel problem,” *Mathematical Methods of Operations Research*, vol. 62, pp. 99–122, 2005.
- [34] J. A. Filar, A. Gupta, and S. K. Lucas, “Connected co-spectral graphs are not necessarily both Hamiltonian,” *The Australian Mathematical Society Gazette*, vol. 32, no. 3, p. 193, 2005.
- [35] J. A. Filar and D. Krass, “Hamiltonian cycles and Markov chains,” *Mathematics of Operations Research*, vol. 19, pp. 223–237, 1994.
- [36] J. A. Filar and J.-B. Lasserre, “A non-standard branch and bound method for the Hamiltonian cycle problem,” *ANZIAM Journal*, vol. 42, no. E, pp. C556–C577, 2000.
- [37] J. A. Filar and K. Liu, “Hamiltonian cycle problem and singularly perturbed Markov decision process,” in *Statistics, Probability and Game Theory: Papers in Honor of David Blackwell*, IMS Lecture Notes — Monograph Series, USA, 1996.
- [38] J. A. Filar and K. Vrieze, *Competitive Markov Decision Processes*. New York: Springer-Verlag, 1996.
- [39] J. Gondzio, “Multiple centrality corrections in a primal-dual method for linear programming,” *Computational Optimization and Applications*, vol. 6, pp. 137–156, 1996.
- [40] R. R. Greenlaw and R. Petreschi, “Cubic graphs,” *ACM Computational Survey*, vol. 27, pp. 471–495, 1995.
- [41] C. M. Grinstead and J. L. Snell, *Introduction to Probability*. American Mathematical Society, Second ed., 2003.
- [42] A. Hordijk, R. Dekker, and L. C. M. Kallenberg, “Sensitivity-analysis in discounted Markovian decision problems,” *OR Spektrum*, vol. 7, pp. 143–151, 1985.
- [43] A. Hordijk and L. C. M. Kallenberg, “Transient policies in discrete dynamic programming: Linear programming including suboptimality tests and additional constraints,” *Mathematical Programming*, vol. 30, pp. 46–70, 1984.
- [44] J. J. Hunter, “Mixing times with applications to perturbed Markov chains,” *Linear Algebra and its Applications*, vol. 417, no. 1, pp. 108–123, 2006.

- [45] L. C. M. Kallenberg, *Linear programming and Finite Markovian Control Problems*. Mathematical Centre Tracts 148, Amsterdam, 1983.
- [46] L. C. M. Kallenberg, "Survey of linear programming for standard and non-standard Markovian control problems, Part I: Theory," *ZOR — Methods and Models in Operations Research*, vol. 40, pp. 1–42, 1994.
- [47] R. Karp, "Probabilistic analysis of partitioning algorithms for the travelling-salesman problem in the plane," *Mathematics of Operations Research*, vol. 2, no. 3, pp. 209–224, 1977.
- [48] T. Kato, *Perturbation Theory for Linear Operators*. Berlin: Springer-Verlag, 1966.
- [49] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. New York: Springer, 1983.
- [50] W. Kocay and P.-C. Li, "An algorithm for finding a long path in a graph," *Utilitas Mathematica*, vol. 45, pp. 169–185, 1994.
- [51] M. Kreuzer and L. Robbiano, *Computational Commutative Algebra 1*. New York: Springer, 2000.
- [52] E. L. Lawler, J. K. Lenstra, A. R. Kan, and D. Shmoys, "The travelling salesman problem: A guided tour of combinatorial optimization," Wiley Series in *Discrete Mathematics and Optimization*. New York: John Wiley and Sons, 1985.
- [53] N. Litvak and V. Ejov, "Optimality of the Hamiltonian cycle," preprint, University of Twente, Enschede, The Netherlands, 2007.
- [54] M. Meringer, "Fast generation of regular graphs and construction of cages," *Journal of Graph Theory*, vol. 30, pp. 137–146, 1999.
- [55] P. Mněv, *Discrete Path Integral Approach to the Trace Formula for Regular Graphs*. Preprint, Steklov Math. Inst., St. Petersburg, 2006.
- [56] W. Murray and K.-M. Ng, "Algorithms for global optimization and discrete problems based on methods for local optimization," in *Handbook of Global Optimization — Volume 2: Heuristic Approaches*, (P. Pardalos and H. E. Romeijn, eds.), Kluwer Academic Publishers, 2002.
- [57] W. Murray and K.-M. Ng, *An Algorithm for Nonlinear Optimization Problems with Discrete Variables*. preprint, 2003.
- [58] I. Parberry, "An efficient algorithm for the Knight's tour problem," *Discrete Applied Mathematics*, vol. 73, pp. 251–260, 1997.
- [59] J. W. Pitman, "Occupation measures for Markov chains," *Advances in Applied Probability*, vol. 9, pp. 69–86, 1997.
- [60] L. Posa, "Hamiltonian circuits in random graphs," *Discrete Mathematics*, vol. 14, pp. 359–364, 1976.
- [61] M. L. Puterman, *Markov Decision Processes*. New York: John Wiley & Sons, 1994.
- [62] R. Robinson and N. Wormald, "Almost all regular graphs are Hamiltonian," *Random Structures and Algorithms*, vol. 5, no. 2, pp. 363–374, 1994.
- [63] R. Y. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," *Methodology and Computing in Applied Probability*, vol. 2, pp. 127–190, 1999.

- [64] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. New York: Springer-Verlag, 2004.
- [65] P. J. Schweitzer, “Perturbation theory and finite Markov chains,” *Journal of Applied Probability*, vol. 5, pp. 401–413, 1968.
- [66] P. J. Schweitzer, “Perturbation series expansion of nearly completely-decomposable Markov chains,” in *Teletraffic Analysis and Computer Performance Evaluation*, (O. J. Boxma, J. W. Cohen, and H. C. Tijms, eds.), pp. 319–328, Elsevier Science Publishers B.V. (North Holland), 1986.
- [67] R. S. Varga, *Matrix Iterative Analysis*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1962.
- [68] A. F. Veinott, Jr., “Discrete dynamic programming with sensitive discount optimality criteria,” *Annals of Mathematical Statistics*, vol. 40, pp. 1635–1660, 1969.
- [69] S. J. Wright, *Primal-Dual Interior-Point Methods*. Philadelphia: SIAM, 1997.