MODEL REDUCTION OF MARKOV CHAINS WITH APPLICATIONS TO BUILDING SYSTEMS

BY

KUN DENG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mechanical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Professor Prashant G. Mehta, Chair and Director of Research
Professor Sean P. Meyn, University of Florida
Professor Carolyn Beck
Professor Geir E. Dullerud
Professor Srinivasa M. Salapaka

Abstract

Markov chain serves as an important modeling framework in applied science and engineering. e.g., Markov Chain Monte Carlo methods and Markov Decision Processes. A fundamental problem of Markov chain models is that the dimension of the problem could be very large in practice. This causes immense difficulties in manipulation and analysis of Markov chain models. Hence, the model reduction of Markov chains is an important problem that is relevant to many applications.

In the first part of this thesis, the model reduction problem of Markov chains is studied and investigated. An information theoretic method is proposed to reduce Markov chains via the *aggregation of states*. The Kullback-Leibler (K-L) divergence rate, a commonly used pseudo metric in statistics and information theory, is employed to measure the differences between two Markov chains. The proposed framework reveals a connection to the spectral properties of Markov chains. In particular, the significance of the second eigenvector is explained in information theoretic terms for the first time. This result leads to a practical recursive model-reduction algorithm based on spectral analysis, and a limited set of error bounds for the model reduction of Markov chains.

Besides using the spectral method, a simulation-based method is also proposed to perform state aggregation of the Markov chain. The main idea is to recast the model reduction problem as an infinite-horizon average cost optimal control problem. An optimal policy corresponds to an optimal aggregation of the state space. The optimal control problem is simplified in an approximate dynamic programming (ADP) approach. A relaxation of the policy space is performed, and based on this a parameterization of the set of optimal policies is introduced. This makes possible a stochastic approximation approach to compute the best policy within a given parameterized class. Convergence of stochastic approximation approach is established using the ODE method.

The aggregation-based model reduction method of Markov chains is extended to hidden Markov models (HMM), which is a special Markov chain model with unobserved state process. Similarly, the state space is aggregated or partitioned to reduce the complexity of the HMM. The optimal aggregation is obtained by minimizing the K-L divergence rate *between the laws of the observation process*. The optimal aggregated HMM is given as a function of the partition function of the state space. The optimal partition is obtained by using a recursive stochastic approximation learning algorithm, which can be implemented through a single sample path of the HMM. Convergence of the algorithm is

established using ergodicity of the filtering process and standard stochastic approximation arguments.

In the second part of this thesis, the modeling and control problems of building systems are investigated. Firstly, A nonlinear resistor-capacitor (RC) network model of a multi-zone building is established to capture the building thermal dynamics. Then an aggregation-based model reduction method is proposed to preserve the structure of the building-thermal model, that is, the reduced building thermal model is still a nonlinear RC-network. This is achieved by obtaining super-nodes via aggregation, and determining the super-capacitance for each super-node and super-resistance for each edge between two adjacent super-nodes. The aggregation-based approach proposed here is based on model reduction method of Markov chains that was described in the first part of this thesis. The main idea is to connect the linear portion of the multi-zone thermal model to a continuous-time Markov chain, and extend the model reduction framework for Markov chains to the nonlinear full-order building thermal model.

A decentralized optimal control strategy is also proposed for a multi-zone building in the second part, where model complexity is mitigated by using a two pronged approach. First, we use the aggregation-based model reduction technique introduced in this part to construct a reduced-order model of the multi-zone building's thermal dynamics. Second, we use the mean-field intuition from statistical mechanics so that the effect of other zones on a particular zone is captured though a mean-field model. Then the whole model does not have to be used in computing the controls over short time scales. A local optimal zonal control law is designed based on the local model of thermal dynamics, and its interaction with the building via the mean-field. The methodology is shown to yield distributed control laws that can be easily implemented on large-scale problems.

In the third part of this thesis, the modeling, analysis, and control of occupancy evolution in a large building are studied. The main concern is efficient evacuation of a building in the event of emergency. Complexity arises due to building topology, uncertainty regarding the distribution of occupants, and the uncertain behavior of occupants. The relaxation techniques borrowed from queueing theory is employed to address complexity issues. These techniques are used to model occupancy evolution during evacuation, obtain lower bounds on evacuation time, and construct control policies to instruct occupants in order to efficiently evacuate the building. The control solutions are based on recent generalizations of the MaxWeight policy for decentralized routing. These results are illustrated with the aid of simulations using realistic building models.

To my beloved family and country.

Acknowledgments

Above all I would like to thank my advisor, Professor Prashant G. Mehta, for his infinite support, enthusiasm, encouragement and guidance through my doctoral study at University of Illinois at Urbana-Champaign. Special thanks also goes to Professor Sean P. Meyn for his directions and insights on most of the research works included in my thesis. I feel lucky and privileged to have worked with them for the past few years. I greatly appreciate the guidance from and interaction with my collaborator Professor Prabir Barooah at University of Florida. He created a great environment for sharing ideas and working together toward a common goal. It was a pleasure and an honor to work with him on my PhD project. I would like to thank Professor Carolyn Beck, Professor Geir E. Dullerud, and Professor Srinivasa M. Salapaka for serving on my committee and provide valuable suggestions. I feel honored to have them as my committee.

I would like to thank Professor Mathukumalli Vidyasagar for hosting me at the University of Texas at Dallas in the spring of 2011. I would like to thank Professor Mustafa Khammash for hosting me at the University of California at Santa Barbara in the summer of 2010. I have thoroughly enjoyed working with them on many interesting problems. I also enjoyed my spare time at Dallas and Santa Barbara.

I would like to thank the Coordinated Science Laboratory (CSL) for providing me the working space and research facilities. The CSL is an exceptional place to be a graduate student. The researches in the lab are interdisciplinary and spectacular. I would like to thank the staff and students of CSL for maintaining such a friendly atmosphere. I especially wish to thank Rebecca Lonberger for her helps these years. I am also grateful to the Department of Mechanical Science and Engineering for accepting me to its PhD program. I would like to thank Kathryn A. Smith for her patience to answer all my questions related to the graduate students and the PhD program.

I would not survive the long journey without my friends along the way. I would like to thank Wei Chen, Yushan Chen, Siddarth Goyal, Dayu Huang, and Yu Sun for collaborating and interacting with me on research and paper writing. I would like to thank the brain trust whom I sat with and bounced ideas off of on a daily basis. Thank you Geng Huang, Adam Tilton, Tao Yang, Huibing Yin, and Chi Zhang. I wold also like to thank Vikas Chandan, Chun Ge, Xun Gong, Bin Li, Yanen Li, Yun Li, Chao Ma, Gayathri Mohans, Yunwen Xu, and many many others. The friendship they offered is greatly appreciated.

Last but not the least I would like to thank my parents, Xiaosong Deng and Yuzhen Yao, for their long and unconditional love all along the way. They also teach me the value of hard work, efficiency, and self-sufficiency. I would like to give special thanks to my wife, Xiangqian Chen, for being incredibly supportive and patient during the course of my PhD life. Their supports and encouragements always give me strength to deal with any difficulties I encountered in my research.

The material of this thesis is based upon work supported in part by the National Science Foundation under Grant (CMS 05-56352, CMS 07-20541, ECCS-0925534, CNS-0931416), and in part by the Air Force Office of Scientific Research under Grant (FA9550-09-1-0190). Any opinions, findings, and conclusions or recommendations expressed in this thesis are those of the authors and do not necessarily reflect the views of the National Science Foundation or Air Force Office of Scientific Research.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Model Reduction of Markov Chains

Markov chains serves as an important modeling framework in applied science and engineering. For example, in *biology*, Markov models are used in genetics and population modeling [67]. In *social science*, social mobility models are described in Markovian terms [63]. *Physics* is another field, where Markov models are used, e.g., Ehrenfest diffusion models, the annealing models, and the Ising-Peierls model of phase transition [65]. In *electrical engineering* and *operation research*, Markov chains are used in the performance analysis of multiple access communications protocols and communication networks, coded modulation, and image processing and segmentation [10]. In recent years, Markov chain theory has received additional attention because of growing importance of Markov Chain Monte Carlo (MCMC) numerical methods [35, 94].

Markov chain models are so useful because they serve as relatively simple building blocks for modeling variety of complex stochastic and complex phenomena. A fundamental problem, however, is that in many applications the dimension of the state space can be very large. This causes immense difficulties in manipulation and analysis of Markov chain models. Hence, the model reduction of Markov chains is an important problem thats is relevant to many applications.

Model reduction for Markov chains has a long history, with many different approaches to both the construction of simplified models, and analysis of the resulting models. Elegant results have been obtained in an asymptotic setting known as singular perturbation [91, 126]. As a particular small parameter approaches zero, the model can be described as a *nearly completely decomposable Markov chain* (NCDMC). A reduced order model is obtained by exploiting this property to construct groups of states with strong interactions. The strongly interacting states within each group are treated as an aggregated super-state, and a Markov model is obtained that describes the transitions from one super-state to another.

1.1.1 Related work

Perhaps the most effective approach to aggregation in singular perturbation models is through spectral theory. It is widely known that the second eigenvalue of an ergodic Markov chain on a finite state space captures the rate of convergence to stationarity, and there are now many extensions to general state space models [83]. In the pioneering work of Wentzell and Freidlin it is shown that an optimal aggregation is obtained through the sign structure of the second eigenvector (or eigenfunction for models on a continuous state space) [32, 60, 100, 120]. In more recent work it is shown that this aggregation technique does not require a singular perturbation framework – an aggregation based on the second eigenfunction is optimal in a certain sense for multi-dimensional diffusions under very general conditions [57].

Markov spectral theory can be regarded as a generalization of spectral graph theory, as described in [21,40,52,106]. In this prior work, a particular graph cut is obtained by consideration of the sign structure of the second eigenvector of the graph Laplacian. This cut is then used to solve a bi-partition problem to break the graph into two sub-graphs with small interaction. The connection with Markov spectral theory is through construction: For a given undirected graph, a Markov chain is obtained as a random walk on the vertices of the graph, where transition probabilities are defined by the weights assigned to the edges. The graph Laplacian is in fact the infinitesimal generator of the Markov transition semigroup in continuous time.

Because of the enormous flexibility of Markov and graphical models, these techniques have broad application: [107] develops theory for aggregation of Markov chains for application in econometric models; Control applications are contained in [91]; Clustering applications are contained in [103]; Estimation and filtering applications are contained in [34, 121]; Applications to queueing models are contained in [23, 126]; Markov spectral theory is applied to nonlinear chaotic dynamics of Chua's circuit model in [60]; [57, 100] concerns applications to molecular models; and [76] contains a treatment of transport phenomena in building systems. There is also a related body of work, in the reinforcement learning literature, for aggregating Markov Decision Processes (MDP) by using the spectral properties of the underlying state transition graph [74]. To date, the range of application of spectral graph theory is even greater (see [21] and references therein).

1.1.2 Overview of contributions

The objective of the this part of the thesis is to examine the decomposition, aggregation, and model reduction techniques of Markov chains. In the following, we survey the major contributions of this part of the thesis.

Optimal aggregated Markov chain

A novel information-theoretic aggregation framework is proposed to obtain the optimal reduced Markov model via aggregation. The metric adopted here is the Kullback-Leibler (K-L) divergence rate for Markov chains [24, 92], which coincides with the Donsker-Varadhan rate function appearing in the large deviations theory of Markov chains [36]. A lifting technique is proposed in this part to extend the K-L divergence rate for two Markov chains defined on different state spaces. It is shown in Theorem 2.3.1 that the optimal lifting is obtained by using the invariant distribution of the original Markov chain, which is consistent with the lifting considered in [38, 101, 117].

The optimal model reduction problem with respect to the K-L divergence rate is tackled in two steps: In the first step, the optimal model reduction problem is solved for a given *fixed* partition function. The solution to the model reduction problem in this step is called the optimal aggregated Markov chain model. The optimal aggregated model is consistent with the optimal reduced model obtained in the optimal prediction theory [7, 20, 117]. The second step pertains to selection of the optimal partition function. Two aggregation methods are proposed in this part to search for the optimal partition function, which are presented below.

Spectral-based aggregation method

A spectral-based aggregation method is proposed for the selection of the optimal partition function. We consider a special case of the bi-partition problem, which is a combinatorial optimization problem. A certain non-convex relaxation is implemented to obtain a continuous nonlinear program to approximate the bi-partition problem. The solution to the relaxed bi-partition problem is shown to be related to an eigenvector problem, consistent with those used in the spectral theory of Markov models. The solution admits a beautiful geometric structure and the significance of the second eigenvector is explained in information theoretic terms for the first time. Some heuristics are proposed for model reduction based on the consideration of the bi-partition problem.

Simulation-based aggregation method

A simulation-based aggregation method is proposed to perform state aggregation of the Markov chain based on observations of a single sample path. Model reduction with respect to the K-L divergence rate is recast as an infinite-horizon average cost optimal control problem. In this way an optimal policy corresponds to an optimal partition of the state space.

The major contribution here is that we simplify the optimal control problem in an approximate dynamic programming (ADP) framework: A relaxation of the policy space is performed, and based on this a parameterization of the set of optimal policies is introduced. This makes possible a stochastic approximation approach to compute the best policy within a given parameterized class. The algorithm can be implemented using a single sample path of the Markov chain.

Convergence is established using the ODE method. Examples illustrate the theoretical results, and show remarkably low variance and fast convergence.

Extension to hidden Markov model

The aggregation-based model reduction method of Markov chains is extended to hidden Markov models (HMM), which is a special Markov chain model with unobserved state process. Similarly, the state space is aggregated or partitioned to reduce the complexity of the HMM. The optimal aggregation is obtained by minimizing the K-L divergence rate *between the laws of the observation process.* The optimal aggregated HMM is given as a function of the partition function of the state space. The optimal partition is obtained by using a recursive stochastic approximation learning algorithm, which can be implemented through a single sample path of the HMM. Convergence of the algorithm is established using ergodicity of the filtering process and standard stochastic approximation arguments.

1.2 Model Reduction of Building Systems

In 2009, commercial and residential buildings accounted for 42% of the total energy usage and 75% of total electricity consumption in the United States [115, Table A.2]. Among all energy consumers of buildings, Heating, Ventilation, and Air Conditioning (HVAC) account for a large share. A large fraction of the energy delivered to buildings is wasted because of inefficient building technologies [3, 109]. Interests in methods for controlling building HVAC systems to reduce their energy usage or cost have been on the increase in recent years; particularly in model-based approaches such as Model Predictive Control (MPC) [46, 71, 73, 90, 128].

Accurate models of temperature evolution in a building are required for real-time prediction and control, especially in model-based control methods. The focus of this part is on the model order reduction of buildings with multiple zones. A zone here refers to a single space (room, hallway, etc.) that is serviced by a single "terminal box" with supply air diffuser and return air grilles.

The physical processes that determine thermal dynamics in buildings, which are governed by a set of coupled partial differential equations, are complex. In principle, Computational Fluid Dynamics (CFD) can be used to solve these equations. CFD models are, however, computationally intensive [19] and sensitive to boundary condition specifications [54]. Complexity issues of CFD models have led to development of simplified models in the past few decades. In this framework, the air in each zone is assumed to be well mixed with a uniform temperature. The thermal response of a zone and conduction between zones that are separated by solid surfaces (walls, floors, ceiling, windows, partitions, etc.) are modeled by capacitances and resistances, respectively. Such resistor-capacitor models have been extensively used to construct dynamic models of zone temperatures in the HVAC and building modeling literature;

see, for instance [14, 44, 102, 124]. The resistances and capacitances are carefully chosen to model the combined effect of conduction between the air masses separated by the surface, as well as long wave radiation and convection between the surface and the air mass in contact with it [44, 89], [4, Chapters 4, 15, & 25]. In addition, there is thermal interaction between each zone and the outside due to the ventilation air that is supplied to and extracted from the zone that has to be accounted for.

A complete model of the entire multi-zone building's thermal response can then be constructed by using (i) resistor-capacitor networks for combined conduction-convection-radiation through surfaces, and (ii) heat balance equations to account for the enthalpy exchange between a zone and the outside due to the ventilation air. The resulting lumped parameter model is called the *full-order building thermal model* in this chapter. This approach of constructing multi-zone thermal models have been pursued previously in [45, 61, 85, 119, 125]. The full-order model we consider here is from [45]. The ventilation heat exchange terms make the thermal dynamic model nonlinear.

A fundamental problem with the full-order models is that they quickly explode in complexity as the number of zones increases. For a large commercial building with hundreds of zones, the number of nodes are of the order of several hundreds, or even thousands. This is a cause of concern for optimization-based control schemes such as MPC, particularly if the optimization is to be performed with a day-long prediction horizon to take advantage of slow thermal responses of buildings as well as daily variations in environment and energy prices. Thus, model reduction methods are required for successful implementation of advanced control schemes in realistic buildings by reducing the computational complexity. Large model complexity is also an issue even for off-line predictions during the building design phase, when a large number of parametric studies are to be performed using building energy prediction software such as EnergyPlus or DOE-2. These design iterations require yearly energy consumption predictions, which need whole-year building simulations. As a result, speedy simulation that comes from low model complexity is important in design iterations. Using a small number of zones to reduce computation time is common practice, and in fact recommended for EnergyPlus [53]. Thus, model reduction techniques can aid in the building design phase as well.

1.2.1 Related work

Due to the nonlinear nature of the building thermal model, the number of available techniques for model reduction is limited. Balanced truncation method for nonlinear systems has been introduced by Scherpen in [98], which uses controllability and observability energy functions of a system to balance the realization. Related methods [1, 22, 47] has also been developed for bilinear systems. These energy functions however are difficult to compute in practice. Lall *et al.* in [69] use empirical Gramians to determine the importance of a particular subspace in terms of its contribution to the input-output behavior. These empirical Gramians are calculated by simulation or experimental data generated within the system's expected operating region, in which some of the nonlinear behavior is captured by resulting

Gramians. Hahn and Edgar [49] propose a hybrid method by introducing controllability and observability covariance matrices, which can be computed from data along system trajectories. The balanced truncation method is further extended in [95] by introducing the so-called extended Gramians to improve error bounds and enforce structural constraints. The recent paper [45] proposes a method for reduction of multi-zone building thermal models of the type considered here. The method in [45] is also based on balanced truncation; it applies a balanced transformation computed from the linear part of the dynamics to the nonlinear full-order model and then performs truncation. The sparsity pattern of the nonlinear terms are exploited to reduce loss of predictive ability in performing the truncation. However, the resulting states of the reduced model have no physical meaning, unlike the states of the original model that relate to temperatures of the zones and internal nodes of walls. The same is true for all model reduction methods mentioned above: the reduced order models do not retain the structure and the physical intuition of the original full-order model.

There is a growing interest in optimal control methods to minimize building-wide energy consumption based on dynamic models [18, 73, 84, 87]. Such control techniques require a model of the transient thermal dynamics of the building that relates the control signals to the space temperature of each zone. A challenge in developing such techniques is the *complexity* of the underlying models due to large dimension of state-space and a large number of control objectives.

1.2.2 Overview of contributions

Model reduction of building thermal model

An *aggregation-based approach* is proposed to preserve the structure of the original model, that is, the reduced building thermal model is still a nonlinear RC-network. This is achieved by obtaining super-nodes via aggregation, and determining the super-capacitance for each super-node and super-resistance for each edge between two adjacent super-nodes. The aggregation-based approach proposed in this chapter is based on model reduction method of Markov chains that was described in the first part of this thesis. The idea here is to connect the linear portion of the multi-zone thermal model to a continuous-time Markov chain, and extend the model reduction procedure for Markov chains to the nonlinear full-order building thermal model.

There are several advantages of the model reduction method of multi-zone building thermal models proposed here compared to the existing general model-reduction methods mentioned above. Unlike the empirical Gramian based methods, we do not need simulation data from full order models to construct the reduced order model; it is obtained directly from the model description. The proposed method does not suffer from the computational difficulty of the energy function based methods. The building thermal model reduction method of [45] has the limitation that the minimum number of states in the reduced model is equal to the number of zones in the building. In contrast, any user

specified reduction in the model order is possible with the proposed method. This makes the proposed method more attractive for MPC-type control schemes, and for performing off-line control design and analysis studies for a building with a large number of zones, when a large reduction in model order is called for. However, for the same (reduced) model order, it turns out that the reduced-order model obtained by the truncation-based method has a slightly smaller prediction error than that by the aggregation-based method.

The key difference of the proposed method over existing work comes from the fact that, unlike all the previously mentioned methods, the method proposed here is *structure-preserving* in the sense that that the reduced model of a nonlinear RC-network is still a nonlinear RC-network. Thus, the parameters and nodes of the reduced model retain the same physical meaning of the parameters and nodes of the original RC-network model. A number of zones can be reduced to a smaller number of "super-zones" with the proposed method. This makes the model reduction method proposed in this chapter is of potential use in the design of buildings as well.

The proposed method also benefits the use in architectural design. Note that it is common in the building design stage to combine a number of zones into a large "super-zones". This is done to reduce simulation time [53]. For instance, a building with 15 zones was reduced to 7 zones for reduction in computation complexity in [2]. However, the process of combining multiple zones into single super zone is done manually; we are not aware of any formal method to perform such aggregation *automatically*. Frequently, such aggregation is done in an an-doc fashion. For instance, all office spaces scattered throughout a building are combined into one zone in [64]. The method proposed in this chapter provides a formal method to perform such aggregation automatically, which benefits building design studies.

Mean-field control of energy efficient building

A *decentralized* optimal control strategy is proposed for a multi-zone building, where model complexity is mitigated by using a two pronged approach. First, we use the aggregation-based model reduction technique introduced in this part to construct a reduced-order model of the multi-zone building's thermal dynamics. Second, we use the mean-field intuition from statistical mechanics so that the effect of other zones on a particular zone is captured though a *mean-field* model [70]. Then the whole model does not have to be used in computing the controls over short time scales.

By using the mean-field idea, we cast the control problem as a game, whereby each zone has its own control objective modeled as set-point tracking of the local (zonal) temperature. In general, the control problem quickly becomes intractable for even a moderate number of competing objectives. In order to mitigate complexity, we employ the Nash Certainty Equivalence principle to obtain a mean-field description [56]. The mean-field here represents the *net effect* of the entire building envelope on any individual zone. A local optimal zonal control is designed based on the

local model of thermal dynamics and its interaction with the building via the mean-field. A consistency relationship is used to enforce the mean-field in a self-consistent manner. The methodology is shown to yield distributed control laws that can be easily implemented on large-scale problems.

1.3 Optimal Evacuation of Large Buildings

This part is concerned with the modeling, analysis, and control of occupancy evolution in a large building. There is currently great interests in these questions for various applications, including energy conservation, and to aid first-responders arriving to a burning building.[1]

The focus of this part is *egress*: We restrict to a transient regime in which the occupants are exiting the building. Although, this is a daily routine for millions of people who egress from a building at the end of work day, we are especially motivated by situations arising due to an emergency, such as a fire or security threat. In such cases, the evacuation time is a critical factor whose reduction can help save lives by both evacuating people faster and providing earlier access to first responders.

In the simplest case in which there is a single occupant in the building, this person will leave the building using the preferred/nearest exit. Modeling this behavior requires an understanding of the preferences of building occupants, which involves considerable uncertainty. Nevertheless, the behavior of an individual in a building can be recorded to construct a probabilistic model. This model can be extended to model a building with several occupants, provided they do not interact — a realistic model of occupancy evolution during evening hours or over the weekend. Predictions from such a model will be similar to the case of a single occupant: individuals will leave the building using their preferred exits.

Congestion makes the evacuation problem much more interesting: If there are a large number of occupants, and a large number of possible routes and exits, then the best route to the nearest exit may not be optimal for the overall evacuation problem. This is a key motivation for the research reported in this chapter.

1.3.1 Related work

For modeling occupancy evolution, *grid-based models* of movement including cellular automata are computationally attractive [122]. Following this literature, we use the term *agent* and *occupant* interchangeably. Such discrete-time discrete-state models have been used for carrying out efficient computer simulations of traffic phenomenon in buildings, planes, and outdoor walkways [48, 51, 97, 104, 105]. The power of these methods lies in their ability to simulate agent behavior by having each agent follow a set of heuristic rules that determine how agents move and interact with

[1] For resources see the Illinois Fire Service Institute website: `www.fsi.uiuc.edu/content/research/agenda.cfm`.

one another. Even though simulations of the grid-based models have been used for performance assessments and off-line designs [15], the application of these models for real-time control and estimation has been limited on account of complexity. Moreover, these approaches offer little insight regarding the properties of good policies.

Fluid and Markovian queueing networks have been used previously in applications to building systems. In particular, Smith and Towsley [108] consider variants of standard Markovian networks to estimate occupancy evolution during evacuation, and related models are used for similar purposes in [28, 80, 88, 99, 114]. Such models are used for occupancy estimation in [82], and anomaly detection in buildings [82].

1.3.2 Overview of contributions

The starting point of this part is the introduction of queueing models for occupancy evolution in the spirit of these previous works. We then apply several complexity mitigation techniques for the purposes of control-oriented modeling and optimization. Specifically, we adapt workload relaxation techniques from queueing theory to formulate appropriate reduced order dynamic models, and to address performance bounds and optimization problems in the context of building evacuation. The major contributions of this part are summarized in the following.

Modeling occupancy evolution

A queueing model is proposed to model agent dynamics such as preferred walking speeds, and preference for exits in the building. The model includes capacity constraints that represent flow and occupancy constraints, such as the restricted rate of flow resulting from a narrow hallway or stairwell. The queueing models proposed in this part provide a flexible framework for modeling occupancy evolution. Two general classes of queueing model are considered in this chapter. The first model class is based on the *Controlled Random Walk (CRW) model* of [80]; it is favored for its simplicity. The second model class is even more idealized, the *fluid model* [80, 86] that models control but not variability.

Performance analysis

The main result of [26] is applied in Proposition 8.2.1 of this part to show that the mean evacuation time for the stochastic model is bounded below by the corresponding quantity for the fluid model. The minimal evacuation time for the fluid model can be obtained by solving a linear program, and can be expressed as a piecewise linear function of the initial condition through associated workload vectors.

Control of evacuation

The policies introduced here are based on the h-MaxWeight policy of [81], which is a generalization of the MaxWeight policy of Tassiulas and Ephremides [113]. For any stochastic queueing model, the h-MaxWeight policy is stabilizing and approximately average-cost optimal under appropriate conditions on the function h and the network [81]. The proof of approximate optimality is also based on the workload relaxation technique. In this part, the h-MaxWeight policy is shown to have considerable reduction in evacuation time over the baseline policy where each occupant uses their nearest exit. The results also compare favorably against the lower bound on evacuation obtained using the workload analysis based on the fluid model.

The remaining of the thesis are organized as follows: The work of model reduction of Markov chains is presented in Chapters 2 to 5. The modeling and control of building systems are presented in Chapter 6 and 7. The optimal evacuation of large buildings is presented in Chapter 8. All proofs are summarized in appendices.

Part I

Model Reduction of Markov Chains

Chapter 2

Optimal Aggregation of Markov Chains

2.1 Introduction

In this chapter, a novel information-theoretic aggregation framework is proposed to obtain the optimal reduced Markov model via aggregation of states.

The metric adopted here is the Kullback-Leibler (K-L) divergence rate for Markov chains [24,92], which coincides with the Donsker-Varadhan rate function appearing in the large deviations theory of Markov chains [36]. The K-L divergence rate measures the distance between two Markov chains, defined on the same state space. However, in the model reduction problems considered here, the reduced order model is constructed on a state space of reduced size. In order to measure the K-L divergence rate between the original and reduced order models, the idea proposed here is to first *lift* the reduced order model to the original state space, and then define the metric as the minimal K-L divergence rate between the original and the lifted Markov chains, over all possible liftings. Here the lifting idea is similar to the aggregation/disaggregation scheme in the linear algebra literature which seeks to compute the invariant distribution of the Markov chain, see [101] for instance. A reduced order model is lifted to the original state space by averaging the dynamics according to a given probability distribution. It is shown in Theorem 2.3.1 that the optimal lifting is obtained by using the invariant distribution of the original Markov chain, which is consistent with the lifting considered in [38, 101].

The optimal model reduction problem with respect to the K-L metric is tackled in two steps: In the first step, the optimal model reduction problem is solved for a given *fixed* partition function. The solution to the model reduction problem in this step is presented in Theorem 2.4.1 of this chapter. The optimal aggregated model is consistent with the optimal reduced model obtained in the optimal prediction theory [7, 20]. The second step pertains to selection of the optimal partition function, which are the topics of the next two chapters.

The remainder of this chapter is organized as follows: In Section 2.2, we introduce the preliminaries of the Markov chain models and K-L metric. In Section 2.3, we extend the K-L metric for comparing two Markov chain defined on different state spaces using the lifting technique. In Section 2.4 the model reduction optimization problem is formulated, and its solution is obtained for a fixed partition function. All the numerical results are contained in

Section 3.4, and all proofs appear in the Appendix A.

2.2 Preliminaries and Notations

2.2.1 Markov chain models

We consider a discrete-time Markov chain $\{X_k\}_{k\geq 0}$ defined on a finite state space $\mathcal{N} = \{1, 2, \ldots, n\}$ (see [83] for terminology). The following notations are adopted throughout the thesis: The state value at time k is denoted as X_k, the sequence $\{X_0, X_2, \ldots, X_l\}$ is denoted as $X_0^l \in \mathcal{N}^{l+1}$, where $\mathcal{N}^{l+1} := \mathcal{N} \times \cdots \times \mathcal{N}$ is the $(l+1)$-fold Cartesian product, and $\mathsf{P}(\mathcal{N}^{l+1})$ denotes the space of probability distributions on $\mathcal{N}^{l+1}$. The transition probability between states is described by a $n \times n$ stochastic matrix P whose ijth entry is given by

$$P_{ij} = \mathsf{P}(X_{k+1} = j \mid X_k = i), \quad i, j \in \mathcal{N}.$$

Assumption 2.2.1 *All Markov chains considered in this thesis are assumed to be* irreducible *and* aperiodic.

Under Assumption 2.2.1, there is a unique *invariant probability distribution* that we denote as π. The following notations are adopted throughout the thesis [24]:

$$h(\pi) := -\sum_{i\in\mathcal{N}} \pi_i \log \pi_i, \tag{2.1}$$

$$H(P) := -\sum_{i,j\in\mathcal{N}} \pi_i P_{ij} \log P_{ij},$$

$$I(P) := h(\pi) - H(P) \tag{2.2}$$

to denote the *entropy* of the invariant distribution, the *entropy rate* of the Markov chain, and the *mutual information*, respectively. The mutual information is between X_k and X_{k+1} assuming stationarity of the process. We adopt the usual convention $0 \log 0 = 0$.

2.2.2 Kullback-Leibler metric

For the model reduction problem, one needs a metric to quantify the differences between the full and the reduced-order models. In this thesis, the *Kullback-Leibler (K-L) divergence* is employed as a "probability metric" to quantify the difference between distributions. Let ξ and ζ denote two probability distributions defined on the same space $\mathcal{N}$. The

K-L divergence between ξ and ζ is defined as:

$$D(\xi\|\zeta) := \begin{cases} \sum_{i\in\mathcal{N}} \xi_i \log(\xi_i/\zeta_i), & \text{if} \quad \xi \prec \zeta \\ +\infty, & \text{otherwise} \end{cases} \tag{2.3}$$

where $\xi \prec \zeta$ means that ξ is *absolutely continuous* with respect to ζ, i.e., $(\zeta_i = 0) \Rightarrow (\xi_i = 0)$ for all $i \in \mathcal{N}$. Note that K-L divergence is not a true metric since it is not symmetric and it does not satisfy the triangle inequality [24]. However, K-L divergence is a natural pseudo-metric for comparing probability distributions and is widely used in statistics, information theory, and control theory [30, 111]. It has two useful properties:

(i) K-L divergence is a *pre-metric*:

$$D(\xi\|\zeta) \geq 0, \quad \text{with equality if and only if } \xi = \zeta.$$

(ii) If the K-L divergence between two distributions is small then the two distributions are also close in the sense of the standard L_1 metric [24]:

$$D(\xi\|\zeta) \geq \frac{1}{2}\|\xi - \zeta\|_1^2, \quad \text{with } \|\xi - \zeta\|_1 := \sum_{i\in\mathcal{N}} |\xi_i - \zeta_i|.$$

To compare two Markov chains, we consider the *K-L divergence rate*, which is a direct generalization of the K-L divergence to the stochastic processes. For two Markov chains with transition matrices P and Q defined on the same state space $\mathcal{N}$, the K-L divergence rate is defined as:

$$R(P\|Q) := \lim_{l\to\infty} \frac{1}{l+1} D(\mathsf{P}_P(X_0^l)\|\mathsf{P}_Q(X_0^l)) \tag{2.4}$$

where $\mathsf{P}_P(\cdot)$ and $\mathsf{P}_Q(\cdot)$ denote the joint probability distribution defined on the cartesian product space $\mathcal{N}^{l+1}$ associated with the two Markov chains. Under Assumption 2.2.1, we have a closed-form formula for the K-L divergence rate [92]:

$$R(P\|Q) = \sum_{i,j\in\mathcal{N}} \pi_i P_{ij} \log\left(\frac{P_{ij}}{Q_{ij}}\right). \tag{2.5}$$

The divergence rate $R(P\|Q)$ is finite provided P is absolutely continuous with respect to Q, i.e., $Q_{ij} = 0 \Rightarrow P_{ij} = 0$.

2.3 Extension of Kullback-Leibler Metric

The goal of this thesis is to construct a simplified Markov model using state aggregation. For aggregation-based approach, it is of interest to compare two probability distributions or two Markov chains defined on different state spaces.

Let ξ and ζ denote two probability distributions defined on $\mathcal{N} = \{1, \ldots, n\}$ and $\mathcal{M} = \{1, \ldots, m\}$ ($m \leq n$), respectively. The relationship between $\mathcal{N}$ and $\mathcal{M}$ is described by a *partition function* ϕ:

Definition 2.3.1 *Let $\mathcal{N} = \{1, 2, \ldots, n\}$ and $\mathcal{M} = \{1, 2, \ldots, m\}$ be two finite state spaces with $m \leq n$. A* partition function $\phi : \mathcal{N} \mapsto \mathcal{M}$ *is a surjective function from $\mathcal{N}$ onto $\mathcal{M}$. For $k \in \mathcal{M}$, $\phi^{-1}(k)$ denotes the kth group of states in $\mathcal{N}$.*

Since ξ and ζ are not defined on the same space, one can not directly use the formula (2.3) to compute the K-L divergence between ξ and ζ. One strategy is to *lift* ζ from the space $\mathcal{M}$ to the space $\mathcal{N}$ based upon a partition function ϕ and a *lifting distribution* μ, where $\mu_i \geq 0$ and $\sum_{i \in \mathcal{V}} \mu_i = 1$. The lifted probability distribution is defined as

$$\widehat{\zeta}_i^{(\mu)}(\phi) := \frac{\mu_i}{\sum_{k \in \psi(i)} \mu_k} \zeta_{\phi(i)}, \quad i \in \mathcal{N}$$

where $\psi(i) := \phi^{-1} \circ \phi(i) \subset \mathcal{N}$ denotes the set of states belonging to the same group as the ith state. The K-L divergence is then extended for two distributions defined on different state spaces using the lifting technique:

$$D_\phi(\xi \| \zeta) := \min_{\mu} D(\xi \| \widehat{\zeta}^{(\mu)}(\phi)).$$

Similar to the extension of K-L divergence for probability distributions, we also extend the K-L divergence rate for Markov chains using the lifting technique. Consider two Markov chains with transition matrices P and Q defined on differen state spaces $\mathcal{N}$ and $\mathcal{M}$, respectively. The relationship between $\mathcal{N}$ and $\mathcal{M}$ is described by a partition function ϕ. To compare two chains with different state spaces, we use the partition function ϕ to *lift* the transition matrix Q from the state space $\mathcal{M}$ to $\mathcal{N}$; the lifted transition matrix is denoted as $\widehat{Q}$. Then we can apply the metric (2.5) to obtain the K-L divergence rate between P and $\widehat{Q}$. The lifted matrix $\widehat{Q}$ is obtained by averaging the probability of Q based upon the partition function ϕ and a probability distribution μ:

Definition 2.3.2 *Let Q be a transition matrix defined on the state space $\mathcal{M}$. Let ϕ be a partition function on $\mathcal{N}$ with range $\mathcal{M}$, and μ be a probability distribution on $\mathsf{P}(\mathcal{N})$. The* μ-lifting of Q *with the partition function ϕ is the*

transition matrix on $\mathcal{N}$ defined as

$$\widehat{Q}^{(\mu)}_{ij}(\phi) = \frac{\mu_j}{\sum_{k\in\psi(j)}\mu_k} Q_{\phi(i)\phi(j)}, \quad i,j \in \mathcal{N} \tag{2.6}$$

where $\psi(j) = \phi^{-1}\circ\phi(j) \subset \mathcal{N}$ denotes the set of states belonging to the same group as the jth state.

The definition of K-L divergence rate is extended to two chains defined on different state spaces using the lifting technique:

Definition 2.3.3 *Consider two Markov chains defined on $\mathcal{N}$ and $\mathcal{M}$, each satisfying Assumption 2.2.1, with respective transition matrices P and Q. Then the K-L divergence rate between these two chains is defined as:*

$$R_\phi(P\|Q) := \min_{\mu\in\mathcal{P}(\mathcal{N})} R(P\|\widehat{Q}^{(\mu)}(\phi)). \tag{2.7}$$

The following theorem states that we can take the probability distribution $\mu = \pi$ to achieve the minimum of the K-L divergence rate (2.7). The proof of Theorem 2.3.1 is given in Appendix A.1.

Theorem 2.3.1 *For any partition function $\phi : \mathcal{N} \mapsto \mathcal{M}$, the invariant distribution π of P is a minimizer of* (2.7)*. The minimum of* (2.7) *can be expressed as:*

$$R_\phi(P\|Q) = R(P\|\widehat{Q}^{(\pi)}(\phi)) = T_\phi(P\|Q) - S_\phi(\pi) \tag{2.8}$$

where

$$T_\phi(P\|Q) = \sum_{i,j\in\mathcal{N}} \pi_i P_{ij} \log\left(\frac{P_{ij}}{Q_{\phi(i)\phi(j)}}\right), \quad S_\phi(\pi) = \sum_{j\in\mathcal{N}} \pi_j \log\frac{\pi_j}{\sum_{k\in\psi(j)}\pi_k}.$$

Note that the theorem does not say that $\mu = \pi$ is the unique minimizer of (2.7). A probability distribution μ minimizes (2.7) if and only if there exist constants $\{K_l,\ l \in \mathcal{M}\}$ satisfying

$$\frac{\pi_j}{\mu_j} = K_l, \qquad \forall j \in \phi^{-1}(l), \quad l \in \mathcal{M}. \tag{2.9}$$

The corresponding matrix $\widehat{Q}^{(\mu)}(\phi)$ then coincides with $\widehat{Q}^{(\pi)}(\phi)$. An example to illustrate this is discussed in Section 2.5.1.

For any $\mu \in \mathsf{P}(\mathcal{N})$, the lifted matrix $\widehat{Q}^{(\mu)}(\phi)$ satisfies the important property of *lumpability* (see e.g, [62, Theo-

rem 6.3.2]). The matrices $\widehat{Q}^{(\mu)}(\phi)$ and Q can be related by the following diagram:

$$\begin{array}{ccc} \hat{\xi}^{(t)} & \xrightarrow{\widehat{Q}} & \hat{\xi}^{(t+1)} \\ \Phi \downarrow & & \uparrow \Psi \\ \xi^{(t)} & \xrightarrow{Q} & \xi^{(t+1)} \end{array}$$

where $\hat{\xi}^{(t)}, \hat{\xi}^{(t+1)} \in \mathsf{P}(\mathcal{N})$ and $\xi^{(t)}, \xi^{(t+1)} \in \mathsf{P}(\mathcal{M})$ denote the probability distributions of two Markov chains at times t and $t+1$. The two transformations $\Phi : \mathsf{P}(\mathcal{N}) \mapsto \mathsf{P}(\mathcal{M})$ and $\Psi : \mathsf{P}(\mathcal{M}) \mapsto \mathsf{P}(\mathcal{N})$ are defined as

$$\Phi_k(\hat{\xi}) = \sum_{i \in \phi^{-1}(k)} \hat{\xi}_i, \quad k \in \mathcal{M}; \qquad \Psi_i(\xi) = \frac{\mu_i}{\sum_{j \in \psi(i)} \mu_j} \xi_{\phi(i)}, \quad i \in \mathcal{N}.$$

2.4 Optimal Aggregated Markov Chain

2.4.1 Problem statement

The *m-partition problem* is to find the optimal partition function $\phi : \mathcal{N} \mapsto \mathcal{M}$, and the optimal aggregated Markov chain with transition matrix Q such that $R_\phi(P \parallel Q)$ is minimized:

$$\begin{array}{ll} \min\limits_{\phi, Q} & R_\phi(P \| Q) \\ \text{s.t.} & \sum_{l \in \mathcal{M}} Q_{kl} = 1, \quad k \in \mathcal{M} \\ & Q_{kl} \geq 0, \quad k, l \in \mathcal{M} \end{array} \tag{2.10}$$

where $R_\phi(P\|Q)$ is given in (2.8), and the constraints characterize the condition that Q is a Markov transition matrix.

The optimization problem (2.10) is a *mixed-integer nonlinear program* due to the fact that ϕ is integer valued, Q is real valued, and $R_\phi(P\|Q)$ is a nonlinear function of ϕ and Q. The problem is non-convex on account of the non-convex constraints on ϕ. Although it can be solved using numerical techniques, the computational complexity increases explosively with the size of the state space.

2.4.2 Optimal solution of Q for fixed ϕ

The main difficulty in solving (2.10) lies in constructing the optimal partition function. The following theorem shows that, for any fixed partition function ϕ, the problem (2.10) is convex, and the matrix Q that solves (2.10) is easily obtained. The proof of Theorem 2.4.1 appears in Appendix A.2.

Theorem 2.4.1 *Let ϕ be a partition function defined on $\mathcal{N}$ with range $\mathcal{M}$. Consider the optimization problem (2.10) subject to the further constraint that ϕ is fixed. Then*

(i) *The optimal solution is given by*

$$Q_{kl}(\phi) = \frac{\sum_{i\in\phi^{-1}(k)} \sum_{j\in\phi^{-1}(l)} \pi_i P_{ij}}{\sum_{i\in\phi^{-1}(k)} \pi_i}, \quad k,l \in \mathcal{M}. \tag{2.11}$$

(ii) *The invariant distribution of $Q(\phi)$ is given by*

$$\theta_k(\phi) = \sum_{i\in\phi^{-1}(k)} \pi_i, \quad k \in \mathcal{M}. \tag{2.12}$$

(iii) *The invariant distribution of the lifted chain with transition matrix $\widehat{Q}^{(\pi)}(\phi)$ coincides with that of the original chain.*

This model reduction result is illustrated with the aid of an example in Section 2.5.2. The formula (2.11) for aggregated transition matrix is well-known (see [23, 101] for early references), and is referred to as the optimal prediction model in several recent papers [7, 38, 96].

Using Theorem 2.4.1, we simplify the m-partition problem (2.10) to only finding the *optimal partition function* ϕ^* such that

$$\phi^* \in \arg\min_{\phi:\mathcal{N}\mapsto\mathcal{M}} R_\phi(P\|Q(\phi)). \tag{2.13}$$

The following lemma shows that *minimizing the K-L metric* $R_\phi(P\|Q(\phi))$ is equivalent to *maximizing the mutual information* $I(Q(\phi))$ of the aggregated chain as defined in (2.2). The proof appears in Appendix A.3.

Lemma 2.4.2 *If the partition function ϕ is fixed, the K-L divergence rate between the original chain and the optimal aggregated chain is given by*

$$R_\phi(P \parallel Q(\phi)) = I(P) - I(Q(\phi)). \tag{2.14}$$

The value of Lemma 2.4.2 is that it allows us to transform the minimization problem (2.10) to the following maximization problem over ϕ:

$$\max_{\phi:\mathcal{N}\mapsto\mathcal{M}} \quad I(Q(\phi)). \tag{2.15}$$

This has the interesting implication that the optimal reduced order model is the one that is most "predictable". For $m = |\mathcal{M}| = 2$, the two deterministic Markov chains with transition matrices,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

have the maximum possible mutual information, $\log(2)$ in either case. This means that if a partition function $\phi^* : \mathcal{N} \to \{1, 2\}$ exists such that $Q(\phi^*)$ in (2.11) is one of these transition matrices, then such a $Q(\phi^*)$ is the optimal solution of (2.10).

2.5 Numerical Results

2.5.1 An Example illustrating Theorem 2.3.1

On $\mathcal{N} = \{1, 2, 3\}$, the Markov chain is described by the transition matrix

$$P = \begin{bmatrix} 0.97 & 0.01 & 0.02 \\ 0.02 & 0.48 & 0.50 \\ 0.01 & 0.75 & 0.24 \end{bmatrix}$$

whose invariant distribution $\pi = [0.3471, 0.3883, 0.2646]$.

The partition function is chosen as $\phi = [1, 2, 2]$, i.e. the state $\{1\}$ is one group, and the states $\{2, 3\}$ are aggregated into another group. Consider a Markov chain on $\mathcal{M} = \{1, 2\}$ with transition matrix

$$Q = \begin{bmatrix} 0.9700 & 0.0300 \\ 0.0159 & 0.9841 \end{bmatrix}.$$

The lifted transition matrix is denoted by $\widehat{Q}^{(\mu)}(\phi)$ (see Definition 2.7). Fig. 2.1 depicts the values of $R(P\|\widehat{Q}^{(\mu)}(\phi))$ as a function of the probability distribution $\mu = [\mu_1, \mu_2, \mu_3]$. We set $\mu_1 = 1 - \mu_2 - \mu_3$, and vary μ_2 and μ_3 in the range $[0, 1]$. The plot shows that $R(P\|\widehat{Q}^{(\mu)}(\phi))$ takes the minimum value whenever

$$\frac{\mu_2}{\pi_2} = \frac{\mu_3}{\pi_3} = \text{const.}$$

This is consistent with (2.9) in Theorem 2.3.1.

Using (2.6), the π-lifting of Q to $\mathcal{N}$ is the 3×3 transition matrix

$$\widehat{Q}^{(\pi)}(\phi) = \begin{bmatrix} 0.9700 & 0.0178 & 0.0122 \\ 0.0159 & 0.5853 & 0.3988 \\ 0.0159 & 0.5853 & 0.3988 \end{bmatrix}.$$

One can verify that $\widehat{Q}^{(\pi)}(\phi)$ is lumpable with Q as its lumped counterpart.

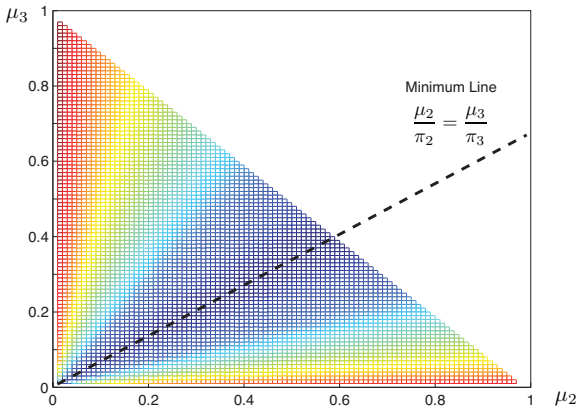


Figure 2.1: Color-plot of $R(P\|\widehat{Q}^{(\mu)}(\phi))$ as a function of μ_2, μ_3. The dashed line indicates points which minimize $R(P\|\widehat{Q}^{(\mu)}(\phi))$.

2.5.2 Optimal model reduction for a fixed partition

Consider a Markov chain parameterized by a parameter $\varepsilon \in (0, 0.5)$:

$$P = \begin{bmatrix} 0.5 & 0.5-\varepsilon & 0 & \varepsilon \\ 0.5-\varepsilon & 0.5 & \varepsilon & 0 \\ 0 & \varepsilon & 0.5 & 0.5-\varepsilon \\ \varepsilon & 0 & 0.5-\varepsilon & 0.5 \end{bmatrix}.$$

For all values of ε, P is a symmetric matrix and thus its invariant distribution $\pi = [0.25, 0.25, 0.25, 0.25]$. Note that for $\varepsilon \approx 0$ or $\varepsilon \approx 0.5$, P is *nearly completely decomposable* (NCD) with partition functions $\phi = [1, 1, 2, 2]$ or $\phi = [1, 2, 2, 1]$, respectively.

We consider three candidate partition functions:

$$\phi^{(1)} = [1, 1, 2, 2], \quad \phi^{(2)} = [1, 2, 1, 2], \quad \phi^{(3)} = [1, 2, 2, 1]. \tag{2.16}$$

The corresponding transition matrices are obtained via application of (2.11) in Theorem 2.4.1,

$$Q(\phi^{(1)}) = \begin{bmatrix} 0.9-\varepsilon & \varepsilon \\ \varepsilon & 0.9-\varepsilon \end{bmatrix}, \quad Q(\phi^{(2)}) = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}, \quad Q(\phi^{(3)}) = \begin{bmatrix} 0.5+\varepsilon & 0.5-\varepsilon \\ 0.5-\varepsilon & 0.5+\varepsilon \end{bmatrix}.$$

Figure 2.2 depicts the K-L divergence rate as a function of the parameter ε for these three partition choices. For $\varepsilon < 0.25$, $R_{\phi^{(1)}}(P\|Q(\phi^{(1)})) < R_{\phi^{(2)}}(P\|Q(\phi^{(2)})), R_{\phi^{(3)}}(P\|Q(\phi^{(3)}))$, and $\phi^{(1)}$ is the optimal partition function. For $\varepsilon > 0.25$, $R_{\phi^{(3)}}(P\|Q(\phi^{(3)})) < R_{\phi^{(1)}}(P\|Q(\phi^{(1)})), R_{\phi^{(2)}}(P\|Q(\phi^{(2)}))$, and $\phi^{(3)}$ is the optimal partition function.

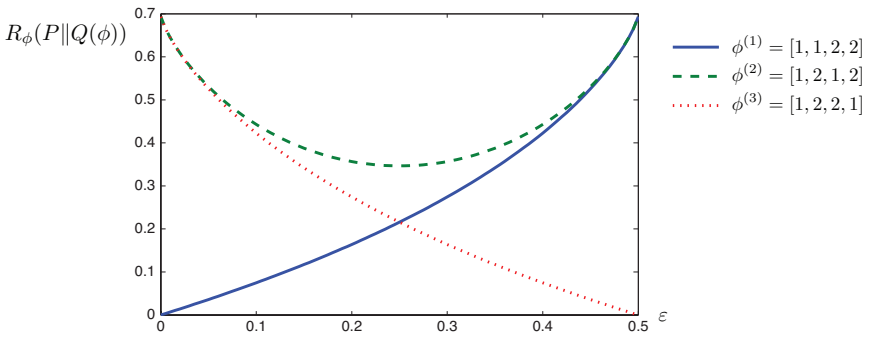


Figure 2.2: K-L divergence rate as a function of the parameter ε for the three partition choices in (2.16).

The optimal aggregation is thus consistent with the NCD aggregation. We note that even though P is lumpable for $\phi = \phi^{(1)}$ (see e.g. [62] for the definition of lumpable matrix), the lifting $\widehat{Q}$ is not equal to P for any $\varepsilon \in (0, 0.5)$. In particular,

$$P - \widehat{Q}^{(\pi)}(\phi^{(1)}) = \begin{bmatrix} 0.05 + \varepsilon/2 & 0.05 - \varepsilon/2 & -\varepsilon/2 & \varepsilon/2 \\ 0.05 - \varepsilon/2 & 0.05 + \varepsilon/2 & \varepsilon/2 & -\varepsilon/2 \\ -\varepsilon/2 & \varepsilon/2 & 0.05 + \varepsilon/2 & 0.05 - \varepsilon/2 \\ \varepsilon/2 & -\varepsilon/2 & 0.05 - \varepsilon/2 & 0.05 + \varepsilon/2 \end{bmatrix}.$$

2.6 Conclusions

In this chapter, we proposed a model reduction method to aggregate the state space of the complex Markov chain models. The K-L divergence rate was employed as a metric to measure the distance between two Markov chains. The lifting technique was used to generalize the metric for chains defined on different state spaces. The optimal lifting has been obtained by considering the invariant distribution of the original Markov chain. Then an optimization problem is posed to minimize this metric and the optimal aggregated Markov chain has been obtained as a function of the partition function. The minimizing K-L divergence rate has been shown to be equivalent to maximizing the mutual information of the optimal aggregated Markov chain over all partition functions.

Chapter 3

Spectral-based Aggregation Method

3.1 Introduction

In Chapter 2, we transform the optimal aggregation problem of Markov chains to a problem for only finding the optimal partition function (see (2.15)). In this chapter, we introduce a spectral-based aggregation method for the selection of the optimal partition function. We consider a special case of the bi-partition problem, which is a combinatorial optimization problem. In Section 3.3, a certain non-convex relaxation is implemented to obtain a continuous nonlinear program to approximate the bi-partition problem. The solution to the relaxed bi-partition problem is shown to be related to an eigenvector problem, consistent with those used in the spectral theory of Markov models. The solution admits a beautiful geometric structure that is illustrated in Figure 3.2. In Section 3.3.4, we describe the error bounds for the bi-partition problem. To construct a reduced order model with m super-states, a heuristic algorithm based on recursive bi-partitions is proposed and illustrated with an example in Section 3.4.2. All numerical results are contained in Section 3.4, and roofs appear in the Appendix B.

3.2 Bi-partition Problem

3.2.1 Setup of the bi-partition problem

As a first step, we introduce a n-by-1 column vector v whose $i\text{th}$ entry is given by

$$v_i = \begin{cases} 1, & \text{if } \phi(i) = 1 \\ 0, & \text{if } \phi(i) = 2. \end{cases} \tag{3.1}$$

Rather than optimizing over ϕ we optimize over functions v. The relaxation is obtained by allowing the entries of v to take on arbitrary values in $\mathbb{R}$. For a given vector $v \in \mathbb{R}^n$ we define a partition function ϕ via,

$$\phi(i) = \begin{cases} 1, & \text{if } v_i > 0 \\ 2, & \text{if } v_i \leq 0. \end{cases} \tag{3.2}$$

A real-valued v is called a *bi-partition function.* A binary-valued v (with values in $\{0,1\}$) is referred to as an *indicator bi-partition function.*

For a given indicator bi-partition function v, we use the notation $Q(v)$ to denote $Q(\phi)$. It is a 2×2 matrix whose entries are obtained using Theorem 2.4.1. The following proposition highlights the special structure in the case $m = 2$. The proof appears in Appendix B.1.

Proposition 3.2.1 *Let v be an indicator bi-partition function defined on $\mathcal{N}$ with range $\mathcal{M} = \{1, 2\}$.*

(i) *The transition matrix (2.11) can be expressed*

$$Q(v) = \begin{bmatrix} \frac{\alpha(v)}{\beta(v)} & \frac{\beta(v)-\alpha(v)}{\beta(v)} \\ \frac{\beta(v)-\alpha(v)}{1-\beta(v)} & \frac{1-2\beta(v)+\alpha(v)}{1-\beta(v)} \end{bmatrix} \tag{3.3}$$

where, with $\Pi = diag(\pi)$,

$$\alpha(v) = v'\Pi P v, \quad \beta(v) = v'\Pi v.$$

(ii) *The invariant distribution of $Q(v)$ is given by*

$$\theta(v) = [\beta(v), 1 - \beta(v)].$$

(iii) *The following constraints hold:*

$$0 \leq \alpha(v) \leq \beta(v) \leq \frac{1 + \alpha(v)}{2} \leq 1. \tag{3.4}$$

(iv) *The two eigenvalues of Q are given by*

$$\kappa_1 = 1, \quad \kappa_2 = \frac{\alpha(v) - \beta(v)^2}{\beta(v)(1 - \beta(v))}. \tag{3.5}$$

Since $Q(v)$ and $\theta(v)$ can be represented in terms of $\alpha = \alpha(v)$ and $\beta = \beta(v)$, the objective function in (2.15) can be expressed as

$$F(v) := I(Q(v)) = h(\theta(v)) - H(Q(v)) = \widetilde{F}(\alpha(v), \beta(v)) \tag{3.6}$$

where the function

$$\widetilde{F}(\alpha, \beta) = \alpha \log \alpha + (1 - 2\beta + \alpha) \log(1 - 2\beta + \alpha) + 2(\beta - \alpha) \log(\beta - \alpha) - 2\beta \log \beta - 2(1 - \beta) \log(1 - \beta). \tag{3.7}$$

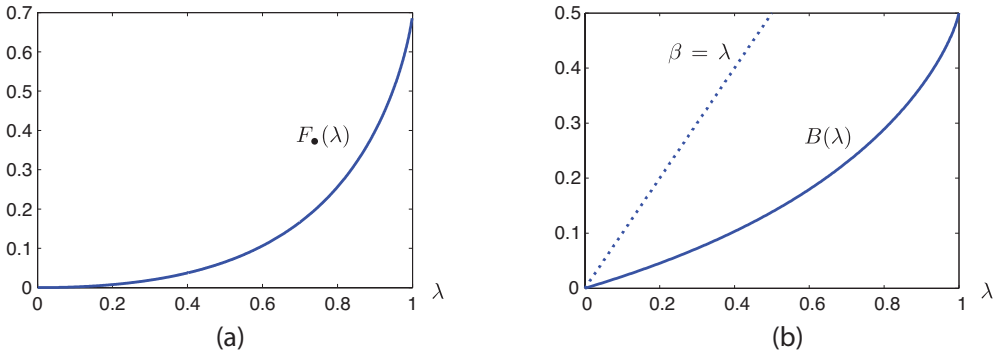


Figure 3.1: Plots of (a) the function $F_\bullet(\lambda)$ for $\lambda \in [0, 1]$ and (b) two solutions of the equation (3.15).

The constraints (3.4) on α and β are necessary and sufficient for $F(v)$ to be well-defined and for $Q(v)$ to be a transition matrix. The constraints are satisfied automatically by an indicator bi-partition function but not necessarily by an arbitrary bi-partition function. It can be shown that $F(v) \in [0, \log(2)]$ for any real-valued bi-partition function v for which the constraints (3.4) hold (see Proposition 3.3.5 in Section 3.3.4).

3.3 Spectral Relaxation of Bi-partition Problem

3.3.1 Non-convex relaxation

We now describe a non-convex relaxation of the optimization problem (2.15). We are interested in choosing a v that solves

$$\max_{v \in \{0,1\}^n} \quad F(v). \tag{3.8}$$

An exact solution of (3.8) may be obtained by searching over 2^n possibilities for vectors v, corresponding to all possible indicator partition functions. We instead consider the relaxation obtained by optimizing over a subset of $v \in \mathbb{R}^n$.

We impose restrictions on v to avoid trivial solutions, such as $v_i > 0$ for each i. This is resolved by assuming that v has mean zero: Denote,

$$\mathbb{S} := \{v \in \mathbb{R}^n : v' \Pi \mathsf{1} = 0\}. \tag{3.9}$$

To avoid a second class of meaningless solutions we impose a nonlinear constraint on v. Define,

$$\mathbb{S}_\bullet := \{v \in \mathbb{S} : \alpha(v) \geq \beta(v)^2\}. \tag{3.10}$$

The relaxation is then defined by the nonlinear program,

$$\begin{aligned} \max \quad & F(v) \\ \text{s.t.} \quad & v \in \mathbb{S}_\bullet. \end{aligned} \tag{3.11}$$

Justification for the constraint $v \in \mathbb{S}_\bullet$ is provided in Section 3.3.2 and the discussion that follows.

To obtain a solution for the relaxation (3.11), we first consider the critical points of the function F in $\mathbb{R}^n$. The critical points are defined as points $v \in \mathbb{R}^n$ where the derivative vanishes. The proof of Lemma 3.3.1 appears in Appendix B.2.

Lemma 3.3.1 *For $v \in \mathbb{R}^n$ to be a critical point of the function F defined in (3.6), one of the following two conditions must hold:*

(i)

$$\check{P} v = \lambda v \tag{3.12}$$

where $\check{P} = \frac{1}{2}(P + \Pi^{-1} P' \Pi)$ and

$$\lambda = \frac{\log \frac{(1-\beta)^2(\beta-\alpha)^2}{\beta^2(1-2\beta+\alpha)^2}}{\log \frac{(\beta-\alpha)^2}{\alpha(1-2\beta+\alpha)}}; \tag{3.13}$$

(ii)

$$\alpha = \beta^2$$

where $\alpha = \alpha(v)$ and $\beta = \beta(v)$.

3.3.2 Spectral characterization

A critical point v that satisfies (3.12) is called a *spectral critical point*; otherwise it is called a *non-spectral critical point*. Proposition 3.3.5 in Section 3.3.4 shows that the non-spectral critical points correspond to the *minimum* value of the function $F(v)$, and are hence uninteresting.

The spectral critical points are obtained by solving the generalized eigenvalue problem (3.12). We summarize the properties of the matrix $\check{P}$:

(i) $\check{P}$ is called the *additive reversiblization* of P [41]. $\check{P}$ is a stochastic matrix with invariant distribution π. If P is reversible with respect to π, i.e., if $\Pi P = P' \Pi$, then $P = \check{P}$.

(ii) The eigenvalues of $\check{P}$ are all real-valued. We denote these eigenvalues as $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, sorted in a decreasing order, and $\{u^{(1)}, u^{(2)}, \ldots, u^{(n)}\}$ the corresponding eigenvectors.

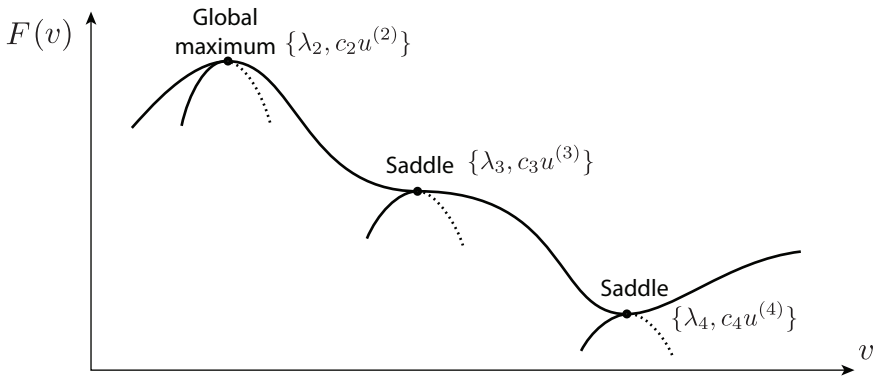


Figure 3.2: Geometric picture for the spectral critical points of the function F on $\mathbb{S}_\bullet$.

(iii) Under Assumption 2.2.1, we have $\lambda_1 = 1$, and $\lambda_i < 1$ for $i \geq 2$. The eigenvectors are chosen to be orthonormal to each other with respect to the π-norm inner product:

$$\langle u^{(i)}, u^{(j)} \rangle_\pi := (u^{(i)})' \Pi u^{(j)}, \quad i, j \in \mathcal{N}.$$

(iv) For a given partition function v, the same reduced order model $Q(v)$ is obtained with either P or $\check{P}$. This is because $v' \Pi P v = v' \Pi \check{P} v$.

The spectral critical points are given by $v = c_k u^{(k)}$, where c_k is some suitable constant for $k \in \{1, \ldots, n\}$ (see Proposition 3.3.2 below).

In the remained of this subsection and the next subsection we make the following two assumptions:

Assumption 3.3.1 *All eigenvalues of $\check{P}$ are non-negative, i.e. $\lambda_k \geq 0$ for all k.*

Assumption 3.3.2 *The second eigenvalue κ_2 of $Q(v)$ is non-negative.*

We can now justify the introduction of the nonlinear constraint in (3.10). Using formula (3.5) for κ_2, we can establish the following two properties for the eigenvalues of $Q(v)$:

(i) Eigenvalues of $Q(v)$ are non-negative when $v \in \mathbb{S}_\bullet$.

(ii) The second eigenvalue κ_2 is positive in the interior of the set $\mathbb{S}_\bullet$, and zero on the boundary of this set.

Also, the constraint (3.4) is automatically satisfied for all $v \in \mathbb{S}_\bullet$. A proof showing this is contained in Appendix B.3.

The following proposition describes the spectral critical points associated with non-negative eigenvalues. Before stating the proposition, we define two functions of a non-negative eigenvalue λ. Observe that if v is a spectral critical

point then

$$\alpha(v) = \frac{v'(\Pi P + P'\Pi)v}{2} = v'\Pi\check{P}v = \lambda v'\Pi v = \lambda\beta(v). \tag{3.14}$$

By substituting $\alpha = \lambda\beta$ into (3.13), we obtain an implicit relation between λ and β:

$$\lambda = \frac{\log \frac{(1-\beta)^2(1-\lambda)^2}{(1-2\beta+\lambda\beta)^2}}{\log \frac{(1-\lambda)^2\beta}{\lambda(1-2\beta+\lambda\beta)}}. \tag{3.15}$$

For each fixed value of $\lambda \in [0,1]$, there are two possible implicit solutions $\beta = \lambda$ and $\beta = B(\lambda)$ shown in Figure 3.1.

The first function of interest is $\beta = B(\lambda)$. The other function

$$F_\bullet(\lambda) := \widetilde{F}(\lambda B(\lambda), B(\lambda)) \tag{3.16}$$

which is depicted in Figure 3.1 for $\lambda \in [0,1]$.

We now state the main conclusion regarding the spectral critical points. The proof appears in Appendix B.5.

Proposition 3.3.2 *Let v denote a spectral critical point associated with the eigenvalue λ_k for $k \in \{2,3,\ldots,n\}$. Then*

(i) $v = c_k u^{(k)} \in \mathbb{S}_\bullet$ *with* $\beta(v) = B(\lambda_k)$, $\alpha(v) = \lambda_k B(\lambda_k)$, *and* $c_k = \pm\sqrt{\frac{B(\lambda_k)}{u^{(k)'}\Pi u^{(k)}}}$,

(ii) $F(v) = F_\bullet(\lambda_k)$,

(iii)

$$Q(v) = \begin{bmatrix} \lambda_k & 1-\lambda_k \\ \frac{B(\lambda_k)-\lambda_k B(\lambda_k)}{1-B(\lambda_k)} & \frac{1-2B(\lambda_k)+\lambda_k B(\lambda_k)}{1-B(\lambda_k)} \end{bmatrix}. \tag{3.17}$$

The function $F_\bullet(\lambda)$ is a monotonically increasing function of $\lambda \in [0,1]$, with $F_\bullet(0) = 0$ and $F_\bullet(1) = \log(2)$ (see Figure 3.1, or Appendix B.4 for a formal proof). As a result, the maximum of $F_\bullet(\lambda)$ is obtained at the largest value of $\lambda = \lambda_2 \geq 0$.

3.3.3 Second variation analysis of spectral critical points

Analysis of the second variation of the function F is used next to show that the spectral critical point associated with λ_2 gives a global maximum of the function $F(v)$ for $v \in \mathbb{S}_\bullet \subset \mathbb{S}$. The spectral critical points associated with other eigenvalues λ_k for $k \in \{3,4,\ldots,n\}$ are saddle points of the function F. In stating the theorem, the following

orthogonal decomposition of the subspace $\mathbb{S} \subset \mathbb{R}^n$ will be useful. For any $k \in \{2, 3, \ldots, n\}$,

$$\mathbb{S} = \Omega_I^{(k)} \oplus \Omega_D^{(k)}$$

where

$$\Omega_D^{(k)} := \text{span}\{u^{(\tilde{k})}, \ldots, u^{(n)}\} \tag{3.18}$$

and $\tilde{k}$ is the largest integer such that $\lambda_{\tilde{k}-1} < \lambda_k$ and $\lambda_{\tilde{k}} = \lambda_k$. For $k = 2$, we define $\Omega_I^{(2)} = \emptyset$ and $\Omega_D^{(2)} = \mathbb{S}$, and for $k = 3, 4, \ldots, n$,

$$\Omega_I^{(k)} := \text{span}\{u^{(2)}, \ldots, u^{(\tilde{k}-1)}\}.$$

The following theorem summarizes the geometric picture illustrated in Figure 3.2. The proof appears in Appendix B.6.

Theorem 3.3.3 *Let v denote a spectral critical point with the eigenvalue λ_v. Denote the Hessian matrix $W(v) := d^2F/dv^2(v)$. Then*

(i) *If $\lambda_v = \lambda_k$ with $k \in \{3, \ldots, n\}$, the spectral critical point $v = c_k u^{(k)}$ defined in Proposition 3.3.2 corresponds to a* saddle point *of the function $F(v)$. We have*

$$\tilde{u}' W(v) \tilde{u} \geq 0, \quad \forall \tilde{u} \in \Omega_I^{(k)}; \qquad \tilde{u}' W(v) \tilde{u} \leq 0, \quad \forall \tilde{u} \in \Omega_D^{(k)}. \tag{3.19}$$

(ii) *If $\lambda_v = \lambda_2 \in (\lambda_3, 1)$ is a simple eigenvalue, the spectral critical point $v = c_2 u^{(2)}$ corresponds to the* global maximum *of the function $F(v)$ for $v \in \mathbb{S}_\bullet \subset \mathbb{S}$.*

The spectral critical point $v = c_2 u^{(2)}$ associated with the eigenvalue λ_2 thus provides a solution to the relaxed problem (3.11). The partition function ϕ is obtained using (3.2).

Recall that the maximum value of mutual information for the bi-partition problem is $\log(2)$ (see Proposition 3.3.5). The next theorem shows that mutual information of the aggregated model is close to the maximum possible value for a nearly completely decomposable Markov chain.

Theorem 3.3.4 *Consider a family of Markov chains parameterized by their transition matrices $\{P_\varepsilon\}$ such that Assumption 2.2.1 holds, with $\lambda_2(\check{P}_\varepsilon) \to 1$ as $\varepsilon \to 0$. Then*

$$\lim_{\varepsilon \to 0} \max_{v \in \mathbb{S}_\bullet} F(v) = \lim_{\varepsilon \to 0} F_\bullet(\lambda_2(\check{P}_\varepsilon)) = \log(2).$$

For each fixed value of ε, the maximizer is the eigenvector for the second eigenvalue of $\check{P}_\varepsilon$.

A procedure for obtaining the spectral critical points is illustrated with the aid of an example in Section 3.4.1.

3.3.4 Performance bounds

In this section, we presents bounds on mutual information for the bi-partition problem. For the relaxed problem, where $v \in \mathbb{R}^n$, these bounds appear in the following proposition. The proof appears in Appendix B.7.

Proposition 3.3.5 *Suppose $v \in \mathbb{R}^n$ and constraints (3.4) hold, then*

$$0 \leq F(v) \leq \log(2)$$

where the lower bound is reached for $\alpha(v) = \beta(v)^2$, and the upper bound is reached for $\alpha(v) = \beta(v) = \frac{1}{2}$.

The main interest here is to obtain bounds for indicator bi-partition functions, i.e., where $v \in \{0,1\}^n$. To distinguish from the real-valued case, we reserve $\bar{v}$ to denote an indicator bi-partition function. In particular, $\bar{v}^{(2)}$ is defined as: For $i \in \mathcal{N}$,

$$\bar{v}_i^{(2)} := \begin{cases} 1, & \text{if } \; u_i^{(2)} \geq 0 \\ 0, & \text{if } \; u_i^{(2)} < 0 \end{cases} \tag{3.20}$$

where $u^{(2)}$ denotes the second eigenvector of $\check{P}$.

For $\bar{v} \in \{0,1\}^n$, we note that both $\bar{v}$ and $(\mathsf{1} - \bar{v})$ represent the same bi-partition of the state space, and $\beta(\bar{v}) + \beta(\mathsf{1} - \bar{v}) \equiv 1$. So, without loss of generality, we restrict $\bar{v}$ to the following subset:

$$\mathbb{T} = \{\bar{v} \in \{0,1\}^n : 0 < \beta(\bar{v}) \leq \frac{1}{2}\}.$$

For the indicator bi-partition function, the following proposition gives bounds on mutual-information. The bounds are given in terms of the following function:

$$F_\circ(\lambda, \beta) := \widetilde{F}(\lambda\beta(1-\beta) + \beta^2, \beta)$$

where the function $\widetilde{F}$ is defined in (3.7). The proof of Proposition 3.3.6 appears in Appendix B.8.

Proposition 3.3.6 *Suppose $\bar{v} \in \mathbb{T}$, then*

$$0 \leq F_\circ(\lambda_n, \underline{\pi}) \leq F(\bar{v}) \leq F_\circ(\lambda_2, \frac{1}{2}) \leq \log(2) \tag{3.21}$$

where λ_n denotes the smallest eigenvalue of $\check{P}$, and $\underline{\pi} := \min_i \pi_i$.

The bounds given in Proposition 3.3.6 do not depend upon the choice of $\bar{v}$, and hence are quite conservative. The following proposition gives better bounds for $\bar{v} = \bar{v}^{(2)}$. Before stating the proposition, we define a variable $\rho(\bar{v})$ that was first introduced in [58]. It is defined as

$$\rho(\bar{v}) := \langle \mathcal{D}\bar{v}, \mathcal{D}\bar{v} \rangle_\pi \tag{3.22}$$

where the projection operator $\mathcal{D}$ is given by

$$\mathcal{D}\bar{v} := \frac{\langle u^{(2)}, \bar{v} \rangle_\pi}{\langle \bar{v}, \bar{v} \rangle_\pi} \bar{v} + \frac{\langle u^{(2)}, \mathsf{1} - \bar{v} \rangle_\pi}{\langle \mathsf{1} - \bar{v}, \mathsf{1} - \bar{v} \rangle_\pi} (\mathsf{1} - \bar{v}).$$

It is straightforward to verify that $\rho(\bar{v}) \in [0, 1]$. Furthermore, $\rho(\bar{v}) = 1$ if and only if $u^{(2)}$ is piecewise constant on each component of the partition defined by $\bar{v}$. The proof of the following proposition appears in Appendix B.9.

Proposition 3.3.7 *Suppose $\bar{v}^{(2)}$ is as defined in (3.20), then*

$$F_\circ(\rho(\bar{v}^{(2)})\lambda_2, \beta(\bar{v}^{(2)})) \le F(\bar{v}^{(2)}) \le F_\circ(\lambda_2, \beta(\bar{v}^{(2)})). \tag{3.23}$$

The next theorem shows that the lower bound approaches the upper bound for a class of NCDMC. The proof appears in Appendix B.10.

Theorem 3.3.8 *Consider a family of Markov chains parameterized by their transition matrices $\{P_\varepsilon\}$ with $\lambda_2(\check{P}_\varepsilon) \to 1$ as $\varepsilon \to 0$. Supposed the indicator bi-partition function $\bar{v}_\varepsilon^{(2)}$ is obtained using (3.20). If $\beta(\bar{v}_\varepsilon^{(2)}) \to \beta^*$ as $\varepsilon \to 0$, then*

$$\lim_{\varepsilon \to 0} F_\circ(\rho(\bar{v}_\varepsilon^{(2)})\lambda_2(\check{P}_\varepsilon), \beta(\bar{v}_\varepsilon^{(2)})) = \lim_{\varepsilon \to 0} F_\circ(\lambda_2(\check{P}_\varepsilon), \beta(\bar{v}_\varepsilon^{(2)})) = h([\beta^*, 1 - \beta^*])$$

where $h(\cdot)$ denotes the entropy function (see (2.1)). In particular, if the limit $\beta^* = \frac{1}{2}$, then

$$\lim_{\varepsilon \to 0} \max_{\bar{v} \in \mathbb{T}} F(\bar{v}) = \lim_{\varepsilon \to 0} F(\bar{v}_\varepsilon^{(2)}) = \log(2).$$

Finally, we make some remarks regarding the implication of these bounds on modeling error:

(i) For an NCDMC, the modeling error

$$I(P) - I(Q(\bar{v}^{(2)})) \approx I(P) - h([\beta, 1 - \beta])$$

where $\beta = \langle \bar{v}^{(2)}, \bar{v}^{(2)} \rangle_\pi$. In particular, if $\beta = \frac{1}{2}$, the the modeling error approaches its smallest possible value $I(P) - \log(2)$.

(ii) For the general case,

$$I(P) - F_\circ(\lambda_2, \beta) \leq I(P) - I(Q(\bar{v}^{(2)})) \leq I(P) - F_\circ(\lambda_2 \rho, \beta)$$

where $\beta = \langle \bar{v}^{(2)}, \bar{v}^{(2)} \rangle_\pi$ and $\rho = \langle \mathcal{D}\bar{v}^{(2)}, \mathcal{D}\bar{v}^{(2)} \rangle_\pi$. In addition to β, the modeling error depends upon the variable ρ, whose value depends upon the geometric structure (piecewise constant on each component) of the second eigenvector $u^{(2)}$.

In general, the computation of bounds require the knowledge of $\bar{v}^{(2)}$ as well as the value of λ_2. Note that the reduced order model itself depends upon $\alpha(\bar{v}^{(2)})$ and $\beta(\bar{v}^{(2)})$ (see (3.3)). So, one may as well compute the reduced order model and the associated modeling error directly.

The utility of these bounds thus lies in illuminating the conditions under which one should expect to use the spectral bi-partition for model reduction. These conditions are as follows:

(i) For an NCDMC, the mutual information and hence the modeling error depends upon the value of β. The modeling error is small (close to its best possible value $\log(2)$) provided $\beta \approx \frac{1}{2}$. This means that, with respect to the π-norm inner product, the two components of the partition have roughly the same number of states.

(ii) For the general non-NCDMC case, the modeling error can deteriorate by an amount that depends upon the geometric structure of the second eigenvector. The best case scenario is where the second eigenvector is piecewise constant on each component of the partition (as in the NCDMC limit).

The condition (ii) suggests that the spectral bi-partition based model reduction is applicable primarily to NCDMC. The condition (i) suggests that even for NCDMC, the modeling error can be arbitrarily poor if β is close to zero. Note that the mutual information estimate for the relaxed problem is misleading in this case (see Theorem 3.3.4).

In summary, the spectral bi-partition based model reduction is useful for NCDMC whose second eigenvector splits the state space into two roughly equal components. In this case, the modeling error is reduced approximately by $\log(2)$, the maximum possible value. For other cases, one should look for alternate approaches including the use of multi-partitioning techniques.

3.3.5 Relationship to other's work

The spectral methods have a rich history for aggregation of Markov chain models [32, 33, 120], as well as related dynamical system models [57, 60], and graphical models [40, 106]. For the bi-partition problem, a sub-optimal aggregation is defined in terms of the sign structure of the second eigenvector (see (3.20)). Under certain conditions,

the optimality is established via a singular perturbation analysis [91, 126] or by obtaining probabilistic bounds on exit time from a superstate [57, 60].

The extension to the multi-partition case has been considered by either employing the bi-partition algorithm recursively (see Section 3.4.2 and also [52, 121]), or via the analysis of the sign structure of multiple eigenvectors [32, 33]. An example of the latter approach is the so-called Perron Cluster Analysis (PCA) for aggregation of a reversible NCDMC. In PCA, the sign structure of the k dominant eigenvectors is used to partition the state space into k superstates.

In literature, the superstates formed via spectral aggregation are referred to as almost invariant sets, and probabilistic bounds yield a form of stability referred to as metastability of the Markov process [58]. The sign-structure heuristic proposed here is directly motivated by the above literature. The result of Theorem 3.3.8 may be viewed as an information-theoretic bound for metastability.

3.4 Numerical Results

3.4.1 Using the second eigenvector for optimal bi-partition

Consider a Markov chain parameterized by a parameter $\varepsilon \in (0, 0.95)$:

$$P = \begin{bmatrix} 0.95-\varepsilon & \varepsilon & 0 & 0.05 \\ \varepsilon & 0.95-\varepsilon & 0.05 & 0 \\ 0 & 0.05 & 0.95-\varepsilon & \varepsilon \\ 0.05 & 0 & \varepsilon & 0.95-\varepsilon \end{bmatrix}.$$

Once again, P is a symmetric matrix and thus its invariant distribution is $\pi = [0.25, 0.25, 0.25, 0.25]$ for any ε. The largest eigenvalue of P is 1, and the other three eigenvalues can be directly obtained as a function of the parameter ε,

$$\lambda_2 = 0.9, \quad \lambda_3 = 1 - 2\varepsilon, \quad \lambda_4 = 0.9 - 2\varepsilon.$$

Note that λ_3 and λ_4 are negative for $\varepsilon > 0.5$. The associated eigenvectors are given by

$$u^{(2)} = [0.5, 0.5, -0.5, -0.5], \quad u^{(3)} = [0.5, -0.5, -0.5, 0.5], \quad u^{(4)} = [0.5, -0.5, 0.5, -0.5].$$

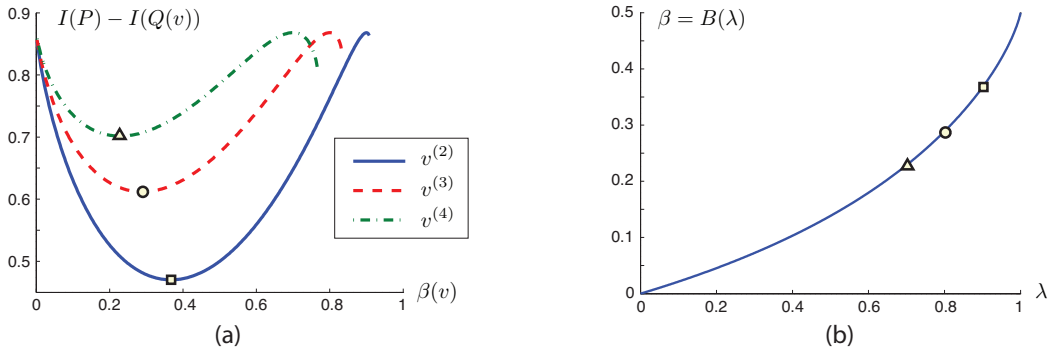


Figure 3.3: (a) $I(P) - I(Q(v))$ as function of $\beta(v)$ for three eigenvalues, and (b) the graph of $\beta = B(\lambda)$ along with the three minimizing points for the three eigenvalues.

Using (3.2), the sign-structure of the three eigenvectors yield three partition functions:

$$\phi^{(2)} = [1, 1, 2, 2], \quad \phi^{(3)} = [1, 2, 2, 1], \quad \phi^{(4)} = [1, 2, 1, 2]. \tag{3.24}$$

We begin by discussing the nearly completely decomposable case where $\varepsilon = 0.1$. The eigenvalues $\lambda_2 = 0.9$, $\lambda_3 = 0.8$, and $\lambda_4 = 0.7$ are all positive. For $i = 2, 3, 4$, we set $v^{(i)} = \gamma u^{(i)}$. Then $\beta(v^{(i)}) = v^{(i)'} \Pi v^{(i)} = \gamma^2 \|u^{(i)}\|_\pi^2$ and $\alpha(v^{(i)}) = \lambda_i \beta(v^{(i)})$. The formula (3.3) in Proposition 3.2.1 gives the following aggregated transition matrices in each case:

$$Q(v^{(i)}) = \begin{bmatrix} \lambda_i & 1 - \lambda_i \\ \frac{\beta(v^{(i)}) - \lambda_i \beta(v^{(i)})}{1 - \beta(v^{(i)})} & \frac{1 - 2\beta(v^{(i)}) + \lambda_i \beta(v^{(i)})}{1 - \beta(v^{(i)})} \end{bmatrix}. \tag{3.25}$$

Figure 3.3 (a) depicts $I(P) - I(Q(v^{(i)}))$ as function of $\beta(v^{(i)})$ for the three cases, where $\beta(v^{(i)})$ is varied by changing the proportionality constant γ. For each case, the minimizer β is a function of λ and falls on the graph $(\lambda, B(\lambda))$ obtained according to the solution of (3.15) (see Figure 3.3 (b)). Figure 3.3 (a) also shows that the second eigenvector yields the smallest value of $I(P) - I(Q(v))$. Substituting $\beta(v^{(2)}) = B(\lambda_2)$ in (3.25) gives

$$Q(v^{(2)}) = \begin{bmatrix} \lambda_2 & 1 - \lambda_2 \\ \frac{B(\lambda_2) - \lambda_2 B(\lambda_2)}{1 - B(\lambda_2)} & \frac{1 - 2B(\lambda_2) + \lambda_2 B(\lambda_2)}{1 - B(\lambda_2)} \end{bmatrix} = \begin{bmatrix} 0.900 & 0.100 \\ 0.058 & 0.942 \end{bmatrix}$$

which is consistent with (3.17) in Proposition 3.3.2. For this case, $I(P) - I(Q(v^{(2)})) = I(P) - F_\bullet(\lambda_2) = 0.4702$.

For the partition function $\phi^{(2)}$, formula (3.3) in Proposition 3.2.1 gives the aggregated transition matrix,

$$Q(\phi^{(2)}) = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix}.$$

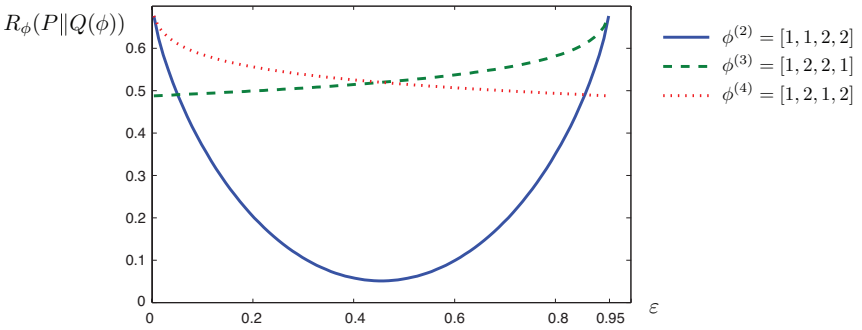


Figure 3.4: K-L divergence rate as a function of the parameter ε for the three partition choices in (3.24).

The K-L divergence rate $R_{\phi^{(2)}}(P\|Q(\phi^{(2)})) = I(P) - I(Q(\phi^{(2)})) = 0.3735$.

As a function of ε, the optimal partition can be obtained by directly comparing $R_{\phi^{(i)}}(P\|Q(\phi^{(i)}))$ for $i = 2, 3, 4$. The comparison is shown in Figure 3.4. We have the following cases:

(i) For $\varepsilon \in (0, 0.05)$, all the eigenvalues are positive and λ_3 is the second largest eigenvalue. The partition function $\phi^{(3)}$ yields the optimal partition.

(ii) For $\varepsilon \in (0.05, 0.45)$, all the eigenvalues are positive and λ_2 is the second largest eigenvalue. The partition function $\phi^{(2)}$ yields the optimal partition.

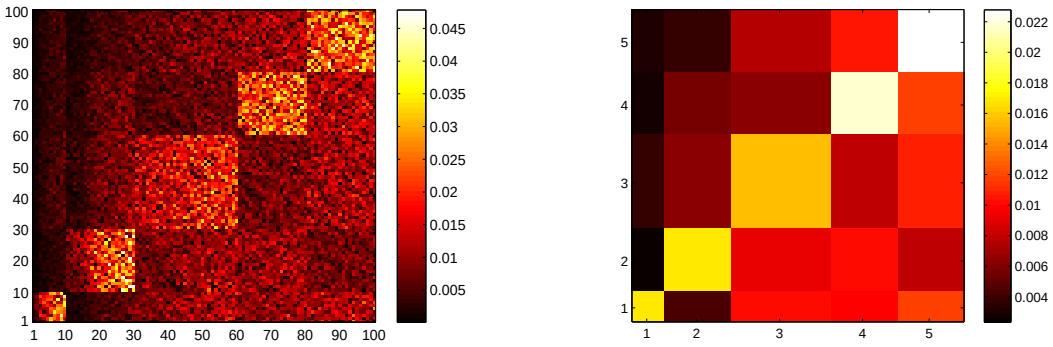
(iii) For $\varepsilon \in (0.45, 0.9)$, some of the eigenvalues are negative but $\lambda_2 > |\lambda_3|, |\lambda_4|$. The partition function $\phi^{(2)}$ yields the optimal partition.

(iv) For $\varepsilon \in (0.9, 0.95)$, λ_3, λ_4 are negative but $|\lambda_4| > \lambda_2, |\lambda_3|$. The partition function $\phi^{(4)}$ yields the optimal partition.

The cases (i) and (ii) are well-supported by theory presented in this paper. Case (iii) can be partially justified by Proposition 3.3.2 for positive eigenvalue λ_2. The geometric picture described in Theorem 3.3.3 is no longer applicable because some of the eigenvalues are negative. The conclusion of case (iv) requires further development of the theory for negative eigenvalues. This example suggests that with negative eigenvalues, the optimal partition function is obtained by considering the sign-structure of the eigenvector associated with the eigenvalue whose modulus is largest.

We close this example with further discussion of case (iv). For $\varepsilon = 0.95$,

$$P = \begin{bmatrix} 0 & 0.95 & 0 & 0.05 \\ 0.95 & 0 & 0.05 & 0 \\ 0 & 0.05 & 0 & 0.95 \\ 0.05 & 0 & 0.95 & 0 \end{bmatrix}.$$



(a) (b)

Figure 3.5: Color plots of (a) the 100-state transition matrix P and (b) the aggregated 5-state transition matrix obtained using the recursive algorithm.

With partition function $\phi^{(4)}, \phi^{(2)}$, the reduced order models are, respectively,

$$Q(\phi^{(4)}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Q(\phi^{(2)}) = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix}.$$

The K-L divergence rates are $R_{\phi^{(4)}}(P\|Q(\phi^{(4)})) = 0.4946$ and $R_{\phi^{(2)}}(P\|Q(\phi^{(2)})) = 0.6931$.

The two partition functions represent two different coordinate choices that capture two distinct types of dynamic features in P. The model $Q(\phi^{(4)})$ is periodic and $Q(\phi^{(2)})$ is nearly completely decomposable. With respect to the K-L divergence rate, $Q(\phi^{(4)})$ is the optimal reduced order model. The numerical values, however, show that $Q(\phi^{(2)})$ is also a reasonable sub-optimal model.

3.4.2 A recursive algorithm to obtain m-partitions

We conclude this chapter with an example that illustrates the recursive bi-partition algorithm AlgoBIPA summarized in Appendix C. The 100-state Markov chain for this example is taken from [77]. Figure 3.5 (a) depicts the transition probabilities. The cold colors indicate weak interactions (small transition probability), and warm colors indicate strong interactions (large transition probability) between states. The color plot suggests that the Markov chain is nearly completely decomposable with five groups.

With $m = 1$, all states are aggregated into a single group and $R_{\phi^{(1)}}(P \parallel Q(\phi^{(1)})) = 0.247$. The bi-partition problem ($m = 2$) is solved by considering the sign-structure of the second eigenvector for the generalized eigenvalue problem (3.12). For the resulting 2-state Markov model, the K-L divergence rate $R_{\phi^{(2)}}(P\|Q(\phi^{(2)})) = 0.176$. The recursive algorithm correctly identifies the five groups in the fifth recursion. Fig. 3.5 (b) depicts the transition probabilities for the resulting reduced order model.

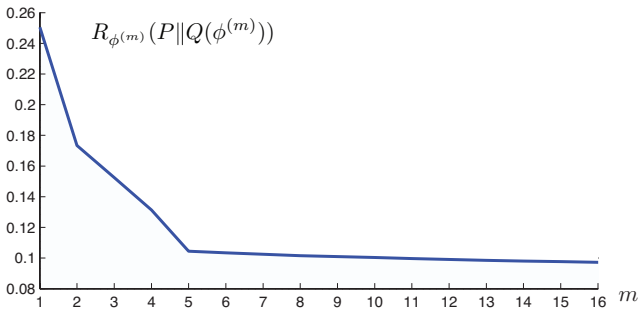


Figure 3.6: K-L divergence rate as a function of the number of aggregated states m.

Figure 3.6 depicts K-L divergence rate as a function of the number of aggregated states $m = |\mathcal{M}|$. The plot shows that $R_{\phi^{(m)}}(P\|Q(\phi^{(m)}))$ decreases rapidly as m increases from 1 to 5. With $m = 5$, $R_{\phi^{(5)}}(P\|Q(\phi^{(5)})) = 0.105$. After five strongly interacting groups have been identified, additional states in the reduced order model ($m > 5$) do not significantly decrease the K-L divergence rate.

3.5 Conclusions

We have shown in Theorem 2.4.1 of the last chapter that the optimal aggregated model is easily obtained, provided a partition function is fixed in advance. The main difficulty in model reduction based on aggregation thus lies in selecting an effective partition. In this chapter we have presented several results intended to approximate this problem, or provide qualitative insight. The main result has been illustrated for the optimal bi-partition problem. The optimal solution is characterized by an associated eigenvalue problem, whose form is similar to the eigenvalue problems considered in Markov spectral theory for model reduction. This result is the basis of a heuristic proposed for the m-ary partition problem, resulting in a practical recursive bi-partition algorithm. Finally, several examples have been presented to illustrate the theoretical results.

Chapter 4

Simulation-based Aggregation Method

4.1 Introduction

In Chapter 2, we transform the optimal aggregation problem of Markov chains to a problem for only finding the optimal partition function (see (2.15)). In Chapter 3, we introduce a spectral-based aggregation method for the selection of the optimal partition function. However, the problem with a spectral-based solution of the aggregation problem is that it requires eigenvectors of the Markov transition matrix, a difficult task for large dimensional problems. A recursive bi-partition algorithm is also introduced to search for the sub-optimal multi-partition functions. The complexity of the algorithm largely increases as the size of the state space. In many applications, the Markov chain is so large that it can only be used for simulations from an initial condition. Due to the high dimensionality of the state space, it may not be even possible to store the eigenvector let alone compute it. This motivates methods for model reduction via simulation and learning, the topic of this chapter.

To confront the complexity issues, in this paper, we formulate the model reduction problem as an infinite-horizon average cost optimal control problem. The control objective is to obtain the partition function (policy) that minimizes the K-L metric (average cost). For small problems, model reduction can be obtained by directly using the methods of dynamic programming (DP). The key advantage of a DP based formulation, however, is that a number of simulation-based approximation methods exist to approximate the optimal value function and learn an optimal partition for large problems [9].

Based on these methods, the model reduction problem is simplified in an approximate dynamic programming (ADP) framework. First, a relaxation of the policy space is performed, and based on this a parameterization of the set of optimal policies is introduced. This makes possible a stochastic approximation approach to compute the best policy within a given parameterized class. The algorithm can be implemented online by simulating a single sample path. At each step of the iteration, the algorithm requires only a small number of calculations even for large problems. Convergence properties are established using standard stochastic approximation arguments (in particular, the ODE method), and illustrated with the aid of examples.

The outline of this chapter is as follows. Section 4.2 introduces the dynamic programming formulation of the K-L

metric based model reduction problem. In Section 4.3, a simulation-based algorithm for the DP is outlined along with a discussion of its convergence properties. In Sections 4.4, numerical results are described.

4.2 Dynamic Programming Formulation

In this chapter, we also consider the discrete time finite Markov chains satisfying Assumption 2.2.1. We use a tuple (π, P) to denote a Markov chain with the transition matrix P and the invariant distribution π. Consider a stationary Markov chain (π, P) defined on a finite state space $\mathcal{N}$ with $n =\mid \mathcal{N} \mid$. Let $\Phi = \{\phi_1, \phi_2, \ldots, \phi_L\}$ denote the collection of all possible partition functions defined on $\mathcal{N}$. For a m-partition problem, we have $L = m^n$.

4.2.1 K-L metric as an average cost

At state $i \in \mathcal{N}$, we define the *one-step cost* as

$$g_i(\phi) = \sum_{j \in \mathcal{N}} P_{ij} \log \left(\frac{P_{ij}}{\widehat{Q}_{ij}^{(\pi)}(\phi)} \right) \tag{4.1}$$

where $\phi \in \Phi$, $\widehat{Q}^{(\pi)}(\phi)$ denotes the π-lifting of $Q(\phi)$ (see (2.6)), and $Q(\phi)$ denotes the optimal aggregated Markov transition matrix (2.11).

The *average cost* is defined as

$$\lambda(x_0; \phi) = \lim_{T \to \infty} \frac{1}{T} \mathsf{E} \left[\sum_{t=0}^{T-1} g_{X(t)}(\phi) \right]$$

where $X(t) \in \mathcal{N}$ denotes the state of Markov chain at time t with $X(0) = x_0$.

For a stationary Markov chain (π, P), the average cost $\lambda(x_0; \phi)$ is well defined for any partition function $\phi \in \Phi$ and does not depend on the initial state x_0. It is given by

$$\lambda(\phi) := \sum_{i \in \mathcal{N}} \pi_i g_i(\phi). \tag{4.2}$$

Substituting (4.1) into (4.2), we find that the average cost $\lambda(\phi)$ is just the K-L metric $R_\phi(P \| Q)$ (see (2.8)).

The optimization problem (2.13) now becomes

$$\phi^* \in \arg\min_{\phi \in \Phi} \lambda(\phi) \tag{4.3}$$

where optimal partition function ϕ^* is the one that achieves the minimum, and $\lambda^* := \lambda(\phi^*)$ denotes the *optimal average cost*.

4.2.2 Dynamic programming approach

We recast the model reduction problem (4.3) as an optimal control problem: The control objective is to obtain a stationary policy, the partition function ϕ^*, that achieves the optimal average cost.

For any state $i \in \mathcal{N}$ and policy $\phi \in \Phi$, we define the *differential cost function* $h_i(\phi)$ as

$$h_i(\phi) = \mathsf{E}\left[\sum_{t=0}^{T-1}(g_{X(t)}(\phi) - \lambda(\phi)) \mid X(0) = i\right]$$

where $\lambda(\phi))$ is given in (4.2), and $T = \min\{t > 0 \mid X(t) = s\}$ is the first future time that *reference state* s is visited. With this definition, we always have $h_s(\phi) = 0$. Since (π, P) is a stationary Markov chain, the reference state s can be taken as any fixed state in $\mathcal{N}$. The differential cost function $h_i(\phi)$ captures the relative difference of starting the process in state i, rather than in the reference state s. It follows that the vector $h(\phi) := [h_1(\phi), h_2(\phi), \ldots, h_n(\phi)]$ is the unique solution to the following Poisson equation (see e.g. [83]),

$$g(\phi) = \lambda(\phi)e + (I - P)h$$

where $e := [1, 1, \ldots, 1]$, and I denotes the identity matrix.

For the optimization problem (4.3), although there may exist several optimal polices, there only exists a unique vector h^* such that, for all optimal polices ϕ^*, we have $h(\phi^*) = h^*$. We refer to h^* as the *optimal differential cost vector*. The theory of dynamic programming asserts that the optimal differential cost vector h^* is the unique solution to the following Bellman equation

$$h_i = \min_{\phi \in \Phi}\left\{g_i(\phi) - \lambda^* + \sum_{j \in \mathcal{N}} P_{ij}h_j\right\}, \quad i \in \mathcal{N}\backslash\{s\} \tag{4.4}$$

and $h_s = 0$, where $\lambda^* = \lambda(\phi^*)$ is the optimal average cost.

The optimal differential cost vector as well as the optimal policies can be obtained using standard dynamic programming methods, i.e., value or policy iteration algorithms [9]. For a stationary Markov chain, these algorithms are guaranteed to converge to an optimum. However, a direct implementation of these algorithms is impractical because of the curse of dimensionality [9]. The curse here arises not only due to the large size n of the state space $\mathcal{N}$ but also the even larger size L of the partition function space Φ.

4.3 Aggregation via Learning

There are two separate complexity related issues. The first issue is the large number L of partition functions. The second issue is the large size n of the state space. To confront the first issue, a parametric representation is described to represent partition functions in terms of a small number of parameters. For the second issue, a simulation-based method is described based on a single sample path of the Markov chain.

4.3.1 Parameterizations of randomized partition policy

A randomized partition policy is defined as a mapping

$$\eta : \mathcal{N} \to [0,1]^L$$

where component $\eta_\phi(i)$ is the probability that the partition function ϕ is assigned to state i. We have $\sum_\phi \eta_\phi(i) = 1$ for any $i \in \mathcal{N}$. The partition policy is said to be deterministic if for every state i, there is a single partition function $\phi^{(i)}$ such that $\eta_{\phi^{(i)}}(i) = 1$. If the function $\phi^{(i)}$ is the same for all i then the policy η yields a partition of the space $\mathcal{N}$.

The problem is that L is very large, i.e. $L = m^n$ for the m-partition problem. So, following the considerations of [75], we introduce a parameter vector $\theta := [\theta_1, \theta_2, \ldots, \theta_K] \in \mathbb{R}^K$ where K is of moderate size.

For $\theta \in \mathbb{R}^K$, we associate a randomized partition policy $\eta_\phi(i,\theta)$ where $\sum_\phi \eta_\phi(i,\theta) = 1$. For every $i \in \mathcal{N}$ and $\theta \in \mathbb{R}^K$, the expected cost per stage is defined as

$$g_i(\theta) = \sum_{\phi \in \Phi} \eta_\phi(i,\theta) g_i(\phi).$$

Then the average cost is defined as

$$\lambda(\theta) = \lim_{T\to\infty} \frac{1}{T} \mathsf{E}\left[\sum_{t=0}^{T-1} g_{X(t)}(\theta)\right].$$

For the Markov chain (π, P), we have

$$\lambda(\theta) = \sum_i \pi_i g_i(\theta). \tag{4.5}$$

Then the optimization problem can be expressed as

$$\theta^* \in \arg\min_{\theta \in \mathbb{R}^K} \lambda(\theta), \tag{4.6}$$

and θ^* defines an optimal randomized policy $\eta_\phi(i,\theta^*)$. Furthermore, if η_ϕ is deterministic and independent of i then a partition function can be uniquely obtained from η_ϕ.

In practice, the policy is independent of i because the Markov chain transition matrix P does not depend upon the partition function ϕ. A numerical solution will in general, however, only lead to a partition function with high probability determined by η_ϕ. This is discussed in a greater detail in Section 4.4 with the aid of numerical examples.

4.3.2 Parametric representation for the bi-partition problem

Let $\mathcal{M} = \{1, 2\}$ denote the aggregated state space with two superstates. For the bi-partition problem, a partition function $\phi : \mathcal{N} \to \mathcal{M}$ can take only two values, 1 and 2 for any state $i \in \mathcal{N}$. We consider the parameter vector $\theta := [\theta_1, \ldots, \theta_n] \in \mathbb{R}^n$, where θ_i decides the group assignment for the state $i \in \mathcal{N}$. In particular, we use $\frac{1}{1+\exp(M\theta_i)}$ to reflect the probability that $\phi(i) = 1$, where $M > 0$ is some positive constant. This gives a randomized partition policy

$$\eta_\phi(i, \theta) = \frac{1}{1 + \exp(M\theta_i)} \mathbb{1}_{\phi(i)=1} + \frac{\exp(M\theta_i)}{1 + \exp(M\theta_i)} \mathbb{1}_{\phi(i)=2} \tag{4.7}$$

where $\mathbb{1}_{\phi(i)=1}$ is 1 if $\phi(i) = 1$ and 0 otherwise, and similarly for $\mathbb{1}_{\phi(i)=2}$.

In the remainder of this paper, we will consider only the bi-partition problem with the representation $\eta_\phi(i, \theta)$ given in (4.7). The general case is similarly handled but the bi-partition case allows for a simpler discussion of the main ideas.

With the parametric representation (4.7), the function $\lambda(\theta)$ defined in (4.5) is bounded from below (i.e. $\lambda(\theta) \geq 0$), twice differentiable with respect to θ, and has bounded first and second derivatives for any $\theta \in \mathbb{R}^n$. Thus, the gradient of $\lambda(\theta)$ is well defined (we denote $\nabla := \nabla_\theta$ for short)

$$\nabla\lambda(\theta) = \sum_{i \in \mathcal{N}} \pi_i \nabla g_i(\theta)$$

where

$$\nabla g_i(\theta) = \sum_{\phi \in \Phi} \nabla\eta_\phi(i, \theta) g_i(\phi). \tag{4.8}$$

4.3.3 An idealized gradient-descent algorithm

Since the gradient of $\lambda(\theta)$ can be exactly computed, we could use the gradient-based method to solve the unconstrained optimization problem (4.6) in the whole real space $\mathbb{R}^n$. An idealized gradient-decent algorithm is given here to update the parameter θ

$$\theta^{(t+1)} = \theta^{(t)} - \gamma_t \nabla\lambda(\theta^{(t)}). \tag{4.9}$$

With the parametric representation (4.7), the gradient $\nabla\lambda(\theta)$ is a bounded and globally Lipschitz continuous function on $\mathbb{R}^n$. By choosing suitable stepsize γ_t, one can show that $\lim_{t\to\infty} \nabla\lambda(\theta^{(t)}) = 0$ and $\theta^{(t)}$ converges to a

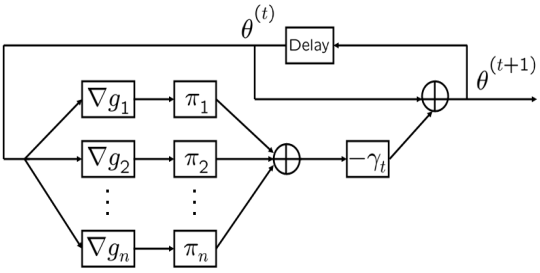


Figure 4.1: Idealized gradient-descent algorithm: The parameters are updated using the gradient of the average cost evaluated on the entire state space.

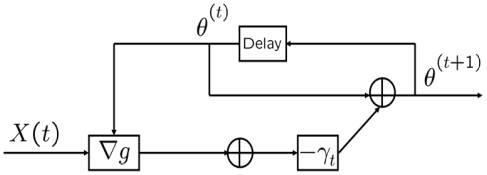


Figure 4.2: Simulation-based gradient-descent algorithm: The parameters are updated using the gradient of the one-step cost evaluated only on the current state $X(t)$.

finite value (see [9, Prop. 4.3.2]).

However, each iteration of the idealized gradient-descent algorithm involves a burdening computation of the averaged gradient term $\nabla\lambda(\theta^{(t)})$ based on each gradient $\nabla g_i(\theta^{(t)})$ over the entire state space (see Figure 4.1). Significant computer resources must be allocated in order to store a large number of iteration values and compute this averaged gradient.

4.3.4 A simulation-based gradient-descent algorithm

A simulation-based (stochastic-approximation) gradient-descent algorithm is obtained by dropping the averaging operation in the idealized gradient algorithm (4.9). In this algorithm, the true (averaged) gradient $\nabla\lambda(\theta)$ is approximated by the one-step cost gradient $\nabla g_{X(t)}(\theta)$ (see Figure 4.2).

Let $\{X(t)\}_{t=0}^{\infty}$ be a single sample path generated from the simulation of a stationary Markov chain (π, P). Then the simulation-based gradient-descent algorithm for updating the parameter vector θ is given by

$$\theta^{(t+1)} = \theta^{(t)} - \gamma_t \nabla g_{X(t)}(\theta^{(t)}) \tag{4.10}$$

where the value $\theta^{(t)}$ is assumed to be available from the previous iteration, and $\nabla g_{X(t)}(\theta)$ is computed using (4.8) for any $X(t) \in \mathcal{N}$. In addition, another stochastic approximation algorithm for updating the average cost is run in parallel

$$\tilde{\lambda}^{(t+1)} = \tilde{\lambda}^{(t)} + \gamma_t (g_{X(t)}(\theta^{(t)}) - \tilde{\lambda}^{(t)}) \tag{4.11}$$

where $\tilde{\lambda}^{(t)}$ is the estimated average cost and parameter $\theta^{(t)}$ comes from (4.10).

It is assumed throughout that the stepsize γ_t satisfies the standard stochastic approximation conditions [11]:

Assumption 4.3.1 *The stepsize values $\{\gamma_t\}$ are nonnegative and satisfy*

$$\sum_{t=1}^{\infty} \gamma_t = \infty, \quad \sum_{t=1}^{\infty} \gamma_t^2 < \infty.$$

In the simulations described in this paper, $\gamma_t = \frac{1}{t}$ is chosen to satisfy the above assumption.

The convergence of the simulation-based algorithm is established using an ODE method. The ODE is obtained by first considering the difference equation obtained by taking an expectation on both sides of (4.10) and (4.11)

$$\begin{aligned}
\theta^{(t+1)} &= \theta^{(t)} - \gamma_t \nabla \lambda(\theta^{(t)}) \\
\tilde{\lambda}^{(t+1)} &= \tilde{\lambda}^{(t)} + \gamma_t (\lambda(\theta^{(t)}) - \tilde{\lambda}^{(t)}).
\end{aligned}$$

The ODE is then the differential equation analog

$$\begin{aligned}
\dot{\theta}_t &= -\nabla \lambda(\theta_t) \\
\dot{\tilde{\lambda}}_t &= \lambda(\theta_t) - \tilde{\lambda}_t.
\end{aligned} \tag{4.12}$$

By construction, $\lambda(\theta_t) \geq 0$ is bounded from below (see (4.5)). It is also a non-increasing function because $\dot{\lambda}(\theta_t) = -\|\nabla \lambda(\theta_t)\|^2 \leq 0$. Thus $\lambda(\theta_t)$ must converge to some non-negative limit and $\nabla \lambda(\theta_t)$ must converge to the zero vector as $t \to \infty$. Using (4.12), we conclude that $\tilde{\lambda}_t$ must also converge to the same limit as $\lambda(\theta_t)$.

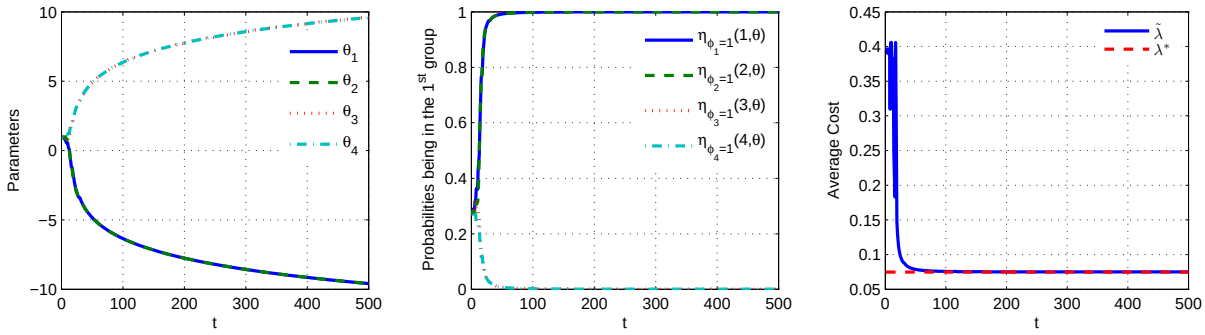
The analogous statements for the stochastic approximation recursion are contained in the following proposition. The proof follows from stability of the ODE and standard arguments (see e.g. Theorem 2 of Chapter 6 of [11]).

Proposition 4.3.1 *Let Assumption 5.4.3 hold, and assume that the parameter vector sequence $\{\theta^{(t)}\}$ and average cost sequence $\{\tilde{\lambda}^{(t)}\}$ are updated according to (4.10) and (4.11) respectively. Then, the sequence $\lambda(\theta^{(t)})$ converges almost surely and*

$$\nabla \lambda(\theta^{(t)}) \xrightarrow{a.s.} 0, \quad \text{as } t \to \infty.$$

Moreover, $\tilde{\lambda}^{(t)}$ also converges almost surely to the same limit, i.e.

$$\tilde{\lambda}^{(t)} \xrightarrow{a.s.} \lambda(\theta^{(t)}), \quad \text{as } t \to \infty.$$



(a) (b) (c)

Figure 4.3: Plots of (a) $\theta^{(t)}$, (b) $\eta_\phi(\cdot, \theta^{(t)})$, and (c) $\tilde{\lambda}^{(t)}$ for 4-state Markov chain with the idealized gradient algorithm (4.9).

4.4 Numerical Results

4.4.1 4-state Markov chain

Consider a stationary 4-state Markov chain with transition matrix

$$P = \begin{bmatrix} 0.5 & 0.4 & 0.0 & 0.1 \\ 0.4 & 0.5 & 0.1 & 0.0 \\ 0.0 & 0.1 & 0.5 & 0.4 \\ 0.1 & 0.0 & 0.4 & 0.5 \end{bmatrix}$$

and the invariant distribution $\pi = [0.25, 0.25, 0.25, 0.25]$.

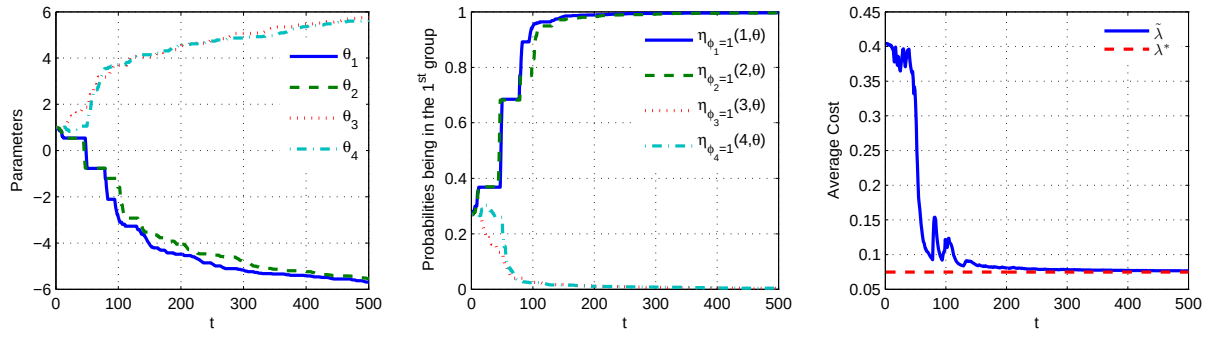
The optimal bi-partition function for this Markov chain is $\phi^* = [1, 1, 2, 2]$, i.e. the state $\{1, 2\}$ are aggregated into one group and the states $\{3, 4\}$ are aggregated into another group. The optimal aggregated Markov chain is computed according to (2.11) as

$$Q^* = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$$

and the invariant distribution $\varpi^* = [0.5, 0.5]$.

1) Dynamic programming: Since the state space is relatively small, the dimension of bi-partition function space is $L = 2^4$. The optimal partition function ϕ^* is numerically obtained by using the policy iteration algorithm for the solution of (4.4). The optimal average cost $\lambda^* = 0.0749$.

2) Idealized gradient algorithm: For this small-scale example, we can efficiently compute $\nabla\lambda(\theta)$ for any given θ. We can thus implement the idealized gradient algorithm (4.9). The evolution of this algorithm, starting with $\theta^{(0)} = [1, 1, 1, 1]$, is shown in Figure 4.3. After 100 iterations, parameter vector $\theta^{(100)} = [-6.3411, -6.3464, 6.3575, 6.3580]$,



(a) (b) (c)

Figure 4.4: Plots of (a) $\theta^{(t)}$, (b) $\eta_\phi(\cdot, \theta^{(t)})$, and (c) $\tilde{\lambda}^{(t)}$ for 4-state Markov chain with the simulation-based algorithm (4.10) and (4.11).

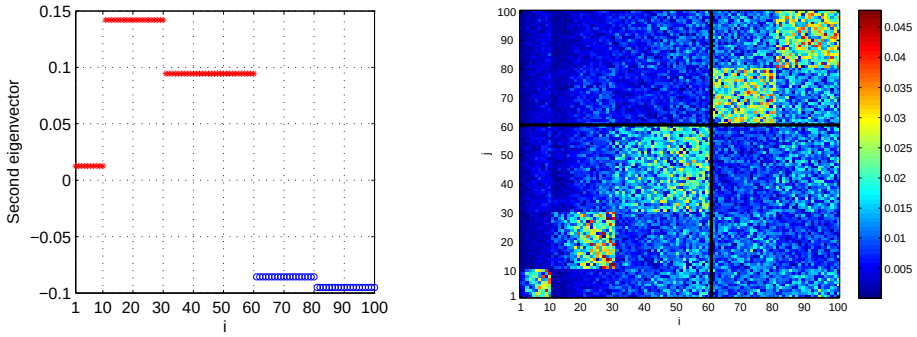
and the probabilities of states being in the first group are $\eta_{\phi=[1,1,1,1]}(\cdot, \theta^{(100)}) = [0.9982, 0.9983, 0.0017, 0.0017]$. Thus, we obtain the optimal partition function as $\phi^* = [1, 1, 2, 2]$ with high probability. The corresponding estimated average cost is given by $\tilde{\lambda}^{(100)} = 0.0757$, which is very close to the optimal value.

3) Simulation-based algorithm: We consider a single sample path of the Markov chain simulated according to the transition matrix P. The updating algorithms (4.10) and (4.11) are implemented at every time step along the sample path. We start with the same initial parameter vector $\theta^{(0)} = [1, 1, 1, 1]$. The evolution of parameters and the average cost are give in Figure 4.4. After 500 iterations, the simulation-based algorithm has comparable performance to the idealized gradient algorithm after 100 iterations. The parameter vector $\theta^{(500)} = [-5.6893, -5.5390, 5.7600, 5.6252]$, and the probabilities of states being in the first group are $\eta_{\phi=[1,1,1,1]}(\cdot, \theta^{(500)}) = [0.9966, 0.9961, 0.0031, 0.0036]$. From this, the optimal partition function $\phi^* = [1, 1, 2, 2]$ can be determined with high probability. The corresponding estimated average cost is equal to $\tilde{\lambda}^{(500)} = 0.0766$, which is a little larger than the optimal value.

4.4.2 100-state Markov chain

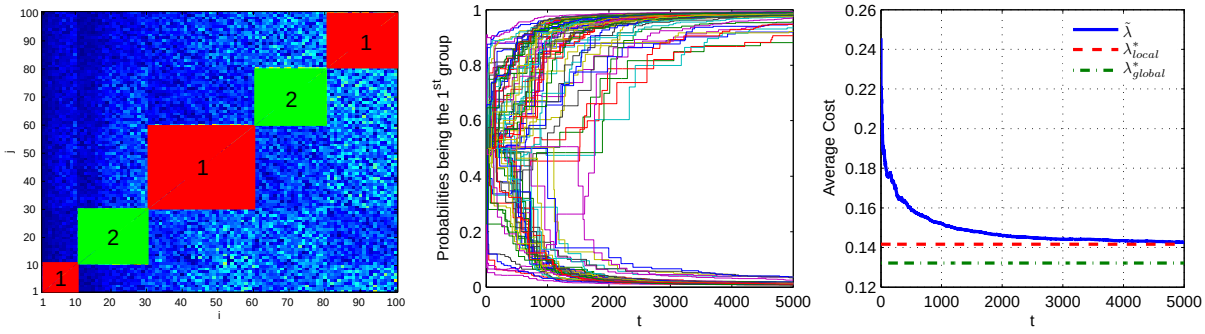
We consider the same 100-state Markov chain as described in Section 3.4.2. It is shown that a spectral relaxation of the bi-partition problem gives a solution according to the sign-structure of the second eigenvector, depicted in Figure 4.5 (a). This optimal (spectral) bi-partition of the state space is: States $1 - 60$ are aggregated as the first group and states $61 - 100$ are aggregated as the second group. It is shown in Figure 4.5 (b). The corresponding K-L metric (average cost) is given by $\lambda^*_{global} = 0.1321$. The subscript *"global"* refers to the fact that the spectral bi-partition gives the global minimum of the K-L metric for the bi-partition problem (4.3).

For the simulation-based algorithm, a single sample path of the Markov chain was obtained according to the transition matrix P. θ is now a 100-dimensional parameter vector initialized as $\theta^{(0)} = [0, 0, \ldots, 0]$. At time t, the vector $\theta^{(t)}$ is updated according to (4.10) and the estimated average cost $\tilde{\lambda}^{(t)}$ is updated according to (4.11).



(a) (b)

Figure 4.5: (a) A plot of the second eigenvector of the 100×100 transition matrix P. (b) The graph of transition probabilities for a 100-state nearly complete decomposable Markov chain (see [31, 77]). The optimal bi-partition of the state space is indicted by two bold lines: States $1 - 60$ are aggregated as the first group and states $61 - 100$ are aggregated as the second group.



(a) (b) (c)

Figure 4.6: Plots of (a) the bi-partition obtained using $\theta^{(5000)}$, (b) $\eta_\phi(\cdot, \theta^{(t)})$, and (c) $\tilde{\lambda}^{(t)}$ for 100-state Markov chain with the simulation-based algorithm.

Figure 4.6 (b) depicts $\frac{1}{1+\exp(M\theta_i^{(t)})}$ as a function of t for $i \in \mathcal{N}$. It reflects the probability that i^{th} state is assigned to the 1^{st} group (see Section 4.3.2). For large t, each state is assigned with high probability to one of the two groups. The resulting two-group assignments at time $t = 5000$ are obtained with high probability and depicted in Figure 4.6 (a): The first group is formed by states $1 - 10$ (first block), $31 - 60$ (third block) and $81 - 100$ (fifth block); The second group is formed by states $11 - 30$ (second block) and $61 - 80$ (fourth block). The partition found using the simulation-based algorithm is *not* the global optimum for the bi-partition problem (compare Figure 4.5 with Figure 4.6 (a)). The algorithm converges to a local minimum. In Figure 4.6 (c), we depict the plots of estimated average cost $\tilde{\lambda}^{(t)}$, the average cost $\lambda^*_{local} = 0.1416$ for the local optimal partition shown in Figure 4.6 (a), and the average cost λ^*_{global} for the global optimal partition shown in Figure 4.5.

The convergence to a local optimum is due to the non-convexity of the average cost $\lambda(\theta)$ with respect to the parameter θ. In general, one is guaranteed to obtain only a local minimum with a gradient-descent scheme [11]. For the Markov chain shown in Figure 4.5, one can also obtain multiple partitions by recursively applying the bi-partition algorithm (see Section 3.4.2 for more details).

4.4.3 A discussion of computational requirements

Even for the Markov chain with large state space, the computation burden of the simulation-based algorithm is manageable. For this algorithm, an update at time t only requires the evaluation of a single term

$$\sum_{j \in \mathcal{N}} P_{X(t)j} \log \left(\frac{P_{X(t)j}}{\widehat{Q}^{(\pi)}_{X(t)j}(\phi)} \right)$$

for entries with $P_{X(t)j} > 0$, where $X(t)$ is the state at time t. For many applications, $P_{X(t)j} > 0$ only for states j in a small neighborhood of $X(t)$. Such a requirement is also natural for simulation purpose. The computational complexity for evaluating this single term does not grow even if the size of the state space $\mathcal{N}$ increases.

4.5 Conclusions

In this chapter, we proposed a simulation-based algorithm to aggregate a large-scale Markov chain to obtain a reduced-order Markov model. There are two main concepts in the chapter: 1) the dynamic programming formulation of the model reduction problem; 3) its solutions using a simulation-based (stochastic-approximation) algorithm. The proposed framework requires only a single sample path of the Markov chain, and has a small computation burden even for very large problems.

Chapter 5

Model Reduction of Hidden Markov Models

5.1 Introduction

A fundamental problem for Hidden Markov Models (HMMs) that arise in applications is the large size of the underlying state space [16]. For example, the genome of an organism is an enormously long sequence of symbols from the four-symbol alphabet {A (Adenine), C (Cytosine), G (Guanine), or T (Thymine)}. The hidden Markov model that arises in the gene finding problem usually has a large number of states, 4^k, where k is the order of the underlying Markov chain [118].

In Chapter 2 to Chapter 4, we introduced an information-theoretic framework to reduce the Markov chain models via aggregation of the state space. By using the fact that the joint state and observation process of HMM is Markovian, we extended this aggregation framework for the model reduction of HMMs based on the *K-L divergence rate between laws of the joint process* [29, 116]. Similar to the model reduction of Markov chains, we obtain the optimal representation of the aggregated HMM for any fixed partition function.

However, the problem with an aggregation based on joint process is that two HMMs may have very similar laws of the observation process, while the K-L divergence rate of their joint laws might be very large or even unbounded. In this chapter, we employ the *K-L divergence rate between laws of the observation process* as the "probability distance" to compare two HMMs. If this distance is zero, then two HMMs are equivalent in the probability sense up to a permutation of the state space (see Section 5.2.1 for more details). This K-L divergence rate pseudo-metric has been studied in statistics [72], speech recognition [59], bioinformatics [118], and control theory [111, 123].

One of the problems with using the K-L divergence rate based on the observation process alone is that, in general, it does not have an explicit expression in terms of the parameters of HMMs [123]. The K-L divergence rate can be approximated by using a nonlinear filter along a sample path of the observation process of the HMM [37, 72]. However, the computational complexity is usually high for HMMs with large state spaces [66]. The computational barrier makes the optimization and estimation using K-L divergence rate a challenging problem in practice.

The goal of this chapter is to find a reduced model of HMM via aggregation of the state space by minimizing the K-L divergence rate between original and reduced laws of the observation process. There are two main ideas:

One, we use the optimal representation of the aggregated HMM derived in our earlier work [29, 116] as a structured model for optimization. We take advantage of the optimal representation to overcome some of the complexity issues in computing the K-L divergence rate. The second idea is to generate observations from the original HMM with large state space, but to recursively evaluate the filter only for the aggregated HMM with much smaller state space. The aggregated HMM is represented in terms of parameters from the original HMM and the partition function, which needs to be optimized.

In the third part, we extend the simulation-based aggregation method described in Chapter 4 to the HMMs. We first parameterize the discrete partition function space into a smaller real parameter space [29], and then convert the optimal partition problem to an optimal estimation problem, in fact, the Maximum Likelihood Estimation (MLE) problem of the HMM [37, 72]. We employ a gradient-based simulation algorithm to solve the MLE problem. The algorithm is recursively updated based on the stochastic gradient of the nonlinear filter evaluated using the aggregated HMM model. The convergence of the algorithm is established based on the stochastic approximation arguments as well as the the ergodicity of the filtering process.

The outline of this chapter is as follows: Preliminaries and notations are contained in Section 5.2. In Section 5.3, we briefly review the model reduction of HMM via aggregation of the state space, and connect the optimal partition problem to the maximum likelihood estimation problem. The recursive learning algorithm is discussed in Section 5.4, and illustrated with numerical examples in Section 5.5.

5.2 Preliminaries and Notations

5.2.1 Hidden Markov Model

In this chapter, we consider a discrete-time HMM $\{X_n, Y_n\}_{n\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Without loss of generality, we assume that $(\Omega, \mathcal{F}, \mathsf{P})$ is a canonical probability space and the $\{X_n, Y_n\}_{n\geq 0}$ is a coordinate process taking values on the product space $\mathcal{N} \times \mathcal{O}$, where finite sets $\mathcal{N} = \{1, \ldots, N\}$ and $\mathcal{O} = \{1, \ldots, O\}$ denote the state space and observation space, respectively.

The unobserved *state process* $\{X_n\}_{n\geq 0}$ is a time-homogeneous Markov chain with the initial distribution μ and the transition matrix A. For any time $n \geq 0$ and any $i, j \in \mathcal{N}$,

$$\mathsf{P}(X_0 = i) = \mu_i, \quad \mathsf{P}(X_{n+1} = j | X_n = i) = A_{ij}.$$

The n-step distribution of the chain is then given by $\mathsf{P}(X_n = i) = (\mu A^n)_i$.

The *observation process* $\{Y_n\}_{n\geq 0}$ are mutually independent conditioned on the state process of the Markov chain,

i.e., for any time $n \geq 0$, any $i_0, \ldots, i_n \in \mathcal{N}$, and any $r_0, \ldots, r_n \in \mathcal{O}$,

$$\mathsf{P}(Y_n = r_n, \ldots, Y_0 = r_0 | X_n = i_n, \ldots, X_0 = i_0) = \prod_{k=0}^{n} \mathsf{P}(Y_k = r_k | X_k = i_k).$$

The conditional distribution of Y_n only depends on X_n, which can be described by the transition matrix C. For any time $n \geq 0$, any $i \in \mathcal{N}$, and any $r \in \mathcal{O}$,

$$\mathsf{P}(Y_n = r | X_n = i) = C_{ir}.$$

For any $r \in \mathcal{O}$, denote the diagonal matrix $B(r) := \mathrm{diag}(b_i(r))$, where the vector $b(r) = [C_{1r}, C_{2r}, \ldots, C_{Nr}]^T$.

The complete statistics of the HMM $\{X_n, Y_n\}_{n\geq 0}$ are fully characterized by a model, denoted by $\xi = (\mu, A, C)$. For an HMM with the parameter set ξ, we denote the probability measure and associated expectation as P_ξ and E_ξ, respectively. The law of the observations does not change up to *a permutation of the state space*: Let $\Sigma_\mathcal{N}$ be the symmetric group containing all permutations of the set $\mathcal{N}$. For any $\sigma \in \Sigma_\mathcal{N}$, consider a bijection φ_σ mapping one HMM $\xi = (\mu, A, C)$ to another HMM $\varphi_\sigma(\xi) = (\sigma\mu, \sigma A, \sigma C)$, where $(\sigma\mu)_i := \mu_{\sigma(i)}$, $(\sigma A)_{ij} := A_{\sigma(i)\sigma(j)}$, $(\sigma C)_{ir} := C_{\sigma(i)r}$ for $i, j \in \mathcal{N}$ and $r \in \mathcal{O}$. Observe that $\mathsf{P}_{\varphi_\sigma(\xi)} = \mathsf{P}_\xi$ for almost all $\{Y_n\}_{n\geq 0}$.

Throughout the chapter it is assumed that:

Assumption 5.2.1 (Ergodicity) *All underlying Markov chains are assumed to be* irreducible *and* aperiodic.

Under Assumption 5.2.1, there exists a unique *invariant distribution* π such that $\pi = \pi A$. In fact, the chain is *geometrically ergodic*, i.e., the n-step distribution of the chain converges geometrically fast to the invariant distribution π in total variation sense [16].

Assumption 5.2.2 (Nondegeneracy) *The transition matrix C is strictly positive, i.e., $C_{ir} > 0$ for any $i \in \mathcal{N}$ and $r \in \mathcal{O}$.*

Under Assumption 5.2.2, the unobserved state process $\{X_n\}_{n\geq 0}$ can be statistically inferred from any sample path of observations of the observed process $\{Y_n\}_{n\geq 0}$.

5.2.2 Filter recursion and its stability

For an HMM, an important problem is to compute the *prediction filter*: For any time $n \geq 0$ and any $i \in \mathcal{N}$,

$$p_n(i) := \mathsf{P}(X_n = i | Y_{n-1}, \ldots, Y_0)$$

where we take $p_0 = \mu$. The prediction filter is used to obtain the *predictive distribution* of the observations: For any $n \geq 0$,

$$\mathsf{P}(Y_n|Y_{n-1}, \ldots, Y_0) = b^T(Y_n)p_n. \tag{5.1}$$

The solution to the HMM filtering problem is *recursive* in nature. For any time $n \geq 0$,

$$p_{n+1} = \frac{A^T B(Y_n)p_n}{b^T(Y_n)p_n}. \tag{5.2}$$

The recursive nature of the filter is inherited from the Markovian nature of the state process, and is computationally very convenient for on-line estimation.

The recursive filter defined in (5.2) is *exponentially stable* for general HMMs [16, 42] under egodicity and nondegeneracy assumptions. Here we state the results of [16] for HMMs defined on finite state and observation spaces.

Proposition 5.2.1 *Suppose Assumption 5.2.1 and Assumption 5.2.2 hold. Then, for any two distributions μ and ν, there exists constants $0 < C_1 < \infty$, $0 < C_2 < \infty$, and $0 < \rho < 1$ such that*

(i) *For any $n \geq 0$,*

$$\|p_{n+1}^{\mu} - p_{n+1}^{\nu}\|_{\mathrm{TV}} \leq C_1(1-\rho)^n \|\mu - \nu\|_{\mathrm{TV}}$$

where p_{n+1}^{μ} and p_{n+1}^{ν} denote two filter recursions defined in (5.2) starting with initial distributions μ and ν, respectively.

(ii) *For any $0 \leq k \leq n$,*

$$\|\mathsf{P}^{\mu}(X_{n+1}|Y_0^n) - \mathsf{P}^{\mu}(X_{n+1}|Y_k^n)\|_{\mathrm{TV}} \leq C_2(1-\rho)^{n-k}$$

where P^{μ} denotes the probability measure with the initial distribution μ.

The stability of the filter implies that the extended Markov chain $\{X_n, Y_n, p_n\}_{n\geq 0}$ is geometrically ergodic [42]. Thus, the initial distributions are forgotten exponentially fast and are hence asymptotically not important in the analysis of the filtering process.

5.2.3 Probability distance between HMMs

In this section, we define the probability distance between two HMMs using the Kullback-Leibler divergence rate. For two HMMs $\xi = (\mu, A, C)$ and $\bar{\xi} = (\bar{\mu}, \bar{A}, \bar{C})$ defined on the same observation space $\mathcal{O}$ (but not necessarily on the

same state space), we consider the K-L metric between laws of the observations [59]:

$$R(\xi\|\bar{\xi}) := \lim_{n\to\infty} \frac{1}{n} D\left(\mathsf{P}_\xi(Y_0^n)\|\mathsf{P}_{\bar{\xi}}(Y_0^n)\right) = \lim_{n\to\infty} \frac{1}{n}\mathsf{E}_\xi\left[\log \frac{\mathsf{P}_\xi(Y_0^n)}{\mathsf{P}_{\bar{\xi}}(Y_0^n)}\right].$$

As shown in [37, 123], the following asymptotic results can be established under Assumption 5.2.1 and Assumption 5.2.2: There exist finite constants $H(\xi,\xi)$ and $H(\xi,\bar{\xi})$ such that the following limits exist in P_ξ-a.s. sense:

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}_\xi(Y_0^n) = \lim_{n\to\infty} \frac{1}{n}\mathsf{E}_\xi\left[\log \mathsf{P}_\xi(Y_0^n)\right] = H(\xi,\xi), \tag{5.3}$$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathsf{P}_{\bar{\xi}}(Y_0^n) = \lim_{n\to\infty} \frac{1}{n}\mathsf{E}_\xi\left[\log \mathsf{P}_{\bar{\xi}}(Y_0^n)\right] = H(\xi,\bar{\xi}). \tag{5.4}$$

The convergence of (5.3) follows directly from the Shannon-McMillan-Breiman theorem for finite-valued stationary ergodic process [123] and the limit $H(\xi,\xi)$ is equal to the entropy rate of the observation process $\{Y_n\}_{n\geq 0}$. The convergence of (5.4) was first established in [6] for finite-valued stationary ergodic HMMs. Alternatively, we also note that

$$\frac{1}{n} \log \mathsf{P}_{\bar{\xi}}(Y_0^n) = \frac{1}{n}\sum_{k=0}^{n} \log \mathsf{P}_{\bar{\xi}}(Y_k|Y_0^{k-1})$$

where the predictive distribution $\mathsf{P}_{\bar{\xi}}(Y_k|Y_0^{k-1})$ can be expressed as a measurable function of an ergodic Markov chain $\{X_n, Y_n, p_n\}_{n\geq 0}$ (see (5.1) and comments after Proposition 5.2.1). Then the convergence of (5.4) follows from the ergodic theorem of Markov chains [83].

Thus, the probability distance between two HMMs is well-defined through the K-L metric between laws of the observations:

$$R(\xi\|\bar{\xi}) = H(\xi,\xi) - H(\xi,\bar{\xi}). \tag{5.5}$$

In general, we do not have an explicit expression for $R(\xi\|\bar{\xi})$ in terms of parameters of HMMs ξ and $\bar{\xi}$. The prediction filter is usually employed to approximate the K-L metric given a sufficient number of observations [37, 72].

5.3 Model Reduction of HMM

5.3.1 Reduction via aggregation of state space

The goal of this chapter is to obtain a reduced order HMM through the aggregation of the state space. The model reduction error is expressed in terms of the K-L metric between laws of observations for the original and reduced models.

Consider the HMM $\xi = (\mu, A, C)$ defined on the state space $\mathcal{N}$ and the observation space $\mathcal{O}$. We want to

find another HMM $\bar{\xi} = (\bar{\mu}, \bar{A}, \bar{C})$ defined on the state space $\mathcal{M} = \{1, \ldots, M\}$ with cardinality $M \leq N$ and the observation space $\mathcal{O}$ such that the probability distance $R(\xi \| \bar{\xi})$ is minimized. Additionally, we want the reduced HMM $\bar{\xi}$ to be obtained by aggregating the state space of the HMM ξ. Then any state of the HMM $\bar{\xi}$ can be thought of as a *super-state* of the HMM ξ, which capture the coarse behavior of a group of states of ξ. The relationship between $\mathcal{N}$ and $\mathcal{M}$ is described by a partition function ϕ (see Definition 2.3.1). Let Φ denote all M-partition functions from $\mathcal{N}$ to $\mathcal{M}$. Note that Φ is a finite set with $L = M^N$ elements and we can write it as $\Phi = \{\phi_1, \phi_2, \ldots, \phi_L\}$.

As shown in our prior work [29, 116], an optimal representation of the aggregated HMM, (5.6)–(5.8) below, is obtained by minimizing the *K-L metric between joint laws of the states and observations* together. Given the focus vision to K-L metric between laws of the observations, it would have been ideal to construct an optimal model based on the observations alone. This however is a difficult problem. Instead, we use the representation (5.6)–(5.8) for the aggregated HMM. The problem of optimal partition selection is based on the K-L metric between laws of the observations.

For any fixed partition function $\phi \in \Phi$, the aggregated HMM $\bar{\xi}(\phi) = (\bar{\mu}(\phi), \bar{A}(\phi), \bar{C}(\phi))$ is represented as a function of ϕ (see e.g. Theorem 2 of [29]).

$$\bar{\mu}_k(\phi) = \sum_{i \in \phi^{-1}(k)} \mu_i, \quad k \in \mathcal{M} \tag{5.6}$$

$$\bar{A}_{kl}(\phi) = \frac{\sum_{i \in \phi^{-1}(k)} \pi_i \sum_{j \in \phi^{-1}(l)} A_{ij}}{\sum_{i \in \phi^{-1}(k)} \pi_i}, \quad k, l \in \mathcal{M} \tag{5.7}$$

$$\bar{C}_{kr}(\phi) = \frac{\sum_{i \in \phi^{-1}(k)} \pi_i C_{ir}}{\sum_{i \in \phi^{-1}(k)} \pi_i}, \quad k \in \mathcal{M},\ r \in \mathcal{O}. \tag{5.8}$$

The invariant state distribution of $\bar{\xi}(\phi)$ is given by

$$\bar{\pi}_k(\phi) = \sum_{i \in \phi^{-1}(k)} \pi_i, \quad k \in \mathcal{M}.$$

For any fixed $\phi \in \Phi$, we observe that the aggregated HMM $\bar{\xi}(\phi)$ satisfies both Assumption 5.2.1 and Assumption 5.2.2, i.e., the underlying aggregated Markov chain with the transition matrix $\bar{A}(\phi)$ is ergodic and the transition matrix $\bar{C}(\phi)$ is non-degenerate. Thus the probability distance $R(\xi \| \bar{\xi}(\phi))$ is well-defined for any $\phi \in \Phi$. We also observe that:

$$\begin{aligned}
\bar{\mu}_k(\phi) &= \mathsf{P}_\xi^\mu(X_0 \in \phi^{-1}(k)) \\
\bar{A}_{kl}(\phi) &= \mathsf{P}_\xi^\pi(X_{n+1} \in \phi^{-1}(l) | X_n \in \phi^{-1}(k)) \\
\bar{C}_{kr}(\phi) &= \mathsf{P}_\xi^\pi(Y_n = r | X_n \in \phi^{-1}(k))
\end{aligned}$$

where P_ξ^μ and P_ξ^π denote probability measures with initial distributions μ and π, respectively. This result is consistent with the optimal prediction theory from the statistical mechanics [38].

5.3.2 Maximum likelihood estimation formulation

For a fixed partition function $\phi \in \Phi$, the aggregated HMM is represented as $\bar{\xi}(\phi) = (\bar{\mu}(\phi), \bar{A}(\phi), \bar{C}(\phi))$. The problem then is to find the optimal ϕ^* such that

$$\phi^* \in \arg\min_{\phi \in \Phi} R(\xi \| \bar{\xi}(\phi))$$

which, after using (5.5), is equivalent to the following maximization problem:

$$\phi^* \in \arg\max_{\phi \in \Phi} H(\xi, \bar{\xi}(\phi)). \tag{5.9}$$

Due to the almost sure convergence of log-likelihood function to the limit $H(\xi, \bar{\xi}(\phi))$ (see (5.4)), we instead consider the following stochastic counterpart of (5.9):

$$\widehat{\phi}_n \in \arg\max_{\phi \in \Phi} l_n(\phi) \tag{5.10}$$

where the *log-likelihood rate* is defined as

$$l_n(\phi) := \frac{1}{n} \log \mathsf{P}_{\bar{\xi}(\phi)}(y_0^n) \tag{5.11}$$

with observations $\{y_0, \ldots, y_n\}$ generated from the HMM ξ.

The optimization problem (5.10) is the *maximum likelihood estimation* in statistics: In effect, we select the partition function which gives the highest probability of the observations generated from the true model. Note that the objective function (5.10) converges to the objective function of (5.9) in P_ξ-a.s. sense. One may wonder whether $\widehat{\phi}_n \to \phi^*$ P_ξ-a.s. as $n \to \infty$. The answer to this question is affirmative due to the fact that the partition function space Φ is a finite set.

Proposition 5.3.1 *Let Φ denote a finite partition function space and consider an equivalent class in Φ*

$$\Phi^e := \{\phi \in \Phi : \mathsf{P}_{\bar{\xi}(\phi)} = \mathsf{P}_{\bar{\xi}(\phi^*)} \text{ for almost all } \{Y_n\}_{n \geq 0}\}.$$

Then P_ξ-a.s.,

(i) For any $\phi \in \Phi$, we have $H(\xi, \bar{\xi}(\phi)) \leq H(\xi, \bar{\xi}(\phi^*))$ where the equality holds if and only if $\phi \in \Phi^e$.

(ii) Maximum likelihood estimation is consistent: $\widehat{\phi}_n \to \phi^e$ *as* $n \to \infty$ *for some* $\phi^e \in \Phi^e$.

Remark 5.3.1 As shown at the beginning of the paper, a permutation of the state space of the HMM will generate the same law of observations. Thus the equivalent class Φ^e is non-empty in general.

5.3.3 Hypothesis testing-based approach for optimal partition selection

The observations $\{y_0, y_1, \ldots, y_n\}$ are generated according to the original HMM defined on the state space $\mathcal{N}$ and the observation space $\mathcal{O}$. For a fixed partition function $\phi \in \Phi$, the log-likelihood rate $l_n(\phi)$ is recursively computed using the filtering recursion (5.2) with the aggregated HMM $\bar{\xi}(\phi)$ defined on the state space $\mathcal{M}$ and the observation space $\mathcal{O}$. The complexity of the filtering recursion only depends on the dimension of the aggregated HMM. Even if the original HMM has a huge state space, we can still efficiently evaluate the filtering recursion using aggregated HMMs with much smaller state space.

Since the partition function space Φ is a finite set, the optimization problem (5.10) can in practice be approached through the *hypothesis testing*: We are given $|\Phi|$ different hypotheses (or $|\Phi|$ different aggregated HMMs), and our goal is to decide on the basis of observations alone which of the hypotheses holds true (or which of the aggregated HMM is with the maximum log-likelihood rate). If the set Φ is of moderate size, then the maximum log-likelihood rate hypothesis can be found efficiently. All we need to do is to compute $|\Phi|$ different filters, one for each partition function. For any fixed-length observations $\{y_0, y_1, \ldots, y_n\}$, we choose the n-step hypothesis $\widehat{\phi}_n$ as the one with the largest log-likelihood rate. Then $\widehat{\phi}_n$ asymptotically converges to the global maximum ϕ^* as $n \to \infty$ (see Proposition 5.3.1).

5.4 Recursive Learning Algorithm

In general, the optimization problem (5.10) is intractable because of the curse of dimensionality. The curse here arises due to the large size of the partition function space, e.g., $L = |\Phi| = M^N$ for the M-partition of the N-state space. To confront this complexity issue, a parametric representation is used to represent the partition function in terms of a small number of parameters. A recursive learning algorithm is described to adaptively update the parameters based on a sample path of the HMM.

5.4.1 Parameterization of the partition function space via randomization

The randomization of the partition function gives us greater flexibility to solve the optimization problem (5.10). A *randomized partition policy* is defined as a mapping,

$$\eta : \mathcal{N} \to [0,1]^L$$

with the component $\eta_\phi(i)$ such that $\sum_{\phi\in\Phi}\eta_\phi(i)=1$ for every $i\in\mathcal{N}$. Under a policy η, the partition function ϕ is assigned to the state i with the probability $\eta_\phi(i)$, independent of everything else.

The policy is said to be deterministic if for every state i, there is a single partition function $\phi^{(i)}$ such that $\eta_{\phi^{(i)}}(i)=1$. If the function $\phi^{(i)}$ is the same for all i then the policy η yields a consistent partition of the space $\mathcal{N}$. If $\eta(\cdot)$ is a degenerate probability distribution (i.e., a dirac delta in the probability simplex of Φ), then a partition function can be uniquely obtained from $\eta(\cdot)$. In practice, a numerical method will in general only lead to a partition function with high probability determined by $\eta(\cdot)$.

The combinatorial optimization problem (5.10) involves a very large partition space Φ. Following the consideration of [29], we consider the randomized policies $\eta(\cdot;\theta)$ which are described in terms of a parameter vector $\theta=(\theta(1),\ldots,\theta(K))^T$, where the dimension K is chosen much smaller than L, the dimension of Φ.

The following assumption is made for the ease of the optimization over the parameter θ:

Assumption 5.4.1 *The parameter space Θ is a compact subset of a K-dimensional real vector space $\mathbb{R}^K$. For any $i\in\mathcal{N}$, the randomized and parameterized policy $\eta(i;\theta)$ is twice differentiable with respect to θ, and has bounded first and second derivatives for all $\theta\in\Theta$.*

5.4.2 Parametric representation of the MLE problem

For any $\theta\in\Theta$, we consider a randomized partition policy $\eta(\cdot;\theta)$ such that for every $i\in\mathcal{N}$, $\eta(i;\theta)$ depends smoothly on θ, $\eta_\phi(i;\theta)\geq 0$, and $\sum_{\phi\in\Phi}\eta_\phi(i;\theta)=1$. We associate a probability measure $\mathsf{P}_{\eta(\cdot;\theta)}$ and the corresponding expectation $\mathsf{E}_{\eta(\cdot;\theta)}$ with the policy $\eta(\cdot;\theta)$. For any measurable function $f(\phi)$, we define

$$\mathsf{E}_{\eta(\cdot;\theta)}[f(\phi)]:=\sum_{\phi\in\Phi}\eta_\phi(\cdot;\theta)f(\phi).$$

The parameterized one-step log-likelihood can also be defined: For any $n\geq 0$,

$$g_n(\theta):=\mathsf{E}_{\eta(X_n;\theta)}\left[\log\left(\mathsf{P}_{\bar{\xi}(\phi)}(Y_n|Y_0^{n-1})\right)\right]$$

where X_n is the hidden state associated with the observation Y_n generated from the HMM ξ.

The parameterized maximization problem is defined as

$$\theta^*\in\arg\max_{\theta\in\Theta}\widetilde{H}(\theta) \tag{5.12}$$

where the parameterized average cost is given by

$$\widetilde{H}(\theta) = \lim_{n\to\infty} \frac{1}{n} \mathsf{E}_\xi \left[\sum_{k=0}^{n} g_k(\theta) \right].$$

The parameterized maximum likelihood estimation (MLE) is the stochastic counterpart of (5.12):

$$\widehat{\theta}_n \in \arg\max_{\theta\in\Theta} \tilde{l}_n(\theta) \tag{5.13}$$

where the parameterized log-likelihood rate is defined as

$$\tilde{l}_n(\theta) = \frac{1}{n} \sum_{k=0}^{n} g_k(\theta).$$

5.4.3 Recursive learning algorithm and its convergence

Under Assumption 5.2.1–5.4.1, one can show that the MLE $\widehat{\theta}_n$ converge to θ^* P_ξ-a.s as $n \to \infty$. However, the maximum of (5.12) or (5.13) with respect to θ is typically very difficult to compute. Instead, we describe a recursive learning algorithm that searches for a maximum along the gradient-ascent direction of the log-likelihood rate $\tilde{l}_n(\theta)$.

In order to compute the gradient of $\tilde{l}_n(\theta)$, we employ the simulation to produce a sample-based estimate $\nabla h_n(\theta)$ of $\nabla\tilde{l}_n(\theta)$ (we denote $\nabla := \nabla_\theta$ for short). At every time step n, the estimate h_n is computed using the current observation as well as finite length of past observations: For any time $n \geq 0$,

$$h_n(\theta) := \frac{1}{\lfloor m_n \rfloor + 1} \left(\sum_{k=n-\lfloor m_n \rfloor}^{n} \widetilde{g}_k(\theta) \right) \tag{5.14}$$

where the finite-length log-likelihood

$$\widetilde{g}_k(\theta) := \mathsf{E}_{\eta(X_k;\theta)} \left[\log \left(\mathsf{P}_{\bar{\xi}(\phi)}(Y_k | Y_{n-\lfloor m_n \rfloor}^{k-1}) \right) \right]$$

and the averaging sequence $\{m_n\}_{n\geq 0}$ satisfies the following assumption:

Assumption 5.4.2 *For any $n \geq 0$,*

$$0 \leq m_0 \leq m_1 \leq \ldots \leq m_{n-1} \leq m_n \leq n$$

and as $n \to \infty$, $m_n \to \infty$.

Given any partition function ϕ and the observations $\{Y_{n-\lfloor m_n \rfloor}, \ldots, Y_n\}$, the estimate h_n can be computed through

the filter recursion (5.2) of $\{\mathsf{P}_{\bar{\xi}(\phi)}(Y_{n-\lfloor m_n \rfloor}), \ldots, \mathsf{P}_{\bar{\xi}(\phi)}(Y_n | Y_{n-\lfloor m_n \rfloor}^{n-1})\}$. Due to the ergodicity of the filter (see Proposition 5.2.1 (ii)), the recursion can be started with an arbitrary initial distribution $\bar{\mu}$ on $\mathcal{M}$.

The estimate $\nabla h_n(\theta)$ asymptotically converges to $\nabla \tilde{l}_n(\theta)$ as $n \to \infty$ and the convergence is geometrically fast due to the ergodicity of the filter. By choosing the sequence $\{m_n\}_{n \geq 0}$ alternatively, we compute $h_n(\theta)$ efficiently: e.g., one can take $m_n = n^\alpha$ where selecting $\alpha \in (0, 1]$ allows one to tradeoff between the computation efficiency and the estimation performance.

A recursive learning algorithm is employed to approach the optimization problem (5.12). Let $\{x_n, y_n\}_{n \geq 0}$ denote a sample path generated from the HMM ξ. The recursive learning algorithm for updating the parameter vector is given by: For any $n \geq 0$,

$$\bar{\theta}_{n+1} = \bar{\theta}_n + \gamma_n \nabla h_n(\bar{\theta}_n) \tag{5.15}$$

where $\bar{\theta}_0$ is taken to be an arbitrary point in Θ, the value $\bar{\theta}_n$ is assumed to be available from the previous iteration, and $h_n(\theta)$ is computed using (5.14). In addition, another adaptive algorithm for updating the log-likelihood rate is run in parallel

$$\bar{l}_{n+1} = \bar{l}_n + \gamma_n (h_n(\bar{\theta}_n) - \bar{l}_n) \tag{5.16}$$

where $\bar{l}_n$ is the estimated log-likelihood rate and parameter $\bar{\theta}_n$ comes from (5.15). The diminishing stepsize γ_n satisfies the standard stochastic approximation conditions:

Assumption 5.4.3 *The stepsize values $\{\gamma_n\}_{n \geq 0}$ are non-negative and satisfy*

$$\sum_{n=0}^{\infty} \gamma_n = \infty, \quad \sum_{n=0}^{\infty} \gamma_n^2 < \infty.$$

The convergence of the simulation-based algorithm is established using the ODE method and ergodicity of the filtering process:

Proposition 5.4.1 *Suppose,*

(i) *The sample path $\{x_n, y_n\}_{n \geq 0}$ are generated from the HMM ξ, which satisfies Assumption 5.2.1 and Assumption 5.2.2.*

(ii) *The randomized and parameterized policy $\eta(\cdot; \theta)$ satisfies Assumption 5.4.1.*

(iii) *The averaging sequence $\{m_n\}_{n \geq 0}$ and the stepsize sequence $\{\gamma_n\}_{n \geq 0}$ satisfy Assumption 5.4.2 and Assumption 5.4.3, respectively.*

(iv) *The parameter vector sequence $\{\bar{\theta}_n\}_{n\geq 0}$ and the log-likelihood rate sequence $\{\bar{l}_n\}_{n\geq 0}$ are updated according to the recursive learning algorithm (5.15) and (5.16), respectively.*

Then, as $n \to \infty$, the sequence $\tilde{l}_n(\bar{\theta}_n)$ converges to a non-positive limit,

$$\nabla \tilde{l}_n(\bar{\theta}_n) \to \mathsf{0} \quad \text{and} \quad \bar{l}_n \to \tilde{l}_n(\bar{\theta}_n),$$

all in P_ξ-a.s. *sense.*

5.4.4 A simple bi-partition parameterization

Let $\mathcal{M} = \{1, 2\}$ denote the reduced aggregated state space with two superstates. For the bi-partition problem, a partition function ϕ takes only two values, either $\phi(i) = 1$ or $\phi(i) = 2$ for any state $i \in \mathcal{N}$. Let Θ be a sufficiently large compact subset of $\mathbb{R}^N$. We consider a real-valued parameter vector $\theta := (\theta(1), \ldots, \theta(N))^T \in \Theta$, where $\theta(i)$ decides the group assignment for the state $i \in \mathcal{N}$. In particular, we use $\zeta(\theta(i)) := \frac{1}{1+\exp(M\theta(i))}$ to reflect the probability that $\phi(i) = 1$, where $M > 0$ is some positive constant.

At time n, we only need to consider the randomized and parameterized partition policy for the state X_n. Suppose the current stat is $X_n = i \in \mathcal{N}$, and partition function at time $n-1$ is $\widetilde{\phi}$. The policy is defined for all $\phi \in \Phi$:

- If $\phi(j) = \widetilde{\phi}(j)$ for every $j \in \mathcal{N}/\{i\}$, then

$$\eta_\phi(i; \theta) = \zeta(\theta(i))\mathbb{1}_{\{\phi(i)=1\}} + (1 - \zeta(\theta(i))\mathbb{1}_{\{\phi(i)=2\}}.$$

- Otherwise, $\eta_\phi(i; \theta) = 0$.

One can easily verify that the policy satisfies the Assumption 5.4.1. At each time step, the policy only affects or changes the probability of the group assignment for the state X_n and keep others unchanged. Thus this policy can save a lot of computations at each time-step, which makes it more suitable for on-line estimation.

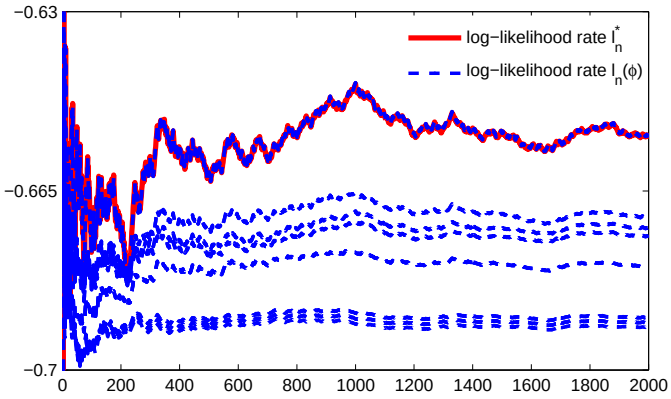


Figure 5.1: The original log-likelihood rate l_n^* is compared with the 8 different aggregated log-likelihood rates $l_n(\phi)$.

5.5 Numerical Results

In this section, we use a simple HMM ξ to illustrate the theoretical results and algorithms described in this paper. The HMM $\xi = (\mu, A, C)$ has 4 states and 2 observations. The transition matrices

$$A = \begin{bmatrix} 0.500 & 0.200 & 0.225 & 0.075 \\ 0.200 & 0.500 & 0.135 & 0.165 \\ 0.030 & 0.270 & 0.500 & 0.200 \\ 0.150 & 0.165 & 0.185 & 0.500 \end{bmatrix}, \quad C = \begin{bmatrix} 0.15 & 0.85 \\ 0.05 & 0.95 \\ 0.89 & 0.11 \\ 0.88 & 0.12 \end{bmatrix}$$

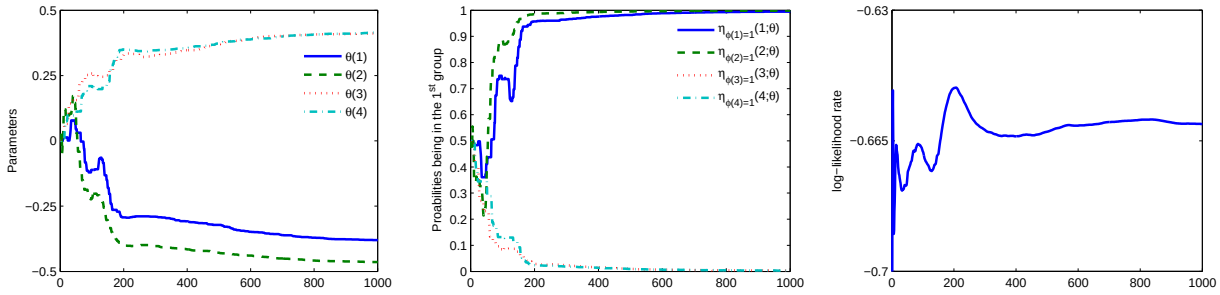
with the initial distribution $\mu = \pi$, the invariant distribution of A. We consider the bi-partition problem of the HMM ξ here, i.e., the state space $\mathcal{N} = \{1, 2, 3, 4\}$ is aggregated into the state space $\mathcal{M} = \{1, 2\}$.

5.5.1 Hypothesis testing approach for a simple HMM

Note that the partition function space Φ is of a moderate size ($|\Phi| = 2^4 = 16$). Thus the hypothesis testing method is employed in this subsection to find the optimal partition function as described in Section 5.3.3.

First, a sample path of $n = 2000$ observations $\{y_0, \ldots, y_n\}$ is generated according to the HMM ξ. The original log-likelihood rate $l_n^* = n^{-1} \log \mathsf{P}_\xi(y_0^n)$ is computed based on the recursive filer of the HMM ξ (see Section 5.2.2 for more details).

Second, for any fixed $\phi \in \Phi$, the aggregated HMM $\bar{\xi}(\phi)$ is obtained using the representation (5.6)–(5.8). Then we compute the aggregated log-likelihood rate $l_n(\phi) = n^{-1} \log \mathsf{P}_{\bar{\xi}(\phi)}(y_0^n)$ (5.11) for every aggregated HMM $\bar{\xi}(\phi)$ based on the recursive filter of $\bar{\xi}(\phi)$. Note that if the partition functions ϕ_1 and ϕ_2 are symmetric (e.g., $\phi_1 = [1, 2, 2, 2]$ and $\phi_2 = [2, 1, 1, 1]$ are symmetric), then the probability laws $\mathsf{P}_{\bar{\xi}(\phi_1)} = \mathsf{P}_{\bar{\xi}(\phi_2)}$ for almost all observations. Based on the



(a) (b) (c)

Figure 5.2: Plots of (a) the estimated parameter vector $\bar{\theta}_n$, (b) probabilities of the states being in the first group $\eta_{\phi=[1,1,1,1]}(\cdot;\bar{\theta}_n)$, and (c) the estimated log-likelihood rate $\bar{l}_n$ for the HMM ξ with the recursive learning algorithm (5.15) and (5.16).

symmetry of the problem, we only need to consider 8 partition functions for the hypothesis testing. In Figure 5.1, we depict the original log-likelihood rate l_n^* as well as 8 different aggregated log-likelihood rates $l_n(\phi)$ (two symmetric partition functions correspond to the same log-likelihood rate).

Finally, we choose the optimal partition function corresponding to the largest log-likelihood rate. For this example, the optimal partition functions is $\phi^* = [1,1,2,2]$ or $\phi^* = [2,2,1,1]$. The two corresponding aggregated HMMs are equivalent up to the permutation of the state space. We also note that for this special example the optimal aggregated log-likelihood rate is almost the same as the original one.

5.5.2 Recursive learning approach

From the hypothesis testing of all partition functions, we know that $\phi = [1,1,2,2]$ is the optimal bi-partition of the HMM ξ. In this subsection, we apply the recursive learning algorithm (5.15) and (5.16) to find the optimal partition function based on a single sample-path $\{x_n, y_n\}_{n\geq 0}$ of the HMM ξ.

The randomized and parameterized bi-partition policy, with the constant $M = 15$, is chosen for the recursive learning algorithm as described in Section 5.4.4. The averaging sequence is taken as $m_n = n^{0.8}$ and the stepsize sequence is taken as $\gamma_n = \frac{1}{n+1}$ for $n \geq 0$. The parameter space Θ is a sufficiently large compact subset of $\mathbb{R}^N$ and the algorithm is initialized with the parameter vector $\bar{\theta}_0 = [0,0,0,0]$.

In Figure 5.2, we depict a typical run of the recursive learning algorithm for the 1000 iterations. After $n = 1000$ iterations, the estimated parameter vector $\bar{\theta}_n = [-0.3802, -0.4643, 0.4117, 0.4154]$, and the probabilities of states being in the first group are $\eta_{\phi=[1,1,1,1]}(\cdot;\bar{\theta}_n) = [0.9948, 0.9984, 0.0034, 0.0032]$. From this, the optimal partition function $\phi = [1,1,2,2]$ can be determined with high probability. The corresponding estimated log-likelihood rate is equal to $\bar{l}_n = -0.6605$, which is close to maximum log-likelihood rate depicted in Figure 5.1. The recursive learning algorithm thus recovers the optimal partition function for this example.

5.6 Conclusions

In this chapter, we proposed a recursive learning algorithm for model reduction of Hidden Markov Models. The state space is aggregated to reduce the complexity of the HMM. The optimal aggregation is obtained by minimizing the K-L divergence rate between the laws of the observation process. The optimal aggregated HMM is represented as a function of the partition function of the state space. The optimal partition can be obtained by the hypothesis testing method, or in large-scale problem by using the recursive learning algorithm based on the stochastic approximation. The algorithm can be implemented only through a single sample path of the HMM.

Part II

Modeling and Control of Building Systems

Chapter 6

Structure-Preserving Model Reduction of Nonlinear Building Thermal Models

6.1 Introduction

Accurate models of temperature evolution in a building are required for real-time prediction and control, especially in model-based control methods [46, 71, 73, 90, 128]. This chapter focuses on model order reduction of buildings with multiple zones. A zone in this chapter refers to a single space (room, hallway, etc.) that is serviced by a single "terminal box" with supply air diffuser and return air grilles. Figure 6.1 shows a four-zone building HVAC system, where each zone refers to a single room here; terminologies and more details appear in Section 6.2.

A complete model of the entire multi-zone building's thermal response can be constructed by using (i) resistor-capacitor networks for combined conduction-convection-radiation through surfaces, and (ii) heat balance equations to account for the enthalpy exchange between a zone and the outside due to the ventilation air. The resulting lumped parameter model is called the *full-order building thermal model* in this chapter. This approach of constructing multi-zone thermal models have been pursued previously in [45, 61, 85, 119, 125]. The full-order model we consider here is from [45]. The ventilation heat exchange terms make the thermal dynamic model nonlinear; more modeling details appear in Section 6.2.

A fundamental problem with the full-order models is that they quickly explode in complexity as the number of zones increases. For example, the full-order model of the downstream part of a four-zone building, shown in Figure 6.1, has 37 nodes (more details appear in Section 6.5). For a large commercial building with hundreds of zones, the number of nodes are of the order of several hundreds, or even thousands.

In this chapter, we propose an *aggregation-based approach* that preserves the structure of the original model, that is, the reduced building thermal model is still a nonlinear RC-network. This is achieved by obtaining super-nodes via aggregation, and determining the super-capacitance for each super-node and super-resistance for each edge between two adjacent super-nodes. The aggregation-based approach proposed in this chapter is based on model reduction method of Markov chains that was described in the first part of this thesis. The Kullback-Leibler (K-L) divergence rate is used as a "metric" to reduce Markov chains via aggregation of states in (see Chapter 2). The idea of this chapter is to connect the linear portion of the multi-zone thermal model to a continuous-time Markov chain, and extend the

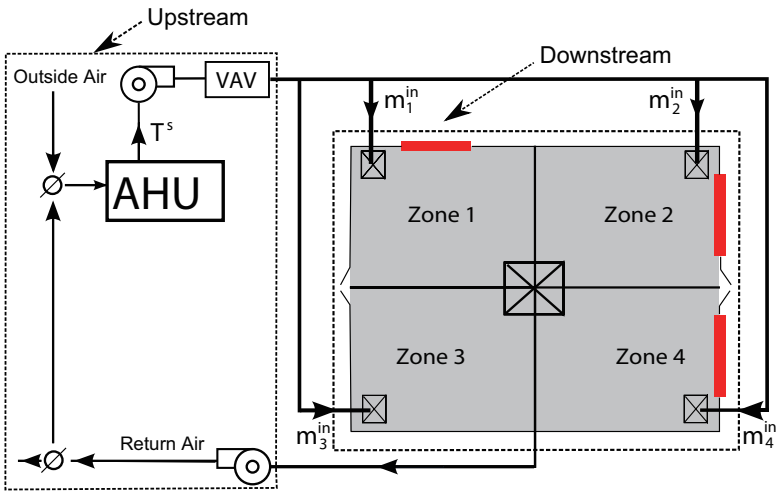


Figure 6.1: The configuration of a four-zone building HVAC system.

model reduction procedure for Markov chains to the nonlinear full-order building thermal model. The degree of reduction can be specified by the user, and the full-order model with n nodes can be reduced to a model with m super-nodes, with m being any integer between 1 and n. Simulations reported in this chapter show that the proposed method produces reduced-order models that well approximate the time-domain predictions of the original model. As one would expect, the prediction accuracy decreases as the specified degree of reduction increases.

The rest of the chapter is organized as follows. In Section 6.2, the full-order model is described and the model reduction problem is stated. In Section 6.3, the Markov chain analogy of the building thermal dynamics is presented. In Section 6.4, the aggregation-based methodology is applied to reduce the building thermal model. In Section 6.5, theoretical results are illustrated by numerical simulations.

6.2 Full-order Building Thermal Model

A typical HVAC system consists of AHUs, supply ducts, and terminal boxes; see Figure 6.1 for an example. The AHU (Air Handling Unit) supplies conditioned air (usually cold and dry) to terminal boxes at so-called leaving-air temperature and humidity. Each terminal box delivers air to one or more zones. When the box is equipped with a reheat coil (a common configuration), the supply air temperature downstream of the box can be increased beyond the AHU leaving temperature. In a VAV (Variable-Air-Volume) system, the terminal box may vary the supply air mass flow rate through dampers, but not in a CAV (constant air volume) system. A controller at each terminal box can be used to maintain the temperature of a zone at a specified value by controlling the mass flow rate of air supplied to the zone.

The dynamics of the building with its HVAC system can be divided into upstream and downstream parts (see Fig-

ure 6.1). The upstream part includes the AHU dynamics and the downstream part includes the thermal dynamics of the zones. The focus here is on modeling the downstream thermal dynamics. The reasons for ignoring the AHU dynamics are twofold. First, the dimension of the downstream model increases quickly with the number of zones and internal thermal nodes, while the dimension of the upstream model increases only with the number of AHUs. The later is typically small even for a large building. Second, the AHU has fast dynamics in the HVAC system, with a time constant of about a minute [12], whereas the thermal dynamics of the zones are relatively slow with time constants in tens of minutes [112] to hours [43]. As a result, the dynamics of the AHUs are replaced by static gains in this chapter without significant loss of accuracy. From now on, "building thermal dynamics" would mean dynamics of the downstream part. Variations of temperature within a zone are neglected; each zone is characterized by a single temperature variable.

6.2.1 RC-network model of building thermal dynamics

A building thermal model is constructed by combining lumped parameter models of thermal interaction between two zones separated by a solid surface (e.g., walls, windows, ceilings, and floors). For the sake of simplicity, here we ignore the inter-zone convective heat transfer that occurs through the open doors and hallways. A lumped parameter model of combined heat flow across a surface is modeled as a simple RC-network, with current and voltage being analogous of heat flow and temperature. In this modeling framework, the capacitances are used to model the total thermal capacity of the wall, and the resistances are used to represent the total resistance that the wall offers to the flow of heat from one side to the other. In [44], Gouda *et. al.* showed that a second-order RC-network model with 3 resistors and 2 capacitors, which we will call 3R2C model, is sufficient to capture the conductive dynamic interaction between two spaces through a single wall; see Example 6.2.1 at the end of this subsection.

For a building consisting of a number of surface elements (e.g., walls, windows, ceilings, and floors), 3R2C models for surface elements can be inter-connected to obtain a RC-network model of the entire building. The resulting model can be represented as an *undirected graph* $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}{:=}\{1, \ldots, n{+}1\}$ denotes the set of *nodes* of the graph. A node may represent a physical zone (e.g., a room, a hallway, or "the outside"), or some point inside a wall. For the sake of simplicity of the description, the nodes are assumed to be re-indexed so that the first N nodes correspond to the temperatures of zones $1, \ldots, N$; these are called the *zone nodes*. The next $(n - N)$ nodes correspond to the temperatures internal to the surfaces that appear due to the 3R2C networks; these are called the *internal nodes*. The last, i.e. $(n + 1)$th node, corresponds to the outside. Each node $i \in \mathcal{V}$ corresponds to a temperature T_i and each node $i \in \mathcal{V}/\{n + 1\}$ has an associated capacitance C_i. The set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ denotes the set of all *edges*. Edges represent pathways for conductive heat transports. More specifically, each edge $(i, j) \in \mathcal{E}$ represents the conductive thermal interaction between the nodes i and j and there exists a resistance $R_{ij} \in \mathbb{R}_+$ connecting them directly. Since the graph

is undirected, $R_{ji} = R_{ij}$ by convention.

The states and inputs of the building thermal model are summarized below:

$$\text{States} : T_1, \ldots, T_N, T_{N+1}, \ldots, T_n, T_{n+1}$$
$$\text{Inputs} : T_o, T^s; \dot{m}_i^{in}, \dot{Q}_i^r, \dot{Q}_i^{int}, \dot{Q}_i^{ext}, i = 1, \ldots, N$$

where $T_1, \ldots, T_N$ denote the space temperature of the zones, and $T_{N+1}, \ldots, T_n$ denote temperature of the points internal to the surface elements, and T_{n+1} & T_o denote the same quantity – the outside temperature (it is denoted as T_{n+1} if considered as a state, and T_o if considered as an input), T^s denotes the *supply temperature*, that is the temperature of the air supplied by the AHU, $\dot{m}_i^{in}$ denotes the mass flow rate of the supply air entering the ith zone, $\dot{Q}_i^r$ denotes the heat gain due to reheating that may occur at the VAV box of the ith zone, $\dot{Q}_i^{int}$ denotes the *internal heat gain*, i.e., the rate of heat generated by occupants, equipments, lights, etc. in the ith zone, and $\dot{Q}_i^{ext}$ denotes the *external heat gain*, i.e., the rate of solar radiation entering the ith zone.

The supplied air temperature T^s is usually constant for a VAV system, at least over short intervals of time [4]. All other inputs are time varying. In this chapter, it is assumed that (i) the supply air temperature T^s is given as a constant, (ii) the (estimation of) the outside temperature T_0 and the heat gains $\dot{Q}^r, \dot{Q}^{int}, \dot{Q}^{ext}$ are available based on historical data, weather predictions, and various sensors.

The thermal dynamics of a multi-zone building, described by a graph G, are modeled by the following coupled nonlinear differential equations: For $i = 1, \ldots, n$,

$$C_i \frac{dT_i}{dt}(t) = \sum_{j \in \mathcal{N}_i} (T_j(t) - T_i(t))/R_{ij} + \dot{Q}_i(t) + \Delta H_i(t) \tag{6.1}$$

where $\mathcal{N}_i := \{j \in \mathcal{V} : j \neq i, (i,j) \in \mathcal{E}\}$ denotes the set of *neighbors* connecting to the node i (note that the outside node $n+1$ may belong to the set $\mathcal{N}_i$ for some node i), and the terms $\dot{Q}_i, \Delta H_i$ are described below:

- The *heat gain* term $\dot{Q}_i$ is the rate of thermal energy entering the node i from all sources other than ventilation air and conduction from neighboring nodes. It is non-zero only for zone nodes:

$$\begin{cases} \dot{Q}_i(t) = \dot{Q}_i^r(t) + \dot{Q}_i^{int}(t) + \dot{Q}_i^{ext}(t), & i = 1, \ldots, N \\ \dot{Q}_i(t) = 0, & i = N+1, \ldots, n. \end{cases}$$

- The *ventilation heat exchange* term ΔH_i is the rate of thermal energy entering the node i due to ventilation. It

is non-zero only for the zone nodes:

$$\begin{cases} \Delta H_i(t) = C_{pa}\dot{m}_i^{in}(t)(T^s - T_i(t)), & i = 1, \ldots, N \\ \Delta H_i(t) = 0, & i = N+1, \ldots, n \end{cases}$$

where C_{pa} is the specific heat capacitance of the supplied air at constant pressure. Recall that T^s is the supply air temperature.

The coupled ordinary differential equation model (6.1) so obtained is nonlinear because of the presence of the bilinear term $\dot{m}_i^{in}T_i$ in defining the ventilation heat exchange ΔH_i. Note that the term $\dot{Q}_i + \Delta H_i$ can be together interpreted as a current source injected into (or extracted from) the node i of the RC-network, except that the source strength depends on the "voltage" T_i of node i as well: The full-order model can be thought as a RC-network model with additional current sources, where the source strengths are dependent on the voltage of the nodes they are connected to.

In the following sections, a compact state-space representation of the building thermal model (6.1) is used. The outside temperature is taken as a "virtual state" T_{n+1} to the system. We assign a very large "virtual capacitance" to the outside node: $C_{n+1} \gg C_i$, for $i = 1, \ldots, n$. Letting $C_{n+1} \to \infty$, the system of equations (6.1) is expressed as a state-space representation:

$$\frac{dT}{dt} = AT + L(T, U, \dot{Q}) \tag{6.2}$$

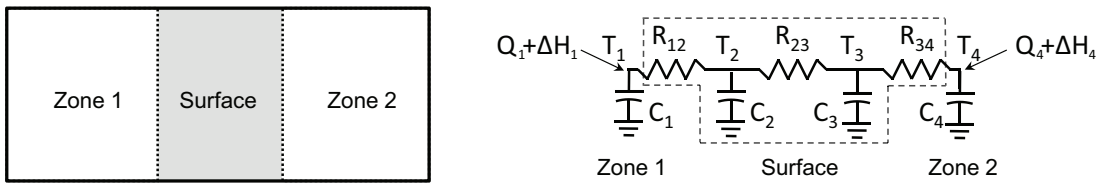
where the state vector $T := [T_1, \ldots, T_{n+1}]^T$, the control vector $U := [\dot{m}_1^{in}, \ldots, \dot{m}_N^{in}, 0, \ldots, 0]^T$, and the heat gain vector $\dot{Q} := [\dot{Q}_1, \ldots, \dot{Q}_N, 0, \ldots, 0]^T$. The *transition rate matrix* A is an $(n+1) \times (n+1)$ matrix with entries given by

$$A_{ij} := \begin{cases} 0, & \text{if } j \neq i,\ (i,j) \notin \mathcal{E} \\ 1/(C_i R_{ij}), & \text{if } j \neq i,\ (i,j) \in \mathcal{E} \\ -\sum_{k \neq i} A_{ik}, & \text{if } j = i,\ (i,j) \in \mathcal{E} \end{cases} \tag{6.3}$$

and the nonlinear function:

$$\begin{cases} L_i(T, U, \dot{Q}) = \dfrac{C_{pa}U_i(T^s - T_i) + \dot{Q}_i}{C_i}, & i = 1, \ldots, N \\ L_i(T, U, \dot{Q}) = 0, & i = N+1, \ldots, n \\ L_i(T, U, \dot{Q}) = \eta, & i = n+1 \end{cases}$$

where $\eta(t) \in \mathbb{R}$ is chosen such that $\eta(t) = \dot{T}_o(t)$. To see the equivalence between (6.1) and (6.2), note that the



(a) (b)

Figure 6.2: (a) Two zones separated by a single surface, and (b) its lumped RC-network model.

entries in the last row of A approach 0 as $C_{n+1} \to \infty$ (since they are of the form $1/(C_{n+1}R_{n+1,j})$). In the limit, $\dot{T}_{n+1} = \eta(t)$, which gives $T_{n+1}(t) = T_o(t)$ for all $t \geq 0$.

Example 6.2.1 (A simple two-zone building) Consider the simplest example where two zones are separated by a single wall/surface as shown in Fig. 6.2 (a). Here it is assumed that two zones have no thermal interaction with other zones or the outside. A 3R2C network model is used to model the surface as shown in Fig. 6.2 (b). Ventilation air enters each zone at temperature T^s, and leaves the zone at the same temperature as that of the zone. There are 4 building nodes, two zone nodes plus two internal nodes. T_1 and T_4 denote the space temperatures of zone 1 and zone 2, respectively. T_2 and T_3 denote temperatures of the points internal to the surface that arise due to the 3R2C model of the surface. The parameters C_1 and C_4 are the thermal capacitances of the two zones, while C_2, C_3 are thermal capacitances for the 3R2C model of the surface. The dynamics for the RC-network model are described by the following differential equations obtained by using the heat balance:

$$\begin{aligned}
C_1\frac{dT_1}{dt} &= -\frac{1}{R_{12}}T_1 + \frac{1}{R_{12}}T_2 + \dot{Q}_1 + \Delta H_1 \\
C_2\frac{dT_2}{dt} &= -(\frac{1}{R_{21}} + \frac{1}{R_{23}})T_2 + \frac{1}{R_{21}}T_1 + \frac{1}{R_{23}}T_3 \\
C_3\frac{dT_3}{dt} &= -(\frac{1}{R_{32}} + \frac{1}{R_{34}})T_3 + \frac{1}{R_{32}}T_2 + \frac{1}{R_{34}}T_4 \\
C_4\frac{dT_4}{dt} &= -\frac{1}{R_{43}}T_4 + \frac{1}{R_{43}}T_3 + \dot{Q}_4 + \Delta H_4
\end{aligned} \tag{6.4}$$

where R_{12}, R_{23}, R_{34} are thermal resistances, the $\dot{Q}_i = \dot{Q}_i^r + \dot{Q}_i^{int} + \dot{Q}_i^{ext}$ is the heat gain, and $\Delta H_i = C_{pa}\dot{m}_i^{in}(T^s - T_i)$ is the heat exchange due to ventilation for $i = 1, 4$, with C_{pa} being the specific heat capacitance of the supply air. ■

6.2.2 Problem statement

For a building with N zones, the number of states in the full-order model (6.2) described above is of the order of $7N$, usually more. A medium size commercial building has close to 100 zones and a larger building can have several hundreds. The dimension of the full-order model thus can be quite large. The goal of this chapter is to obtain a

reduced-order model of smaller dimension such that model reduction produces the RC-network physical structure.

To achieve this goal, an aggregation methodology is considered: Mathematically, suppose the goal is to reduce the state dimension from n to m, where $m \leq n$ is the (user-specified) number of super-nodes. The first step is to choose a *partition function* $\phi : \mathcal{V} \to \bar{\mathcal{V}}$, where $\bar{\mathcal{V}} = \{1, \ldots, m+1\}$ denotes the set of "super-nodes" for the reduced-order model, and recall that $\mathcal{V} = \{1, \ldots, n+1\}$ denotes the set of nodes for the full-order model. A partition function is an onto function but possibly many-to-one. The elements of $\bar{\mathcal{V}}$ are the super-nodes, and for every $k \in \bar{\mathcal{V}}$, the inverse mapping $\phi^{-1}(k) \subset \mathcal{V}$ denotes the group of nodes in the full-order model that are aggregated into the kth super-node using the partition function ϕ. The second step is to define a graph for $\mathcal{V}$ by defining "super-edges" between super-nodes. The third step is to define appropriate super-capacitances and super-resistances.

To perform such a structure-preserving model reduction, we need to answer the following questions:

- *Q1: How to choose the partition function?*
- *Q2: Given a partition function, how to find the super-capacitances and super-resistances, and how to aggregate the nonlinear terms of the full-order model?*
- *Q3: How to compare the full and reduced-order models?*

The rest of the chapter is about answering these questions. A brief outline of the approach is provided below.

- *Partition by Markov chain aggregation:* It is shown in Section 6.3 that the linear thermal dynamics is analogous to a continuous-time Markov chain. A recently developed aggregation method for Markov chains is then employed to obtain a (sub)-optimal partition function. In this method, a measure of optimality of the aggregation is defined in terms of the Kullback-Leibler divergence rate. Solving the optimal partition problem is shown to be hard, and a recursive bi-partition algorithm is proposed here to obtain sub-optimal partitions. Details of the algorithm are summarized in Appendix C.
- *Finding super-quantities and the reduced-order model:* The super-capacitances and super-resistances are obtained directly based on the Markov chain analogy. Details appears in Section 6.4.1. Due to the current source interpretation of the nonlinear part, the current sources connecting to the same group of the aggregated nodes are directly added up to form a super-current source for the corresponding super-node. It is shown that the reduced-order model is a RC-network model defined with super-quantities. Details appear in Section 6.4.2.
- *Comparison between full and reduced-order models:* We obtain the reduced-order model by aggregating the nodes into a smaller number of super-nodes. Then we employ the *lifting technique* to lift the reduced-order model to the one with the same number of nodes as the full-order model. Details of the lifting technique is

summarized in Section 6.4.1. We can now directly compare the temperatures of zone nodes for the full and reduced-order models.

6.3 Markov Chain Analogy and Aggregation

In this section, it is shown that the linear part of the building thermal model (6.2) is analogous to a *continuous-time Markov chain*. The *linear dynamics* of the building thermal model (6.2) are given by:

$$\frac{dT}{dt} = AT. \tag{6.5}$$

Due to the special structure of the matrix A (see (6.3)), the linear thermal model (6.5) is *conservative*. Specifically, a scalar-valued function $V(t) := \sum_{i\in\mathcal{V}} C_i T_i(t)$ is conserved for all time:

$$\begin{aligned}
\frac{dV}{dt}(t) &= \sum_{i\in\mathcal{V}} C_i \frac{dT_i}{dt}(t) \\
&= \sum_{j\in\mathcal{V}} T_j(t) \sum_{i\in\mathcal{V}} C_i A_{ij} && \text{(6.6a)} \\
&= \sum_{j\in\mathcal{V}} C_j T_j(t) \sum_{i\in\mathcal{V}} A_{ji} && \text{(6.6b)} \\
&= 0 && \text{(6.6c)}
\end{aligned}$$

where the equality (6.6a) follows from (6.5), the equality (6.6b) is due to the fact that $C_i A_{ij} = C_j A_{ji}$ for all $i, j \in \mathcal{V}$ (see (6.3)), and the equality (6.6c) uses the fact that each row sum of the matrix A is zero. We denote $V_0 := V(0) = \sum_{i\in\mathcal{V}} C_i T_i(0)$ as the *invariant quantity* of the linear thermal model (6.5).

6.3.1 Analogy to a Markov chain

Based upon the conservative property of the linear thermal model (6.5), define the *thermal distribution* as a row vector, denoted by f, where

$$f_i = \frac{C_i}{V_0} T_i, \quad i \in \mathcal{V}.$$

Note that $\sum_{i\in\mathcal{V}} f_i(t) \equiv 1$ for all $t \geq 0$.

On differentiating f_i with respect to t, and using (6.5),

$$\frac{df_i}{dt} = \frac{C_i}{V_0}\frac{dT_i}{dt} = \sum_{j\in\mathcal{V}} A_{ij} \frac{C_i}{V_0} T_j. \tag{6.7}$$

By substituting (6.3) in (6.7), we have

$$\begin{aligned}\frac{df_i}{dt} &= A_{ii}\frac{C_i}{V_0}T_i + \sum_{j\neq i}\frac{1}{C_iR_{ij}}\frac{C_i}{V_0}T_j \\ &= A_{ii}\frac{C_i}{V_0}T_i + \sum_{j\neq i}\frac{1}{C_jR_{ji}}\frac{C_j}{V_0}T_j \qquad (6.8a)\\ &= \sum_{j\in\mathcal{V}} f_jA_{ji} \qquad (6.8b)\end{aligned}$$

where the fact that $R_{ij} = R_{ji}$ is used in deriving the equality (6.8a). Using matrix notation for representing (6.8b), we obtain the dynamics of the thermal distribution and its solution

$$\frac{df}{dt} = fA \qquad \Rightarrow \qquad f(t) = f(0)e^{At}, \quad \forall t \geq 0.$$

The Markov chain analogy is now clear. Note that each row sum of A is zero, its diagonal entries are negative, and its non-diagonal entries are non-negative (see (6.3)). Thus, the transition rate matrix A is the *infinitesimal generator* of a *transition semigroup* $\{e^{At}\}_{t\geq 0}$: For any $t, s \geq 0$, (i) $e^{A0} = I$, (ii) e^{At} is a stochastic matrix (That is e^{At} is a nonnegative matrix whose row sums are equal to one), and (iii) $e^{A(t+s)} = e^{At}e^{As}$.

Consider now a continuous-time Markov chain $\{X(t)\}_{t\geq 0}$ on the state space $\mathcal{V}$ with the transition semigroup $\{e^{At}\}_{t\geq 0}$ [110]. Let $g(t)$ denote the *probability distribution* at time t, i.e.,

$$g_i(t) = \Pr(X(t) = i), \quad i \in \mathcal{V}.$$

Using the transition semigroup property, we have

$$g_i(t) = \sum_{j\in\mathcal{V}} \Pr(X(0) = j)\Pr(X(t) = i \mid X(0 = j)) = \sum_{j\in\mathcal{V}} g_j(0)(e^{At})_{ji}.$$

If $f(0)$ is the initial distribution of $\{X(t)\}_{t\geq 0}$, i.e., $g(0) = f(0)$, then

$$g(t) = g(0)e^{At} = f(0)e^{At} = f(t). \qquad (6.9)$$

Thus, starting from the same initial distribution, the probability distribution of the continuous-time Markov chain $\{X(t)\}_{t\geq 0}$ is equal to the thermal distribution of the linear thermal model (6.5). For more details on continuous-time Markov chains, we refer the reader to [93, 110] and references therein.

For any ergodic Markov chain, there exists a unique *stationary distribution* π (obtained as a solution to $\pi A = 0$),

whereby starting from any initial distribution

$$\lim_{t\to\infty} g(t) = \pi.$$

For linear thermal model (6.5), the associated Markov chain is shown to be ergodic in [27], and the stationary distribution is given by:

$$\pi_i = \frac{C_i}{\sum_{j\in\mathcal{V}} C_j}, \quad i \in \mathcal{V}. \tag{6.10}$$

6.3.2 Discretization of the continuous-time Markov chain

In practice, it is more convenient to work with discrete-time Markov chains (DTMC) instead of continuous-time Markov chains (CTMC). The DTMC $\{X(k\Delta t)\}_{k\geq 0}$ is obtained by discretizing the CTMC $\{X(t)\}_{t\geq 0}$ with a step-size Δt. Let $\xi(k)$ denote the probability distribution of the DTMC at kth time-step, i.e., $\xi_i(k) = \Pr(X(k\Delta t) = i)$ for $i \in \mathcal{V}$. Using (6.9),

$$\xi(k) = \xi(0)P^k(\Delta t), \quad k \geq 0$$

where the *transition matrix* is defined as

$$P(\Delta t) := e^{A\Delta t}. \tag{6.11}$$

For any $t \geq 0$, there exists an integer $k \geq 0$ such that $k\Delta t \leq t < (k+1)\Delta t$ and $e^{At} \approx e^{Ak\Delta t}$ for small enough Δt. Thus the CTMC $\{X(t)\}_{t\geq 0}$ with the transition semigroup $\{e^{At}\}_{t\geq 0}$ is approximated by the DTMC $\{X(k\Delta t)\}_{k\geq 0}$ with the transition matrix $P(\Delta t)$ as $\Delta t \to 0$. One can verify that π given in (6.10) is also the stationary distribution of the DTMC, i.e., $\lim_{k\to\infty} \xi(k) = \pi$.

6.3.3 Kullback-Leibler metric

The *Kullback-Leibler divergence rate* is employed as a metric to compare Markov chains (see Chapter 2 for more details). Here we only summarize the major formulae that are useful in this chapter: Consider two Markov chains (π, P) and $(\bar{\pi}, \bar{P})$ defined on different state spaces $\mathcal{V} = \{1, \ldots, n+1\}$ and $\bar{\mathcal{V}} = \{1, \ldots, m+1\}$, respectively. Without loss of generality, we let $m \leq n$. The relationship between $\mathcal{V}$ and $\bar{\mathcal{V}}$ is described by a partition function $\phi : \mathcal{V} \mapsto \bar{\mathcal{V}}$. To compare two Markov chains in the same state space, we lift the Markov transition matrix $\bar{P}$ to another one defined on the state space $\mathcal{V}$ based on partition function ϕ and the invariant distribution π:

$$\widehat{P}^{(\pi)}_{ij}(\phi) = \frac{\pi_j}{\sum_{k\in\psi(j)} \pi_k} \bar{P}_{\phi(i)\phi(j)}, \quad i, j \in \mathcal{V}.$$

The K-L divergence rate between two Markov chains is given by:

$$R_\phi(P\|\bar{P}) := R(P\|\widehat{P}^{(\pi)}(\phi)).$$

6.3.4 Aggregation of Markov chain

Here we summarize the basic results for optimally aggregating Markov chains from Chapter 2 and Chapter 3. Let (π, P) denote a discrete-time Markov chain defined on the state space $\mathcal{V} = \{1, \ldots, n+1\}$ with the transition matrix P and the stationary distribution π. The model reduction problem is to find an optimal aggregated Markov chain, denoted by $(\bar{\pi}, \bar{P})$, defined on the state space $\bar{\mathcal{V}} = \{1, \ldots, m+1\}$, where $m \leq n$, such that the KL divergence rate $R_\phi(P\|\bar{P})$ between two Markov chains is minimized.

The *m-partition problem* is to find a partition function $\phi : \mathcal{V} \mapsto \bar{\mathcal{V}}$ and an aggregated transition matrix $\bar{P}$ that solves the following optimization problem:

$$\begin{array}{ll} \min\limits_{\phi,\bar{P}} & R_\phi(P\|\bar{P}) \\ \text{s.t.} & \bar{P}\mathbf{1} = \mathbf{1},\ \bar{P} \geq 0. \end{array}$$

As shown in Theorem 2.4.1, for a fixed (whether optimal or not) partition function ϕ, the optimal aggregated Markov chain $(\bar{\pi}(\phi), \bar{P}(\phi))$ is given by:

$$\bar{P}_{kl}(\phi) = \frac{\sum_{i\in\phi^{-1}(k)} \sum_{j\in\phi^{-1}(l)} \pi_i P_{ij}}{\sum_{i\in\phi^{-1}(k)} \pi_i}, \quad k, l \in \bar{\mathcal{V}} \tag{6.12}$$

where the stationary distribution of $\bar{P}(\phi)$ is given by

$$\bar{\pi}_k(\phi) = \sum_{i\in\phi^{-1}(k)} \pi_i, \quad k \in \bar{\mathcal{V}}. \tag{6.13}$$

As a result, the m-partition problem reduces to finding only an *optimal partition function* $\phi^* : \mathcal{V} \to \bar{\mathcal{V}}$ such that

$$\phi^* \in \arg\min_{\phi}\ R_\phi(P\|\bar{P}(\phi)). \tag{6.14}$$

It is shown in Chapter 2 that solving the optimization problem (6.14) exactly is difficult for $m > 2$, but a sub-optimal solution for $m = 2$ can be easily computed. This leads to a sub-optimal solution for arbitrary $m \geq 2$ through the *recursive bi-partition algorithm* AlgoBIPA, which is described in Appendix C. Here we summarize the basic ideas: First, AlgoBIPA is used to obtain $\mathcal{V}_1, \mathcal{V}_2$ so that $\mathcal{V}_1 \cup \mathcal{V}_2 = \mathcal{V}$ and $\mathcal{V}_1 \cap \mathcal{V}_2 = \Phi$, where Φ denotes the empty set.

Then, AlgoBIPA is used on the set $\mathcal{V}_1$ to obtain $\mathcal{V}_{11}$ and $\mathcal{V}_{12}$ such that $\mathcal{V}_{11} \cup \mathcal{V}_{12} = \mathcal{V}_1$ and $\mathcal{V}_{11} \cap \mathcal{V}_{12} = \Phi$. If the model order is less than m, AlgoBIPA is used on the set $\mathcal{V}_2$ to obtain partitions $\mathcal{V}_{21}$ and $\mathcal{V}_{22}$ such that $\mathcal{V}_{21} \cup \mathcal{V}_{22} = \mathcal{V}_2$ and $\mathcal{V}_{21} \cap \mathcal{V}_{22} = \Phi$. This procedure is repeated until m partitions are obtained. At each step the number of partition increases by one.

6.3.5 Analogy to thermal dynamics

Based on the Markov chain analogy for the linear thermal dynamics (see Section 6.3.1), the model reduction framework for Markov chains is extended to building thermal models:

- *Metric for comparing thermal distributions*: The K-L divergence is employed as a metric to compare two thermal distributions f and g defined on the same building node set $\mathcal{V}$:

$$D(f\|g) = \sum_{i\in\mathcal{V}} f_i \log(f_i/g_i).$$

  For the model reduction problem, it is of interest to compare two thermal distributions defined on building graphs of different cardinalities. Let f and $\bar{f}$ denote two thermal distributions defined on $\mathcal{V}$ and $\bar{\mathcal{V}}$, respectively. The low-dimensional distribution $\bar{f}$ is lifted to a high-dimensional distribution $\widehat{f}$ defined on $\mathcal{V}$ by using partition function ϕ and lifting distribution μ:

$$\widehat{f}_i^{(\mu)}(\phi) = \frac{\mu_i}{\sum_{k\in\psi(i)}\mu_k}\bar{f}_{\phi(i)}, \quad i \in \mathcal{V}. \tag{6.15}$$

  The lifting may be viewed as a linear transformation that conserves the total heat. The K-L metric is then used to compare the two thermal distributions f and $\widehat{f}$ on the same node set $\mathcal{V}$.

- *Metric for comparing thermal models*: The K-L divergence rate is used as a measure to compare two building thermal models. In particular, suppose full-order model is simulated starting from an initial distribution $f(0)$. Denote the resulting trajectory of the thermal distribution as $\{f(k\Delta t)\}_{0\le k\le N}$. Now, suppose the reduced-order model is also simulated starting from the initial distribution

$$\bar{f}_l(0) = \sum_{i\in\phi^{-1}(l)} f_i(0), \quad l \in \mathcal{V}.$$

  Denote the resulting trajectory of thermal distribution as $\{\bar{f}(k\Delta t)\}_{0\le k\le N}$, which evolves over reduced graph $\bar{\mathcal{V}}$. The trajectory $\{\bar{f}(k\Delta t)\}_{0\le k\le N}$ is lifted to the full building graph by using (6.15), and denoted by $\{\widehat{f}(k\Delta t)\}_{0\le k\le N}$.

The K-L divergence rate between full and reduced-order models is given by,

$$\frac{1}{N}\sum_{k=1}^{N} D(f(k\Delta t)\|\widehat{f}(k\Delta t)). \tag{6.16}$$

Thus, the K-L divergence rate is a measure of average distance between trajectories generated from simulating two thermal models.

- *Bi-partition*: An optimal bi-partition of a given model produces a 2-state reduced-order model that is closest to the full-order model in the sense of distance (6.16). Since the distance is a time average, discrepancies between two models at the slow(est) time-scales contribute more to the error compared to the fast transients. The choice of metric thus leads to a 2-state model that approximates the full-order model on the slowest time-scale.

- *Recursive bi-partition*: The recursive application of bi-partition algorithm produces a reduced-order model that progressively captures multiple time-scales in the problem. The first bi-partition results in splitting of the graph into two clusters, and a 2-state model that captures the slowest time-scale. The next bi-partition further splits one of the two clusters so as to capture the slowest time-scale in that cluster and so on. In effect after m-applications of the algorithm, the reduced-order model describes the m slowest time-scales of the full-order model.

6.4 Aggregated Building Thermal Model

In this section, the aggregation methodology is applied to obtain a reduced-order model for building thermal model (6.2). We first describe the reduced-order model for the linear part of the building thermal model (6.2), and then the reduced-order model for the nonlinear part of (6.2).

6.4.1 Aggregated linear thermal dynamics

For the linear thermal model (6.5), the goal is to aggregate the node set $\mathcal{V} = \{1, \ldots, n+1\}$ into a smaller super-node set $\bar{\mathcal{V}} = \{1, \ldots, m+1\}$ where $m \leq n$. For each super-node $k \in \bar{\mathcal{V}}$, we introduce the super-temperature $\bar{T}_k$, super-capacitance $\bar{C}_k$, and super-resistance $\bar{R}_{kl}$. For a given partition function ϕ, the reduced-order model for (6.5) has the form:

$$\frac{d\bar{T}}{dt} = \bar{A}(\phi)\bar{T}, \tag{6.17}$$

where $\bar{T} = [\bar{T}_1, \ldots, \bar{T}_{m+1}]^T$ denotes the super-temperature vector, and $\bar{A}(\phi)$ denotes the $(m+1) \times (m+1)$ super-transition-rate matrix. The Markov chain analogy also works for the reduced-order model with the associated transition semigroup $\{e^{\bar{A}(\phi)t}\}_{t\geq 0}$. Discretizing with a small step-size Δt, one obtains the transition matrix for the aggregated

Markov chain defined on $\bar{\mathcal{V}}$:

$$\bar{P}(\Delta t) := e^{\bar{A}(\phi)\Delta t}.$$

Recall that the transition matrix for the discrete-time Markov chain associated with the full-order linear thermal dynamics is denoted by $P(\Delta t)$ (see (6.11)). The goal is to choose $\bar{A}(\phi)$ so that the aggregated Markov chain with the transition matrix $\bar{P}(\Delta t)$ optimally approximates the original Markov chain with the transition matrix $P(\Delta t)$. The aggregation method described in Section 6.3.4 is employed to determine the formula for the optimal aggregated transition matrix $\bar{A}(\phi)$. According to (6.12), the formula for the optimal aggregated Markov transition matrix is given by:

$$\bar{P}_{kl}(\Delta t) = \frac{\sum_{i\in\phi^{-1}(k)}\sum_{j\in\phi^{-1}(l)}\pi_i P_{ij}(\Delta t)}{\sum_{i\in\phi^{-1}(k)}\pi_i}, \quad k,l\in\bar{\mathcal{V}}. \tag{6.18}$$

By expressing $P(\Delta t)$ and $\bar{P}(\Delta t)$ in the form

$$P(\Delta t) = I + A\Delta t + O(\Delta t^2)$$
$$\bar{P}(\Delta t) = I + \bar{A}(\phi)\Delta t + O(\Delta t^2)$$

the equation (6.18) becomes

$$\begin{aligned}\mathbb{1}_{\{k=l\}} + \bar{A}_{kl}(\phi)\Delta t + O(\Delta t^2) &= \frac{\sum_{i\in\phi^{-1}(k)}\sum_{j\in\phi^{-1}(l)}\pi_i(\mathbb{1}_{\{i=j\}} + A_{ij}\Delta t + O(\Delta t^2))}{\sum_{i\in\phi^{-1}(k)}\pi_i} \\ &= \mathbb{1}_{\{k=l\}} + \frac{\sum_{i\in\phi^{-1}(k)}\sum_{j\in\phi^{-1}(l)}\pi_i A_{ij}}{\sum_{i\in\phi^{-1}(k)}\pi_i}\Delta t + O(\Delta t^2).\end{aligned} \tag{6.19}$$

By matching terms on both sides of (6.19), we obtain the formula for the optimal super-transition-rate matrix

$$\bar{A}_{kl}(\phi) = \frac{\sum_{i\in\phi^{-1}(k)}\sum_{j\in\phi^{-1}(l)}\pi_i A_{ij}}{\sum_{i\in\phi^{-1}(k)}\pi_i}, \quad k,l\in\bar{\mathcal{V}}. \tag{6.20}$$

By substituting (6.3) and (6.10) into (6.20), one can verify that $\bar{A}(\phi)$ is indeed a *transition-rate matrix* for any partition function ϕ, i.e., the row sums of $\bar{A}(\phi)$ are zeros, diagonal entries are negative, and non-diagonal entries are non-negative:

$$\begin{cases} \bar{A}_{kl}(\phi) = \dfrac{\sum_{i\in\phi^{-1}(k)}\sum_{j\in\phi^{-1}(l)} 1/R_{ij}}{\sum_{i\in\phi^{-1}(k)} C_i}, \quad k\neq l\in\bar{\mathcal{V}} \\ \bar{A}_{kk}(\phi) = -\displaystyle\sum_{l\neq k}\bar{A}_{kl}(\phi), \quad k\in\bar{\mathcal{V}} \end{cases} \tag{6.21}$$

The super-capacitances and super-resistances can also be expressed in terms of C_i and R_{ij}:

- According to (6.10), the stationary distribution of the aggregated Markov chain has the form:

$$\bar{\pi}_k(\phi) = \frac{\bar{C}_k(\phi)}{\sum_{l\in\bar{\mathcal{V}}} \bar{C}_l(\phi)}, \quad k \in \bar{\mathcal{V}} \tag{6.22}$$

  where $\bar{C}_k(\phi)$ denotes the *super-capacitances* for the kth node. By substituting (6.10) into (6.13), we obtain formula for the the optimal stationary distribution:

$$\bar{\pi}_k(\phi) = \sum_{i\in\phi^{-1}(k)} \pi_i = \frac{\sum_{i\in\phi^{-1}(k)} C_i}{\sum_{l\in\bar{\mathcal{V}}} \sum_{j\in\phi^{-1}(l)} C_j}, \quad k \in \bar{\mathcal{V}}. \tag{6.23}$$

  By comparing (6.22) and (6.23), we obtain the formulae for the super-capacitances:

$$\bar{C}_k(\phi) = \sum_{i\in\phi^{-1}(k)} C_i, \quad k \in \bar{\mathcal{V}}. \tag{6.24}$$

- By using (6.21) and (6.24), we obtain the formulae for the *super-resistances*:

$$\bar{R}_{kl}(\phi) = \frac{1}{\bar{C}_k(\phi)\bar{A}_{kl}(\phi)} = \frac{1}{\sum_{i\in\phi^{-1}(k)} \sum_{j\in\phi^{-1}(l)} 1/R_{ij}}, \quad k \neq l \in \bar{\mathcal{V}}. \tag{6.25}$$

Thus, the reduced-order linear model (6.17) corresponds to a reduced RC-network with super-capacitances and super-resistances given by (6.24) and (6.25), respectively. The super-capacitance $\bar{C}_k(\phi)$, given in (6.24), is the equivalent capacitance of parallel configuration of all capacitors in the kth partition. Similarly, the super-resistance $\bar{R}_{kl}(\phi)$ given in (6.25) is the equivalent resistance of parallel configuration of all resistors connecting the kth partition and the lth partition. These observations also serve to provide an intuitive justification of the aggregation approach.

Similar to the full-order model (6.2), the reduced-order model (6.17) is also *conservative* because of the fact that $\bar{A}(\phi)$ is a super-transition-rate matrix. The invariant quantity for the reduced-order model is given by

$$\bar{V}_0 := \sum_{k\in\bar{\mathcal{V}}} \bar{C}_k \bar{T}_k(0).$$

If one chooses the initial condition for the reduced-order model (6.17) as

$$\bar{T}_k(0) = \sum_{i\in\phi^{-1}(k)} (C_i/\bar{C}_k(\phi)) T_i(0), \quad k \in \bar{\mathcal{V}} \tag{6.26}$$

then

$$\bar{V}_0 = \sum_{k\in\bar{\mathcal{V}}} \sum_{i\in\phi^{-1}(k)} C_i T_i(0) = \sum_{i\in\mathcal{V}} C_i T_i(0) = V_0.$$

This implies that the invariant quantity of the linear thermal dynamics is unchanged after the aggregation. The *aggregated thermal distribution* is defined as

$$\bar{f}_k = \frac{\bar{C}_k}{\bar{V}_0}\bar{T}_k, \quad k \in \bar{\mathcal{V}}. \tag{6.27}$$

Recall that we introduce the lifting technique to compare the low and high-dimensional distributions. The *lifted thermal distribution* is defined as

$$\widehat{f}_i = \frac{C_i}{V_0}\widehat{T}_i, \quad i \in \mathcal{V} \tag{6.28}$$

where $\widehat{T}_i$ is called the *lifted temperature* for the node i. Using (6.15) and choosing the lifting distribution as $\mu = \pi$, we obtain

$$\widehat{f}_i = \frac{\pi_i}{\sum_{j\in\psi(i)}\pi_j}\bar{f}_{\phi(i)}. \tag{6.29}$$

Substituting (6.10) and (6.27) into (6.29), we have

$$\widehat{f}_i = \frac{C_i}{\sum_{j\in\psi(i)} C_j}\frac{\bar{C}_{\phi(i)}}{\bar{V}_0}\bar{T}_{\phi(i)} = \frac{C_i}{V_0}\bar{T}_{\phi(i)} \tag{6.30}$$

where we use the fact that $\bar{C}_{\phi(i)} = \sum_{j\in\psi(i)} C_j$ and $\bar{V}_0 = V_0$. By comparing (6.28) and (6.30), we have the explicit expression for the lifted temperature

$$\widehat{T}_i = \bar{T}_{\phi(i)}, \quad i \in \mathcal{V}.$$

Note that the lifted temperature $\widehat{T}_i$ of the node i is indeed a temperature quantity for the corresponding aggregated node $\phi(i)$. Thus, we can compare the full and reduced-order models by directly comparing T_i and $\bar{T}_{\phi(i)}$ for each node i.

6.4.2 Aggregated building thermal model

Recall that the outside node is taken as a virtual $(n+1)$th node in the full-order building thermal model (6.2), and the outside temperature is denoted as T_{n+1}. We also take the outside node as a virtual $(m+1)$th node in the reduced-order model and we denote its temperature as $\bar{T}_{m+1}$. That is, for any given partition function ϕ, the building node set $\{1,\ldots,n\}$ is aggregated into the super-node set $\{1,\ldots,m\}$, and the $(n+1)$th outside node has a one-to-one correspondence to the $(m+1)$th super-node.

Due to the current source interpretation of nonlinear thermal dynamics $L(T,U,Q)$ (see Section 6.2.1), the current sources connecting to the same group of the aggregated nodes are directly added up to form a super-current source for the corresponding super-node:

- For $k = 1, \ldots, m$, the aggregated nonlinear thermal dynamics is given by:

$$\widetilde{L}_k(T, U, \dot{Q}) = \sum_{i \in \phi^{-1}(k)} C_i L_i(T, U, Q) / \bar{C}_k(\phi) = (C_{pa}(T^s \bar{U}_k(\phi) - \widetilde{W}_k(\phi)) + \dot{\bar{Q}}_k(\phi)) / \bar{C}_k(\phi) \tag{6.31}$$

  where

$$\bar{U}_k(\phi) := \sum_{i \in \phi^{-1}(k)} U_i, \quad \dot{\bar{Q}}_k(\phi) := \sum_{i \in \phi^{-1}(k)} \dot{Q}_i, \quad \widetilde{W}_k(\phi) := \sum_{i \in \phi^{-1}(k)} U_i T_i. \tag{6.32}$$

- For $k = m + 1$, the aggregated nonlinear thermal dynamics is given by:

$$\widetilde{L}_{m+1}(T, U, \dot{Q}) = \eta$$

  where $\eta(t) = \dot{T}_o(t)$ and $T_o(t)$ denotes the outside temperature. The construction here is to make sure $\bar{T}_{m+1}(t) = T_o(t)$ for all $t \geq 0$ in the reduced-order model (6.34) described later.

Since U and Q are external inputs to the full-order model, we can also take $\bar{U}(\phi)$ and $\dot{\bar{Q}}(\phi)$ defined in (6.32) as the super-inputs to the reduced-order model. One problem is that the term $\widetilde{W}(\phi)$ defined in (6.32) depends on T, which is the state vector of the full-order model. We use $\bar{T}_k$ (the temperature of the kth super-node) to approximate T_i (the temperature of the ith node that belongs to the kth group) in $\widetilde{W}_k(\phi)$:

$$\bar{W}_k(\phi) := \sum_{i \in \phi^{-1}(k)} U_i \bar{T}_k = \bar{U}_k(\phi) \bar{T}_k, \quad k = 1, \ldots, m.$$

Replacing $\widetilde{W}_k(\phi)$ by $\bar{W}_k(\phi)$ in (6.31), we approximate $\widetilde{L}_k(T, U, \dot{Q})$ by

$$\bar{L}_k(\bar{T}, \bar{U}(\phi), \dot{\bar{Q}}(\phi)) = \left(C_{pa} \bar{U}_k(\phi)(T^s - \bar{T}_k) + \dot{\bar{Q}}_k(\phi) \right) / \bar{C}_k(\phi) \tag{6.33}$$

for $k = 1, \ldots, m$ and $\bar{L}_{m+1}(\bar{T}, \bar{U}(\phi), \dot{\bar{Q}}(\phi)) = \eta$. Note that the aggregated nonlinear thermal dynamics in (6.33) only depends on super-quantities for the reduced-order model.

By combining the aggregated linear thermal dynamics (6.17) with the aggregated nonlinear thermal (6.33), we obtain the state-space representation of the reduced-order building thermal model:

$$\frac{d\bar{T}}{dt} = \bar{A}(\phi)\bar{T} + \bar{L}(\bar{T}, \bar{U}(\phi), \dot{\bar{Q}}(\phi)). \tag{6.34}$$

The model reduction method proposed in this chapter preserves the RC-network structure of the original building model, that is, the reduced-order model (6.34) is still a RC-network defined with super-nodes with super-edges connecting these super-nodes. According to state-space representation (6.34), the aggregated building thermal dynamics

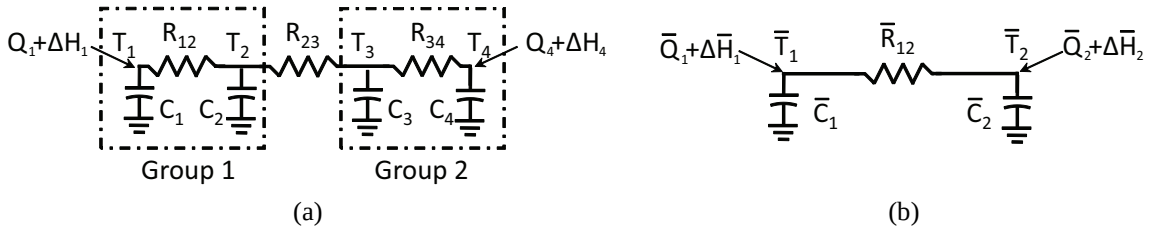


Figure 6.3: The RC-network representation of (a) the full-order thermal model and (b) the reduced-order thermal model for the single surface separating two zones.

can be also expressed by the following coupled differential equations: For each $k = 1, \ldots, m$,

$$\bar{C}_k(\phi)\frac{d\bar{T}_k}{dt}(t) = \sum_{l \in \bar{\mathcal{N}}_k} (\bar{T}_l(t) - \bar{T}_k(t))/\bar{R}_{kl}(\phi) + \dot{\bar{Q}}_k(\phi)(t) + \Delta\bar{H}_k(\phi)(t) \tag{6.35}$$

where $\bar{T}_k$ is the temperature of the kth super-node, $\bar{\mathcal{N}}_k \subset \bar{\mathcal{V}}$ denotes the set of neighbors of the kth super-node, $\dot{\bar{Q}}_k(\phi)$ denotes the heat gain for the kth super-node, and the ventilation heat exchange $\Delta\bar{H}_k(\phi)$ for the kth super-node is given by

$$\Delta\bar{H}_k(\phi)(t) = C_{pa}\dot{\bar{m}}_k^{in}(\phi)(t)(T^s - \bar{T}_k(t))$$

with the mass flow rate entering the kth super-node given by $\dot{\bar{m}}_k^{in}(\phi) = \sum_{i \in \phi^{-1}(k)} \dot{m}_i^{in}$. The initial condition of the reduced-order model (6.35) is chosen as (6.26).

The reduced-order model so far depends on the choice of the partition function ϕ. The sub-optimal partition function ϕ^* is obtained by using the recursive bi-partition algorithm AlgoBIPA. However, one can also directly choose a sub-optimal ϕ^* based on physical intuition (e.g., floor plans in a multi-zone building), or some other kinds of expert-based heuristics.

Example 6.4.1 (Bi-partition of the two-zone building) Consider the model reduction problem for the two-zone building shown in Example 6.2.1. The goal is to find a reduced-order model with two super-states. The thermal dynamics of this two-zone building is described by (6.4). To perform model reduction, we consider the linear dynamics of (6.2) first. The linear model is given by (6.5) with the transition rate matrix

$$A = \begin{bmatrix} -\frac{1}{C_1R_{12}} & \frac{1}{C_1R_{12}} & 0 & 0 \\ \frac{1}{C_2R_{21}} & -\frac{R_{21}+R_{23}}{C_2R_{21}R_{23}} & \frac{1}{C_2R_{23}} & 0 \\ 0 & \frac{1}{C_3R_{32}} & -\frac{R_{32}+R_{34}}{C_3R_{32}R_{34}} & \frac{1}{C_3R_{34}} \\ 0 & 0 & \frac{1}{C_4R_{43}} & -\frac{1}{C_4R_{43}} \end{bmatrix}$$

where the capacitances and resistances are given by $C_1 = 0.1$, $C_2 = 0.15$, $C_3 = 0.2$, $C_4 = 0.25$, $R_{12} = 0.15$,

$R_{34} = 0.15$, and $R_{23} = 1.5$.

Using the bi-partition algorithm AlgoBIPA described in Appendix C, the sub-optimal bi-partition is obtained according to the sign structure of the second largest eigenvalue of the Markov transition matrix $P(\Delta t) = e^{A\Delta t}$, where the step-size $\Delta t = 0.01$. The second eigenvector of $P(\Delta t)$ is given by

$$u^{(2)} = [-0.639, -0.601, +0.312, +0.366].$$

The sign-structure of $u^{(2)}$ suggests the optimal bi-partition function $\phi^* = [1, 1, 2, 2]$. That is, the nodes $\{1\}$ and $\{2\}$ are aggregated to form one super-node, and the states $\{3\}$ and $\{4\}$ are aggregated to form the other super-node. The optimal bi-partition is shown in Figure 6.3 (a).

Using (6.20) with the optimal bi-partition function ϕ^*, we obtain the super-transition-rate matrix

$$\bar{A} = \begin{bmatrix} -\frac{1}{\bar{C}_1 \bar{R}_{12}} & \frac{1}{\bar{C}_1 \bar{R}_{12}} \\ \frac{1}{\bar{C}_2 \bar{R}_{21}} & -\frac{1}{\bar{C}_2 \bar{R}_{21}} \end{bmatrix}$$

where the super-capacitances and super-resistances are obtained according to (6.24) and (6.25)

$$\bar{C}_1 = C_1 + C_2 = 0.25, \quad \bar{C}_2 = C_3 + C_4 = 0.45, \quad \bar{R}_{12} = \bar{R}_{21} = R_{23} = 1.5.$$

The RC-network representation of the reduced-order model is shown in Figure 6.3 (b). Let $\bar{T}_k$ denote the temperatures of super-nodes k for $k = 1, 2$. Using (6.35), we represent the reduced-order building thermal model by the following differential equations:

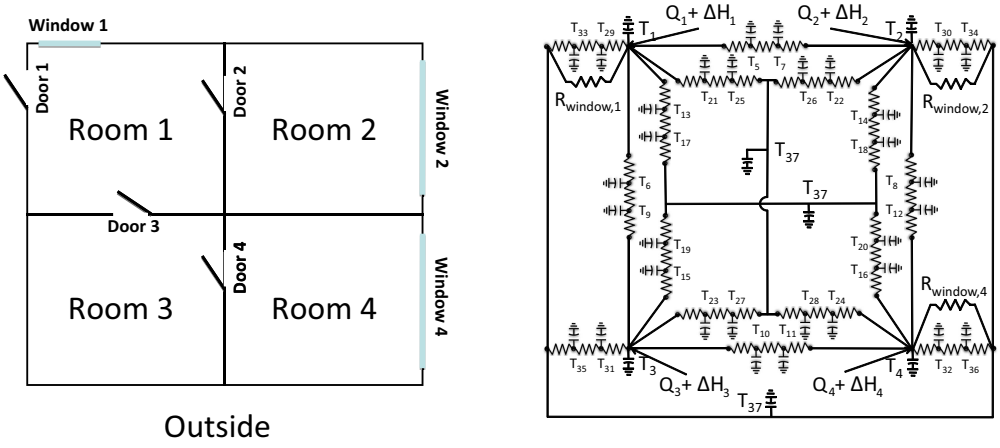
$$\bar{C}_1 \frac{d\bar{T}_1}{dt} = -\frac{1}{\bar{R}_{12}} \bar{T}_1 + \frac{1}{\bar{R}_{12}} \bar{T}_2 + \dot{\bar{Q}}_1 + \Delta \bar{H}_1$$
$$\bar{C}_2 \frac{d\bar{T}_2}{dt} = -\frac{1}{\bar{R}_{21}} \bar{T}_2 + \frac{1}{\bar{R}_{21}} \bar{T}_1 + \dot{\bar{Q}}_2 + \Delta \bar{H}_2$$

where the heat gains $\dot{\bar{Q}}_1 = \dot{Q}_1$ and $\dot{\bar{Q}}_2 = \dot{Q}_4$, the heat exchanges $\Delta \bar{H}_1 = \Delta H_1$ and $\Delta \bar{H}_2 = \Delta H_4$.

The main observations from this model reduction example are as follows:

- The thermal interaction between the group of nodes $\{1, 2\}$ and the group $\{3, 4\}$ is much weaker compared to the interactions within each group since the resistance R_{23} separating the two groups is 10 times larger than the resistances R_{12} and R_{34} within each group. The optimal bi-partition produced by the algorithm AlgoBIPA thus is consistent with the heuristic that would partition the building based on the thermal interactions among the nodes.

- The reduced-order model is also a RC-network, just as the original model is. ■



(a) (b)

Figure 6.4: (a) The layout of the four-zone building of the HVAC system shown in Figure 6.1, and (b) its RC-network representation.

6.5 Numerical Results

In this section, we apply the aggregation-based model reduction method to the downstream part of the four-zone building HVAC system shown in Figure 6.1, where each zone is serviced by a single terminal box. The layout of the four-zone building is shown in Figure 6.4 (a). Each of the four rooms/zones has an equal floor area of $5\mathrm{m} \times 5\mathrm{m}$ and each wall is $3\mathrm{m}$ tall, which provides a volumetric area of $75\mathrm{m}^3$ for each room. Room 1 has a small window ($5\mathrm{m}^2$) on the north facing wall, whereas rooms 2 and 4 have larger windows ($7\mathrm{m}^2$ each) on the east facing wall. Room 3 does not have a window.

The HVAC system used for simulation is designed to supply maximal mass flow rate of 0.25 kg/s per room. The supplied air temperature is fixed at $T^s = 12.8°\mathrm{C}$. Here we assume that there is no return air and 100% of the outside air is sent to the AHU. The number of occupants in each room is uniformly generated as a random integer between 0 and 4. Outside temperature and outside solar radiation data is obtained for a summer day $(05/24/1996)$ of Gainesville, FL [68]. Numerical results presented here are simulated using ode45 function in Matlab for 24 hours with the time step size chosen as 10 minutes. All temperatures are initialized at $24°\mathrm{C}$, respectively. The mass flow rates entering four zones are given by $\dot{m}_1^{in} = 0.15\ \mathrm{kg/sec}$, and $\dot{m}_2^{in} = \dot{m}_3^{in} = \dot{m}_4^{in} = 0\ \mathrm{kg/sec}$. Figure 6.5 shows the other two inputs: outside temperature T_0 and the heat gains $\dot{Q}$.

6.5.1 Recursive bi-partition of building graph

The RC-network representation of the four-zone building is shown in Figure 6.4 (b). There are total 36 building nodes plus 1 outside node for the model of this four-zone building:

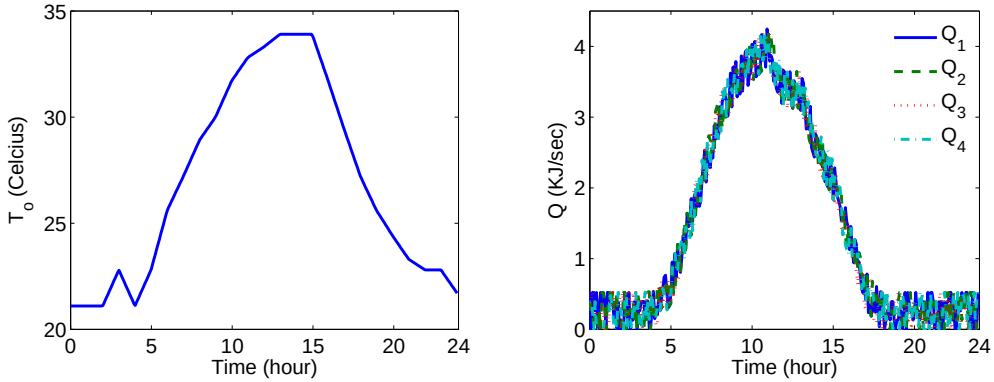


Figure 6.5: (a) The outside temperature T_0, and (b) the total heat gains $\dot{Q}_i$, for room i ($i = 1, \ldots, 4$).

- 4 zone nodes $\{1, \ldots, 4\}$;
- 8 internal-wall nodes $\{5, \ldots, 12\}$;
- 8 internal-floor nodes $\{13, \ldots, 20\}$;
- 8 internal-ceiling nodes $\{21, \ldots, 28\}$;
- 8 external wall nodes $\{29, \ldots, 36\}$;
- 1 outside node $\{37\}$.

Each node is assigned with a thermal capacitance, and two adjacent nodes are connected with a thermal resistance. The windows are modeled as single resistors since they have relatively little capacitance. The values of capacitances and resistances used for simulation are obtained from commercially available software Carrier Hourly Analysis Program [17]. The outside node is assumed to have a very large capacitance $C_{37} = 10^{10}$ KJ/(m^{2}K).

The recursive bi-partition algorithm AlgoBIPA, described in Appendix C, is used to find sub-optimal partitions of the building graph based on the analysis of the linear thermal dynamics. The first iteration of the algorithm AlgoBIPA divides the node set into two groups: the first group contains all building nodes: $\{1, 2, \ldots, 36\}$, and the second group contains only the outside node: $\{37\}$. Such a 2-partition result makes sense since it captures the slowest time-scale of building thermal dynamics.

The second iteration of the algorithm AlgoBIPA leads to a 3-partition, which divides the nodes into three groups: the first group consists of all (zone, wall, ceiling, window, and floor) nodes associated with the room 3, the second group contains all other building nodes associated with the rooms $1, 2, 4$, and the third group contains only the outside node: $\{37\}$. Compared with the 3-partition results, the 4-partition identifies a new group containing all nodes associated with room 1. For the 5-partition, the algorithm returns five groups of nodes with clear physical intuition: group

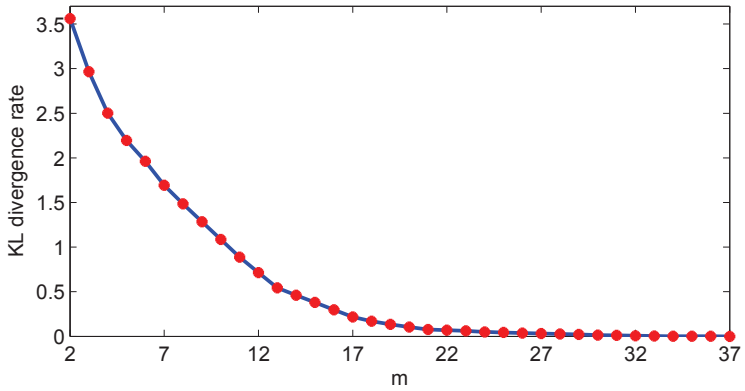


Figure 6.6: Modeling error (K-L divergence rate) in aggregating the linear thermal dynamics vs. number of partitions.

i contains all nodes corresponding to room i, for $i = 1, \ldots, 4$, and group 5 consists of the single outside node $\{37\}$! For $m > 5$, the m-partition further partitions the nodes associated with individual rooms. The largest possible m is 37, which corresponds to no reduction in model order.

Recall that the K-L divergence rate (6.16) is used as a measure of the modeling error for aggregating the linear thermal dynamics. Figure 6.6 depicts the K-L divergence rate with respect to the number of partitions m for $2 \leq m \leq 37$. We observe from Figure 6.6 that the modeling error monotonically decreases to the zero as the number of partitions increases to the dimension of the full-order model, and that there is little additional improvement beyond a model order of around 18. Although the K-L divergence rate is only applicable to the linear part of the model, one can still use it as a conservative guideline for the reduction of the nonlinear model. In that case, we can guess that for good prediction accuracy, the reduced order model should have about 18 states. This is verified by simulations we report next.

6.5.2 Simulation of full and reduced-order models

The full-order model (6.2) is used to describe the full building thermal dynamics, with 36 building nodes plus 1 outside node. The multiple partition results obtained in Section 6.5.1 are used to construct the reduced-order models through aggregation of building nodes into groups, where each group of nodes is represented by a super-node. For $k = 1, \ldots, 36$, the kth-order reduced model (6.34) is used to describe the reduced building thermal dynamics with k super-nodes plus 1 outside node. For comparison, we lift the reduced model to one with 36 building nodes plus 1 outside node (see Section 6.4.1 for more details). That allows direct comparison between the temperature of a zone predicted by the full and reduced-order models.

All simulations reported here are open-loop simulations: the same mass flow rates ($\dot{m}_1^{in} = 0.15$ kg/sec and $\dot{m}_2^{in} = \dot{m}_3^{in} = \dot{m}_4^{in} = 0$ kg/sec) are used as inputs in conducting simulations for both full and reduced-order models;

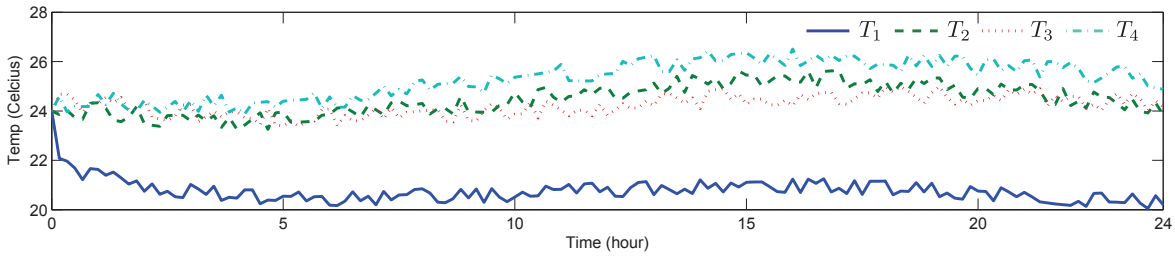


Figure 6.7: Four zone temperatures $T_1, \ldots, T_4$ simulated by the full 36th-order model.

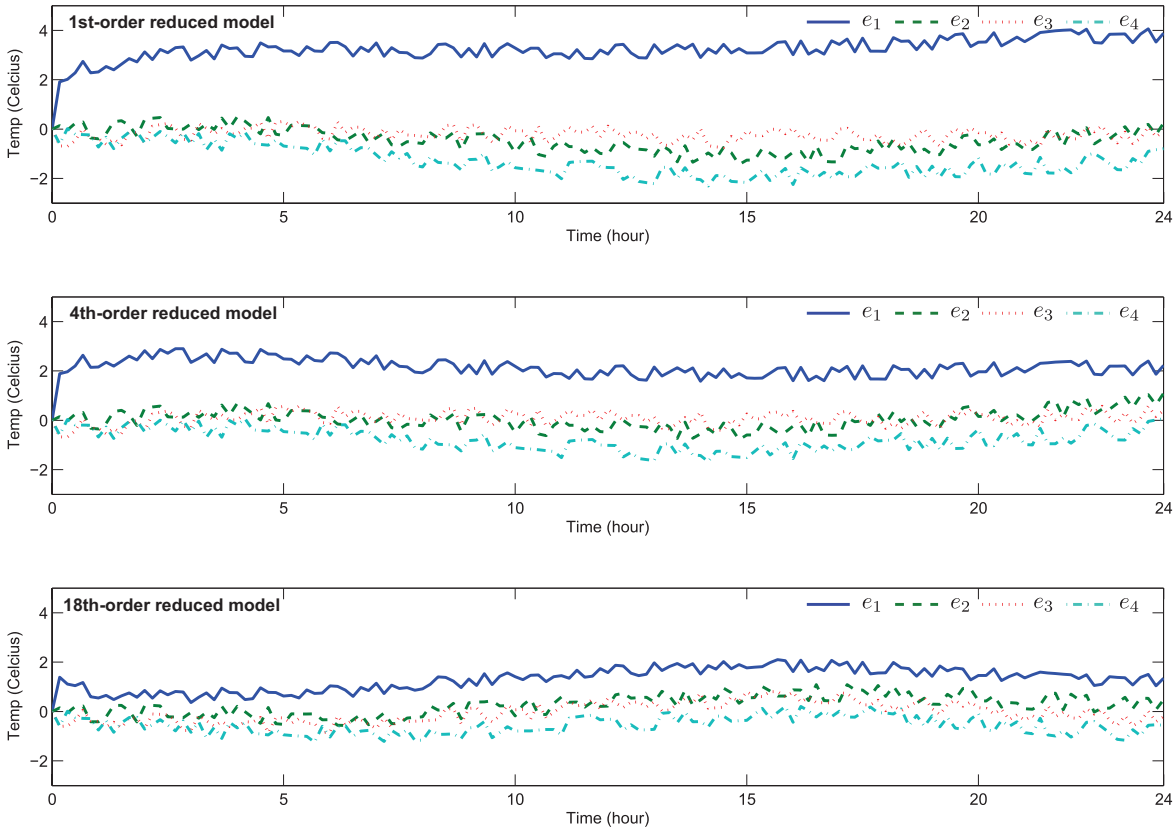


Figure 6.8: Four zone temperature simulation errors are given by $e_i = \widehat{T}_i - T_i$ for $i = 1, \ldots, 4$, where T_i is the temperature simulated by the full-order model and $\widehat{T}_i$ is the lifted temperature simulated by the (1st-order, 4th-order, and 18th-order) reduced models.

the inputs are shown in Figure 6.5. Note that the inputs are aggregated accordingly to obtain the super-inputs for the reduced-order model (see Section 6.4.2 for more details). To test the goodness of the reduced-order models, we compare the four zone temperatures simulated by the full and reduced-order models. When simulated by the full-order model, the temperature of room i is denoted by T_i. When simulated by the reduced-order model, the lifted temperature of room i is denoted by $\widehat{T}_i$ (see Section 6.4.1 for more details). The ith zone temperature prediction error is denoted by $e_i = \widehat{T}_i - T_i$.

Figure 6.7 shows the temperatures of the four zones predicted by the full-order model. Figure 6.8 shows the temperature prediction errors corresponding to reduced-order models with varying degree of reduction: (i) 1st-order reduced model (1 super-node corresponding to all building nodes), (ii) 4th-order reduced model (4 super-nodes corresponding to 4 groups of nodes associated with 4 zones), and (iii) 18th-order reduced model. Note that a kth-order reduced model corresponds to the $(k+1)$-partition described in Section 6.5.1 with k super-building nodes and 1 outside node.

We observe from Figure 6.8 that, as expected, prediction errors decrease as the order of the reduced model increases. In addition, the conjecture based on K-L divergence rate that the 18th-order model will have predictions close to that of the full-order model turns out to be true. In the 18th-order model, the prediction error for the zone with the maximum error (zone 1 here) has a mean of $1.30^{\circ}\mathrm{C}$ and standard deviation of $0.46^{\circ}\mathrm{C}$. Note that even in a building that meets ASHRAE thermal comfort standards, the temperature inside a zone may vary by up to $3^{\circ}\mathrm{C}$ [5]. A lumped model that uses the well-mixed air assumption therefore is fundamentally limited to about a $1.5^{\circ}\mathrm{C}$ prediction error.

We also observe from the Figure 6.8 that *except for zone 1*, temperature prediction with even the 4th-order model (middle plot), which represents a nine-fold reduction in model order, is quite accurate, where the prediction error for the zone with the maximum error (zone 4 here) has a mean of $-0.77^{\circ}\mathrm{C}$ and the standard deviation of $0.44^{\circ}\mathrm{C}$. However, the mean and standard deviation of prediction error for zone 1 with the 4th-order model are $2.14^{\circ}\mathrm{C}$ and $0.37^{\circ}\mathrm{C}$, which is much larger. Thus, large reduction in the model order is not likely to be useful for control design and analysis studies. However, we expect such low order models to be still useful in preliminary building and HVAC system design studies. The higher error in the temperature prediction of zone 1 could be due to the the method's inability to accurately approximate enthalpy dynamics due to the ventilation (note that zone 1 is the only one with the ventilation), or due to the error introduced in lifting the reduced model to a full order model for comparison purpose.

The computation time for executing the Matlab simulation code increases as the order of the reduced model increases. For the case study considered here, the computation times for simulation are $6.829s$, $6.988s$, $7.623s$, $349.86s$ for the 1st-order, 4th-order, 18th-order, and full 36th-order models, respectively. In practice, one can make a tradeoff between the accuracy and complexity of the reduced order model by choosing an appropriate order of the reduced model.

6.6 Conclusions

A method to reduce the order of a building thermal model is proposed via aggregation of states. The original model is a large number of coupled nonlinear differential equations. Structurally, it is an RC-network with nonlinear terms due to ventilation air enthalpy. The heat conservation property of the system is used to obtain an analogy between the linear portion of the thermal dynamics and the continuous-time Markov chains. The aggregation-based model reduction technique for Markov chains can now be applied to the linear portion, with the associated K-L divergence rate serving as a metric for the modeling error. Extension of the aggregation method to the nonlinear building thermal model is then carried out by aggregating inputs accordingly into the super-inputs.

A key advantage of the proposed aggregation technique over existing model reduction techniques is that it is structure-preserving by design. The reduced model is also a RC-network model with nonlinearities with the same structure as those in the original model. Thus, the reduced model directly corresponds to a building with smaller number of zones: groups of zones in the original building are now aggregated to form super-zones. The degree of reduction can be controlled by the user. This property makes the method applicable to not only real-time optimization-based control techniques and off-line control design and analysis, but also to parametric studies during building design phase.

Chapter 7

Mean-Field Control for Energy Efficient Buildings

7.1 Introduction

Buildings are one of the primary consumers of energy. In the United States, buildings are responsible for 30% of energy consumption, and 71% of electricity consumption, while accounting for 33% of CO_2 emissions [3]. A large amount of the energy consumed in buildings is wasted. A major reason for this wastage is *inefficiencies* in the building technologies, particularly in operating the HVAC (heating, ventilation and air conditioning) systems. According to a study commissioned by the U.S. Department of Energy, the current building systems are only $20 - 30\%$ efficient in energy usage [13]. These inefficiencies are in turn caused by the manner in which HVAC systems are currently operated. The temperature in each zone is controlled by a local controller, without regards to the effect that other zones may have on it or the effect it may have on others. Substantial improvement may be possible if inter-zone interactions are taken into account in designing control laws for individual zones.

In this chapter, we propose a *decentralized* optimal control strategy for the zones of a multi-zone building where model complexity is mitigated by using a two pronged approach. First, we use the aggregation-based model reduction technique described in Chapter 6 to construct a reduced-order model of the multi-zone building's thermal dynamics. Second, we use the mean-field intuition from statistical mechanics so that the effect of other zones on a particular zone is captured though a *mean-field* model [70]. Then the whole model (even the reduced model) does not have to be used in computing the controls over short time scales.

By using the mean-field idea, we cast the control problem as a game, whereby each zone has its own control objective modeled as set-point tracking of the local (zonal) temperature. In general, the control problem quickly becomes intractable for even a moderate number of competing objectives. In order to mitigate complexity, we employ the Nash Certainty Equivalence principle to obtain a mean-field description [56]. The mean-field here represents the *net effect* of the entire building envelope on any individual zone. A local optimal zonal control is designed based on the local model of thermal dynamics and its interaction with the building via the mean-field. A consistency relationship is used to enforce the mean-field in a self-consistent manner. The methodology is shown to yield distributed control laws that can easily be implemented on large-scale problems.

We compare the performance of the proposed controller with that of a PI controller. Controllers currently used in commercial buildings use a combination of discrete logic and PID type controllers. Simulations show that the proposed scheme achieves comparable temperature tracking performance while reducing energy consumption by reducing the mass-flow rates entering the zones.

The outline of this chapter is as follows: In Section 7.2, we summarize the baseline and reduced models of building thermal dynamics. In Section 7.3, the mean-field control methodology is described for the building temperature regulation problem. The simulation results and comparisons with PI control appear in Section 7.4.

7.2 Building Thermal Model

7.2.1 Full-order building thermal model

In this chapter, we consider the same four-zone building equipped with a Variable-Air-Volume (VAV) system as depicted in Figure 6.1. The detailed configuration of the system and the modeling of building thermal dynamics are contained in Section 6.2. Here we only summarize the notations and setup of the nonlinear building thermal model used in this chapter.

The building thermal model can be described by an *undirected graph* $G = (\mathcal{V}_0, \mathcal{E})$, where $\mathcal{V}_0 := \{0\} \cup \mathcal{V}$ denotes the set of *nodes* of the graph. The node $\{0\}$ denote the outside and the set $\mathcal{V} := \{1, \ldots, n\}$ denotes the building nodes. The nodes are so indexed that the first N nodes of $\mathcal{V}$ correspond to the zones $1, \ldots, N$, and these are called the *zone nodes*. The next $(n - N)$ nodes of $\mathcal{V}$ correspond to the internal points of the surfaces. These are called the *internal nodes*. An edge (i, j) exists between nodes i and j if there is a resistance connecting them directly. The set $\mathcal{E} \subset \mathcal{V}_0 \times \mathcal{V}_0$ is the set of all *edges*. Therefore each edge (i, j) has an associated *thermal resistance* $R_{ij} \in \mathbb{R}_+$. Since the graph is undirected, $R_{ji} = R_{ij}$ by convention. Each node $i \in \mathcal{V}$ has an associated thermal capacitance C_i.

The states and inputs of the building thermal model are summarized below:

$$\begin{aligned} &\text{States} : T_1, \ldots, T_N, T_{N+1}, \ldots, T_n \\ &\text{Inputs} : T_0, T^s; \dot{m}_i^{in}, \dot{Q}_i^r, \dot{Q}_i^{int}, \dot{Q}_i^{ext}, i = 1, \ldots, N \end{aligned}$$

where $T_1, \ldots, T_N$ denote the space temperature of the zones, $T_{N+1}, \ldots, T_n$ denote the temperature of the points internal to the surface elements, T_0 denotes the outside temperature, T^s denotes the temperature of the air supplied by the AHU, $\dot{m}_i^{in}$ denotes the mass flow rate of the supply air entering the ith zone, $\dot{Q}_i^r$ denotes the rate of heat due to reheat, $\dot{Q}_i^{int}$ denotes the rate of heat generated by occupants, equipments, and lights in the ith zone, and $\dot{Q}_i^{ext}$ denotes the rate of solar radiation entering the ith zone. The supplied air temperature T^s is usually constant for a VAV system, at least over short intervals of time. All other inputs are time varying. In this chapter we assume that (estimates of) the

outside temperature T_0 and the heat gains $\dot{Q}_i^r, \dot{Q}_i^{int}, \dot{Q}_i^{ext}$ are available based on historical data, weather predictions, and various sensors.

The thermal dynamics of a multi-zone building is represented by the following coupled differential equations: For each $i \in \mathcal{V}$,

$$C_i \dot{T}_i(t) = \dot{Q}_i(t) + \Delta H_i(t) + \sum_{j \in \mathcal{N}_i} (T_j(t) - T_i(t))/R_{ij}$$

where $\mathcal{N}_i := \{j \in \mathcal{V}_0 : j \neq i, (i,j) \in \mathcal{E}\}$ denotes the set of *neighbors* of the node i. The *heat gain* $\dot{Q}_i$ is the rate of thermal energy entering the node i from external sources, other than ventilation air and conduction from neighboring nodes:

$$\begin{cases} \dot{Q}_i(t) = \dot{Q}_i^r(t) + \dot{Q}_i^{int}(t) + \dot{Q}_i^{ext}(t), & i = 1, \ldots, N \\ \dot{Q}_i(t) = 0, & i = N+1, \ldots, n. \end{cases}$$

The *ventilation heat exchange* ΔH_i is the rate of thermal energy entering the node i due to ventilation:

$$\begin{cases} \Delta H_i(t) = C_{pa} \dot{m}_i^{in}(t)(T^s - T_i(t)), & i = 1, \ldots, N \\ \Delta H_i(t) = 0, & i = N+1, \ldots, n \end{cases}$$

where C_{pa} is the specific heat capacitance of the supplied air at constant pressure. In the following, the mass flow rate $\dot{m}^{in}$ is the control input u, i.e., $u_i = \dot{m}_i^{in}$ for $i = 1, \ldots, N$.

7.2.2 Reduced-order building thermal model

In this section, we describe a reduced-order building thermal model by using an aggregation technique proposed in [27] (see Chapter 6 for more details). The reduced models will be used in Section 7.3 to develop the mean field control strategies. To obtain the reduced model, we aggregate a subset of nodes into *super-nodes*. Mathematically, suppose we want to reduce the state space dimension from n to m, where $m \leq n$ is the (user-specified) number of super-nodes. The first step is to determine a *partition function* $\phi : \mathcal{V} \to \bar{\mathcal{V}}$, where $\bar{\mathcal{V}} := \{1, \ldots, m\}$ such that ϕ is onto but possibly many-to-one. The elements of $\bar{\mathcal{V}}$ are the super-nodes, and for every $k \in \bar{\mathcal{V}}$, the node set $\phi^{-1}(k) \subset \mathcal{V}$ includes the nodes in the baseline model that are aggregated into the kth super-node. Similar to the baseline model, we let $\{0\}$ denote the outside node and define the set $\bar{\mathcal{V}}_0 := \{0\} \cup \bar{\mathcal{V}}$.

Given a fixed m-partition function ϕ, we introduce the following quantities for the reduced model:

- The *super-capacitance* of the kth partition is the combination of all capacitances of the nodes in kth partition:

$$\bar{C}_k^{(\phi)} := \sum_{i \in \phi^{-1}(k)} C_i, \quad k \in \bar{\mathcal{V}}.$$

- The *super-resistance* between kth and lth partitions is the parallel-equivalence of all resistances connecting the nodes between two partitions:

$$\bar{R}_{kl}^{(\phi)} := \frac{1}{\sum_{i\in\phi^{-1}(k)}\sum_{j\in\phi^{-1}(l)} 1/R_{ij}}, \quad k \neq l \in \bar{\mathcal{V}}.$$

- The *super-load* of the kth partition is the combination of all thermal loads for the zones in the kth partition:

$$\dot{\bar{Q}}_k^{(\phi)}(t) := \sum_{i\in\phi^{-1}(k)} \dot{Q}_i(t), \quad k \in \bar{\mathcal{V}}.$$

The reduced-order model is also a RC-network defined on super-nodes with super-edges connecting these super-nodes. Its thermal dynamics is represented by the following coupled differential equations: For each $k \in \bar{\mathcal{V}}$,

$$\bar{C}_k^{(\phi)}\dot{\bar{T}}_k(t) = \dot{\bar{Q}}_k^{(\phi)}(t) + \Delta\bar{H}_k^{(\phi)}(t) + \sum_{l\in\bar{\mathcal{N}}_k} (\bar{T}_l(t) - \bar{T}_k(t))/\bar{R}_{kl}^{(\phi)} \tag{7.1}$$

where $\bar{T}_k$ is the temperature of the kth super-node, $\bar{\mathcal{N}}_k \subset \bar{\mathcal{V}}_0$ denotes the set of neighbors of the kth super-node, and the ventilation heat exchange for the kth super-node is given by

$$\Delta\bar{H}_k^{(\phi)}(t) := \sum_{i\in\phi^{-1}(k)\cap\{1,\ldots,N\}} \Delta H_i(t) = \sum_{i\in\phi^{-1}(k)\cap\{1,\ldots,N\}} C_{pa}\dot{m}_i^{in}(t)(T^s - T_i(t)).$$

The initial condition of the reduced model (7.1) at initial time t_0 is defined as

$$\bar{T}_k(t_0) = \sum_{i\in\phi^{-1}(k)} (C_i/\bar{C}_k^{(\phi)})T_i(t_0), \quad k \in \bar{\mathcal{V}}.$$

Note that the reduced model (7.1) requires for its inputs: the mass flow rate $\dot{m}_i^{in}$ and the zone temperature T_i for $i = 1, \ldots, N$. These are assumed to be available, or are measured.

The reduced model described so far depends on the choice of the partition function ϕ. We should note that any m-partition function ϕ induces a reduced model with m super-states. In Appendix C, we proposed a recursive bi-partition algorithm to search for the sub-optimal m-partition function ϕ^*. However, one can also directly choose a sub-optimal ϕ^* based on physical intuition (e.g., floors in a multi-zone building), or some kind of expert-based heuristics. The goodness of the reduced model (7.1) with ϕ^* can be verified in practice. In this chapter, we will not discuss the algorithms for choosing the optimal partition function ϕ^*. In the following, we assume that ϕ^* has already already properly specified, and we mainly focus on how to design optimal control laws by taking advantage of the reduced building model.

Example 7.2.1 (Reduced model with one super-node) For the reduced model (6.35), one may pick any number of super-nodes (less than n) to reduce the complexity. In the extreme case, with $m = 1$, we only have one super-node to represent the entire building thermal dynamics. The thermal dynamics of this single super-node is represented by the following differential equation:

$$\bar{C}\dot{\bar{T}}(t) = \dot{\bar{Q}}(t) + \Delta\bar{H}(t) + (T_0 - \bar{T}(t))/\bar{R}_0 \tag{7.2}$$

where two super-quantities are defined as

$$\bar{C} := \sum_{i=1}^{n} C_i, \quad \bar{R}_0 := \frac{1}{\sum_{i\in\mathcal{N}_0} 1/R_{i0}}$$

and two super-inputs are given by

$$\dot{\bar{Q}}(t) = \sum_{i=1}^{N} \dot{Q}_i(t), \quad \Delta\bar{H}(t) = C_{pa} \sum_{i=1}^{N} \dot{m}_i^{in}(t)(T^s - T_i(t)).$$

The initial condition of (7.2) is given by

$$\bar{T}(t_0) = \sum_{i=1}^{n} (C_i/\bar{C}) T_i(t_0).$$

The single super-node model described by (7.2) is an extremely simplified version of the building thermal dynamics described by (6.1). The dynamics are very slow due to the very large capacitance $\bar{C}$. The input of (7.2) requires the combination of mass flow rates of all zones. In the following of this chapter, we use the reduced model (7.2) to describe the *mass-behavior of the building*. ■

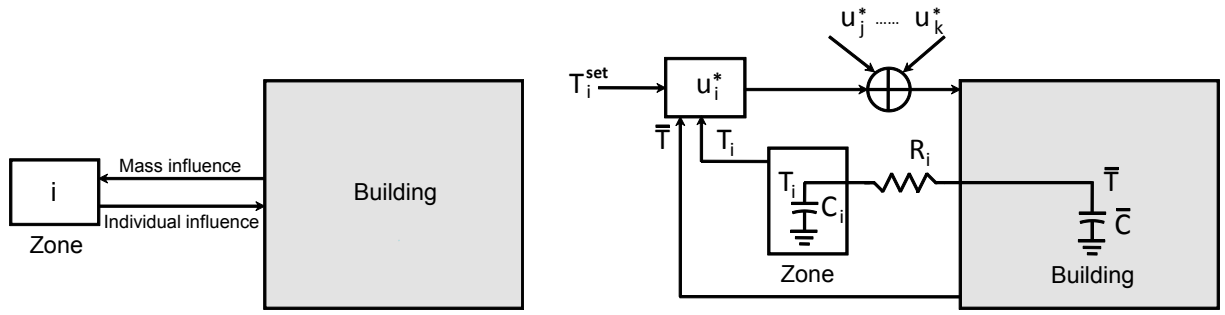
7.3 Mean-Field Control

We consider N zones, each with its local set-point tracking control objective. The dynamics of the ith zone is given by

$$\dot{T}_i = l_i^\circ(T_i; T_{-i}) + b_i(T_i)u_i + d_i \tag{7.3}$$

where

$$l_i^\circ(T_i; T_{-i}) := -\sum_{j\in\mathcal{N}_i} (T_i - T_j)/(C_i R_{ij}), \quad b_i(T_i) := C_{pa}(T^s - T_i)/C_i, \quad d_i := \dot{Q}_i/C_i$$



(a) (b)

Figure 7.1: Conceptual illustration of mean-field control: (a) individual (zone) playing with mass (whole building); (b) mean-field control diagram, where the whole building thermal dynamics is represented by a reduced model with one super-node.

with $T_{-i} := (T_j)_{j \neq i}$. The control problem for the ith zone is to minimize the finite-horizon cost function

$$J_i^\circ(u_i; u_{-i}) = \int_{t_0}^{t_1} c(T_i, u_i) dt \tag{7.4}$$

where

$$c(T_i, u_i) := \frac{1}{2}\Delta T_i^2 + \frac{1}{2} r u_i^2 \tag{7.5}$$

with the tracking error $\Delta T_i := T_i - T_i^{set}$, and a given scalar $r > 0$ as control penalty. A *Nash equilibrium* in control policies is given by $\{u_i^*\}_{i=1}^N$ such that u_i^* minimizes $J_i^\circ(u_i; u_{-i}^*)$ for $i = 1, \ldots, N$.

We denote $R_i := (\sum_{j \in \mathcal{N}_i} 1/R_{ij})^{-1}$ and define the time constant for the ith zone as $\tau_i := C_i R_i$. The individual zones are distinguished by their initial conditions $T_i(0)$, set-point T_i^{set}, loads d_i and the time-constants τ_i. We introduce a parameter $\omega := (T(0), T^{set}, \tau, d)$, and consider a large number N of zones, where ω is sampled from a given distribution $\rho(\omega)$. For each zone i, the parameter ω_i is assumed to be i.i.d., with common distribution $\omega_i \sim \rho(\omega)$.

We seek a control solution that is decentralized and of the following form: For each $i \in \mathcal{V}$ and $t_0 \leq t \leq t_1$, the control input $u_i(t)$ depends only on *local information* $\{T_i(s) : t_0 \leq s \leq t\}$, and perhaps some *aggregate information*. This amounts to a dynamic game, whose exact solution is infeasible for large N.

Instead we construct an approximation of the form described in [55]. This approximation is based on the aggregated models described in Section 7.2.2 with the following steps:

(i) We identify a small number of super-nodes that describe the slow evolution of the thermal dynamics of the building.

To simplify the introduction of the mean-field control method, we consider here only the simplest case: we use a single super-node to represent the entire building.

(iii) We consider an approximation of the interaction between a single zone and the entire building. Motivated by

the consideration of the physics of thermal interactions (large time constants for interactions) and the separable nature of the control objectives (e.g., (7.4)), we consider an approximation based on replacing $T_{-i}(t)$ by $F(t)$, a known function of time. In particular, $l_i^\circ(T_i(t); T_{-i}(t))$ in (7.3) is replaced by

$$\bar{l}_i(T_i(t); F(t)) := -\frac{T_i(t) - F(t)}{\tau_i}. \tag{7.6}$$

Comparison of l_i° and $\bar{l}_i$ suggests the following approximation

$$F(t) \approx \bar{T}(t). \tag{7.7}$$

(iii) For the local model (7.3) with $l_i^\circ(T_i; T_{-i})$ replaced by $\bar{l}_i(T_i; F)$, the game reduces to decentralized optimal control problems. The individual zones are "oblivious" to the state of the entire system and make their control decisions based only on local state variables.

(iv) A form of *self-consistency* is required: oblivious actions of individual zones reproduce the evolution of $\bar{T}$ as described by the aggregated model.

The conceptual diagram of mean-field control is depicted in Figure 7.1. In the following subsection we develop the "oblivious" solution described in (iii). We then turn to the self-consistent aggregated model in (i) that defines the approximate interaction (7.7) in (ii). Mathematically, we obtain a fixed-point problem.

7.3.1 Local optimal control of a single zone

Suppose the interaction function $F(t)$ is given, possibly in a time-dependent form for $t \in [t_0, t_1]$. We consider the following dynamics for the single-zone:

$$\dot{T}_i = \bar{l}_i(T_i; F) + b_i(T_i)u_i + d_i$$

where $\bar{l}_i$ is given by (7.6).

The control problem for single zone model is to choose the control law u_i so as to minimize the finite-horizon cost function

$$J_i(u_i; F) = \int_{t_0}^{t_1} c(T_i, u_i)dt.$$

The solution of the optimal control problem with the cost function $J_i(u_i; F)$ is standard. It is given in terms of the

optimal cost-to-go function or *value function*:

$$J_i^*(T_i, t) = \min_{u_i} \left\{ \int_t^{t_1} c_i(T_i, u_i) dt \right\}.$$

The value function J_i^* is known to satisfy the Hamilton-Jacobi-Bellman (HJB) equation

$$\frac{\partial J_i^*}{\partial t} + \min_{u_i} \left\{ H_i \left(T_i, u_i, \frac{\partial J_i^*}{\partial T_i} \right) \right\} = 0 \tag{7.8}$$

with the boundary condition $J_i^*(T_i, t_1) = 0$. The *Hamiltonian* in (7.8) is defined for $\lambda \in \mathbb{R}$

$$H_i(T_i, u_i, \lambda) := c_i(T_i, u_i) + \lambda \left(l_i(T_i; F) + b_i(T_i) u_i \right)$$

with

$$l_i(T_i; F) := \bar{l}_i(T_i; F) + d_i.$$

The optimal control in (7.8) is explicitly obtained as

$$u_i^*(T_i; F) = -\frac{b_i(T_i)}{r} \left(\frac{\partial J_i^*}{\partial T_i}(T_i; F) \right). \tag{7.9}$$

Substituting (7.9) into (7.8), we obtain the HJB equation for $J_i^*(T_i, t)$:

$$\frac{\partial J_i^*}{\partial t} = \frac{1}{2} \frac{b_i^2(T_i)}{r} \left(\frac{\partial J_i^*}{\partial T_i} \right)^2 - l_i(T_i; F) \left(\frac{\partial J_i^*}{\partial T_i} \right) - \frac{1}{2} \Delta T_i^2. \tag{7.10}$$

7.3.2 Coupled model

We now provide a complete description of the *coupled model* that is intended to approximate the game model for large N. This model is based on the interaction function $F(t)$ introduced in the preceding section. A value function function $J^*(T, t; \omega)$ for the large N model is defined by the following differential equation identical to the HJB equation (7.10) for the single-zone model.

$$\frac{\partial J^*}{\partial t} = \frac{1}{2} \frac{b^2(T)}{r} \left(\frac{\partial J^*}{\partial T} \right)^2 - l(T; F) \left(\frac{\partial J^*}{\partial T} \right) - \frac{1}{2} \Delta T^2.$$

The associated optimal feedback control law is then defined by

$$u^*(T; F) = -\frac{b(T)}{r} \left(\frac{\partial J^*}{\partial T}(T; F) \right). \tag{7.11}$$

Given the feedback control law (7.11), the differential equation that defines the evolution of the super-temperature $\bar{T}$ is given by

$$\bar{C}\dot{\bar{T}}(t) = (T_0(t) - \bar{T}(t))/\bar{R}_0 + \bar{U}(t)$$

where

$$\bar{U}(t) = N \int (\dot{Q}(t;\omega) + C_{pa} u^*(T(t;\omega); F)(T^s - T(t;\omega)))\rho(\omega) d\omega.$$

The only difference thus far is notational: $J_i^*(T,t)$ is the value function for a single zone with parameter ω_i, and $J^*(T,t;\omega)$ is the value function for a large number of zones, distinguished by their own ω. Such is the case because we have *assumed* $F(t)$ is a known deterministic function that is consistent across the population. All that remains is to specify $F(t)$ in a self-consistent manner. The consistency enforced here is inspired by the approximation given in (7.7). The two PDEs are coupled through this integral that defines the relationship between the interaction function F and the mean temperature $\bar{T}$:

$$F(t) = \bar{T}(t).$$

In summary, the coupled PDE model is given by: For $t \in [t_0, t_1]$,

$$\frac{\partial J^*}{\partial t} = \frac{1}{2}\frac{b^2(T)}{r}\left(\frac{\partial J^*}{\partial T}\right)^2 - l(T;F)\left(\frac{\partial J^*}{\partial T}\right) - \frac{1}{2}\Delta T^2 \tag{7.12a}$$

$$\bar{C}\dot{\bar{T}}(t) = (T_0(t) - \bar{T}(t))/\bar{R}_0 + \bar{U}(t) \tag{7.12b}$$

$$F(t) = \bar{T}(t) \tag{7.12c}$$

with boundary conditions $J^*(T, t_1; \omega) = 0$ and $\bar{T}(t_0) = \sum_{i=1}^n (C_i/\bar{C}) T_i(t_0)$.

Numerically, the optimal control may be obtained by iteratively solving the equations (7.12) over a sufficiently long time-horizon. A waveform relaxation algorithm for solving such equations appear in our earlier paper [127]. In the following, we propose an approximate solution based on the observation that the value function is known to approximately become a constant for large terminating times [9].

7.3.3 Approximate local optimal control

In this section, we propose an approximation approach to the solution of coupled PDE (7.12) by considering the the *equilibrium solutions*. By setting $\partial J^*/\partial t \approx 0$ and letting $F = \bar{T}$ in (7.12a), we consider the equilibrium solution to (7.12a):

$$k(T)\left(\frac{\partial J^*}{\partial T}\right)^2 - 2m(T,\bar{T})\left(\frac{\partial J^*}{\partial T}\right) - n(\Delta T) = 0 \tag{7.13}$$

where we define

$$k(T) := b^2(T),\ m(T,\bar{T}) := rl(T;\bar{T}),\ n(\Delta T) := r\Delta T^2.$$

Since T^s (the temperature of supplied air) is always strictly less/more than T (the temperature of zone) when cooling/heating, then we always have

$$k(T) = (C_{pa}(T^s - T)/C)^2 > 0.$$

Thus (7.13) is a second order equation, whose solutions are given by,

$$\left(\frac{\partial J^*}{\partial T}\right)_{\pm} = \frac{m(T,\bar{T})}{k(T)} \pm \frac{\sqrt{m^2(T,\bar{T}) + k(T)n(\Delta T)}}{k(T)}.$$

For any T, ΔT, and $\bar{T}$, we can check

$$\left(\frac{\partial J^*}{\partial T}\right)_{+} \geq 0, \quad \left(\frac{\partial J^*}{\partial T}\right)_{-} \leq 0.$$

We would like to construct a "value function" $\widehat{J}^*$ such that it is approximately convex with respect to T with minimum achieved at $\Delta T = T - T^{set} = 0$. One possible choice is

$$\frac{\partial \widehat{J}^*}{\partial T} = \begin{cases} \left(\partial J^*/\partial T\right)_{+}, & \text{if } \Delta T \geq 0 \\ \left(\partial J^*/\partial T\right)_{-}, & \text{if } \Delta T < 0. \end{cases} \tag{7.14}$$

However, such a choice of $(\partial \widehat{J}^*/\partial T)$ is not a smooth function of T, and neither is the associated control law (7.11). Note that the mass flow rate (the control) is usually varied continuously to regulate the zone temperature for a building. To obtain a smooth control law, here we consider a smooth approximation to the sign function,

$$\operatorname{sgn}(x) \approx \tanh(cx) = \frac{1 - e^{-2cx}}{1 + e^{-2cx}}, \quad \text{for} \quad c \gg 1. \tag{7.15}$$

Then we modify (7.14) to obtain the following smooth approximation:

$$\frac{\partial \widehat{J}^*}{\partial T} = \frac{m(T,\bar{T})}{k(T)} + \tanh(c\Delta T)\frac{\sqrt{m^2(T,\bar{T}) + k(T)n(\Delta T)}}{k(T)}$$

The approximate local optimal control law is chosen as

$$\widehat{u}^*(T;\bar{T}) = -\frac{b(T)}{r}\left(\frac{\partial \widehat{J}^*}{\partial T}(T;\bar{T})\right). \tag{7.16}$$

By setting $\dot{\bar{T}} \approx 0$ in (7.12b) and substituting (7.16) into (7.12b), we can obtain the equilibrium solution $\bar{T}^s > 0$ by solving a second order equation (we omit the details here). Finally, the stationary local optimal control law is given by

$$\widehat{u}^{*,s}(T) = -\frac{b(T)}{r}\left(\frac{\partial \widehat{J}^*}{\partial T}(T;\bar{T}^s)\right). \tag{7.17}$$

7.3.4 Mean-field Control of Linearized System

In this section, we describe an optimal control law based on linearizing the building model along its nominal trajectory. Note that the nominal values of state (the zone temperatures) and control (the mass flow rates) can be directly measured in practice.

By assuming the nominal state and the control values T^o and u^o, we linearize (7.8) and obtain the following linearized model: For $i = 1, \ldots, N$,

$$\dot{T}_i = \tilde{a}_i T_i + \tilde{b}_i u_i + \tilde{l}_i(F) + d_i$$

where we define the following quantities

$$\tilde{a}_i := -1/\tau_i - C_{pa} u_i^o / C_i,$$
$$\tilde{b}_i := C_{pa}(T^s - T_i^o)/C_i,$$
$$\tilde{l}_i(F) := F/\tau_i + C_{pa} T_i^o u_i^o / C_i.$$

We consider the same cost function (7.8) for each zone node. The local optimal control for the ith zone is explicitly obtained as:

$$u_i^*(T_i; F) = -\frac{\tilde{b}_i}{r_i}(\tilde{p}_i T_i - \tilde{g}_i)$$

where $\widetilde{p}_i$ is the positive solution to the Ricatti differential equation

$$\dot{\tilde{p}}_i = -2\tilde{p}_i\tilde{a}_i + \tilde{b}_i^2\tilde{p}_i^2/r - 1$$

with the boundary condition $\tilde{p}_i(t_1) = 0$, and $\tilde{g}_i$ is the solution of the linear differential equation

$$\dot{\tilde{g}}_i = -(\tilde{a}_i - \tilde{b}_i^2\tilde{p}_i/r)\tilde{g}_i - T_i^{set} + \tilde{p}_i(\tilde{l}_i(F) + d_i)$$

with the boundary condition $\tilde{g}_i(t_1) = 0$.

Similarly to Section 7.3.2, we can obtain the coupled PDE model by letting $F(t) \approx \bar{T}(t)$. We consider the case that zone parameters are drawn from certain distribution $\rho(\omega)$ and consider the reduced model (7.12) to evolve $\bar{T}$. As

before, we can also find the approximate local optimal control by considering the equilibrium solutions:

$$0 = -2\tilde{p}\tilde{a} + \tilde{b}^2\tilde{p}^2/r - 1$$
$$0 = -(\tilde{a} - \tilde{b}^2\tilde{p}/r)\tilde{g} - T^{set} + \tilde{p}(\tilde{l}(\bar{T}) + d)$$

which leads to following equilibrium solutions

$$\tilde{p}^s = \frac{r\tilde{a} + \sqrt{r^2\tilde{a}^2 + r\tilde{b}^2}}{\tilde{b}^2}, \quad \tilde{g}^s(\bar{T}) = \frac{\tilde{p}^s(\tilde{l}(\bar{T}) + d) - T^{set}}{\tilde{a} - \tilde{b}^2\tilde{p}^s/r}.$$

The approximate local optimal control is then chosen as:

$$\widehat{u}^*(T;\bar{T}) = -\frac{\tilde{b}}{r}\left(\tilde{p}^s T - \tilde{g}^s(\bar{T})\right).$$

By substituting (7.18) into (7.12b), we can obtain the equilibrium solution $\bar{T}^s$, and therefore the stationary local optimal control law $\widehat{u}^{*,s}(T)$.

7.4 Numerical Results

7.4.1 Basic setup

Simulations are carried out for the four-zone building shown in Figure 6.1: All four zones/rooms have an equal floor area of $5m \times 5m$ and each wall is $3m$ tall, which provides a volumetric area of $75m^3$ for each room. Room 1 has a small window ($5m^2$) on the north facing wall, whereas rooms 2 and 4 have larger windows ($7m^2$ each) on the east facing wall. Room 3 does not have a window. The RC-network representation of the four-zone building has totally 36 building nodes plus 1 outside node [27]. Each building node is assigned with a thermal capacitance, two adjacent nodes are connected with a thermal resistance. The windows are modeled as single resistors since they have relatively little capacitance. The values of the capacitances and resistances are obtained from Carrier's Hourly Analysis Program [17].

The HVAC system used for simulation is designed to supply maximal mass flow rate of 0.25 kg/s per zone. The mass flow rate m_i^{in} for $i = 1, \ldots, 4$ for four zones can be adjusted based on designed control laws. The supplied air temperature is fixed at $T^s = 12.8°C$. Here we assume there is no return air and 100% of the outside air is sent to chiller. Number of people in each zone is uniformly generated as a random integer ranging between 0 and 4. Outside temperature and outside solar radiation data is obtained for a summer day ($05/24/1996$) of Gainesville, FL [68]. The outside temperature and the heat gains (due to solar radiation and people occupancy) of each zone are depicted as the same in Figure 6.5 (a) and Figure 6.5 (b), respectively.

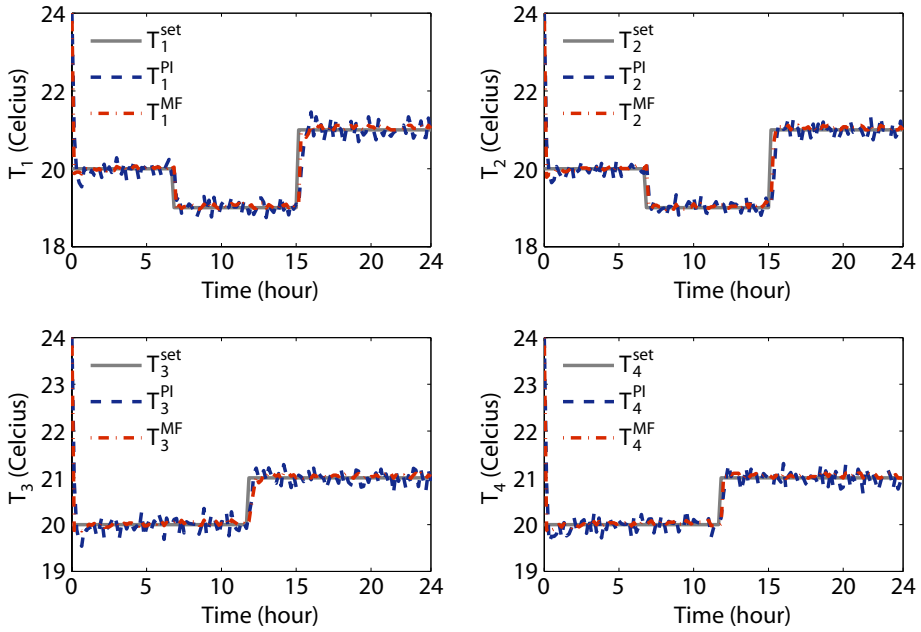


Figure 7.2: Comparison of zone temperatures simulated by using PI control law and mean-field control law with $r = 10$.

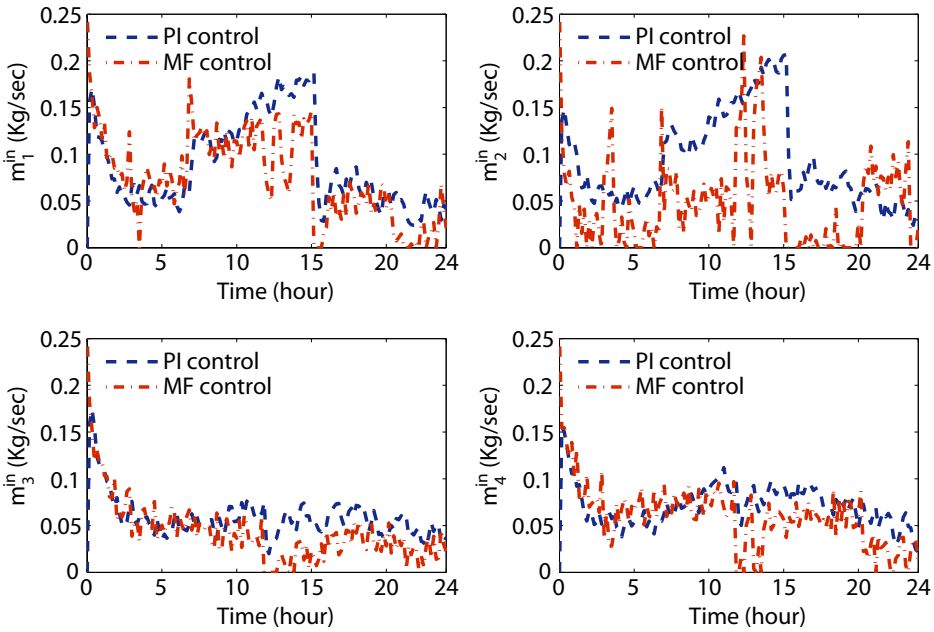


Figure 7.3: Comparison of zone mass-flow rates obtained by using PI control law and mean-field control law with $r = 10$.

Numerical results presented in the following are obtained using ode45 function in Matlab for 24 hours with the time step size chosen as 10 minutes. All temperatures of the building nodes are initialized at $24^{\circ}C$, respectively. The desired zone temperatures T_i^{set} for $i = 1, \ldots, 4$ are varying with time and are depicted as solid lines in Figure 7.2 and Figure 7.4. In the following, we apply the mean-field and PI control laws to regulate zone temperatures to track the desired values.

7.4.2 Simulation results

To compare performance of the proposed controller with existing control algorithms commonly used in commercial buildings, we consider the following decentralized PI control law: For $i = 1, \ldots, 4$,

$$m_i^{in}(t) = K_p \Delta T_i(t) + K_i \int_0^t \Delta T_i(s) ds \tag{7.18}$$

where the tuned proportional gain $K_p = -0.00005$, the tuned integration gain $K_i = 0.0001$, and the temperature tracking error $\Delta T_i := T_i^{set} - T_i$. We next compare the performance of the mean-field control law introduced in this chapter with that of the PI controller.

The mean-field control is implemented based on the stationary nonlinear policy (7.17). For the four-room building, we only consider the reduced model with one super-node. The control performance becomes slightly better by adding more super-nodes into the reduced model. One may expect larger performance improvement by considering more super-nodes for more complex building topologies.

First, we take $r = 10$ in the individual cost function (7.5), and we take $c = 5$ for smooth approximation of sign function in (7.15). We apply mean-field control law (7.17) and PI control law (7.18) to each zone, respectively. The comparison results of simulated zone temperatures are depicted in Figure 7.2. The comparison results of mass-flow rates associated with two control laws are depicted in Figure 7.3. We observe that the mean-field control has better temperature tracking performance than that for PI control (see Figure 7.2).

The total energy consumption of each zone can be computed based on the mass-flow rate entering each zone [45]. Here the total energy consumption is the combination of fan power and the chiller power consumptions. For PI control law, the energy consumption (kWh) for each zone is 69.3, 72.5, 45.2, and 55.9, and the total energy consumption for all four zones is 242.9; For mean-field control law, the energy consumption for each zone is 59.6, 38.9, 34.0, and 47.7, and the total energy consumption for all four zones is 179.2. In this case, the mean-field control thus reduces total energy consumption by 25% over the PI control.

Then, we take $r = 60$ in the cost function (7.5). The comparison results of simulated zone temperatures are depicted in Figure 7.4. The comparison results of mass-flow rates associated two control laws are depicted in Figure 7.5.

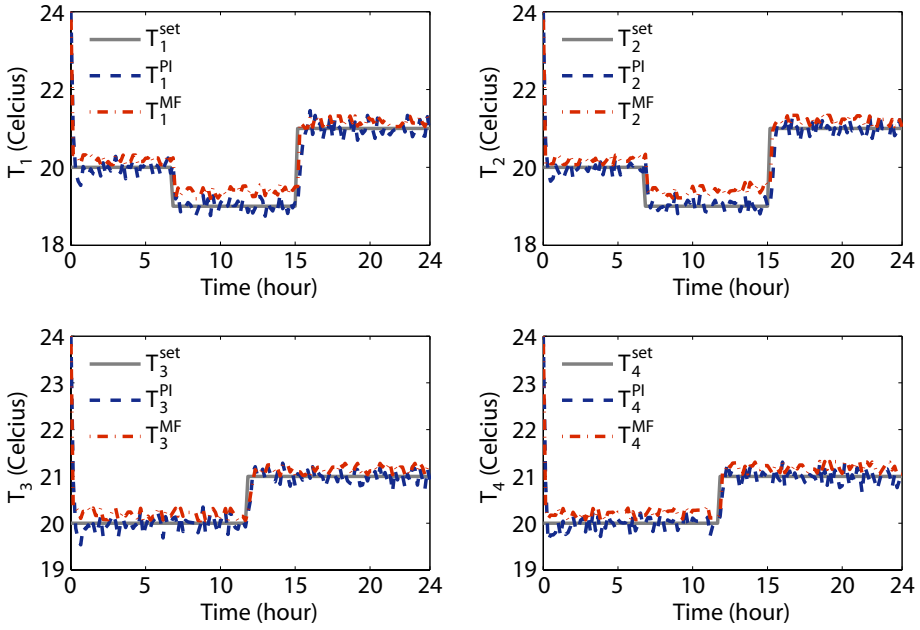


Figure 7.4: Comparison of zone temperatures simulated using PI control law and mean-field control law with $r = 60$.

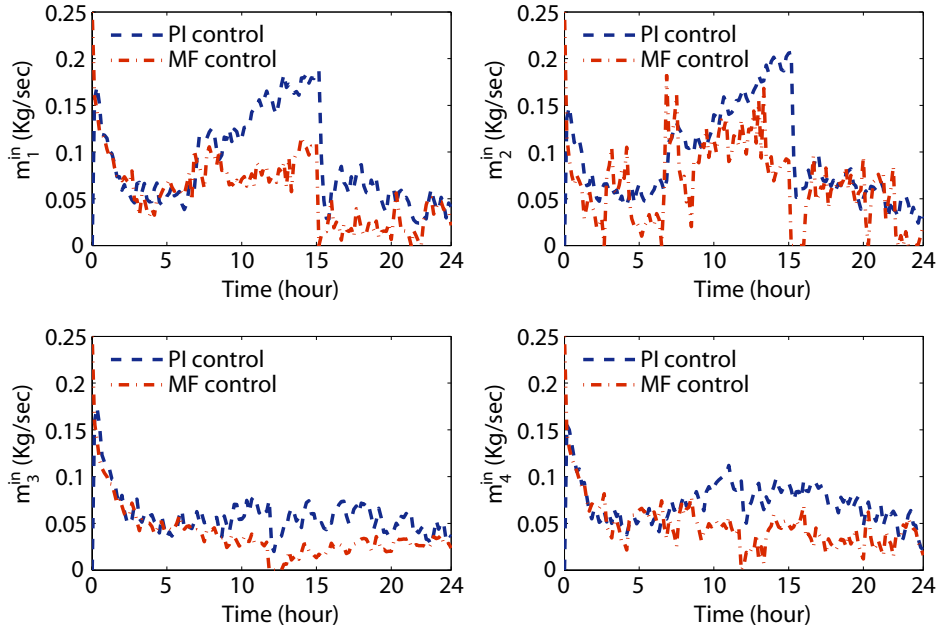


Figure 7.5: Comparison of zone mass-flow rates obtained using PI control law and mean-field control law with $r = 60$.

For mean-field control law, the total energy consumption (kWh) for each zone is 45.9, 52.1, 30.0, and 37.7, and the total energy consumption for all four zones is 165.7. We observe that the temperature tracking performance for mean-field control is slightly worse than that for PI control (see Figure 7.4). But in this case the mean-field control reduces energy consumption by 31% over the PI control, larger than in the previous case (compare Figure 7.3 with Figure 7.5). This is due to a larger control penalty parameter r.

In practice, one can make a tradeoff between temperature tracking performance and the energy consumption by adjusting the control penalty parameter r. The simulation results for linearized control policy (7.18) are omitted in this paper on account of space. In simulations, the linearized policy consumes more energy while maintaining comparable temperature tracking performance, as compared to the nonlinear control policy (7.17). These inefficiencies are caused due to the unmodeled nonlinearities.

7.5 Conclusions

In this chapter, a mean-field methodology is proposed as a means to mitigate complexity associated with large-scale control problems in buildings. Rather than solving the large-scale centralized problem, we explore distributed game-theoretic solution approaches that work by optimizing with respect to the mean-field. Simulation results show that the proposed mean-field scheme achieves comparable temperature tracking performance while reducing energy consumption in the operation of HVAC systems. Moreover, the tradeoff between tracking performance and the energy savings can be made by adjusting the control penalty parameter in the mean-field scheme.

Part III

Optimal Evacuation of Large Buildings

Chapter 8

Optimal Evacuation of Large Buildings via Workload Relaxation

8.1 Introduction

The purpose of this chapter is to introduce tools for the modeling, analysis, and control of occupancy evolution in a large building. The starting point of this chapter is the introduction of queueing models for occupancy evolution in the spirit of the previous works. We then apply several complexity mitigation techniques for the purposes of control-oriented modeling and optimization. Specifically, we adapt workload relaxation techniques from queueing theory to formulate appropriate reduced order dynamic models, and to address performance bounds and optimization problems in the context of building evacuation. The main ingredients of this chapter are surveyed in Section 8.1.1 – Section 8.1.4.

8.1.1 Modeling occupancy evolution

A queueing model is proposed to model agent dynamics such as preferred walking speeds, and preference for exits in the building. The model includes capacity constraints that represent flow and occupancy constraints, such as the restricted rate of flow resulting from a narrow hallway or stairwell. In the presence of congestion, queues build up, increasing the time to evacuate the building. Resource allocation, as defined in queueing theory, corresponds in this application to directions for occupants on how best to exit the building. This might be realized through controllable LED signs at corridors indicating which direction to go, based on current building conditions.

The queueing models proposed in this chapter provide a flexible framework for modeling occupancy evolution. For example, one can easily include upper bounds on node occupancy to model physical constraints. The building population need not be homogeneous – different types of agents (e.g., handicapped or older adults) can be modeled by including more than one queue at a node, for example.

Two general classes of queueing model are considered in this chapter. The first model class is based on the *Controlled Random Walk (CRW) model* of [80]; it is favored for its simplicity. A more refined stochastic model based on renewal or Markov renewal primitives (e.g., [25]) may be more useful for statistical applications such as forecasting of occupancy behavior. The second model class is even more idealized, the *fluid model* [80, 86] that models control

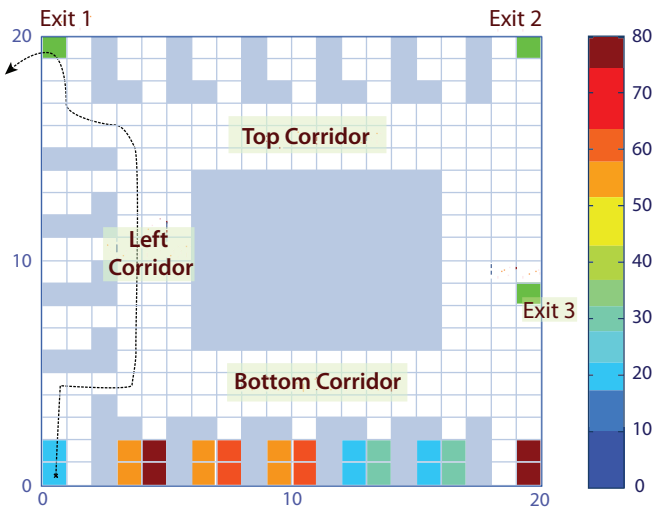


Figure 8.1: Layout of a single-floor building containing numerous offices, and three exits. The selected path indicates the trajectory of a single occupant moving from her office to her preferred/nearest exit.

but not variability. The rationale for studying this model is explained at length in [80]. The main results of [26, 78] provide the most compelling motivation: The value function obtained in transient or average-cost optimal control of a stochastic model is approximated by associated value functions for the fluid model. We apply this insight to construct policies for approximately optimal evacuation.

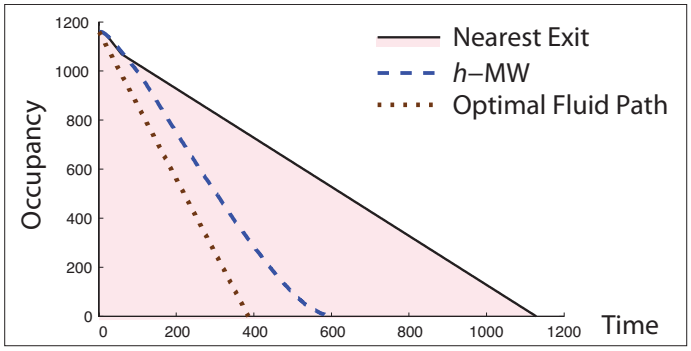


Figure 8.2: Comparison of the evacuation times obtained using two policies for a single floor building shown in Figure 8.1. Also shown is an optimal trajectory for the fluid model. The h-MaxWeight (h-MW) policy most closely reaches the performance of the optimal fluid trajectory.

Shown in Figure 8.2 is a comparison of the h-MaxWeight policy proposed in this chapter for evacuation, and an optimal path from the fluid model. Although the simulation involved a stochastic model, the evolution of occupancy appears to be nearly deterministic. The gap between the fluid and stochastic paths is explained by delays in the stochastic model that are not represented in the fluid approximation. Also shown in Figure 8.2 is the evolution of occupancy in a stochastic model when each occupant goes to its nearest exit. More simulation details are contained in Section 8.4.1.

8.1.2 Performance analysis

The main result of [26] is applied in Proposition 8.2.1 to show that the mean evacuation time for the stochastic model is bounded below by the corresponding quantity for the fluid model. The minimal evacuation time for the fluid model is obtained by solving a linear program, and can be expressed as a piecewise linear function of the initial condition through associated workload vectors,

$$T^*(x) = \max_{1 \le i \le n_r} \langle \xi^i, x \rangle, \tag{8.1}$$

where x is a vector whose entries equal occupancy at the zones in the building, $\{\xi^i, 1 \le i \le n_r\}$ are workload vectors (obtained in the dual of the linear program for the fluid model which is described in Section 8.2.3), and $\langle \cdot, \cdot \rangle$ denotes the vector inner product. We find that each workload vector defines a partition of the building graph: a region N_+ corresponding to zones where the workload vector entry is non-zero, and its complement N_- where entries are zero.

If i^* is a unique maximizer in (8.1), then the corresponding region in the building $N_+ = \{k : \xi_k^{i^*} > 0\}$ is congested in the sense that the evacuation time is sensitive to the total occupancy in this region, while sensitivity to occupancy in the region N_- is zero. This intuition is used to guide the construction of evacuation policies that concentrate on driving agents from congested regions towards non-congested regions, and finally the exits.

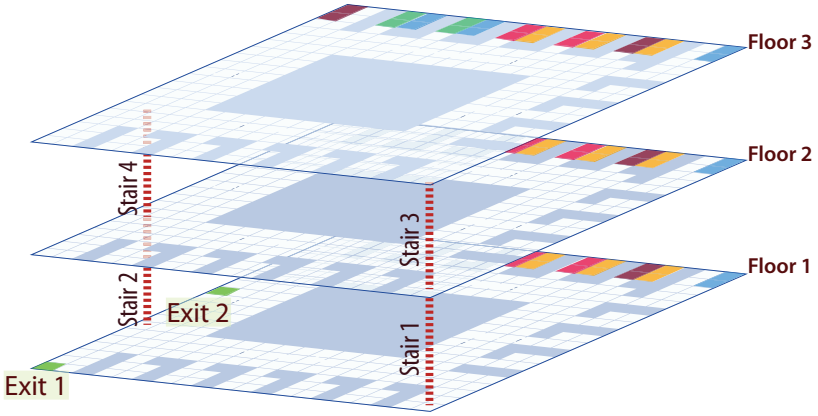


Figure 8.3: Layout of a three-floor office building. If there are many occupants on the second floor, then optimal evacuation requires resource pooling at the two lower stairs.

A secondary value of workload is inspiration on the source of delayed evacuation. The dual of the linear program that determines $T^*(x)$ is based on two sets of variables — the workload vectors $\{\xi^i\}$, and also resource pooling variables. The latter represents the resources that must work at capacity during the entire evacuation process. For example, in a three-floor building schematic shown in Figure 8.3, for some initial conditions for occupancy it is necessary for stairs 1 and 2 to be fully used during the entire evacuation to ensure minimal evacuation time (see examples in Section 8.4.2).

8.1.3 Control of evacuation

The policies introduced here are based on the h-MaxWeight policy of [81], which is a generalization of the MaxWeight policy of Tassiulas and Ephremides [113]. For any stochastic queueing model, the h-MaxWeight policy is stabilizing and approximately average-cost optimal under appropriate conditions on the function h and the network [81]. The proof of approximate optimality is also based on the workload relaxation technique used in this chapter.

As seen in one set of simulations illustrated in Figure 8.2, the h-MaxWeight policy shows considerable reduction in evacuation time over the baseline where each occupant uses their nearest exit. The results also compare favorably against the lower bound on evacuation obtained using the workload analysis based on the fluid model.

8.1.4 Building re-design

The resource pooling variables associated with the workload vectors provide information regarding congested regions in the building. Application of this insight to building re-design is illustrated in Section 8.4.2. For example, the value of the stairwell denoted *Stair 4* in Figure 8.3 can be assessed through a workload analysis. For a general building topology, the re-design procedure can be carried out by iteratively analyzing dominant workload vectors under typical building conditions. At each step of iteration, additional resources are included that will best reduce the overall evacuation time. This procedure is illustrated with the aid of the three-floor building example.

The remainder of this chapter is organized as follows. In Section 8.2, we describe the queueing models and discuss the workload analysis. In Section 8.3, we use these models for the purposes of optimization using the h-MaxWeight policy based on workload relaxations of various dimensions. Section 8.4 contains results from simulation experiments for the single-floor building and the three-floor building, respectively.

8.2 Control Model for Occupancy

We begin with a focus on control. Consider the three-floor building whose floor plan is illustrated in Figure 8.3. The goal of control is to evacuate the building in minimal time. This is achieved by sending instructions to occupants through monitors and/or audio instructions. The evolution of agents in the building is conceptualized as a multiclass queueing model.

8.2.1 Stochastic models

Initially we focus on just one source of variability – that originating from variability in walking speed. The model allows for multiple *classes* of agents, classified by walking speed and variability. One class may represent young office

workers, and another class may represent persons with disabilities, or children. To simplify discussion we restrict most of our attention to the simpler single-class setting in which differences in agent types are disregarded.

We suppose that the building plane is broken into n zones. We let $Q_i(t)$ denote the number of agents in zone i at time t. In addition there is a network structure described by an associated graph with $(n+1)$ nodes. Each zone defines a node in a graph, and the $(n+1)$th node represents the world outside of the building. There is a link between two nodes in the graph if the corresponding zones are connected by a passageway. In modeling a stairwell, the stair region can form one zone *or* this can represent a passageway between two zones on the two connecting floors. This is a modeling choice that will depend on such factors as time scale, and the total building occupancy. The choice of granularity in the definition of zones depends on these factors, as well as computational convenience.

The Controlled Random Walk (CRW) model is defined in discrete time. We let $M_{ij}(t)$ denote the number of agents that move from node i to node j at time t, provided agents are directed to do so. For simplicity we take $M_{ij}(t)$ to be a Bernoulli random variable – at most one agent travels between neighboring nodes at each time step. The network structure implies that $M_{ij}(t) \equiv 0$ for i, j that are not directly connected by a link in the graph. The *direction* to agents is the control, which is modeled as a collection of binary variables $\{U_{ij}(t)\}$. If $U_{ij}(t) = 1$ then agents at node i are directed to node j; otherwise ($U_{ij}(t) = 0$) no agent at node i is directed to node j. Since we are modeling evacuation, it is assumed that there are no external arrivals to the building. Hence we are ignoring the possible ingress of emergency personnel.

Occupancy in the CRW model evolves as follows: Given the initial occupancy vector $Q(0) \in \mathbb{Z}_+^n$, we have for $t \geq 0$, and $1 \leq i \leq n$,

$$Q_i(t+1) = Q_i(t) + \sum_{k=1}^{n} U_{ki}(t) M_{ki}(t+1) - \sum_{j=1}^{n+1} U_{ij}(t) M_{ij}(t+1). \tag{8.2}$$

Note that if $j = n+1$, then the event $M_{ij}(t) = 1$ means that an agent can transition from node i to an exit at time t.

It is assumed throughout that the process $\boldsymbol{M} := \{M(t)\}_{t \geq 0}$ is i.i.d.. We do not specify the correlation structure of $\{M_{ij}(t) : 1 \leq i, j \leq n\}$.

The goal of the control problem considered here is to choose the sequence $\boldsymbol{U} := \{U(t)\}_{t \geq 0}$, based on causal observations of $\boldsymbol{Q} := \{Q(t)\}_{t \geq 0}$, to minimize the mean of the evacuation time defined by,

$$\mathcal{T}_0 = \min\{t \geq 1 : Q(t) = 0\}. \tag{8.3}$$

In general this optimal control problem is intractable for large queueing networks. Fortunately, a fluid model has an explicit solution, and provides intuitive and approximately optimal solutions for the CRW model.

8.2.2 Fluid models

The fluid model is motivated by considering the average behavior of the CRW model. On denoting $\bar{q}_i(t) = \mathsf{E}[Q_i(t)]$ and $\bar{u}_i(t) = \mathsf{E}[U_i(t)]$ we obtain under the i.i.d. assumption on $\boldsymbol{M}$,

$$\bar{q}_i(t+1) = \bar{q}_i(t) + \sum_{k=1}^{n} \bar{u}_{ki}(t)\mu_{ki} - \sum_{j=1}^{n+1} \bar{u}_{ij}(t)\mu_{ij}$$

where μ_{ij} is the mean of $M_{ij}(t)$, assumed independent of t. Letting $\bar{z}(t)$ denote the mean *cumulative* allocation process,

$$\bar{z}_{ij}(t) = \sum_{s=0}^{t-1} \bar{u}_{ij}(s), \qquad t \geq 1$$

we obtain the alternate representation that includes explicitly the initial condition,

$$\bar{q}_i(t) = \bar{q}_i(0) + \sum_{k=1}^{n} \bar{z}_{ki}(t)\mu_{ki} - \sum_{j=1}^{n+1} \bar{z}_{ij}(t)\mu_{ij}.$$

The fluid model obeys identical equations in continuous time. Its state process $\boldsymbol{q} := \{q(t)\}_{t\geq 0}$ evolves on $\mathbb{R}_+^n$ along with a cumulative allocation process $\boldsymbol{z} := \{z(t)\}_{t\geq 0}$. For a given initial condition $q(0) \in \mathbb{R}_+^n$, the occupancy vector at time t satisfies

$$q_i(t) = q_i(0) + \sum_{k=1}^{n} z_{ki}(t)\mu_{ki} - \sum_{j=1}^{n+1} z_{ij}(t)\mu_{ij}. \tag{8.4}$$

We adopt the following more compact notation,

$$q(t) = q(0) + Bz(t) \tag{8.5}$$

in which the matrix B is defined consistently with (8.4), and $z(t)$ is the vector of cumulative allocations. The dimension of $z(t)$ coincides with the number of links in the graph, which we denote n_l. The matrix B is thus an $n \times n_l$ matrix, whose $i,(kj)$ entry is the difference $\mu_{ki}\mathbb{1}\{i = j\} - \mu_{ij}\mathbb{1}\{i = k\}$.

The continuous time process $\boldsymbol{z}$ is interpreted as the control for the fluid model. It is subject to numerous constraints. First, there are indirect constraints due to the fact that $\boldsymbol{q}$ is constrained. It is assumed that $q_i(t)$ is non-negative for each i and t. We may also impose upper bounds on each of these occupancy variables. The set of feasible values for $q(t)$ is denoted X, assumed to be a convex subset of $\mathbb{R}_+^n$ containing the origin. There are direct constraints modeled via a *constituency matrix* denoted C. These constraints are linear, of the following form: For each $0 \leq t_0 \leq t_1 < \infty$,

$$z(t_1) - z(t_0) \geq \mathsf{0}, \quad C(z(t_1) - z(t_0)) \leq (t_1 - t_0)\mathsf{1} \tag{8.6}$$

where $\mathsf{1}$ and $\mathsf{0}$ are column vectors of ones and zeros, respectively. In the special case $C = I$ we are simply echoing the previous constraint that $U_{ij}(t)$ is equal to zero or one. More generally, additional constraints may arise from the geometry of the building. We let n_m denote the number of rows in the matrix C.

The fluid model (8.5) is also expressed as the *controlled differential equation*,

$$\tfrac{d}{dt}q(t) = B\zeta(t), \qquad t \geq 0 \tag{8.7}$$

where $\zeta(t) = \frac{d}{dt}z(t)$ denotes the *allocation rate*. The constraints given in (8.6) imply that $\boldsymbol{\zeta} := \{\zeta(t)\}_{t\geq 0}$ is subject to polyhedral constraints of the form $\zeta(t) \in \mathsf{U}$, with

$$\mathsf{U} := \{\zeta \in \mathbb{R}_+^{n_l} : \zeta \geq \mathsf{0}, \quad C\zeta \leq \mathsf{1}\}.$$

We let $\mathsf{U}(x)$ denote the subset of values for which constraints on $\boldsymbol{q}$ are not violated. For example, if only non-negativity constraints are imposed, then we have

$$\mathsf{U}(x) := \{\zeta \in \mathsf{U} : (B\zeta)_i \geq 0, \text{ when } x_i = 0\}. \tag{8.8}$$

A third representation of the fluid model dynamics are through its *velocity* $v(t) := \frac{d}{dt}q(t)$. On denoting the *velocity space* by

$$\mathsf{V} := \{v \in \mathbb{R}^n : v = B\zeta, \quad \zeta \in \mathsf{U}\}$$

the fluid model is expressed as the *differential inclusion*,

$$\tfrac{d}{dt}q(t) \in \mathsf{V}, \quad q(t) \in \mathsf{X}, \qquad t \geq 0.$$

We assume throughout the chapter that the fluid model is *stabilizable*: For each $x \in \mathsf{X}$, there exists a feasible pair $(\boldsymbol{q}, \boldsymbol{z})$ and time $T < \infty$ such that $q(t) = 0$ for $t \geq T$.

Since V contains the origin, it follows that it can be expressed as intersection of half spaces. Proposition 6.1.3 of [80] gives the particular form,

$$\mathsf{V} = \{v \in \mathbb{R}^n : \langle \xi^i, v\rangle \geq -o_i, \quad 1 \leq i \leq n_v\} \tag{8.9}$$

where the constraints $\{o_i, 1 \leq i \leq n_v\}$ take binary values, and $\xi^i \in \mathbb{R}^n$ for each i. The set $\{\langle \xi^i, v\rangle = -o_i\}$ is a face of the polyhedron V. If $o_i = 0$, the face passes through the origin; If $o_i = 1$, the corresponding vector ξ^i is called the

workload vector. The number of distinct workload vectors is denoted by n_r. By reordering, we assume that $o_i = 1$ for $1 \le i \le n_r$.

8.2.3 Workload and control

Given the initial condition $q(0) = x$, the *evacuation time* $T(x)$ is defined as the first time the the total occupancy reaches zero in the fluid model. This of course depends on the choice of $\boldsymbol{z}$. We denote by $T^*(x)$ the *minimal* evacuation time, over all feasible $\boldsymbol{z}$. This can be expressed as the solution to a linear program:

$$\begin{aligned} T^*(x) = \min\ & T \\ \text{s.t. } & x + Bz = 0 \\ & Cz \le T\mathsf{1},\ z \ge 0,\ T \ge 0. \end{aligned} \tag{8.10}$$

On considering the dual of (8.10) we obtain the representation (8.1) for $T^*(x)$ (see eq. (8.12) below and Proposition 6.1.5 of [80]).

Observe that state space constraints are ignored in this linear program. This is justified by convexity of the model: If $\boldsymbol{q}$ is any feasible trajectory which reaches zero at time $T(x) < \infty$, then we define $\bar{\zeta} = z(T(x))/T(x)$, along with the linear trajectory,

$$\tfrac{d}{dt}\bar{q}(t) = B\bar{\zeta}, \quad 0 \le t < T(x), \quad \bar{q}(0) = q(0) = x.$$

It is easy to see that for each i,

$$\bar{q}_i(t) = \bar{q}_i(0)\frac{T(x) - t}{T(x)}, \qquad 0 \le t < T(x).$$

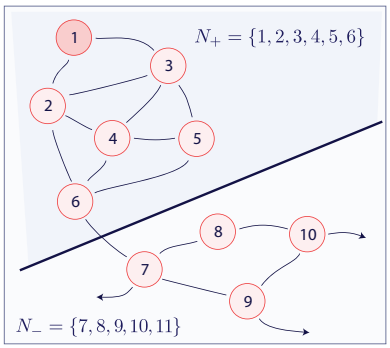
Since X is convex and contains the origin, we must have $\bar{q}(t) \in \mathsf{X}$ for all t, and evidently $\bar{q}(t) \to 0$ as $t \uparrow T(x)$.

The optimal evacuation time $T^*(x)$ for the fluid model also provides a lower bound for the mean evacuation time for the CRW model. Recall that the evacuation time $\mathcal{T}_0$ is defined in (8.3). The proof of the following proposition appears in Appendix D.1.

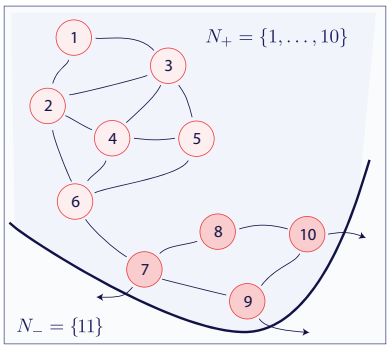
Proposition 8.2.1 *For any initial condition $Q(0) = x$, and any feasible allocation sequence for the CRW model (8.2), the following lower bound holds for the mean evacuation time:*

$$\mathsf{E}[\mathcal{T}_0] \ge T^*(x). \tag{8.11}$$

Resource pooling There are in fact two sets of dual variables associated with the linear program (8.10): Variables corresponding to the equality constraint $x + Bz = 0$, denoted $\xi \in \mathbb{R}^n$, and variables corresponding to the inequality



(a) $q(0) = (100, 0, \ldots, 0)^T$



(b) $q(0) = (0, \ldots, 0, 15, 20, 10, 10)^T$

Figure 8.4: The congested region N_+, the non-congested region N_- and the cut R_* for a 10-node building model with three exits (nodes $7, 9, 10$). The node 11 denotes the world outside of the building, which is not depicted in the figure. (a) For initial condition $x = q(0) = (100, 0, \ldots, 0)^T$, i.e., the agents start in node 1, and (b) For initial condition $x = q(0) = (0, \ldots, 0, 15, 20, 10, 10)^T$, i.e., the agents start in nodes $7, 8, 9, 10$.

constraint $Cz \le T\mathsf{1}$, which are denoted ν. The latter variables have non-negative entries $\nu \in \mathbb{R}_+^{n_m}$, where n_m denotes the number of rows in the matrix C.

The dual of (8.10) can be expressed as follows as a maximization over the variables (ξ, ν):

$$\begin{aligned} T^*(x) &= \max \langle \xi, x \rangle \\ \text{s.t.} \quad & -B^T\xi - C^T\nu \le \mathsf{0} \\ & \mathsf{1}^T\nu \le 1, \ \nu \ge \mathsf{0}. \end{aligned} \tag{8.12}$$

If (ξ^*, ν^*) is an optimizer then we do have $T^*(x) = \langle \xi^*, x \rangle$. If the maximizer (ξ^*, ν^*) is unique, then ξ_j^* is the sensitivity of the minimal evacuation time with respect to occupancy at zone j.

Complementary slackness of (8.12) implies that $\sum_{i=1}^{n_m} \nu_i^* = 1$, so that ν^* can be interpreted as a probability on the *resources* represented by the rows of C (see Proposition 6.1.7 of [80]). Complementary slackness of (8.10) also implies the following necessary condition for time-optimality of an allocation rate ζ:

$$\big(C\zeta(t)\big)_i = 1, \ \text{ whenever } \nu_i^* > 0, \qquad 0 \le t < T^*(x).$$

Following the terminology in the stochastic networks literature, we say that there is *resource pooling* among resources $R_* := \{j : \nu_j^* > 0\}$ [50, 80].

Analogous to concepts from the graph theory [8, 80], the set R_* is interpreted as a *cut* on the building graph, separating a subset of nodes in the graph and the exit node.

Figure 8.4 uses a simple graph network model to illustrate the concepts introduced here. The region $N_+ := \{j : \xi_j^* > 0\}$ corresponds to a congested region of the building, in the sense that the evacuation time is sensitive to

occupancy in this region, while sensitivity to the occupancy in the region $N_- := N_+^c$ is zero. In Figure 8.4 (a) the resource pooling vector ν^* will have an entry that is strictly positive – the entry index corresponding to the row of C that represents the link constraint between nodes 6 and 7. In the alternate situation illustrated in Figure 8.4 (b) there is resource pooling among the three exits shown.

8.3 Evacuation policies

The policies described here are of the state feedback form $U(t) = \phi(Q(t))$. Although $Q(t)$ is restricted to the integer lattice, we take the domain of ϕ to be the convex state space for the fluid model, X. The range of ϕ is $\mathsf{U}_\diamond$, the set of vectors $u \in \mathsf{U}$ with binary entries. For any $x \in \mathsf{X}$ we must also satisfy the constraint $\phi(x) \in \mathsf{U}_\diamond(x)$, where

$$\mathsf{U}_\diamond(x) := \{u \in \mathsf{U}_\diamond : (Bu)_i \geq 0, \text{ when } x_i = 0\}.$$

Our goal goes somewhat beyond minimizing the mean evacuation time. Observe that for the fluid model there are many time-optimal policies, and some policies are more attractive than others. Given this lack of uniqueness we seek to also minimize a certain *cost* on the occupancy distribution. We let $c(x)$ denote the cost associated with $x \in \mathsf{X}$. It is assumed to be linear, of the specific form $c(x) = c^{\scriptscriptstyle T} x$, with

$$c_i = \textit{distance to the nearest exit from node } i. \tag{8.13}$$

We may also take any monotone function of distance. The motivation for this cost function is clear: At a given time, if there are ten people left in the building, then the cost will be lowest if they are all near an exit.

We begin with some general techniques for policy synthesis.

8.3.1 Myopic policies

Given any function $h\colon \mathbb{R}_+^n \to \mathbb{R}_+$, one policy for the CRW model is the *h-myopic policy* defined as the minimum,

$$\phi(x) := \arg\min_{u \in \mathsf{U}_\diamond(x)} \mathsf{E}[h(Q(t+1) \mid Q(t) = x,\ U(t) = u]. \tag{8.14}$$

The motivation for these policies comes in part from the dynamic programming equations from optimal control. Consider the total cost optimality criterion,

$$h^*(x) := \min \mathsf{E}\Big[\sum_{t=0}^{\infty} c(Q(t)) \mid Q(0) = x\Big] \tag{8.15}$$

where the minimum is over all feasible allocation sequences $\boldsymbol{U}$. This value function satisfies the dynamic programming equation:

$$\min_{u\in\mathsf{U}_\diamond(x)} \mathsf{E}[h^*(Q(1)\mid Q(0)=x,\, U(0)=u] = h^*(x) - c(x)$$

and the h^*-myopic policy is optimal, in the sense that it achieves the minimum in (8.15).

Returning to (8.14), we assume throughout the chapter that the function h appearing in this chapter is convex, monotone, and vanishes only at the origin. For example in the special case where h is linear, say $h(x) = c(x) = c^{\scriptscriptstyle T}x$, the model description (8.2) combined with the definition of B in (8.5) gives,

$$\phi(x) = \arg\min_{u\in\mathsf{U}_\diamond(x)} \langle c, Bu\rangle, \qquad x\in\mathsf{X}.$$

This is similar to a myopic policy for the fluid model that we consider next.

Suppose that the function $h\colon \mathbb{R}_+^n \to \mathbb{R}_+$ is continuously differentiable (C^1). We then define the *h-myopic policy* for the fluid model (8.7) by

$$\phi^{\text{F}}(x) := \arg\min_{\zeta\in\mathsf{U}(x)} \langle \nabla h(x), B\zeta\rangle \tag{8.16}$$

where $\mathsf{U}(x)$ is defined in (8.8). The definition is analogous to the definition used for the CRW model: Under the h-myopic policy, for each t, the rate $\zeta(t)$ is chosen so that $\frac{d}{dt}h(q(t))$ is minimized.

One example of a convex, monotone function is the total cost for the fluid model,

$$J^*(x) := \min \int_{t=0}^{\infty} c(q(t))\, dt, \qquad q(0)=x \tag{8.17}$$

where the minimum is over all feasible $\boldsymbol{z}$. Analogous to the stochastic model, the principle of optimality holds: For each $T>0$,

$$\min\Big\{\int_0^T c(q(t))\, dt + J^*(q(T))\Big\} = J^*(x)$$

which implies a dynamic programming equation, provided J^* is smooth:

$$\min_{\zeta\in\mathsf{U}(x)} \langle \nabla J^*(x), B\zeta\rangle = -c(x).$$

Moreover, the J^*-myopic policy is optimal, achieving the minimal total cost in (8.17).

The optimality equations bring us to further motivation for consideration of the fluid model: It follows from the main result of [78] that the two value functions are similar for large x:

$$\frac{h^*(x)}{J^*(x)} \to 1, \qquad \text{as } \|x\|\to\infty. \tag{8.18}$$

The function J^* is piecewise quadratic [81], and we will see that approximations are easily constructed.

8.3.2 MaxWeight policies

The h-myopic policy is stabilizing for the fluid model under the assumptions on h imposed here [80]. Results for the CRW model are far less positive. Performance may be very poor in the CRW model under an h-myopic policy due to *starvation of resources* resulting from the more highly constrained minimum in (8.14). In particular, consider the network shown in Figure 8.4 (a) in which all of the occupants are initially at node 1. In an h-myopic policy or a time-optimal policy for the fluid model, it is possible that $q_i(t) = 0$ for all t and all $i \neq 1$. In the CRW model the situation is very different: It is essential that occupants move to node 6 so that capacity is not wasted on the link connecting nodes 6 and 7.

To avoid starvation we impose further structure on the function h. Assume that for any $x \in \mathbb{R}^n_+$,

$$\frac{\partial}{\partial x_i} h\,(x) = 0, \quad \text{whenever } x_i = 0. \tag{8.19}$$

This means that the *marginal disutility* at node i vanishes as the occupancy population reaches zero. It is shown in [80, 81] that under this condition, an h-myopic policy for the fluid model is *feasible* for the CRW model, in the sense that the minimum in (8.16) can be taken over the restricted set $\mathsf{U}_\diamond(x)$. Based on this simple observation, remarkably strong performance bounds are obtained in [81] for a general class of scheduling models.

Returning to Figure 8.4 (a), if nodes 2 and 3 are vacant, and $q_1(t) \sim 100$, then there will be a strong incentive to move agents to these two downstream nodes when (8.19) holds.

We can now define the *h-MaxWeight policy*: Assuming that h is C^1, convex, and monotone, and that the boundary conditions (8.19) hold, this is defined in analogy with the myopic policy (8.16) for the fluid model:

$$\phi^{\text{MW}}(x) := \arg\min_{u \in \mathsf{U}_\diamond(x)} \langle \nabla h(x), Bu \rangle. \tag{8.20}$$

A special case is a diagonal quadratic, $h(x) = \sum h_i x_i^2$, with $h_i > 0$ for each i. The resulting h-MaxWeight policy is precisely the MaxWeight policy of Tassiulas and Ephremides [113].

The boundary condition (8.19) is not satisfied for any quadratic functions other than diagonal quadratics, and of course (8.19) cannot hold when h is linear. To construct a large class of functions we apply the following perturbation: Suppose we are given a function $h_0 \colon \mathbb{R}^n_+ \to \mathbb{R}_+$ satisfying the general assumptions imposed earlier: h_0 is C^1, convex, monotone, and vanishes only at the origin. Suppose that for each i we have a function $\tilde{x}_i(x)$ that is also C^1, convex,

monotone, and vanishes only when $x_i = 0$. We then define $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)^T \in \mathbb{R}^n_+$, and

$$h(x) := h_0(\tilde{x}), \qquad x \in \mathbb{R}^n_+. \tag{8.21}$$

The chain rule then implies that (8.19) is satisfied, and clearly the function h satisfies all of the other required assumptions. Throughout the chapter we restrict to the following special case: For fixed $\theta > 0$ we take,

$$\tilde{x}_i = x_i \log(1 + \theta^{-1} x_i), \qquad 1 \le i \le n.$$

The gradient of h is then given by:

$$\nabla h(x) = L_\theta(x) \nabla h_0(\tilde{x})$$

where

$$L_\theta(x) = \operatorname{diag}\left((\theta + x_i)^{-1} x_i + \log\left(1 + \theta^{-1} x_i\right) \right).$$

The next section concerns the choice of h_0.

8.3.3 Workload relaxations

We wish to construct a function $h_0 : \mathbb{R}^n_+ \to \mathbb{R}_+$ that is convex, monotone and vanishing only at the origin. This is interpreted as an approximate value function for the associated optimal control problem for the CRW model. Insight regarding the choice of h_0 comes from consideration of the fluid model, motivated by the approximation (8.18). However, the choice $h_0 = J^*$ is not convenient because J^* may be complex, and it may not be C^1. Instead, we choose an approximation $\widehat{J}^*$ based on a relaxation of the fluid model dynamics.

The relaxation is based on the velocity space V defined in (8.9). Choose an integer $\hat{n}_r \le n_r$, with n_r denoting the number of distinct workload vectors, and define the relaxed velocity space by

$$\widehat{\mathsf{V}} := \{v \in \mathbb{R}^n : \langle \xi^i, v \rangle \ge -1, \quad 1 \le i \le \hat{n}_r\}.$$

The relaxation is defined as the differential inclusion

$$\tfrac{d}{dt} \hat{q}(t) \in \widehat{\mathsf{V}}, \quad \hat{q}(t) \in \mathsf{X}, \qquad t \ge 0. \tag{8.22}$$

The minimal evacuation time for the relaxation can be expressed as the lower bound on T^*,

$$\widehat{T}^*(x) = \max_{1\le i\le \hat{n}_r} \langle \xi^i, x\rangle. \tag{8.23}$$

Note that $\widehat{T}^*(x) \le T^*(x)$ for all $x \in \mathbb{R}^n_+$, but equality may hold for a large portion of the state space [80].

The associated fluid value function is denoted

$$\widehat{J}^*(x) := \min \int_{t=0}^{\infty} c(\hat{q}(t))\, dt, \qquad \hat{q}(0) = x \tag{8.24}$$

where the minimum is over all solutions to (8.22). Observe that although $\widehat{T}^*$ is a function of $x \in \mathbb{R}^n_+$, it can be expressed as a function of the $\hat{n}_r$ *workload values* $\{w_i = \langle \xi^i, x\rangle\}$. This is also true for the value function $\widehat{J}^*$, as seen through the following arguments.

In the relaxation, if the vector $x^+ \in \mathbb{R}^n_+$ satisfies $\langle \xi^i, x^+\rangle \ge \langle \xi^i, x\rangle$ for each $1 \le i \le \hat{n}_r$, then the time required to reach x^+ from $q(0) = x$ is *zero*. Consequently, $\widehat{J}^*(x)$ coincides with the minimum over all such x^+,

$$\widehat{\overline{J}}^*(w) = \min\big\{\widehat{J}^*(x^+) : \langle \xi^i, x^+\rangle \ge w_i,\ 1 \le i \le \hat{n}_r\big\}$$

where $w \in \mathbb{R}^{\hat{n}_r}$ is the vector with components $w_i = \langle \xi^i, x\rangle$. Moreover, we may regard the workload process as the control model, and define the control problem directly on $\boldsymbol{w}$ rather than $\boldsymbol{q}$. The dynamics are trivial: For each $t \ge 0$,

$$\tfrac{d}{dt}\hat{w}_i(t) \ge -1,\ \ 1 \le i \le \hat{n}_r, \qquad \hat{w}(t) \in \widehat{\mathsf{W}} \tag{8.25}$$

where $\widehat{\mathsf{W}}$ is a polyhedron arising due to the state space constraints:

$$\widehat{\mathsf{W}} := \big\{w : \textit{for some } x \in \mathsf{X},\ w_i = \langle \xi^i, x\rangle,\ \ 1 \le i \le \hat{n}_r\big\}.$$

To complete the description of the workload model we define an *effective cost* function $\bar{c}\colon \widehat{\mathsf{W}} \to \mathbb{R}_+$ as the solution to the linear program

$$\begin{aligned} \bar{c}(w) = \min\ & c(x) \\ \text{s.t. } & \widehat{\Xi} x = w,\ x \in \mathsf{X} \end{aligned} \tag{8.26}$$

where $\widehat{\Xi} = [\xi^1 \mid \xi^2 \mid \cdots \mid \xi^{\hat{n}_r}]^{\mathrm{T}}$. That is, $\bar{c}(w)$ is the cost associated with the "cheapest" state $x \in \mathsf{X}$ satisfying the given workload values. With this workload model and effective cost function, the value function (8.24) can be

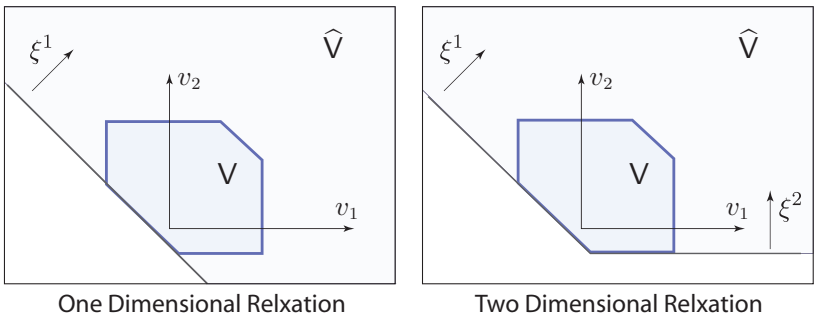


Figure 8.5: Relaxation of the velocity space V.

expressed as,

$$\widehat{J}^*(x) = \widehat{\overline{J}}^*(w) = \min \int_{t=0}^{\infty} \bar{c}(\hat{w}(t))\, dt, \quad \hat{w}(0) = w \tag{8.27}$$

where the minimum is over all solutions to (8.25).

Motivation for considering this relaxation is contained in [80]. The mathematical justification is through a *separation of time scales* that is made precise in the following special cases.

One-dimensional relaxation Consider the one-dimensional relaxation obtained with $\hat{n}_r = 1$, in which case $\widehat{\mathsf{V}}$ is the half-space,

$$\widehat{\mathsf{V}} = \{v \in \mathbb{R}^n : \langle \xi^1, v \rangle \geq -1\}.$$

The relationship between V and $\widehat{\mathsf{V}}$ is illustrated in the left of Figure 8.5.

If $\mathsf{X} = \mathbb{R}^n_+$, so that upper bounds on occupancy are ignored, then the effective cost $\bar{c}$ is linear, the value function $\widehat{\overline{J}}^*$ is a quadratic function of w, and have the following explicit expressions: On setting $\bar{c}^* = \min_i c_i/\xi^1_i$,

$$\bar{c}(w) = \bar{c}^* w, \qquad \widehat{\overline{J}}^*(w) = \tfrac{1}{2}\bar{c}^* w^2. \tag{8.28}$$

Suppose for example that ξ^1 corresponds to the network cut as shown in Figure 8.4 (a). Then $\xi^1_i = 1/\mu_{6,7}$ for $i \in N_+ = \{1, \ldots, 6\}$, and $\xi^1_i = 0$ for $i \in N_- = \{7, \ldots, 10\}$. Under the restriction (8.13) on the cost function, the parameter that defines the effective cost is $\bar{c}^* = c_6/\xi^1_1$. Moreover, the minimizing value of x in (8.26) is given by $x^*_6 = w/\xi^1_1$, and $x^*_i = 0$ for $i \neq 6$. That is, all agents crowd near the corridor connecting zones 6 and 7.

Two-dimensional relaxation Consider now a two dimensional relaxation for the ten-node building model shown in Figure 8.4, in which case $\widehat{\mathsf{V}}$ is an intersection of two half-spaces,

$$\widehat{\mathsf{V}} = \{v \in \mathbb{R}^n : \langle \xi^i, v \rangle \geq -1, \quad i = 1, 2\}.$$

The relationship between V and $\widehat{\mathsf{V}}$ is illustrated in the right of Figure 8.5. We take ξ^1 as described in one-dimensional relaxation, and ξ^2 as the workload vector corresponding to the network cut as shown in Figure 8.4 (b). We have $\xi^2_i = (\mu_{7,11} + \mu_{9,11} + \mu_{10,11})^{-1}$ for every $1 \le i \le 10$. The workload space is given by

$$\widehat{\mathsf{W}} = \{w : w_2/\xi^2_1 \ge w_1/\xi^1_1 \ge 0\}.$$

Assuming again that $\mathsf{X} = \mathbb{R}^n_+$, the solution x^* to (8.26) can be taken as a basic feasible solution, for which $x^*_i = 0$ for all but at most two values of i. If $w_1 > 0$, then one value i^*_1 will be in $N^1_+ = \{1, \dots, 6\}$, and from the foregoing we again find that $i^*_1 = 6$. The other value i^*_2 will be one of the values 7, 9, or 10, assuming c_i is constant for these three values of i. Let $i^*_2 = 7$, and consequently,

$$w_1 = x^*_6 \xi^1_1, \qquad w_2 = (x^*_6 + x^*_7)\xi^2_1$$

giving $x^*_6 = w_1/\xi^1_1$ and $x^*_7 = w_2/\xi^2_1 - w_1/\xi^1_1$. The effective cost is again a linear function of workload in this example,

$$\bar{c}(w) = c^{\scriptscriptstyle T} x^* = c_6 w_1/\xi^1_1 + c_7(w_2/\xi^2_1 - w_1/\xi^1_1).$$

As shown in Proposition 2.1 of [79], the value function $\widehat{\overline{J}}^*$ is a C^1 and piecewise-quadratic function of $w \in \widehat{\mathsf{W}}$. The detailed calculation of $\widehat{\overline{J}}^*$ appears in Appendix D.2.

The choice of h_0 Based on the preceding analysis, we might take $h_0 = \widehat{J}^*$, and then use the perturbation (8.21) to define h in the h-MaxWeight policy (8.20). However, in some cases this approximation is weak. For example, in the one-dimensional relaxation leading to the value function in (8.28), we have $\widehat{J}^*(x) = 0$ if $x_i = 0$ for $i \le 6$. To address this we introduce a penalty term that introduces a cost when $c(x)$ is far from $\bar{c}(\widehat{\Xi}x)$.

In this chapter, we consider the following class of functions based on the multi-dimensional relaxations (say we consider $\hat{n}_r$ distinct workload vectors) [81],

$$h_0(x) = \widehat{\overline{J}}^*(w) + \gamma\big[c(x) - \bar{c}(w)\big]^2 \tag{8.29}$$

where $\gamma > 0$ is constant, $\widehat{\overline{J}}^*(w)$ is the associated fluid value function defined in (8.27), $c(x)$ is the cost function, and $\bar{c}(w)$ with $w = \widehat{\Xi}x$ is the effective cost obtained from the linear program (8.26).

In general $\widehat{\overline{J}}^*(w)$ is a piecewise quadratic function of w (see workload relaxation examples above and in [79]). The exact expression of $\widehat{\overline{J}}^*(w)$ is usually difficult to obtain when considering a large number of workload vectors for

complex queueing networks [79]. The following function h_0 is taken as an approximation to (8.29) in this chapter:

$$h_0(x) = \widetilde{J}(w) + \gamma \left[c(x) - \bar{c}(w)\right]^2 \tag{8.30}$$

where $\widetilde{J}$ is a quadratic or a piecewise quadratic approximation of $\widehat{\overline{J}}^*$.

One heuristic for choosing $\widetilde{J}$ is through the construction of a sub-optimal value function for the relaxation. Consider the policy that drains the network linearly, so that $q(t)$ heads towards the origin along a straight line, and reaches the origin in minimal time. The draining time for relaxation is $\widehat{T}^*(x)$ obtained from (8.23), so we have,

$$\hat{q}(t) = \hat{q}(0)\frac{\widehat{T}^*(x) - t}{\widehat{T}^*(x)}, \qquad 0 \le t \le \widehat{T}^*(x)$$

when $\hat{q}(0) = x$. For this policy, the value function for the relaxation is given by

$$\widetilde{J}(w) = \tfrac{1}{2}\widehat{T}^*(x)\bar{c}(w) \tag{8.31}$$

where $w = \widehat{\Xi}x$, $\widehat{T}^*(x)$ is also a function of w, and $\bar{c}(w)$ is the effective cost obtained from the the linear program (8.26).

The function $\widetilde{J}(w)$ is a monotone, quadratic function of $w_i = \langle \xi^i, x \rangle$ for $i = 1, \ldots, \hat{n}_r$. Thus, it is a good heuristic to construct the function h_0. Moreover, the function $\widetilde{J}$ given in (8.31) is easy to compute: We only need to use (8.23) to obtain $\widehat{T}^*(x)$, and evaluate the linear program (8.26) to obtain $\bar{c}(w)$ at every time step.

8.4 Numerical Results

We first describe the occupancy model used in the simulation of egress from a building. We then consider several simulation experiments with a single-floor building and a three-floor building. The layout of these buildings appear in Figure 8.1 and Figure 8.3, respectively. We only consider one to three workload vectors for the building layouts considered here. With more complicated building layouts, one could expect greater improvement of evacuation performance by including more workload vectors.

8.4.1 Single-floor building

The single-floor building is simulated with an initial condition of 1160 agents, distributed in the building as illustrated in Figure 8.1. All agents are assumed to have homogeneous cooperative behavior: At each time, an agent has only one choice from four possible routes. The routing matrix is given by $B = [B^{(0)} \mid B^{(1)} \mid B^{(2)} \mid B^{(3)}]$, where $B^{(m)}$

for $m = 1, 2, 3$ is obtained by constructing a route out of the building via Exit-m (see Figure 8.1). The matrix $B^{(0)}$ is the routing matrix whereby an agent chooses the nearest exit route. Since the agent can only pick one route from four possible ones, this induces the constraints for the control (allocation) sequence. The constraints is modeled by taking the constituency matrix as $C = [I \mid I \mid I \mid I]$, where I is the identity matrix.

The statistics of the CRW model is chosen as follows: For the mth constructed route with $m = 0, 1, 2, 3$, the transition probability of an agent moving from ith node to jth node is taken as $\mu_{ij} = 1/2$ if $i \to j$ is along the route towards the exit, and $\mu_{ij} = 1/4$ otherwise. To remove the uncertainty of the simulation results from the CRM model, twenty simulation runs are used to obtain the average value of evacuation time in the following.

By solving the linear program (8.12), we obtain the minimal evacuation time $T^* = 386$ for the deterministic fluid model and the workload vector $\{\xi\}$. The cut $R_* = \{i : \nu_i^* > 0\}$ corresponds to pooling of three building exits 1, 2, and 3, which are the main resources for evacuating the building. The workload vector ξ is supported on the entire building, i.e. $\xi_i > 0$ for each i. Thus, by definition, the congested region N_+ includes all nodes inside the building. The non-congested region N_- is just the region outside the building. This single workload vector is used to obtain the h-MaxWeight evacuation policy for simulations.

Comparison of evacuation policies

(i) *NearestExit policy*. With this policy there is no feedback control and each agent leaves the building via the nearest exit (the routing matrix $B^{(0)}$ is chosen for each agent). This policy performs poorly because it does not efficiently utilize all available resources. After a short transient of relatively rapid decrease in the number of agents, both Exit-1 and Exit-3 become congested, as do the left and bottom corridors. In contrast, the top corridor and Exit-2 are never used. Figure 8.6 depicts the *flow direction* of the agent movement. The flow direction plot is obtained by drawing a directed arrow between neighboring nodes with most frequent single-hops over the simulation run. The average evacuation time using this policy is $T = 1112$ which is almost three times the lower bound $T^*(x) = 386$ obtained for the fluid model (see Figure 8.2).

These results provide a motivation for the use of optimization and feedback control for the purposes of better utilization of resources in the building.

(ii) *h-MaxWeight policy*. The simulation results are based on the quadratic function h_0 defined using the single workload vector $\{\xi\}$. The h-MaxWeight policy is implemented according to (8.29). The corresponding resource pooling vector ν is only supported on three exits which define the resource pooling R_*. It captures the intuition that three exits must be utilized to facilitate the evacuation.

Figure 8.7 depicts the flow direction of the agent movement obtained using the h-MaxWeight policy. At the beginning of the simulation, the agents starting in the left region of the building move towards Exit-1 while the remaining agents move towards Exit-3. This is just like the NearestExit policy as shown in Figure 8.6. As the

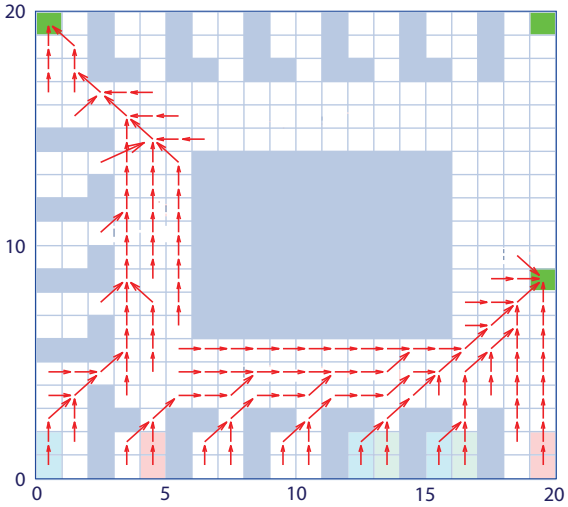


Figure 8.6: Flow direction of agent movement with the NearestExit Policy.

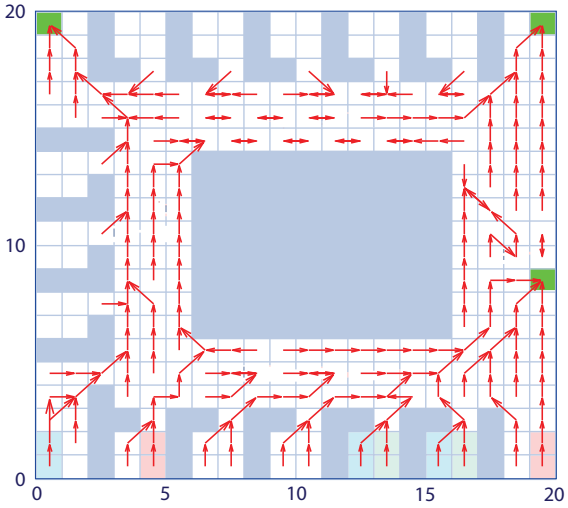


Figure 8.7: Flow direction of agent movement with the h-MaxWeight policy.

simulation proceeds under the h-MaxWeight policy, agents begin to move towards Exit-2 as the first and third exits become congested. Thus the utilization of resources is better than the NearestExit policy. The average evacuation time using the h-MaxWeight policy is $T = 528$ which compares favorably to the NearestExit policy (see Figure 8.2).

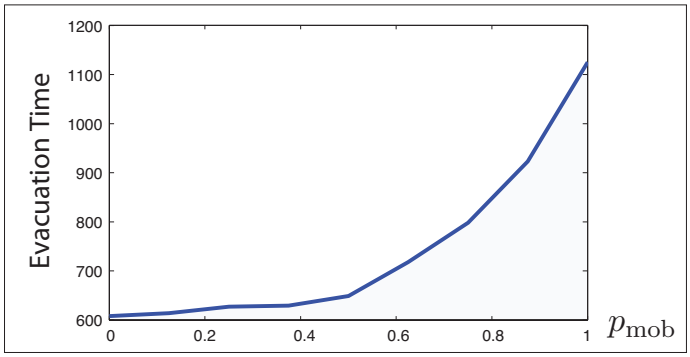


Figure 8.8: Average evacuation time as a function of the fraction of mob agents p_{mob}.

Non-cooperative behavior To test the robustness of the h-MaxWeight policy, we carried out simulations where some of the agents do not obey the h-MaxWeight policy. For each simulation, the agents were divided into two groups: *sheep agents* who move according to the h-MaxWeight policy, and *mob agents* who move according to the selfish NearestExit policy.

For each of the simulation runs, the initial occupancy is the same as in the previous simulation (see Figure 8.1). The parameter $p_{\text{mob}} \in [0, 1]$ is used to specify the fraction of mob agents. If $p_{\text{mob}} = 0$, there are only sheep agents and one recovers the h-MaxWeight case illustrated in Figure 8.7; If $p_{\text{mob}} = 1$, all agents move according to NearestExit policy (see Figure 8.6). Figure 8.8 depicts the average evacuation time T as a function of the parameter p_{mob}. For each value of p_{mob}, twenty simulation runs were used to obtain the average value of the evacuation time. From Figure 8.8, we find that the average evacuation time increases gracefully as long as $p_{\text{mob}} < 0.5$, i.e., there are more sheep agents than mob agents. Figure 8.9 shows the occupancy evolution for three values of $p_{\text{mob}} = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$. For $p_{\text{mob}} = \frac{1}{4}$ (Figure 8.9 (a)), the sheep agents exit the building at an average rate greater than the mob agents. For $p_{\text{mob}} = \frac{3}{4}$ (Figure 8.9 (c)), the evacuation process suffers from the starvation of resources caused by the mob agents (the Exit-1 and Exit-3 are effectively occupied by mob agents). For $p_{\text{mob}} = \frac{1}{2}$ (Figure 8.9 (b)), both sheep and mob agents evacuate at approximately the same rate.

8.4.2 Three-floor building

In this section we present the simulation results for a three-floor building whose schematic is shown in Figure 8.3. Each floor in the building has the same layout as the single-floor building shown in Figure 8.1. There are two exits

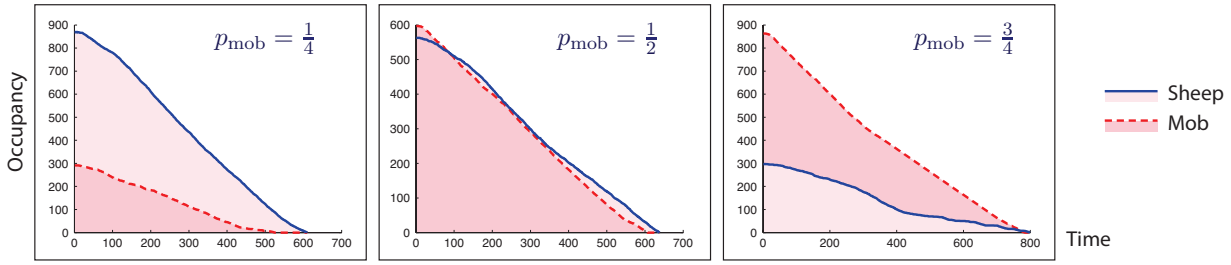


Figure 8.9: Occupancy evolution plots, where a fraction of agents (P_{mob}) uses the selfish NearestExit policy.

out of the building, both of which are on the first floor and labeled as Exit-1 and Exit-2. The building topology is such that an agent must take one of three stairs to move between floors: Stair-1 and Stair-2 are between the first and second floors, Stair-3 is between the second and third floors. Each stair comprises ten steps and is modeled as 10 queues in the CRW model. Note that the Stair-4 is not available here but it will be considered later in the *building re-design* section as a new resource.

The routing matrix is given by $B = [B^{(0)} \mid B^{(1)} \mid \cdots \mid B^{(4)}]$, where $B^{(m)}, m = 1, 2, 3, 4$ is obtained by constructing the m^{th} possible exiting route. For example, $B^{(1)}$ is related to the route: Stair-3→Stair-1→Exit-1. The matrix $B^{(0)}$ is the routing matrix whereby an agent always chooses the nearest exit on each floor. At each time, an agent has only one choice of five routes. Accordingly, the constituency matrix is given by $C = [I \mid I \mid \cdots \mid I]$. The statistics and simulation setups of the CRW model is the same as what we described for the single-floor building.

The simulation is initialized with 232 agents in the offices in each floor, just as in the single-floor building example. Initially, there are no agents on the stairs. By solving the linear program (8.10), we obtain the minimal evacuation time $T^* = 348$ for the fluid model. The three workload vectors are denoted by $\{\xi^1, \xi^2, \xi^3\}$, where ξ^1 is supported (i.e., $\xi^1_i > 0$) on the entire building, ξ^2 is supported on the second and third floors, and ξ^3 is supported on the third floor. The cuts for the three dominant workload are: R^1_* corresponds to pooling of two building exits 1 and 2, R^2_* corresponds to pooling of stairways 1 and 2, and R^3_* corresponds to pooling of stairway 3. These three workload vectors are used to construct the h-MaxWeight policy (as described in Section 8.3).

Comparison of evacuation policies Figure 8.10 depicts a comparison of the performance with the NearestExit policy, the h-MaxWeight policy (8.29) using $\{\xi^1\}$, and the h-MaxWeight policy (8.29) using $\{\xi^1, \xi^2, \xi^3\}$. As in the single-floor example, the performance with the NearestExit policy is poor due to the inefficient utilization of available resources. The two MaxWeight policies utilize the resources better, leading to better performance.

In the simulations, we find that incorporation of more workload vectors (i.e. $\{\xi^2, \xi^3\}$) alleviates the inefficient circulation of agent movement. The cut R^2_* divides the building into two regions: $N^2_+ = \{j : \xi^2_j > 0\}$ includes all nodes of the second and third floors and $N^2_- = \{j : \xi^2_j = 0\}$ includes the nodes of the first floor. Incorporation of the second workload vector $\{\xi^2\}$ into the h-MaxWeight policy directs agents across the cut R^2_* from N^2_+ to N^2_-.

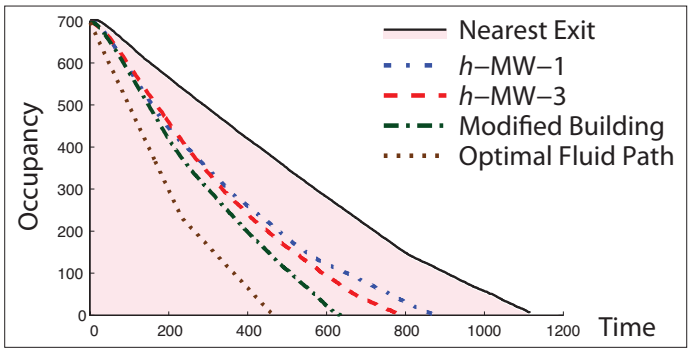


Figure 8.10: Comparison of average evacuation time obtained for the CRW model, using different polices for the three-floor building. Also shown is an optimal trajectory for the deterministic fluid model.

Incorporation of the third workload vector $\{\xi^3\}$ further improves the evacuation performance.

The average evacuation times are given as follows: NearestExit policy: $T = 1116$, h-MaxWeight policy with ξ^1 alone: $T = 877$, and h-MaxWeight policy with $\{\xi^1, \xi^2, \xi^3\}$: $T = 790$ (see Figure 8.10).

Building re-design Resource pooling vectors ν^i associated with the dominant workload vector ξ^i provide congestion information. This information can be used for building re-design, in which additional resources (stairs, exits or corridors) are added for faster evacuation. We illustrate re-design with the aid of the three-floor building example.

For the three-floor building, the workload analysis gives three dominant workload vectors $\{\xi^1, \xi^2, \xi^3\}$ with associated resource pooling vectors $\{\nu^1, \nu^2, \nu^3\}$. The corresponding cuts are denoted by $\{R^1_*, R^2_*, R^3_*\}$. Using the h-MaxWeight policy, one can evacuate at most 2 agents per unit time-step with $\{\xi^1\}$ or $\{\xi^2\}$, and at most 1 agent per unit time-step with $\{\xi^3\}$. Thus, the maximal flow for the cut R^3_* is half of that for R^1_* or R^2_*. During simulation runs, this leads to a slower (by a factor of 2) flow out of the third floor even when the first two floors are relatively empty.

This suggests that evacuation time may be decreased by increasing the maximal flow on the cut R^3_*. This may be accomplished by either re-designing the Stair-3 or by adding another stair from the third floor to the second floor. For example, in Figure 8.3, we add a new resource *Stair*-4, connecting third and the second floor to the building. Then the serving rate for the cut R^3_* is doubled and is equal to that for R^1_* or R^2_*. Figure 8.10 depicts a comparison of the evacuation time obtained using h-MaxWeight policy with and without re-design. Three workload vectors are employed for the h-MaxWeight policy in both cases. The average evacuation time with re-design for the modified building is obtained as $T = 608$, while the average evacuation time for the original building is $T = 790$. Thus this simple building re-design leads to a significant reduction for the overall evacuation time.

For a general building topology, the re-design procedure may be carried out by iteratively analyzing the workload vectors. At each step of iteration, the goal is to increase the smallest serving rate amongst all cuts. Design modifications would involve adding additional resources (stairs, exits, or corridors) that serves to increase the evacuation rate.

Multiclass models In a fully controlled setting, the extension to multi-class models in which, for example, the young and the elderly are given different direction, extensions of the policy synthesis procedure described in this section is straightforward. In this case it is necessary to extend the model, introducing buffers at each zone to account for the different agent types. The cost function must be modified to take into account priorities for different agents. We leave this for the future study.

8.5 Conclusions

We have demonstrated that workload relaxation techniques from queueing theory can be applied to address performance bounds and optimization for the problem of building evacuation. Two models were considered: a stochastic queueing model based on the Controlled Random Walk model and a fluid model that describes the average behavior of the stochastic model.

The h-MaxWeight policy showed considerable improvement in evacuation time over the baseline where each agent uses their best route. The results also compared favorably against the lower bound based on the deterministic fluid model.

Part IV

Appendices

Appendix A

Proofs for Chapter 2

A.1 Proof of Theorem 2.3.1

The K-L divergence rate (2.5) is expressed as

$$\begin{aligned} R(P\|\widehat{Q}^{(\mu)}(\phi)) &= \sum_{i,j\in\mathcal{N}} \pi_i P_{ij} \left(\log P_{ij} - \log \frac{\mu_i}{\sum_{k\in\psi(j)} \mu_k} Q_{\phi(i)\phi(j)} \right) \\ &= T_\phi(P\|Q) - \sum_{i,j\in\mathcal{N}} \pi_i P_{ij} \log \frac{\mu_j}{\sum_{k\in\psi(j)} \mu_k} \\ &= T_\phi(P\|Q) - \sum_{j\in\mathcal{N}} \pi_j \log \frac{\mu_j}{\sum_{k\in\psi(j)} \mu_k} \end{aligned} \tag{A.1}$$

where the third equality follows because $\pi P = \pi$.

Consequently, minimization of $R(P\|\widehat{Q}^{(\mu)}(\phi))$ over $\mu \in \mathsf{P}(\mathcal{N})$ is equivalent to maximizing the second term on the right hand side of (A.1). Setting $l = \phi(j)$ in (A.1) and using the the *log-sum inequality* (see [24]), we have

$$\begin{aligned} R(P\|\widehat{Q}^{(\pi)}(\phi)) - R(P\|\widehat{Q}^{(\mu)}(\phi)) &= \sum_{l\in\mathcal{M}} \Big(\sum_{k\in\phi^{-1}(l)} \pi_k \Big) \log \frac{\sum_{k\in\phi^{-1}(l)} \pi_k}{\sum_{k\in\phi^{-1}(l)} \mu_k} - \sum_{j\in\mathcal{N}} \pi_j \log \frac{\pi_j}{\mu_j} \\ &\le \sum_{j\in\mathcal{N}} \pi_j \log \frac{\pi_j}{\mu_j} - \sum_{j\in\mathcal{N}} \pi_j \log \frac{\pi_j}{\mu_j} = 0 \end{aligned}$$

with equality if and only if there exist constants $\{K_l,\ l \in \mathcal{M}\}$ such that

$$\frac{\pi_j}{\mu_j} = K_l, \quad \forall j \in \phi^{-1}(l), \quad l \in \mathcal{M}.$$

This establishes (2.9). Thus, $R(P\|\widehat{Q}^{(\pi)}(\phi)) \le R(P\|\widehat{Q}^{(\mu)}(\phi))$ for all $\mu \in \mathsf{P}(\mathcal{N})$.

A.2 Proof of Theorem 2.4.1

Using (2.8), we know $R_\phi(P\|Q) = T_\phi(P\|Q) - S_\phi(\pi)$ is a convex function with respect to the entries of Q. We introduce the Lagrangian for the optimization problem (2.10)

$$L = T_\phi(P\|Q) - S_\phi(\pi) + \sum_{k\in\mathcal{M}} \eta_k (\sum_{l\in\mathcal{M}} Q_{kl} - 1)$$

where $\{\eta_k,\ k \in \mathcal{M}\}$ are Lagrange multipliers. On taking the derivative with respect to Q_{kl}, we have

$$\frac{\partial L}{\partial Q_{kl}} = -\frac{\sum_{i\in\phi^{-1}(k)} \sum_{j\in\phi^{-1}(l)} \pi_i P_{ij}}{Q_{kl}} + \eta_k.$$

Setting the right hand side of above equation equal to zero, we obtain

$$Q_{kl} = \frac{\sum_{i\in\phi^{-1}(k)} \sum_{j\in\phi^{-1}(l)} \pi_i P_{ij}}{\eta_k}, \quad k, l \in \mathcal{M}. \tag{A.2}$$

The Lagrange multipliers $\{\eta_k,\ k \in \mathcal{M}\}$ are obtained by using the constraints

$$1 = \sum_{l\in\mathcal{M}} Q_{kl} = \frac{\sum_{i\in\phi^{-1}(k)} \pi_i}{\eta_k} \sum_{j\in\mathcal{N}} P_{ij}$$

where we used the fact that $\phi : \mathcal{N} \mapsto \mathcal{M}$ is a surjective function. Since P is a stochastic matrix, we have

$$1 = \frac{\sum_{i\in\phi^{-1}(k)} \pi_i}{\eta_k} \quad \Rightarrow \quad \eta_k = \sum_{i\in\phi^{-1}(k)} \pi_i. \tag{A.3}$$

Substituting (A.3) into (A.2), we obtain the optimal matrix $Q(\phi)$ shown in (2.11).

The formulae for the invariant distributions of $Q(\phi)$ and $\widehat{Q}^{(\pi)}(\phi)$ can be verified using a straightforward calculation.

A.3 Proof of Lemma 2.4.2

Using (2.8), we have

$$R_\phi(P\|Q) = T_\phi(P\|Q) - S_\phi(\pi). \tag{A.4}$$

In the following, we simplify the two terms on the right hand side separately.

For the first term,

$$\begin{aligned} T_\phi(P\|Q) &= \sum_{i,j\in\mathcal{N}} \pi_i P_{ij} \log P_{ij} - \sum_{i,j\in\mathcal{N}} \pi_i P_{ij} \log Q_{\phi(i)\phi(j)}(\phi) \\ &= -H(P) - \sum_{k,l\in\mathcal{M}} \sum_{i\in\phi^{-1}(k)} \sum_{j\in\phi^{-1}(l)} \pi_i P_{ij} \log Q_{kl}(\phi) \\ &= -H(P) - \sum_{k,l\in\mathcal{M}} \theta_k(\phi) Q_{kl}(\phi) \log Q_{kl}(\phi) \\ &= -H(P) + H(Q(\phi)). \end{aligned} \tag{A.5}$$

For the second term,

$$\begin{aligned} S_\phi(\pi) &= \sum_{j\in\mathcal{N}} \pi_j \log \frac{\pi_j}{\sum_{k\in\psi(j)} \pi_k} \\ &= -h(\pi) - \sum_{j\in\mathcal{N}} \pi_j \log \sum_{k\in\phi^{-1}(l)} \pi_k \\ &= -h(\pi) - \sum_{l\in\mathcal{M}} \theta_l(\phi) \log \theta_l(\phi) \\ &= -h(\pi) + h(\theta(\phi)). \end{aligned} \tag{A.6}$$

Substituting (A.5) and (A.6) into (A.4), we obtain (2.14).

Appendix B

Proofs for Chapter 3

B.1 Proof of Proposition 3.2.1

Using (3.1), we know that if v is an indicator on a subset (i.e., the first group of states) $\mathcal{N}_1 \subset \mathcal{N}$, then $\mathbb{1} - v$ is also an indicator but on the complementary subset $\mathcal{N}_2 = \mathcal{N}/\mathcal{N}_1$. Using (2.11), we have

$$Q(v) = \begin{bmatrix} \frac{v'\Pi P v}{v'\Pi v} & \frac{v'\Pi P(\mathbb{1}-v)}{v'\Pi v} \\ \frac{(\mathbb{1}-v)'\Pi P v}{(\mathbb{1}-v)'\Pi(\mathbb{1}-v)} & \frac{(\mathbb{1}-v)'\Pi P(\mathbb{1}-v)}{(\mathbb{1}-v)'\Pi(\mathbb{1}-v)} \end{bmatrix}.$$

Noting that $v'\Pi P(\mathbb{1}-v) = (\mathbb{1}-v)'\Pi P v = \beta(v) - \alpha(v)$ and $(\mathbb{1}-v)'\Pi(\mathbb{1}-v) = 1 - \beta(v)$, we obtain the formula (3.3) for $Q(v)$. The invariant distribution of Q directly follows from (6.13) in Theorem 2.4.1. The constraints (3.4) follows from the fact that $Q(v)$ is a stochastic matrix.

The first eigenvalue of Q is $\kappa_1 = 1$ due to the fact that Q is a stochastic matrix. The other eigenvalue can be directly computed by using the fact that the trace of a matrix is equal to the summation of all its eigenvalues.

B.2 Proof of Lemma 3.3.1

An application of the chain rule shows that:

$$\begin{aligned} \frac{dF}{dv}(v) &= \frac{\partial \widetilde{F}}{\partial \alpha}(\alpha,\beta)\frac{d\alpha}{dv}(v) + \frac{\partial \widetilde{F}}{\partial \beta}(\alpha,\beta)\frac{d\beta}{dv}(v) \\ &= -\left(\log \frac{(\beta-\alpha)^2}{\alpha(1-2\beta+\alpha)}\right)(\Pi P + P'\Pi)v + \left(\log \frac{(1-\beta)^2(\beta-\alpha)^2}{\beta^2(1-2\beta+\alpha)^2}\right)2\Pi v. \end{aligned} \tag{B.1}$$

The critical points are obtained by setting (B.1) equal to zero. We have the following two cases:

(i) If $\alpha = \beta^2$, then

$$\log \frac{(\beta-\alpha)^2}{\alpha(1-2\beta+\alpha)} = \log \frac{(1-\beta)^2(\beta-\alpha)^2}{\beta^2(1-2\beta+\alpha)^2} = 0$$

Thus, $\frac{dF}{dv}(v) = 0$.

(ii) If $\alpha \neq \beta^2$, then $\log \frac{(\beta-\alpha)^2}{\alpha(1-2\beta+\alpha)} \neq 0$ and

$$\lambda = \frac{\log \frac{(1-\beta)^2(\beta-\alpha)^2}{\beta^2(1-2\beta+\alpha)^2}}{\log \frac{(\beta-\alpha)^2}{\alpha(1-2\beta+\alpha)}}$$

is well defined. If the vector v is chosen to solve $\check{P}v = \lambda v$, we also have

$$\frac{dF}{dv}(v) = 2\left(\log \frac{(\beta-\alpha)^2}{\alpha(1-2\beta+\alpha)}\right)(\lambda \Pi v - \Pi \check{P} v) = 0.$$

B.3 Constraints (3.4) hold for all $v \in \mathbb{S}_\bullet$

By the generalized Rayleigh quotient, for $v \in \mathbb{S}$, we have

$$\frac{\alpha(v)}{\beta(v)} = \frac{v'\Pi\check{P}v}{v'\Pi v} \leq \lambda_2.$$

So, for $v \in \mathbb{S}_\bullet \subset \mathbb{S}$:

$$0 \leq \beta(v)^2 \leq \alpha(v) \leq \lambda_2 \beta(v) \leq \beta(v). \tag{B.2}$$

Next,

$$\beta(v) \leq \frac{1+\beta(v)^2}{2} \leq \frac{1+\alpha(v)}{2} \leq \frac{1+\beta(v)}{2} \leq 1 \tag{B.3}$$

where in the last inequality we used the bound,

$$\beta(v)^2 \leq \alpha(v) \leq \lambda_2 \beta(v) \quad \Rightarrow \quad \beta(v) \leq \lambda_2 < 1.$$

Combining (B.2) and (B.3), we have

$$0 \leq \alpha(v) \leq \beta(v) \leq \frac{1+\alpha(v)}{2} \leq 1.$$

Thus, the constraints (3.4) hold for all $v \in \mathbb{S}_\bullet$.

B.4 $F_\bullet(\lambda)$ is an increasing function of λ

By direct differentiation,

$$\frac{dF_\bullet}{d\lambda}(\lambda) = -B(\lambda)\log\frac{(1-\lambda)^2 B(\lambda)}{\lambda(1-2B(\lambda)+\lambda B(\lambda))}.$$

From the graph of $B(\lambda)$ (see Figure 3.1), we have $B(\lambda) \geq 0$ and $B(\lambda) \leq \lambda$. Then,

$$B(\lambda)(1-2\lambda+\lambda^2) \leq \lambda(1-2B(\lambda)+\lambda B(\lambda)). \tag{B.4}$$

Thus,

$$\log\frac{(1-\lambda)^2 B(\lambda)}{\lambda(1-2B(\lambda)+\lambda B(\lambda))} \leq 0.$$

Therefore, $\frac{dF_\bullet}{d\lambda}(\lambda) \geq 0$ where equality holds for $\lambda = 0$.

B.5 Proof of Proposition 3.3.2

At the spectral critical points, the function F is also a function of the non-negative eigenvalue λ. To see this, left-multiplying (3.12) by v' on both sides, we obtain (3.14). By substituting $\alpha = \lambda\beta$ into (3.13), we obtain the implicit relation (3.15) between λ and β. We then use the (3.15) to express β in terms of λ. For each fixed value of $\lambda \in [0,1]$, there are two possible implicit solutions: The first solution is $\beta = \lambda$. Together with (3.14), this implies that $\alpha = \lambda\beta = \beta^2$. At such points, $F(v) = 0$, the minimum value of F (see Proposition 3.3.5). The other solution is

$$\beta = B(\lambda) \tag{B.5}$$

where $B(\lambda)$ is a monotonically increasing function of λ (depicted in Figure 3.1), with $B(0) = 0$ and $B(1) = \frac{1}{2}$. For this solution,

$$\beta(v) = B(\lambda) \leq \lambda$$

where the equality holds only at $\lambda = 0$. Using (3.14), this also implies $\alpha(v) \geq \beta(v)^2$, i.e., $v \in \mathbb{S}_\bullet$.

In summary, the spectral critical points are given by $v = c_k u^{(k)} \in \mathbb{S}_\bullet$ with $\beta(v) = B(\lambda_k)$ and $\alpha(v) = \lambda_k B(\lambda_k)$ (from (3.14) and (B.5)). The constant c_k is obtained from

$$\beta(v) = v'\Pi v = c_k^2 (u^{(k)})'\Pi u^{(k)} = B(\lambda_k).$$

At such spectral critical points, both the function F and the matrix Q can be expressed as a function of the eigenvalue $\lambda \in \{\lambda_2, \ldots, \lambda_n\}$, i.e., $F(v) = F_\bullet(\lambda)$ (see (3.16)) and

$$Q(v) = \begin{bmatrix} \lambda & 1-\lambda \\ \frac{B(\lambda)-\lambda B(\lambda)}{1-B(\lambda)} & \frac{1-2B(\lambda)+\lambda B(\lambda)}{1-B(\lambda)} \end{bmatrix}. \tag{B.6}$$

Note that (B.6) is a Markov transition matrix because $\lambda \in [0,1]$ and $B(\lambda) \in [0, \frac{1}{2}]$. The eigenvalues are given by $\kappa_1 = 1$ and $\kappa_2 = \frac{\lambda - B(\lambda)}{1-B(\lambda)} \geq 0$.

B.6 Proof of Theorem 3.3.3

The proof of Theorem 3.3.3 is based on the analysis of eigenvalues of the Hessian matrix at the spectral critical points. A formula for these eigenvalues is given in the following Lemma. The proof of the Lemma appears at the end of this section.

Lemma B.6.1 *Suppose v is a spectral critical point of the function $F(v)$, i.e. it is the eigenvector of $\check{P}$ with the corresponding eigenvalue λ_v. We denote the Hessian matrix $W(v) := \frac{d^2F}{dv^2}(v)$. If u is the eigenvector of $\check{P}$ with the corresponding eigenvalue λ_u, then u also satisfies the following eigen-equation:*

$$W(v)u = \tilde{\lambda}_{uv} \Pi u \tag{B.7}$$

where the formula for $\tilde{\lambda}_{uv}$ is given by:

(i) If $u \in$ span$\{v\}$, then

$$\tilde{\lambda}_{uv} = -\frac{4(\lambda_v(\beta(v)+1) - 2\beta(v))}{(1-\beta(v))(1-2\beta(v)+\lambda_v\beta(v))}. \tag{B.8}$$

(ii) If $u \notin$ span$\{v\}$, then

$$\tilde{\lambda}_{uv} = 2(\lambda_v - \lambda_u) \log\left(\frac{(1-\lambda_v)^2\beta(v)}{\lambda_v(1-2\beta(v)+\lambda_v\beta(v))}\right). \tag{B.9}$$

Proof of Theorem 3.3.3: From Lemma B.6.1, we know that the Hessian matrix $W(v)$ shares the same eigenvectors as $\check{P}$. For the spectral critical point v, we use the notation $\tilde{\lambda}_{uv}$ to denote the eigenvalue of $W(v)$ (see (B.7)) and λ_u to denote the eigenvalue of $\check{P}$ (see (3.12)), both with the same eigenvector u. To obtain the decomposition (3.19) on the subspace $\mathbb{S}$, we only need to consider the signs of $\tilde{\lambda}_{uv}$ for different u. We have the following three cases:

(i) *If* $\lambda_u > \lambda_v$, *then* $\tilde{\lambda}_{uv} \geq 0$. This follows by using the formula (B.9) for $\tilde{\lambda}_{uv}$. At the spectral critical point v, $\beta(v) = B(\lambda_v)$. Using (B.4),

$$\log \frac{(1-\lambda_v)^2 B(\lambda_v)}{\lambda_v(1-2B(\lambda_v)+\lambda_v B(\lambda_v))} \leq 0$$

where equality holds only at $\lambda_v = 0$. Using (B.9), we obtain $\tilde{\lambda}_{uv} \geq 0$.

(ii) *If* $\lambda_u < \lambda_v$, *then* $\tilde{\lambda}_{uv} \leq 0$. The proof of this case is identical to Case (i).

(iii) If $\lambda_u = \lambda_v$, there are two cases to consider:

- *If* $u \notin$ *span*$\{v\}$, *then* $\tilde{\lambda}_{uv} = 0$ by using (B.9).
- *If* $u \in$ *span*$\{v\}$, *then* $\tilde{\lambda}_{uv} \leq 0$. The formula for $\tilde{\lambda}_{uv}$ is given by (B.8). For $\lambda_v \in [0,1]$, we have $\beta(v) = B(\lambda_v) \in [0, \frac{1}{2}]$ (see (B.5)). Then $(1-\beta(v)) \geq 0$, $(1-2\beta(v)+\lambda_v\beta(v)) \geq 0$, and $(\lambda_v(B(\lambda_v)+1) - 2B(\lambda_v)) \geq 0$. Thus, $\tilde{\lambda}_{uv} \leq 0$.

We thus arrive at the decomposition (3.19).

If we assume $\lambda_2 \in (\lambda_3, 1)$ is a simple eigenvalue, then $\Omega_I^{(2)} = \emptyset$ (see discussion after definition (3.18)). For $\tilde{u} \in \Omega_D^{(2)} = \mathbb{S}$, $\tilde{u}' W(v) \tilde{u} < 0$. Thus $v = c_2 u^{(2)}$ corresponds to a local maximum of the function F on the subspace $\mathbb{S}$. The decomposition (3.19) also implies that the spectral critical points are saddle points for $k \in \{3, \ldots, n\}$.

In the interior of the constraint set $\mathbb{S}_\bullet \subset \mathbb{S}$, a critical point necessarily is a spectral critical point $v = c_k u^{(k)}$ with the function value $F(v) = F_\bullet(\lambda_k)$ for $k = 2, 3, \ldots, n$. In the constrained set $\mathbb{S}_\bullet$, the maximum of $F_\bullet(\lambda)$ is obtained at the second largest eigenvalue $\lambda = \lambda_2 \geq 0$ (see Figure 3.1 and surrounding discussion). So, to conclude that the local maximizer $v = c_2 u^{(2)}$ is a global maximizer on the constraint set $\mathbb{S}_\bullet$, we only need to check the boundary of $\mathbb{S}_\bullet$. It is given by

$$\partial \mathbb{S}_\bullet := \{v \in \mathbb{S}_\bullet : \alpha(v) = \beta(v)^2\}.$$

For any $v \in \partial \mathbb{S}_\bullet$, we have $F(v) = 0$, the minimum value of the function F (see Proposition 3.3.5). Thus $F_\bullet(\lambda_2)$ is the global maximum of the function F on $\mathbb{S}_\bullet$. ■

Proof of Lemma B.6.1: By direct differentiation, the formula for the Hessian matrix is given by

$$\begin{aligned}\frac{d^2F}{dv^2}(v) = \frac{\partial \widetilde{F}}{\partial \alpha}(\alpha,\beta)\frac{d^2\alpha}{dv^2}(v) + \frac{\partial \widetilde{F}}{\partial \beta}(\alpha,\beta)\frac{d^2\beta}{dv^2}(v) + \frac{d\alpha}{dv}(v)\left(\frac{\partial^2 \widetilde{F}}{\partial \alpha^2}(\alpha,\beta)\frac{d\alpha}{dv}(v) + \frac{\partial^2 \widetilde{F}}{\partial \alpha \partial \beta}(\alpha,\beta)\frac{d\beta}{dv}(v)\right)' \\ + \frac{d\beta}{dv}(v)\left(\frac{\partial^2 \widetilde{F}}{\partial \beta \partial \alpha}(\alpha,\beta)\frac{d\alpha}{dv}(v) + \frac{\partial^2 \widetilde{F}}{\partial \beta^2}(\alpha,\beta)\frac{d\beta}{dv}(v)\right)'.\end{aligned}$$

Suppose v is a spectral critical point with the eigenvalue $\lambda = \lambda_v$, then using (3.12), we have

$$\frac{d\alpha}{dv}(v) = \lambda \frac{d\beta}{dv}(v). \tag{B.10}$$

By substituting (B.10), we obtain

$$\frac{d^2F}{dv^2} = \left(\frac{\partial \widetilde{F}}{\partial \alpha}\right)\frac{d^2\alpha}{dv^2} + \left(\frac{\partial \widetilde{F}}{\partial \beta}\right)\frac{d^2\beta}{dv^2} + \left(\frac{\partial^2 \widetilde{F}}{\partial \alpha^2}\lambda^2 + \frac{\partial^2 \widetilde{F}}{\partial \alpha \partial \beta}2\lambda + \frac{\partial^2 \widetilde{F}}{\partial \beta^2}\right)\frac{d\beta}{dv}\left(\frac{d\beta}{dv}\right)'. \tag{B.11}$$

By substituting (3.14), we can simplify the coefficient of the last term of (B.11) as

$$\frac{\partial^2 \widetilde{F}}{\partial \alpha^2}\lambda^2 + \frac{\partial^2 \widetilde{F}}{\partial \alpha \partial \beta}2\lambda + \frac{\partial^2 \widetilde{F}}{\partial \beta^2} = -\frac{\lambda(\beta(v)+1) - 2\beta(v)}{\beta(v)(1-\beta(v))(1-2\beta(v)+\lambda\beta(v))} \leq 0. \tag{B.12}$$

The nonpositivity of (B.12) is by the fact that $\lambda \in [0,1]$, $\beta(v) = B(\lambda) \in [0, 0.5]$ (see (B.5)), and $\lambda(\beta+1) - 2\beta = \lambda(B(\lambda)+1) - 2B(\lambda) \geq 0$.

Using (B.1) and (B.12), we obtain the Hessian matrix at the critical point v:

$$W(v) = -\frac{4(\lambda(\beta(v)+1) - 2\beta(v))}{\beta(v)(1-\beta(v))(1-2\beta(v)+\lambda\beta(v))}\Pi v v' \Pi - 2\log\left(\frac{(1-\lambda)^2\beta(v)}{\lambda(1-2\beta(v)+\lambda\beta(v))}\right)(\Pi \check{P} - \lambda\Pi).$$

To obtain the eigenvalues of $W(v)$, we first note that the vector v is an eigenvector of (3.12) with eigenvalue λ_v. Let u be an eigenvector of (3.12) with eigenvalue λ_u. We have two cases to consider:

If $u \notin \text{span}\{v\}$, then $v'\Pi u = 0$. So,

$$\begin{aligned} W(v)u &= 2\log\left(\frac{(1-\lambda_v)^2\beta(v)}{\lambda_v(1-2\beta(v)+\lambda_v\beta(v))}\right)\Pi(\lambda_v u - \check{P}u) \\ &= 2(\lambda_v - \lambda_u)\log\left(\frac{(1-\lambda_v)^2\beta(v)}{\lambda_v(1-2\beta(v)+\lambda_v\beta(v))}\right)\Pi u \end{aligned}$$

where we use the eigen-equation $\check{P}u = \lambda_u u$ in the last step.

If $u \in \text{span}\{v\}$, then $u = cv$ and $\check{P}u = \lambda_v u$. So,

$$\begin{aligned} W(v)u &= \frac{4(2\beta(v) - \lambda_v(\beta(v)+1))}{\beta(v)(1-\beta(v))(1-2\beta(v)+\lambda_v\beta(v))}\Pi v v' \Pi c v - 2\log\left(\frac{(1-\lambda_v)^2\beta(v)}{\lambda_v(1-2\beta(v)+\lambda_v\beta(v))}\right)\Pi(\check{P}u - \lambda_v u) \\ &= -\frac{4(\lambda_v(\beta(v)+1) - 2\beta(v))}{(1-\beta(v))(1-2\beta(v)+\lambda_v\beta(v))}\Pi u. \end{aligned}$$

■

B.7 Proof of Proposition 3.3.5

Using (3.6), we have

$$F(v) = h(\theta(v)) - H(Q(v)). \tag{B.13}$$

Denote $\alpha = \alpha(v)$ and $\beta = \beta(v)$. The entropy of aggregated invariant distribution

$$h(\theta(v)) = -\beta \log(\beta) - (1-\beta)\log(1-\beta) \leq \log(2) \tag{B.14}$$

where the upper bound is reached at $\beta = \frac{1}{2}$. The entropy rate of the aggregated Markov chain

$$-H(Q(v)) = -\beta h(\xi_1) - (1-\beta)h(\xi_2) \leq 0 \tag{B.15}$$

where ξ_1 and ξ_2 are probability distributions:

$$\xi_1 = [\frac{\alpha}{\beta}, \frac{\beta-\alpha}{\beta}], \quad \xi_2 = [\frac{\beta-\alpha}{1-\beta}, \frac{1-2\beta+\alpha}{1-\beta}].$$

The upper bound of $-H(Q(v))$ is reached at $\alpha = \beta$. Using (B.13)-(B.15), we obtain the upper bound of $F(v)$.

The lower bound of $F(v)$ is obtained by the application of the *log-sum inequality*:

$$-H(Q(v)) \geq \beta \log \beta + (1-\beta)\log(1-\beta) = -h(\theta(v))$$

where the equality holds if and only if $\alpha = \beta^2$.

The upper bound is reached for all v such that $\alpha(v) = \beta(v) = \frac{1}{2}$ (e.g., $v = \pm\frac{1}{\sqrt{2}}\mathbf{1}$). The lower bound is reached for all v such that $\alpha(v) = \beta(v)^2$ (e.g., $v = \mathbf{0}$ and $v = \mathbf{1}$).

B.8 Proof of Proposition 3.3.6

The following two lemmas are useful to obtain the bounds on mutual information given in Proposition 3.3.6. The first of these lemmas is borrowed form [58].

Lemma B.8.1 (Theorem 2 of [58]) *Suppose $\bar{v} \in \mathbb{T}$, then*

$$1 + \lambda_n \leq Q_{11}(\bar{v}) + Q_{22}(\bar{v}) \leq 1 + \lambda_2$$

where λ_2 and λ_n denote the second largest and the smallest eigenvalues of $\check{P}$, respectively.

The following lemma establishes the monotonicity properties of the functions $\widetilde{F}$ and $F_\circ$. The proof appears at the end of this section.

Lemma B.8.2 (i) *For any fixed $\beta_0 \in [0, \frac{1}{2}]$, the function $\widetilde{F}(\alpha, \beta_0)$ is a non-decreasing function of variable $\alpha \in [\beta_0^2, \beta_0]$.*

(ii) *For any fixed $\lambda_0 \in [0, 1]$, the function $F_\circ(\lambda_0, \beta)$ is a non-decreasing function of $\beta \in [0, \frac{1}{2}]$.*

Proof of Proposition 3.3.6: For $\bar{v} \in \mathbb{T}$, we have $\beta(\bar{v}) \in (0, \frac{1}{2}]$. The exact lower bound on β is given by

$$\beta(\bar{v}) = \sum_{i \in \mathcal{N}} \pi_i \bar{v}_i^2 \geq \min_{i \in \mathcal{N}} \pi_i > 0$$

where the equality holds with $\bar{v} = [0, \ldots, 0, 1, 0, \ldots, 0]$, where the 1 corresponds to the minimum π_i. Therefore, we have $\beta(\bar{v}) \in [\underline{\pi}, \frac{1}{2}]$.

Using Lemma B.8.1, we bound the trace of $Q(\bar{v})$

$$1 + \lambda_n \leq Q_{11}(\bar{v}) + Q_{22}(\bar{v}) \leq 1 + \lambda_2.$$

By using the formula for $Q(\bar{v})$, we have

$$\begin{aligned} 1 + \lambda_n &\leq \tfrac{\alpha}{\beta} + \tfrac{1 - 2\beta + \alpha}{1 - \beta} \leq 1 + \lambda_2 \\ \therefore, \quad \lambda_n \beta(1 - \beta) + \beta^2 &\leq \alpha \leq \lambda_2 \beta(1 - \beta) + \beta^2. \end{aligned} \tag{B.16}$$

For any fixed $\beta_0 \in [\underline{\pi}, \frac{1}{2}]$, we can write (B.16) as

$$\beta_0^2 \leq \lambda_n \beta_0(1 - \beta_0) + \beta_0^2 \leq \alpha \leq \lambda_2 \beta_0(1 - \beta_0) + \beta_0^2 \leq \beta_0.$$

Using Lemma B.8.2 (i), we then have

$$F_\circ(\lambda_n, \beta_0) \leq \widetilde{F}(\alpha, \beta_0) \leq F_\circ(\lambda_2, \beta_0). \tag{B.17}$$

Note that $F(\bar{v}) = \widetilde{F}(\alpha(\bar{v}), \beta(\bar{v}))$. The lower bound of $F(\bar{v})$ is obtained by minimizing the lower bound of (B.17) over all β_0. The upper bound of $F(\bar{v})$ is obtained by maximizing the upper bound of (B.17) over all β_0. Using Lemma B.8.2 (ii), we obtain the lower and the upper bound of $F(\bar{v})$ by setting $\beta_0 = \underline{\pi}$ and $\beta_0 = \frac{1}{2}$, respectively. ■

Proof of Lemma B.8.2: The monotonicity properties are established by explicitly evaluating the partial derivatives:

$$\frac{\partial \widetilde{F}}{\partial \alpha}(\alpha,\beta) = \log \frac{\alpha(1-2\beta+\alpha)}{(\beta-\alpha)^2}$$
$$\frac{\partial F_\circ}{\partial \beta}(\lambda,\beta) = 2\log\frac{(1-\lambda)(1-\beta)}{(1-(1-\lambda)\beta)} + (\lambda+2(1-\lambda)\beta)\log\frac{(1-(1-\lambda)\beta)(\lambda+(1-\lambda)\beta)}{\beta(1-\beta)(1-\lambda)^2}.$$

■

B.9 Proof of Proposition 3.3.7

The following lemma, adapted from [58], is used to obtain the bound on mutual information given in Proposition 3.3.7.

Lemma B.9.1 (Corollary 3 of [58]) *Suppose $\bar{v} \in \mathbb{T}$, then*

$$1 + \rho(\bar{v})\lambda_2 \le Q_{11}(\bar{v}) + Q_{22}(\bar{v}) \le 1 + \lambda_2$$

where $\rho(\bar{v})$ is defined in (3.22).

Proof of Proposition 3.3.7: The proof is similar to the proof of Proposition 3.3.6. Denote $\bar{\rho} = \rho(\bar{v}^{(2)})$, $\bar{\beta} = \beta(\bar{v}^{(2)})$. Using the formula for $Q(\bar{v})$ (see (3.3)), we have

$$1 + \bar{\rho}\lambda_2 \le \frac{\alpha}{\bar{\beta}} + \frac{1-2\bar{\beta}+\alpha}{1-\bar{\beta}} \le 1 + \lambda_2.$$

Then,

$$\bar{\rho}\lambda_2\bar{\beta}(1-\bar{\beta}) + \bar{\beta}^2 \le \alpha \le \lambda_2\bar{\beta}(1-\bar{\beta}) + \bar{\beta}^2. \tag{B.18}$$

The proof then follows by using Lemma B.8.2. ■

B.10 Proof of Theorem 3.3.8

As $\varepsilon \to 0$, $\lambda_2(\check{P}_\varepsilon) \to 1$ and the corresponding eigenvector $u^{(2)}(\check{P}_\varepsilon)$ approaches a piecewise constant vector (see [33]). This implies that $\rho(\bar{v}_\varepsilon^{(2)}) \to 1$. Hence, the lower bound approaches the upper bound of $F(\bar{v}_\varepsilon^{(2)})$ as $\varepsilon \to 0$, which proves the first part of Theorem 3.3.8. The second part follows from the straightforward substitution

$$F_\circ(1,\beta^*) = -(\beta^*)\log(\beta^*) - (1-\beta^*)\log(1-\beta^*) = h(\beta^*).$$

Appendix C

AlgoBIPA: A Recursive Bi-partition Algorithm

In this chapter, we provide a heuristic algorithm to address the optimal partition problem (2.13) or (2.15). Recall that the bi-partition problem is approximated by considering the second eigenvector in Chapter 3. The heuristic here is to employ a recursive bi-partition algorithm to obtain a sub-optimal solution for the m-partition problem.

Let $u^{(2)}$ denote the eigenvector corresponding to the second largest eigenvalue of $\check{P}$ (see (3.12)). A sub-optimal bi-partition function ϕ^* is obtained by considering the sign-structure of $u^{(2)}$ (see Chapter 3 for more details):

$$\phi_i^* = \begin{cases} 1, & \text{if } u_i^{(2)} \geq 0 \\ 2, & \text{otherwise.} \end{cases}$$

A recursive bi-partition algorithm (AlgoBIPA) is proposed to obtain the $m \geq 2$ partitions in a sub-optimal way: In the mth iteration of the algorithm, we assume that a partition with m groups (or super-states) is given. The objective of the mth-iteration is to obtain a refinement that has $(m+1)$ groups. For $i = 1, \ldots, m$, we denote by $P^{(i)}$ the sub-Markov transition matrix that describes the transition probabilities within the ith group. The ith group is split into two sub-groups according to the sign-structure of the second eigenvector for the eigenvalue problem associated with $\check{P}^{(i)}$. The spectral split of the ith group alone provides a partition of the states into $(m+1)$ groups. We denote this partition as $\phi^{(i)}$, and use it to evaluate the optimal aggregated transition matrix $Q(\phi^{(i)})$ according to (2.11). From the resulting m possible choices of $(m+1)$-partitions, we select the one that minimizes $R_{\phi^{(i)}}(P\|Q(\phi^{(i)}))$, i.e.,

$$i_{\min} = \underset{i \in \{1,\ldots,m\}}{\arg\min}\ R_{\phi^{(i)}}(P\|Q(\phi^{(i)})).$$

The associated aggregated transition matrix is chosen as $Q(\phi^{(i_{\min})})$.

The recursive algorithm is a heuristic based on the consideration of the bi-partition problem. A termination strategy can be based on a threshold value for the modeling error, e.g., the algorithm can be terminated if

$$\left| R_{\phi_m}(P\|Q(\phi_m)) - R_{\phi_{m-1}}(P\|Q(\phi_{m-1})) \right| \leq \varepsilon_{\text{tol}}$$

where $\phi_m := \phi^{(i_{\min}^{(m)})}$, $\phi_{m-1} := \phi^{(i_{\min}^{(m-1)})}$, and ε_{tol} is a pre-specified tolerated error.

Remark C.0.1 In certain cases, the heuristic may be used in conjunction with other methods. For instance, if the number of partitions is a priori known (e.g., $m = 2$ in this paper), one could consider employing a stochastic local search, starting with a spectral sub-optimal partition, and considering neighboring partitions (e.g. by swapping states between partitions) [38]. One can also consider including a regularization term in the objective function to favor smaller partitions [39].

Appendix D

Proofs for Chapter 8

D.1 Proof of Proposition 8.2.1

The proof relies on the dynamic programming equation for the fluid model,

$$\min_{\zeta} \tfrac{d}{dt} T^*(q(t)) = -1.$$

This implies that for any time t, and any feasible control for the fluid model,

$$T^*(q(t+1)) \geq T^*(q(t)) - 1. \tag{D.1}$$

We now turn to the CRW model. Since the function T^* is convex, then by Jensen's inequality,

$$\begin{aligned}\mathsf{E}[T^*(Q(t+1)) \mid Q(t) = x,\, U(t) = u] &\geq T^*\big(\mathsf{E}[Q(t+1) \mid Q(t) = x,\, U(t) = u]\big) \\ &= T^*(x + Bu).\end{aligned}$$

Returning to the fluid model, if $q(t) = x$, then $q(t+s) = x + (Bu)s$, $0 \leq s \leq 1$ is feasible for the fluid model. The bound (D.1) then gives,

$$\mathsf{E}[T^*(Q(t+1)) \mid Q(t) = x,\, U(t) = u] \geq T^*(x) - 1.$$

We conclude that the process defined below is a sub-martingale,

$$M(t) := T^*(Q(t)) + t, \qquad t = 0, 1, 2, \ldots$$

The submartingale property gives, for any initial condition $Q(0) = x$,

$$\mathsf{E}_x[M(t \wedge \mathcal{T}_0)] \geq M(0) = T^*(x). \tag{D.2}$$

If the mean of $\mathcal{T}_0$ is infinite, (8.11) is straightforward. Now suppose the mean of $\mathcal{T}_0$ is finite, then the collection of random variables $\{M(t \wedge \mathcal{T}_0) : t \geq 0\}$ is uniformly integrable. To see this we first bound these random variables as follows,

$$\begin{aligned} M(t \wedge \mathcal{T}_0) &= M(0) + \sum_{i=0}^{t \wedge \mathcal{T}_0 - 1} \big(M(i+1) - M(i)\big) \\ &\leq M(0) + \sum_{i=0}^{\mathcal{T}_0 - 1} \big|M(i+1) - M(i)\big|. \end{aligned}$$

The right hand side has finite mean provided the mean of $\mathcal{T}_0$ is finite, since $\big|M(i+1) - M(i)\big|$ is a bounded sequence. Uniform integrability follows.

Consequently, we can let $t \to \infty$ in (D.2) to obtain,

$$\mathsf{E}_x[T^*(Q(\mathcal{T}_0)) + \mathcal{T}_0] = \mathsf{E}_x[M(\mathcal{T}_0)] \geq M(0) = T^*(x).$$

This establishes the desired bound since $T^*(Q(\mathcal{T}_0)) = 0$.

D.2 Value function of two-dimensional relaxation

To obtain the expression of $\widehat{\overline{J}}^*$, we denote the upper boundary, the upper interior, the lower boundary, and the lower interior of $\widehat{\mathsf{W}}$ as:

$$\begin{aligned} \partial\widehat{\mathsf{W}}_1 &:= \{w : w_1 = 0, w_2 \geq 0\} \\ \widehat{\mathsf{W}}_1^\circ &:= \{w : w_2 \geq w_1 > 0\} \\ \partial\widehat{\mathsf{W}}_2 &:= \{w : w_2/\xi_1^2 = w_1/\xi_1^1 \geq 0\} \\ \widehat{\mathsf{W}}_2^\circ &:= \{w : w_1/\xi_1^2 > w_2/\xi_1^2 > w_1/\xi_1^1 \geq 0\}. \end{aligned}$$

Let M^1 and M^2 denote two 2×2 symmetric matrices to be specified latter. The value function is expressed as:

$$\widehat{\overline{J}}^*(w) = \begin{cases} \frac{1}{2} w^\intercal M^1 w, & \text{if } w \in \widehat{\mathsf{W}}_1^\circ \cup \partial\widehat{\mathsf{W}}_1 \\ \frac{1}{2} w^\intercal M^2 w, & \text{if } w \in \widehat{\mathsf{W}}_2^\circ \cup \partial\widehat{\mathsf{W}}_2. \end{cases}$$

Computation of $\widehat{\overline{J}}^*$ is performed as follows: First, we find the optimal drift vector δ^i and the effective cost $\bar{c}(w) = \langle \bar{c}^i, w \rangle$ for $i = 0, 1, 2$ for different parts of workload space $\widehat{\mathsf{W}}$. For $w \in \widehat{\mathsf{W}}_1^\circ \cup \widehat{\mathsf{W}}_2^\circ$, $\delta^0 := (-1, -1)^\intercal$, $\bar{c}^0 := ((c_6 - c_7)/\xi_1^1, c_7/\xi_1^2)^\intercal$. For $w \in \partial\widehat{\mathsf{W}}_1$, $\delta^1 := (0, -1)^\intercal$, $\bar{c}^1 := (0, c_7/\xi_1^2)^\intercal$. For $w \in \partial\widehat{\mathsf{W}}_2$, $\delta^2 := (-1, -\xi_1^2/\xi_1^1)^\intercal$, $\bar{c}^2 := (c_6/\xi_1^1, 0)^\intercal$.

Second, the value function is obtained by solving the following dynamic programming:

$$\min_{\delta \geq -1} \langle \nabla \widehat{\overline{J}}^* (w),\ \delta \rangle = -\bar{c}(w).$$

For $w \in \widehat{\mathsf{W}}_1^\circ \cup \partial \widehat{\mathsf{W}}_1$, we have

$$\langle M^1 w, \delta^0 \rangle = -\langle \bar{c}^0, w \rangle, \qquad \langle M^1 w, \delta^1 \rangle = -\langle \bar{c}^1, w \rangle.$$

It follows that $M^1 \delta^i = -\bar{c}^i$ for $i = 0, 1$, which implies

$$M^1 = \begin{bmatrix} (c_6 - c_7)/\xi_1^1 & 0 \\ 0 & c_7/\xi_1^2 \end{bmatrix}.$$

For $w \in \widehat{\mathsf{W}}_2^\circ \cup \partial \widehat{\mathsf{W}}_2$, we have

$$\langle M^2 w, \delta^0 \rangle = -\langle \bar{c}^0, w \rangle, \qquad \langle M^2 w, \delta^2 \rangle = -\langle \bar{c}^2, w \rangle.$$

It follows that $M^2 \delta^i = -\bar{c}^i$ for $i = 0, 2$, which implies

$$M^2 = \frac{1}{\xi_1^1 - \xi_1^2} \begin{bmatrix} c_6 - (c_6 - c_7)\xi_1^2/\xi_1^1 & -c_7 \\ -c_7 & c_7 \xi_1^1/\xi_1^2 \end{bmatrix}.$$

It is straightforward to verify that the gradient of $\widehat{\overline{J}}^*$ is continuous on the boundary between two regions $\widehat{\mathsf{W}}_1^\circ$ and $\widehat{\mathsf{W}}_2^\circ$, i.e., $M^1 w = M^2 w$ for $w_1 = w_2$. Thus the C^1 property of $\widehat{\overline{J}}^*$ is satisfied for all $w \in \widehat{\mathsf{W}}$.

References

[1] S. A. Al-Baiyat, M. Bettayeb, and U. M. Al-Saggaf. New model reduction scheme for bilinear systems. *International Journal of Systems Science*, 25(10):1631–1642, 1994.

[2] F. Alghimlas. Building load simulation and validation of an office building. In *Second International Conference for Enhanced Building Operations*, 2002.

[3] American Physical Society. *Energy Future: Think Efficiency*, September 2008.

[4] American Society of Heating, Refrigerating and Air-Conditioning Engineers. *ASHRAE Handbook Fundamentals (SI Edition)*, 2009.

[5] American Society of Heating, Refrigerating and Air-Conditioning Engineers. *ASHRAE Standard 55-2010, Thermal Environmental Conditions for Human Occupancy*, 2010.

[6] L. E. Baum and T. Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.*, 37(6):1559–1563, 1966.

[7] C. L. Beck, S. Lall, T. Liang, and M. West. Model reduction, optimal prediction, and the Mori-Zwanzig representation of Markov chains. In *Proceedings of IEEE Conference on Decision and Control*, pages 3282–3287, Shanghai, China, 2009.

[8] D. P. Bertsekas and R. Gallager. *Data Networks*. Prentice-Hall, Englewood Cliffs, NJ, 1992.

[9] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, Bemont, MA, 1996.

[10] G. Bolch, S. Greiner, H. Meer, and K. S. Trivedi. *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. John Wiley & Sons, Inc., Hoboken, NJ, 2006.

[11] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Hindustan Book Agency and Cambridge University Press (jointly), Delhi, India and Cambridge, UK, 2008.

[12] J. P. Bourdouxhe, M. Grodent, and J. Lebrun. *Reference Guide for Dynamic Models of HVAC Equipment*. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc, Atlanta, 1998.

[13] M. R. Brambley, D. Hansen, P. Haves, D. R. Holmberg, S. C. McDonald, K. W. Roth, and P. Torcellini. Advanced sensors and controls for building applications: Market assessment and potential R&D pathways. *DOE under Contract DE-AC05-76RL01830*.

[14] J. E. Braun, K. W. Montgomery, and N. Chaturvedi. Evaluating the performance of building thermal mass control strategies. *Heating Ventillation Air Conditioning and Refrigeration Research*, 7(4):403–428, 2004.

[15] C. Burstedde, A. Kirchner, K. Klauck, A. Schadschneider, and J. Zittartz. Cellular automation approach to pedestrian dynamics-applications. In Michael Schreckenberg and Som Deo Sharma, editors, *Pedestrian and Evacuation Dynamics*, pages 87–97. Springer, 2002.

[16] O. Cappé, E. Moulines, and T. Ryden. *Inference in Hidden Markov Models (Springer Series in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, 2005.

[17] Carrier Corporation. *HAP Quick Reference Guide*, 2003.

[18] V. Chandan and A. G. Alleyne. Optimal control architecture selection for thermal control of buildings. In *Proceedings of American Control Conference*, pages 2071 – 2076, San Francisco, CA, 2011.

[19] Q. Chen. Ventilation performance prediction for buildings: A method overview and recent applications. *Building and Environment*, 44(4):848–858, 2009.

[20] A. J. Chorin, O. H. Hald, and R. Kupferman. Optimal prediction and the Mori-Zwanzig representation of irreversible processes. *PNAS*, 97(7):2968–2973, 2000.

[21] F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, Providence, RI, 1997.

[22] I. J. Couchman, E. C. Kerrigan, and C. Bohm. New model reduction scheme for bilinear systems. *Model reduction of homogeneous-in-the-state bilinear systems with input constraints*, 47(4):761–768, 2011.

[23] P. J. Courtois. *Decomposability: Queueing and Computer System Applications*. Academic Press, New York, NY, 1977.

[24] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., New York, NY, first edition, 1991.

[25] J. G. Dai. On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. *Ann. Appl. Probab.*, 5(1):49–77, 1995.

[26] J. G. Dai and G. Weiss. A fluid heuristic for minimizing makespan in job shops. *Operations Res.*, 50(4):692–707, 2002.

[27] K. Deng, P. Barooah, P. G. Mehta, and S. Meyn. Building thermal model reduction via aggregation of states. In *Proceeding of American Control Conference*, pages 5118–5123, Baltimore, MD, 2010.

[28] K. Deng, W. Chen, P. G. Mehta, and S. P. Meyn. Resource pooling for optimal evacuation of a large building. In *Proc. of the 47th IEEE Conference on Decision and Control*, pages 5565–5570, Cancun, Mexico, 2008.

[29] K. Deng, P. G. Mehta, and S. P. Meyn. Aggregation-based model reduction of a Hidden Markov Model. In *Proceedings of IEEE Conference on Decision and Control*, Atlanta, GA, 2010.

[30] K. Deng, P. G. Mehta, and S. P. Meyn. Optimal Kullback-Leibler aggregation via the spectral theory of Markov chains. *IEEE Trans. Automat. Contr.*, 56(12):2793–2808, 2011.

[31] K. Deng, Y. Sun, P. G. Mehta, and S. P. Meyn. An information-theoretic framework to aggregate a Markov chain. In *Proceedings of American Control Conference*, pages 731–736, St. Louis, MO, 2009.

[32] P. Deuflhard, W. Huisinga, A. Fischer, and C. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Linear Algebra Appl.*, 315(1-3):39–59, 2000.

[33] P. Deuflhard and M. Weber. Robust Perron cluster analysis in conformation dynamics. *Linear Algebra and its Applications*, 398(15):161–184, 2005. Special Issue on Matrices and Mathematical Biology.

[34] S. Dey. Reduced-complexity filtering for partially observed nearly completely decomposable Markov chains. *IEEE Trans. Signal Processing*, 48(12):3334–3344, 2000.

[35] P. Diaconis. The Markov chain Monte Carlo revolution. *Bull. Amer. Math. Soc.*, 46(2):179–205, 2009.

[36] M. D. Donsker and S. R. S. Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Comm. Pure Appl. Math.*, 28:1–47; ibid. **28** (1975), 279–301, 1975.

[37] R. Douc, E. Moulines, J. Olsson, and R. van Handel. Consistency of the maximum likelihood estimator for general hidden markov models. *Ann. Statist.*, 39(1):474–513, 2011.

[38] W. E, T. Li, and E. Vanden-Eijnden. Optimal partition and effective dynamics of complex networks. *PNAS*, 105(23):7907–7912, 2008.

[39] M. Fazel. *Matrix rank minimization with applications*. PhD thesis, Stanford University, 2002.

[40] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Math. J.*, 25(100):619–633, 1975.

[41] J. A. Fill. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab.*, 1(1):62–87, 1991.

[42] F. Le Gland and L. Mevel. Exponential forgetting and geometric ergodicity in hidden Markov models. *Math. of Control Signals and Systems*, 13(1):63–93, 2000.

[43] M. M. Gouda, S. Danaher, and C. P. Underwood. Low-order model for the simulation of a building and its heating system. *Building services energy research technology*, 21(3):199–208, 2000.

[44] M. M. Gouda, S. Danaher, and C.P. Underwood. Building thermal model reduction using nonlinear constrained optimization. *Building and Environment*, 37(12):1255–1265, 2002.

[45] S. Goyal and P. Barooah. A method for model-reduction of nonlinear building thermal dynamics. *Energy and Buildings*, 47:332–340, 2012.

[46] S. Goyal, H. Ingley, and P. Barooah. Zone-level control algorithms based on occupancy information for energy efficient buildings. In *American Control Conference*, Montreal, Canada, June 2012. accepted.

[47] M. S. Gray and J. Mesko. Energy functions and algebraic gramians for bilinear systems. In *Proceedings of the 4th IFAC Nonlinear Control Systems Design Symposium*, The Netherlands, 1998.

[48] S. Gwynne, E. R. Galea, M. Owen, P. J. Lawrence, and L. Filippidis. A systematic comparison of building EXODUS predictions with experimental data from the Stapelfeldt trials and the Milburn House evacuation. *Applied Mathematical Modelling*, 29:818–851, 2005.

[49] J. Hahn and T. F. Edgar. An improved method for nonlinear model reduction using balancing of empirical gramians. *Computers & Chemical Engineering*, 26(10):1379–1397, 2002.

[50] J. M. Harrison and M. J. López. Heavy traffic resource pooling in parallel-server systems. *Queueing Syst. Theory Appl.*, 33(4):339–368, 1999.

[51] D. Helbing, L. Buzna, A. Johansson, and T. Werner. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation Science*, 39(1):1–24, February 2005.

[52] B. Hendrickson and R. Leland. Multidimensional spectral load balancing. Technical report, Sandia National Laboratories, Albuquerque, NM, 1993. SAND93-0074.

[53] T. Hong, F. Buhl, and P. Haves. EnergyPlus run time analysis, 2008. California PIER Project # 500-07-008.

[54] J. M. Horan and D. P. Finn. CFD reliability issues in the prediction of airflows in a naturally ventilated building. *International Journal of Ventilation*, 4(3):255–268, 2005.

[55] M. Huang, P. E. Caines, and R. P. Malhame. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria. *IEEE Transactions on Automatic Control*, 52(9):1560–1571, 2007.

[56] M. Huang, R. P. Malhame, and P. E. Caines. Nash certainty equivalence in large population stochastic dynamic games: Connections with the physics of interacting particle systems. In *Procs. of IEEE Conference on Decision and Control*, pages 4921 –4926, San Diego, CA, 2006.

[57] W. Huisinga, S. P. Meyn, and C. Schütte. Phase transitions and metastability in Markovian and molecular systems. *Ann. Appl. Probab.*, 14(1):419–458, 2004.

[58] W. Huisinga and B. Schmidt. Metastability and dominant eigenvalues of transfer operators. In *New Algorithms for Macromolecular Simulation*, volume 49 of *Lecture Notes in Computational Science and Engineering*, pages 167–182. Springer Berlin Heidelberg, 2006.

[59] B. H. Juang and L. R. Rabiner. A probabilistic distance measure for Hidden Markov Models. *AT&T Tech. J.*, 64(2):391–408, 1985.

[60] O. Junge and M. Dellnitz. Almost invariant sets in Chua's circuit. *Int. J. Bif. and Chaos*, 7(11):2475–2485, 1997.

[61] J. H. Kampf and D. Robinson. A simplified thermal model to support analysis of urban resource flows. *Energy and buildings*, 39(4):445–453, 2007.

[62] J. Kemeny and J. Snell. *Finite Markov Chains*. D. Van Nostrand, New York, NY, 1960.

[63] J. Kemeny and J. Snell. *Mathematical Models in the Social Sciences*. The MIT Press, Cambridge, MA, 1978.

[64] M. K. Khattar and H. I. Henderson. Experiences with modeling supermarket energy use and performance. In *IEA Supermarket Refrigeration Workshop*, October 2000.

[65] R. Kindermann and J. Snell. *Markov Random Fields and Their Applications*. American Mathematical Society, Providence, RI, 1980.

[66] V. Krishnamurthy and G. G. Yin. Recursive algorithms for estimation of Hidden Markov Models and autoregressive models with Markov regime. *IEEE Trans. Info. Thy.*, 48(2):458–476, 2002.

[67] A. Krogh, M. Brown, I. S. Mian, K. Sjolander, and D. Haussler. Hidden Markov Models in computational biology: applications to protein modelling. *Journal of Molecular Biology*, 235(5):1501–1531, 1994.

[68] National Renewable Energy Laboratory. National solar radiation data base (NSRDB), 2005.

[69] S. Lall, J. E. Marsden, and S. Glavaski. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *Journal of Robust and Nonlinear Control*, 12(6):519–526, 2002.

[70] J. M. Lasry and P. L. Lions. Mean field games. *Japan. J. Math.*, 2(1):229–260, 2007.

[71] K. H. Lee and J. E. Braun. Model-based demand-limiting control of building thermal mass. *Building and Environment*, 43(10):1755–1768, 2008.

[72] B. G. Leroux. Maximum-likelihood estimation for hidden Markov models. *Stochastic Process. Applic.*, 40(1):127–143, 1992.

[73] Y. Ma, G. Anderson, and F. Borrelli. A distributed predictive control approach to building temperature regulation. In *Proceedings of American Control Conference*, pages 2089 – 2094, San Francisco, CA, 2011.

[74] S. Mahadevan and M. Maggioni. Value function approximation using diffusion wavelets and laplacian eigenfunctions. In *Advances in Neural Information Processing Systems*, Cambridge, MA, 2006. MIT Press.

[75] P. Marbach and J. N. Tsitsiklis. Simulation-based optimization of Markov reward processes. *IEEE Transactions on Automatic Control*, 46(2):191–209, 2001.

[76] P. G. Mehta, M. Dorobantu, and A. Banaszuk. Graph-based multi-scale analysis of building system transport models. In *Procs. of American Controls Conference*, pages 1110–1115, Minneapolis, 2006.

[77] M. Meila and L. Xu. Multiway cuts and spectral clustering. Technical report, University of Washington, May 2003.

[78] S. P. Meyn. Stability and optimization of queueing networks and their fluid models. In *Mathematics of stochastic manufacturing systems (Williamsburg, VA, 1996)*, pages 175–199. Amer. Math. Soc., Providence, RI, 1997.

[79] S. P. Meyn. Workload models for stochastic networks: Value functions and performance evaluation. *IEEE Trans. Automat. Control*, 50(8):1106–1122, 2005.

[80] S. P. Meyn. *Control Techniques for Complex Networks*. Cambridge University Press, Cambridge, 2007.

[81] S. P. Meyn. Stability and asymptotic optimality of generalized MaxWeight policies. *SIAM J. Control Optim.*, 47(6):3259–3294, 2009.

[82] S. P. Meyn, A. Surana, Y. Lin, S. M. Oggianu, S. Narayanan, and T. A. Frewen. A sensor-utility-network method for estimation of occupancy in buildings. In *Proceedings of the* 48*th IEEE Conference on Decision and Control*, pages 1494–1500, Shanghai, China, 2009.

[83] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, Cambridge, second edition, 2009.

[84] P. D. Morosan, R. Bourdais, and D. Dumur. A distributed MPC strategy based on benders decomposition applied to multi-source multi-zone temperature regulation. *Journal of Process Control*, 21(5):729–737, 2011.

[85] M. Mossolly, K. Ghalib, and N. Ghaddar. Optimal control strategy for a multi-zone air conditioning system using a genetic algorithm. *Energy*, 34(1):58–66, 2009.

[86] G. F. Newell. *Applications of queueing theory*. Monographs on Statistics and Applied Probability. Chapman & Hall, London, second edition, 1982.

[87] T. X. Nghiem and G. J. Pappas. Receding-horizon supervisory control of green buildings. In *Proceedings of American Control Conference*, pages 4416 – 4421, San Francisco, CA, 2011.

[88] J. Niedbalski, K. Deng, P. G. Mehta, and S. P. Meyn. Model reduction for reduced order estimation in traffic models. In *Proceedings of American Control Conference*, pages 914–919, Seattle, WA, 2008.

[89] T. R. Nielsen. Simple tool to evaluate energy demand and indoor environment in the early stages of building design. *Solar Energy*, 78(1):73–83, January 2005.

[90] F. Oldewurtel, A. Parisio, C. N. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth. Energy efficient building climate control using stochastic model predictive control and weather predictions. In *American Control Conference*, pages 5100–5105, Baltimore, MD, 2010.

[91] R. G. Phillips and P. V. Kokotovic. A singular perturbation approach to modeling and control of Markov chains. *IEEE Trans. Automat. Contr.*, 26(5):1087–1094, 1981.

[92] Z. Rached, F. Alalaji, and L. L. Campbell. The Kullback-Leibler divergence rate between Markov sources. *IEEE Trans. Info. Thy.*, 50(5):917–921, 2004.

[93] D. Revuz. *Markov Chains*. Elsevier, New York, NY, second edition, 1984. North-Holland Mathematical Library, vol. 11, North-Holland, Amsterdam.

[94] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, New York, NY, second edition, 2004.

[95] H. Sandberg. An extension to balanced truncation with application to structured model reduction. *IEEE Trans. Automat. Contr.*, 55(4):1038–1043, 2010.

[96] M. Sarich, C. Schütte, and E. Vanden-Eijnden. Optimal fuzzy aggregation of networks. *Multiscale Model. Simul.*, 8(4):1535–1561, 2010.

[97] A. Schadschneider. Cellular automation approach to pedestrian dynamics-theory. In Michael Schreckenberg and Som Deo Sharma, editors, *Pedestrian and Evacuation Dynamics*, pages 75–86. Springer, 2002.

[98] J. Scherpen. Balancing for nonlinear systems. *Systems & Control Letters*, 21(2):143–153, 1993.

[99] M. Schreckenberg and S. D. Sharma, editors. *Pedestrian and Evacuation Dynamics*. Springer, Berlin, 2002.

[100] C. Schütte, A. Fischer, W. Huisinga, and P. Deuflhard. A direct approach to conformational dynamics based on hybrid Monte Carlo. *J. Comput. Phys., Special Issue on Computational Biophysics*, 151(1):146–168, 1999.

[101] P. J. Schweitzer. A survey of aggregation-disaggregation in large Markov chains. In W. J. Stewart, editor, *Numerical Solution of Markov Chains*, pages 63–88. Marcel Dekker Inc, New York, NY, 1991.

[102] J. E. Seem, S. A. Klein, W. A. Beckman, and J. W. Mitchell. Transfer functions for efficient calculation of multidimensional transient heat transfer. *Journal of Heat Transfer*, 111(1):5–12, 1989.

[103] P. Sharma, S. Salapaka, and C. L. Beck. Entropy-based framework for dynamic coverage and clustering problems. *IEEE Trans. on Automatic Control*, 57(1):135–150, 2012.

[104] S. Sharma and S. Gifford. Using RFID to evaluate evacuation behavior models. In *Annual Conference of the North American Fuzzy Information Processing Society-NAFIPS*, pages 804–808, 2005.

[105] T. S. Shen. ESM: a building evacuation simulation model. *Building and Environment*, 40(5):671–680, 2005.

[106] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[107] H. A. Simon and A. Ando. Aggregation of variables in dynamical systems. *Econometrica*, 29(2):111–138, 1961.

[108] J. M. Smith and D. Towsley. The use of queuing networks in the evaluation of egress from buildings. *Environment and Planning B: Planning and Design*, 8(2):125–139, 1981.

[109] M. Sourbrona, R. De Herdtb, T. Van Reetb, W. Van Passelb, M. Baelmansa, and L. Helsena. Efficiently produced heat and cold is squandered by inappropriate control strategies: A case study. *Energy and Buildings*, 41(10):1091–1098, 2009.

[110] D. W. Stroock. *An Introduction to Markov Processes*. Springer, New York, NY, 2005. Volume 230 of Graduate texts in mathematics.

[111] Y. Sun and P. G. Mehta. The Kullback-Leiber rate pseudo-metric for comparing dynamical systems. *IEEE Transactions on Automatic Control*, 55(7):1585–1598, 2010.

[112] B. Tashtoush, M. Molhim, and M. Al-Rousan. Dynamic model of an HVAC system for control analysis. *Energy*, 30(10):1729–1745, 2005.

[113] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Automat. Control*, 37(12):1936–1948, 1992.

[114] R. Tomastik, S. Narayanan, A. Banaszuk, and S. P. Meyn. Model-based Real-Time Estimation of Building Occupancy During Emergency Egress. 4th International Conference on Pedestrian and Evacuation Dynamics., February 27–29 2008.

[115] United States Energy Information Administration. *Annual energy outlook*, 2011.

[116] M. Vidyasagar. Reduced-order modeling of Markov and Hidden Markov Processes via aggregation. In *Proceedings of IEEE Conference of Decision and Control*, pages 1810–1815, Atlanta, GA, 2010.

[117] M. Vidyasagar. A metric between probability distributions on finite sets of different cardinalities and applications to order reduction. *IEEE Trans. Automat. Contr.*, 57(10):2464–2477, 2012.

[118] M. Vidyasagar, S. S. Mande, C. V. S. K. Reddy, and V. V. R. Rao. The 4M (Mixed Memory Markov Model) algorithm for finding genes in prokaryotic genomes. *IEEE Trans. Automat. Contr.*, 53:26–37, 2008. Special Issue.

[119] S. Wang. Dynamic simulation of building VAV air-conditioning system and evaluation of emcs on-line control strategies. *Building and Environment*, 36(6):681–705, 1999.

[120] A. D. Wentzell. The asymptotic behavior of the largest eigenvalue of a second order elliptic differential operator with a small parameter multiplying the highest derivatives. *Dokl. Akad. Nauk SSSR*, 202:19–22, 1972.

[121] L. White, R. Mahony, and G. Brushe. Lumpable Hidden Markov Models – model reduction and reduced complexity filtering. *IEEE Trans. Automat. Contr.*, 45(12):2297–2306, 2000.

[122] S. Wolfram. *Theory and applications of cellular automata*. World Scientific, Singapore, 1986.

[123] L. Xie, V. Ugrinovskii, and I. R. Petersen. Probabilistic distances between finite-state finite-alphabet Hidden Markov Models. *IEEE Trans. Automat. Contr.*, 50(4):505–511, 2005.

[124] X. Xu and S. Wang. A simplified dynamic model for existing buildings using CTF and thermal network models. *International Journal of Thermal Sciences*, 47(9):1249–1262, 2008.

[125] X. Xu, S. Wang, Z. Sun, and F. Xiao. A model-based optimal ventilation control strategy of multi-zone VAV air-conditioning systems. *Applied Thermal Engineering*, 29(1):91–104, 2009.

[126] G. G. Yin and Q. Zhang. *Continuous-time Markov Chains and Applications: A Singular Perturbation Approach*, volume 37 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, New York, 1998.

[127] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Synchronization of coupled oscillators is a game. In *American Control Conference*, pages 1783–1790, Baltimore, MD, 2010.

[128] M. Zaheer-uddin and G. R. Zheng. Optimal control of time-scheduled heating, ventilating and air conditioning processes in buildings. *Energy Conversion and Management*, 41(1):49–60, 2000.