Reach Probability Estimation of Rare Events in Stochastic Hybrid Systems

Ma, H.

DOI
10.4233/uuid:49eaed4b-ff4b-450d-97c9-8ed5dc5e7f22

Publication date
2023

Document Version
Final published version

Citation (APA)
Ma, H. (2023). *Reach Probability Estimation of Rare Events in Stochastic Hybrid Systems*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:49eaed4b-ff4b-450d-97c9-8ed5dc5e7f22

Reach Probability Estimation of Rare Events in Stochastic Hybrid Systems

Reach Probability Estimation of Rare Events in Stochastic Hybrid Systems

DISSERTATION

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus, Prof.dr.ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates,
to be defended publicly on
Wednesday, 29 November 2023 at 15.00 o'clock

by

Hao MA

Master of Science in Electrical Engineering,
Northwestern Polytechnical University, China,
born in Xianyang, China.

This dissertation has been approved by the promotors:

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof.dr.ir. H.A.P. Blom	Delft University of Technology, promotor
Dr.ir. B.F. Lopes dos Santos	Delft University of Technology, copromotor

Independent members:

Prof. dr. P.H.A.J.M. van Gelder	Delft University of Technology
Prof.dr.ir. Max Mulder	Delft University of Technology
Prof.dr. K. Pattipati	University of Connecticut
Prof.dr. B. Xu	Northwestern Polytechnical University
Dr.ir. M.H.C. Everdij	NLR, NL

This research is supported by the Chinese Scholarship Council (CSC), NO. 201606290141

Keywords: Interacting Particles, Factorization, Rare event, Reach Probability,
Stochastic Hybrid System
Printed by: Ipskamp Printing (www.ipskampprinting.nl)
Cover by: Hao MA

ISBN 978-94-6384-501-4
An electronic copy of this dissertation is available http://repository.tudelft.nl/.

To my family and my friends

献给我的家人和我的朋友

Acknowledgements

The process of pursuing a doctoral degree is like a Brownian motion, full of uncertainty. Thanks to the support of my respected mentor, inspiring colleagues, valuable friends, and family, I finally completed my doctoral study.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Henk Blom. It is through your guidance and advice that I was able to complete my doctoral thesis. Your rigorous scientific research attitude has deeply impressed me. In our discussions on my research plan, manuscripts, and even every formula, you made me realize that there should be no uncertainty on the path of scientific research. Your tireless teachings will always be remembered by me and will serve as a valuable foundation for my future research. At the same time, I would also like to thank my respected mentor Dr. Bruno Santos. Your words of encouragement gave me the strength to continue, especially when I faced challenges in my doctoral life.

Thanks to the partners in the Air Traffic Operation (ATO) group. Matthieu Vert, you are my best friend to share. You always support me! Juseong Lee, my favourite bald friend, you will always be remembered as my best study object. Ingeborg de Pater, your encouragement has added confidence to me. Iordanis Tseremoglou thank you for practicing Greek with me. Chengpeng Jiang, thank you for discussing scientific research issues with me day and night. With your help and encouragement, I was able to persevere until now. Mike Zoutendijk and Simon van Oosterom, thank you for giving me various suggestions for studying and living in the Netherlands. I am grateful to other students at ATO—Marie Bieber, Malte von der Burg, Mahdi Noorafza, Ilias Parmaksizoglou, Thomas Pioger, Haonan Li and Dr. Gülçin Ermiş. In addition, I would like to thank ATO alumni who helped me settle down in Delft. Thank you Dr. Qichen Deng, Dr. Alessandro Bombelli, Dr. Wenhua Qu, Dr. Raissa Li, Dr. Stef Janssen, Dr. Vis Dhanisetty, Hemmo Koornneef and Dr. VinhHo-Huu. Finally, thank you to the professors and staff at ATO: Paul Roling, Elise Bavelaar, Dr. Márcia Baptista and Nathalie Zoet.

In addition to research, I would also like to thank all my Chinese friends who have been constantly encouraging me. Thank you Joshua Tsang, Quan Guo, Zhi Jiang, Xiaoyan Sun, Xingyu He, Haotian Niu, Xiaoqing Hu, Jin Zhao and Dr. Maolong Lv. Most importantly, I want to thank Weijia Zhu.

Then I want to dedicate all the honors to my family. My parents, Cunbao Ma and Binbin Wang, gave me selfless love, unconditional support, and sincere advice. When I felt depressed and desperate and wanted to give up, they silently cared for me and enlightened me. When I published a paper, they celebrated for me. When I had emotional problems, your dedication made me deeply feel that there is a place that belongs to me in China. In addition, my lovely sister. Thank her for caring about me when I was down.

Finally, I want to thank myself! Hard work may not get you everything, but at least it will make you better!

Summary

Reach Probability Estimation of Rare Events in Stochastic Hybrid Systems

Hao MA

This thesis conducts a series of interrelated research studies on reach probability estimation of rare events for stochastic hybrid systems. Chapter 1 explains that the motivation for these studies stems from the need to assess safety and capacity of a design for a future Air Traffic Management (ATM) concept of operations (ConOps). The safety/capacity of an ATM ConOps can be expressed in terms of the amount of traffic that can be handled in such a way that the probability of rare events remains sufficiently low. Chapter 1 also explains that the dynamic and stochastic behaviours in an ATM ConOps design can be captured by a General Stochastic Hybrid System (GSHS) model, and that the rare events to be studied can be defined as events that the state of a GSHS model reaches an unsafe set. In ATM safety studies, an unsafe set often considered is the closed subset in the GSHS state space where the physical shapes of two aircraft overlap. The state of a GSHS model consists of two components: i) a Euclidean valued component, and ii) a discrete valued component. The evolution of these two components influence each other; therefore a GSHS model can capture various types of dynamic and stochastic behaviours, including Brownian motion and spontaneous jumps. In contrast to forced jumps, that happen when the GSHS state reaches a boundary in the hybrid state space, spontaneous jumps occur according to a Poisson point process. A mathematically important property of GSHS, is that a GSHS execution satisfies the strong Markov property.

A straightforward approach in estimating the Reach Probability of an unsafe set by a GSHS model is to conduct a large amount of Monte Carlo (MC) simulation runs, and calculate the fraction of runs in which the unsafe set is reached. For realistic application of such MC based rare event estimation approach, there is need for analytical methods that allow to accelerate the simulation. Literature on such acceleration distinguishes two main approaches: Importance Sampling (IS) and Importance Splitting (ISp). The IS approach is to draw random samples from a reference stochastic process model instead of the original process model, and to compensate the estimated reach probability through an analytically derived factor to compensate for sampling from the reference model instead of the original model. The ISp approach embeds the unsafe set by an increasing sequence of nested subsets, and then estimates the Reach Probability

of the unsafe set as a product of conditional probabilities of reaching the next inner subset. The mathematically best developed ISp approach for a strong Markov process makes use of an Interacting Particle System (IPS). For ATM ConOps evaluation, the IPS approach has demonstrated that it may yield a very large acceleration factor. However, to also assess the effect of GSHS model parameter changes on the reach probability, there is need for further improvements. Therefore, the overall aim of this thesis is to develop significant improvements in simulation based Reach Probability estimation for a GSHS.

Chapter 2 investigates a multi-dimensional diffusion process using the IPS framework. In this study, the IPS performances is analysed for four splitting strategies: multinomial resampling (MR), multinomial splitting (MS), residual multinomial splitting (RMS), and fixed assignment splitting (FAS), when employing a finite number of particles. These strategies differ in how they sample the new set of particles from the set of successful particles. The study proves that IPS using FAS dominates in variance reduction over IPS using MR, MS and RMS.

Chapter 3 extends the Chapter 2 results, for a multi-dimensional diffusion process, to a GSHS. In applying IPS to a GSHS, in literature, there are two simulation approaches. The formal approach is to simulate a GSHS according to its formal execution rules. The popular approach is to first transform the spontaneous jumps of a GSHS to forced jumps, and then to simulate this transformed version. Chapter 3 shows that the popular approach leads to a loss of the strong Markov property of the process defined by the original GSHS. Subsequently, Chapter 3 also proves that this loss of the strong Markov property has a negative effect on the acceleration factor of IPS for a GSHS with spontaneous jumps.

Chapter 4 studies an improvement of the IPS approach when a GSHS has mode values that occur at low probability, such as rare system failure conditions. To improve this situation, IPS is incorporated with sampling per mode, denoted as IPSmode. The IPSmode algorithm is combined with four mode-directed splitting strategies: MRmode, MSmode, RMSmode, and FASmode, each employing a finite number of particles. In contrast to the studies in Chapters 2 and 3, the mode-directed splitting strategies employing RMS and MS outperform the FAS approach. The explanation is that for the FASmode splitting approach it is more demanding to take proper account of the effect of particle weights in mode-dependent splitting.

Chapter 5 studies IS for GSHS. The motivation to do so was triggered by a recent development of IS theory for Piecewise Deterministic Markov Process (PDMP), which is a GSHS without Brownian motion. First, the optimal IS strategy for a PDMP is extended to a GSHS; this shows that Brownian motion plays a key role in the derivation, and in the optimal strategy. Second, the approximated IS strategy that has been developed for PDMP is extended to a GSHS. This approximated IS strategy assumes that the PDMP/GSHS consists of a number of subsystems in parallel redundancy, that are subject to failure and repair; and that the Euclidean valued process has no discontinuities. Under these conditions, chapter 5 shows that the approximated IS strategy developed for a PDMP also applies to a GSHS, i.e. the influence of Brownian motion disappears.

Chapter 6 draws conclusions regarding the results obtained through the series of studies. First conclusions are given from the novel contributions to the literature on rare event estimation for stochastic hybrid systems. Secondly, it is explained what this means for specific use in modelling and risk assessment of a future ATM ConOps. Finally, directions for follow-on research are mentioned.

Contents

1

Introduction

This chapter introduces rare event simulation, stochastic behaviors and research gaps addressed in this thesis. It describes the thesis goal and objectives. Furthermore, the thesis overview will be clarified by means of short chapter descriptions which explain how each individual chapter is linked to the overall aim.

1.1 Motivation

This PhD thesis studies rare event estimation using Monte Carlo (MC) simulation. The motivation for these studies stems from the increased need to evaluate a design of a future Air Traffic Management (ATM) Concept of Operations (ConOps) on safety and capacity.

In 2022, the Federal Aviation Administration (FAA) of the United States served approximately 16,405,000 flights, averaging over 45,000 flights and 2.9 million passengers per day (FAA, 2023). The annual growth rate for airline passengers is expected to be 5.5% over the next 20 years (FAA, 2021). As a result, this exponential growth will lead to more crowded airspace, posing challenges for effectively resolving future air traffic congestion. This will be exacerbated by the introduction of unmanned aerial vehicles operating at different altitudes.

Commercial air transportation has attained a very high level of safety, characterized by a notably low frequency of accidents, such as mid-air collision. This very high safety level has been reached through decades of learning from accidents and subsequent improvement of the air transport operations. Complementary to learning from accidents, safety risk assessment methods are in use for safety evaluation of changes of sub-systems in the overall air transportation system. For instance, the FAA has developed an "Air Traffic Services Safety Management Handbook" to conduct a safety risk analysis of changes to sub-systems in use by ATM. However, safety evaluation of sub-systems is inadequate to understand the impact of interactions between different sub-systems and human actors, such as pilots and air traffic controllers. To analyse the safety and capacity effects of these interactions, there is a need for systematic modelling and simulation of a design of a future ATM ConOps, and to provide feedback on the safety/capacity findings to the design team, as shown in Figure 1.1. Modelling and simulation during the design phase of a future ATM ConOps on safety/capacity enable designers and engineers to learn unknown behaviour and to develop design improvements before system deployment, thereby reducing costly errors and risks.

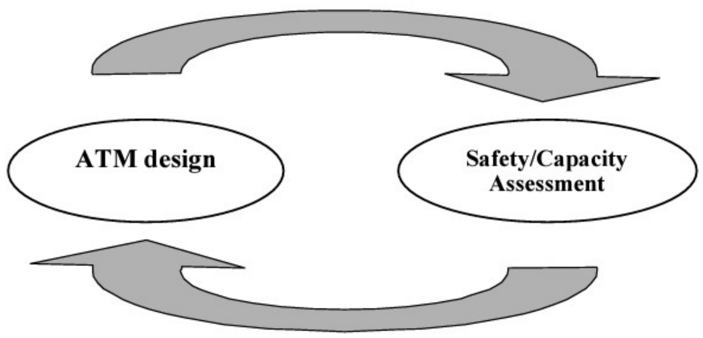

Figure 1.1. Safety/capacity assessment and feedback to ATM design [Blom et al., 2001]

An illustrative example of safety/capacity assessment through modelling and simulation of a design of a future ATM ConOps is found in the work conducted by Blom et al. (2007). They considered a given design of a next-generation ATM ConOps, and assessed mid-air collision risk as a function of increasing levels of air traffic demand. The modelling and simulation covered socio-technical issues such as crew reaction times, the reliability of various subsystems such as GNSS, ADS-B, and ASAS, as well as environmental factors like randomly varying wind conditions.

This illustrative example demonstrates that there are two key challenges. On the one hand, there is the modelling of the complex overall socio-technical system as defined by an ATM ConOps design. On the other hand, accurate assessment of mid-air collision probability asks for a mathematically proven rare event simulation method that applies to the developed model of the ATM ConOps design. In mathematics, rare event simulation is studied in terms of proven variance reduction methods.

In the next subsections, these two key challenges in modelling and rare event simulation are illustrated for the Free Flight ATM ConOps design studied by [Blom et al., 2007].

1.2 Socio-technical modelling of a future ATM ConOps Design

In current ATM ConOps, air traffic controllers on the ground are responsible for keeping aircraft well-separated from each other. In a free-flight design, the separation management responsibility is moved to pilots [RTCA, 1995]. Pilots flying in free-flight airspace are allowed to optimize their trajectories and have greater freedom in choosing paths and flight altitudes. Obviously, the safety implications arising from this level of freedom deserve serious attention in the design phase. This section illustrates the agent-based simulation model of [Blom et al., 2007] that has been developed to assess the safety risk of an early Free Flight design.

A simulation of a Free Flight operation involves a large number N of aircraft. In the agent-based model of [Blom et al., 2007], for each of the N aircraft, there are five active agents: one physical aircraft agent, one Pilot-Flying (PF) agent, one Pilot-not-Flying (PNF) agent, one agent for the Airborne Guidance, Navigation and Control (AGNC) system, and one agent for the Airborne Separation Assistance System (ASAS). As depicted in Figure 1.2, outside of these N aircraft-related agents, there is one common agent for the Communication, Navigation and Surveillance (CNS) systems that facilitate communication between agents for different aircraft. As shown in Table 1.1, each agent in Figure 1.2 consists of multiple interacting sub-systems. Table 1.1 also shows how the total number of sub-systems increases with the number N of aircraft flights in the airspace for which the ATM ConOps has to be evaluated for safety/capacity.

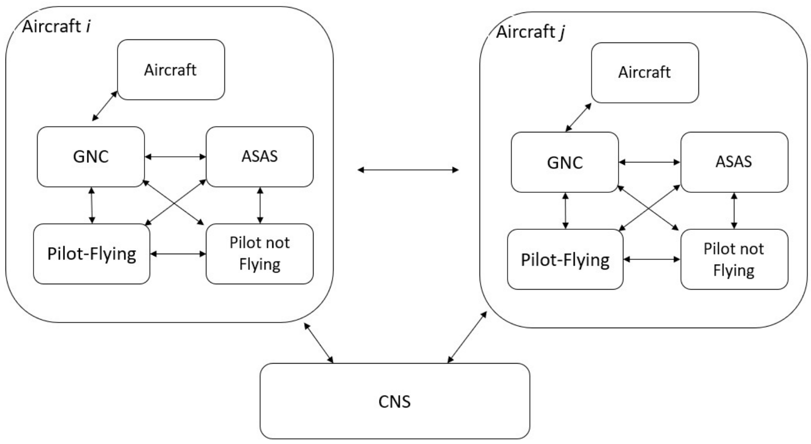

Figure 1.2. Agents and their interactions in the Free-Flight model of [Blom et al., 2007].

Table 1.1. Agent Types and number of subsystems in in the free flight operation, where *N* indicates the number of aircraft [Blom et al., 2007]

Agent Type	Number of agents	Number of subsystems per agent	Total number of subsystems
Aircraft	*N*	4	4*N*
Pilot-Flying (PF)	*N*	6	6*N*
Pilot-Not-Flying (PNF)	*N*	2	2*N*
ASAS	*N*	8	8*N*
AGNC	*N*	18	18*N*
Global CNS	1	3	3
Total	**5*N*+1**	**41**	**38*N* +3**

To properly specify an agent-based model that involves such large number of interacting agents and sub-systems, a compositional model specification method has to be used. Such method allows one to start with the model specification of each sub-system, followed by the specification of the interactions between the sub-systems within an agent, and finally the specification of the interactions between subsystems of different agents.

For the compositional specification of the agent-based model of the Free Flight ConOps design, [Blom et al., 2007] has used the compositional specification method of Stochastically and Dynamically Coloured Petri Nets (SDCPN) [Everdij et al., 2006]. During the first step of SDCPN model specification, for each sub-system a Local Petri Net (LPN) model is specified. During subsequent steps, the interactions between LPN's are specified. For example, the ASAS agent is composed of subsystems for: 1) Processing, 2) Alerting, 3)Audio Alerting, 4) Surveillance, 5) System mode, 6) Priority switch mode, 7) Anti-priority switch mode, and 8) Predictive alerting. Each of these subsystems is modelled as an LPN. Subsequently, within each ASAS agent, the ADS-B information received from other aircraft is processed by the Surveillance LPN. The Processing LPN subsequently uses the outcome of the Surveillance LPN to perform conflict detection and resolution with the own aircraft state, which information is provided by the GNC agent. The outcomes from the Processing LPN are subsequently used as input by the LPN's Alerting and Audio Alerting. Non-nominal events are also captured in the model. For example, the LPN System mode has three discrete states: i) Nominal working; ii) Failed; and iii) Corrupted. If the LPN System mode is Failed or Corrupted, then the performance of the Processing LPN is influenced. Once each agent is specified, then the interactions between agents are also systematically defined. For example, the PF agent may receive a conflict alert from the ASAS agent (Audio Alerting LPN).

The resulting SDCPN model specification of the Free Flight ConOps considered, is subsequently implemented as a Monte Carlo simulator that can run an air traffic scenario involving *N* flights. To accurately estimate the mid-air collision probability for such a Free Flight ConOps design, this scenario has to be simulated a large number (N_{runs}) of times, with randomly varying initial traffic conditions. In practice this asks for impractically large computer simulation times. To bring the computer simulation down by orders in magnitude, an effective rare event simulation method is needed.

1.3 Rare Event Simulation

The mathematical problem setting of rare event simulation is known as Reach Probability estimation, e.g., [Prandini et al., 2011]. For a continuous-time stochastic process $\{x_t\}$, which evolves in a state space X, reach probability is the probability that $\{x_t\}$ reaches a rare set $D \subset X$ within a finite time interval $[0,T]$. For mid-air collision in ATM, the rare set D is the area where two 3-dimensional aircraft shapes are in overlap.

A straightforward approach to Reach Probability estimation is to conduct many Monte Carlo (MC) runs and calculate the fraction of runs that reached the rare set D. For safety assessment of an air traffic scenario involving multiple aircraft, this would take an unrealistically long duration of computer simulation time. To reduce the simulation time, there is a need to apply mathematically sound variance reduction methods. Particular consideration is required to avoid a potential underestimation of the reach probability. For air traffic, heuristic variance reduction can easily lead to a systematic underestimation of mid-air collision risk.

Importance Sampling (IS) and multilevel splitting are the most powerful variance reduction approaches. IS allows rare events to occur more frequently by modifying the underlying probability distribution and then correcting the biased estimator by multiplying the result by the corresponding likelihood ratio [Bucklew, 2004; Glasserman, 2004]. The effectiveness of IS depends on finding the correct measure transformation, and improper operation may result in worse results than direct simulation. Typically, the rough asymptotic behaviour of the rare event probability needs to be determined in order to find the correct measure transformation. In complex dynamic models, this type of analysis may be challenging (e.g. Botev and Kroese, 2008; L'Ecuyer et al., 2009; Rubinstein, 2010; Morio and Balesdent, 2016).

Multilevel Splitting, also referred to as Importance Splitting (ISp), is a well-developed method for estimating the reach probability. The multilevel setting allows one to formulate the expression of the Reach probability of the rare set D as a product of larger reach probabilities for a sequence of nested subsets $\supset D$, e.g. [Glasserman et al., 1999]. Cérou et al. (2005, 2006) embed this multilevel factorization in the Feynman-Kac factorization for strong Markov processes (Del Moral, 2004). This Feynman-Kac setting subsequently supports the evaluation of the Reach probability through sequential Monte Carlo simulation in the form of an Interacting Particle System (IPS), including a proof of convergence (Cérou et al., 2006). The mathematical background of IPS requires that the stochastic process considered is a strong Markov process, i.e., that the Markov property applies not only to fixed times but also to stopping times.

IPS has been used in the Free Flight example [Blom et al., 2007]. Verification that the strong Markov property is satisfied has been accomplished as follows. For the SDCPN specification method used, [Everdij and Blom, 2006] has proven how the many stochastic variables in the complete SDCPN model defines a GSHS, for which the strong Markov property has been proven by [Bujorianu and Lygeros, 2006].

The use of IPS in [Blom et al., 2007] yields a large acceleration factor. However, to also conduct sensitivity analysis, i.e., to evaluate the effect of model parameter changes on the estimated rare event probabilities, the acceleration factor has to be increased by two extra orders of magnitude. Hence, a mathematically formulated motivation of the studies conducted in this PhD thesis is to investigate directions in realizing the two extra orders of the magnitude in variance reduction that IPS realizes in the illustrative Free Flight model of [Blom et al., 2007].

1.4 Stochastic behaviours in rare event simulation for future ATM

As has been explained in the previous subsection, the use of the SDCPN specification method defines a strong Markov process, which is a requirement the mathematical foundation of the IPS method. In a further study of variance reduction, it is also needed to understand which types of stochastic behaviours are used in the agent-based modelling of an ATM ConOps design. The study of variance reduction tends to be more demanding with the increasing complexity of the stochastic behaviours involved.

The first column in Table 1.2 lists the stochastic behaviour types that are relevant to characterise different classes of strong Markov processes [Lygeros and Prandini, 2010]:

- Hybrid state space: A state space that is a Kronecker product of a discrete set and an Euclidean space. For example, the hybrid aircraft state also covers flight mode and aircraft type in addition to position and velocity.
- Ordinary Differential Equation (ODE): The solution of an ODE is a deterministic flow. An example is to model the rate of a non-exponential spontaneous jump rate as a function of passed delay, that is the solution of an ODE.
- Stochastic Differential Equation (SDE): In addition to a deterministic flow, there also is a random influence by Brownian motion processes. An example is to use an SDE model to represent the uncertain and time-varying nature of wind, affecting the aircraft's velocity.
- Forced jumps: Jumps that occur when the Euclidean-valued state reaches the boundary of the hybrid state space. An example is a forced jumps from climb mode to level flight mode when the aircraft altitude reaches its intended cruise level.
- Spontaneous jumps: Jumps that occur according to a Poisson process and therefore are not predictable. An example is a sudden failure of a technical aircraft system.
- Hybrid spontaneity: This means that the rate of spontaneous jumps depends both on the discrete-valued mode and on the Euclidean-valued state. For example, the failure rate of a technical system also depends on the aircraft's altitude.
- Hybrid jumps: Jump in the Euclidean-valued state happens simultaneously with a transition in the discrete-valued state. An example is an aircraft switching from straight flight mode to a turn mode and, at the same time, making a jump in the aircraft’s bank angle.

The ATM examples given above for each stochastic behaviour type illustrate that the strong Markov process class to be studied should capture all the above stochastic behaviour types.

The first row in Table 1.2 lists relevant classes of strong Markov processes; e.g. [Lygeros and Prandini, 2010; Yin & Zhu, 2010]; these range from Diffusion to General Stochastic Hybrid System (GSHS). The increasing complexity of these strong Markov process classes can best be described through a description of stepwise extension of stochastic behaviour types, i.e. starting with Diffusion, and ending with GSHS:

Table 1.2. Overview of stochastic behaviour types and their support in various classes of strong Markov processes

Stochastic Behaviour Type	Diffusion	Jump-diffusion	CTMC	Switching Diffusion	Hybrid Switching Diffusion	Hybrid Switching Diffusion / hybrid jumps	PDMP	GSHS
Hybrid State	-	-	-	X	X	X	X	X
ODE	-	-	-	-	-	-	X	X
SDE	X	X	-	X	X	X	-	X
Forced Jumps	-	-	-	-	-	-	X	X
Spontaneous Jumps	-	X	X	X	X	X	X	X
Hybrid spontaneity	-	-	-	-	X	X	X	X
Hybrid jumps	-	-	-	-	-	X	X	X

- A Diffusion is an Euclidean-valued solution of an SDE driven by Brownian motion.
- A Jump-diffusion is a Euclidean-valued solution of an SDE driven by Brownian motion and Poisson random measure [Glassermann, 2004; Oksendal and Sulem, 2005], where the Poisson random measure triggers spontaneous jumps.
- A Continuous Time Markov Chain (CTMC) is a discrete-valued process that makes spontaneous jumps in its discrete state space.
- A Switching Diffusion is the solution of an SDE, the coefficients of which are a function of an independent CTMC [Ghosh et al., 1997; Mao and Yuan, 2006]. Hence a switching diffusion has a hybrid state space.
- Hybrid Switching Diffusion [Yin & Zhu, 2010]: Relative to a Switching Diffusion, the extra stochastic behaviour is that the transition rates in the CTMC are now a function of the Euclidean-valued state component.
- Hybrid Switching Diffusion with Hybrid Jumps [Hespanha, 2005]: Relative to Hybrid Switching Diffusion, the extra stochastic behaviour is that a jump in the Euclidean state component can happen simultaneously with a transition in the discrete-valued state component.
- Piecewise Deterministic Markov Process (PDMD) [Davis, 1984]. Relative to Hybrid Switching Diffusion, the extra stochastic behaviour is forced hybrid jumps that occur when the Euclidean-valued state hits predefined boundaries in the Euclidean sub-space. In a PDMP, this extension could be handled under the limitation that the flow within the Euclidean sub-space is deterministic. This implies restriction to an ODE instead of SDE.
- General Stochastic Hybrid System (GSHS) [Bujorianu and Lygeros, 2006]: Relative to PDMP, an SDE now replaces the ODE.

Based on the comprehensive understanding of the stochastic behaviours employed in agent-based modelling for the Free Flight example provided by [Blom et al., 2007], it is evident that the GSHS framework emerges as the most suitable formalism. The GSHS framework has the ability to effectively capture all stochastic behaviours, including the strong Markov property. Therefore, GSHS can be confidently recommended as the appropriate formalism for addressing the intricate stochastic behaviours encountered in future ATM ConOps designs.

1.5 Research Gaps in Variance Reduction for Stochastic Behaviours in GSHS

1.5.1. Need for a better understanding of IPS based variance reduction

IPS is a variance reduction method that makes use of multi-level splitting. This allows to express the small reach probability of the inner level set as a product of larger reach probabilities for the sequence of enclosing subsets, e.g. (Glasserman et al. 1999). Cérou et al. (2005, 2006) embedded this multi-level splitting approach within the Feynman-Kac factorization equation for strong Markov processes (Del Moral, 2004).

In the IPS approach, multiple particles are simultaneously simulated to estimate the reach probability of the next level (or subset), if started from the preceding level (or subset). Once a particle enters a subset in between, it splits/copies into many independent sub-paths. Many splitting implementations are available in the literature: a fixed-splitting implementation (L'Ecuyer et al., 2006); a fixed-effort implementation (L'Ecuyer et al., 2007); a fixed success implementation (Le Gland and Oudjane, 2006); a fixed probability of success (Cérou and Guyader, 2007). Moreover, (Gerber et al., 2019; Garvels, 2000) compared different splitting strategies. (Garber and Chopin, 2015) introduced a method to improve the efficiency of Monte Carlo methods by combining quasi-random sequences to improve convergence. Cérou et al. (2005, 2006) embedded this multilevel splitting in the Feynman-Kac factorization equation for strong Markov processes (Del Moral, 2004). The IPS approach seems to be the most suitable for rare event estimation in stochastic dynamical systems (Krystul, 2006).

In spite of these results on the analysis of multilevel splitting and IPS, relevant issues remain for a better understanding. One issue is that in order to simplify the problem, researchers assume that each sequence is independent (e.g., Garvels, 2000). However, this assumption does not hold in practice. Another issue is that IPS variance estimate is based on the assumption that the number of simulated particles tends to infinity. Also this assumption is impossible to realize in practice.

1.5.2. Understanding the effect of modelling spontaneous jumps as forced jumps

The continuous-time executions of a GSHS evolve in a hybrid state space under influence of combinations of diffusions, spontaneous jumps and forced jumps. As explained by Lygeros and Prandini (2010), a spontaneous jump in a GSHS can be transformed to a forced jump. This is done as follows. An auxiliary Markov state component q_t, starts at each exit time as an exponentially distributed random variable, subsequently evolves as $dq_t = -\lambda(\theta_t, x_t)dt$ and defines a new exit time upon reaching value zero. Replacement of spontaneous jumps in $\{x_t, \theta_t\}$ with forced jumps when $q_{t-} = 0$ and resampling of $q_{\tau'}$ from an exponential distribution with rate 1 upon reaching the extended exit boundary at stopping time τ'. Common practice in GSHS model specification is to adopt this transformation. As a result the GSHS model specified has no or fewer spontaneous jumps. [Blom et al., 2018] has demonstrated through IPS simulations, for example, that the transformation of a spontaneous jump to a forced jump may have negative effect on the variance reduction performance of IPS. The specific example considered was the probability of a car hitting a wall scenario, and involved a random reaction delay by a human. Two ways in modelling this human reaction delay have been simulated: i) direct simulation of a GSHS execution and ii) transforming the spontaneous jumps of a GSHS into forced jumps, followed by simulating the executions of this transformed version.

The findings for this example show there is a need for a better understanding of the effect of transforming spontaneous jumps in GSHS to forced jumps.

1.5.3. Error Analysis of Sampling per mode strategies in IPS

Switching Diffusion is a subclass of GSHS. It has received increasing attention recently. A prominent feature of these systems is the coexistence of continuous dynamics and discrete events. Applying the multilevel splitting method to Switching Diffusion cannot produce reasonable estimates within reasonable simulation time. The reason is that there may be very few particles in a mode of low probability (such mode is referred to as a "light" mode). This happens because each resampling step tends to make particle copies for modes with high probabilities, as a result of which few or no particle copies are made for "light" modes. Increasing the number of particles should improve but at the cost of significantly increasing simulation time.

To avoid this situation, (Krystul, 2006) proposed the sampling per mode algorithm, which draws a fixed number of N_j particles in each mode j. Using the law of large numbers and the central limit theorem, (Krystul, 2006) have analyzed the convergence of the sampling per mode algorithm. However, the law of large numbers or the central limit theorem only holds when the number of simulated particles tends to infinity. This theoretical issue can have an unpredictable impact on accuracy.

1.5.4. Importance Sampling in Estimation of Reach Probability for GSHS

In the literature, IS has been studied for CTMC [Shahabuddin, 1994] and Diffusions [Glasserman, 2004; Dupuis et al., 2012; Zhang et al., 2014]. Only recently, IS has been developed for PDMP [Chraibi et al., 2019]. This invaluable development has been identified thanks to participation of Dr. Chraibi in a benchmark competition between tools and methods for rare event estimation for stochastic hybrid systems (Abate et al., 2021). The research question to be addressed is if and how these IS results for PDMP can be extended to GSHS, i.e. a PDMP that involves Brownian motion.

1.6. Research Objectives

The above identified research gaps have inspired the overall aim of this thesis and it is:

To develop significant improvements in rare event simulation for GSHS

The following four research objectives are addressed in this thesis to address this statement. These objectives will be solved in each of the chapters of this thesis.

Objective 1. Error Analysis of Multilevel Splitting

Garvels (2000) has proven that fixed assignment splitting works better or is equal to multinomial resampling by assuming that the sets of particles at different levels are independent of each other. Objective 1 is to prove that using fixed assignment splitting in reach probability, IPS dominates in variance reduction over the random assignment methods, Multinomial Resampling, Multinomial Splitting, and Residual Multinomial Splitting without making use of the independence assumption.

Objective 2. Understanding the Effect of Transforming Spontaneous Jumps to Forced Jumps

There are two approaches to simulating GSHS execution: direct simulation and transformation of spontaneous jumps into forced jumps, followed by simulation of the transformed version. Blom et al. (2018) found that the latter approach can produce unexpected effects, such as particle impoverishment. However, the mechanisms behind these effects needed to be sufficiently understood to make a meaningful contribution to simulating a GSHS. Therefore, Objective 2 is to investigate the effects of using the transformed version in an arbitrary GSHS.

Objective 3. Error Analysis of sampling per mode within IPS

Straightforward application of the IPS approach of (Cérou et al., 2002) to rarely switching diffusions has certain limitations, particularly for a few particles in a mode with small conditional probability, i.e., a "light" mode. In such a case, the possible switching between modes is not properly taken into account, which badly affects estimator performance. In order to improve this, (Krystul, 2006) developed the sampling per mode algorithm to cope with large differences in mode weights and proved the accuracy through asymptotic analysis.

Objective 3 is to develop an error analysis approach for IPS that uses sampling per mode and to use this to develop an improvement of sampling per mode strategy within IPS.

Objective 4. Extending Chraibi's IS results for PDMP to GSHS

IS has been well studied in the field of rare event estimation for CTMC and diffusions. These studies address three main issues. The first issue is to characterize the optimal IS strategy. The second issue is to use the characterization of the optimal IS strategy for the development of a parametric family of approximated IS strategies. The third issue is to optimize the parameter values in this family through a minimization of the Kullback-Leibler divergence between the probability laws of the optimal and the approximated IS strategies. Recently, Chraibi et al. (2019) addressed these three steps for a PDMP, which is a GSHS without diffusion. Therefore, Objective 4 is to extend the IS developments by (Chraibi et al., 2019) to a GSHS.

1.7 Thesis Overview

The thesis consists of six chapters. The contents of the remaining chapters are briefly summarized below.

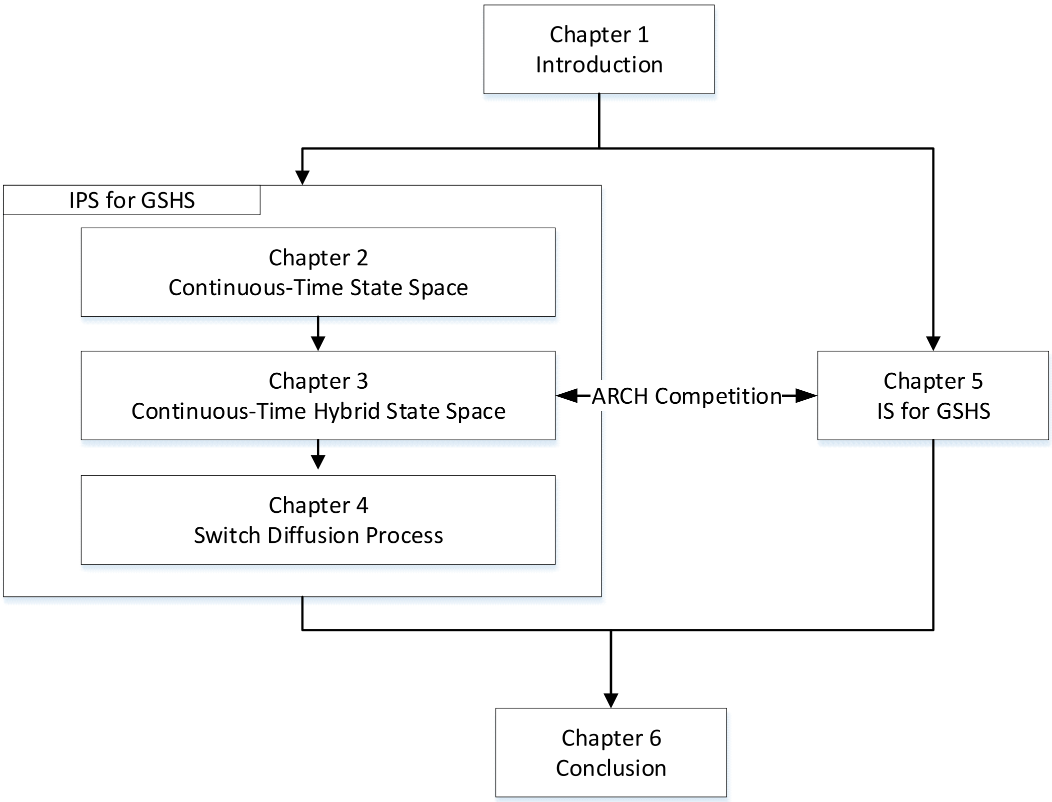

Figure 1.3. The overview of the thesis

Chapter 2 reviews the background of IPS based reach probability estimation for a multi-dimensional diffusion process and characterizes the conditional variances of IPS-based reach probability estimation under four different splitting strategies: multinominal resampling, multinominal splitting, residual multinominal splitting, and fixed assignment splitting. Subsequently, the variance estimates of these four strategies are compared. Numerical evaluations and comparisons of four splitting strategies within IPS are shown. This chapter is based on Ma and Blom (2021).

Chapter 3 examines the effect of transforming spontaneous jumps in a GSHS to forced jumps (referred to as "the transformed version" for convenience). This chapter begins with an overview of the IPS setting for a GSHS. Two methods for simulating GSHS execution are presented: direct simulating GSHS execution and the transformed version. These approaches are then formulated within the IPS framework. The comparison of these two approaches reveals the mechanisms behind the unexpected effects of the transformed version. This chapter is based on the work by Ma and Blom (2022a)

Chapter 4 focuses on the analysis of the sampling per mode strategy for reach probability in GSHS. Firstly, the IPS setting for a GSHS is summarized, followed by the representation of the sampling per mode algorithm within IPS (Krystul et al., 2012). Next, a slightly improved version of Krystul's algorithm is developed, which ensures that the total particle number remains same. This improved version is then compared with the classical IPS approach. Furthermore, several improved versions of the sampling per mode algorithm using different splitting strategies are analysed and compared. This chapter is based on the work by Ma and Blom (2022b).

Chapter 5 extends the optimal IS strategy by Chraibi et al. (2019) for a PDMP to a GSHS. This chapter formulates the reach probability estimation for a GSHS using IS and characterizes the optimal IS strategy. However, this strategy is only of theoretical use in practice. As a result, a parametric family of approximated IS strategies for a GSHS is developed using this characterization . With the Kullback-Leibler divergence, the best parameter value in this family is determined. This chapter is based on Ma and Blom (2023).

Finally, Chapter 6 presents conclusions and possible directions for future research.

1.8 References

Abate, A., Blom, H.A.P., Bouissou, M., Cauchi, N., Chraibi, H., Delicaris, J., Haesaert, S., Hartmanns, A., Khaled, M., Lavaei, A., Ma, H., Mallik, K., Niehage, M., Remke, A., Schupp, S., Shmarov, F., Soudjani, S., Thorpe, A., Turcuman, V., and Zuliani, P., (2021). ARCH-COMP21 category report: Stochastic models. In Goran Frehse and Matthias Althoff, editors, 8th International Workshop on Applied Verification of Continuous and Hybrid Systems (ARCH21), Volume 80 of EPiC Series in Computing, pp. 55–89.

Blom, H.A.P., Bakker, G.J., Blanker, P.J.G., Daams, J., Everdij, M.H.C., Klompstra, M.B., (2001), Accident risk assessment for advanced ATM, In: Air Transportation Systems Engineering, edited by G.L. Donohue and A.G. Zellweger, Vol. 193 in Progress in Astronautics and Aeronautics, Paul Zarchan, Editor-in-Chief, Chapter 29, pp. 463-480.

Blom, H.A.P., Ma, H. and Bakker, G.J., (2018). Interacting particle system-based estimation of reach probability for a generalized stochastic hybrid system. Proc. Conference Analysis and Design of Hybrid Systems (ADHS), Oxford, UK, 11-13 July 2018, IFAC Papers Online 51-16 (2018), pp. 79-84.

Blom, H.A.P., Krystul, J., and Bakker, G.J., Klompstra, M.B., Klein Obbink, B., (2007). Free flight collision risk estimation by sequential MC simulation. Eds: C.G, Cassandras and J. Lygeros, Stochastic hybrid systems, Taylor & Francis, pp.249-281.

Bucklew, J.A., (2004). Introduction to rare event simulation, Springer, New York.

Botev, Z.I. and Kroese, D.P., (2008). An efficient algorithm for rare-event probability estimation, combinatorial optimization, and counting. Methodology and Computing in Applied Probability 10(4), pp. 471-505.

Bujorianu, M.L., (2012). Stochastic Reachability Analysis of Hybrid Systems. Springer Science & Business Media.

Bujorianu, M.L. and Lygeros, J., (2006). Toward a general theory of stochastic hybrid systems. Eds: Blom, H.A.P. and Lygeros, J. Stochastic hybrid systems, Springer, Berlin, pp. 3-30.

Cérou, F. and Guyader, A., (2007). Adaptive multilevel splitting for rare event analysis. Stochastic Analysis and Applications, 25(2), pp. 417–443.

Cérou, F., Legland, F., Del Moral, P., and Lezaud, P., (2005). Limit theorems for the multilevel splitting algorithm in the simulation of rare events. In Proceedings of Winter Simulation Conference. Orlando, FL, pp. 682-691.

Cérou, F., Del Moral, P., Legland, F., and Lezaud, P., (2006). Genetic genealogical models in rare event analysis. Latin American J. of Probability & Mathematical Statistics, 1, pp. 181-203.

Chraibi, H., Dutfoy, A., Galtier, T. and Garnier, J., (2019). On the optimal importance process for piecewise deterministic Markov process. ESAIM: Probability and Statistics, 23, pp. 893-921.

Davis, M.H.A., (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. Journal of the Royal Statistical Society: Series B (Methodological), 46(3), pp.353-376.

Del Moral, P., (2004). Feynman-Kac formulae: genealogical and interacting particle systems with applications, Springer.

Everdij, M.H.C. and Blom, H.A.P., (2006). Hybrid Petri Nets with diffusion that have into-mappings with Generalized Stochastic Hybrid Processes, In: Stochastic Hybrid Systems, eds: H.A.P. Blom and J. Lygeros, Springer, pp. 31-63.

Everdij, M.H.C., Klompstra, M.B., Blom, H.A.P. and Klein Obbink, B.K., (2006). Compositional specification of a multi-agent system by stochastically and dynamically coloured Petri nets. In: Stochastic Hybrid Systems: Theory and safety critical applications, eds: Blom and Lygers, Springer, pp. 325-350.

FAA (2021), National Airspace System Capital Investment Plan Fiscal Years 2023-2027, FAA Air Traffic Organization.

FAA (2023), Air Traffic by the numbers, FAA Air Traffic Organization.

Garvels, M.J.J., (2000). The splitting method in rare event simulation, Ph.D. Thesis, University of Twente, The Netherlands.

Gerber, M. and Chopin, N., (2015). Sequential quasi monte carlo. Journal of the Royal Statistical Society Series B: Statistical Methodology, 77(3), pp.509-579.

Gerber, M., Chopin, N, and Whiteley, N., (2019). Negative association, ordering and convergence of resampling methods. The Annals of Statistics 47.4, pp. 2236-2260.

Ghosh, M.K., Arapostathis, A., Marcus, S.L., (1997). Ergodic Control of Switching Diffusions. SIAM J. Control Optim. 35(6), pp. 1952-1988.

Glasserman, P., Heidelberger, P., Shahabuddin, P., and Zajic, T., (1999). Multilevel Splitting for estimating rare event probabilities. *Operations Research*, *47*(4), pp. 585-600.

Glasserman, P., (2004). Monte Carlo methods in financial engineering. Springer, New York.

Hespanha J., (2005). A model for stochastic hybrid systems with application to communication networks. Nonlinear Anal 62(8), pp. 1353–1383.

Hu, J., Lygeros, J. and Sastry, S., (2000). Towards a theory of stochastic hybrid systems. In *Hybrid Systems: Computation and Control: Third International Workshop, HSCC 2000 Pittsburgh, PA, USA, March 23–25, 2000,* Springer, Berlin, pp. 160-173.

Krystul, J., (2006). Modelling of Stochastic Hybrid Systems with applications to accident risk assessment, PhD Thesis, University of Twente, The Netherlands, 2006.

Krystul, J., Le Gland, F., Lezaud, P., (2012). Sampling per mode for rare event simulation in switching diffusions, Stochastic Processes and their Applications, Vol. 122, pp. 2939-2667.

Le Gland, F. and Oudjane, N., (2006) A sequential algorithm that keeps the particle system alive. In: Stochastic Hybrid Systems: Theory and Safety Critical Applications, eds: Blom and Lygeros, Lecture Notes in Control and Information Sciences, Vol. 337, Springer, Berlin, pp. 351–389.

Lygeros, J. and Prandini, M., (2010). Stochastic hybrid systems: a powerful framework for complex, large scale applications. European Journal of Control, 16(6), pp.583-594.

L'Ecuyer, P., Demers, V., and Tuffin, B., (2006). Splitting for rare-event simulation. In Proceedings of the 2006 Winter Simulation Conference, Monterey, pp. 137–148.

L'Ecuyer, P., Demers, V., and Tuffin, B., (2007). Rare events, splitting, and quasi-Monte Carlo. ACM Transactions on Modeling and Computer Simulation, 17(2). Article 9.

L'Ecuyer, P., LeGland, F., Lezaud, P. and Tuffin, B., (2009), Splitting techniques, Eds: G. Rubino, B. Tuffin, Rare event simulation using Monte Carlo methods, Wiley, pp. 39-61.

Ma, H. and Blom, H.A.P., (2022). Random Assignment Versus Fixed Assignment in Multilevel Importance Splitting for Estimating Stochastic Reach Probabilities. Methodol Comput Appl Probab 24, 2313–2338. https://doi.org/10.1007/s11009-021-09892-4

Ma, H. and Blom, H.A.P., (2023). Interacting particle system based estimation of reach probability of general stochastic hybrid systems. *Nonlinear Analysis: Hybrid Systems*, *47*, pp.101-303.

Mao X. and Yuan C., (2006). Stochastic differential equations with Markovian switching. Imperial College Press.

Morio, J. and Balesdent, M. (eds.) (2016). Estimation of Rare Event Probabilities in Complex Aerospace and Other Systems, Woodhead Publishing.

Oksendal, B., and Sulem, A., (2005), Applied Stochastic Control of Jump Diffusions, Springer.

Prandini, M., Blom, H.A.P. and Bakker, G.J., (2011). Air traffic complexity and the interacting particle system method: An integrated approach for collision risk estimation. Proceedings of 2011 American Control Conference, pp. 2154-2159. IEEE.

Shahabuddin, P., (1994). Importance sampling for the simulation of highly reliable Markovian systems. Management Science, 40(3), pp.333-352.

Stroeve, S.H., Bakker, G.J., Everdij, M.H.C., Trezza, M., Cañizares, C.V., Calle, I.M., and García, D.R., (2021), D(emo)-CRAT Demonstration and Final Report, NLR

Railsback, S.F., and Grimm, V., (2012), Agent-based and individual based modeling: A practical introduction, Princeton, NJ: Princeton University Press.

RTCA (1995), Final report of RTCA Task Force 3; Free Flight implementation, RTCA Inc., Washington DC, October.

Rubinstein, R., (2010), Randomized algorithms with splitting: Why the classic randomization algorithms do not work and how to make them work, Methodology and Computing in Applied Probability, Vol. 12, pp. 1-50.

Yin, G.G. and Zhu, C., (2010). Hybrid Switching Diffusions. Springer New York, NY

2

Random Assignment vs. Fixed Assignment in Multilevel Importance Splitting for Estimating Stochastic Reach Probabilities

This chapter focuses on estimating Reach Probability of a closed unsafe set by a stochastic process. A well-developed approach is to make use of multi-level MC simulation, which consists of encapsulating the unsafe set by a sequence of increasing closed sets and conducting a sequence of MC simulations to estimate the reach probability of each inner set from the previous set. An essential step is to copy (split) particles that have reached the next level (inner set) prior to conducting an MC simulation to the next level. The aim of this chapter is to prove that the variance of the multi-level MC estimated reach probability under fixed assignment splitting is smaller than or equal to that under random assignment splitting methods. The approaches are illustrated for a geometric Brownian motion example.

This chapter has been published as "Ma, H., Blom, H.A.P., Random Assignment Versus Fixed Assignment in Multilevel Importance Splitting for Estimating Stochastic Reach Probabilities. Methodology and Computing in Applied Probability 24, 2313–2338 (2022). https://doi.org/10.1007/s11009-021-09892-4"

2.1. Introduction

Evaluating the reach probability of an unsafe set is well-studied in the domains of control and safety verification of complex safety critical system designs. In the control domain, the focus is on synthesizing a control policy such that a safety critical systems stays away from the unsafe set with a high probability (Alur et al., 2000; Prandini and Hu, 2007). From this control synthesis perspective, it makes good sense to adopt model abstractions in combination with an over-approximation of the unsafe set (Julius and Pappas, 2009; Abate et al., 2011; Di Benedetto et al., 2015). In safety verification of complex safety critical system design, reach probability of the unsafe set is commonly evaluated using statistical simulation techniques, e.g. air traffic (Blom et al., 2006, 2007a), actuarial risks (Asmussen, and Albrecher, 2010), random graphs (Bollobás, 2010), communication network reliability (Robert, 2003).

To evaluate very small reach probabilities, common practice is to make use of methods to reduce variance for a given computational effort. Literature on variance reduction distinguishes two main approaches: Importance Sampling (IS) and Importance Splitting (ISp). IS draws samples from a reference stochastic system model in combination with an analytical compensation for sampling from the reference model instead of the intended model. Bucklew (2004) gives an overview of IS and analytical compensation mechanisms. For complex models, analytical compensation mechanisms typically fall short and multi-level ISp is the preferred approach (e.g. Botev and Kroese, 2008; L'Ecuyer et al., 2009; Rubinstein, 2010; Morio and Balesdent, 2016).

In multi-level splitting, the safe set, or target set, i.e. the set for which the reach probability has to be estimated, is enclosed by a series of strictly increasingly (nested/enclosing) subsets. This multi-level setting allows one to express the small reach probability of the inner level set as a product of larger reach probabilities for the sequence of enclosing subsets (see e.g. Glasserman et al, 1998, 1999). Cérou et al. (2005, 2006) embedded this multi-level splitting in the Feynman-Kac factorization equation for strong Markov processes (Del Moral, 2004). This Feynman-Kac setting subsequently supported the evaluation of the reach probability through sequential Monte Carlo simulation in the form of an Interacting Particle System (IPS), including characterization of asymptotic behaviour (Cérou et al., 2006).

Particle splitting (copying) of N_S successful particles to $N_P \geq N_S$ particles can be done in multiple ways (e.g. Garvels and Kroese, 1998; Cérou et al., 2006; L'Ecuyer et al., 2007; L'Ecuyer et al., 2009). The classical approach is Multinomial Resampling, i.e. drawing the N_P particles at random, with replacement, from the N_s successful particles. Cérou et al. (2006) propose the alternative of adding to the set of N_S successful particles, $N_P - N_S$ random drawings (with replacement) from the N_s successful particles; this we refer to as Multinomial Splitting. A third approach is fixed assignment splitting, i.e. copying each of the N_S successful particles as much as possible the same number of times. Following (L'Equyer et al., 2009), fixed assignment splitting is accomplished in two steps. During the first step each successful particle is copied $\lfloor N_p / N_s \rfloor$ times. During the second step, the residual $N_p - N_s \lfloor N_p / N_s \rfloor$ particles are randomly chosen (without replacement) from the set of successful particles, and these are added to the set of copies from the first step. A fourth approach is residual multinomial splitting, i.e. after the first step of fixed assignment splitting, the residual $N_p - N_s \lfloor N_p / N_s \rfloor$ particles are randomly chosen (with replacement) from the N_s successful particles.

Under restrictive assumptions, Garvels (2000) has proven that fixed assignment splitting works better or equal to multinomial resampling. The key assumption is that the sets of particles at different levels are independent of each other. In IPS for filtering studies, e.g. (Del Moral et al., 2001; Gerber et al., 2019), multi-level Feynman-Kac analysis has been used to make variance comparisons between different particle resampling methods. Through mapping the filtering IPS results of Del Moral et al. (2001) to the reach probability IPS, Cérou et al. (2006) argue that multinominal resampling adds extra randomness to multinomial splitting, as a result of which the multinomial splitting has a variance advantage over multinomial resampling. Through mapping the filtering IPS results of Gerber et al. (2019) to the reach probability IPS, it is clear that residual multinomial splitting has a variance advantage over multinomial resampling. Gerber et al. (2019) also conclude that existing multi-level Feynman-Kac analysis falls short in handling random drawings without replacement, as is done in the second step of fixed assignment splitting.

The main objective of this chapter is to prove that using fixed assignment splitting in reach probability IPS dominates in variance reduction over the random assignment methods: Multinomial Resampling, Multinomial Splitting and Residual Multinomial Splitting. These proofs do not make use of the independence assumption of Garvels (2000). The stochastic process considered is a multi-dimensional diffusion process that is pathwise continuous. The effect of different splitting methods is also illustrated in reach probability estimation for a geometric Brownian motion example.

This chapter is organized as follows. Section 2.2 reviews the background of IPS based reach probability estimation for a multi-dimensional diffusion process. Section 2.3 characterizes the conditional variances of IPS based reach probability estimation under multinominal resampling, multinominal splitting, residual multinominal splitting and fixed assignment splitting. Section 2.4 proves that fixed assignment splitting has a variance advantage over these other three ways of splitting. Section 2.5 presents a case study based on a geometric Brownian motion for evaluating and comparing multinominal resampling, multinomial splitting and fixed assignment splitting. Section 2.6 draws conclusions.

2.2. IPS based reach probability estimation

2.2.1. Reach probability of multi-dimensional diffusion

For the rest of the chapter, we define all stochastic processes on a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$. The problem is to estimate the probability γ that a $\mathbb{R}^n$-valued pathwise continuous diffusion process $\{x_t\}$ reaches a closed subset $D \subset \mathbb{R}^n$ within finite period $[0,T]$, i.e.

$$\gamma = P(\tau < T) \tag{2.1}$$

with τ the first hitting time of D by $\{x_t\}$:

$$\tau = \inf\{t > 0, x_t \in D\} \tag{2.2}$$

Remark: Cérou et al. (2006) and L'Equyer et al. (2009) also address the more general situation that T is a P-a.s. finite stopping time.

2.2.2. Multi-level factorization

If the reach probability γ in (2.1) is too small, then a straightforward MC estimator requires a considerable amount of samples. To overcome this, we introduce a nested sequence of closed subsets D_k of $\mathbb{R}^n$ to factorize the reach probability γ, such that

$D = D_m \subset D_{m-1} \subset \cdots \subset D_1 \subset D_0 = \mathbb{R}^n$ and $P\{x_0 \in D_1\} = 0$. Let τ_k be the first moment in time that $\{x_t\}$ reaches D_k, i.e.

$$\tau_k = \inf\{t > 0;\ x_t \in D_k \vee t \geq T\} \tag{2.3}$$

Then, we define $\{0,1\}$-valued random variables $\{\chi_k, k = 0,..,m\}$ as follows:

$$\begin{aligned} \chi_k &= 1, \text{ if } \tau_k < T \text{ or } k = 0 \\ &= 0, \text{ else} \end{aligned} \tag{2.4}$$

By using this χ_k definition, the factorization becomes (Cérou et al, 2006):

$$\gamma = \prod_{k=1}^{m} \gamma_k \tag{2.5}$$

with $\gamma_k \triangleq \mathrm{P}(\chi_k = 1 | \chi_{k-1} = 1) = \mathrm{P}(\tau_k < T | \tau_{k-1} < T)$.

2.2.3. Recursive estimation of the multi-level factors

By using the strong Markov property of $\{x_t\}$, we can develop a recursive estimation of γ using the factorization in (2.5). First, we define $\xi_k \triangleq (\tau_k, x_{\tau_k})$, $Q_k \triangleq (0,T) \times D_k$, for $k = 1,\ldots,m$, and the following conditional probability measure $\pi_k(B)$ for an arbitrary Borel set B of $\mathbb{R}^{n+1}$:

$$\pi_k(B) \triangleq P(\xi_k \in B \mid \xi_k \in Q_k)$$

Cérou et al. (2006) show that π_k is a solution of the following recursion of transformations:

$$\pi_{k-1}(\cdot) \xrightarrow{\text{I. mutation}} p_k(\cdot) \xrightarrow{\text{III. selection}} \pi_k(\cdot)$$
$$\downarrow \text{ II. conditioning}$$
$$\gamma_k$$

where $p_k(B)$ is the conditional probability measure of $\xi_k \in B$ given $\xi_{k-1} \in Q_{k-1}$ i.e.

$$p_k(B) \triangleq P(\xi_k \in B \mid \xi_{k-1} \in Q_{k-1})$$

Because $\{x_t\}$ is a strong Markov process, $\{\xi_k\}$ is a Markov sequence. Therefore the mutation transformation (I) satisfies a Chapman-Kolmogorov equation prediction for ξ_k:

$$p_k(B) = \int_{\mathbb{R}^{n+1}} p_{\xi_k|\xi_{k-1}}(B \mid \xi)\pi_{k-1}(d\xi) \text{ for all } B \in \beta(\mathbb{R}^{n+1}) \tag{2.6}$$

For the conditioning transformation (II) this means:

$$\gamma_k = \mathrm{P}(\tau_k < T \mid \tau_{k-1} < T) = \int_{\mathbb{R}^{n+1}} 1_{\{\xi \in Q_k\}} p_k(d\xi). \tag{2.7}$$

Hence, selection transformation (III) satisfies:

$$\pi_k(B) = \frac{\int_B 1_{\{\xi \in Q_k\}} p_k(d\xi)}{\int_{\mathbb{R}^{n+1}} 1_{\{\xi' \in Q_k\}} p_k(d\xi')} = [\int_B 1_{\{\xi \in Q_k\}} p_k(d\xi)] / \gamma_k \tag{2.8}$$

With this, the γ_k terms in (2.5) are characterized as solutions of a recursive sequence of mutation equation (2.6), conditioning equation (2.7), and selection equation (2.8).

2.2.4. IPS algorithmic steps

Following Cérou et al (2006), equations (2.5)-(2.8) yield the IPS algorithmic steps for the numerical estimation of γ:

$$\bar{\pi}_{k-1}(\cdot) \xrightarrow{\text{I. mutation}} \bar{p}_k(\cdot) \xrightarrow{\text{III. selection}} \tilde{\pi}_k(\cdot) \xrightarrow{\text{IV. splitting}} \bar{\pi}_k(\cdot)$$

$$\downarrow \text{ II. conditioning}$$

$$\bar{\gamma}_k$$

A set of N_P particles is used to form empirical density approximations $\bar{\gamma}_k$, $\bar{p}_k$ and $\bar{\pi}_k$ of γ_k, p_k and π_k respectively. By increasing the number N_P of particles in a set, the errors in these approximations will decrease. When simulating particles from Q_{k-1} to Q_k, a fraction $\bar{\gamma}_k$ of the simulated particle trajectories only will reach Q_k within the time period $[0,T]$ considered; these particles form $\tilde{\pi}_k$. Prior to starting the next IPS cycle with N_P particles, $(N_P - N_{S_k})$ copies (also called splittings) from the N_{S_k} successful particles in $\tilde{\pi}_k$ are added to $\bar{\pi}_k$. In the next sections, we consider four ways of splitting: multinominal resampling, multinominal splitting, residual multinominal splitting and fixed assignment splitting.

Under Multinomial Splitting, Cérou et al. (2006) prove that $\bar{\gamma}$ forms an unbiased γ estimate, i.e.

$$\mathbb{E}\{\bar{\gamma}\} = \mathbb{E}\left\{\prod_{k=1}^{m} \bar{\gamma}_k\right\} = \prod_{k=1}^{m} \mathbb{E}\{\bar{\gamma}_k\} = \prod_{k=1}^{m} \gamma_k = \gamma \tag{2.9}$$

Moreover, Cérou et al. (2006) derive second and higher order asymptotic bounds for the error $(\bar{\gamma} - \gamma)$ based on multi-level Feynman Kac analysis, e.g. Del Moral (2004; Theorem 12.2.2).

2.3. Conditional variance characterizations

In this section, conditional characterizations of the variance of $\bar{\gamma}_k$ are developed for IPS using multinomial resampling (MR), multinomial splitting (MS), residual multinomial splitting (RMS) and fixed assignment splitting (FAS), respectively.

2.3.1. IPS using multinomial resampling

In IPS using multinomial resampling, N_P offspring are cloned randomly from $\tilde{\pi}_k$. The resulting algorithm of IPS with multinomial resampling, starting from ξ_{k-1}^i, is described in Algorithm 2.1 below.

In order to gain a better understanding of the probabilistic characteristics of the particles that reached a level, we now characterize the conditional distribution of particles that reach level k+1, given that at level k the i-th successful particle $\tilde{\xi}_k^i$ is copied K_k^i times, $i = 1,..,N_{S_k}$.

Proposition 2.1: If $N_{S_k} > 0$ and K_k^i, with i=1,2,...,N_{S_k}, denote the number of particles that copies $\tilde{\xi}_k^i$ at level k. Then the number $Y_{k+1}^{k,i}$, of the K_k^i particle copies of $\tilde{\xi}_k^i$ that reach level k+1, has a conditional Binomial distribution of size K_k^i and success probability $\gamma_{k+1}(\tilde{\xi}_k^i)$, i.e.

$$p_{Y_{k+1}^{k,i}|K_k^i,\tilde{\xi}_k^i}(n; K_k^i, \tilde{\xi}_k^i) = \text{Bin}(n; K_k^i, \gamma_{k+1}(\tilde{\xi}_k^i)) \tag{2.10}$$

with

$$\gamma_{k+1}(\tilde{\xi}_k^i) \triangleq \text{P}(\tau_{k+1} < T \mid \xi_k = \tilde{\xi}_k^i) \tag{2.11}$$

Proof: See Appendix 2.A.

Algorithm 2.1. IPS using multinomial resampling

Input: Initial measure π_0, end time T, decreasing sequence of closed subsets $D_k = \{x_t \in \mathbb{R}^n\}$, $D_{k-1} \supset D_k$, $k = 1,..,m.$ Also $D_0 = \mathbb{R}^n$, $Q_k = (0,T) \times D_k$ and number of particles N_P.

Output: Estimated reach probability $\overline{\gamma}$ collects

0. Initiation: Generate N_P particles $\xi_0^i \sim \pi_0$, $i = 1,..,N_P$, i.e., $\overline{\pi}_0(\cdot) = \sum_{i=1}^{N_p} \frac{1}{N_p} \delta_{\{\xi_0^i\}}(\cdot)$, with Dirac δ. Set $k = 1$.

I. Mutation: $\overline{p}_k(.) = \sum_{i=1}^{N_p} \frac{1}{N_p} \delta_{\{\overline{\xi}_k^i\}}(\cdot)$, where $\overline{\xi}_k^i$ is obtained through simulating the strong Markov process starting from ξ_{k-1}^i.

II. Conditioning: $N_{S_k} = \sum_{i=1}^{N_p} 1(\overline{\xi}_k^i \in Q_k)$ and $\overline{\gamma}_k = \frac{N_{S_k}}{N_p}$. If $N_{S_k} = 0$, then $\overline{\gamma}_{k'} = 0$, $k' \in \{k,..,m\}$ and go to Step V.

III. Selection: $\tilde{\pi}_k(.) = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \delta_{\{\tilde{\xi}_k^i\}}(\cdot)$, with $\{\tilde{\xi}_k^j\}_{j=1}^{N_{S_k}}$ the collection of $\overline{\xi}_k^i \in Q_k$, $i = 1,..,N_P$.

IV. Splitting: $\overline{\pi}_k(.) = \sum_{i=1}^{N_p} \frac{1}{N_p} \delta_{\{\xi_k^i\}}(.)$, with $\xi_k^i \sim \tilde{\pi}_k(.)$.

V. If $k < m$, then $k := k+1$ and go to Step I, else $\overline{\gamma} = \prod_{k=1}^{m} \overline{\gamma}_k$

Theorem 2.1: If $N_{S_k} \geq 1$ and K_k^i, $i = 1,..N_{S_k}$, denotes the number of copies made of the *i*-th successful particle $\tilde{\xi}_k^i$ during the splitting step at level *k* of the IPS algorithm, then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p} \sum_{i=1}^{N_{S_k}} \left[\mathbb{E}\left\{K_k^i \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} \gamma_{k+1}(\tilde{\xi}_k^i) \right] \tag{2.12}$$

$$\begin{aligned} \mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} &= \frac{1}{N_p^2} \sum_{i=1}^{N_{S_k}} \left[\mathbb{E}\left\{K_k^i \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} \gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right) \right] \\ &+ \frac{1}{N_p^2} \sum_{i=1}^{N_{S_k}} \left[\mathrm{Var}\left\{K_k^i \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} \gamma_{k+1}(\tilde{\xi}_k^i)^2 \right] \\ &+ \frac{1}{N_p^2} \sum_{i=1}^{N_{S_k}} \sum_{\substack{i'=1 \\ i' \neq i}}^{N_{S_k}} \left[\mathrm{Cov}\left\{K_k^i, K_k^{i'} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} \gamma_{k+1}(\tilde{\xi}_k^i) \gamma_{k+1}(\tilde{\xi}_k^{i'}) \right] \end{aligned} \tag{2.13}$$

Proof: See Appendix 2.A.

Proposition 2.2: If $N_{S_k} \geq 1$, and we use multinomial resampling at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \gamma_{k+1}(\tilde{\xi}_k^i) \tag{2.14}$$

$$\begin{aligned} \mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} &= \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right) \right] \\ &+ \frac{1}{N_p N_{S_k}} \left[\sum_{i=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2 \right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i) \gamma_{k+1}(\tilde{\xi}_k^{i'})\right) \right] \right] \end{aligned} \tag{2.15}$$

Proof: See Appendix 2.A.

2.3.2. IPS using multinomial splitting

IPS using multinomial splitting follows the steps of Algorithm 2.1, except for splitting step IV. Now each particle in $\tilde{\pi}_k(\cdot)$ is first copied once, and then $\left(N_p - N_{S_k}\right)$ offspring are cloned randomly from $\tilde{\pi}_k(\cdot)$ (Cérou et al., 2006, Section 3.2, p189). This multinomial splitting in IPS step IV is specified in Algorithm 2.2.

Algorithm 2.2. Multinomial Splitting in IPS step IV

IV. Splitting: $\xi_k^i = \tilde{\xi}_k^i$ for $i = 1,...,N_{S_k}$; then $\xi_k^i \sim \tilde{\pi}_k(.)$ for $i = N_{S_k}+1,...,N_p$. Each particle receives weight $1/N_p$.

In IPS using multinomial splitting, all particles have the same weight at any given level. Each particle is simulated until it reaches the first subset Q_1. Then $\sum_{i=1}^{N_p} 1(\overline{\xi}_1^i \in Q_1)$ is the number of particles that have reached the first subset Q_1. The fraction $\overline{\gamma}_1 = \sum_{i=1}^{N_p} \frac{1}{N_p} 1(\overline{\xi}_1^i \in Q_1)$ is an unbiased estimate of $\gamma_1 = \mathrm{P}(\tau_1 < T)$. To maintain a sufficiently large population of particles, in IPS step IV $(N_P - N_{S_k})$ copies of these $\sum_{i=1}^{N_p} 1(\overline{\xi}_1^i \in Q_1)$ successful particles are added to the set of N_{S_k} successful particles. During the next IPS cycle each new particle is simulated until it reaches the second subset Q_2. Again, the fraction $\overline{\gamma}_2 = \sum_{i=1}^{N_p} \frac{1}{N_p} 1(\overline{\xi}_2^i \in Q_2)$ of $\sum_{i=1}^{N_p} 1(\overline{\xi}_2^i \in Q_2)$ particles that reach the second subset Q_2 is a natural estimate of $\gamma_2 = \mathrm{P}(\tau_2 < T \mid \tau_1 < T)$. This cycle is repeated until particles reach the last subset Q_m. The fraction $\overline{\gamma}_k = \sum_{i=1}^{N_p} \frac{1}{N_p} 1(\overline{\xi}_k^i \in Q_k)$ of particles that have timely reached the k-th subset from the preceding subset is an unbiased estimate of $\gamma_k = \mathrm{P}(\tau_k < T \mid \tau_{k-1} < T)$. From eq. (2.9) we know that the product of these m fractions is an unbiased estimate of $\gamma = \mathrm{P}(\tau_m < T)$.

It is straightforward to verify that Proposition 2.1 and Theorem 2.1 also hold true using multinomial splitting in IPS step IV.

Proposition 2.3: If $N_{S_k} \geq 1$, and we use multinomial splitting at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \gamma_{k+1}(\tilde{\xi}_k^i) \tag{2.16}$$

$$\begin{aligned}\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} &= \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] \\ &+ \frac{\left(N_p - N_{S_k}\right)}{N_p{}^2 N_{S_k}} \left[\sum_{i=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right]\right]\end{aligned} \tag{2.17}$$

Proof: See Appendix 2.A.

2.3.3. IPS using residual multinomial splitting

IPS using residual multinomial splitting follows the steps of Algorithm 2.1 with a new Step IV. Now each successful particle is first copied $\alpha_k = \lfloor N_p / N_{S_k} \rfloor$ times, and then residual $\left(N_p \bmod N_{S_k}\right)$ particles are randomly drawn from $\tilde{\pi}_k(\cdot)$. The residual multinomial splitting step IV is specified in Algorithm 2.3 below.

Algorithm 2.3. Residual multinomial splitting in IPS step IV

IV. Splitting: $\xi_k^i = \tilde{\xi}_k^i$ for $i = 1,...,N_{S_k}$;

$\xi_k^{i+N_{S_k}} = \tilde{\xi}_k^i$ for $i = 1,...,N_{S_k}$;

...

$\xi_k^{i+\left\lfloor \frac{N_p}{N_{S_k}} - 1 \right\rfloor N_{S_k}} = \tilde{\xi}_k^i$ for $i = 1,...,N_{S_k}$;

$\xi_k^{i+\left\lfloor \frac{N_p}{N_{S_k}} \right\rfloor N_{S_k}} \sim \tilde{\pi}_k(.)$ for $i = 1,...,N_p - \left\lfloor \frac{N_p}{N_{S_k}} \right\rfloor N_{S_k}$.

Each particle receives weight $1/N_p$.

Straightforward verification shows that Proposition 2.1 and Theorem 2.1 also hold true when using residual multinomial splitting in IPS step IV.

Proposition 2.4: If $N_{S_k} \geq 1$, and we use residual multinomial splitting at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \gamma_{k+1}(\tilde{\xi}_k^i) \tag{2.18}$$

$$\begin{aligned}\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} &= \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] \\ &+ \frac{\left(N_p \bmod N_{S_k}\right)}{{N_p}^2 N_{S_k}} \cdot \left[\sum_{i=1}^{N_{S_k}}\left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right]\right]\end{aligned} \tag{2.19}$$

Proof: See Appendix 2.A.

2.3.4. IPS using fixed assignment splitting

When using fixed assignment splitting, each particle in $\tilde{\pi}_k(\cdot)$ is copied as much as possible the same number of times. This is applied by first copying each particle $\lfloor N_p/N_{S_k} \rfloor$ times, and then making $\left(N_p \bmod N_{S_k}\right)$ copies from distinct particles chosen at random (without replacement). So the chosen particles would be copied $\lfloor N_p/N_{S_k} \rfloor + 1$ times (L'Ecuyer et al., 2006; L'Ecuyer et al., 2007). The Fixed Assignment splitting Step IV is specified in Algorithm 2.4 below.

Algorithm 2.4. Fixed assignment splitting in IPS step IV

IV. Splitting: $\{\tilde{\tilde{\xi}}_k^j\}_{j=1}^{N_{S_k}}$ is a random permutations of $\{\tilde{\xi}_k^j\}_{j=1}^{N_{S_k}}$.

Copy: $\xi_k^i = \tilde{\tilde{\xi}}_k^i$ for $i = 1,...,N_{S_k}$;

$\xi_k^{i+N_{S_k}} = \tilde{\tilde{\xi}}_k^i$ for $i = 1,...,N_{S_k}$;

...

$\xi_k^{i+\left\lfloor \frac{N_p}{N_{S_k}} - 1 \right\rfloor N_{S_k}} = \tilde{\tilde{\xi}}_k^i$ for $i = 1,...,N_{S_k}$;

$\xi_k^{i+\left\lfloor \frac{N_p}{N_{S_k}} \right\rfloor N_{S_k}} = \tilde{\tilde{\xi}}_k^i$ for $i = 1,...,N_p - \left\lfloor \frac{N_p}{N_{S_k}} \right\rfloor N_{S_k}$.

Each particle receives weight $1/N_p$.

Straightforward verification shows that Proposition 2.1 and Theorem 2.1 also hold true using fixed assignment splitting in IPS step IV.

Proposition 2.5: If $N_{S_k} \geq 2$, and we use fixed assignment splitting at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \gamma_{k+1}(\tilde{\xi}_k^i) \tag{2.20}$$

$$\begin{aligned} \mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} &= \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] \\ &+ \frac{\left(N_p \bmod N_{S_k}\right)\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right]}{N_p^{\,2} N_{S_k}\left(N_{S_k}-1\right)} . \left[\sum_{i=1}^{N_{S_k}}\left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right]\right] \end{aligned} \tag{2.21}$$

with mod representing modulo operation.

Proof: See Appendix 2.A.

2.4. Comparison of variances

This section proves that IPS using fixed assignment splitting has variance advantage over IPS under each of the three Random assignment splitting methods MR, MS and RMS. This is accomplished through a sequence of three of Theorems. Theorem 2.2 compares the four splitting strategies at a single level only. Theorem 2.3 considers multiple levels, with difference in splitting strategies at a single level and no differences in splitting strategy at the other levels. Theorem 2.4 uses Theorem 2.3 to complete the comparison of IPS under different ways of splitting.

Theorem 2.2: Given successful particles $\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}$ at IPS level *k* with $N_{S_k} \geq 1$. The dominance of the four splitting methods (MR, MS, RMS, FAS) in terms of $\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}$ is:

$$\mathrm{V}_{FAS}^k \leq \mathrm{V}_{RMS}^k \leq \mathrm{V}_{MS}^k \leq \mathrm{V}_{MR}^k \tag{2.22}$$

Theorem 2.3: If IPS levels 1 to *k*-1 make use of the same type of splitting (either MR, MS, RMS or FAS), then the dominance of the four splitting methods at level *k*, in terms of $\mathrm{Var}\left\{\prod_{k'=1}^{k}\overline{\gamma}_{k'}\right\}$ satisfies:

$$\mathrm{V}_{FAS\_k} \leq \mathrm{V}_{RMS\_k} \leq \mathrm{V}_{MS\_k} \leq \mathrm{V}_{MR\_k} \tag{2.23}$$

Theorem 2.4: Under the same type of Splitting (either MR, MS, RMS or FAS) at all levels, then the dominance of the four splitting methods in terms of $\mathrm{Var}\{\overline{\gamma}\}$ satisfies:

$$\mathrm{V}_{FAS} \leq \mathrm{V}_{RMS} \leq \mathrm{V}_{MS} \leq \mathrm{V}_{MR} \tag{2.24}$$

Proof of Theorem 2.2:

From inequality of arithmetic and geometric means we know:

$$\sum_{i=1}^{N_{S_k}}\left(\gamma_{k+1}(\tilde{\tilde{\xi}}_k^i)\right)^2 \geq \frac{1}{N_{S_k}}\left(\sum_{i=1}^{N_{S_k}}\gamma_{k+1}(\tilde{\tilde{\xi}}_k^i)\right)^2 \quad (2.25)$$

The right hand term equals:

$$\left(\sum_{i=1}^{N_{S_k}}\gamma_{k+1}(\tilde{\tilde{\xi}}_k^i)\right)^2 = \sum_{i=1}^{N_{S_k}}\sum_{i'=1}^{N_{S_k}}\gamma_{k+1}(\tilde{\tilde{\xi}}_k^i)\gamma_{k+1}(\tilde{\tilde{\xi}}_k^{i'}) \quad (2.26)$$

Substituting this in (2.25) yields:

$$\left[\sum_{i=1}^{N_{S_k}}\left[\left(\gamma_{k+1}(\tilde{\tilde{\xi}}_k^i)\right)^2\right] - \frac{1}{N_{S_k}}\sum_{i=1}^{N_{S_k}}\sum_{i'=1}^{N_{S_k}}\left[\gamma_{k+1}(\tilde{\tilde{\xi}}_k^i)\gamma_{k+1}(\tilde{\tilde{\xi}}_k^{i'})\right]\right] \geq 0 \quad (2.27)$$

If $N_{S_k} = 1$, then all four splitting methods do the same. If $N_{S_k} \geq 2$, we have to compare variances in Propositions 2.2, 2.3, 2.4 and 2.5. Due to (2.27) for $\text{Var}_{MS} \leq \text{Var}_{MR}$, $\text{Var}_{RMS} \leq \text{Var}_{MS}$ and $\text{Var}_{FAS} \leq \text{Var}_{RMS}$ this means we have to verify:

$$\frac{\left(N_p - N_{S_k}\right)}{N_p N_{S_k}} \leq \frac{1}{N_{S_k}} \quad (2.28)$$

$$\left(N_p \bmod N_{S_k}\right) \leq \left(N_p - N_{S_k}\right) \quad (2.29)$$

$$\frac{\left(N_p \bmod N_{S_k}\right)\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right]}{\left(N_{S_k} - 1\right)} \leq \left(N_p \bmod N_{S_k}\right) \quad (2.30)$$

From $\frac{1}{N_{S_k}} - \frac{1}{N_p} \leq \frac{1}{N_{S_k}}$ follows that inequality (2.28) holds true.

Substituting $\alpha_k N_{S_k} + \left(N_p \bmod N_{S_k}\right) = N_p$ in $\left(N_p \bmod N_{S_k}\right) \leq \left(N_p \bmod N_{S_k}\right) + N_{S_k}(\alpha_k - 1)$ and subsequent rearrangement of terms proves (2.29).

Because $\left(N_p \bmod N_{S_k}\right)^2 \geq \left(N_p \bmod N_{S_k}\right)$ we get:

$$\left(N_p \bmod N_{S_k}\right)\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right] \leq \left(N_p \bmod N_{S_k}\right)\left(N_{S_k} - 1\right)$$

Dividing both sides by $\left(N_{S_k} - 1\right)$ proves (2.30). **Q.E.D.**

Proof of Theorem 2.3:

By defining the notation $\overline{\gamma}_k^\pi \triangleq \prod_{k'=1}^k \overline{\gamma}_{k'}$, we can write $\overline{\gamma}_k^\pi = \overline{\gamma}_k \overline{\gamma}_{k-1}^\pi$. By also defining the sigma algebra $C_{\tilde{\xi}}^k = \sigma\{\tilde{\xi}_{k'}^i; i = 1,...,N_{S_{k'}}, k' = 1,..,k\}$, we can subsequently derive:

$$\mathbb{E}\{\overline{\gamma}_{k-1}^\pi \overline{\gamma}_k \mid C_{\tilde{\xi}}^{k-1}\} \overset{a}{=} \overline{\gamma}_{k-1}^\pi \mathbb{E}\{\overline{\gamma}_k \mid C_{\tilde{\xi}}^{k-1}\} \quad (2.31)$$

where equality (a) holds because $C_{\tilde{\xi}}^{k-1} \subset \sigma\{\overline{\gamma}_{k-1}^\pi\}$.

If $N_{S_k} = 0$, then we have $\overline{\gamma}_k = \frac{N_{S_k}}{N_p} = 0$ and

$$\mathrm{Var}\left\{\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k\right\} = \mathrm{Var}\left\{\overline{\gamma}_{k-1}^{\pi}\cdot 0\right\} = 0 \tag{2.32}$$

Thus if $N_{S_k} = 0$, then $\mathrm{Var}\left\{\overline{\gamma}_k^{\pi}\right\}$ is the same under any of the four splitting methods.

If $N_{S_k} \geq 1$, we can derive as follows by using the law of total variance:

$$\begin{aligned}\mathrm{Var}\left\{\overline{\gamma}_k^{\pi}\right\} &= \mathrm{Var}\left\{\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k\right\} = \mathbb{E}\left\{\mathrm{Var}\left\{\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\} + \mathrm{Var}\left\{\mathbb{E}\left\{\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\} \\ &= \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2 \mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\} + \mathrm{Var}\left\{\mathbb{E}\left\{\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\}\end{aligned} \tag{2.33}$$

By using $\mathrm{Var}\{X\} = \mathbb{E}\{X^2\} + \{\mathbb{E}(X)\}^2$, (2.33) becomes:

$$\mathrm{Var}\left\{\overline{\gamma}_k^{\pi}\right\} = \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2 \mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\} + \mathbb{E}\left\{\left[\mathbb{E}\left\{\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right]^2\right\} - \left\{\mathbb{E}\left[\mathbb{E}\left(\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k \mid C_{\xi}^{k-1}\right)\right]\right\}^2 \tag{2.34}$$

Using the property of the conditional expectation, we can derive:

$$\mathrm{Var}\left\{\overline{\gamma}_k^{\pi}\right\} = \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2 \mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\} + \mathbb{E}\left\{\left[\overline{\gamma}_{k-1}^{\pi}\mathbb{E}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right]^2\right\} - \left\{\mathbb{E}\left(\overline{\gamma}_{k-1}^{\pi}\overline{\gamma}_k\right)\right\}^2 \tag{2.35}$$

Further evaluation of (2.35) yields:

$$\begin{aligned}&\mathrm{Var}\left\{\overline{\gamma}_k^{\pi}\right\} \\ &= \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2 \mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\} + \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2\left[\mathbb{E}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right]^2\right\} - \left\{\gamma_k^{\pi}\right\}^2 \\ &= \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2 \mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\} + \left(\overline{\gamma}_{k-1}^{\pi}\right)^2\left[\mathbb{E}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right]^2\right\} - \left\{\gamma_k^{\pi}\right\}^2 \\ &= \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2 \mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right\} + \mathbb{E}\left\{\left(\overline{\gamma}_{k-1}^{\pi}\right)^2\left[\mathbb{E}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right]^2\right\} - \left\{\gamma_k^{\pi}\right\}^2 \\ &= \int p_{\overline{\gamma}_{k-1}^{\pi}}(y)\, y^2 \mathbb{E}\left\{\mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\} \mid \overline{\gamma}^{k-1} = y\right\} dy \\ &\quad + \int p_{\overline{\gamma}_{k-1}^{\pi}}(y)\, y^2 \mathbb{E}\left\{\left[\mathbb{E}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}\right]^2 \mid \overline{\gamma}^{k-1} = y\right\} dy - \left\{\gamma_k^{\pi}\right\}^2\end{aligned} \tag{2.36}$$

To complete the proof we have to compare (2.36) under the four splitting methods. $\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_{\xi}^{k}\right\}$ equals under each of the four splitting methods. To compare the variance term we denote by V_{FAS}^k, V_{RMS}^k, V_{MS}^k and V_{MR}^k the $\mathrm{Var}\left\{\overline{\gamma}_k \mid C_{\xi}^{k-1}\right\}$ under FAS, RMS, MS and MR respectively. From Theorem 2.2, we know $\mathrm{V}_{FAS}^k \leq \mathrm{V}_{RMS}^k \leq \mathrm{V}_{MS}^k \leq \mathrm{V}_{MR}^k$ at level k. Due to the monotonicity of conditional expectation, this implies $\mathbb{E}\left\{\mathrm{V}_{FAS}^k \mid \overline{\gamma}_{k-1}^{\pi} = y\right\} \leq \mathbb{E}\left\{\mathrm{V}_{RMS}^k \mid \overline{\gamma}_{k-1}^{\pi} = y\right\} \leq \mathbb{E}\left\{\mathrm{V}_{MS}^k \mid \overline{\gamma}_{k-1}^{\pi} = y\right\} \leq \mathbb{E}\left\{\mathrm{V}_{RMS}^k \mid \overline{\gamma}_{k-1}^{\pi} = y\right\}$. For (2.36) this means that if $N_{S_k} \geq 1$, then $\mathrm{Var}\left\{\overline{\gamma}_k^{\pi}\right\}$ under FAS, RMS, MS and MR satisfy inequality (2.24). **Q.E.D.**

Proof of Theorem 2.4:

Theorem 2.3 shows that it is advantageous to use FAS at level *k,* whatever splitting types are used at level 1 to level *k*-1. For *k*=*m*, this implies an advantage to use FAS at level *m*. The same reasoning shows that it also is advantageous to use FAS at level *k*=*m*-1. This reasoning can be repeated for level m-2, m-3, …, k=2. At level k=1, there is no difference between the two splitting strategies. Therefore, we can conclude that if all levels make use of FAS, then $\mathrm{Var}\{\bar{\gamma}\}$ is less than or equal to that when all levels make use of RMS, i.e. $V_{FAS} \le V_{RMS}$. This reasoning is also be applied for RMS relative to MS and MR, which yields $V_{RMS} \le V_{MS}$. Finally this reasoning is applied to MS relative to MR, which yields $V_{MS} \le V_{MR}$. **Q.E.D.**

2.5. Simulation example

2.5.1 Geometric Brownian motion example

Following Krystul (2006, pp. 22-26) in this section we apply IPS for the estimation reach probability for a Geometric Brownian motion, and compare the results under Fixed Assignment splitting versus those under multinomial splitting versus those under multinomial resampling. The SDE of Geometric Brownian motion satisfies:

$$dX_t = (\mu + \frac{\sigma^2}{2})X_t dt + \sigma X_t dW_t \tag{2.37}$$

where $\mu > 0$, $\sigma > 0$ and $X_0 \ge 1$. We want to estimate the probability $\mathbb{P}\{\tau < T\}$ with $\tau \triangleq \inf\{t > 0 : X_t \ge L\}$.

2.5.2. Analytical and MC simulation results

Thanks to (Tuckwell and Wan, 1984; Karlin and Taylor, 1975, p363, Theorem 5.3), we can use the following equation to evaluate reach probabilities:

$$\gamma = \mathbb{P}(\tau < T) = \int_0^T \frac{\ln(L / X_0)}{\sqrt{2\pi\sigma^2 t^3}} \exp\left\{ \frac{-(\ln(L / X_0) - \mu t)^2}{2\sigma^2 t} \right\} dt \tag{2.38}$$

For this example, we use (2.38) to set the levels $\{L_k, k = 1,...,m\}$, such that the conditional probabilities between successive levels are equal to 1/10 for Table 2.1. Table 2.1 shows the resulting L_k level values for k=1,2,…, as well as the analytical γ_k and γ results for these levels. The right columns in Table 2.1 also show the $\bar{\gamma}_{MC}$ results obtained through straightforward Monte Carlo (MC) simulation using 10000 runs with numerical integration time step $\Delta = 2 \times 10^{-3} s$. The results in Table 2.1 show that straightforward MC simulation based estimation of γ fails to work beyond k=4. Instead of stopping the simulation of the i-th particle at each stopping times τ_k^i, we stop it at t_k^i, i.e. the end of the first integration time step that $x_{t_k^i}^i$ is at or has passed level k. The implication is that it remains to be verified if the numerical time step Δ of the IPS simulation is small enough.

Table 2.1. Analytical and MC estimated γ and γ_k, $k=1..10$, for geometric Brownian motion example, with $\mu=1$, $\sigma=1$, $X_0=1$, T=1s and $L=1717.25$. The MC estimated $\bar{\gamma}_{MC}$ used 10000 runs with $\Delta=2\times10^{-3}s$

k	L_k	γ_k	γ	$\bar{\gamma}_{MC}$
1	12.27	0.09998	0.09998	0.0957
2	33.038	1.000×10^{-1}	1.000×10^{-2}	0.0085
3	69.09	1.000×10^{-1}	1.000×10^{-3}	6.000×10^{-4}
4	127.45	1.001×10^{-1}	1.001×10^{-4}	1.000×10^{-4}
5	217.5	1.000×10^{-1}	1.000×10^{-5}	0
6	351.445	1.000×10^{-1}	1.000×10^{-6}	0
7	545.14	1.000×10^{-1}	1.000×10^{-7}	0
8	818.935	1.000×10^{-1}	1.000×10^{-8}	0
9	1198.75	1.000×10^{-1}	1.000×10^{-9}	0
10	1717.25	1.000×10^{-1}	1.000×10^{-10}	0

2.5.3. IPS simulation results

In this subsection we apply IPS under Multinomial resampling, under multinomial splitting and under fixed assignment splitting. By repeating IPS N_{IPS} times estimates of the rate of surviving IPS, ρ_S, and Normalized root-mean-square error, $\hat{c}_{\hat{\gamma},NRMSE}$. The results are shown in Table 2.2 with $\Delta=2\times10^{-3}s$ and Table 2.3 with $\Delta=4\times10^{-4}s$, for $N_p=1000$ and $N_{IPS}=1000$. The measures $\hat{\gamma}$, ρ_S and $\hat{c}_{\hat{\gamma},NRMSE}$ are defined as follows:

$$\hat{\gamma}=\frac{\sum_{i=1}^{N_{IPS}}\bar{\gamma}^i}{N_{IPS}} \tag{2.39}$$

$$\rho_S=\frac{\sum_{i=1}^{N_{IPS}}1_{\bar{\gamma}^i>0}}{N_{IPS}} \tag{2.40}$$

$$\hat{c}_{\hat{\gamma},NRMSE}=\frac{RMSE}{\gamma}\times100\% \tag{2.41}$$

with $1_{\bar{\gamma}^i>0}=\begin{cases}1, \text{ if } & \bar{\gamma}^i>0\\ 0, \text{if } & \bar{\gamma}^i=0\end{cases}$ and

$$RMSE=\sqrt{\frac{\sum_{i=1}^{N_{IPS}}(\bar{\gamma}^i-\gamma)^2}{N_{IPS}}} \tag{2.42}$$

where $\bar{\gamma}^i$ denotes the estimated reach probability for the i-th IPS simulation.

Table 2.2 Multiple times IPS simulation results under Multinomial Resampling vs. Multinomial splitting vs. Fixed Assignment splitting for the setting of Table 2.1, $\Delta = 2\times10^{-3} s$, $N_p = 1000$ and $N_{IPS} = 1000$

	Multinomial Resampling			Multinomial splitting			Fixed Assignment splitting		
k	$\hat{\gamma}$	ρ_S	$\hat{c}_{\hat{\gamma},NRMSE}$	$\hat{\gamma}$	ρ_S	$\hat{c}_{\hat{\gamma},NRMSE}$	$\hat{\gamma}$	ρ_S	$\hat{c}_{\hat{\gamma},NRMSE}$
1	9.51×10^{-2}	100%	11%	9.60×10^{-2}	100%	10%	9.51×10^{-2}	100%	11%
2	9.29×10^{-3}	100%	19%	9.33×10^{-3}	100%	19%	9.31×10^{-3}	100%	18%
3	9.17×10^{-4}	100%	30%	9.21×10^{-4}	100%	29%	9.11×10^{-4}	100%	28%
4	9.10×10^{-5}	100%	46%	9.11×10^{-5}	100%	46%	9.04×10^{-5}	100%	43%
5	8.96×10^{-6}	100%	71%	8.95×10^{-6}	100%	69%	8.89×10^{-6}	100%	65%
6	8.90×10^{-7}	100%	110%	8.85×10^{-7}	100%	102%	8.67×10^{-7}	100%	95%
7	9.07×10^{-8}	99%	176%	8.67×10^{-8}	99%	158%	8.42×10^{-8}	100%	134%
8	9.31×10^{-9}	96%	292%	8.60×10^{-9}	95%	253%	8.06×10^{-9}	97%	182%
9	9.72×10^{-10}	86%	488%	8.59×10^{-10}	88%	398%	7.81×10^{-10}	88%	243%
10	1.10×10^{-10}	69%	912%	8.36×10^{-11}	73%	559%	7.29×10^{-11}	73%	292%

Table 2.3 Multiple times IPS simulation results under Multinomial Resampling vs. Multinomial splitting vs. Fixed Assignment splitting for the setting of Table 2.1, $\Delta = 4\times10^{-4} s$, $N_p = 1000$ and $N_{IPS} = 1000$

	Multinomial Resampling			Multinomial splitting			Fixed Assignment splitting		
k	$\hat{\gamma}$	ρ_S	$\hat{c}_{\hat{\gamma},NRMSE}$	$\hat{\gamma}$	ρ_S	$\hat{c}_{\hat{\gamma},NRMSE}$	$\hat{\gamma}$	ρ_S	$\hat{c}_{\hat{\gamma},NRMSE}$
1	9.81×10^{-2}	100%	10%	9.77×10^{-2}	100%	10%	9.78×10^{-2}	100%	10%
2	9.75×10^{-3}	100%	18%	9.73×10^{-3}	100%	17%	9.70×10^{-3}	100%	17%
3	9.63×10^{-4}	100%	29%	9.72×10^{-4}	100%	29%	9.58×10^{-4}	100%	28%
4	9.49×10^{-5}	100%	46%	9.76×10^{-5}	100%	46%	9.43×10^{-5}	100%	44%
5	9.47×10^{-6}	100%	74%	9.81×10^{-6}	100%	73%	9.18×10^{-6}	100%	66%
6	9.56×10^{-7}	100%	114%	9.95×10^{-7}	100%	112%	8.98×10^{-7}	100%	94%
7	9.86×10^{-8}	99%	173%	1.00×10^{-7}	99%	169%	8.65×10^{-8}	100%	131%
8	1.03×10^{-8}	95%	261%	1.02×10^{-8}	96%	241%	8.23×10^{-9}	96%	179%
9	1.09×10^{-9}	86%	390%	1.04×10^{-9}	89%	376%	7.68×10^{-10}	89%	235%
10	1.14×10^{-10}	72%	554%	1.02×10^{-10}	74%	498%	7.17×10^{-11}	75%	301%

The results in Table 2.2 and Table 2.3 show that the Normalized Root-Mean-Square Error, $\hat{c}_{\hat{\gamma},NRMSE}$, is under Fixed Assignment splitting better than under Multinomial splitting which is better than under Multinomial Resampling. This difference in $\hat{c}_{\hat{\gamma},NRMSE}$ increases with the *k*.

2.6. Conclusion

This chapter has studied the estimation of the reach probability of an unsafe set by a multi-dimensional diffusion process using the Interacting Particle System (IPS) framework of Cérou et al. (2006). More specifically it has been proven that IPS using fixed assignment splitting dominates in variance reduction over IPS using multinomial resampling (MR), multinomial splitting (MS), residual multinomial splitting RMS), or fixed assignment splitting (FAS).

First, in section 2.3, a novel characterization has been derived for the conditional variance at level k in Theorem 2.1. This has been elaborated in Propositions 2.2, 2.3, 2.4 and 2.5 for MR, MS, RMS and FAS respectively. Subsequently, the conditional variances are compared in section 2.4 through Theorems 2.2, 2.3 and 2.4. Theorem 2.2 proves the aimed results for an arbitrary single level k, given the same set of survived particles at the beginning of this level. Subsequently Theorem 2.3 proves the aimed results for an arbitrary single level k, under the condition that there are no differences in splitting strategy used at all earlier levels. Finally Theorem 2.4 completes the proof by induction using Theorem 2.3.

The difference in IPS performances under different splitting methods has been illustrated for a one-dimensional geometric Brownian motion example for which the reach probabilities are analytically known.

2.7 References

Abate, A., D'Innocenzo, A. and Di Benedetto, M.D. (2011). Approximate abstractions of stochastic hybrid systems. IEEE Transactions on Automatic Control, 56(11), pp.2688-2694.

Alur, R., Henzinger, T., Laffierriere and G., Pappas, G. (2000). Discrete abstractions of hybrid systems, Proc. of the IEEE, Vol. 88, pp. 971-984.

Asmussen, S. and H. Albrecher, H. (2010). Ruin probabilities. Advanced Series on Statistical Science & Applied Probability, 14. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second edition.

Bowsher, C.G. and Swain, P.S. (2012). Identifying sources of variation and the flow of information in biochemical networks. Proceedings of the National Academy of Sciences, 109(20), pp.E1320-E1328.

Blom, H.A.P., Krystul, J. and Bakker, G.J. (2006). A particle system for safety verification of free flight in air traffic, Proc. IEEE Conf. Decision and Control, San Diego, CA, pp. 1574-1579.

Blom, H.A.P., Krystul, J., and Bakker, G.J., Klompstra, M.B., Klein Obbink, B. (2007a). Free flight collision risk estimation by sequential MC simulation. Eds: C.G, Cassandras and J. Lygeros, Stochastic hybrid systems, Taylor & Francis, pp.249-281.

Bollobás, B. (2001), Random graphs, volume 73 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, second edition.

Botev, Z.I. and Kroese, D.P. (2008). An efficient algorithm for rare-event probability estimation, combinatorial optimization, and counting. Methodology and Computing in Applied Probability 10(4), pp. 471-505.

Bucklew, J.A. (2004). Introduction to rare event simulation, Springer, New York.

Cérou, F., Legland, F., Del Moral, P., and Lezaud, P. (2005). Limit theorems for the multilevel splitting algorithm in the simulation of rare events. In Proceedings of Winter Simulation Conference. Orlando, FL, pp. 682-691.

Cérou, F., Del Moral, P., Legland, F., and Lezaud, P. (2006). Genetic genealogical models in rare event analysis. Latin American J. of Probability & Mathematical Statistics, 1, pp. 181-203.

Del Moral, P. (2004). Feynman-Kac formulae: genealogical and interacting particle systems with applications, Springer.

Del Moral, P., Kouritzin, M. and Miclo, L. (2001). On a class of discrete generation interacting particle systems. Electronic Journal of Probability (6), pp1-26

Di Benedetto, M.D. Di Gennaro, S., D'Innocenzo, A. (2015). Hybrid Systems and Verification by Abstraction, in Hybrid Dynamical Systems, Springer, Switzerland, M. Djemai and M. Defoort (eds.), Lecture Notes in Control and Information Sciences 457, pp. 1-25.

Garvels, M.J. and Kroese, D.P. (1998). A comparison of RESTART implementations. In Proceedings of the 30th conference on Winter simulation, pp. 601-608. IEEE Computer Society Press.

Garvels, M.J.J. (2000). The splitting method in rare event simulation. PhD Thesis, University of Twente, The Netherlands.

Gerber, M., Chopin, N. and Whiteley, N., 2019. Negative association, ordering and convergence of resampling methods. Annals of Statistics, 47(4), pp.2236-2260.

Glasserman, P., Heidelberger, P., Shahabuddin, P. and Zajic, T. (1998). A large deviations perspective on the efficiency of multilevel splitting. IEEE Transactions on Automatic Control, 43(12), pp.1666-1679.

Glasserman, P., Heidelberger, P., Shahabuddin, P. and Zajic, T. (1999). Multilevel splitting for estimating rare event probabilities, Operations Research, Vol. 47, pp. 585-600.

Julius, A.A. and Pappas, G.J. (2009). Approximations of stochastic hybrid systems. IEEE Transactions on Automatic Control, 54(6), pp.1193-1203.

Karlin, S. and Taylor, H. M. (1975). Chapter 7 - Brownian Motion. in Karlin, S. and Taylor, H. M. (eds), A First Course in Stochastic Processes, Academic Press, pp. 340-391

Krystul, J. (2006). Modelling of stochastic hybrid systems with applications to accident risk assessment, PhD Thesis, University Twente, Enschede.

L'Ecuyer, P., Demers, V., and Tuffin, B. (2006), Splitting for rare-event simulation. Proceedings of the 38th conference on Winter simulation. Winter Simulation Conference. Monterey, California, pp. 137-148.

L'Ecuyer, P., Demers, V. and Tuffin, B. (2007), Rare events, splitting, and quasi-Monte Carlo. ACM Transactions on Modeling and Computer Simulation (TOMACS), 17(2), p.9.

L'Ecuyer, P., LeGland, F., Lezaud, P. and Tuffin, B. (2009), Splitting techniques, Eds: G. Rubino, B. Tuffin, Rare event simulation using Monte Carlo methods, Wiley, pp. 39-61.

Morio, J. and Balesdent, M. (2016). Estimation of Rare Event Probabilities in Complex Aerospace and Other Systems, Woodhead Publishing.

Prandini, M. and Hu, J. (2007). Stochastic Reachability: Theory and numerical approximation. In Cassandras, C.G. and Lygeros, J. (eds.), Stochastic hybrid systems, Taylor & Francis, pp. 107-137.

Robert, P. (2003). Stochastic networks and queues, volume 52 of Applications of Mathematics (New York). Springer-Verlag, Berlin, french edition. Stochastic Modelling and Applied Probability.

Rubinstein, R. (2010), Randomized algorithms with splitting: Why the classic randomization algorithms do not work and how to make them work, Methodology and Computing in Applied Probability, Vol. 12, pp. 1-50.

Tuckwell, H.C. and Wan, F.Y. (1984). First-passage time of Markov processes to moving barriers. Journal of applied probability, 21(4), pp.695-709.

Appendix 2.A. Conditional variance derivations

Proof of Proposition 2.1

If we consider the particles ξ_k^j copies from $\tilde{\xi}_k^i$ as a group, then for IPS step II in Algorithm 2.1 at level $k+1$, $\overline{\gamma}_{k+1}$ can be written as follows:

$$\overline{\gamma}_{k+1} = \frac{1}{N_p}\sum_{i=1}^{N_p} 1(\overline{\xi}_{k+1}^i \in Q_{k+1}) = \frac{1}{N_p}\sum_{i=1}^{N_{S_k}} Y_{k+1}^{k,i} \tag{2.43}$$

with $Y_{k+1}^{k,i}$ the number of particles that have reached level k+1 after mutation of the K_k^i copies from $\tilde{\xi}_k^i$. Hence, $Y_{k+1}^{k,i}$ has a conditional Binomial distribution with size K_k^i and success probability $\gamma_{k+1}(\tilde{\xi}_k^i)$ given K_k^i and $\tilde{\xi}_k^i$. Therefore, the pdf of $Y_{k+1}^{k,i}$ can be expressed as (2.10) and (2.11). **Q.E.D.**

Proof of Theorem 2.1

Let us define $C_{\tilde{\xi}}^k$ and $C_{\tilde{\xi},K}^k$ as follows:

$$C_{\tilde{\xi}}^k \triangleq \sigma\{\tilde{\xi}_{k'}^i; i = 1,\ldots,N_{S_{k'}}, k' = 1,..,k\} \tag{2.44}$$

$$C_{\tilde{\xi},K}^k \triangleq \sigma\left\{\tilde{\xi}_{k'}^i, K_{k'}^i; i = 1,\ldots,N_{S_{k'}}, k' = 1,..,k;\right\} \tag{2.45}$$

Substitution of (2.43) in $\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^k\right\}$ and subsequent evaluation yields:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^k\right\} = \mathbb{E}\left\{\frac{1}{N_p}\sum_{i=1}^{N_{S_k}} Y_{k+1}^{k,i} \mid C_{\tilde{\xi}}^k\right\} = \frac{1}{N_p}\sum_{i=1}^{N_{S_k}} \mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi}}^k\right\} \overset{a}{=} \frac{1}{N_p}\sum_{i=1}^{N_{S_k}} \mathbb{E}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\} \tag{2.46}$$

where equality (a) holds because $C_{\tilde{\xi},K}^k \supset C_{\tilde{\xi}}^k$ and $\mathbb{E}\{X \mid Y\} = \mathbb{E}\{\mathbb{E}\{X \mid Y,Z\} \mid Y\}$.

In a similar way, we can derive:

$$\begin{aligned}
&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^k\right\} = \mathrm{Var}\left\{\frac{1}{N_p}\sum_{i=1}^{N_{S_k}} Y_{k+1}^{k,i} \mid C_{\tilde{\xi}}^k\right\} = \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}} \mathrm{Var}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi}}^k\right\} + \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1\\ i'\neq i}}^{N_{S_k}} \mathrm{Cov}\left\{Y_{k+1}^{k,i}, Y_{k+1}^{k,i'} \mid C_{\tilde{\xi}}^k\right\}\\
&\overset{a}{=} \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{\mathrm{Var}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\} + \mathrm{Var}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\}\right] + \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1\\ i'\neq i}}^{N_{S_k}}\left[\mathbb{E}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i} Y_{k+1}^{k,i'} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\}\right.\\
&\left.-\mathbb{E}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\}\mathbb{E}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i'} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\}\right]
\end{aligned} \tag{2.47}$$

where equality (a) holds because of the law of total conditional variance (Bowsher and Swain, 2012).

Further evaluation of (2.47) yields:

$$\begin{aligned}
\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^k\right\} &\overset{a}{=} \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{\mathrm{Var}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\} + \mathrm{Var}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\}\right]\\
&+\frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1\\ i'\neq i}}^{N_{S_k}}\left[\mathbb{E}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\}\mathbb{E}\left\{Y_{k+1}^{k,i'} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\} - \mathbb{E}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\}\mathbb{E}\left\{\mathbb{E}\left\{Y_{k+1}^{k,i'} \mid C_{\tilde{\xi},K}^k\right\} \mid C_{\tilde{\xi}}^k\right\}\right]
\end{aligned} \tag{2.48}$$

where equality (a) holds because $Y_{k+1}^{k,i}$ and $Y_{k+1}^{k,i'}$ are conditionally independent given $C_{\tilde{\xi},K}^{k}$.

Since each $Y_{k+1}^{k,i}$ has a conditional Binomial distribution with size K_k^i and success probability $\gamma_{k+1}(\tilde{\xi}_k^i)$. Then, by using Binomial distribution properties, we get:

$$\mathbb{E}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^{k}\right\} = \mathbb{E}\left\{Y_{k+1}^{k,i} \mid K_k^j, \tilde{\xi}_k^j, j=1,\dots,N_{S_k}\right\} \overset{a}{=} \mathbb{E}\left\{Y_{k+1}^{k,i} \mid K_k^i, \tilde{\xi}_k^i\right\} = K_k^i \cdot \gamma_{k+1}(\tilde{\xi}_k^i) \tag{2.49}$$

and

$$\operatorname{Var}\left\{Y_{k+1}^{k,i} \mid C_{\tilde{\xi},K}^{k}\right\} = \operatorname{Var}\left\{Y_{k+1}^{k,i} \mid K_k^j, \tilde{\xi}_k^j, j=1,\dots,N_{S_k}\right\} \overset{a}{=} \operatorname{Var}\left\{Y_{k+1}^{k,i} \mid K_k^i, \tilde{\xi}_k^i\right\} = K_k^i \gamma_{k+1}(\tilde{\xi}_k^i)(1-\gamma_{k+1}(\tilde{\xi}_k^i)) \tag{2.50}$$

where equality (a) holds because $Y_{k+1}^{k,i}$ is conditionally dependent of K_k^i and $\tilde{\xi}_k^i$, but conditionally independent of $K_k^{i'}$ and $\tilde{\xi}_k^{i'}$ for $i' \neq i$.

Substituting (2.49) into (2.46) and subsequent evaluation yields:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^{k}\right\} = \frac{1}{N_p}\sum_{i=1}^{N_{S_k}} \mathbb{E}\left\{K_k^i \cdot \gamma_{k+1}(\tilde{\xi}_k^i) \mid C_{\tilde{\xi}}^{k}\right\} \overset{a}{=} \frac{1}{N_p}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}\gamma_{k+1}(\tilde{\xi}_k^i)\right] \overset{b}{=} \frac{1}{N_p}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i \mid \tilde{\xi}_k^j, \text{ all } j\right\}\gamma_{k+1}(\tilde{\xi}_k^i)\right] \tag{2.51}$$

where equality (a) holds because $\mathbb{E}\{f(Z)Y|Z\} = f(Z)\mathbb{E}\{Y|Z\}$; equality (b) holds because of Markov property of $\{\tilde{\xi}_k^i\}$.

Similarly, substituting (2.49) and (2.50) into (2.48) and subsequent evaluation yields:

$$\begin{aligned}
\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^{k}\right\} &= \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i.\gamma_{k+1}(\tilde{\xi}_k^i)(1-\gamma_{k+1}(\tilde{\xi}_k^i)) \mid C_{\tilde{\xi}}^{k}\right\} + \operatorname{Var}\left\{K_k^i.\gamma_{k+1}(\tilde{\xi}_k^i) \mid C_{\tilde{\xi}}^{k}\right\}\right] \\
&+ \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1 \\ i'\neq i}}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i \cdot \gamma_{k+1}(\tilde{\xi}_k^i) K_k^{i'} \cdot \gamma_{k+1}(\tilde{\xi}_k^{i'}) \mid C_{\tilde{\xi}}^{k}\right\} - \mathbb{E}\left\{K_k^i \cdot \gamma_{k+1}(\tilde{\xi}_k^i) \mid C_{\tilde{\xi}}^{k}\right\}\mathbb{E}\left\{K_k^{i'} \cdot \gamma_{k+1}(\tilde{\xi}_k^{i'}) \mid C_{\tilde{\xi}}^{k}\right\}\right] \\
&\overset{a}{=} \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}\mathbb{E}\left\{\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right) \mid C_{\tilde{\xi}}^{k}\right\}\right] \\
&+ \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{\left(K_k^i\right)^2 \mid C_{\tilde{\xi}}^{k}\right\}\mathbb{E}\left\{\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2 \mid C_{\tilde{\xi}}^{k}\right\} - \mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}^2 \mathbb{E}\left\{\gamma_{k+1}(\tilde{\xi}_k^i) \mid C_{\tilde{\xi}}^{k}\right\}^2\right] \\
&+ \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1 \\ i'\neq i}}^{N_{S_k}}\left[\left(\mathbb{E}\left\{K_k^i \cdot K_k^{i'} \mid C_{\tilde{\xi}}^{k}\right\} - \mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}\mathbb{E}\left\{K_k^{i'} \mid C_{\tilde{\xi}}^{k}\right\}\right).\mathbb{E}\left\{\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'}) \mid C_{\tilde{\xi}}^{k}\right\}\right]
\end{aligned} \tag{2.52}$$

where equality (a) is thanks to $C_{\tilde{\xi}}^{k}$-conditional independence of $\tilde{\xi}_k^i$ and K_k^i.

Further evaluation of (2.52) yields:

$$\begin{aligned}
&\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^{k}\right\} \\
&= \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] + \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\left(\mathbb{E}\left\{\left(K_k^i\right)^2 \mid C_{\tilde{\xi}}^{k}\right\} - \mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}^2\right)\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] \\
&+ \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{i'=1}^{N_{S_k}}\left[\left(\mathbb{E}\left\{K_k^i \cdot K_k^{i'} \mid C_{\tilde{\xi}}^{k}\right\} - \mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}\mathbb{E}\left\{K_k^{i'} \mid C_{\tilde{\xi}}^{k}\right\}\right).\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right] \\
&= \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] + \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\operatorname{Var}\left\{K_k^i \mid C_{\tilde{\xi}}^{k}\right\}\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] \\
&+ \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1 \\ i'\neq i}}^{N_{S_k}}\left[\operatorname{Cov}\left\{K_k^i, K_k^{i'} \mid C_{\tilde{\xi}}^{k}\right\}.\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right]
\end{aligned} \tag{2.53}$$

Due to the strong Markov property of $\{\tilde{\xi}_k^i\}$, the $C_{\tilde{\xi}}^k$ conditioning in (2.53) can be replaced by the condition $\{\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\}$. **Q.E.D.**

Proof of Proposition 2.2

For multinomial resampling, the vector $(K_k^1, K_k^2,...,K_k^{N_{S_k}})$ follows a multinomial distribution with the number of trials equal to N_p, which means $K_k^1 + K_k^2+...+K_k^{N_{S_k}} = N_p$, and with equal success probabilities $\frac{1}{N_{S_k}}$. Using multinomial distribution properties, we know for $i=1,2,...,N_{S_k}$:

$$\mathbb{E}\left\{K_k^i \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = N_p \cdot \frac{1}{N_{S_k}} = \frac{N_p}{N_{S_k}} \tag{2.54}$$

$$\mathrm{Var}\left\{K_k^i \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = N_p \frac{1}{N_{S_k}}(1-\frac{1}{N_{S_k}}) = \frac{N_p}{N_{S_k}}(1-\frac{1}{N_{S_k}}) \tag{2.55}$$

For K_k^i and $K_k^{i'}$ ($i \neq i'$), we can derive:

$$\mathrm{Cov}\left\{K_k^i K_k^{i'} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = -N_p \frac{1}{N_{S_k}} \frac{1}{N_{S_k}} \tag{2.56}$$

Substituting (2.54) into (2.51) and substituting (2.54), (2.55) and (2.56) into (2.53) yield:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\right] \tag{2.57}$$

$$\begin{aligned}\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^k\right\} &= \frac{1}{N_p^2}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] \\ &+\frac{1}{N_p^2}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}(1-\frac{1}{N_{S_k}})\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] + \frac{1}{N_p^2}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1\\ i'\neq i}}^{N_{S_k}}\left[\frac{-N_p}{N_{S_k}^2}\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right]\end{aligned} \tag{2.58}$$

Elaboration of (2.57) and (2.58) yields the equations of Proposition 2.2. **Q.E.D.**

Proof of Proposition 2.3

For Multinomial Splitting, the vector $(K_k^1-1, K_k^2-1,...,K_k^{N_{S_k}}-1)$ follows a multinomial distribution with the number of trials equal to $N_p - N_{S_k}$, which means $\left(K_k^1-1\right)+\left(K_k^2-1\right)+...+\left(K_k^{N_{S_k}}-1\right) = N_p - N_{S_k}$, and with equal success probabilities $\frac{1}{N_{S_k}}$. Using multinomial distribution properties, we know for $i=1,2,...,N_{S_k}$:

$$\mathbb{E}\left\{K_k^i - 1 \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{N_p - N_{S_k}}{N_{S_k}} \tag{2.59}$$

$$\mathrm{Var}\left\{K_k^i - 1 \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \left(N_p - N_{S_k}\right)\frac{1}{N_{S_k}}(1 - \frac{1}{N_{S_k}})$$
$$= \frac{N_p - N_{S_k}}{N_{S_k}}(1 - \frac{1}{N_{S_k}}) \tag{2.60}$$

From (2.59) and (2.60), we can derive:

$$\mathbb{E}\left\{K_k^i \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \frac{N_p}{N_{S_k}} \tag{2.61}$$

and

$$\mathrm{Var}\left\{K_k^i \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \frac{N_p - N_{S_k}}{N_{S_k}}(1 - \frac{1}{N_{S_k}}) \tag{2.62}$$

For K_k^i and $K_k^{i'}$ ($i \neq i'$), we can derive:

$$\mathrm{Cov}\left\{K_k^i K_k^{i'} \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \mathrm{Cov}\left\{\left(K_k^i - 1\right)\left(K_k^{i'} - 1\right) \mid \tilde{\xi}_k^j,\ \mathrm{all}\, j\right\} \overset{a}{=} \frac{-\left(N_p - N_{S_k}\right)}{{N_{S_k}}^2} \tag{2.63}$$

where equality (a) holds because of the multinomial distribution property on the covariance. Substituting (2.61) into (2.51) and substituting (2.61), (2.62) and (2.63) into (2.53) yield:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\right] \tag{2.64}$$

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{{N_p}^2}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\left(1 - \gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] + \frac{1}{{N_p}^2}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p - N_{S_k}}{N_{S_k}}(1 - \frac{1}{N_{S_k}})\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right]$$
$$+ \frac{1}{{N_p}^2}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1 \\ i' \neq i}}^{N_{S_k}}\left[\frac{-\left(N_p - N_{S_k}\right)}{{N_{S_k}}^2}\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right] \tag{2.65}$$

Elaboration of (2.64) and (2.65) yields the equations of Proposition 2.3. **Q.E.D.**

Proof of Proposition 2.4

For residual multinomial splitting, the vector $(K_k^1 - \alpha_k, K_k^2 - \alpha_k, ..., K_k^{N_{S_k}} - \alpha_k)$ follows a multinomial distribution with the number of trials equal to $\left(N_p \bmod N_{S_k}\right)$, and with equal success probabilities $\frac{1}{N_{S_k}}$. Using multinomial distribution properties, we know:

$$\mathbb{E}\left\{K_k^i - \alpha_k \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \frac{N_p \bmod N_{S_k}}{N_{S_k}} \tag{2.66}$$

$$\mathrm{Var}\left\{K_k^i - \alpha_k \mid \tilde{\xi}_k^1, ..., \tilde{\xi}_k^{N_{S_k}}\right\} = \frac{\left(N_p \bmod N_{S_k}\right)}{N_{S_k}}(1 - \frac{1}{N_{S_k}}) \tag{2.67}$$

From (2.66) and (2.67), we can derive:

$$\mathbb{E}\left\{K_k^i \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{\left(N_p \bmod N_{S_k}\right)}{N_{S_k}} + \alpha_k = \frac{N_p}{N_{S_k}} \tag{2.68}$$

and

$$\mathrm{Var}\left\{K_k^i \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{\left(N_p \bmod N_{S_k}\right)}{N_{S_k}}\left(1-\frac{1}{N_{S_k}}\right) \tag{2.69}$$

For K_k^i and $K_k^{i'}$ ($i \neq i'$), we can derive:

$$\mathrm{Cov}\left\{K_k^i K_k^{i'} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = -\left(N_p \bmod N_{S_k}\right)\frac{1}{N_{S_k}}\frac{1}{N_{S_k}} \tag{2.70}$$

Substituting (2.68) into (2.51) and substituting (2.68), (2.69) and (2.70) into (2.53) yield:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\right] \tag{2.71}$$

$$\begin{aligned}\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^k\right\} = {} & \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] + \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\frac{\left(N_p \bmod N_{S_k}\right)}{N_{S_k}}\left(1-\frac{1}{N_{S_k}}\right)\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] \\ & + \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1 \\ i'\neq i}}^{N_{S_k}}\left[\frac{-\left(N_p \bmod N_{S_k}\right)}{N_{S_k}^{\,2}}\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right]\end{aligned} \tag{2.72}$$

Elaboration of (2.71) and (2.72) yields the equations of Proposition 2.4. **Q.E.D.**

Proof of Proposition 2.5

To evaluate K_k^i, $i=1,2,...,N_{S_k}$, we define scalar parameter α_k as follows:

$$\alpha_k \triangleq \left\lfloor \frac{N_p}{N_{S_k}} \right\rfloor \tag{2.73}$$

with floor function $\lfloor x \rfloor \triangleq \max\{i \in \mathbb{Z} \mid i \leq x\}$.

For Fixed Assignment Splitting, the vector $(K_k^1-\alpha_k, K_k^2-\alpha_k,...,K_k^{N_{S_k}}-\alpha_k)$ follows a multivariate hypergeometric distribution with the number of trials equal to $\left(N_p \bmod N_{S_k}\right)$, and with equal success probabilities $\frac{1}{N_{S_k}}$. Using multivariate hypergeometric distribution properties, we know for $i=1,2,...,N_{S_k}$:

$$\mathbb{E}\left\{K_k^i-\alpha_k \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{N_p \bmod N_{S_k}}{N_{S_k}} \tag{2.74}$$

$$\begin{aligned}\mathrm{Var}\left\{K_k^i-\alpha_k \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} &= \left(N_p \bmod N_{S_k}\right)\frac{\left[N_{S_k}-\left(N_p \bmod N_{S_k}\right)\right]}{N_{S_k}-1}\frac{1}{N_{S_k}}\left(1-\frac{1}{N_{S_k}}\right) \\ &= \frac{\left(N_p \bmod N_{S_k}\right)}{N_{S_k}}\frac{\left[N_{S_k}-\left(N_p \bmod N_{S_k}\right)\right]}{N_{S_k}}\end{aligned} \tag{2.75}$$

From (2.74) and (2.75), we can derive:

$$\mathbb{E}\left\{K_k^i \mid \tilde{\xi}_k^1,\ldots,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{N_p \bmod N_{S_k}}{N_{S_k}} + \alpha_k = \frac{N_p}{N_{S_k}} \tag{2.76}$$

and

$$\operatorname{Var}\left\{K_k^i \mid \tilde{\xi}_k^1,\ldots,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{\left(N_p \bmod N_{S_k}\right)}{N_{S_k}}\frac{\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right]}{N_{S_k}} \tag{2.77}$$

For K_k^i and $K_k^{i'}$ ($i \neq i'$), we derive:

$$\begin{aligned}\operatorname{Cov}\left\{\left(K_k^i\right)\left(K_k^{i'}\right) \mid \tilde{\xi}_k^1,\ldots,\tilde{\xi}_k^{N_{S_k}}\right\} &= \operatorname{Cov}\left\{\left(K_k^i - \alpha_k\right)\left(K_k^{i'} - \alpha_k\right) \mid \tilde{\xi}_k^1,\ldots,\tilde{\xi}_k^{N_{S_k}}\right\} \\ &\overset{a}{=} -\left(N_p \bmod N_{S_k}\right)\frac{\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right]}{N_{S_k} - 1}\frac{1}{N_{S_k}}\frac{1}{N_{S_k}}\end{aligned} \tag{2.78}$$

where equality (a) holds because of the multivariate hypergeometric distribution property on the covariance.

Substituting (2.76) into (2.51) and substituting (2.76), (2.77) and (2.78) into (2.53) yield:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,\ldots,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\right] \tag{2.79}$$

$$\begin{aligned}\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi}}^k\right\} &= \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\frac{N_p}{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] \\ &+ \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\frac{\left(N_p \bmod N_{S_k}\right)}{N_{S_k}}\frac{\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right]}{N_{S_k}}\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] \\ &+ \frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1 \\ i'\neq i}}^{N_{S_k}}\left[\frac{-\left(N_p \bmod N_{S_k}\right)}{N_{S_k}^{\,2}}\frac{\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right]}{N_{S_k} - 1}\cdot\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right]\end{aligned} \tag{2.80}$$

Elaboration of (2.79) and (2.80) yields the equations of Proposition 2.5. **Q.E.D.**

3

Interacting Particle System based Estimation of Reach Probability of General Stochastic Hybrid Systems

For diffusions, a well-developed approach in rare event estimation is to introduce a suitable factorization of the reach probability and then to estimate these factors through simulation of an Interacting Particle System (IPS). This chapter studies IPS based reach probability estimation for General Stochastic Hybrid Systems (GSHS). The continuous-time executions of a GSHS evolve in a hybrid state space under influence of combinations of diffusions, spontaneous jumps and forced jumps. In applying IPS to a GSHS, simulation of the GSHS execution plays a central role. From literature, two basic approaches in simulating GSHS execution are known. One approach is direct simulation of a GSHS execution. An alternative is to first transform the spontaneous jumps of a GSHS to forced transitions, and then to simulate executions of this transformed version. This chapter will show that the latter transformation yields an extra Markov state component that should be treated as being unobservable for the IPS process. To formally make this state component unobservable for IPS, this chapter also develops an enriched GSHS transformation prior to transforming spontaneous jumps to forced jumps. The expected improvements in IPS reach probability estimation are also illustrated through simulation results for a simple GSHS example.

This chapter has been published as "Ma, H. and Blom, H.A.P., Interacting particle system based estimation of reach probability of general stochastic hybrid systems. Nonlinear Analysis: Hybrid Systems, 47, 2023, p.101303. https://doi.org/10.1016/j.nahs.2022.101303"

3.1. Introduction

A Stochastic Hybrid System (SHS) as defined by (Hu et al., 2000) involves two dynamically interacting state components, i.e. a discrete-valued θ_t and a continuous-valued x_t . The θ_t component may switch when x_t hits a θ_t -dependent boundary. The x_t component evolves under influence of θ_t -dependent Brownian motion and forced jumps at moments of hitting a θ_t -dependent boundary. Bujorianu and Lygeros (2006) define a General SHS (GSHS) by extending an SHS with spontaneous jumps, the rate λ of which depends on the joint state (x_t, θ_t). Well-known sub-classes of GSHS executions are solutions of SDE's driven by Brownian motion and spontaneous jumps generated by Poisson random measure. Specific subclasses are Markov switching diffusions (Mao and Yuan, 2006), hybrid switching diffusions (Yin and Zhu, 2010) and hybrid switching jump-diffusions (Kunwai and Zhu, 2020). These developments include methods for the numerical integration of both spontaneous jumps and Brownian motion. Teel et al. (2014) provide an in-depth survey regarding stability analysis of GSHS and various sub-classes.

A GSHS can be transformed to an SHS of (Hu et al., 2000) by capturing each spontaneous jump as a forced jump at an exit time condition (Lygeros and Prandini, 2010). More specifically, an auxiliary state component q_t, representing "remaining local time", starts at each exit time as an exponentially distributed random variable, subsequently evolves as $dq_t = -\lambda(\theta_t, x_t)dt$, and defines a new exit time upon reaching value zero. As shown in the stochastic hybrid systems survey by Lygeros and Prandini (2010), the mainstream of stochastic hybrid control developments address diffusion and forced jumps only; e.g. Bensoussan and Menaldi (2000), Koutsoukos (2004). A key exception is optimal control of a Markov switching diffusion via its SDE coefficients and spontaneous jump rate (Ghosh et al., 1993).

As will be shown in this chapter, there may be unexpected effects when transforming spontaneous jumps in a GSHS to forced jumps in an SHS. This chapter studies the role played by these unexpected effects in estimating stochastic reach probability for a GSHS using the Interacting Particle System (IPS) approach of Cérou et al. (2006). The objective is to understand the effect on IPS of transforming spontaneous jumps to forced jumps.

Bujorianu (2012) provides an in-depth overview of stochastic reachability analysis for hybrid systems, including GSHS. Stochastic reach probability estimation is a safety verification problem (e.g. Prandini and Hu, 2007; Abate et al., 2009; Lavaei et al., 2021) that has been well studied in the control systems domain and in the safety domain. In the control domain the focus is on developing an (approximate) abstraction of the system for which it can be shown that the reach probability problem is sufficiently similar (Alur et al., 2000; Julius and Pappas, 2009). Approximate abstractions typically make use of a finite partition of the state space (e.g. Prandini and Hu, 2007; Abate et al., 2011; Di Benedetto et al., 2015).

In the safety domain, reach probability is evaluated using a finite partition method or statistical simulation. For realistic applications, the latter requires support from analytical methods to reduce variance. Literature on such variance reduction distinguishes two main approaches: importance sampling (IS) and multi-level importance splitting (ISp). IS draws samples from a reference stochastic system model in combination with analytical compensation for sampling from the reference model instead of the intended model. Bucklew (2004) gives an overview of IS and analytical compensation mechanisms. For complex models analytical compensation mechanisms typically fall short and multi-level ISp is the preferred approach (e.g. Botev and Kroese, 2008; L'Ecuyer et al., 2009; Rubinstein, 2010; Morio and Balesdent, 2016).

The basic idea of multi-level ISp is to enclose the target set, i.e., the set for which the reach probability has to be estimated by a series of nested/enclosing subsets. Each time a simulated particle hits one of the nested subsets, the particle may be split into multiple copies. This multi-level setting allows one to express the small reach probability of the inner level set as a product of larger reach probabilities for the sequence of nested subsets (see, e.g., Glasserman et al, 1999). Cérou et al. (2005, 2006) embedded this multi-level factorization in the Feynman-Kac factorization equation for strong Markov processes (Del Moral, 2004). This Feynman-Kac setting subsequently supported the evaluation of the reach probability through sequential Monte Carlo simulation in the form of an Interacting Particle System (IPS), including proof of convergence (Cérou et al., 2006). Krystul et al. (2012) have used the Feynman-Kac setting to prove convergence of IPS using sampling per mode for a switching diffusion.

Because the theoretical setting of IPS (Cérou et al., 2006) includes strong Markov processes, and a GSHS execution is strong Markov (Bujorianu and Lygeros, 2006), IPS theory applies to GSHS. Blom et al. (2006, 2007a) apply IPS to rare event estimation for an SHS model of an advanced air traffic scenario, which is obtained through applying a Lygeros and Prandini (2010) type of transformation to the underlying GSHS. The hybrid state space of this SHS model is very large, i.e., 490 discrete states and a 28-dimensional Euclidean state space. To prevent particle depletion or impoverishment, a very large number of particles is used. In an attempt to improve the quality of the set of particles, Blom et al. (2007b, 2009) develop and apply a further IPS extension for an SHS with a large number of modes. Complementarily, Prandini et al. (2011) investigate the integration of air traffic complexity model with IPS. For a true GSHS setting, Blom et al. (2018) showed that the use of different numerical integration methods in applying IPS to a true GSHS may have unexpected effects on reach probability estimation. However, these studies did not lead to a basic understanding of the underlying mechanisms. This chapter aims to close this gap in basic understanding.

This chapter is organized as follows. Section 3.2 presents background of GSHS and the transformation to SHS. Section 3.3 reviews IPS theory and presents the algorithmic steps and particle splitting options for an arbitrary GSHS. Section 3.4 specifies three IPS-FAS algorithms for GSHS, two of which make use of the transformation to SHS of Hu et al. (2000). Section 3.5 illustrates results of IPS-FAS algorithms from Section 3.4 applied to a simple GSHS example. Section 3.6 draws conclusions.

3.2. General Stochastic Hybrid System (GSHS)

Throughout this and the following sections, all stochastic processes are defined on a complete stochastic basis $(\Omega,\mathcal{F},\mathbb{F},P,\mathcal{T})$ with $(\Omega,\mathcal{F},P)$ being a complete probability space and $\mathbb{F}$ an increasing sequence of sub-σ-algebras on the time line $\mathcal{T}=\mathbb{R}_+$,i.e., $\mathbb{F}\triangleq\{\mathcal{J},(\mathcal{F}_t,t\in\mathbb{R}_+),\mathcal{F}\}$, with $\mathcal{J}$ containing all P-null sets of $\mathcal{F}$ and $\mathcal{J}\subset\mathcal{F}_s\subset\mathcal{F}_t\subset\mathcal{F}$ for every $s<t$.

3.2.1 GSHS definition

(Bujorianu and Lygeros, 2006) formalized the concept of GSHS or general stochastic hybrid automata as follows:

Definition 1 (GSHS). A GSHS is a collection $(\Theta,d,X,f,g,Init,\lambda,R)$ where

- Θ is a countable set of discrete-valued variables;
- $d : \Theta \to \mathbb{N}$ is a map giving the dimensions of the continuous state spaces;
- $X : \Theta \to \mathbb{R}^{d(\cdot)}$ maps each $\theta \in \Theta$ into an open subset X^{θ} of $\mathbb{R}^{d(\theta)}$;
- $f : \Xi \to \mathbb{R}^{d(\cdot)}$ is a vector field, where $\Xi \triangleq \bigcup_{\theta \in \Theta} \{\theta\} \times X^{\theta}$;
- $g : \Xi \to \mathbb{R}^{d(\cdot) \times m_{\dim}}$ is an $X^{(\cdot)}$-valued matrix, $m_{\dim} \in \mathbb{N}$;
- $\mathrm{Init} : \beta(\Xi) \to [0,1]$ an initial probability measure on Ξ;
- $\lambda : \bar{\Xi} \to \mathbb{R}^{+}$ is a transition rate function;
- $R : \bar{\Xi} \times \beta(\bar{\Xi}) \to [0,1]$ is a transition measure.

3.2.2 GSHS execution

Definition 2 (GSHS Execution). A stochastic process $\{\theta_t, x_t\}$ is called a solution of GSHS execution if there exists a sequence of stopping times $s_0 = 0 < s_1 < s_2 < \ldots$ such that:

- (θ_0, x_0) is a Ξ-valued random variable satisfying the probability measure *Init*;
- For $t \in [s_{j-1}, s_j)$, $j \geq 1$, $\{\theta_t, x_t\}$ is a solution of the SDE:

$$\begin{aligned} d\theta_t &= 0 \\ dx_t &= f(\theta_t, x_t)dt + g(\theta_t, x_t)dW_t \end{aligned} \tag{3.1}$$

  with W_t m -dimensional standard Brownian motion;

- s_j is the minimum of the following two stopping times: i) first hitting time $> s_{j-1}$ of the boundary of $X^{\theta_{s_{j-1}}}$ by the phase process $\{x_t\}$; and ii) first moment $> s_{j-1}$ of a transition event to happen at rate $\lambda(\theta_t, x_t)$.
- At stopping time s_j the novel hybrid state $\{\theta_{s_j}, x_{s_j}\}$ satisfies the conditional probability measure $p_{\theta_{s_j}, x_{s_j} | \theta_{s_j-}, x_{s_j-}}(A | \theta, x) = R((\theta, x), A)$ for any $A \in \beta(\Xi)$.

In order to assure that a GSHS execution has a solution the following assumptions are adopted:

A1 (non-Zeno property): $E\{s_j - s_{j-1}\} > 0$, P -a.s.

A2: For each $(\theta_0, x_0) \in \Xi$, equation (3.1) has a pathwise unique solution on a finite time interval $[0, T]$.

A3 λ is measurable and finite valued.

A4 $Init(\Xi) = 1$, and $R((\theta, x), \Xi) = 1$ for each $(\theta, x) \in \bar{\Xi}$.

Bujorianu and Lygeros (2006) show that the stochastic process $\{\theta_t, x_t\}$ generated by execution of a GSHS satisfies the strong Markov property.

3.2.3 Stochastic analysis background of GSHS execution

Complementary to the probabilistic characterizations of GSHS (Bujorianu and Lygeros, 2006; Bujorianu, 2012), various subclasses of GSHS have been studied as solutions of stochastic differential equations on a hybrid state space that are driven by Brownian motion and Poisson random measure. These studies derive conditions for the existence of pathwise unique solutions, continuity of solutions relative to initial condition (Feller property), and convergent numerical integration schemes.

The best known subclass is Markov switching diffusion (Mao and Yuan, 2006); which forms a GSHS subclass satisfying the following restrictions:

i) There are no boundary hittings, i.e. $X^{\theta} = \mathbb{R}^{d(\theta)}$;

ii) Transition measure R does not support jumps in $\{x_t\}$, i.e. $R(\theta,x;\Theta,dy)) = 0$ if $\{x\} \cap dy = 0$; and

iii) Transition rate function $\lambda(\theta,x)$ is $x-$ invariant.

By dropping the third restriction, we get the subclass of hybrid switching diffusions [Yin and Zhu, 2010]. As is well addressed by Yin and Zhu (2010), the dependency of the mode process $\{\theta_t\}$ on the phase process $\{x_t\}$ asks for complementary derivations regarding existence of pathwise unique solutions and Feller property. (Yin and Zhu, 2010) also show weak converge of an adapted Euler-Maryuama integration scheme to hybrid switching diffusions.

By dropping both restriction ii) and iii), the subclass of hybrid switching diffusions emerges. Pathwise unique solutions have been derived by (Blom, 2003; Ghosh and Bagchi, 2004; Xi et al., 2019). Feller property has been derived by (Krystul et al., 2011; Xi et al., 2019; Kunwai and Zhu, 2020; Blom, 2022). Convergent numerical integration has been addressed by (Krystul, 2006, chapter 4), including approximation of the first hitting time of a boundary. The final step is to also drop restriction i). This allows the generation of instantaneous jumps upon hitting boundaries of X^{θ}; pathwise unique solutions have been addressed by (Krystul et al., 2007).

3.2.4 Probabilistic transformation to an SHS

As explained by Lygeros and Prandini (2010) a GSHS can be transformed to an SHS of Hu et al. (2000). This transformation consists of the following four changes: i) An auxiliary state component q_t, representing "remaining local time", starts at an applicable stopping time τ at initial condition $q_\tau \sim \exp(1)$, and subsequently evolves as $dq_t / dt = -\lambda(\theta_t, x_t)$; ii) The exit boundary of X^{θ} is extended with an extra boundary of the form $q_{t-} = 0$; and iii) Spontaneous probabilistic jumps in $\{x_t, \theta_t\}$ are replaced by forced probabilistic jumps at moment $q_{t-} = 0$; and iv) Upon reaching the extended exit boundary at stopping time τ' the "remaining local time" is resampled, i.e. $q_{\tau'} \sim \exp(1)$.

Hence, transformation of GSHS $(\Theta, d, X, f, g, Init, \lambda, R)$ to SHS $(\Theta^*, d^*, X^*, f^*, g^*, Init^*, R^*)$ works as follows:

- $\Theta^* = \Theta$
- $d^* = d + 1$
- $X^* = X \times (0,\infty)$
- $f^*(\theta,x,.) = \left[f(\theta,x) \quad -\lambda(\theta,x) \right]^T$
- $g^*(\theta,x,.) = \left[g(\theta,x) \quad 0 \right]^T$
- $\text{Init*} = \left[\text{Init} \quad q_0 \right]^T$ with $q_0 \sim \exp(1)$;

- $R^*((\theta, x, .); A \times dq) = R((\theta, x); A) \times e^{-q}\, dq$

Execution of this SHS yields the SHS execution process $\{\theta_t^*, x_t^*, q_t^*\}$, which is a strong Markov process relative to its underlying increasing sequence of sigma algebras $\sigma\{\theta_s^*, x_s^*, q_s^*; s \in [0,t]\}$, $t \in \mathcal{T}$.

It should be noticed that from a stochastic perspective the process $\{\theta_t^*, x_t^*\}$ differs from the process $\{\theta_t, x_t\}$. The key difference is that the sigma algebra $\sigma\{\theta_s^*, x_s^*, q_s^*; s \in [0,t]\}$ includes "remaining local time", which implies (partial) information about the next hitting time of the boundary 0 of $(0, \infty)$, while the sigma algebra $\sigma\{\theta_s, x_s; s \in [0,t]\} \subset \mathcal{F}_t$, i.e. it does not include any information about such future event. To avoid abusing the extra information, the "remaining local time" component $\{q_t^*\}$ should be treated as being unobservable for other processes that depend on the GSHS execution.

3.3. IPS based reach probability estimation

3.3.1. GSHS reach probability

The problem is to estimate the probability γ that $\{\theta_t, x_t\}$ reaches a closed subset $D \subset \Xi$ within finite period $[0,T]$, i.e.

$$\gamma = P(\tau < T) \tag{3.2}$$

with τ being the first hitting time of D by $\{\theta_t, x_t\}$:

$$\tau = \inf\{t > 0, (\theta_t, x_t) \in D\} \tag{3.3}$$

Remark: Cérou et al. (2006) and L'Equyer et al. (2009) also address the more general situation that *T* is a P-a.s. finite stopping time.

Cérou et al. (2006) developed the IPS theory and algorithmic steps for estimating reach probability for a strong Markov process on a general Polish state space. Thanks to the strong Markov property of the process $\{\theta_t, x_t\}$ defined by the execution of the GSHS in section 3.2, the IPS approach applies to the estimation of GSHS reach probability.

3.3.2. Multi-level factorization of reach probability

The principle in factorizing the reach probability $\gamma = P(\tau < T)$ is to introduce a sequence D_k, $k = 0,..,m$, of nested closed subsets of Ξ, i.e. $D = D_m \subset D_{m-1} \subset \cdots \subset D_1 \subset D_0 = \Xi$, with D_1 such that $P\{(\theta_0, x_0) \in D_1\} = 0$. Let τ_k be the first moment in time that $\{\theta_t, x_t\}$ reaches D_k, i.e.

$$\tau_k = \inf\{t > 0;\ (\theta_t, x_t) \in D_k \vee t \geq T\} \tag{3.4}$$

Next, we define $\{0,1\}$-valued random variables $\{\chi_k, k = 0,..,m\}$ as follows:

$$\begin{aligned} \chi_k &= 1, \text{ if } \tau_k < T \text{ or } k = 0 \\ &= 0, \text{ else} \end{aligned} \tag{3.5}$$

By using this χ_k definition we get the desired factorization.

Proposition 3.1:

The reach probability satisfies the factorization:

$$\gamma = \prod_{k=1}^{m} \gamma_k \tag{3.6}$$

where $\gamma_k \triangleq E\{\chi_k = 1 \big| \chi_{k-1} = 1\} = \mathrm{P}(\tau_k < T \big| \tau_{k-1} < T)$.

Proof: Because $D_{k-1} \supset D_k$ we have:

$$\inf\{t > 0;\ (\theta_t, x_t) \in D_{k-1} \vee t \geq T\} \leq \inf\{t > 0;\ (\theta_t, x_t) \in D_k \vee t \geq T\}$$

Substituting (3.4) at left and at right yields: $\tau_{k-1} \leq \tau_k$.

Hence we can derive:

$$\begin{aligned}
\gamma &= \mathrm{P}(\tau < T) = \mathrm{P}\left(\tau_m < T\right) \\
&= \mathrm{P}\left(\tau_m < T \ \wedge\ \tau_{m-1} \leq \tau_m\right) = \mathrm{P}\left(\tau_m < T \ \wedge\ \tau_{m-1} < T\right) \\
&= P\left(\tau_m < T \mid \tau_{m-1} < T\right) P\left(\tau_{m-1} < T\right) \\
&= \prod_{k=1}^{m} \mathrm{P}\left(\tau_k < T \mid \tau_{k-1} < T\right) P\left(\tau_0 < T\right) \\
&= \prod_{k=1}^{m} \mathrm{P}\left(\tau_k < T \mid \tau_{k-1} < T\right) \\
&= \prod_{k=1}^{m} E\{\chi_k = 1 \mid \chi_{k-1} = 1\} = \prod_{k=1}^{m} \gamma_k
\end{aligned}$$

Q.E.D.

3.3.3. Recursive estimation of the multi-level factors

By using the strong Markov property of $\{\theta_t, x_t\}$, we develop a recursive estimation of γ using the factorization in (3.6). First we define $\Xi' \triangleq \mathbb{R} \times \Xi$, $\xi_k \triangleq (\tau_k, \theta_{\tau_k}, x_{\tau_k})$, $Q_k \triangleq (0,T) \times D_k$, for $k = 1,\ldots,m,$ and the following conditional probability measure $\pi_k(B)$ for an arbitrary Borel set B of Ξ':

$$\pi_k(B) \triangleq P(\xi_k \in B \mid \xi_k \in Q_k)$$

Cérou et al. (2006) show that π_k is a solution of the following recursion of transformations:

$$\begin{array}{ccccc}
\pi_{k-1}(\cdot) & \xrightarrow{\text{I. mutation}} & p_k(\cdot) & \xrightarrow{\text{III. selection}} & \pi_k(\cdot) \\
 & & \downarrow \scriptstyle{\text{II. conditioning}} & & \\
 & & \gamma_k & &
\end{array}$$

where $p_k(B)$ is the conditional probability measure of $\xi_k \in B$ given $\xi_{k-1} \in Q_{k-1}$, i.e.,

$$p_k(B) \triangleq P(\xi_k \in B \mid \xi_{k-1} \in Q_{k-1})$$

Because $\{\theta_t, x_t\}$ is a strong Markov process, $\{\xi_k\}$ is a Markov sequence. Hence, the mutation transformation (I) satisfies a Chapman-Kolmogorov equation prediction for ξ_k:

$$p_k(B) = \int_{\Xi'} p_{\xi_k|\xi_{k-1}}(B \mid \xi)\pi_{k-1}(d\xi) \text{ for all } B \in \beta(\Xi') \tag{3.7}$$

For the conditioning transformation (II) this means:

$$\gamma_k = \mathrm{P}(\tau_k < T \mid \tau_{k-1} < T) = \int_{\Xi'} 1_{\{\xi \in Q_k\}} p_k(d\xi). \tag{3.8}$$

Hence, selection transformation (III) satisfies:

$$\pi_k(B) = \frac{\int_B 1_{\{\xi \in Q_k\}} p_k(d\xi)}{\int_{\Xi'} 1_{\{\xi' \in Q_k\}} p_k(d\xi')} = [\int_B 1_{\{\xi \in Q_k\}} p_k(d\xi)] / \gamma_k \tag{3.9}$$

With this, the γ_k terms in (3.6) are characterized as solutions of a recursive sequence of mutation equation (3.7), conditioning equation (3.8) and selection equation (3.9).

3.3.4. IPS algorithmic steps for a GSHS

Following Cérou et al. (2006), equations (3.6)-(3.9) yield the IPS algorithmic steps for the numerical estimation of γ :

$$\bar{\pi}_{k-1}(\cdot) \xrightarrow{\text{I. mutation}} \bar{p}_k(\cdot) \xrightarrow{\text{III. selection}} \tilde{\pi}_k(\cdot) \xrightarrow{\text{IV. splitting}} \bar{\pi}_k(\cdot)$$
$$\bar{p}_k(\cdot) \xrightarrow{\text{II. conditioning}} \bar{\gamma}_k$$

A set of N_P particles is used to form empirical density approximations $\bar{\gamma}_k$, $\bar{p}_k$ and $\bar{\pi}_k$ of γ_k, p_k and π_k respectively. By increasing the number N_P of particles in a set, the errors in these approximations decrease. When simulating particles from Q_{k-1} to Q_k, only a fraction $\bar{\gamma}_k$ of the simulated particle trajectories will reach Q_k within the time period $[0,T]$ considered; these particles form $\tilde{\pi}_k$. In order to start the next IPS cycle with N_P particles, the classical way is to perform a multinomial resampling (MR) of $\tilde{\pi}_k$ to produce $\bar{\pi}_k$. More effective splitting methods are: multinomial splitting (MS), residual multinomial splitting RMS) and fixed assignment splitting (FAS). MS generates $\bar{\pi}_k$ by starting with the particles in $\tilde{\pi}_k$, and subsequently adding randomly selected particles from $\tilde{\pi}_k$ (with replacement). RMS first makes $\lfloor 1/\bar{\gamma}_k \rfloor$ copies from each particle in $\tilde{\pi}_k$, and subsequently complements the residual number $N_P\left(1-\bar{\gamma}_k \lfloor 1/\bar{\gamma}_k \rfloor\right)$ by randomly selected particles from $\tilde{\pi}_k$ (with replacement). FAS also follows the two step approach of RMS, though during the second step the random selection from $\tilde{\pi}_k$ is done without replacement.

Cérou et al. (2006) prove that using IPS with multinomial splitting (MS) for a strong hybrid state Markov process, $\bar{\gamma}$ forms an unbiased γ estimate, i.e.

$$\mathbb{E}\{\bar{\gamma}\} = \mathbb{E}\left\{\prod_{k=1}^{m} \bar{\gamma}_k\right\} = \prod_{k=1}^{m} \mathbb{E}\{\bar{\gamma}_k\} = \prod_{k=1}^{m} \gamma_k = \gamma \tag{3.10}$$

Moreover, Cérou et al. (2006) derive second and higher order asymptotic bounds for the error $(\bar{\gamma} - \gamma)$ based on the multi-level Feynman Kac analysis, e.g. Del Moral (2004; Theorem 12.2.2).

For a diffusion process $\{x_t\}$, Ma and Blom (2022) have proven that IPS using FAS yields a lower or equal variance in the estimated reach probability $\bar{\gamma}$ than IPS using MR, MS or RMS. In the next section we extend these results for an IPS applied to a GSHS.

3.4. IPS algorithmic steps for GSHS

3.4.1. IPS application for a GSHS.

The algorithmic steps of IPS application for a GSHS are specified in Algorithm 3.1 below. For the splitting step IV, use is made of FAS.

Algorithm 3.1; IPS-FAS algorithmic steps for a GSHS

Input: Initial measure π_0, end time T, decreasing sequence of closed subsets $D_k = \{(\theta_t, x_t) \in \Xi\}$, $D_{k-1} \supset D_k$, $k = 1,..,m$. Also $D_0 = \Xi$, $Q_k = (0,T) \times D_k$ and number of particles N_P.

Output: Estimated reach probability $\bar{\gamma}$

0. Initiation: Generate N_P particles $\xi_0^i \sim \pi_0$, $i = 1,..,N_P$, i.e. $\bar{\pi}_0(\cdot) = \sum_{i=1}^{N_p} \frac{1}{N_p} \delta_{\{\xi_0^i\}}(\cdot)$, with Dirac δ. Set $k = 1$.

I. Mutation: $\bar{p}_k(.) = \sum_{i=1}^{N_p} \frac{1}{N_p} \delta_{\{\bar{\xi}_k^i\}}(\cdot)$, where $\bar{\xi}_k^i$ is obtained by simulating the GSHS execution starting from ξ_{k-1}^i.

II. Conditioning: $\bar{\gamma}_k = \frac{N_{S_k}}{N_p}$ with $N_{S_k} = \sum_{i=1}^{N_p} 1(\bar{\xi}_k^i \in Q_k)$. If $N_{S_k} = 0$, then $\bar{\gamma}_{k'} = 0, k' \in \{k,...,m\}$ and go to Step V.

III. Selection: $\tilde{\pi}_k(.) = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \delta_{\{\tilde{\xi}_k^i\}}(\cdot)$, with $\{\tilde{\xi}_k^j\}_{j=1}^{N_{S_k}}$ the collection of $\bar{\xi}_k^i \in Q_k$, $i = 1,..,N_P$.

IV. Splitting: $\{\tilde{\tilde{\xi}}_k^j\}_{j=1}^{N_{S_k}}$ is a random permutations of $\{\tilde{\xi}_k^j\}_{j=1}^{N_{S_k}}$.

Copy:

$$\xi_k^i = \tilde{\tilde{\xi}}_k^i \text{ for } i = 1,...,N_{S_k};$$

$$\xi_k^{N_{S_k}+i} = \tilde{\tilde{\xi}}_k^i \text{ for } i = 1,...,N_{S_k};$$

$$\ldots$$

$$\xi_k^{(\lfloor N_P/N_{S_k} \rfloor - 1)N_{S_k}+i} = \tilde{\tilde{\xi}}_k^i \text{ for } i = 1,...,N_{S_k};$$

$$\xi_k^{\lfloor N_P/N_{S_k} \rfloor N_{S_k}+i} = \tilde{\tilde{\xi}}_k^i \text{ for } i = 1,...,N_p - \lfloor N_p / N_{S_k} \rfloor N_{S_k}.$$

Each particle receives weight $1/N_p$.

V. If $k < m$, then $k := k+1$ and go to step I, else $\bar{\gamma} = \prod_{k=1}^{m} \bar{\gamma}_k$

By extending the results of (Ma and Blom, 2022) for IPS application to a diffusion, in Appendix 3.A we proof the following regarding the use of different splitting methods in IPS application to GSHS.

Theorem 3.1: Replacing the FAS splitting step IV in algorithm 3.1 by RMS splitting, MS splitting or MR splitting has the following effects on the variance $\mathrm{V}\{\bar{\gamma}\}$:

$$V_{FAS}\{\overline{\gamma}\} \le V_{RMS}\{\overline{\gamma}\} \le V_{MS}\{\overline{\gamma}\} \le V_{MR}\{\overline{\gamma}\} \quad (3.11)$$

Proof: See Appendix 3.A.

Next we address the details of mutation step I of Algorithm 3.1, i.e. the Monte Carlo simulation of the GSHS from particle state ξ_{k-1}^i to particle state $\overline{\xi}_k^i$. Subsection 3.4.2 addresses simulation of the execution of an SHS transformed version of GSHS within IPS. Subsection 3.4.3 develops an algorithm that takes into account that the "remaining local time" process $\{q_t^*\}$ should be unobservable for the IPS process. For reference purpose, subsection 3.4.3 addresses the more demanding direct simulation of the execution of a GSHS, i.e. without using the transformation to SHS.

3.4.2. Simulation of execution of SHS transformed version of GSHS in mutation step I

The process $\{\theta_t, x_t\}$ is assumed to be the SHS transformed version of the GSHS, i.e. $\{\theta_t, x_t\} = \{\theta_t^*, x_t^*, q_t^*\}$ as defined in subsection 3.2.3. Then in step I of Algorithm 3.1, the evolution of $\{\theta_t, x_t\} = \{\theta_t^*, x_t^*, q_t^*\}$ is executed on interval $[\tau_{k-1}^i, \tau_k^i]$, starting with ξ_{k-1}^i and delivering ξ_k^i. Mutation step I is conducted using Euler-Maruyama integration of eq. (3.1) along small time steps $\vartriangle$, i.e.

$$\begin{aligned} \theta_{t+\vartriangle} &= \theta_t \\ x_{t+\vartriangle} &= f(\theta_t, x_t)\vartriangle + g(\theta_t, x_t)(W_{t+\vartriangle} - W_t) \end{aligned} \quad (3.12)$$

The algorithm for the execution of an SHS transformed version of GSHS within mutation step I is specified below.

Remark: Convergence of the Euler-Maruyama integration scheme (3.12) is guaranteed iff the SDE coefficients satisfy certain Lipschitz conditions, e.g. Hutzenthaler et al. (2011).

Algorithm 3.2. Simulating the execution of SHS transformed version of GSHS in step I of Algorithm 3.1

Input: *i*-th particle vector $\xi_{k-1}^i = (\tau_{k-1}^i, \theta_{k-1}^{*i}, x_{k-1}^{*i}, q_{k-1}^{*i})$, and the SHS elements $(\Theta^*, d^*, X^*, f^*, g^*, Init^*, R^*)$ and $Q_k^* = Q_k \times \mathbb{R}$.

Output: Estimated particle $\overline{\xi}_k^i = (\tau_k^i, \overline{\theta}_k^{*i}, \overline{x}_k^{*i}, \overline{q}_k^{*i})$

1. Set $t := \tau_{k-1}^i$ and $\overline{\varsigma} := (\theta_{k-1}^{*i}, x_{k-1}^{*i}, q_{k-1}^{*i})$

2. Evaluate equation (1) and $dq_t / dt = -\lambda(\theta_t, x_t)$ from $\overline{\varsigma}$ at t until $t_+ = \min\{t + \vartriangle, \overline{s}_t, \overline{\tau}_k\}$; this yields $\overline{\varsigma}_+$. Here $\overline{s}_t$ is the first time $> t$ that this solution hits the boundary of X^*; and $\overline{\tau}_k$ is the first time that this solution hits Q_k^*.

3. If $t_+ \ge \overline{\tau}_k$ then stop with output $\overline{\xi}_k^i = (\overline{\tau}_k, \overline{\theta}_k^{*i}, \overline{x}_k^{*i}, \overline{q}_k^{*i})$, where $(\overline{\theta}_k^{*i}, \overline{x}_k^{*i}, \overline{q}_k^{*i}) \sim R^*(\overline{\varsigma}_+, .)$ if $\overline{s}_t = \overline{\tau}_k$, else $(\overline{\theta}_k^{*i}, \overline{x}_k^{*i}, \overline{q}_k^{*i}) := \overline{\varsigma}_+$.

4. If $t_+ \ge \overline{s}_t$ then $\overline{\varsigma} \sim R^*(\overline{\varsigma}_+, .)$, set $t := t_+$ and repeat from step 2.

If during any of the small time steps Δ one of the boundaries of X^* or Q_k^* is passed, then additional MC simulation steps may be conducted to get a better approximation $\bar{s}_t$ or $\bar{\tau}_k$ of the first hitting time. As an alternative for using a lower Δ value, Glasserman (2004, p. 367) proposes an interpolation of the solution of equation (3.1) on the Δ interval considered, by simulating a Brownian bridge between the already simulated Brownian motion points W_t and $W_{t+\Delta}$. The resulting Brownian bridge yields a more accurate approximation of the first hitting time.

3.4.3 Accounting for unobservability of remaining local time

As has been identified at the end of subsection 3.2.4, the "remaining local time" process $\{q_t^*\}$ of the SHS transformed version of a GSHS should be treated as being unobservable for the IPS process. To formalize this, the transformation to SHS is applied to a enriched version of the original GSHS. The GSHS enrichment consists of adding IPS hitting levels Q_k, $k=1,..,m$, to the original GSHS, with reset $(\theta_{\tau_k}, x_{\tau_k}) = (\theta_{\tau_k -}, x_{\tau_k -})$, at a hitting time τ_k. Thanks to the continuity of the latter reset, the execution of the enriched GSHS yields the same pathwise solutions as execution of the original GSHS does. Subsequent application of the transformation of Prandini and Lygeros (2010) to this enriched GSHS yields a SHS, that also resets the remaining local time upon reaching an IPS hitting level Q_k, $k=1,..,m$. For algorithm 3.2 this means that it can be improved by adding a reset of local remaining time at the beginning of each IPS cycle; this is specified in algorithm 3.3 below. Hence, at the begin of mutation step I within an IPS cycle, the remaining local time value of each particle is freshly sampled from exp(1).

Algorithm 3.3. Simulating execution of SHS version of modified GSHS in step I of Algorithm 3.1

Input: *i*-th particle vector $\xi_{k-1}^i = (\tau_{k-1}^i, \theta_{k-1}^{*i}, x_{k-1}^{*i}, q_{k-1}^{*i})$, and the SHS elements $(\Theta^*, d^*, X^*, f^*, g^*, Init^*, R^*)$ and $Q_k^* = Q_k \times \mathbb{R}$.
Output: Estimated particle $\bar{\xi}_k^i = (\tau_k^i, \bar{\theta}_k^{*i}, \bar{x}_k^{*i}, \bar{q}_k^{*i})$

1. Set $t := \tau_{k-1}^i$ and $\bar{\varsigma} := (\theta_{k-1}^{*i}, x_{k-1}^{*i}, \bar{q})$, with $\bar{q} \sim \exp(1)$
2. = step 2 in algorithm 3.2.
3. = step 3 in algorithm 3.2.
4. = step 4 in algorithm 3.2.

The combination of algorithms 3.1&3.3 starts at each IPS cycle with N_p particles, each of which has a different sample of remaining local time q_{k-1}^{*i}, $k=1,..m$. This differs significantly from the combination of algorithm combination 3.1&3.2, where the N_p particles having different remaining local time q_0^{*i} applies at the start of the first IPS cycle only. Hence, with increasing IPS level *k*, under algorithm combination 3.1&3.3 particle diversity will gain relative to particle diversity under algorithm combination 3.1&3.2.

3.4.4. Simulation of original GSHS execution in mutation step I

For reference purpose, we also specify an algorithm for the simulation of the original GSHS execution. For this we follow the numerical integration scheme of Krystul (2004, Chapter 4). In addition to fixed small time steps Δ, random time steps are generated at which potential jumps may happen. Realizations of the these random time steps are obtained through Monte Carlo sampling of an in-homogeneous Poisson process on $[0,T]\times[0,\bar{\lambda}]$, with $\bar{\lambda} \geq \sup_{(\theta,x)\in\Xi} \lambda(\theta,x)$. Subsequently the potential Poisson points are thinned by rejecting points that lie above the graph of $\lambda(\theta_t,x_t)$. The remaining points, i.e., those at or below the graph of $\lambda(\theta_t,x_t)$, are projected onto the *time*-axis $[0,T]$. The resulting execution of the GSHS within an IPS cycle, starting from ξ^i_{k-1}, on the interval $[\tau^i_{k-1},\tau^i_k]$ is specified in algorithm 3.4.

Algorithm 3.4. Simulating GSHS execution (step I of algorithm 3.1)

Input: *i*-th particle vector $\xi^i_{k-1}=(\tau^i_{k-1},\theta^i_{k-1},x^i_{k-1})$, Q_k , the GSHS elements $(\Theta,d,X,f,g,Init,\lambda,R)$, and the transition rate maximum $\bar{\lambda}$.

Output: Estimated particle $\bar{\xi}^i_k=(\bar{\tau}^i_k,\bar{\theta}^i_k,\bar{x}^i_k)$

1. Set $t=\tau^i_{k-1}$ and $\bar{\varsigma}:=(\theta^i_{k-1},x^i_{k-1})$
2. Generate $u\sim \mathrm{U}(0,1)$, and set $\Delta_t:=-(\ln u)/\bar{\lambda}$
3. Evaluate equation (3.1) from $\bar{\varsigma}$ at t until $t_+=\min\{t+\Delta,t+\Delta_t,\bar{s}_t,\bar{\tau}_k\}$; this yields $\bar{\varsigma}_+$. Here $\bar{s}_t$ is the first time $>t$ that this solution hits the boundary of X; and $\bar{\tau}_k$ is the first time that this solution hits Q_k.
4. If $t_+\geq\bar{\tau}_k$ then stop with output $\bar{\xi}^i_k=(\bar{\tau}^i_k,\bar{\theta}^i_k,\bar{x}^i_k)$, where $(\bar{\theta}^i_k,\bar{x}^i_k)\sim R(\bar{\varsigma}_+,.)$ if $\bar{s}_t=\bar{\tau}_k$, else $(\bar{\theta}^i_k,\bar{x}^i_k):=\bar{\varsigma}_+$
5. If $t_+\geq\bar{s}_t$ then $\bar{\varsigma}\sim R(\bar{\varsigma}_+,.)$, set $t:=t_+$ and repeat from step 2
6. If $t_+\geq t+\Delta_t$ then generate $v\sim \mathrm{U}(0,1)$
7. If $\lambda(\bar{\varsigma}_+)\geq v\bar{\lambda}$, then generate $\bar{\varsigma}\sim R(\bar{\varsigma}_+,(.,.))$, else $\bar{\varsigma}:=\bar{\varsigma}_+$
8. Set $t:=t_+$ and repeat from step 2.

In case of a stop during step 4 of GSHS algorithm 3.4, there is a "remaining integration time" $t+\Delta_t-\bar{\tau}_k$. Because this "remaining integration time" does not make part of the Markov state $\xi_{\bar{\tau}_k}$, it does not influence the GSHS execution during the next IPS cycle. The latter coincides with ignoring "remaining local time" in algorithm 3.3. Hence it is expected that algorithm combination 3.1&3.4 estimates reach probability similarly well as algorithm combination 3.1&3.3 does.

3.5. Application of IPS to GSHS example

3.5.1. Hypothetical car example

A car driver in dense fog is heading to a wall at position d_{wall}. If the car is at distance d_{fog} from the wall, then the driver sees the wall for the first time. Then, it takes the driver a random reaction delay to start braking, with a density $p_{delay}(s)$. During the reaction delay, the velocity of the car does not change; after the reaction delay, the car decelerates at constant value $a_{\min}$. The aim is to estimate the probability γ that the car hits the wall.

From the moment that the car reaches distance d_{fog} from the wall at velocity v_0, it takes the sum of reaction delay T_{delay} and the time of deceleration $T_{dec}=-v_0/a_{\min}$ until the car is at a standstill. This implies

$$\gamma=\mathrm{P}\{v_0T_{delay}+v_0T_{dec}+\tfrac{1}{2}a_{\min}T^2_{dec}\ \geq\ d_{fog}\}\tag{3.13}$$

Elaboration of (3.13) yields:

$$\gamma = \mathrm{P}\{T_{delay} \geq \tfrac{1}{2} v_0 / a_{\min} + d_{fog} / v_0\} \tag{3.14}$$

If we assume a Rayleigh density $p_{delay}(s) = \frac{s}{\mu^2} e^{-s^2/(2\mu^2)}$, and we write $T_C = \tfrac{1}{2} v_0 / a_{\min} + d_{fog} / v_0$, evaluation of (3.14) yields:

$$\gamma = \int_{T_C}^{+\infty} \frac{t}{\mu^2} e^{-t^2/(2\mu^2)} dt = -e^{-t^2/(2\mu^2)} \Big|_{t=T_C}^{+\infty} = e^{-t^2/(2\mu^2)} \big|_{t=T_C} \tag{3.15}$$

Table 3.1 gives the analytically obtained γ results for various mean reaction delays μ, and parameter settings $d_{wall} = 300m,\ d_{fog} = 120m,\ v_0 = 72 \,{}^{km}\!/_{h} = 20 \,{}^{m}\!/_{s},\ a_{\min} = -4 \,{}^{m}\!/_{s^2}$.

Table 3.1 Analytical γ results for various μ

μ (s)	γ
0.9	5.19976×10^{-4}
0.8	6.97696×10^{-5}
0.7	3.72665×10^{-6}
0.6	4.08284×10^{-8}

For this example, subsection 3.5.2 specifies the GSHS model and the transformation of subsection 3.2.4 to an SHS model. Subsection 3.5.3 estimates γ using straightforward MC simulation and IPS-FAS algorithm combinations 3.1&3.2, 3.1&3.3 and 3.1&3.4.

3.5.2. GSHS model and transformation to SHS model

For this example, the discrete set of the GSHS is:

$$\Theta = \{-1, 0, delay, stop, hit\} \tag{3.16}$$

where -1 indicates decelerating mode, 0 indicates uniform mode, $delay$ is a reaction delay mode, $stop$ indicates stopping mode, and hit indicates the wall has been hit. A transition diagram representing the transitions between these modes is given in Figure 3.1.

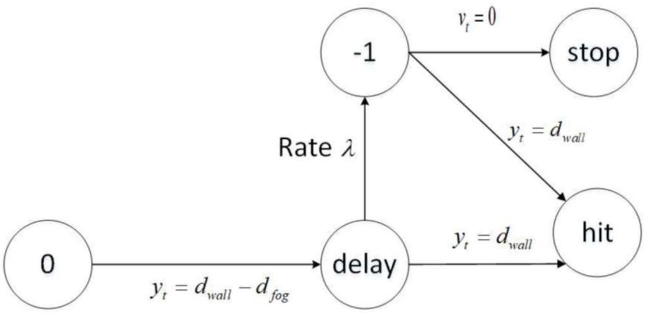

Figure 3.1. State transition diagram of GSHS model.

The continuous state components are $x_t = Col(z_t, y_t, v_t)$, where z_t is the amount of time passed since the driver could see the wall for the first time, y_t is the position of the car at time t, and v_t is the velocity at time t. Hence, the dimension of the continuous state space is $d(.) = 3$. The subsets X^{θ} are defined as follows:

$$
\begin{aligned}
X^0 &= \mathbb{R}\times(-\infty, d_{wall}-d_{fog})\times\mathbb{R} \\
X^{-1} &= \mathbb{R}\times(-\infty, d_{wall})\times(0,\infty) \\
X^{delay} &= \mathbb{R}\times(-\infty, d_{wall})\times\mathbb{R} \\
X^{stop} &= \mathbb{R}\times(-\infty, d_{wall})\times 0 \\
X^{hit} &= \mathbb{R}^3
\end{aligned} \tag{3.17}
$$

The initial measure $Init$ generates $\theta_0=0,\ z_0=0,\ y_0=0$. Between switching moment of $\{\theta_t\}$, x_t evolves as (3.1) with $f\left(\theta,[z,y,v]^T\right)=[1,v,1\{\theta=-1\}a_{\min}]^T$ and $g(\theta,.)=[0,g_2,0]^T$ if $\theta\in\{0,delay,-1\}$, else $g(\theta,.)=[0,0,0]^T$. The analytical results in Table 3.1 apply for $g_2=0$, i.e. no Brownian motion.

The instantaneous transition rate $\lambda(\theta_t,(z_t,y_t,v_t))$ satisfies:

$$
\lambda(\theta,(z,y,v)) = \chi(\theta=delay)p_{delay}(z)\Big/\int_z^\infty p_{delay}(s)ds \tag{3.18}
$$

The transition measure $R((\theta,(z,y,v)),(.,.))$ satisfies:

$R((-1,(z,y,v)),\{stop\}\times\{0,y,v\})=1$ iff $v=0$

$R((0,(z,y,v)),\{delay\}\times\{0,y,v\})=1$ iff $y=d_{wall}-d_{fog}$

$R((delay,(z,y,v)),\{-1\}\times\{0,y,v\})=1$, iff λ generates a point,

$R((delay,(z,y,v)),\{hit\}\times\{0,y,0\})=1$, iff $y=d_{wall}$

$R((-1,(z,y,v)),\{hit\}\times\{0,y,0\})=1$, iff $y=d_{wall}$.

IPS-FAS algorithm combination 3.1&3.4 makes use of this GSHS model. By applying the transformation from subsection 3.2.3, the above GSHS model transforms to an SHS model. The resulting SHS has continuous state components (z_t,y_t,v_t,q_t), with $\{q_t\}$ evolving as $dq_t=-\lambda(\theta_t,(z_t,y_t,v_t))dt$ in between discontinuities, and $q_s\sim\exp(1)$ at a mode switch and if q_{t-} hits 0.

IPS-FAS algorithm combinations 3.1&3.2 and 3.1&3.3 make use of this SHS transformed version of the GSHS model. Though algorithm combination 3.1&3.3 also refreshes the "remaining time" q_{τ_k-} at the start of a mutation during the next IPS cycle.

3.5.3. Simulation results

By conducting each of the approaches $N_{\bar\gamma}$ times we get $\bar\gamma^i, i=1,..,N_{\bar\gamma}$. These results are used to assess the mean $\hat\gamma$, the percentage ρ_S of successful IPS runs, and the normalized root-mean-square error (RMSE), i.e.

$$
\hat\gamma = \frac{1}{N_{\bar\gamma}}\sum_{i=1}^{N_{\bar\gamma}}\bar\gamma^i \tag{3.19}
$$

$$
\rho_S = \frac{1}{N_{\bar\gamma}}\sum_{i=1}^{N_{\bar\gamma}}1(\bar\gamma^i>0) \tag{3.20}
$$

$$
RMSE = \sqrt{\frac{1}{N_{\bar\gamma}}\sum_{i=1}^{N_{\bar\gamma}}(\bar\gamma^i-\gamma)^2} \tag{3.21}
$$

In the subsequent IPS cycles the following levels are used: $D_k = \{0, delay, hit\} \times \mathbb{R} \times [L_k, \infty) \times \mathbb{R} \cup \{-1, stop\} \times \mathbb{R} \times [d_{wall}, \infty) \times \mathbb{R}$, with the μ -dependent L_k values shown in Table 3.2.

Table 3.2 Values of L_k for various μ values

k \ μ	$0.9s$	$0.8s$	$0.7s$	$0.6s$
1	181	181	181	181
2	217	215	210	205
3	230	230	220	215
4	240	241	230	223
5	300	300	237	230
6			244	236
7			300	243
8				300

Table 3.3. Simulation results for MC and IPS-FAS algorithm combinations 3.1&3.2, 3.1&3.3 and 3.1&3.4 applied to the GSHS example $g_2 = 0$ at simulation settings $\Delta = 0.01s$, $N_p = 1000$, and $N_{\bar{\gamma}} = 100$.

$\mu = 0.9s$	$\hat{\gamma}$	ρ_S	$RMSE/\gamma$
MC (m=1)	5.300×10^{-4}	44%	137.2%
IPS-FAS combination 1&2	3.859×10^{-4}	33%	116.7%
IPS-FAS combination 1&3	5.096×10^{-4}	100%	13.4%
IPS-FAS combination 1&4	5.125×10^{-4}	100%	15.2%
$\mu = 0.8s$	$\hat{\gamma}$	ρ_S	$RMSE/\gamma$
MC (m=1)	4.000×10^{-5}	4%	284.1%
IPS-FAS combination 1&2	3.811×10^{-5}	4%	271.7%
IPS-FAS combination 1&3	6.968×10^{-5}	100%	20.6%
IPS-FAS combination 1&4	6.948×10^{-5}	100%	19.8%
$\mu = 0.7s$	$\hat{\gamma}$	ρ_S	$RMSE/\gamma$
MC (m=1)	/	/	/
IPS-FAS combination 1&2	/	/	/
IPS-FAS combination 1&3	3.605×10^{-6}	100%	20.9%
IPS-FAS combination 1&4	3.757×10^{-6}	100%	20.4%

$\mu = 0.6s$	$\hat{\gamma}$	ρ_S	$RMSE/\gamma$
MC (m=1)	/	/	/
IPS-FAS combination 1&2	/	/	/
IPS-FAS combination 1&3	4.055×10^{-8}	100%	28.30%
IPS-FAS combination 1&4	4.029×10^{-8}	100%	28.47%

Table 3.4. Simulation results for MC and IPS-FAS algorithm combinations 3.1&3.2, 3.1&3.3 and 3.1&3.4 applied to the GSHS example $g_2 = 1$ at simulation settings $\Delta = 0.01s$, $N_p = 1000$, and $N_{\hat{\gamma}} = 100$.

$\mu = 0.9s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC (m=1)	7.000×10^{-4}	50%	120.37%
IPS-FAS combination 1&2	6.306×10^{-4}	94%	110.11%
IPS-FAS combination 1&3	6.829×10^{-4}	100%	13.95%
IPS-FAS combination 1&4	6.832×10^{-4}	100%	15.60%
$\mu = 0.8s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC (m=1)	4.000×10^{-5}	3%	604.15%
IPS-FAS combination 1&2	1.266×10^{-4}	49%	244.33%
IPS-FAS combination 1&3	1.027×10^{-4}	100%	18.62%
IPS-FAS combination 1&4	1.022×10^{-4}	100%	17.37%
$\mu = 0.7s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC (m=1)	1.000×10^{-5}	1%	994.99%
IPS-FAS combination 1&2	1.316×10^{-5}	14%	666.32%
IPS-FAS combination 1&3	6.921×10^{-6}	100%	16.52%
IPS-FAS combination 1&4	7.021×10^{-6}	100%	18.93%
$\mu = 0.6s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC (m=1)	/	/	/
IPS-FAS combination 1&2	/	/	/
IPS-FAS combination 1&3	1.199×10^{-7}	100%	28.34%
IPS-FAS combination 1&4	1.140×10^{-7}	100%	25.39%

For $g_2 = 0$ and $g_2 = 1$, Table 3.3 and Table 3.4 respectively show simulation results of straightforward MC and of IPS-FAS using algorithm combinations 3.1&3.2, 3.1&3.3 and 3.1&3.4. These results show that IPS-FAS combination 3.1&3.2 performs similar or slightly better than straightforward MC simulation. Both in Table 3.3 and in Table 3.4, IPS-FAS combinations 3.1&3.3 and 3.1&3.4 perform far better than MC and IPS-FAS combination 3.1&3.2.

For $g_2 = 0$ and $\mu = 0.8s$, Tables 3.5, 3.6 and 3.7 present average counts of particles per IPS level, over successful IPS-FAS runs of algorithm combinations 3.1&3.2, 3.1&3.3 and 3.1&3.4 respectively. Comparison of Tables 3.5 and 3.6 show a steady increase in particle diversity under algorithm combination 3.1&3.3 relative to combination 3.1&3.2. Comparison of Tables 3.6 and 3.7 show that diversity of particles after mutation step I is similar under algorithm combinations 3.1&3.3 and 3.1&3.4.

Table 3.5. Average counts of particles per level over successful IPS-FAS runs of combination 3.1&3.2, for $g_2 = 0$, $\mu = 0.8s$.

k	No. of particles at start of Step I	No. of different particles at start of Step I	No. of different particles after Step I	No. of survived particles after Step III	No. of different particles after Step III	% of successful IPS runs through level k
1	1000	1	999.99	997.96	997.96	100%
2	1000	997.96	971.46	92.84	92.62	100%
3	1000	92.62	92.15	82.58	7.59	100%
4	1000	7.59	13.26	184.40	1.41	49%
5	1000	1.41	18.51	592.16	1	6%

Table 3.6. Average counts of particles per level over successful IPS-FAS runs of combination 3.1&3.3, for $g_2 = 0$, $\mu = 0.8s$.

k	No. of particles at start of Step I	No. of different particles at start of Step I	No. of different particles after Step I	No. of survived particles after Step III	No. of different particles after Step III	% of successful IPS runs through level k
1	1000	1	999.98	998.25	998.25	100%
2	1000	998.25	974.03	92.14	92.14	100%
3	1000	92.14	929.76	82.10	82.10	100%
4	1000	82.10	906.45	92.28	92.28	100%
5	1000	92.28	889.19	100.17	98.80	100%

Table 3.7. Average counts of particles per level over successful IPS-FAS runs using combination 3.1&3.4, for $g_2 = 0$, $\mu = 0.8s$.

k	No. of particles at start of Step I	No. of different particles at start of Step I	No. of different particles after Step I	No. of survived particles after Step III	No. of different particles after Step III	% of successful IPS runs through level k
1	1000	1	3.15	997.85	1	100%
2	1000	1	909.74	91.26	1	100%
3	1000	1	917.58	83.42	1	100%
4	1000	1	908.38	92.62	1	100%
5	1000	1	999.41	99.01	98.42	100%

For the GSHS example $g_2 = 0,\ \mu = 0.8s$, the differences in particle diversity in Tables 3.5-3.7 correspond with the theory-based expectations in subsections 3.4.3 and 3.4.4.

For the GSHS example $g_2 = 1,\ \mu = 0.8s$, in addition to random delays, Brownian motion creates small differences in the position component of particles, as a result of which almost all particles will differ from each other. As shown in Table 3.4, in spite of this Brownian motion effect, algorithm combination 3.1&3.2 falls short in capturing proper effect on particle diversity and reach probability by the spontaneous jumps in the original GSHS.

3.6. Conclusion

In many application domains, processes have a hybrid state space and their evolution involves diffusion as well as forced and spontaneous jumps. This explains why GSHS and its subclasses play a key role in formal modelling and analysis. However in simulation and control of such systems, common practice is to use an SHS model, i.e. a hybrid system that involves diffusion and forced jumps, though no spontaneous jumps. Hence a relevant question is: "Can a GSHS model be transformed to an SHS model without changing process behavior that is relevant for the application considered?" This chapter has addressed this question in using the Interacting Particle System (IPS) framework of Cérou et al. (2006) for numerically estimating the reach probability γ of an unsafe set *D* in a GSHS model.

In section 3.2 stochastic process executions of GSHS have been defined, as well as their relation to solutions of SDE's on a hybrid space. Also explained is that the transformation of GSHS to an SHS by Lygeros and Prandini (2010) has as side-effect that it produces "remaining local time" information that should be treated as being not observable for other process(es) than the GSHS execution considered.

Section 3.3 explains the IPS setting for a GSHS, by adopting a nested sequence of increasing subsets of *D*, and an implied factorization of the reach probability γ. Because a GSHS may jump over a subset boundary it is shown that this does not hinder the factorization (Proposition 3.1).

Section 3.4 develops IPS algorithms for application to GSHS. First, subsection 3.4.1 specifies the IPS algorithm cycles for a GSHS using Fixed Assignment Splitting (FAS). Theorem 3.4.1 proves that this yields lower or equal variance than using other IPS with splitting options. Subsections 3.4.2 addresses IPS evaluation of a GSHS by using an SHS version, that follows from the Lygeros and Prandini (2010) transformation. The side-effect is that each IPS cycle makes use of the "remaining local time" information that is non-existing in the original GSHS. Subsection 3.4.3 mitigates this side-effect, by an enrichment of the original GSHS, prior to applying the transformation of Lygeros and Prandini (2010) with the first hitting times of the IPS subsets. Thanks to this enrichment, the resulting SHS refreshes "remaining local time" at the start of each next IPS cycle. The latter refreshment induces a significant improvement in particle diversity at the start of each IPS cycle. As a result of this improved particle diversity IPS performance in reach probability estimation is expected to significantly improve when reach probability estimation becomes a challenge. For purpose of comparison, in subsection 3.4.4 an algorithm for the direct simulation of a GSHS execution within IPS cycles is specified. Based on theory, use of this algorithm in IPS for GSHS will yield similar good performance as the algorithm of subsection 3.4.3. In section 3.5, the expected differences in IPS performance have been illustrated for a GSHS example.

The findings in section 3.4 mean that for IPS based reach probability estimation for an arbitrary GSHS model, can be applied to a properly derived SHS version of the GSHS model. The proper way in deriving such SHS consists of three steps. The first step is to specify a GSHS model of the practical system. The second step is to enrich this GSHS with the first hitting times of the IPS subsets, without affecting the pathwise behavior of the GSHS execution. The third step is to apply the transformation by Lygeros and Prandini (2010) to the enriched GSHS from step 2.

In view of this positive finding for the limited scope of IPS application to GSHS, a logical follow-on question is if there also exists an improved transformation of a GSHS to SHS for stochastic control problems. Such transformation would make optimal control policies developed for SHS applicable to GSHS.

ACKNOWLEDGEMENT

The authors would like to thank anonymous reviewers for helpful suggestions in improving the chapter, and Bert Bakker (NLR, Amterdam) for suggesting the comparison of GSHS and SHS in applying IPS.

3.7 References

Abate, A., Prandini, M., Lygeros, J., Sastry, S. (2009). Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems, Automatica, Vol. 44, pp. 2724-2734.

Abate, A, D'Innocenzo, A. and Di Benedetto, M. (2011). Approximate Abstractions of Stochastic Hybrid systems, IEEE Tr. Automatic Control, vol. 56, pp. 2688-2694.

Alur, R., Henzinger, T., Laffierriere, G. and Pappas, G. (2000). Discrete abstractions of hybrid systems, Proc. of the IEEE, Vol. 88, pp. 971-984.

Bensoussan, A. and J.L. Menaldi (2000), Stochastic hybrid control, J. of Mathematical Analysis and Applications, Vol. 249, pp. 261-288.

Blom, H.A.P. (2003). Stochastic hybrid processes with hybrid jumps. In Proc. IFAC Conf. Analysis and Design of Hybrid Systems, Saint-Malo, Brittany, France, 16-18 June 2003.

Blom, H.A.P. (2022). Feller property of regime-switching jump diffusion processes with hybrid jumps. Forthcoming, Stochastic Analysis and Applications.

Blom, H.A.P., Krystul, J. and Bakker, G.J. (2006). A particle system for safety verification of free flight in air traffic, Proc. IEEE Conf. Decision and Control, San Diego, CA, pp. 1574-1579.

Blom, H.A.P., Krystul, J., and Bakker, G.J., Klompstra, M.B., Klein Obbink, B. (2007a). Free flight collision risk estimation by sequential MC simulation. Eds: C.G, Cassandras and J. Lygeros, Stochastic hybrid systems, Taylor & Francis, pp.249-281.

Blom, H.A.P., Bakker, G.J., and Krystul, J. (2007b). Probabilistic reachability analysis for large scale stochastic hybrid systems, Proc. IEEE Conf. on Decision and Control, pp. 3182-3189.

Blom, H.A.P., Bakker, G.J., Krystul, J. (2009). Rare event estimation for a large scale stochastic hybrid system with air traffic application, Eds: G. Rubino and B. Tuffin, Rare event simulation using Monte Carlo methods, Wiley, pp. 193-214.

Blom, H.A.P, Ma, H. and Bakker, G.J. (2018). Interacting Particle System-based Estimation of Reach Probability for a Generalized Stochastic Hybrid System. IFAC-PapersOnLine, 51(16), pp.79-84.

Botev, Z.I. and Kroese, D.P. (2008). An efficient algorithm for rare-event probability estimation, combinatorial optimization, and counting. Methodology and Computing in Applied Probability 10(4), pp. 471-505.

Bujorianu, L.M. (2012). Stochastic Reachability Analysis of Hybrid Systems, Springer, London.

Bujorianu, M.L. and Lygeros, J. (2006). Toward a general theory of stochastic hybrid systems. Eds: H.A.P. Blom and J. Lygeros, Stochastic hybrid systems, Springer, Berlin, pp. 3-30.

Bucklew, J.A. (2004). Introduction to rare event simulation, Springer, New York.

Cérou, F., Legland, F., Del Moral, P., and Lezaud, P. (2005). Limit theorems for the multilevel splitting algorithm in the simulation of rare events. Proc. Winter Simulation Conference. Orlando, FL, pp. 682-691.

Cérou, F., Del Moral, P., Legland, F., and Lezaud, P. (2006). Genetic genealogical models in rare event analysis. Latin American J. of Probability & Mathematical Statistics, 1, pp. 181-203.

Del Moral, P. (2004). Feynman-Kac formulae: genealogical and interacting particle systems with applications, Springer.

Di Benedetto, M.D., Di Gennaro, S. and D'Innocenzo, A. (2015). Hybrid systems and verification by abstraction, Eds: M. Djemai and M. Defoort, Hybrid Dynamical Systems, Springer, pp.1-25.

Ghosh, M.K., A. Arapostathis and S.I. Marcus (1993). Optimal control of switching diffusions with application to flexible manufacturing systems, SIAM J. Control and Optimization, Vol. 31 , pp. 1183-1204.

Ghosh, M.K. and A. Bagchi (2004). Modeling stochastic hybrid systems. In J. Cagnol and J.P. Zolèsio, editors, System Modeling and Optimization. Proc. 21st IFIP TC7 Conference, Sophia Antipolis, France, July 2004, pp. 269-279.

Glasserman, P. (2004). Monte Carlo Methods in Financial Engineering. Springer Science & Business Media.

Glasserman, P., Heidelberger, P., Shahabuddin, P. and Zajic, T. (1999). Multilevel splitting for estimating rare event probabilities, Operations Research, Vol. 47, pp. 585-600.

Hespanha, J.P. (2005). A model for stochastic hybrid systems with application to communication networks, Nonlinear Analysis, Vol. 62 (2005), pp. 1353-1383.

Hu, J, J. Lygeros, S.S. Sastry (2000), Towards a theory of stochastic hybrid systems. Eds: N. Lynch and B.H. Krogh, Proc. Hybrid Systems in Computation and Control 2000, LNCS number 1790, Springer, 2000, pp. 160-173

Hutzenthaler, M, A. Jentzen and P. E. Kloeden (2011), Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients, Proc. R. Soc. A., Vol. 467, pp. 1563–1576.

Julius, A.A. and Pappas, G.J. (2009). Approximations of stochastic hybrid systems, IEEE Trans. Automatic Control, Vol. 54, pp. 1193-1203.

Koutsoukos, X.D. (2004). Optimal control of stochastic hybrid systems based on locally consistent Markov decision processes. Int. J. Hybrid Systems, Vol. (4), pp. 301-318.

Krystul, J. (2006). Modelling of Stochastic Hybrid Systems with applications to accident risk assessment, PhD Thesis, University of Twente, The Netherlands, 2006.

Krystul, J., H.A.P. Blom, and A. Bagchi (2007). Stochastic differential equations on hybrid state spaces. In C. G. Cassandras and J. Lygeros, editors, Stochastic Hybrid Systems: Recent Developments and Research Trends. Taylor & Francis/CRC Press, 2007, pp. 15-45.

Krystul, J., A. Bagchi, and H.A.P. Blom (2011), On strong Markov property of solutions to stochastic differential equations on hybrid state spaces, Report 28th February 2011, U. of Twente, https://repository.tudelft.nl/islandora/object/uuid%3A948fc4dd-66f5-42bc-82e6-8770d371add3

Krystul, J., Le Gland, F., Lezaud, P. (2012). Sampling per mode for rare event simulation in switching diffusions, Stochastic Processes and their Applications, Vol. 122, pp. 2939-2667.

Kunwai, K. and C. Zhu (2020). On Feller and strong Feller properties and irreducibility of regime-switching jump diffusion processes with countable regimes, Nonlinear Analysis: Hybrid Systems, Vol. 38, article 100946, 21 pages.

Lavaei, A, Soudjani, S, Abate, A. and Zamani, M. (2021) Automated verication and synthesis of stochastic hybrid systems: A survey. arXiv preprint arXiv:2101.07491. pp. 1-59

L'Ecuyer, P., LeGland, F., Lezaud, P. and Tuffin, B. (2009). Splitting techniques, Eds: G. Rubino, B. Tuffin, Rare event simulation using Monte Carlo methods, Wiley, pp. 39-61.

Lygeros, J. and Prandini, M.. (2010). Stochastic hybrid systems: a powerful framework for complex, large scale applications. European Journal of Control, 16(6), pp.583-594.

Ma, H. and Blom, H.A.P., 2022. Random Assignment Versus Fixed Assignment in Multilevel Importance Splitting for Estimating Stochastic Reach Probabilities. Methodol Comput Appl Probab 24, 2313–2338. https://doi.org/10.1007/s11009-021-09892-4

Mao, X. and C. Yuan (2006). Stochastic Differential Equations with Markovian Switching, Imperial College Press, 2006.

Morio, J. and Balesdent, M. (eds.) (2016). Estimation of Rare Event Probabilities in Complex Aerospace and Other Systems, Woodhead Publishing.

Prandini, M. and Hu, J. (2007). Stochastic Reachability: Theory and numerical approximation. Eds: C.G. Cassandras and J. Lygeros, Stochastic hybrid systems, Taylor & Francis, pp. 107-137.

Prandini, M., Blom, H.A.P. and Bakker, G.J. (2011). Air traffic complexity and the interacting particle system method: An integrated approach for collision risk estimation, Proc. American Control Conf., San Francisco, CA, pp. 2154-2159.

Rubinstein, R. (2010). Randomized algorithms with splitting: Why the classic randomization algorithms do not work and how to make them work, Methodology and Computing in Applied Probability, Vol. 12, pp. 1-50.

Teel, A.R., A. Subbaraman and A. Sferlazza (2014). Stability analysis for stochastic hybrid systems: A survey. Automatica, Vol. 50, pp. 2435-2456.

Xi, F., G. Yin, and C. Zhu (2019), Regime-switching jump diffusions with non-Lipschit coefficients and countably many switching states: existence and uniqueness, Feller, and strong Feller properties. In: G. Yin and Q. Zhang (eds), Modeling, stochastic control, optimization and applications, Springer, pp. 571-599.

Yin, G.G. and C. Zhu (2010). Hybrid switching diffusions; properties and applications, Springer, 2010.

Appendix 3.A: Proof of Theorem 3.4.1

In this appendix we compare the variance of applying IPS to GSHS under FAS versus multinomial resampling (MR), multinomial splitting (MS), and residual multinomial splitting (RMS). In doing so it becomes clear that the earlier comparison by Ma and Blom (2022) for diffusion process extends to GSHS executions.

The first proof starts with a characterization of the conditional distribution of particles that reach level k+1, given that at level k the i-th successful particle $\tilde{\xi}_k^i$ is copied K_k^i times, $i=1,..,N_{S_k}$.

Proposition 3.A.1: If $N_{S_k}>0$ and K_k^i, with i=1,2,...,N_{S_k}, denote the number of particles that copies $\tilde{\xi}_k^i$ at level k. Then the number $Y_{k+1}^{k,i}$, of the K_k^i particle copies of $\tilde{\xi}_k^i$ that reach level k+1, has a conditional Binomial distribution of size K_k^i and success probability $\gamma_{k+1}(\tilde{\xi}_k^i)$, i.e.

$$p_{Y_{k+1}^{k,i}|K_k^i,\tilde{\xi}_k^i}(n;K_k^i,\tilde{\xi}_k^i)=\mathrm{Bin}(n;K_k^i,\gamma_{k+1}(\tilde{\xi}_k^i)) \tag{3.26}$$

with

$$\gamma_{k+1}(\tilde{\xi}_k^i)\triangleq \mathrm{P}(\tau_{k+1}<T\,|\,\xi_k=\tilde{\xi}_k^i) \tag{3.27}$$

Proof: Similar to the proof of Proposition 1 in (Ma and Blom, 2022).

Theorem 3.A.1: If $N_{S_k}\geq 1$ and K_k^i, $i=1,..N_{S_k}$, denotes the number of copies made of the i-th successful particle $\tilde{\xi}_k^i$ during the splitting step at level k of the IPS algorithm, then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1}\,|\,\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}=\frac{1}{N_p}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i\,|\,\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}\gamma_{k+1}(\tilde{\xi}_k^i)\right] \tag{3.28}$$

$$\begin{aligned}\mathrm{Var}\left\{\overline{\gamma}_{k+1}\,|\,\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}=&\frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathbb{E}\left\{K_k^i\,|\,\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right]\\&+\frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\left[\mathrm{Var}\left\{K_k^i\,|\,\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}\gamma_{k+1}(\tilde{\xi}_k^i)^2\right]\\&+\frac{1}{N_p^{\,2}}\sum_{i=1}^{N_{S_k}}\sum_{\substack{i'=1\\ i'\neq i}}^{N_{S_k}}\left[\mathrm{Cov}\left\{K_k^i,K_k^{i'}\,|\,\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}\gamma_{k+1}(\tilde{\xi}_k^i)\,\gamma_{k+1}(\tilde{\xi}_k^{i'})\right]\end{aligned} \tag{3.29}$$

Proof: Similar to the proof of Theorem 1 in (Ma and Blom, 2022).

Proposition 3.A2: If $N_{S_k}\geq 1$, and we use multinomial resampling at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1}\,|\,\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}=\frac{1}{N_{S_k}}\sum_{i=1}^{N_{S_k}}\gamma_{k+1}(\tilde{\xi}_k^i) \tag{3.30}$$

$$\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] + \frac{1}{N_p N_{S_k}} \left[\sum_{i=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right]\right] \tag{3.31}$$

Proof: Similar to the proof of Proposition 2 in (Ma and Blom, 2022).

Proposition 3.A3: If $N_{S_k} \geq 1$, and we use multinomial splitting at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \gamma_{k+1}(\tilde{\xi}_k^i) \tag{3.32}$$

$$\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] + \frac{\left(N_p - N_{S_k}\right)}{N_p^2 N_{S_k}} \left[\sum_{i=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right]\right] \tag{3.33}$$

Proof: Similar to the proof of Proposition 3 in (Ma and Blom, 2022).

Proposition 3.A4: If $N_{S_k} \geq 1$, and we use residual multinomial splitting at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \gamma_{k+1}(\tilde{\xi}_k^i) \tag{3.34}$$

$$\begin{aligned} &\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} \\ &= \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] \\ &+ \frac{\left(N_p \bmod N_{S_k}\right)}{N_p^2 N_{S_k}} . \left[\sum_{i=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right)\right]\right] \end{aligned} \tag{3.35}$$

Proof: Similar to the proof of Proposition 4 in (Ma and Blom, 2022).

Proposition 3.A5: If $N_{S_k} \geq 2$, and we use fixed assignment splitting at IPS level k then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \gamma_{k+1}(\tilde{\xi}_k^i) \tag{3.36}$$

$$\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\} = \frac{1}{N_p N_{S_k}} \sum_{i=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\left(1-\gamma_{k+1}(\tilde{\xi}_k^i)\right)\right] + \frac{\left(N_p \bmod N_{S_k}\right)\left[N_{S_k} - \left(N_p \bmod N_{S_k}\right)\right]}{N_p^2 N_{S_k}\left(N_{S_k}-1\right)} . \left[\sum_{i=1}^{N_{S_k}} \left[\left(\gamma_{k+1}(\tilde{\xi}_k^i)\right)^2\right] - \frac{1}{N_{S_k}} \sum_{i=1}^{N_{S_k}} \sum_{i'=1}^{N_{S_k}} \left[\gamma_{k+1}(\tilde{\xi}_k^i)\gamma_{k+1}(\tilde{\xi}_k^{i'})\right]\right] \tag{3.37}$$

Proof: Similar to the proof of Proposition 5 in (Ma and Blom, 2022).

Theorem 3.A2: Given successful particles $\tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}$ at IPS level k with $N_{S_k} \geq 1$. The dominance of the four splitting methods (MR, MS, RMS, FAS) in terms of $\text{Var}\left\{\overline{\gamma}_{k+1} \mid \tilde{\xi}_k^1,...,\tilde{\xi}_k^{N_{S_k}}\right\}$ is:

$$\text{V}_{FAS}^k \leq \text{V}_{RMS}^k \leq \text{V}_{MS}^k \leq \text{V}_{MR}^k \tag{3.38}$$

Proof: Similar to the proof of Theorem 2 in (Ma and Blom, 2022).

Theorem 3.A3: If IPS levels 1 to *k*-1 make use of the same type of splitting (either MR, MS, RMS or FAS), then the dominance of the four splitting methods at level *k*, in terms of $\text{Var}\left\{\prod_{k'=1}^{k} \overline{\gamma}_{k'}\right\}$ satisfies:

$$\text{V}_{FAS\_k} \leq \text{V}_{RMS\_k} \leq \text{V}_{MS\_k} \leq \text{V}_{MR\_k} \tag{3.39}$$

Proof: Similar to the proof of Theorem 3 in (Ma and Blom, 2022).

Theorem 3.A4: Under the same type of Splitting (either MR, MS, RMS or FAS) at all levels, then the dominance of the four splitting methods in terms of variance $\text{V=Var}\{\overline{\gamma}\}$ satisfies:

$$\text{V}_{FAS} \leq \text{V}_{RMS} \leq \text{V}_{MS} \leq \text{V}_{MR} \tag{3.40}$$

Proof: Similar to the proof of Theorem 4 in (Ma and Blom, 2022).

4

Sampling per mode strategies in rare event simulation of stochastic hybrid systems

This chapter studies sampling per mode strategies in a multi-level splitting approach to estimating reach probability for stochastic hybrid systems. In the literature, the theoretical framework of multi-level splitting based rare event simulation for diffusions has been well extended to stochastic hybrid systems. A critical issue is potential particle depletion for safety-critical modes of a hybrid stochastic system; then a multi-level splitting approach may run out of relevant particles prior to reaching the unsafe set. To improve this situation, Krystul et al. (2012) incorporate sampling per mode in the theoretical framework of multi-level splitting for switching diffusions. The objective of the current chapter is twofold. Firstly, to develop more efficient sampling per mode strategies. Secondly, to characterize and compare these sampling per mode strategies in terms of mean and variance of estimated reach probability. The novel results are also illustrated through rare event simulations for a simple rare event simulation example.

This chapter has been submitted to Statistics and Computing, as H. Ma and H.A.P. Blom, Sampling per mode strategies in rare event simulation of stochastic hybrid systems.

4.1. Introduction

In a multi-level splitting approach to rare event simulation, the set for which a reach probability has to be estimated, is enclosed by a series of strictly increasingly (nested/enclosing) subsets. This allows one to express the small reach probability of the inner level set as a product of larger reach probabilities for the sequence of enclosing subsets (Glasserman et al, 1999; Au and Beck, 2003; Botev and Kroese, 2008; Rubinstein, 2010). Embedding of this multi-level setting in the Feynman-Kac framework (Del Moral, 2004) has enabled a systematic evaluation of reach probability through sequential Monte Carlo simulation of an interacting particle system (IPS), including characterization of asymptotic behaviour (Cérou et al., 2006). Cérou et al. (2012, 2019) provide overviews of rare event simulation developments.

A well known issue of IPS is the possibility of particle depletion prior to reaching the unsafe set. To mitigate such particle depletion, LeGland and Oudjane (2006) develop an IPS version that keeps the particle system alive, at the possible cost of having to make a too large number of particle copies. The issue of particle depletion typically plays an even larger role for safety-critical modes in stochastic hybrid systems. A possible reason is that there may be few or no particles in modes (discrete-valued state components) with small probabilities (i.e., "light" modes). This happens because each splitting step tends to sample more "heavy" particles from modes with higher probabilities, thus, "light" particles in the "light" modes tend to be discarded. To address this problem, for switching diffusions, Krystul et al. (2012) extend IPS with a sampling per mode strategy, including the embedding in the Feynman-Kac framework.

The objective of the current chapter is twofold: i) to develop more efficient sampling per mode strategies; and ii) to characterize and compare these strategies in terms of mean and variance of estimated reach probability.

For the realization of the first objective, use is made of the sampling strategy background from literature on rare event simulation (Garvels and Kroese, 1998; Cérou et al., 2006; L'Ecuyer et al., 2007; L'Ecuyer et al., 2009) and on particle filtering (Del Moral et al., 2001; Gerber et al., 2019). This yields four main splitting strategies for use in IPS: i) Multinomial Resampling (MR); ii) Multinomial Splitting (MS); iii) Remainder Multinomial Splitting (RMS); and iv) Fixed Assignment Splitting (FAS). MR is the classical method of drawing N_P random samples, with replacement, from the set of N_S successful particles. MS adds to the set of N_S successful particles, $N_P - N_S$ random samples, with replacement from the set of N_S successful particles. RMS makes of each succesfull particle $\lfloor N_P / N_S \rfloor$ copies, and subsequently adds $N_P - \lfloor N_P / N_S \rfloor$ random samples, with replacement, from the set of successful particles. FAS also makes of each succesfull particle $\lfloor N_P / N_S \rfloor$ copies, though subsequently adds $N_P - \lfloor N_P / N_S \rfloor$ random samples, without replacement, from the set of successful particles. In (Ma and Blom, 2023), it is shown that for rare event estimation, in multi-dimensional diffusions, in terms of variance of the estimated reach probability, FAS tends to score best, RMS second, MS third, and MR last. Subsequently, (Ma and Blom, 2023) have extended this result to general stochastic hybrid systems (GSHS).

The sampling per mode strategy developed by Krystul et al. (2012) is of Multinomial Resampling (MR) type. In contrast to classical IPS, in IPS using sampling per mode the weights of the copied particles have to be taken into account in such a way that the reach probability estimator is unbiased. The current chapter develops novel IPS sampling per mode versions of MR MS, RMS and FAS and corresponding weighting mechanisms such that this unbiasedness condition is satisfied. In realizing the second objective, i.e. characterizing mean and variance of reach probability estimates, the corresponding weighting mechanisms will explicitly be taken into account.

The remainder of this chapter is organized as follows. Section 4.2 summarizes IPS setting for a General Stochastic Hybrid System (GSHS), and presents the algorithmic steps in the sampling per mode IPS algorithm of Krystul et al. (2012). Section 4.3 develops a slightly improved version of this algorithm, and compares mean and variance of estimated reach probability with those from normal IPS. Section 4 specifies MS, RMS and FAS versions of sampling per mode IPS and characterizes mean and variance of estimated reach probabilities. Section 4.5 compares means and variances, and proves performance dominance relations under specific conditions. Section 4.6 compares simulation results for the various sampling per mode versions for a simple GSHS example. Section 4.7 draws conclusions.

4.2. IPS based reach probability estimation

4.2.1. Reach probability of GSHS

Throughout this and the following sections, all stochastic processes are defined on a complete stochastic basis $(\Omega,\mathcal{F},\mathbb{F},\mathrm{P},\mathcal{T})$ with $(\Omega,\mathcal{F},\mathrm{P})$ being a complete probability space and $\mathbb{F}$ an increasing sequence of sub-σ-algebras on the time line $\mathcal{T}=\mathbb{R}_+$,i.e.,$\mathbb{F}\triangleq\{\mathcal{J},(\mathcal{F}_t,t\in\mathbb{R}_+),\mathcal{F}\}$, with $\mathcal{J}$ containing all P-null sets of $\mathcal{F}$ and $\mathcal{J}\subset\mathcal{F}_s\subset\mathcal{F}_t\subset\mathcal{F}$ for every $s<t$.

Following (Bujorianu and Lygeros, 2006), we consider the execution process $\{\theta_t,x_t\}$ of a General Stochastic Hybrid System (GSHS). The latter is a 8-tuple $(\Theta,d,X,f,g,Init,\lambda,R)$ where Θ is a countable set of discrete-valued variables, $d:\Theta\to\mathbb{N}$ is a map giving the dimensions of the continuous state spaces, $X:\Theta\to\mathbb{R}^{d(\cdot)}$ maps each $\theta\in\Theta$ into an open subset X^θ of $\mathbb{R}^{d(\theta)}$, $f:\Xi\to\mathbb{R}^{d(\cdot)}$ is a vector field, where $\Xi\triangleq\bigcup_{\theta\in\Theta}\{\theta\}\times X^\theta$, $g:\Xi\to\mathbb{R}^{d(\cdot)\times m_{\dim}}$ is an $X^{(\cdot)}$-valued matrix, $m_{\dim}\in\mathbb{N}$, $\mathrm{Init}:\beta(\Xi)\to[0,1]$ an initial probability measure on Ξ, $\lambda:\bar{\Xi}\to\mathbb{R}^+$ is a transition rate function, $R:\bar{\Xi}\times\beta(\bar{\Xi})\to[0,1]$ is a transition measure.

The problem is to estimate the probability γ that $\{\theta_t,x_t\}$ reaches a closed subset $D\subset\Xi$ within finite period $[0,T]$, i.e.

$$\gamma=P(\tau<T) \tag{4.1}$$

with τ being the first hitting time of D by $\{\theta_t,x_t\}$:

$$\tau=\inf\{t>0,(\theta_t,x_t)\in D\} \tag{4.2}$$

4.2.2. Multi-level factorization of reach probability

The principle in factorizing the reach probability $\gamma=P(\tau<T)$ is to introduce a sequence D_k, $k=0,..,m,$ of nested closed subsets of Ξ, i.e. $D=D_m\subset D_{m-1}\subset\cdots\subset D_1\subset D_0=\Xi$, with D_1 such that $P\{(\theta_0,x_0)\in D_1\}=0$. Let τ_k be the first moment in time that $\{\theta_t,x_t\}$ reaches D_k, i.e.

$$\tau_k=\inf\{t>0;\ (\theta_t,x_t)\in D_k\vee t\geq T\} \tag{4.3}$$

Next, we define $\{0,1\}$-valued random variables $\{\chi_k,k=0,..,m\}$ as follows:

$$\chi_k = 1, \text{ if } \tau_k < T \text{ or } k = 0 \\ = 0, \text{ else} \tag{4.4}$$

By using this χ_k definition, reach probability γ satisfies the factorization, e.g. (Ma and Blom, 2021, Proposition 3.1):

$$\gamma = \prod_{k=1}^{m} \gamma_k \tag{4.5}$$

where $\gamma_k \triangleq E\{\chi_k = 1 | \chi_{k-1} = 1\} = \mathrm{P}(\tau_k < T | \tau_{k-1} < T)$.

4.2.3. Recursive estimation of the multi-level factors

First we define $\Xi' \triangleq \mathbb{R} \times \Xi$, $\xi_k \triangleq (\tau_k, \theta_{\tau_k}, x_{\tau_k})$, $Q_k \triangleq (0,T) \times D_k$, for $k = 1,..,m$, and the following conditional probability measure $\pi_k(B)$ for an arbitrary Borel set B of Ξ':

$$\pi_k(B) \triangleq P(\xi_k \in B \,|\, \xi_k \in Q_k)$$

Following Cérou et al. (2006) a recursive scheme in evolving π_k and estimating the fractions γ_k involves the following transformations:

$$\pi_{k-1}(\cdot) \xrightarrow{\text{I. mutation}} p_k(\cdot) \xrightarrow{\text{III. selection}} \pi_k(\cdot)$$
$$\downarrow \text{II. conditioning}$$
$$\gamma_k$$

where $p_k(B)$ is the conditional probability measure of $\xi_k \in B$ given $\xi_{k-1} \in Q_{k-1}$, i.e.,

$$p_k(B) \triangleq P(\xi_k \in B \,|\, \xi_{k-1} \in Q_{k-1})$$

Because $\{\theta_t, x_t\}$ is a strong Markov process, $\{\xi_k\}$ is a Markov sequence. Hence, the mutation transformation (I) satisfies a Chapman-Kolmogorov equation prediction for ξ_k:

$$p_k(B) = \int_{\Xi'} p_{\xi_k|\xi_{k-1}}(B \,|\, \xi) \pi_{k-1}(d\xi) \text{ for all } B \in \beta(\Xi') \tag{4.6}$$

For the conditioning transformation (II) this means:

$$\gamma_k = \mathrm{P}(\tau_k < T \,|\, \tau_{k-1} < T) = \int_{\Xi'} 1_{\{\xi \in Q_k\}} p_k(d\xi). \tag{4.7}$$

Hence, selection transformation (III) satisfies:

$$\pi_k(B) = \frac{\int_B 1_{\{\xi \in Q_k\}} p_k(d\xi)}{\int_{\Xi'} 1_{\{\xi' \in Q_k\}} p_k(d\xi')} = [\int_B 1_{\{\xi \in Q_k\}} p_k(d\xi)] / \gamma_k \tag{4.8}$$

With this, the γ_k terms in (4.5) are characterized as solutions of a recursive sequence of mutation equation (4.6), conditioning equation (4.7) and selection equation (4.8).

If a set of N_P particles is used to form empirical density approximations $\bar{\gamma}_k$, $\bar{p}_k$ and $\bar{\pi}_k$ of γ_k, p_k and π_k respectively, then equations (4.5)-(4.8) yield the following recursion of the IPS algorithmic steps for the numerical estimator $\bar{\gamma} = \prod_{k=1}^{m} \bar{\gamma}_k$ of γ:

$$\bar{\pi}_{k-1}(\cdot) \xrightarrow{\text{I. mutation}} \bar{p}_k(\cdot) \xrightarrow{\text{III. selection}} \tilde{\pi}_k(\cdot) \xrightarrow{\text{IV. splitting}} \bar{\pi}_k(\cdot)$$
$$\bar{p}_k(\cdot) \xrightarrow{\text{II. conditioning}} \bar{\gamma}_k$$

Cérou et al. (2006) prove that using IPS with multinomial splitting (MS), $\bar{\gamma}$ forms an unbiased γ estimator, i.e.

$$\mathbb{E}\{\bar{\gamma}\} = \mathbb{E}\left\{\prod_{k=1}^{m} \bar{\gamma}_k\right\} = \prod_{k=1}^{m} \mathbb{E}\{\bar{\gamma}_k\} = \prod_{k=1}^{m} \gamma_k = \gamma \tag{4.9}$$

Moreover, Cérou et al. (2006) derive second and higher order asymptotic bounds for the error $(\bar{\gamma} - \gamma)$ based on the multi-level Feynman Kac analysis, e.g. Del Moral (2004; Theorem 12.2.2).

4.2.4. Multinomial resampling per Mode

To cope with large differences in mode weights, (Krystul et al. 2012) propose an IPS that applies sampling per mode. The resulting algorithm is specified as Algorithm 4.1 below, and referred to as the IPS_{mode} algorithm.

Algorithm 4.1. IPS_{mode} (Krystul et al., 2012)

Input: Initial measure π_0, end time T, decreasing sequence of closed subsets $D_k = \{(\theta_t, x_t) \in \Xi\}$, $D_{k-1} \supset D_k$, $k = 1,..,m$. Also $D_0 = \Xi$, $Q_k = (0,T) \times D_k$ and number of particles N_P.

Output: Estimated reach probability $\bar{\gamma}$

0. Initiation: Generate $N_0^\theta = N_P / M$ particles for each $\eta \in \Theta$: $\xi_0^{\eta,j} = (\tau_0^{\eta,j}, x_0^{\eta,j}, \theta_0^{\eta,j}) = (0, x_0^{\eta,j}, \eta)$ for $j = 1,.., \frac{N_P}{M}$, with $x_0^{\eta,j} \sim \pi_0(0,\cdot,\eta) \Big/ \int \pi_0(0,x',\eta)dx'$ and $\omega_0^{\eta,j} = \frac{\int \pi_0(0,x',\eta)dx'}{N_p/M}$, i.e. $\bar{\pi}_0(0,\cdot,\eta) \approx \sum_{j=1}^{N_p/M} [\omega_0^{\eta,j} \delta_{\{\xi_0^{\eta,j}\}}(0,\cdot,\eta)]$ with Dirac δ. Set $k = 1$.

I. Mutation: $\bar{p}_k(.) = \sum_{\eta \in \Theta} \sum_{j=1}^{N_{k-1}^\theta} \omega_{k-1}^{\eta,j} \delta_{\{\bar{\xi}_k^{\eta,j}\}}(\cdot)$, where $\bar{\xi}_k^{\eta,j}$ is obtained by simulating GSHP execution starting at $\xi_{k-1}^{\eta,j}$.

II. Conditioning: $\bar{\gamma}_k = \sum_{\eta \in \Theta} \sum_{j=1}^{N_{k-1}^\theta} \left[\omega_{k-1}^{\eta,j} 1\left(\bar{\xi}_k^{\eta,j} \in Q_k\right)\right]$.

If $\bar{\gamma}_k = 0$ then $\bar{\gamma}_{k'} = 0$, $k' \in \{k,..,m\}$ and go to Step V.

III. Selection: $J_k^\theta = \{(\eta, j) \in \Theta \times [1, N_{k-1}^\theta]; \bar{\theta}_k^{\eta,j} = \theta, \bar{\xi}_k^{\eta,j} \in Q_k\}$.

$\tilde{\pi}_k(.) = \sum_{\theta \in \Theta} \sum_{(\eta,j) \in J_k^\theta} \varpi_k^{\eta,j} \delta_{\{\bar{\xi}_k^{\eta,j}\}}(\cdot)$ and $\tilde{\pi}_{\theta_k}(\theta) = \sum_{(\eta,j) \in J_k^\theta} \varpi_k^{\eta,j}$

with $\bar{\omega}_k^{\eta,j} = \omega_{k-1}^{\eta,j} 1\left(\bar{\xi}_k^{\eta,j} \in Q_k\right) / \bar{\gamma}_k$.

For each $\theta \in \Theta$, collect $\left\{\bar{\omega}_k^{\eta,j}, \bar{\xi}_k^{\eta,j}\right\}_{(\eta,j)\in J_k^\theta}$ in $\left\{\tilde{\omega}_k^{\theta,i}, \tilde{\xi}_k^{\theta,i}\right\}_{i=1}^{|J_k^\theta|}$.

IV. Splitting: For each mode $\theta \in \underline{\Theta} \triangleq \{\theta \in \Theta; \left|J_k^\theta\right| > 0\}$, draw N_p / M samples $\xi_k^{\theta,j} \sim \frac{\tilde{\pi}_k(\cdot,\cdot,\theta)}{\tilde{\pi}_{\theta_k}(\theta)}$. Set $N_k^\theta = N_p / M$ and $\bar{\pi}_k(.) = \sum_{\theta\in\Theta} \sum_{j=1}^{N_p/M} \omega_k^{\theta,j} \delta_{\{\xi_k^{\theta,j}\}}(\cdot)$ with $\omega_k^{\theta,j} = \frac{\tilde{\pi}_{\theta_k}(\theta)}{N_p/M}$.

V. If $k < m$, then $k := k+1$ and go to step I, else $\bar{\gamma} = \prod_{k=1}^{m} \bar{\gamma}_k$

In this algorithm, J_k^θ is the set of particles that successfully arrive at the *k*-th level under mode $\theta \in \Theta$. Hence at the end of splitting after reaching the *k*-th level, for each $\theta \in \Theta$, for which $\left|J_k^\theta\right| > 0$, there are $N_k^\theta = N_p / M$ multinomial samples, with $M = |\Theta|$.

Krystul et al. (2012) prove that the IPS$_{\text{mode}}$ algorithm applied to a switching diffusion yields an estimated reach probability $\bar{\gamma}$ that is unbiased, and asymptotically converges to γ.

4.3. IPS$_{\text{mode}}$-MR$_{\text{mode}}$ vs. IPS-MR

In this section, the IPS$_{\text{mode}}$ algorithm 4.1 is slightly improved, and the conditional variance of the $\bar{\gamma}$ estimator of this improved IPS$_{\text{mode}}$ version is compared to the variance of $\bar{\gamma}$ from the basic IPS algorithm.

Firstly, in subsection 4.3.1, an improved version of algorithm 4.1 is proposed. Next, in subsection 4.3.2, the conditional variance of $\bar{\gamma}_k$ at the *k*-th level of this improved IPS$_{\text{mode}}$ algorithm is evaluated. Next in subsection 4.3.3 this conditional variance is compared to conditional variance if at the *k*-th level a normal MR is used instead of sampling per mode.

4.3.1. Improvement of splitting step in algorithm 4.1

The splitting step of Krystul et al. (2012), implicitly assume that for each mode there are particles that successfully reach the next level. However if the number M_k of modes that successfully reach the *k*-th level is $< M$, then the number of samples drawn in the splitting step IV of algorithm 4.1 is $M_k \frac{N_p}{M} < N_p$. Hence a straightforward improvement of the splitting step IV is to generate N_p / M_k samples per successful mode. This improved splitting step is specified in Algorithm 4.1* below; and referred to as MR$_{\text{mode}}$ splitting. Together with the other steps of IPS$_{\text{mode}}$ algorithm 4.1, this specifies IPS$_{\text{mode}}$-MR$_{\text{mode}}$.

Algorithm 4.1*. MR$_{\text{mode}}$ splitting step IV in Algorithm 4.1

IV. MR$_{\text{mode}}$ splitting: Set $M_k := \sum_{\theta\in\underline{\Theta}} 1\{\left|J_k^\theta\right| > 0\}$

For each mode $\theta \in \underline{\Theta} \triangleq \{\theta \in \Theta; \left|J_k^\theta\right| > 0\}$, draw N_p / M_k samples $\xi_k^{\theta,j} \sim \frac{\tilde{\pi}_k(\cdot,\cdot,\theta)}{\tilde{\pi}_{\theta_k}(\theta)}$. Set $N_k^\theta = N_p / M_k$ and $\bar{\pi}_k(.) = \sum_{\theta\in\Theta} \sum_{j=1}^{N_k^\theta} \omega_k^{\theta,j} \delta_{\{\xi_k^{\theta,j}\}}(\cdot)$ with $\omega_k^{\theta,j} = \frac{\tilde{\pi}_{\theta_k}(\theta)}{N_p/\bar{M}_k}$.

4.3.2. Conditional variance under IPS$_{mode}$-MR$_{mode}$

For the IPS$_{mode}$ in algorithm 4.1, the conditional mean and variance of the factor $\overline{\gamma}_{k+1}$ for reaching at the k+1-th level, we derive the following characterization.

Theorem 4.1: If we use the IPS$_{mode}$ algorithm of Krystul et al. (2012) as specified in Algorithm 4.1. Then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.10}$$

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \frac{M}{N_p} \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \sum_{j=1}^{|J_k^\theta|} \left[\tilde{\omega}_k^{\theta,i} \tilde{\omega}_k^{\theta,j} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right] \tag{4.11}$$

where $\underline{\Theta}$ indicates $\{\theta: |J_k^\theta| > 0\}$, and $C_k^{\tilde{\xi},\tilde{\omega}}$ the sigma-algebra $C_k^{\tilde{\xi},\tilde{\omega}} \triangleq \sigma\left\{\tilde{\xi}_k^{\theta,j}, \tilde{\omega}_k^{\theta,j}; j=1,\ldots,|J_k^\theta|, \theta\in\underline{\Theta}\right\}$.

Proof: See Appendix 4.A.1.

By replacing M by M_k in Appendix 4.A.1, we get a similar result for IPSmode-MRmode.

Theorem 4.1*: If we use the MR$_{mode}$ splitting of algorithm 4.1* in step IV of Algorithm 4.1. Then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.12}$$

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \frac{M_k}{N_p} \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \sum_{j=1}^{|J_k^\theta|} \left[\tilde{\omega}_k^{\theta,i} \tilde{\omega}_k^{\theta,j} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right] \tag{4.13}$$

Comparison of eqs. (4.11) and (4.13) shows that $\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ is under IPSmodeMRmode smaller than or equal to that under IPSmode of Krystul et al. (2012).

4.3.3. Comparison of IPS$_{mode}$-MR$_{mode}$ versus normal IPS

For completeness, we first specify in Algorithm 4.2 below a normal MR splitting step to replace step IV in Algorithm 4.1. Together with steps I, II, III and V of Algorithm 4.1, this defines IPS$_{mode}$-MR. Derivation of the conditional mean and variance of the fraction $\overline{\gamma}_{k+1}$ for IPS$_{mode}$-MR yields Theorem 4.2.

Theorem 4.2: If we use a normal MR splitting step IV within an IPS$_{mode}$ cycle and assume $|J_k^\theta| > 0$ for all $\theta\in\Theta$, then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.14}$$

$$\operatorname{Var}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \frac{1}{N_p}\left[\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]\left[1-\sum_{\eta\in\underline{\Theta}}\sum_{j=1}^{|J_k^\eta|}\left(\tilde{\omega}_k^{\eta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\right)\right] \quad (4.15)$$

Proof: See Appendix 4.A.2.

Algorithm 4.2; MR splitting step IV of Algorithm 4.1

IV. MR splitting:

Draw N_P samples $\tilde{\tilde{\xi}}_k^j \sim \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\tilde{\omega}_k^{\theta,i}\delta_{\{\tilde{\xi}_k^{\theta,i}\}}(.)$, $j=1,..,N_p$, $\bar{\pi}_k(.) = \sum_j^{N_p}\tilde{\tilde{\omega}}_k^j\delta_{\{\tilde{\tilde{\xi}}_k^j\}}(\cdot)$ with $\tilde{\tilde{\omega}}_k^j = \frac{1}{N_p}$

Count $N_k^\theta = \sum_{j=1}^{N_p}1\{\tilde{\tilde{\theta}}_k^j = \theta\}$ and map the elements in $\left\{\tilde{\tilde{\omega}}_k^j, \tilde{\tilde{\xi}}_k^j\right\}_{j=1}^{N_p}$ one-on-one to $\left\{\omega_k^{\theta,i}, \xi_k^{\theta,i}; \theta\in\Theta, i=1,..,N_k^\theta\right\}$,

Next, we will prove that under specific conditions, the conditional variances of $\bar{\gamma}_k$ at the k-th level is larger under IPS-MR than it is under IPS_{mode}-MR_{mode}. To accomplish this, we start with a comparison of conditional variances of $\bar{\gamma}_k$ at the k-th level under IPS_{mode}-MR and IPS_{mode}-MR_{mode}, in Theorem 4.3.

Theorem 4.3: Given $C_k^{\tilde{\xi},\tilde{\omega}} = \sigma\left\{\tilde{\xi}_k^{\theta,j},\ \tilde{\omega}_k^{\theta,j}, \theta\in\Theta, j=1,...,|J_k^\theta|\right\}$ at IPS_{mode} level k. Let $\tilde{\pi}_{\theta_k}(\theta)$ and $\tilde{\Sigma}_k^\theta \triangleq \sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)$ satisfy for $\theta\in\underline{\Theta}$:

$$\tilde{\Sigma}_k^\theta \le \frac{\mathbb{E}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}}{M_k}, \text{ if } \tilde{\pi}_{\theta_k}(\theta) \ge \frac{1}{M_k} \quad (4.C1)$$

$$\tilde{\Sigma}_k^\theta > \frac{\mathbb{E}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}}{M_k}, \text{ if } \tilde{\pi}_{\theta_k}(\theta) < \frac{1}{M_k} \quad (4.C2)$$

Then the dominance of the IPS_{mode}-MR and IPS_{mode}-MR_{mode} methods in terms of $\operatorname{Var}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ is:

$$\operatorname{Var}_{\text{IPSmode-MRmode}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \le \operatorname{Var}_{\text{IPSmode-MR}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \quad (4.16)$$

Proof: See Appendix 4.B.

Following the kind of reasoning in Ma and Blom (2022), we first extend the results of Theorem 3 to a comparison of IPS_{mode}-MR_{mode} versus IPS-MR. This is accomplished in Theorems 4.4 and 4.5 below.

Theorem 4.4: Suppose IPS$_{\text{mode}}$ levels 1 to *k*-1 make use of the same type of splitting (either MR or MR$_{\text{mode}}$), and conditions (4.C1-4.C2) hold true. Then the dominance of the IPS$_{\text{mode}}$-MR and IPS$_{\text{mode}}$-MR$_{\text{mode}}$ methods at level *k*, in terms of $\text{V}^k = \text{Var}\left\{\prod_{k'=1}^{k} \bar{\gamma}_{k'}\right\}$ satisfies:

$$\text{V}^k_{\text{IPSmode-MRmode}} \leq \text{V}^k_{\text{IPSmode-MR}} \tag{4.17}$$

Proof: Apply the reasoning in the proof of Theorem 3 in (Ma and Blom, 2022) to the results in Theorems 4.1-4.2.

Theorem 4.5: Suppose the same type of Splitting (either MR or MR$_{\text{mode}}$) is used at all levels, and conditions (4.C1-4.C2) hold true for all levels. Then the dominance of the basic IPS-MR and IPS$_{\text{mode}}$-MR$_{\text{mode}}$ methods in terms of variance $\text{V}=\text{Var}\{\bar{\gamma}\}$ satisfies:

$$\text{V}_{\text{IPSmode-MRmode}} \leq \text{V}_{\text{IPS-MR}} \tag{4.18}$$

Proof: Application of the reasoning in the proof of Theorem 4 in (Ma and Blom, 2022) to the result in Theorem 4.4 yields

$$\text{V}_{\text{IPSmode-MRmode}} \leq \text{V}_{\text{IPSmode-MR}} \tag{4.19}$$

It is straightforward to show that application of MR splitting at all levels of IPS$_{\text{mode}}$ coincides with the basic IPS-MR algorithm. Hence inequality (4.19) implies inequality (4.18).

4.4. MS, RMS AND FAS Splitting per Mode

This section studies version of Algorithm 4.1, where step IV is replaced by particle splitting per mode, i.e. of each successful particle one or more copies are being made. Subsection 4.4.1 develops three versions of splitting per mode: MSmode, RMSmode and FASmode. Subsequently, subsection 4.4.2 characterizes conditional mean and variance for each of these three versions.

4.4.1. Splitting per mode versions for step IV

We study replacement of step IV in Algorithm 4.1 by the Algorithms 4.3, 4.4 and 4.5 for MSmode, RMSmode and FASmode respectively. The resulting IPSmode algorithms are indicated as IPS$_{\text{mode}}$-MS$_{\text{mode}}$, IPS$_{\text{mode}}$-RMS$_{\text{mode}}$ and IPS$_{\text{mode}}$-FAS$_{\text{mode}}$ respectively.

In Algorithm 4.3, for IPS$_{\text{mode}}$-MS$_{\text{mode}}$, each successful particle in $\tilde{\pi}_{\theta_k}(\theta)$ is first copied once. This yields $\bar{N}_k := \sum_{\theta\in\Theta} |J_k^\theta|$ particles. Subsequently, $N_p - \bar{N}_k$ additional particles are cloned randomly with replacement from the conditional measure $\frac{\tilde{\pi}_k(\cdot,\cdot,\theta)}{\tilde{\pi}_{\theta_k}(\theta)}$. Hence the total number of particles satsifies:

$$\sum_{\theta\in\Theta}\left((1+\rho_k^\theta)\left|J_k^\theta\right|\right) = \sum_{\theta\in\underline{\Theta}}\left[\left(1+\frac{N_p - N_k}{M_k\left|J_k^\theta\right|}\right)\left|J_k^\theta\right|\right] = \sum_{\theta\in\underline{\Theta}}\left[\left(\left|J_k^\theta\right| + \frac{N_p - N_k}{M_k}\right)\right] = N_k + M_k\left(\frac{N_p - N_k}{M_k}\right) = N_p.$$

Algorithm 4.3. MSmode in splitting step IV of Algorithm 4.1

IV. Set $N_k := \sum_{\theta \in \underline{\Theta}} |J_k^\theta|$ and $M_k := \sum_{\theta \in \underline{\Theta}} 1\{|J_k^\theta| > 0\}$.

For $\theta \in \underline{\Theta}$, set $\rho_k^\theta = \frac{N_p - N_k}{M_k |J_k^\theta|}$. else $\rho_k^\theta = 0$.

Splitting: For $\theta \notin \underline{\Theta}$, $|J_k^\theta| = 0$; hence $\omega_k^{\theta,j} = 0$. For $\theta \in \underline{\Theta}$:

$$\left(\xi_k^{\theta,j}, \omega_k^{\theta,j}\right) = \left(\tilde{\xi}_k^{\theta,j}, \frac{\tilde{\omega}_k^{\theta,j}}{(1+\rho_k^\theta)}\right),\ j = 1,..,|J_k^\theta|$$

The total number of these particles is $\overline{N}_k$.

For $\theta \in \underline{\Theta}$, draw $\rho_k^\theta |J_k^\theta|$ additional multinomial samples:

$$\xi_k^{\theta,|J_k^\theta|+j} \sim \frac{\tilde{\pi}_k(\cdot,\cdot,\theta)}{\tilde{\pi}_{\theta_k}(\theta)},\ \omega_k^{\theta,|J_k^\theta|+j} = \frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|},\ \text{for } j = 1,.,\rho_k^\theta |J_k^\theta|$$

Total number of particles is then $\sum_{\theta \in \Theta} \left((1+\rho_k^\theta)|J_k^\theta|\right) = N_p$.

Set $N_k^\theta = (1+\rho_k^\theta)|J_k^\theta|$ and $\overline{\pi}_k(.) = \sum_{\theta \in \underline{\Theta}} \sum_{j=1}^{(1+\rho_k^\theta)|J_k^\theta|} \omega_k^{\theta,j} \delta_{\{\xi_k^{\theta,j}\}}(\cdot)$

In Algorithm 4.4, for IPS$_{\text{mode}}$-RMS$_{\text{mode}}$, each successful particle in $\tilde{\pi}_{\theta_k}(\theta)$ is first copied as much as possible the same number of times, and then the rest offspring are cloned randomly with replacement from the conditional measure $\frac{\tilde{\pi}_k(\cdot,\cdot,\theta)}{\tilde{\pi}_{\theta_k}(\theta)}$.

Algorithm 4.4. RMS$_{\text{mode}}$ in splitting step IV of Algorithm 4.1

IV. Set $N_k := \sum_{\theta \in \underline{\Theta}} |J_k^\theta|$ and $M_k := \sum_{\theta \in \underline{\Theta}} 1\{|J_k^\theta| > 0\}$.

For $\theta \in \underline{\Theta}$, set $\rho_k^\theta = \frac{N_p - N_k}{M_k |J_k^\theta|}$. else $\rho_k^\theta = 0$. Set $\alpha_k^\theta = \lfloor \rho_k^\theta \rfloor$.

Splitting: For $\theta \notin \underline{\Theta}$, $|J_k^\theta| = 0$; hence $\omega_k^{\theta,j} = 0$. For $\theta \in \underline{\Theta}$:

$$\left(\xi_k^{\theta,j+i|J_k^\theta|}, \omega_k^{\theta,j+i|J_k^\theta|}\right)_{i=0}^{\alpha_k^\theta} = \left(\tilde{\xi}_k^{\theta,j}, \frac{\tilde{\omega}_k^{\theta,j}}{(1+\rho_k^\theta)}\right),\ j = 1,..,|J_k^\theta|$$

The total number of these particles: $\sum_{\theta \in \Theta} \left((1+\alpha_k^\theta)|J_k^\theta|\right)$.

For $\theta \in \underline{\Theta}$, draw $(\rho_k^\theta - \alpha_k^\theta)|J_k^\theta|$ additional multinomial samples:

$$\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j} \sim \frac{\tilde{\pi}_k(\cdot,\cdot,\theta)}{\tilde{\pi}_{\theta_k}(\theta)}, \qquad \omega_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j} = \frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|},$$

$j = 1,.,(\rho_k^\theta - \alpha_k^\theta)|J_k^\theta|$

The total number of particles is then $\sum_{\theta \in \Theta} \left((1+\rho_k^\theta)|J_k^\theta|\right) = N_p$.

Set $N_k^\theta = (1+\rho_k^\theta)|J_k^\theta|$ and $\overline{\pi}_k(.) = \sum_{\theta \in \underline{\Theta}} \sum_{j=1}^{(1+\rho_k^\theta)|J_k^\theta|} \omega_k^{\theta,j} \delta_{\{\xi_k^{\theta,j}\}}(\cdot)$

In algorithm 4.5, for IPS$_{\text{mode}}$-FAS$_{\text{mode}}$, each successful particle in $\tilde{\pi}_{\theta_k}(\theta)$ is first copied as much as possible the same number of times, and then the rest offspring are cloned randomly without replacement from the set of successful particles. Hence the difference of FASmode with

RMSmode is that the remainder particles are obtained without replacement. As shown in Algorithm 4.5, the latter leads to a significant change in the evaluation of the particle weights.

Algorithm 4.5. FASmode in splitting step IV of Algorithm 4.1

IV. Set $N_k := \sum_{\theta\in\underline{\Theta}} |J_k^\theta|$ and $M_k := \sum_{\theta\in\underline{\Theta}} 1\{|J_k^\theta| > 0\}$.

For $\theta \in \underline{\Theta}$, set $\rho_k^\theta = \frac{N_p - N_k}{M_k |J_k^\theta|}$. else $\rho_k^\theta = 0$. Set $\alpha_k^\theta = \lfloor \rho_k^\theta \rfloor$.

Splitting: For $\theta \notin \underline{\Theta}$, $|J_k^\theta| = 0$; hence $\omega_k^{\theta,j} = 0$. For $\theta \in \underline{\Theta}$:

$$\left(\xi_k^{\theta,j+i|J_k^\theta|}\right)_{i=0}^{\alpha_k^\theta} = \left(\tilde{\xi}_k^{\theta,j}\right),\ j = 1,..,|J_k^\theta|$$

The total number of these particles: $\sum_{\theta\in\underline{\Theta}} \left((1+\alpha_k^\theta)|J_k^\theta|\right)$.

For $\theta \in \underline{\Theta}$, draw $(\rho_k^\theta - \alpha_k^\theta)|J_k^\theta|$ additional samples without replacement from the set $\{\tilde{\xi}_k^{\theta,j}, j = 1,..,|J_k^\theta|\}$. This yields additional copies: $\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j}$, $j = 1,.,(\rho_k^\theta - \alpha_k^\theta)|J_k^\theta|$.

The total number of copies is then $\sum_{\theta\in\Theta} \left((1+\rho_k^\theta)|J_k^\theta|\right) = N_p$.

Adaptation of the weights (assuming $\tilde{\xi}_k^{\theta,i} \neq \tilde{\xi}_k^{\theta,i'}$ for $i' \neq i$):

$$\omega_k^{\theta,j} = \sum_{i=1}^{|J_k^\theta|} \left[1\{\xi_k^{\theta,j} = \tilde{\xi}_k^{\theta,i}\}\tilde{\omega}_k^{\theta,i} / K_k^{\theta,i}\right],\ j = 1,..,(1+\rho_k^\theta)|J_k^\theta|,$$

where $K_k^{\theta,i} = \sum_{j=1}^{(1+\rho_k^\theta)|J_k^\theta|} 1\{\xi_k^{\theta,j} = \tilde{\xi}_k^{\theta,i}\},\ i = 1,..,|J_k^\theta|$.

Set $N_k^\theta = (1+\rho_k^\theta)|J_k^\theta|$ and $\bar{\pi}_k(.) = \sum_{\theta\in\underline{\Theta}} \sum_{j=1}^{(1+\rho_k^\theta)|J_k^\theta|} \omega_k^{\theta,j} \delta_{\{\xi_k^{\theta,j}\}}(\cdot)$

4.4.2. Characterization of conditional mean and variance

This subsection derives characterizations of the conditional mean and variance of $\bar{\gamma}_{k+1}$, given the information known at the begin of splitting step IV of IPSmode-MSmode, IPSmode-RMSmode and IPSmode-FASmode. First the derivation is done for IPSmode-RMSmode in Theorem 4.6.

Theorem 4.6: If we use RMS$_{\text{mode}}$ splitting step (Algorithm 4.4) at level k of the IPS$_{\text{mode}}$ algorithm. Then

$$\mathbb{E}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\omega}_k^{\theta,j} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.20}$$

$$\begin{aligned} \operatorname{Var}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} &= \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \left[\frac{\left(1+\alpha_k^\theta\right)}{\left(1+\rho_k^\theta\right)^2} (\tilde{\omega}_k^{\theta,i})^2 \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right] \\ &+ \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \sum_{j=1}^{|J_k^\theta|} \left[\frac{\left(\rho_k^\theta - \alpha_k^\theta\right)}{(1+\rho_k^\theta)^2 |J_k^\theta|} \tilde{\omega}_k^{\theta,i} \tilde{\omega}_k^{\theta,j} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right] \end{aligned} \tag{4.21}$$

Proof: See Appendix 4.C.1.

By setting $\alpha_k^\theta = 0$ in Theorem 4.6, we immediately get the conditional mean and variance for MSmode in the Corollary below.

Corollary 4.7: If we use MS$_{mode}$ splitting step (Algorithm 4.3) at level k of the IPS$_{mode}$ algorithm. Then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.22}$$

$$\begin{aligned}\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = &\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{\left(1+\rho_k^\theta\right)^2}(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right] \\ &+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\rho_k^\theta}{(1+\rho_k^\theta)^2\left|J_k^\theta\right|}\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]\end{aligned} \tag{4.23}$$

Subsequently the characterization for IPSmode-FASmode follows in Theorem 4.8.

Theorem 4.8: If we use FASmode splitting step (Algorithm 4.5) at level k of the IPS$_{mode}$ algorithm, then

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.24}$$

and

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(\alpha_k^\theta+1\right)\left(\alpha_k^\theta+2\right)}\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left(1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right] \tag{4.25}$$

Proof: See Appendix 4.C.2.

Comparison of Th. 4.1, Th. 4.6, Corollary 4.7 and Th. 4.8 shows that $\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ is the same (i.e. unbiased), whereas $\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ changes under different splitting steps IV.

4.5. Comparison of variances

This section compares relative dominance in terms of variances obtained under IPSmode-MRmode, IPSmode-MSmode, IPSmode-RMSmode and IPSmode-FASmode. First, subsection 4.5.1 compares, under specific conditions, the conditional variances when applying MRmode, MSmode, RMSmode and FASmode respectively at level k. Next, subsection 4.5.2 elaborates what this means for the relative dominance.

4.5.1. Comparison of conditional variances

We perform pairwise comparisons of the conditional variances at level k. Theorem 4.9 does so for RMSmode versus MSmode. Subsequently Theorem 4.10 does so for MSmode versus MR$_{mode}$; and Theorem 4.10* for MSmode versus FASmode.

Theorem 4.9: Given $C_k^{\tilde{\xi},\tilde{\omega}} = \sigma\left\{\tilde{\xi}_k^{\theta,j},\ \tilde{\omega}_k^{\theta,j}, \theta\in\Theta, j=1,...,|J_k^\theta|\right\}$ at IPS$_{mode}$ level k. If for every $\theta\in\underline{\Theta}$:

$$\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) < \frac{\tilde{\Sigma}_k^\theta}{\tilde{\pi}_{\theta_k}(\theta)}(1-\Delta_k^\theta),\ \text{ if } \tilde{\omega}_k^{\theta,i} > \overline{\tilde{\omega}}_k^\theta \tag{4.C3}$$

$$\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \geq \frac{\tilde{\Sigma}_k^{\theta}}{\tilde{\pi}_{\theta_k}(\theta)}(1+\Delta_k^{\theta}), \text{ if } \tilde{\omega}_k^{\theta,i} \leq \bar{\tilde{\omega}}_k^{\theta} \tag{4.C4}$$

with $\bar{\tilde{\omega}}_k^{\theta} = \frac{1}{\left|J_k^{\theta}\right|}\sum_j^{|J_k^{\theta}|} \tilde{\omega}_k^{\theta,j}$ and

$$\Delta_k^{\theta} = \frac{\sum_i^{|J_k^{\theta}|}\left[\tilde{\omega}_k^{\theta,i}\left(\tilde{\omega}_k^{\theta,i} - \bar{\tilde{\omega}}_k^{\theta}\right)\right]}{\sum_i^{|J_k^{\theta}|}\left[\left(1\{\tilde{\omega}_k^{\theta,i} > \bar{\tilde{\omega}}_k^{\theta}\} - 1\{\tilde{\omega}_k^{\theta,i} \leq \bar{\tilde{\omega}}_k^{\theta}\}\right)\tilde{\omega}_k^{\theta,i}\left(\tilde{\omega}_k^{\theta,i} - \bar{\tilde{\omega}}_k^{\theta}\right)\right]} \tag{4.26}$$

Then the dominance of IPSmode-RMSmode and IPSmode-MSmode methods in terms of $\text{Var}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ satisfies:

$$\text{Var}_{\text{IPSmode-RMSmode}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \leq \text{Var}_{\text{IPSmode-MSmode}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.27}$$

Proof: See Appendix 4.D.1.

Theorem 4.10: Given $C_k^{\tilde{\xi},\tilde{\omega}} = \sigma\left\{\tilde{\xi}_k^{\theta,j},\ \tilde{\omega}_k^{\theta,j}, \theta \in \Theta, j = 1,...,|J_k^{\theta}|\right\}$ at IPS$_{\text{mode}}$ level *k*. Let the conditions (4.C1-4.C2) of Theorem 4.3 and conditions (4.C3-4.C4) of Theorem 4.9 hold true, and let $\left|J_k^{\theta}\right|$ and $\tilde{\pi}_{\theta_k}(\theta)$ satisfy for $\theta \in \Theta$:

$$\left|J_k^{\theta}\right| \geq \frac{N_k}{M_k}, \quad \text{if } \tilde{\pi}_{\theta_k}(\theta) \geq \frac{1}{M_k} \tag{4.C5}$$

$$\left|J_k^{\theta}\right| < \frac{N_k}{M_k}, \quad \text{if } \tilde{\pi}_{\theta_k}(\theta) < \frac{1}{M_k} \tag{4.C6}$$

Then the dominance of IPSmode-MS$_{\text{mode}}$ and IPSmode-MR$_{\text{mode}}$ methods in terms of $\text{Var}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ is:

$$\text{Var}_{\text{IPSmode-MSmode}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \leq \text{Var}_{\text{IPSmode-MRmode}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.28}$$

Proof: See Appendix 4.D.2.

Theorem 4.10*: Given $C_k^{\tilde{\xi},\tilde{\omega}} = \sigma\left\{\tilde{\xi}_k^{\theta,j},\ \tilde{\omega}_k^{\theta,j}, \theta \in \Theta, j = 1,...,|J_k^{\theta}|\right\}$ at IPS$_{\text{mode}}$ level *k*. Let conditions (4.C3-4.C4) of Theorem 4.9 hold true. Then the dominance of IPS$_{\text{mode}}$-MS$_{\text{mode}}$ and IPS$_{\text{mode}}$-FAS$_{\text{mode}}$ methods in terms of $\text{Var}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ is:

$$\text{Var}_{\text{IPSmode-MSmode}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \leq \text{Var}_{\text{IPSmode-FASmode}}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.29}$$

Proof: See Appendix 4.D.3.

Ma and Blom (2022) have shown that, in terms of conditional variance, IPS-FAS dominates both IPS-MS and IPS-RMS. It is remarkable to see from Theorems 4.10 and 4.10* that, in terms of conditional variance, both IPSmode-MSmode and IPSmode-RMSmode dominate IPSmode-

FASmode. The explanation is that for FASmode it is more demanding to take proper account if the effect of particle weigths in splitting step IV.

4.5.2. Comparison of variances

Finally we perform pairwise comparisons of the overall variances, i.e. using the same splitting strategy at each level. Theorem 4.11 considers difference in splitting strategies at the *k*-th level only, and no differences in splitting strategy at the preceding levels. Subsequently Theorem 4.12 extends the results of Theorem 4.11 to cases of using the same splitting strategy at all levels.

Theorem 4.11: Suppose IPS_{mode} levels 1 to *k*-1 make use of the same type of splitting (either RMSmode or MSmode or MR_{mode}), and conditions 4.C1-4.C6 hold true. Then the dominance of the two splitting methods at level *k*, in terms of $V^k = \mathrm{Var}\left\{\prod_{k'=1}^{k} \bar{\gamma}_{k'}\right\}$ satisfies:

$$V^k_{IPS_{mode}\text{-}RMS_{mode}} \leq V^k_{IPS_{mode}\text{-}MS_{mode}} \leq V^k_{IPS_{mode}\text{-}MR_{mode}} \tag{4.30}$$

Proof: Apply the reasoning in the proof of Theorem 3 in (Ma and Blom, 2022) to the results of Theorems 4.9 and 4.10.

Theorem 4.12: Suppose the same type of Splitting (either RMSmode or MSmode or MR_{mode}) is used at all levels, and conditions 4.C1-4.C6 hold true for levels 1 to *m*. Then the dominance of the two splitting methods in terms of variance $V = \mathrm{Var}\{\bar{\gamma}\}$ satisfies:

$$V_{IPS_{mode}\text{-}RMS_{mode}} \leq V_{IPS_{mode}\text{-}MS_{mode}} \leq V_{IPS_{mode}\text{-}MR_{mode}} \tag{4.31}$$

Proof: Apply the reasoning in the proof of Theorem 4 in (Ma and Blom, 2022) to the result of Theorem 4.11.

4.6. Rare event simulation example

4.6.1. Hypothetical car example

A car driver in dense fog is heading to a wall at position d_{wall}. If the car is at distance d_{fog} from the wall, then the driver sees the wall for the first time. Then, it takes the driver a random reaction delay to start braking, with a density $p_{delay}(s)$. During the reaction delay, the velocity of the car does not change; after the reaction delay, the car decelerates at constant value $a_{\min}$. We apply IPS_{mode} to estimate the probability γ that the car hits the wall.

From the moment that the car reaches distance d_{fog} from the wall at velocity v_0, it takes the sum of reaction delay T_{delay} and the time of deceleration $T_{dec} = -v_0 / a_{\min}$ until the car is at a standstill. This implies

$$\gamma = \mathrm{P}\{v_0 T_{delay} + v_0 T_{dec} + \tfrac{1}{2} a_{\min} T_{dec}^2 \geq d_{fog}\} \tag{4.32}$$

Elaboration of (4.32) yields:

$$\gamma = \mathrm{P}\{T_{delay} \geq \tfrac{1}{2} v_0 / a_{\min} + d_{fog} / v_0\} \tag{4.33}$$

If we assume a Rayleigh density $p_{delay}(s) = \frac{s}{\mu^2} e^{-s^2/(2\mu^2)}$, and we write $T_C = \frac{1}{2} v_0 / a_{\min} + d_{fog} / v_0$, evaluation of (4.33) yields:

$$\gamma = \int_{T_C}^{+\infty} \frac{t}{\mu^2} e^{-t^2/(2\mu^2)} dt = -e^{-t^2/(2\mu^2)} \Big|_{t=T_C}^{+\infty} = e^{-t^2/(2\mu^2)} \big|_{t=T_C} \tag{4.34}$$

Table 4.1 gives the analytically obtained γ results for various mean reaction delay values μ, and parameter values $d_{wall} = 300m, d_{fog} = 120m, v_0 = 72 km/h = 20 m/s$, and $a_{\min} = -4 m/s^2$.

Table 4.1 Analytical γ results for various μ

μ (s)	γ
0.9	5.19976×10^{-4}
0.8	6.97696×10^{-5}
0.7	3.72665×10^{-6}
0.6	4.08284×10^{-8}

4.6.2. GSHS model

For this example, the discrete set of the GSHS is:

$$\Theta = \{-1, 0, 1, delay, stop, hit\} \tag{4.35}$$

where -1 indicates decelerating mode, 0 indicates uniform mode, 1 indicates accelerating mode, $delay$ is a reaction delay mode, $stop$ indicates stopping mode, and hit indicates the wall has been hit. A transition diagram representing the transitions between these modes is given in Figure 4.1.

The continuous state components are $x_t = Col(z_t, y_t, v_t)$, where z_t is the amount of time passed since the driver could see the wall for the first time, y_t is the position of the car at time t, and v_t is the velocity at time t. Hence, the dimension of the continuous state space is $d(.) = 3$. The subsets X^θ are defined as follows:

$$\begin{aligned}
X^0 &= \mathbb{R} \times (-\infty, d_{wall} - d_{fog}) \times \mathbb{R} \\
X^1 &= \mathbb{R} \times (-\infty, d_{wall} - d_{fog}) \times (0, v_{\max}) \\
X^{-1} &= \mathbb{R} \times (-\infty, d_{wall}) \times (0, \infty) \\
X^{delay} &= \mathbb{R} \times (-\infty, d_{wall}) \times \mathbb{R} \\
X^{stop} &= \mathbb{R} \times (-\infty, d_{wall}) \times 0 \\
X^{hit} &= \mathbb{R}^3
\end{aligned} \tag{4.36}$$

Between switching moment of $\{\theta_t\}$, x_t evolves as follows:

$$\begin{aligned}
dz_t &= dt \\
dy_t &= v_t dt \\
dv_t &= \theta_t(\theta_t - 1) a_{\min} / 2 + \theta_t(\theta_t + 1) a_{\max} / 2
\end{aligned} \tag{4.37}$$

where $a_{\min}$ is the deceleration value and $a_{\max}$ is the acceleration value. The initial measure $Init$ generates $\theta_0 = 0$, $z_0 = 0$, $y_0 = 0.$.

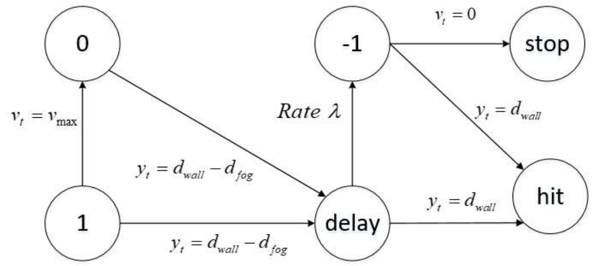

Figure 4.1. State transition diagram of car example GSHS.

The instantaneous transition rate $\lambda(\theta_t,(z_t,y_t,v_t))$ satisfies:

$$\lambda(\theta,(z,y,v)) = \chi(\theta = delay) p_{delay}(z) / \int_z^\infty p_{delay}(s) ds \tag{4.38}$$

The transition measure $R((\theta,(z,y,v)),(.,.))$ satisfies:

$R((1,(z,y,v)),\{0\}\times\{0,y,v\})=1$ iff $v=v_{max}$

$R((-1,(z,y,v)),\{stop\}\times\{0,y,v\})=1$ iff $v=0$

$R((0,(z,y,v)),\{delay\}\times\{0,y,v\})=1$ iff $y=d_{wall}-d_{fog}$

$R((1,(z,y,v)),\{delay\}\times\{0,y,v\})=1$ iff $y=d_{wall}-d_{fog}$

$R((delay,(z,y,v)),\{-1\}\times\{0,y,v\})=1,$ iff λ generates a point,

$R((delay,(z,y,v)),\{hit\}\times\{0,y,0\})=1,$ iff $y=d_{wall}$

$R((-1,(z,y,v)),\{hit\}\times\{0,y,0\})=1,$ iff $y=d_{wall}$.

4.6.3. Simulation results

We adopt the following levels, $D_k = \{0,1,delay,hit\}\times\mathbb{R}\times[L_k,\infty)\times\mathbb{R}\cup\{-1,stop\}\times\mathbb{R}\times[d_{wall},\infty)\times\mathbb{R}$, with L_k values shown in Table 4.2.

Table 4.2. Values of L_k for various μ values

k \ μ	$0.9s$	$0.8s$	$0.7s$	$0.6s$
1	181	181	181	181
2	217	215	210	205
3	230	230	220	215
4	240	241	230	223
5	300	300	237	230
6			244	236
7			300	243
8				300

By conducting IPSmode N_{IPS} times we get $\bar{\gamma}^i, i=1,..,N_{IPS}$. These results are used to assess the mean $\hat{\gamma}$, the percentage ρ_S of successful IPSmode runs, and the normalized root-mean-square error (RMSE), i.e.

$$\hat{\gamma} = \frac{1}{N_{IPS}} \sum_{i=1}^{N_{IPS}} \bar{\gamma}^i \tag{4.39}$$

$$\rho_S = \frac{1}{N_{IPS}} \sum_{i=1}^{N_{IPS}} 1(\bar{\gamma}^i > 0) \tag{4.40}$$

$$RMSE = \sqrt{\frac{1}{N_{IPS}} \sum_{i=1}^{N_{IPS}} (\bar{\gamma}^i - \gamma)^2} \tag{4.41}$$

Table 4.3 shows the estimation simulation results for straightforward MC, IPS-MR, IPSmode, IPSmode-MRmode, IPSmode-MSmode, IPSmode-RMSmode, and IPSmode-FASmode, for mean reaction delay ranging from μ =0.9s till μ =0.6s.

Table 4.3. Simulation results for MC, IPS-MR, IPSmode-MRmode, IPSmode-MSmode, IPSmode-RMSmode and IPSmode-FASmode applied to GSHS model for Rayleigh mean delay for the L_k and μ values in Table 4.2, and $\Delta = 0.01s$, $N_P = 1000$ and $N_{IPS} = 100$

$\mu = 0.9s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC	5.300×10^{-4}	44%	134.61%
IPS-MR	5.124×10^{-4}	100%	17.30%
IPSmode (Krystul et al.)	4.537×10^{-4}	100%	45.77%
IPSmode-MRmode	5.113×10^{-4}	100%	16.86%
IPSmode-MSmode	5.087×10^{-4}	100%	15.00%
IPSmode-RMSmode	5.135×10^{-4}	100%	14.61%
IPSmode-FASmode	5.105×10^{-4}	100%	15.41%
$\mu = 0.8s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC	4.000×10^{-5}	4%	495.54%
IPS-MR	7.074×10^{-5}	100%	23.19%
IPSmode (Krystul et al.)	6.742×10^{-5}	100%	53.41%
IPSmode-MRmode*	6.985×10^{-5}	100%	21.13%
IPSmode-MSmode	6.897×10^{-5}	100%	19.36%
IPSmode-RMSmode	6.910×10^{-5}	100%	19.27%
IPSmode-FASmode	6.946×10^{-5}	100%	19.43%
$\mu = 0.7s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC	0	0%	∞
IPS-MR	3.673×10^{-6}	100%	22.06%
IPSmode (Krystul et al.)	3.309×10^{-6}	100%	53.47%
IPSmode-MRmode	3.642×10^{-6}	100%	20.44%
IPSmode-MSmode	3.686×10^{-6}	100%	19.33%
IPSmode-RMSmode	3.669×10^{-6}	100%	18.27%
IPSmode-FASmode	3.682×10^{-6}	100%	19.57%

$\mu = 0.6s$	$\hat{\gamma}$	ρ_S	$RMSE/\hat{\gamma}$
MC	0	0%	∞
IPS-MR	4.094×10^{-8}	100%	34.20%
IPS_{mode} (Krystul et al.)	3.820×10^{-8}	100%	69.49%
IPS_{mode}-MR_{mode}	4.061×10^{-8}	100%	27.92%
IPS_{mode}-MS_{mode}	4.113×10^{-8}	100%	27.02%
IPS_{mode}-RMS_{mode}	4.008×10^{-8}	100%	26.77%
IPS_{mode}-FAS_{mode}	4.023×10^{-8}	100%	27.89%

The results in Table 4.3 show that for this GSHS example, all IPS versions outperform straightforward MC simulation. IPS_{mode}-RMS_{mode} yields lowest RMSE value, second is IPS_{mode}-MS_{mode}, third is IPS_{mode}-FAS_{mode}, fourth is IPS_{mode}-MR_{mode}, fifth is normal IPS (IPS-MR), and last is IPS_{mode} of (Krystul et al., 2012). We also verified that conditions 4.C1-4.C6 were satisfied for each simulated particle, and at each level. This sequence corresponds with the comparison of the variances for IPS_{mode}-RMS_{mode}, IPS_{mode}-MS_{mode}, IPS_{mode}-FAS_{mode}, and normal IPS (IPS-MR) in Theorem 4.5 and Theorem 4.12.

4.7. Conclusion

This chapter has developed novel sampling per mode strategies for use in IPS based estimation of reach probability for a general stochastic hybrid system (GSHS). The starting point is formed by the IPSmode algorithm of Krystul et al. (2012); this has been described in Section 4.2. In Section 4.3, IPSmode-MRmode has been proposed as a straightforward improvement of this IPSmode algorithm. In addition, it has been shown that under specific conditions, IPSmode-MRmode yields a variance of estimated reach probability that is lower or equal than those of basic IPS. In Section 4.4, three additional sampling per mode strategies have been developed, yielding: IPS_{mode}-MS_{mode}, IPS_{mode}-RMS_{mode} and IPS_{mode}-FAS_{mode}. The crucial part was to capture the effect of particle weights in each sampling per mode strategy such that the estimated reach probability remains unbiased. In section 4.5, it is shown that, under specific conditions, IPS_{mode}-RMS_{mode} performs best, IPSmode-MSmode performs second, while both dominate IPS_{mode}-FAS_{mode} as well as IPS_{mode}-RMS_{mode}. In section 4.6, the various IPS versions have been simulated for a simple GSHS example; the simulation results obtained show similar relative performance of the different IPS versions.

4.8 References

Au, S.K., Beck, J.L. (2003), Subset simulation and its application to seismic risk based on dynamic analysis," J. Eng. Mech., Vol. 129, pp. 901–917.

Blom, H.A.P, Ma, H. and Bakker, G.J. (2018). Interacting Particle System-based Estimation of Reach Probability for a Generalized Stochastic Hybrid System. IFAC-PapersOnLine, 51(16), pp.79-84.

Botev, Z.I. and Kroese, D.P. (2008). An efficient algorithm for rare-event probability estimation, combinatorial optimization, and counting. Methodology and Computing in Applied Probability 10(4), pp. 471-505.

Bujorianu, M.L. and Lygeros, J. (2006). Toward a general theory of stochastic hybrid systems. Eds: H.A.P. Blom and J. Lygeros, Stochastic hybrid systems, Springer, Berlin, pp. 3-30.

Cérou, F., Del Moral, P., Legland, F., and Lezaud, P. (2006). Genetic genealogical models in rare event analysis. Latin American J. of Probability & Mathematical Statistics, 1, pp. 181-203.

Cérou, F., Del Moral, P., Furon, T., Guyader, A. (2012). Sequential Monte Carlo for rare event estimation, Statistics and Computing, Vol. 22, pp. 795-808.

Cérou, F., Guyader, A., Rousset, M. (2019), Adaptive multilevel splitting: Historical perspective and recent results, Chaos: An Interdisciplinary Journal of Nonlinear Science, Vol. 29, pp. 43-108; https://doi.org/10.1063/1.5082247.

Del Moral, P. (2004). Feynman-Kac formulae: genealogical and interacting particle systems with applications, Springer.

Del Moral, P., Kouroutzin, M., Miclo., L. (2001). On a class of discrete generation interacting particle systems, Electron. J. Probab. Vol. 6, pp. 1-26.

Garvels, M.J. and Kroese, D.P. (1998). A comparison of RESTART implementations. In Proceedings of the 30th conference on Winter simulation, pp. 601-608. IEEE Computer Society Press.

Gerber, M., Chopin, N. and Whiteley, N. (2019). Negative association, ordering and convergence of resampling methods. Annals of Statistics, 47(4), pp.2236-2260.

Glasserman, P., Heidelberger, P., Shahabuddin, P. and Zajic, T. (1999). Multilevel splitting for estimating rare event probabilities, Operations Research, Vol. 47, pp. 585-600.

Krystul, J., Le Gland, F., Lezaud, P. (2012). Sampling per mode for rare event simulation in switching diffusions, Stochastic Processes and their Applications, Vol. 122, pp. 2939-2667.

L'Ecuyer, P., LeGland, F., Lezaud, P. and Tuffin, B. (2009). Splitting techniques, Eds: G. Rubino, B. Tuffin, Rare event simulation using Monte Carlo methods, Wiley, pp. 39-61.

L'Ecuyer, P., Demers, V. and Tuffin, B. (2007). Rare events, splitting, and quasi-Monte Carlo. ACM Transactions on Modeling and Computer Simulation. 17(2) pp. 9–es. https://doi.org/10.1145/1225275.1225280

Le Gland, F., and Oudjane, N. (2006). A sequential particle algorithm that keeps the particle system alive, Eds: H.A.P. Blom and J. Lygeros, Stochastic hybrid systems, Springer, Berlin, pp. 351-389.

Ma, H. and Blom, H.A.P. (2022). Fixed Assignment vs. Random Assignment in Multilevel Importance Splitting for Estimating Stochastic Reach Probability, Methodology & Computing in Applied Probability, On-line January 2022, https://doi.org/10.1007/s11009-021-09892-4.

Ma, H. and Blom, H.A.P., 2023. Interacting particle system based estimation of reach probability of general stochastic hybrid systems. Nonlinear Analysis: Hybrid Systems, 47, p.101303. https://doi.org/10.1016/j.nahs.2022.101303

Rubinstein, R. (2010). Randomized algorithms with splitting: Why the classic randomization algorithms do not work and how to make them work, Methodology and Computing in Applied Probability, Vol. 12, pp. 1-50.

Appendix 4.A.1: Proof of Theorem 4.1

For IPS$_{mode}$ step II at level $k+1$, $\overline{\gamma}_{k+1}$ is defined as follows:

$$\overline{\gamma}_{k+1} = \sum_{\theta \in \Theta} \sum_{j=1}^{N_k^\theta} \left[\omega_k^{\theta,j} 1\left(\overline{\xi}_{k+1}^{\theta,j} \in Q_{k+1} \right) \right]$$

If $\left| J_k^\theta \right| = 0$, then $\omega_k^{\theta,j} = 0, \;\; j = 1,...,N_k^\theta$. Hence

$$\overline{\gamma}_{k+1} = \sum_{\theta \in \underline{\Theta}} \sum_{j=1}^{N_k^\theta} \left[\omega_k^{\theta,j} 1\left(\overline{\xi}_{k+1}^{\theta,j} \in Q_{k+1} \right) \right] \tag{4.42}$$

where underlining of Θ indicates $\left| J_k^\theta \right| > 0$.

If $\left| J_k^\theta \right| > 0$ in step IV at level k, $\omega_k^{\theta,j}$, satisfies:

$$\omega_k^{\theta,j} = \tilde{\pi}_{\theta_k}(\theta) M / N_p \qquad \text{if } j = 1,...,N_k^\theta$$

Substitution in (4.42) yields:

$$\begin{aligned}
\overline{\gamma}_{k+1} &= \sum_{\theta \in \underline{\Theta}} \sum_{j=1}^{N_k^\theta} \left[\tilde{\pi}_{\theta_k}(\theta) M / N_p 1\left(\overline{\xi}_{k+1}^{\theta,j} \in Q_{k+1} \right) \right] = \sum_{\theta \in \underline{\Theta}} \sum_{j=1}^{N_k^\theta} \left[\tilde{\pi}_{\theta_k}(\theta) M / N_p 1\left(\mathcal{M}(\xi_k^{\theta,j}) \in Q_{k+1} \right) \right] \\
&= \sum_{\theta \in \underline{\Theta}} \sum_{j=1}^{N_k^\theta} \left[\tilde{\pi}_{\theta_k}(\theta) M / N_p 1\left(\mathcal{M}(\xi_k^{\theta,j}) \in Q_{k+1} \right) \sum_{i=1}^{\left| J_k^\theta \right|} 1\left(\xi_k^{\theta,j} = \tilde{\xi}_k^{\theta,i} \right) \right] \\
&= \sum_{\theta \in \Theta} \sum_{j=1}^{N_k^\theta} \sum_{i=1}^{\left| J_k^\theta \right|} \left[\tilde{\pi}_{\theta_k}(\theta) M / N_p 1\left(\mathcal{M}(\xi_k^{\theta,j}) \in Q_{k+1} \right) 1\left(\xi_k^{\theta,j} = \tilde{\xi}_k^{\theta,i} \right) \right] \\
&= \sum_{\theta \in \underline{\Theta}} \sum_{i=1}^{\left| J_k^\theta \right|} \left[\tilde{\pi}_{\theta_k}(\theta) M / N_p 1\left(\mathcal{M}(\tilde{\xi}_k^{\theta,i}) \in Q_{k+1} \right) \sum_{j=1}^{N_k^\theta} 1\left(\xi_k^{\theta,j} = \tilde{\xi}_k^{\theta,i} \right) \right]
\end{aligned}$$

where $\mathcal{M}(\cdot)$ represents the mutation at Step I.

If we let $\tilde{Y}_{k+1}^{k,\theta,i}$ with $i = 1,...,\left| J_k^\theta \right|$ be the number of the $K_k^{\theta,i}$ particle copies from $\tilde{\xi}_k^{\theta,i}$ that reach Q_{k+1} after mutation, then we have

$$\tilde{Y}_{k+1}^{k,\theta,i} = 1\left(\mathcal{M}(\tilde{\xi}_k^{\theta,i}) \in Q_{k+1} \right) \tilde{K}_k^{\theta,i}$$

with

$$\tilde{K}_k^{\theta,i} = \sum_{j=1}^{N_p/M} 1\left(\xi_k^{\theta,j} = \tilde{\xi}_k^{\theta,i} \right)$$

Combining the above three equations yields:

$$\overline{\gamma}_{k+1} = \sum_{\eta \in \underline{\Theta}} \left[\tilde{\pi}_{\theta_k}(\theta) M / N_p \sum_{i=1}^{\left| J_k^\eta \right|} \left[\tilde{Y}_{k+1}^{k,\theta,i} \right] \right] \tag{4.43}$$

Using eq. (4.43) yields:

$$\mathbb{E}\left\{ \overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}} \right\} = \mathbb{E}\left\{ \sum_{\theta \in \underline{\Theta}} \sum_{i=1}^{\left| J_k^\theta \right|} \left[\left(\tilde{\pi}_{\theta_k}(\theta) M / N_p \right) \tilde{Y}_{k+1}^{k,\theta,i} \right] \mid C_k^{\tilde{\xi},\tilde{\omega}} \right\} = \sum_{\theta \in \underline{\Theta}} \sum_{i=1}^{\left| J_k^\theta \right|} \left[\left(\tilde{\pi}_{\theta_k}(\theta) M / N_p \right) \mathbb{E}\{ \tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}} \} \right] \tag{4.44}$$

In a similar way, we derive:

$$\begin{aligned}
&\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&=\sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\operatorname{Cov}\left\{\left[\left(\tilde{\pi}_{\theta_k}(\theta)M/N_p\right)\tilde{Y}_{k+1}^{k,\theta,i}\right],\left[\left(\tilde{\pi}_{\theta_k}(\eta)M/N_p\right)\tilde{Y}_{k+1}^{k,\eta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&=\sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\eta)M^2}{N_p^{\ 2}}\operatorname{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&=\sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\eta)M^2}{N_p^{\ 2}}\left[\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]
\end{aligned} \tag{4.45}$$

Each $\tilde{Y}_{k+1}^{k,\theta,i}$ has a conditional Binomial distribution with success probability $\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)$ and size $K_k^{\theta,i}$.

Let us define $C_k^{\tilde{\xi},\,\tilde{\omega},K}$ as follows:

$$C_k^{\tilde{\xi},\,\tilde{\omega},K} \triangleq \sigma\left\{\tilde{\xi}_k^{\theta,j},\tilde{\omega}_k^{\theta,j},K_k^{\theta,j};\, j=1,\ldots,|J_k^\theta|,\theta\in\Theta\right\}$$

Then we also get:

$$\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},\tilde{K}}\right\}=K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \tag{4.46}$$

$$\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\,\tilde{\omega}}\right\} \tag{4.47}$$

$$\operatorname{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_{1:k}^{\tilde{\xi},\,\tilde{\omega},\tilde{K}}\right\}=K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.48}$$

Hence

$$\begin{aligned}
&\operatorname{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\,\tilde{\omega}}\right\} \\
&=\mathbb{E}\left\{(\tilde{Y}_{k+1}^{k,\theta,i})^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&=\mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i})^2\} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&=\mathbb{E}\left\{Var\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}+\mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\}^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&=\mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}+\mathbb{E}\left\{[K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)]^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&=\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}+\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\left[\mathbb{E}\left\{[K_k^{\theta,i}]^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2\right] \\
&=\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}+\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2 Var\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}
\end{aligned} \tag{4.49}$$

Substituting (4.47) in eq. (4.44) yields:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\pi}_{\theta_k}(\theta)M/N_p\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\,\tilde{\omega}}\right\}\right] \tag{4.50}$$

If $\theta\neq\eta$, then $\tilde{Y}_{k+1}^{k,\theta,i}$ and $\tilde{Y}_{k+1}^{k,\eta,j}$ are conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega}}$, and eq. (4.45) becomes:

$$\begin{aligned}
&\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&=\sum_{\theta\in\Theta}\sum_{\substack{\eta\in\Theta \\ \eta\neq\theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\eta)M^2}{N_p^{\ 2}}\left[\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&=0
\end{aligned} \tag{4.51}$$

If $\theta = \eta$ and $i \neq j$ then $\tilde{Y}_{k+1}^{k,\theta,i}$ and $\tilde{Y}_{k+1}^{k,\eta,j}$ are conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega},K}$. Hence eq. (4.45) becomes:

$$\begin{aligned}&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\\ &=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\theta)M^2}{N_p{}^2}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]\end{aligned} \quad (4.52)$$

Substitution of (4.46) yields:

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\theta)M^2}{N_p{}^2}\left[\mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)K_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{K_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \quad (4.53)$$

Because $\mathbb{E}\{f(Z) \mid Z\} = f(Z)$ the latter simplifies to:

$$\begin{aligned}&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\\ &=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\theta)M^2}{N_p{}^2}\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\mathbb{E}\left\{K_k^{\theta,i}K_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\mathbb{E}\left\{K_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]\\ &=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2M^2}{N_p{}^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\mathrm{Cov}\left\{K_k^{\theta,i},K_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\end{aligned} \quad (4.54)$$

If $\theta = \eta$ and $i = j$ then eq. (4.45) becomes:

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2M^2}{N_p{}^2}Var\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \quad (4.55)$$

Substitution of (4.49) yields:

$$\begin{aligned}&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\\ &=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2M^2}{N_p{}^2}\left(\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right)\right]+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2M^2}{N_p{}^2}\left(\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mathrm{Var}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right)\right]\end{aligned} \quad (4.56)$$

Combining (4.51), (4.54) and (4.56) yields:

$$\begin{aligned}&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\\ &=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2M^2}{N_p{}^2}\left(\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right)\right]\\ &+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2M^2}{N_p{}^2}\left(\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mathrm{Var}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right)\right]\\ &+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2M^2}{N_p{}^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\mathrm{Cov}\left\{K_k^{\theta,i},K_k^{\theta,j} \mid C_{1:k}^{\tilde{\xi},\tilde{\omega}}\right\}\right]\end{aligned} \quad (4.57)$$

For $\theta \in \underline{\Theta}$, the vector $(K_k^{\theta,1}, K_k^{\theta,2}, ..., K_k^{\theta,|J_k^\theta|})$ has a multinomial distribution with number of trials equal to N_p/M, and with success probabilities $\tilde{\omega}_k^i / \tilde{\pi}_{\theta_k}(\theta)$. Multinomial distribution properties yields

$$\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \frac{N_p}{M}\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)} \tag{4.58}$$

$$\mathrm{Var}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \frac{N_p}{M}\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)}\left(1-\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)}\right) \tag{4.59}$$

$$\mathrm{Cov}\left\{K_k^{\theta,i}K_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \frac{-N_p}{M}\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)}\frac{\tilde{\omega}_k^{\theta,j}}{\tilde{\pi}_{\theta_k}(\theta)} \tag{4.60}$$

Inserting (4.58)-(4.60) into eqs. (4.50) and (4.57) and subsequent evaluation yields eq. (4.10) in Theorem 4.1, and:

$$\begin{aligned}
&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&= \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\pi}_{\theta_k}(\theta)\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left(1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right] + \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\left(\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\tilde{\omega}_k^{\theta,i}\left(\tilde{\pi}_{\theta_k}(\theta)-\tilde{\omega}_k^{\theta,i}\right)\right)\right] - \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right] \\
&= \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\pi}_{\theta_k}(\theta)\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] - \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right] \\
&= \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\left(\sum_{j=1}^{|J_k^\theta|}\tilde{\omega}_k^{\theta,j}\right)\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] - \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right] \\
&= \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,j}\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] - \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right] \\
&= \frac{M}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right]
\end{aligned} \tag{4.61}$$

Rewriting the latter yields eq. (4.11) in Th. 4.1. **Q.E.D.**

Appendix 4.A.2: Proof of Theorem 4.2

For IPS$_{\text{mode}}$ Step II at level $k+1$, $\overline{\gamma}_{k+1}$ becomes:

$$\overline{\gamma}_{k+1} = \sum_{\theta\in\Theta}\sum_{j=1}^{N_k^\theta}\left[\omega_k^{\theta,j}1\left(\overline{\xi}_{k+1}^{\theta,j}\in Q_{k+1}\right)\right] \tag{4.62}$$

For $\theta\in\Theta$ in step IV at level *k*, we have:

$$\tilde{\tilde{\omega}}_k^j = \tfrac{1}{N_p},\ j=1,..,N_p,$$

$$\left\{\tilde{\tilde{\omega}}_k^j,\tilde{\tilde{\xi}}_k^j\right\}_{j=1}^{N_p} \text{ one-on-one to } \left\{\omega_k^{\theta,i},\xi_k^{\theta,i};\theta\in\Theta, i=1,..,N_k^\theta\right\}$$

Thus, $\omega_k^{\theta,j}$, satisfies:

$$\omega_k^{\theta,j} = \frac{1}{N_p} \quad \text{if } j=1,...,N_k^\theta$$

Substitution in (4.62) yields:

$$\bar{\gamma}_{k+1} = \sum_{\theta\in\Theta}\sum_{j=1}^{N_k^\theta}\left[\frac{1}{N_P}1\left(\bar{\xi}_{k+1}^{\theta,j}\in Q_{k+1}\right)\right] = \sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{N_k^\theta}\left[\frac{1}{N_P}1\left(\mathsf{M}(\xi_k^{\theta,j})\in Q_{k+1}\right)\right] = \sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{N_k^\theta}\left[\frac{1}{N_P}1\left(\mathsf{M}(\xi_k^{\theta,j})\in Q_{k+1}\right)\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\eta|}1\left(\xi_k^{\theta,j}=\tilde{\xi}_k^{\eta,i}\right)\right]$$
$$= \sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{N_k^\theta}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\eta|}\left[\frac{1}{N_P}1\left(\mathsf{M}(\xi_k^{\theta,j})\in Q_{k+1}\right)1\left(\xi_k^{\theta,j}=\tilde{\xi}_k^{\eta,i}\right)\right] = \sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{N_k^\theta}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\eta|}\left[\frac{1}{N_P}1\left(\mathsf{M}(\tilde{\xi}_k^{\eta,i})\in Q_{k+1}\right)1\left(\xi_k^{\theta,j}=\tilde{\xi}_k^{\eta,i}\right)\right]$$
$$= \sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\eta|}\left[\frac{1}{N_P}1\left(\mathsf{M}(\tilde{\xi}_k^{\eta,i})\in Q_{k+1}\right)\sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{N_k^\theta}1\left(\xi_k^{\theta,j}=\tilde{\xi}_k^{\eta,i}\right)\right] = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_P}1\left(\mathsf{M}(\tilde{\xi}_k^{\theta,i})\in Q_{k+1}\right)\sum_{\eta\in\underline{\Theta}}\sum_{j=1}^{N_k^\eta}1\left(\xi_k^{\eta,j}=\tilde{\xi}_k^{\theta,i}\right)\right]$$

where $\mathsf{M}(\cdot)$ represents the mutation at Step I.

If we let $\tilde{Y}_{k+1}^{k,\theta,i}$ with $i=1,...,|J_k^\theta|$ be the number of the $K_k^{\theta,i}$ particle copies from $\tilde{\xi}_k^{\theta,i}$ that reach Q_{k+1} after mutation, then we have:

$$\tilde{Y}_{k+1}^{k,\theta,i} = 1\left(\mathsf{M}(\tilde{\xi}_k^{\theta,i})\in Q_{k+1}\right)\tilde{K}_k^{\theta,i}$$

with

$$\tilde{K}_k^{\theta,i} = \sum_{\eta\in\underline{\Theta}}\sum_{j=1}^{N_k^\eta}1\left(\xi_k^{\eta,j}=\tilde{\xi}_k^{\theta,i}\right) \tag{4.63}$$

Combining the above three equations yields:

$$\bar{\gamma}_{k+1} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\tfrac{1}{N_P}\tilde{Y}_{k+1}^{k,\theta,i}\right] \tag{4.64}$$

Using eq. (4.64) yields:

$$\mathbb{E}\left\{\bar{\gamma}_{k+1}\,|\,C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \mathbb{E}\left\{\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\tfrac{1}{N_p}\tilde{Y}_{k+1}^{k,\theta,i}\right]|\,C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\tfrac{1}{N_p}\mathbb{E}\{\tilde{Y}_{k+1}^{k,\theta,i}\,|\,C_k^{\tilde{\xi},\tilde{\omega}}\}\right] \tag{4.65}$$

In a similar way, we derive:

$$\begin{aligned}&\mathrm{Var}\left\{\bar{\gamma}_{k+1}\,|\,C_k^{\tilde{\xi},\tilde{\omega}}\right\}\\ &= \sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\mathrm{Cov}\left\{\left[\tfrac{1}{N_p}\tilde{Y}_{k+1}^{k,\theta,i}\right],\left[\tfrac{1}{N_p}\tilde{Y}_{k+1}^{k,\eta,j}\right]|\,C_k^{\tilde{\xi},\tilde{\omega}}\right\}\\ &= \sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\frac{1}{N_p^{\,2}}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j}\,|\,C_k^{\tilde{\xi},\tilde{\omega}}\right\}\end{aligned} \tag{4.66}$$

Let us define $C_k^{\tilde{\xi},\,\tilde{\omega},K}$ as follows:

$$C_k^{\tilde{\xi},\,\tilde{\omega},K} \triangleq \sigma\left\{\tilde{\xi}_k^{\theta,j},\tilde{\omega}_k^{\theta,j},K_k^{\theta,j};\,j=1,...,|J_k^\theta|,\theta\in\Theta\right\}$$

Then we also get:

$$\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\,|\,C_k^{\tilde{\xi},\tilde{\omega},\tilde{K}}\right\} = K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \tag{4.67}$$

$$\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\,|\,C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mathbb{E}\left\{K_k^{\theta,i}\,|\,C_k^{\tilde{\xi},\,\tilde{\omega}}\right\} \tag{4.68}$$

$$\mathrm{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\,|\,C_k^{\tilde{\xi},\,\tilde{\omega},\tilde{K}}\right\} = K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.69}$$

Hence

$$\begin{aligned}
&\mathrm{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \mathbb{E}\left\{(\tilde{Y}_{k+1}^{k,\theta,i})^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&= \mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i})^2\} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&= \mathbb{E}\left\{Var\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\}^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\mathbb{E}\{(\tilde{Y}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&= \mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \mathbb{E}\left\{[K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)]^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2 \\
&= \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\left[\mathbb{E}\left\{[K_k^{\theta,i}]^2 \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}^2\right] \\
&= \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2 Var\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}
\end{aligned} \tag{4.70}$$

Substituting (4.68) in eq. (4.65) yields:

$$\mathbb{E}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\left(1/N_p\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \tag{4.71}$$

Let us define a set $S_{\eta,j} \triangleq \{(\eta,j)\in\underline{\Theta}\times[1,|J_k^\eta|];\eta\neq\theta \text{ or } j\neq i\}$, then eq. (4.66) can be written as:

$$\mathrm{Var}\left\{\bar{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\frac{1}{N_p^{\,2}}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\notin S_{\eta,j}}\frac{1}{N_p^{\,2}}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.72}$$

If $\theta=\eta$ and $i=j$, i.e., $(\eta,j)\notin S_{\eta,j}$, then we obtain:

$$\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\notin S_{\eta,j}}\frac{1}{N_p^{\,2}}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\frac{1}{N_p^{\,2}}\mathrm{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.73}$$

Substitution of (4.70) yields:

$$\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\notin S_{\eta,j}}\frac{1}{N_p^{\,2}}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\frac{1}{N_p^{\,2}}\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mathrm{Var}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \tag{4.74}$$

If $(\eta, j)\in S_{\eta,j}$, then we obtain:

$$\begin{aligned}
&\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\frac{1}{N_p^{\,2}}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&= \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p^{\,2}}\left[\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&\overset{a}{=} \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p^{\,2}}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&\overset{b}{=} \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p^{\,2}}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]
\end{aligned} \tag{4.75}$$

where equality (a) holds because of the law of total expectation, equality (b) holds because $\tilde{Y}_{k+1}^{k,\theta,i}$ and $\tilde{Y}_{k+1}^{k,\eta,j}$ are conditional independent given $C_k^{\tilde{\xi},\tilde{\omega},K}$.

Substitution of (4.67) yields:

$$\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\frac{1}{N_p^{\,2}}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p^{\,2}}\left[\mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{K_k^{\eta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{K_k^{\eta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \quad (4.76)$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p^{\,2}}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\mathrm{Cov}\left\{K_k^{\theta,i},K_k^{\eta,j}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]$$

Combing (4.72), (4.74) and (4.76) yields:

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p^{\,2}}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{K_k^{\theta,i}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]+\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p^{\,2}}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mathrm{Var}\left\{K_k^{\theta,i}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \quad (4.77)$$
$$+\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p^{\,2}}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\mathrm{Cov}\left\{K_k^{\theta,i},K_k^{\eta,j}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]$$

For any $\theta\in\Theta$, the vector $(K_k^{\theta,1},K_k^{\theta,2},\ldots,K_k^{\theta,|J_k^\theta|})$ has a multinomial distribution with number of trials equal to N_p, which means $\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}K_k^{\theta,i}=N_P$, and with success probabilities $\tilde{\omega}_k^{\theta,i}$. Multinomial distribution properties yields

$$\mathbb{E}\left\{K_k^{\theta,i}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=N_p\tilde{\omega}_k^{\theta,i} \quad (4.78)$$

$$\mathrm{Var}\left\{K_k^{\theta,i}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=N_p\tilde{\omega}_k^{\theta,i}(1-\tilde{\omega}_k^{\theta,i}) \quad (4.79)$$

$$\mathrm{Cov}\left\{K_k^{\theta,i},K_k^{\eta,j}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\left(-N_p\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j} \quad (4.80)$$

Insertion into eqs. (4.71) and (4.77), yields eq. (4.14) of Th. 4.2, and

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\frac{1}{N_p^{\,2}}\left[N_p\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]+\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p^{\,2}}N_p\tilde{\omega}_k^{\theta,i}(1-\tilde{\omega}_k^{\theta,i})\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\right]+\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p^{\,2}}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\left(-N_p\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}\right]$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\tilde{\omega}_k^{\theta,i}\right]+\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\tilde{\omega}_k^{\theta,i}(1-\tilde{\omega}_k^{\theta,i})\right]-\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}\right]$$

Further evaluation yields:

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\tilde{\omega}_k^{\theta,i}\right]-\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\tilde{\omega}_k^{\theta,i}\right]$$
$$+\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\tilde{\omega}_k^{\theta,i}\right]-\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\left(\tilde{\omega}_k^{\theta,i}\right)^2\right] \quad (4.81)$$
$$-\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{(\eta,j)\in S_{\eta,j}}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}\right]$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\tilde{\omega}_k^{\theta,i}\right]-\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{\eta\in\Theta}\sum_{j=1}^{|J_k^\eta|}\left[\frac{1}{N_p}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}\right]$$

Rewriting the latter eq. yields eq. (4.15) in Th. 4.2. **Q.E.D.**

Appendix 4.B: Proof of Theorem 4.3

We have to proof that the variance in eq. (4.15) of Th. 4.2 (MR-normal) is larger or equal to the variance in eq. (4.13) of Th. 4.1*, i.e.

$$\frac{1}{N_p}\left[\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]\left[1-\sum_{\eta\in\underline{\Theta}}\sum_{j=1}^{|J_k^\eta|}\left(\tilde{\omega}_k^{\eta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\right)\right]\geq\frac{M_k}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right] \quad (4.82)$$

By using $\sum_{\theta\in\Theta}\tilde{\pi}_{\theta_k}(\theta)=1$, we can write:

$$\sum_{\theta\in\Theta}\left(M^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)=0$$

This can step-wise be rewritten as follows:

$$\sum_{\theta\in\underline{\Theta}}\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)=0$$

$$\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)<M_k^{-1})\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\right]+\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)\geq M_k^{-1})\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\right]=0$$

$$\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)<M_k^{-1})\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\right]=\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)\geq M_k^{-1})\left(\tilde{\pi}_{\theta_k}(\theta)-M_k^{-1}\right)\right]$$

Multiplication by $\bar{\bar{\Sigma}}_k=\sum_{\theta\in\underline{\Theta}}\tilde{\Sigma}_k^\theta/M_k=\mathbb{E}\left\{\bar{\gamma}_{k+1}\,|\,C_k^{\tilde{\xi},\tilde{\omega}}\right\}/M_k$ yields:

$$\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)<M_k^{-1})\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\bar{\bar{\Sigma}}_k\right]=\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)\geq M_k^{-1})\left(\tilde{\pi}_{\theta_k}(\theta)-M_k^{-1}\right)\bar{\bar{\Sigma}}_k\right]$$

Due to conditions (4.C1) and (4.C2), this implies:

$$\begin{aligned}&\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)<M_k^{-1})\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\tilde{\Sigma}_k^\theta\right]\geq\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)<M_k^{-1})\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\bar{\bar{\Sigma}}_k\right]\\&=\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)\geq M_k^{-1})\left(\tilde{\pi}_{\theta_k}(\theta)-M_k^{-1}\right)\bar{\bar{\Sigma}}_k\right]\geq\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)\geq M_k^{-1})\left(\tilde{\pi}_{\theta_k}(\theta)-M_k^{-1}\right)\tilde{\Sigma}_k^\theta\right]\end{aligned}$$

Hence

$$\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)<M_k^{-1})\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\tilde{\Sigma}_k^\theta\right]-\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_{\theta_k}(\theta)\geq M_k^{-1})\left(\tilde{\pi}_{\theta_k}(\theta)-M_k^{-1}\right)\tilde{\Sigma}_k^\theta\right]\geq 0$$

Stepwise evaluation yields:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(M_k^{-1}-\tilde{\pi}_{\theta_k}(\theta)\right)\tilde{\Sigma}_k^\theta\right]\geq 0$$

$$\sum_{\theta\in\underline{\Theta}}\tilde{\Sigma}_k^\theta-M_k\sum_{\theta\in\underline{\Theta}}\left[\tilde{\pi}_{\theta_k}(\theta)\tilde{\Sigma}_k^\theta\right]\geq 0 \quad (4.83)$$

From inequality of arithmetic and quadratic means we get:

$$\sum_{\theta\in\underline{\Theta}}\left(\tilde{\Sigma}_k^\theta\right)^2\geq\frac{1}{M_k}\left(\sum_{\theta\in\underline{\Theta}}\tilde{\Sigma}_k^\theta\right)^2$$

Together with (4.83) this yields:

$$\sum_{\theta\in\underline{\Theta}}\tilde{\Sigma}_k^\theta-M_k\sum_{\theta\in\underline{\Theta}}\left[\tilde{\pi}_{\theta_k}(\theta)\tilde{\Sigma}_k^\theta\right]+M_k\sum_{\theta\in\underline{\Theta}}\left(\tilde{\Sigma}_k^\theta\right)^2-\left(\sum_{\theta\in\underline{\Theta}}\tilde{\Sigma}_k^\theta\right)^2\geq 0$$

Step-wise evaluation of this inequality yields

$$\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)-M_k\sum_{\theta\in\underline{\Theta}}\left[\tilde{\pi}_{\theta_k}(\theta)\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]+M_k\sum_{\theta\in\underline{\Theta}}\left[\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]^2-\left[\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]^2\geq 0$$

$$\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)-M_k\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,j}\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]+M_k\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]$$
$$-\left[\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]\left[\sum_{\eta\in\underline{\Theta}}\sum_{j=1}^{|J_k^\eta|}\left(\tilde{\omega}_k^{\eta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\right)\right]\geq 0$$

$$\left[\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]\left[1-\sum_{\eta\in\underline{\Theta}}\sum_{j=1}^{|J_k^\eta|}\left(\tilde{\omega}_k^{\eta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\eta,j}\right)\right)\right]-M_k\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]\geq 0$$

Division by N_p, and taking into account that $\tilde{\omega}_k^{\theta,i}=0$ for $\theta\notin\underline{\Theta}$, yields inequality (4.82). **Q.E.D.**

Appendix 4.C.1: Proof of Theorem 4.6

In IPS$_{\text{mode}}$ step II at level $k+1$, $\overline{\gamma}_{k+1}$ is defined as follows:

$$\overline{\gamma}_{k+1}=\sum_{\theta\in\Theta}\sum_{j=1}^{N_k^\theta}\left[\omega_k^{\theta,j}1\left(\overline{\xi}_{k+1}^{\theta,j}\in Q_{k+1}\right)\right]$$

If $\left|J_k^\theta\right|=0$, then $\omega_k^{\theta,j}=0,\ j=1,\ldots,N_k^\theta$. Hence

$$\overline{\gamma}_{k+1}=\sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{N_k^\theta}\left[\omega_k^{\theta,j}1\left(\overline{\xi}_{k+1}^{\theta,j}\in Q_{k+1}\right)\right]\tag{4.84}$$

For $\theta\in\underline{\Theta}$ in step IV at level k, we have:

$$\left(\xi_k^{\theta,j+i\left|J_k^\theta\right|},\omega_k^{\theta,j+i\left|J_k^\theta\right|}\right)_{i=0}^{\alpha_k^\theta}=\left(\tilde{\xi}_k^{\theta,j},\frac{\tilde{\omega}_k^{\theta,j}}{(1+\rho_k^\theta)}\right),\ j=1,..,\left|J_k^\theta\right|$$
$$\xi_k^{\theta,(1+\alpha_k^\theta)\left|J_k^\theta\right|+j'}\sim\frac{\tilde{\pi}_k(\cdot,\cdot,\theta)}{\tilde{\pi}_{\theta_k}(\theta)},\omega_k^{\theta,(1+\alpha_k^\theta)\left|J_k^\theta\right|+j'}=\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|},\ \text{if } j'=1,.,(\rho_k^\theta-\alpha_k^\theta)\left|J_k^\theta\right|$$

Hence, eq. (4.84) can be written as:

$$\overline{\gamma}_{k+1}=\sum_{\theta\in\underline{\Theta}}\sum_{i=0}^{\alpha_k^\theta}\sum_{j=1}^{\left|J_k^\theta\right|}\left[\omega_k^{\theta,j+i\left|J_k^\theta\right|}1\left(\mathcal{M}(\xi_k^{\theta,j+i\left|J_k^\theta\right|})\in Q_{k+1}\right)\right]+\sum_{\theta\in\underline{\Theta}}\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)\left|J_k^\theta\right|}\left[\omega_k^{\theta,(1+\alpha_k^\theta)\left|J_k^\theta\right|+j'}1\left(\mathcal{M}(\xi_k^{\theta,(1+\alpha_k^\theta)\left|J_k^\theta\right|+j'})\in Q_{k+1}\right)\right]$$
$$=\sum_{\theta\in\underline{\Theta}}\sum_{i=0}^{\alpha_k^\theta}\sum_{j=1}^{\left|J_k^\theta\right|}\left[\frac{\tilde{\omega}_k^{\theta,j}}{1+\rho_k^\theta}1\left(\mathcal{M}(\tilde{\xi}_k^{\theta,j})\in Q_{k+1}\right)\right]+\sum_{\theta\in\underline{\Theta}}\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)\left|J_k^\theta\right|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|}1\left(\mathcal{M}(\xi_k^{\theta,(1+\alpha_k^\theta)\left|J_k^\theta\right|+j'}\in Q_{k+1}\right)\right]\tag{4.85}$$

where $\mathcal{M}(\cdot)$ represents the mutation at Step I.

Evaluation of the second term yields:

$$\sum_{\theta\in\Theta}\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|}1\left(\mathsf{M}(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'}\in Q_{k+1}\right)\right]$$

$$=\sum_{\theta\in\Theta}\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|}1\left(\mathsf{M}(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'})\in Q_{k+1}\right)\sum_{j=1}^{|J_k^\theta|}1\left(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'}=\tilde{\xi}_k^{\theta,j}\right)\right]$$

$$=\sum_{\theta\in\Theta}\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|}1\left(\mathsf{M}(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'})\in Q_{k+1}\right)1\left(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'}=\tilde{\xi}_k^{\theta,j}\right)\right]$$

$$=\sum_{\theta\in\Theta}\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|}1\left(\mathsf{M}(\tilde{\xi}_k^{\theta,j})\in Q_{k+1}\right)1\left(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'}=\tilde{\xi}_k^{\theta,j}\right)\right]$$

$$=\sum_{\theta\in\Theta}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|}1\left(\mathsf{M}(\tilde{\xi}_k^{\theta,j})\in Q_{k+1}\right)\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)|J_k^\theta|}1\left(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'}=\tilde{\xi}_k^{\theta,j}\right)\right]$$

If we let $\tilde{Y}_{k+1}^{k,\theta,j}$ with $j=1,\dots,|J_k^\theta|$ be the number of the $(\alpha_k^\theta+1)$ particle copies from $\tilde{\xi}_k^{\theta,j}$ that reach Q_{k+1} after mutation, and $\tilde{\tilde{Y}}_{k+1}^{k,\theta,j}$ with $j=1,\dots,|J_k^\theta|$ be the number of additional multinomial samples $\tilde{\tilde{K}}_k^{\theta,j}$ from $\tilde{\xi}_k^{\theta,j}$ that reach Q_{k+1} after mutation, then we have

$$\tilde{Y}_{k+1}^{k,\theta,j}=\sum_{i=0}^{\alpha_k^\theta}1\left(\mathsf{M}(\tilde{\xi}_k^{\theta,j})\in Q_{k+1}\right) \tag{4.86}$$

$$\tilde{\tilde{Y}}_{k+1}^{k,\theta,j}=1\left(\mathsf{M}(\tilde{\xi}_k^{\theta,j})\in Q_{k+1}\right)\tilde{\tilde{K}}_k^{\theta,j} \tag{4.87}$$

with

$$\tilde{\tilde{K}}_k^{\theta,j}=\sum_{j'=1}^{(\rho_k^\theta-\alpha_k^\theta)|J_k^\theta|}1\left(\xi_k^{\theta,(1+\alpha_k^\theta)|J_k^\theta|+j'}=\tilde{\xi}_k^{\theta,j}\right) \tag{4.88}$$

Eq. (4.86) shows that $\tilde{Y}_{k+1}^{k,\theta,i}$ and $\tilde{Y}_{k+1}^{k,\eta,j}$ are conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega}}$ if $\theta\neq\eta$.

Eq. (4.87) shows that $\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}$ and $\tilde{\tilde{Y}}_{k+1}^{k,\eta,j}$ are conditionally independent given $\tilde{\xi}_k^{\theta,i},K_k^{\theta,i},\tilde{\xi}_k^{\eta,j},K_k^{\eta,j}$ if $\theta=\eta$ and $i\neq j$.

Inserting (4.86) and (4.87) into (4.85) yields:

$$\overline{\gamma}_{k+1}=\sum_{\theta\in\Theta}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,j}}{1+\rho_k^\theta}\tilde{Y}_{k+1}^{k,\theta,j}\right]+\sum_{\theta\in\Theta}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|}\tilde{\tilde{Y}}_{k+1}^{k,\theta,j}\right] \tag{4.89}$$

Using eq. (4.89) yields:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \mathbb{E}\left\{\sum_{\theta\in\Theta}\sum_{j=1}^{|J_k^\theta|}\left[\left(\frac{\tilde{\omega}_k^{\theta,j}}{1+\rho_k^\theta}\right)\tilde{Y}_{k+1}^{k,\theta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$+\mathbb{E}\left\{\sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{|J_k^\theta|}\left[\left(\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|}\right)\tilde{\tilde{Y}}_{k+1}^{k,\theta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.90}$$

$$=\sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{|J_k^\theta|}\left[\left(\frac{\tilde{\omega}_k^{\theta,j}}{1+\rho_k^\theta}\right)\mathbb{E}\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\}\right]+\sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{|J_k^\theta|}\left[\left(\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|}\right)\mathbb{E}\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\}\right]$$

In a similar way, we derive:

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$=\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\mathrm{Cov}\left\{\left[\left(\frac{\tilde{\omega}_k^{\theta,i}}{1+\rho_k^\theta}\right)\tilde{Y}_{k+1}^{k,\theta,i}+\left(\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|}\right)\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}\right],\left[\left(\frac{\tilde{\omega}_k^{\eta,j}}{1+\rho_k^\eta}\right)\tilde{Y}_{k+1}^{k,\eta,j}+\left(\frac{\tilde{\pi}_{\theta_k}(\eta)}{(1+\rho_k^\eta)\left|J_k^\eta\right|}\right)\tilde{\tilde{Y}}_{k+1}^{k,\eta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$=\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\mathrm{Cov}\left\{\left[\left(\frac{\tilde{\omega}_k^{\theta,i}}{1+\rho_k^\theta}\right)\tilde{Y}_{k+1}^{k,\theta,i}\right],\left[\left(\frac{\tilde{\omega}_k^{\eta,j}}{1+\rho_k^\eta}\right)\tilde{Y}_{k+1}^{k,\eta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\mathrm{Cov}\left\{\left[\left(\frac{\tilde{\omega}_k^{\theta,i}}{1+\rho_k^\theta}\right)\tilde{Y}_{k+1}^{k,\theta,i}\right],\left[\left(\frac{\tilde{\pi}_{\theta_k}(\eta)}{(1+\rho_k^\eta)\left|J_k^\eta\right|}\right)\tilde{\tilde{Y}}_{k+1}^{k,\eta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\mathrm{Cov}\left\{\left[\left(\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|}\right)\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}\right],\left[\left(\frac{\tilde{\omega}_k^{\eta,j}}{1+\rho_k^\eta}\right)\tilde{Y}_{k+1}^{k,\eta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\mathrm{Cov}\left\{\left[\left(\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|}\right)\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}\right],\left[\left(\frac{\tilde{\pi}_{\theta_k}(\eta)}{(1+\rho_k^\eta)\left|J_k^\eta\right|}\right)\tilde{\tilde{Y}}_{k+1}^{k,\eta,j}\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$=\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\frac{\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}}{\left(1+\rho_k^\theta\right)\left(1+\rho_k^\eta\right)}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\frac{\tilde{\omega}_k^{\theta,i}\tilde{\pi}_{\theta_k}(\eta)}{\left(1+\rho_k^\theta\right)(1+\rho_k^\eta)\left|J_k^\eta\right|}\mathrm{Cov}\left\{\tilde{Y}_{k+1}^{k,\theta,i},\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\omega}_k^{\eta,j}}{(1+\rho_k^\theta)\left|J_k^\theta\right|\left(1+\rho_k^\eta\right)}\mathrm{Cov}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i},\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.91}$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\eta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|(1+\rho_k^\eta)\left|J_k^\eta\right|}\mathrm{Cov}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i},\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

Further evaluation yields:

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$

$$=\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}}{\left(1+\rho_k^\theta\right)\left(1+\rho_k^\eta\right)}\left[\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\pi}_{\theta_k}(\eta)}{\left(1+\rho_k^\theta\right)(1+\rho_k^\eta)\left|J_k^\eta\right|}\left[\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_{1:k}^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \tag{4.92}$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\omega}_k^{\eta,j}}{(1+\rho_k^\theta)\left|J_k^\theta\right|\left(1+\rho_k^\eta\right)}\left[\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]$$

$$+\sum_{\theta\in\underline{\Theta}}\sum_{\eta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\eta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|(1+\rho_k^\eta)\left|J_k^\eta\right|}\left[\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]$$

Evaluation of $\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ and $\mathrm{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ yields:

$$\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \mathbb{E}\left\{\sum_{i=0}^{\alpha_k^\theta} 1\left(M(\tilde{\xi}_k^{\theta,j}) \in Q_{k+1}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{i=0}^{\alpha_k^\theta} \mathbb{E}\left\{1\left(M(\tilde{\xi}_k^{\theta,j}) \in Q_{k+1}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \left(1+\alpha_k^\theta\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \tag{4.93a}$$

$$\mathrm{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \mathrm{Var}\left\{\sum_{i=0}^{\alpha_k^\theta} 1\left(M(\tilde{\xi}_k^{\theta,j}) \in Q_{k+1}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{i=0}^{\alpha_k^\theta} \mathrm{Var}\left\{1\left(M(\tilde{\xi}_k^{\theta,j}) \in Q_{k+1}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \left(1+\alpha_k^\theta\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.93b}$$

By using $C_k^{\tilde{\xi},\tilde{\omega},K} \triangleq \sigma\left\{\tilde{\xi}_k^{\theta,j}, \tilde{\omega}_k^{\theta,j}, K_k^{\theta,j}; j=1,\ldots,|J_k^\theta|, \theta \in \Theta\right\}$ we get:

$$\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} = \mathbb{E}\left\{1\left(M(\tilde{\xi}_k^{\theta,i}) \in Q_{k+1}\right)\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} = \tilde{\tilde{K}}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \tag{4.94}$$

$$\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \mathbb{E}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \tag{4.95}$$

$$\begin{aligned}\mathrm{Var}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},\tilde{K}}\right\} &= \mathrm{Var}\left\{1\left(M(\tilde{\xi}_k^{\theta,i}) \in Q_{k+1}\right)\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},\tilde{K}}\right\} = \tilde{\tilde{K}}_k^{\theta,i}\mathrm{Var}\left\{1\left(M(\tilde{\xi}_k^{\theta,i}) \in Q_{k+1}\right) \mid C_k^{\tilde{\xi},\tilde{\omega},\tilde{K}}\right\} \\ &= \tilde{\tilde{K}}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\end{aligned} \tag{4.96}$$

Hence

$$\begin{aligned}\mathrm{Var}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} &= \mathbb{E}\left\{\mathrm{Var}\{(\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \mathrm{Var}\left\{\mathbb{E}\{(\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}) \mid C_k^{\tilde{\xi},\tilde{\omega},K}\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\ &= \mathbb{E}\left\{\tilde{\tilde{K}}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \mathrm{Var}\left\{\tilde{\tilde{K}}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\ &= \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mathrm{Var}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\end{aligned} \tag{4.97}$$

Substituting (4.93a) and (4.95) in eq. (4.90) yields:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\Theta}\sum_{j=1}^{|J_k^\theta|}\left[\left(\frac{\tilde{\omega}_k^{\theta,j}}{1+\rho_k^\theta}\right)\left(1+\alpha_k^\theta\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right] + \sum_{\theta\in\Theta}\sum_{j=1}^{|J_k^\theta|}\left[\left(\frac{\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)|J_k^\theta|}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\mathbb{E}\left\{\tilde{\tilde{K}}_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \tag{4.98}$$

If $\theta \neq \eta$, then $\tilde{Y}_{k+1}^{k,\theta,i}$ and $\tilde{Y}_{k+1}^{k,\eta,j}$ are conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega}}$, and eq. (4.92) becomes:

$$\begin{aligned}&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\ &= \sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}}{\left(1+\rho_k^\theta\right)\left(1+\rho_k^\eta\right)}\left[\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &+ \sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\pi}_{\theta_k}(\eta)}{\left(1+\rho_k^\theta\right)(1+\rho_k^\eta)\left|J_k^\eta\right|}\left[\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &+ \sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\omega}_k^{\eta,j}}{(1+\rho_k^\theta)\left|J_k^\theta\right|\left(1+\rho_k^\eta\right)}\left[\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &+ \sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\eta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|(1+\rho_k^\eta)\left|J_k^\eta\right|}\left[\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\eta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &= 0\end{aligned} \tag{4.99}$$

If $\theta = \eta$ and $i \neq j$ then $\tilde{Y}_{k+1}^{k,\theta,i}$ and $\tilde{Y}_{k+1}^{k,\eta,j}$ are conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega},K}$. Hence eq. (4.92) becomes:

$$
\begin{aligned}
&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}}{\left(1+\rho_k^\theta\right)\left(1+\rho_k^\theta\right)}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\pi}_{\theta_k}(\theta)}{\left(1+\rho_k^\theta\right)(1+\rho_k^\theta)\left|J_k^\theta\right|}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i}\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\omega}_k^{\theta,j}}{(1+\rho_k^\theta)\left|J_k^\theta\right|\left(1+\rho_k^\theta\right)}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|(1+\rho_k^\theta)\left|J_k^\theta\right|}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i}\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]
\end{aligned}
\tag{4.100}
$$

Further evaluation yields:

$$
\begin{aligned}
&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}}{\left(1+\rho_k^\theta\right)\left(1+\rho_k^\theta\right)}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\pi}_{\theta_k}(\theta)}{\left(1+\rho_k^\theta\right)(1+\rho_k^\theta)\left|J_k^\theta\right|}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\omega}_k^{\theta,j}}{(1+\rho_k^\theta)\left|J_k^\theta\right|\left(1+\rho_k^\theta\right)}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)\tilde{\pi}_{\theta_k}(\theta)}{(1+\rho_k^\theta)\left|J_k^\theta\right|(1+\rho_k^\theta)\left|J_k^\theta\right|}\left[\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]
\end{aligned}
$$

Substitution of (4.93a) and (4.94) yields:

$$
\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2}{(1+\rho_k^\theta)^2\left|J_k^\theta\right|^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\mathrm{Cov}\left\{\tilde{\tilde{K}}_k^{\theta,i},\tilde{\tilde{K}}_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]
\tag{4.101}
$$

If $\theta=\eta$ and $i=j$ then eq. (4.92) becomes:

$$
\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{(\tilde{\omega}_k^{\theta,i})^2}{\left(1+\rho_k^\theta\right)^2}\mathrm{Var}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|^2}\mathrm{Var}\left\{\tilde{\tilde{Y}}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]
\tag{4.102}
$$

Substitution of (4.93b) and (4.97) yields:

$$
\begin{aligned}
&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{(\tilde{\omega}_k^{\theta,i})^2}{\left(1+\rho_k^\theta\right)^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\left(1+\alpha_k^\theta\right)\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mathrm{Var}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]
\end{aligned}
\tag{4.103}
$$

Combining (4.99), (4.101) and (4.103) yields:

$$
\begin{aligned}
&\mathrm{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\
&= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{(\tilde{\omega}_k^{\theta,i})^2}{\left(1+\rho_k^\theta\right)^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\left(1+\alpha_k^\theta\right)\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\mathbb{E}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mathrm{Var}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{\substack{j=1\\ j\neq i}}^{|J_k^\theta|}\left[\frac{\tilde{\pi}_{\theta_k}(\theta)^2}{(1+\rho_k^\theta)^2\left|J_k^\theta\right|^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\mathrm{Cov}\left\{\tilde{\tilde{K}}_k^{\theta,i},\tilde{\tilde{K}}_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]
\end{aligned}
\tag{4.104}
$$

For $\theta \in \underline{\Theta}$, the vector $(\tilde{\tilde{K}}_k^{\theta,1},\tilde{\tilde{K}}_k^{\theta,2},...,\tilde{\tilde{K}}_k^{\theta,|J_k^\theta|})$ has a multinomial distribution with number of trials equal to $\left(\rho_k^\theta-\alpha_k^\theta\right)\left|J_k^\theta\right|$, and with success probabilities $\tilde{\omega}_k^i/\tilde{\pi}_{\theta_k}(\theta)$. Multinomial distribution properties yields:

$$
\mathbb{E}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\left(\rho_k^\theta-\alpha_k^\theta\right)\left|J_k^\theta\right|\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)}
\tag{4.105}
$$

$$
\mathrm{Var}\left\{\tilde{\tilde{K}}_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\left(\rho_k^\theta-\alpha_k^\theta\right)\left|J_k^\theta\right|\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)}\left(1-\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)}\right)
\tag{4.106}
$$

$$
\mathrm{Cov}\left\{\tilde{\tilde{K}}_k^{\theta,i},\tilde{\tilde{K}}_k^{\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\left(\alpha_k^\theta-\rho_k^\theta\right)\left|J_k^\theta\right|\frac{\tilde{\omega}_k^{\theta,i}}{\tilde{\pi}_{\theta_k}(\theta)}\frac{\tilde{\omega}_k^{\theta,j}}{\tilde{\pi}_{\theta_k}(\theta)}
\tag{4.107}
$$

Inserting (4.105)-(4.107) into (4.98) and (4.104) and subsequent evaluation yields eqs. (4.20) and (4.21) of Th. 4.6. **Q.E.D.**

Appendix 4.C.2: Proof of Theorem 4.8

In IPS$_{\text{mode}}$ step II at level $k+1$, $\overline{\gamma}_{k+1}$ is defined as follows:

$$
\overline{\gamma}_{k+1}=\sum_{\theta\in\Theta}\sum_{j=1}^{N_k^\theta}\left[\omega_k^{\theta,j}1\left(\overline{\xi}_{k+1}^{\theta,j}\in Q_{k+1}\right)\right]
\tag{4.108}
$$

If $\left|J_k^\theta\right|=0$, then $\omega_k^{\theta,j}=0$, $j=1,\dots,N_k^\theta$. Hence

$$
\overline{\gamma}_{k+1}=\sum_{\theta\in\underline{\Theta}}\sum_{j=1}^{N_k^\theta}\left[\omega_k^{\theta,j}1\left(\overline{\xi}_{k+1}^{\theta,j}\in Q_{k+1}\right)\right]
\tag{4.109}
$$

For $\theta\in\underline{\Theta}$ in step IV at level *k*, we have:

$$
\omega_k^{\theta,j}=\sum_{i=1}^{|J_k^\theta|}\left[1\{\xi_k^{\theta,j}=\tilde{\xi}_k^{\theta,i}\}\tilde{\omega}_k^{\theta,i}/K_k^{\theta,i}\right],\ j=1,..,N_k^\theta.
$$

with the $K_k^{\theta,i}$ particle copies from $\tilde{\xi}_k^{\theta,i}$ and $K_k^{\theta,i}=\sum_{j=1}^{N_k^\theta}1\{\xi_k^{\theta,j}=\tilde{\xi}_k^{\theta,i}\}$, $i=1,..,\left|J_k^\theta\right|$.

Replacing $\omega_k^{\theta,j}$ in (4.109) and subsequent evaluation yields:

$$\overline{\gamma}_{k+1}=\sum_{\theta\in\Theta}\sum_{j=1}^{N_k^\theta}\left[\sum_{i=1}^{|J_k^\theta|}\left[1\{\xi_k^{\theta,j}=\tilde{\xi}_k^{\theta,i}\}\frac{\tilde{\omega}_k^{\theta,i}}{K_k^{\theta,i}}\right]1\left(\mathcal{M}(\xi_k^{\theta,j})\in Q_{k+1}\right)\right]$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}}{K_k^{\theta,i}}\sum_{j=1}^{N_k^\theta}\left[1\{\xi_k^{\theta,j}=\tilde{\xi}_k^{\theta,i}\}1\left(\mathcal{M}(\xi_k^{\theta,j})\in Q_{k+1}\right)\right]\right]$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}}{K_k^{\theta,i}}\sum_{j=1}^{N_k^\theta}\left[1\{\xi_k^{\theta,j}=\tilde{\xi}_k^{\theta,i}\}1\left(\mathcal{M}(\tilde{\xi}_k^{\theta,i})\in Q_{k+1}\right)\right]\right] \tag{4.110}$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}}{K_k^{\theta,i}}Y_{k+1}^{k,\theta,i}\right]$$

where $Y_{k+1}^{k,\theta,i}$ is the number of the $K_k^{\theta,i}$ particle copies from $\tilde{\xi}_k^{\theta,i}$ that reach Q_{k+1} after mutation, $\mathcal{M}(\cdot)$ represents the mutation at Step I.

Using eq. (4.110) yields:

$$\mathbb{E}\left\{\overline{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\mathbb{E}\left\{\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\frac{\tilde{\omega}_k^{\theta,i}}{K_k^{\theta,i}}Y_{k+1}^{k,\theta,i}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]$$
$$\overset{a}{=}\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\mathbb{E}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}}\mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \tag{4.111}$$
$$=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\mathbb{E}\left\{\frac{1}{K_k^{\theta,i}}\mathbb{E}\left\{Y_{k+1}^{k,\theta,i}\mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]$$

where equality (a) holds because $C_k^{\tilde{\xi},\tilde{\omega},K}\supset C_k^{\tilde{\xi},\tilde{\omega}}$ and $\mathbb{E}\{X\mid Y\}=\mathbb{E}\{\mathbb{E}\{X\mid Y,Z\}\mid Y\}$.

In a similar way, we can derive:

$$\mathrm{Var}\left\{\overline{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$$
$$=\sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\mathrm{Cov}\left\{\frac{\tilde{\omega}_k^{\theta,i}Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}},\frac{\tilde{\omega}_k^{\eta,j}Y_{k+1}^{k,\eta,j}}{K_k^{\eta,j}}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]$$
$$=\sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{i'=1}^{|J_k^\eta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}\right)\mathrm{Cov}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}},\frac{Y_{k+1}^{k,\eta,j}}{K_k^{\eta,j}}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \tag{4.112}$$
$$=\sum_{\theta\in\Theta}\sum_{\eta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}\right)\left[\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}}\frac{Y_{k+1}^{k,\eta,j}}{K_k^{\eta,j}}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\frac{Y_{k+1}^{k,\eta,j}}{K_k^{\eta,j}}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]$$

Each $Y_{k+1}^{k,\theta,i}$ has a conditional Binomial distribution with success probability $\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)$ and size $K_k^{\theta,i}$.

Let us define $C_k^{\tilde{\xi},\tilde{\omega},K}$ as follows:

$$C_k^{\tilde{\xi},\tilde{\omega},K}\triangleq\sigma\left\{\tilde{\xi}_k^{\theta,j},\tilde{\omega}_k^{\theta,j},K_k^{\theta,j};j=1,\ldots,|J_k^\theta|,\theta\in\Theta\right\}$$

Then we also get:

$$\mathbb{E}\left\{Y_{k+1}^{k,\theta,i}\mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}=K_k^{\theta,i}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \tag{4.113}$$

$$\mathbb{E}\left\{\tilde{Y}_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mathbb{E}\left\{K_k^{\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \tag{4.114}$$

$$\operatorname{Var}\left\{Y_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} = K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right] \tag{4.115}$$

Substituting (4.113) in (4.111) and subsequent evaluation yields:

$$\begin{aligned}\mathbb{E}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\mathbb{E}\left\{\frac{1}{K_k^{\theta,i}}K_k^{\theta,i}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \\ &\overset{a}{=}\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\end{aligned} \tag{4.116}$$

where equality (a) holds because $\mathbb{E}\{f(Z)\mid Z\} = f(Z)$.

If $\theta \neq \eta$, then $Y_{k+1}^{k,\theta,i}$ and $Y_{k+1}^{k,\eta,j}$ are conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega}}$, $K_k^{\theta,i}$ and $K_k^{\eta,j}$ are also conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega}}$, and eq. (4.112) becomes:

$$\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} = \sum_{\theta\in\Theta}\sum_{\substack{\eta\in\Theta\\ \eta\neq\theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\eta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\eta,j}\right)\left[\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\mathbb{E}\left\{\frac{Y_{k+1}^{k,\eta,j}}{K_k^{\eta,j}} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\frac{Y_{k+1}^{k,\eta,j}}{K_k^{\eta,j}} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] = 0 \tag{4.117}$$

If $\theta = \eta$, $i \neq j$ then $Y_{k+1}^{k,\theta,i}$ and $Y_{k+1}^{k,\eta,j}$ are conditionally independent given $C_k^{\tilde{\xi},\tilde{\omega},K}$. Hence eq. (4.112) becomes:

$$\begin{aligned}&\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} \\ &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right)\left[\mathbb{E}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}}\frac{Y_{k+1}^{k,\theta,j}}{K_k^{\theta,j}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,j}}{K_k^{\theta,j}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right)\left[\mathbb{E}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,j}}{K_k^{\theta,j}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,j}}{K_k^{\theta,j}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right)\bullet\left[\mathbb{E}\left\{\frac{1}{K_k^{\theta,i}K_k^{\theta,i'}}\mathbb{E}\left\{Y_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\}\mathbb{E}\left\{Y_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\frac{1}{K_k^{\theta,i}}\mathbb{E}\left\{Y_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\frac{1}{K_k^{\theta,j}}\mathbb{E}\left\{Y_{k+1}^{k,\theta,j} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right)\bullet\left[\mathbb{E}\left\{\frac{1}{K_k^{\theta,i}K_k^{\theta,j}}K_k^{\theta,i}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)K_k^{\theta,j}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} - \mathbb{E}\left\{\frac{1}{K_k^{\theta,i}}K_k^{\theta,i}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\cdot\mathbb{E}\left\{\frac{1}{K_k^{\theta,j}}K_k^{\theta,j}\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right) \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\right)\left[\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right) - \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\cdot\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right] = 0\end{aligned}$$

(4.118)

If $\theta = \eta$, $i = j$ then eq. (4.112) becomes:

$$\begin{aligned}&\operatorname{Var}\left\{\overline{\gamma}_{k+1} \mid C_{\tilde{\xi},\tilde{\omega}}^k\right\} \\ &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\left[\operatorname{Var}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right] \\ &\overset{a}{=}\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\mathbb{E}\left\{\operatorname{Var}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] + \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\operatorname{Var}\left\{\mathbb{E}\left\{\frac{Y_{k+1}^{k,\theta,i}}{K_k^{\theta,i}} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right] \\ &= \sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\left[\mathbb{E}\left\{\left(\frac{1}{K_k^{\theta,i}}\right)^2\operatorname{Var}\left\{Y_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\} + \operatorname{Var}\left\{\frac{1}{K_k^{\theta,i}}\mathbb{E}\left\{Y_{k+1}^{k,\theta,i} \mid C_k^{\tilde{\xi},\tilde{\omega},K}\right\} \mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]\end{aligned}$$

where equality (a) is due to the law of total variance.

Substitution of (4.113)-(4.115) yields:

$$\begin{aligned}&\mathrm{Var}\left\{\bar{\gamma}_{k+1} \mid C^{k}_{\tilde{\xi},\tilde{\omega}}\right\}\\&=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\left[\mathbb{E}\left\{\left(\frac{1}{K_k^{\theta,i}}\right)^2 K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left(1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}+\mathrm{Var}\left\{\frac{1}{K_k^{\theta,i}}K_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]\end{aligned}$$

Further evaluation yields:

$$\begin{aligned}&\mathrm{Var}\left\{\bar{\gamma}_{k+1} \mid C^{k}_{\tilde{\xi},\tilde{\omega}}\right\}\\&=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\left[\mathbb{E}\left\{\left(\frac{1}{K_k^{\theta,i}}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left(1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}+\mathbb{E}\left\{\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}-\left[\mathbb{E}\left\{\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]^2\right]\right]\\&=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left(1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\left[\mathbb{E}\left\{\frac{1}{K_k^{\theta,i}}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right]\right]\end{aligned}\tag{4.119}$$

For FASmode splitting, we have:

$$P\{K_k^{\theta,i}=\alpha_k^\theta+1\}=1-\rho_k^\theta+\alpha_k^\theta \tag{4.120a}$$

$$P\{K_k^{\theta,i}=\alpha_k^\theta+2\}=\rho_k^\theta-\alpha_k^\theta \tag{4.120b}$$

By using (4.120a,b), we derive:

$$\begin{aligned}&\mathbb{E}\left\{\frac{1}{K_k^{\theta,i}}\mid \tilde{\xi}_k^{\theta,j},\tilde{\omega}_k^{\theta,j};\theta\in\Theta, j=1,\dots,|J_k^\theta|\right\}\\&=\frac{1}{\alpha_k^\theta+1}P\{K_k^{\theta,i}=\alpha_k^\theta+1\}+\frac{1}{\alpha_k^\theta+2}P\{K_k^{\theta,i}=\alpha_k^\theta+2\}\\&=\frac{\left(1-\rho_k^\theta+\alpha_k^\theta\right)}{\alpha_k^\theta+1}+\frac{\left(\rho_k^\theta-\alpha_k^\theta\right)}{\alpha_k^\theta+2}\\&=\frac{\left(1-\rho_k^\theta+\alpha_k^\theta\right)\left(\alpha_k^\theta+2\right)}{\left(\alpha_k^\theta+1\right)\left(\alpha_k^\theta+2\right)}+\frac{\left(\rho_k^\theta-\alpha_k^\theta\right)\left(\alpha_k^\theta+1\right)}{\left(\alpha_k^\theta+2\right)\left(\alpha_k^\theta+1\right)}\\&=\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(\alpha_k^\theta+1\right)\left(\alpha_k^\theta+2\right)}\end{aligned}\tag{4.121}$$

Inserting (4.121) into (4.119) yields:

$$\mathrm{Var}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}=\sum_{\theta\in\Theta}\sum_{i=1}^{|J_k^\theta|}\left[\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left(1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(\alpha_k^\theta+1\right)\left(\alpha_k^\theta+2\right)}\right]$$

Q.E.D.

Appendix 4.D.1: Proof of Theorem 4.9

We denote $\mathrm{Var}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}$ under IPSmode-RMSmode and IPSmode-MSmode by $\mathrm{V}_{k+1}^{\text{IPSmode-RMSmode}}$ and $\mathrm{V}_{k+1}^{\text{IPSmode-MSmode}}$. Hence we have to prove $\mathrm{V}_{k+1}^{\text{IPSmode-MSmode}}-\mathrm{V}_{k+1}^{\text{IPSmode-RMSmode}}\geq 0.$ From Theorem 4.6 and Corollary 4.7 we get:

$$\mathrm{V}_{k+1}^{\text{IPSmode-MSmode}} - \mathrm{V}_{k+1}^{\text{IPSmode-RMSmode}}$$

$$= -\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\alpha_k^\theta}{\left(1+\rho_k^\theta\right)^2}\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right] + \sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\alpha_k^\theta\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right] \tag{4.122}$$

Evaluation of Δ_k^θ as defined by eq. (4.26) yields:

$$\sum_i^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\right]+\Delta_k^\theta\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}\le\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\right]-\Delta_k^\theta\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}>\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\right]=0$$

Hence

$$\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}\le\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)(1+\Delta_k^\theta)\right]+\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}>\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)(1-\Delta_k^\theta)\right]=0$$

Hence

$$\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}\le\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)(1+\Delta_k^\theta)\right]=\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}>\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\tilde{\omega}_k^{\theta,i}-\bar{\tilde{\omega}}_k^\theta\right)(1-\Delta_k^\theta)\right]$$

Multiplication by $\tilde{\Sigma}_k^\theta / \tilde{\pi}_{\theta_k}(\theta)$ yields:

$$\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}\le\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\left(\frac{\tilde{\Sigma}_k^\theta}{\tilde{\pi}_{\theta_k}(\theta)}(1+\Delta_k^\theta)\right)\right]=\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}>\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\tilde{\omega}_k^{\theta,i}-\bar{\tilde{\omega}}_k^\theta\right)\left(\frac{\tilde{\Sigma}_k^\theta}{\tilde{\pi}_{\theta_k}(\theta)}(1-\Delta_k^\theta)\right)\right]$$

By using conditions (4.C3) and (4.C4) this yields:

$$\begin{aligned}&\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}\le\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\ge\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}\le\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\left(\frac{\tilde{\Sigma}_k^\theta}{\tilde{\pi}_{\theta_k}(\theta)}(1+\Delta_k^\theta)\right)\right]\\&=\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}>\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\tilde{\omega}_k^{\theta,i}-\bar{\tilde{\omega}}_k^\theta\right)\left(\frac{\tilde{\Sigma}_k^\theta}{\tilde{\pi}_{\theta_k}(\theta)}(1-\Delta_k^\theta)\right)\right]\\&\ge\sum_i^{|J_k^\theta|}\left[1\{\tilde{\omega}_k^{\theta,i}>\bar{\tilde{\omega}}_k^\theta\}\tilde{\omega}_k^{\theta,i}\left(\tilde{\omega}_k^{\theta,i}-\bar{\tilde{\omega}}_k^\theta\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\end{aligned}$$

Rewriting yields:

$$\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\ge 0 \tag{4.123}$$

Inequality of Arithmetic and Quadratic Means yields:

$$\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]^2\ge\frac{1}{\left|J_k^\theta\right|}\left[\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]^2$$

Combining this with inequality (4.123) yields:

$$\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\tilde{\omega}_k^{\theta,i}\left(\bar{\tilde{\omega}}_k^\theta-\tilde{\omega}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]-\left[\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]^2+\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\right]\ge 0$$

Rewriting yields:

$$\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\tilde{\omega}_k^{\theta,i}\bar{\tilde{\omega}}_k^\theta\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]-\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]$$
$$-\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]+\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\right]\geq 0$$

Further rewriting yields:

$$\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]-\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]$$
$$-\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]+\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)^2\right]\geq 0$$

Hence:

$$\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]-\sum_{i=1}^{|J_k^\theta|}\left[\left|J_k^\theta\right|\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]\geq 0 \quad (4.124)$$

Multiplication by $\dfrac{\alpha_k^\theta}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|}$ and subsequent summation over $\theta\in\underline{\Theta}$ yields:

$$\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\alpha_k^\theta\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]-\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{\alpha_k^\theta}{\left(1+\rho_k^\theta\right)^2}\left(\tilde{\omega}_k^{\theta,i}\right)^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]\geq 0$$

(4.125)

Q.E.D.

Appendix 4.D.2: Proof of Theorem 4.10

We have to proof that the variance in eq. (4.13) of Th. 4.1* is larger or equal to the variance in eq. (4.23) of Corollary 4.7, i.e.

$$\begin{aligned}
&\mathrm{V}_{k+1}^{\text{IPSmode-MRmode}}\\
&=\frac{M_k}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]\\
&\geq\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{(\tilde{\omega}_k^{\theta,i})^2}{\left(1+\rho_k^\theta\right)^2}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\rho_k^\theta}{(1+\rho_k^\theta)^2\left|J_k^\theta\right|}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]\\
&=\mathrm{V}_{k+1}^{\text{IPSmode-MSmode}}
\end{aligned} \quad (4.126)$$

For $\tilde{\Sigma}_k^\theta$ defined in Theorem 4.3 applies:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]=\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta<\frac{1}{M_k})\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]+\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta\geq\frac{1}{M_k})\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]$$

Rewriting yields:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-\bar{N}_k}\right)\tilde{\Sigma}_k^\theta\right]-\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta<\frac{1}{M_k})\left(\frac{N_k-M_k\left|J_k^\theta\right|}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]=\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta\geq\frac{1}{M_k})\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]$$

Thanks to conditions (4.C5) and (4.C6) this yields:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]-\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta<\frac{1}{M_k})\left(\frac{N_k-M_kN_k/M}{M_k\left|J_k^\theta\right|+N_p-\bar{N}_k}\right)\tilde{\Sigma}_k^\theta\right]\geq\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta\geq\frac{1}{M_k})\left(\frac{M_kN_k/M_k-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]$$

Evaluation yields:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\tilde{\Sigma}_k^\theta\right]\geq 0$$

Multiplication by $\left(1-\mathbb{E}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right)/M_k$ yields:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\frac{1}{M_k}\left(1-\mathbb{E}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}\right)\tilde{\Sigma}_k^\theta\right]\geq 0$$

Decomposition yields:

$$\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta<\frac{1}{M_k})\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\frac{1}{M_k}-\frac{\mathbb{E}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}}{M_k}\right)\tilde{\Sigma}_k^\theta\right]+\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta\geq\frac{1}{M_k})\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\frac{1}{M_k}-\frac{\mathbb{E}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}}{M_k}\right)\tilde{\Sigma}_k^\theta\right]\geq 0$$

Rewriting yields:

$$\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta\geq\frac{1}{M_k})\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\frac{1}{M_k}-\frac{\mathbb{E}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}}{M_k}\right)\tilde{\Sigma}_k^\theta\right]\geq\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta<\frac{1}{M_k})\left(\frac{N_k-M_k\left|J_k^\theta\right|}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\frac{1}{M_k}-\frac{\mathbb{E}\left\{\bar{\gamma}_{k+1}\mid C_k^{\tilde{\xi},\tilde{\omega}}\right\}}{M_k}\right)\tilde{\Sigma}_k^\theta\right]$$

Thanks to (4.C1) and (4.C2) this yields:

$$\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta\geq\frac{1}{M_k})\left(\frac{M_k\left|J_k^\theta\right|-N_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\tilde{\pi}_k^\theta-\tilde{\Sigma}_k^\theta\right)\tilde{\Sigma}_k^\theta\right]\geq\sum_{\theta\in\underline{\Theta}}\left[1(\tilde{\pi}_k^\theta<\frac{1}{M_k})\left(\frac{N_k-M_k\left|J_k^\theta\right|}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\tilde{\pi}_k^\theta-\tilde{\Sigma}_k^\theta\right)\tilde{\Sigma}_k^\theta\right]$$

Hence:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M\left|J_k^\theta\right|-\bar{N}_k}{M\left|J_k^\theta\right|+N_p-\bar{N}_k}\right)\left(\tilde{\pi}_k^\theta-\tilde{\Sigma}_k^\theta\right)\tilde{\Sigma}_k^\theta\right]>0$$

$$\sum_{\theta\in\underline{\Theta}}\left[\left(1-\frac{N_p}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\tilde{\pi}_{\theta_k}(\theta)-\tilde{\Sigma}_k^\theta\right)\tilde{\Sigma}_k^\theta\right]\geq 0$$

Multiplication by M_k/N_p yields:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k}{N_p}-\frac{M_k}{M_k\left|J_k^\theta\right|+N_p-N_k}\right)\left(\tilde{\pi}_{\theta_k}(\theta)-\tilde{\Sigma}_k^\theta\right)\tilde{\Sigma}_k^\theta\right]\geq 0$$

Substitution of $\dfrac{M_k \left|J_k^\theta\right| + N_p - N_k}{M_k \left|J_k^\theta\right|} = \rho_k^\theta + 1$ yields:

$$\sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \left(\tilde{\pi}_{\theta_k}(\theta) - \tilde{\Sigma}_k^\theta \right) \tilde{\Sigma}_k^\theta \right] \geq 0$$

Stepwise further evaluation yields:

$$\sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \tilde{\pi}_{\theta_k}(\theta) \tilde{\Sigma}_k^\theta \right] - \sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \left(\tilde{\Sigma}_k^\theta \right)^2 \right] \geq 0$$

$$\sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \tilde{\pi}_{\theta_k}(\theta) \sum_{i=1}^{|J_k^\theta|} \left[\left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right] \right] - \sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \left[\sum_{i=1}^{|J_k^\theta|} \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right]^2 \right] \geq 0$$

$$\sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\pi}_{\theta_k}(\theta) \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right] \right] - \sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \left[\sum_{i=1}^{|J_k^\theta|} \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right]^2 \right] \geq 0$$

$$\sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1+\rho_k^\theta}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|} \right) \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\pi}_{\theta_k}(\theta) \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right] \right] - \sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{1}{\left(1+\rho_k^\theta\right)\left|J_k^\theta\right|} \right) \left[\sum_{i=1}^{|J_k^\theta|} \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right]^2 \right] \geq 0$$

$$\begin{aligned} &\sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{\rho_k^\theta}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|} \right) \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\pi}_{\theta_k}(\theta) \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right] \right] - \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \left[\frac{1}{\left(1+\rho_k^\theta\right)^2} \bar{\tilde{\omega}}_k^\theta \tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right] \\ &- \sum_{\theta\in\underline{\Theta}} \left[\left(\frac{M_k}{N_p} - \frac{\rho_k^\theta}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|} \right) \left[\sum_{i=1}^{|J_k^\theta|} \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right]^2 \right] + \sum_{\theta\in\underline{\Theta}} \sum_{i=1}^{|J_k^\theta|} \left[\frac{1}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|} \left[\sum_{i=1}^{|J_k^\theta|} \left(\tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right]^2 \right] \geq 0 \end{aligned}$$

In the proof of Th. 4.9 (RMSmode vs. MSmode), the following has shown to hold true under conditions (4.C3) and (4.C4):

$$\sum_{i=1}^{|J_k^\theta|} \left[\bar{\tilde{\omega}}_k^\theta \tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right] \geq \sum_{i=1}^{|J_k^\theta|} \left[\tilde{\omega}_k^{\theta,i} \tilde{\omega}_k^{\theta,i} \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right]$$

From inequality of arithmetic and quadratic means:

$$\sum_{i=1}^{|J_k^\theta|} \left((\tilde{\omega}_k^{\theta,i}) \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right)^2 \geq \frac{1}{\left|J_k^\theta\right|} \left[\sum_{i=1}^{|J_k^\theta|} \left((\tilde{\omega}_k^{\theta,i}) \gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i} \right) \right) \right]^2$$

These two, together with the previous inequality, yields:

$$\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k}{N_p}-\frac{\rho_k^\theta}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|}\right)\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\pi}_{\theta_k}(\theta)\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]\right]$$
$$-\sum_{\theta\in\underline{\Theta}}\left[\left(\frac{M_k}{N_p}-\frac{\rho_k^\theta}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|}\right)\left[\sum_{i=1}^{|J_k^\theta|}\left(\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right)\right]^2\right]$$
$$-\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{\left(1+\rho_k^\theta\right)^2}\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]+\sum_{\theta\in\underline{\Theta}}\left[\frac{1}{\left(1+\rho_k^\theta\right)^2}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]^2\right]$$
$$\geq 0$$

Evaluation yields:

$$\frac{M_k}{N_p}\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]$$
$$-\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\rho_k^\theta}{\left(1+\rho_k^\theta\right)^2\left|J_k^\theta\right|}\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]$$
$$-\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{\left(1+\rho_k^\theta\right)^2}(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]$$
$$\geq 0$$

This confirms inequality (4.126). **Q.E.D.**

Appendix 4.D.3. Proof of Theorem 4.10*

To compare the variances we denote $\mathrm{Var}\{\overline{\gamma}_{k+1} \mid C_k^{\tilde{\xi},\tilde{\omega}}\}$ under MSmode and FASmode by $\mathrm{V}_{k+1}^{\text{IPSmode-MSmode}}$ and $\mathrm{V}_{k+1}^{\text{IPSmode-FASmode}}$ respectively. Then from Corollary 4.7 and Theorem 4.8 we get:

$$\begin{aligned}
&\mathrm{V}_{k+1}^{\text{IPSmode-MSmode}}-\mathrm{V}_{k+1}^{\text{IPSmode-FASmode}}\\
&=\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{1}{\left(1+\rho_k^\theta\right)^2}(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]\\
&+\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\frac{\rho_k^\theta}{(1+\rho_k^\theta)^2\left|J_k^\theta\right|}\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]\\
&-\sum_{\theta\in\underline{\Theta}}\sum_{i=1}^{|J_k^\theta|}\left[\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)}(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]
\end{aligned} \tag{4.127}$$

Hence, to show $\mathrm{V}_{k+1}^{MS\,\mathrm{mode}} \leq \mathrm{V}_{k+1}^{FAS\,\mathrm{mode}}$ we have to show:

$$\begin{aligned}&\frac{1}{\left(1+\rho_k^\theta\right)^2}\sum_{i=1}^{|J_k^\theta|}\left[(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]\\&+\frac{\rho_k^\theta}{(1+\rho_k^\theta)^2\left|J_k^\theta\right|}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]\\&-\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)}\sum_{i=1}^{|J_k^\theta|}\left[(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]\le 0\end{aligned}\tag{4.128}$$

Evaluation of the double summation term yields:

$$\begin{aligned}&\frac{1}{\left|J_k^\theta\right|}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]\\&=\frac{1}{\left|J_k^\theta\right|}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]-\frac{1}{\left|J_k^\theta\right|}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\\&=\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\overline{\tilde{\omega}}_k^{\theta}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]-\frac{1}{\left|J_k^\theta\right|}\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\\&=\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\overline{\tilde{\omega}}_k^{\theta}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]-\frac{1}{\left|J_k^\theta\right|}\left[\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]^2\end{aligned}\tag{4.129}$$

Inequality of Arithmetic and Quadratic Means yields:

$$\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]^2\ge\frac{1}{\left|J_k^\theta\right|}\left[\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]^2$$

From inequality (4.123) in the proof of Theorem 4.9 we know that due to conditions (4.C3) and (4.C4):

$$\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\overline{\tilde{\omega}}_k^{\theta}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\ge\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]$$

Substitution of the last two inequalities in (4.129) yields:

$$\begin{aligned}\frac{1}{\left|J_k^\theta\right|}\sum_{i=1}^{|J_k^\theta|}\sum_{j=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\tilde{\omega}_k^{\theta,j}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,j}\right)\right]\right]&\ge\sum_{i=1}^{|J_k^\theta|}\left[(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]-\sum_{i=1}^{|J_k^\theta|}\left[\tilde{\omega}_k^{\theta,i}\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]^2\\&=\sum_{i=1}^{|J_k^\theta|}\left[(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]\end{aligned}$$

Substitution in (4.128) yields sufficient condition:

$$\frac{1}{\left(1+\rho_k^\theta\right)^2}\sum_{i=1}^{|J_k^\theta|}\left[(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]$$
$$+\frac{\rho_k^\theta}{(1+\rho_k^\theta)^2}\sum_{i=1}^{|J_k^\theta|}\left[(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]$$
$$-\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)}\sum_{i=1}^{|J_k^\theta|}\left[(\tilde{\omega}_k^{\theta,i})^2\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\left[1-\gamma_{k+1}\left(\tilde{\xi}_k^{\theta,i}\right)\right]\right]\leq 0$$

This yields as sufficient condition:

$$\frac{1}{\left(1+\rho_k^\theta\right)^2}+\frac{\rho_k^\theta}{(1+\rho_k^\theta)^2}-\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)}\leq 0$$

Evaluation yields:

$$\frac{\left(1+\rho_k^\theta\right)}{\left(1+\rho_k^\theta\right)^2}-\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)}=\frac{1}{\left(1+\rho_k^\theta\right)}-\frac{2-\rho_k^\theta+2\alpha_k^\theta}{\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)}\leq 0$$

Hence we have to show:

$$\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)-(1+\rho_k^\theta)\left(2-\rho_k^\theta+2\alpha_k^\theta\right)\leq 0$$

Evaluation yields:

$$\begin{aligned}&\left(2+\alpha_k^\theta\right)\left(1+\alpha_k^\theta\right)-(1+\rho_k^\theta)\left(2-\rho_k^\theta+2\alpha_k^\theta\right)\\&=2+2\alpha_k^\theta+\alpha_k^\theta+(\alpha_k^\theta)^2-2+\rho_k^\theta-2\alpha_k^\theta-2\rho_k^\theta+(\rho_k^\theta)^2-2\alpha_k^\theta\rho_k^\theta\\&=\alpha_k^\theta+(\alpha_k^\theta)^2-2\alpha_k^\theta\rho_k^\theta-\rho_k^\theta+(\rho_k^\theta)^2\\&=\left(\rho_k^\theta-\alpha_k^\theta\right)^2-\left(\rho_k^\theta-\alpha_k^\theta\right)\leq 0\end{aligned}$$

This confirms that inequality (4.128) holds true. **Q.E.D.**

5

Importance Sampling in Rare Event Estimation for General Stochastic Hybrid Systems

In the field of rare event estimation of continuous time stochastic processes, the use of Importance Sampling (IS) within statistical simulation has been well studied for Continuous Time Markov Chains (CTMC) and for diffusions. These studies address three main issues. The first issue is to characterize the optimal IS strategy. By the very nature of optimal IS, this strategy cannot be used in practice. Hence the second issue is to use the characterization of the optimal IS strategy for the development of a parametric family of approximated IS strategies. The third issue is to optimize the parameter values in this family through a minimization of the Kullback-Leibler divergence between the probability laws of the optimal and the approximated IS strategies. These three issues have been well studied for continuous time Markov chains (CTMC) as well as for Diffusion processes. More recently, these three steps have been addressed for a Piecewise Deterministic Markov Process (PDMP), which is a general class of hybrid stochastic processes, though without diffusion. This chapter develops an extension of these IS results to a PDMP that is enriched with diffusion, which is studied as a pathwise unique solution of a General Stochastic Hybrid System (GSHS). This IS extension is illustrated to work well for IS based statistical simulation of an GSHS example that has multiple subsystems in parallel redundancy, each of which is subject to failure and repair.

This chapter has been submitted to Methodology in Computing and Applied Probability, as H. Ma and H.A.P. Blom, Importance Sampling in Estimation of Reach Probability of General Stochastic Hybrid Systems.

5.1. Introduction

A Piecewise Deterministic Markov Process (PDMP) is defined by Davis (1984) as a continuous-time hybrid state Markov process that involves a large variety of stochastic behaviours, except Brownian motion. A PDMP involves two dynamically interacting processes, a discrete-valued mode process $\{\theta_t\}$ and an Euclidean-valued process $\{x_t\}$. A PDMP generates an increasing series of stopping times $\{s_j; j=0,..\}$, and in between two consecutive stopping times the mode $\{\theta_t\}$ does not change, while the process $\{x_t\}$ evolves according to a mode-dependent flow. [Bujorianu and Lygeros, 2006] have defined a General Stochastic Hybrid Systems (GSHS) as an extension of a PDMP by replacing the mode-dependent deterministic flow by a solution of a mode-dependent Stochastic Differential Equation (SDE) which is driven by Brownian motion. This chapter studies rare event estimation for a GSHS using Importance Sampling (IS) as a variance reduction approach.

As has been explained well by Glasserman (2004, p. 277), IS is a variance reduction approach that has the highest potential in rare event estimation, though also is the most complex. The IS idea is to modify the probability law of the process considered, such that the reach probability of the rare event increases. To compensate the increased reach probability value, it has to be multiplied by the likelihood ratio of the rare event to happen under the original process relative to the modified process. Studies of IS for rare event estimation commonly address three main issues. The first issue is to characterize the optimal IS strategy. Because this characterization involves the rare event probability to be estimated, it is of theoretical use only. Hence the second issue is to use the characterization of the optimal IS strategy for the development of a parametric family of approximated IS strategies. The third issue is to optimize the parameter values in this family through a minimization of the Kullback-Leibler divergence between the probability laws of the optimal and the approximated IS strategies. In literature, these three IS issues have mainly be studied for three classes of continuous-time stochastic processes: i) Continuous Time Markov Chains (CTMCs); ii) Diffusions; and iii) Piecewise Deterministic Markov Processes (PDMPs).

IS of CTMC is studied for highly dependable systems, in which multiple subsystems may be subject to failures as well as repairs. In literature, IS studies of such CTMC's typically reduce the problem to IS of the underlying Discrete Time Markov Chain (DTMC), e.g. [Goyal et al. 1992; Shahabuddin, 1994; Heidelberger, 1995; Papadopoulos and Limnios, 2002; Nakayama and Shahabuddin, 2004; Juneja and Shahabuddin, 2006; L'Ecuyer et al., 2010; L'Ecuyer and Tuffin, 2011; Reijsbergen et al., 2012]. A consequence of this approach is that in these studies, IS only modifies the transition probability matrix of the underlying DTMC, though not the rate of leaving the current mode.

IS of Diffusions is well studied in various domains, ranging from finance, e.g. [Glasserman et al., 1999; Glasserman, 2004] to computational physics, e.g. [Dupuis et al., 2012; Zhang et al., 2014]. The IS modification concerns the drift coefficient of a diffusion, in a direction that involves Brownian motion.

Recently, [Chraibi et al., 2019] have studied IS for application to safety and reliability assessment of complex industrial systems. Such systems typically involve Euclidean valued process components (e.g. temperature of a liquid in a tank), the evolution of which satisfies an ordinary differential equation, the coefficients of which depend on multiple subsystems that are subject to failure and repair. As has been well explained by these authors, these complex industrial systems can be modelled well as a PDMP. [Chraibi et al, 2019] have developed solutions for each of the three IS issues. A key novelty is that their optimal and approximate IS strategies do not only modify the transition probabilities but also the rate of leaving the current mode.

For rare event simulation of advanced air traffic designs, these IS developments for Diffusions, CTMC's and PDMP's fall short. The reason is that air traffic involves uncertainty from wind as well as dependency on technical systems that are subject to failure [Blom et al., 2007]. These stochastic effects are well captured by a GSHS, i.e. a PDMP hat is enriched with diffusion. The aim of this chapter is to extend the IS developments by [Chraibi et al., 2019] to GSHS.

This chapter is organized as follows. Section 2 presents background of General Stochastic Hybrid System (GSHS). Section 3 studies reach probability estimation under IS for a GSHS and derives an optimal IS characterization. Section 4 develops an IS approximation strategy for GSHS with failing subsystems in parallel redundancy. Section 5 presents simulation results for this IS approximation strategy for a simple GSHS example. Section 6 draws conclusions.

5.2. Rare event estimation for General Stochastic Hybrid System

Throughout this and the following sections, all stochastic processes are defined on a complete stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P, \mathcal{T})$ with $(\Omega, \mathcal{F}, P)$ being a complete probability space and $\mathbb{F}$ an increasing sequence of sub-σ-algebras on the time line $\mathcal{T} = \mathbb{R}_+$, i.e., $\mathbb{F} \triangleq \{\mathcal{J}, (\mathcal{F}_t, t \in \mathbb{R}_+), \mathcal{F}\}$, with $\mathcal{J}$ containing all P-null sets of $\mathcal{F}$ and $\mathcal{J} \subset \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for every $s < t$.

5.2.1 GSHS definition

(Bujorianu and Lygeros, 2006) formalized the concept of GSHS or general stochastic hybrid automata as follows:

Definition 5.2.1 (GSHS). A GSHS is a collection $(\Theta, d, X, f, g, Init, \lambda, R)$ where

- Θ is a countable set of discrete-valued variables;
- $d : \Theta \to \mathbb{N}$ is a map giving the dimensions of the continuous state spaces;
- $X : \Theta \to \mathbb{R}^{d(\cdot)}$ maps each $\theta \in \Theta$ into an open subset X^θ of $\mathbb{R}^{d(\theta)}$;
- $f : \Xi \to \mathbb{R}^{d(\cdot)}$ is a vector field, where $\Xi \triangleq \bigcup_{\theta \in \Theta} \{\theta\} \times X^\theta$;
- $g : \Xi \to \mathbb{R}^{d(\cdot) \times m_{\text{dim}}}$ is an $X^{(\cdot)}$-valued matrix, $m_{\text{dim}} \in \mathbb{N}$;
- $\text{Init} : \beta(\Xi) \to [0,1]$ an initial probability measure on Ξ;
- $\lambda : \bar{\Xi} \to \mathbb{R}^+$ is a transition rate function;
- $R : \bar{\Xi} \times \beta(\bar{\Xi}) \to [0,1]$ is a transition measure.

Without loss of generality, transition measure $R : \bar{\Xi} \times \beta(\bar{\Xi}) \to [0,1]$ is assumed to have a transiton density $Q_{\theta,x}(\eta, y)$, $(\theta, x, \eta, y) \in \bar{\Xi} \times \Xi$, such that for $B \in \beta(\bar{X}^\theta)$:

$$R_{\theta,x}(\eta \times B) = \int_B Q_{\theta,x}(\eta, y) d\varpi(y)$$

where $\varpi(.)$ is Lebesgue measure.

5.2.2 GSHS execution

Definition 5.2.2 (GSHS Execution). A stochastic process $\{\theta_t, x_t\}$ is called a solution of GSHS execution if there exists a sequence of stopping times $s_0 = 0 < s_1 < s_2 < \dots$ such that:

- (θ_0, x_0) is a Ξ-valued random variable satisfying the probability measure *Init*;
- For $t \in \left[s_{j-1}, s_j\right)$, $j \geq 1$, $\{\theta_t, x_t\}$ is a solution of the SDE:

$$\begin{aligned} d\theta_t &= 0 \\ dx_t &= f(\theta_t, x_t)dt + g(\theta_t, x_t)dw_t \end{aligned} \tag{5.2.1}$$

  with $\{w_t\}$ m-dimensional standard Brownian motion;
- s_j is the minimum of the following two stopping times: i) first hitting time $> s_{j-1}$ of the boundary of $X^{\theta_{s_{j-1}}}$ by the phase process $\{x_t\}$; and ii) first moment $> s_{j-1}$ of a transition event to happen at rate $\lambda(\theta_t, x_t)$.
- At stopping time s_j the hybrid state $\{\theta_{s_j}, x_{s_j}\}$ satisfies the conditional probability measure $p_{\theta_{s_j}, x_{s_j} | \theta_{s_j-}, x_{s_j-}}(\eta, A \mid \theta, x) = R_{\theta,x}(\{\eta\} \times A),\ A \in \beta(X^\eta)$.

In order to assure that a GSHS execution has a solution the following assumptions are adopted:

A1 (non-Zeno property): $E\{s_j - s_{j-1}\} > 0$, P-a.s.

A2: For each $(\theta_0, x_0) \in \Xi$, equation (1) has a pathwise unique solution on a finite time interval $[0, T]$.

A3 λ is measurable and finite valued.

A4 $Init(\Xi) = 1$, and $R_{\theta,x}(\Xi) = 1$ for each $(\theta, x) \in \overline{\Xi}$.

Bujorianu and Lygeros (2006) show that the stochastic process $\{\theta_t, x_t\}$ generated by execution of a GSHS satisfies the strong Markov property.

5.2.3 Transition rates between mode values

Whereas $\lambda(\theta, x)$ specifies the overall jump rate if $\theta_{t-} = \theta$ and $X_{t-} = x$, we can also define for each mode transition the rate $\lambda_{\theta\eta}(x)$ as follows:

$$\lambda_{\theta\eta}(x) = \lambda(\theta, x) R_{\theta,x}(\{\eta\} \times X^\eta). \tag{5.2.2}$$

This implies:

$$\sum_{\eta \in \Theta} \lambda_{\theta\eta}(x) = \sum_{\eta \in \Theta} \left[\lambda(\theta, x) R_{\theta,x}(\{\eta\} \times X^\eta) \right] = \lambda(\theta, x) \tag{5.2.3}$$

Remark 5.2.3: If $R_{\theta,x}(\{\theta\} \times (X^\theta / \{x\})) \neq 0$, then $\lambda_{\theta\theta}(x) \neq 0$. In this case, at arbitrary GSHS jump time s_j there may be a jump in $\{x_t\}$ only, i.e. $\theta_{s_j} = \theta_{s_j-}$ and $x_{s_j} \neq x_{s_j-}$.

Remark 5.2.4: If $R_{\theta,x}(\{\theta\} \times (X^\theta / \{x\})) = 0$, then $\lambda_{\theta\theta}(x) = 0$ and $R_{\theta,x}(\Theta \times \{x\}) = 1$. In this case, at arbitrary GSHS jump time s_j there may be a jump in $\{\theta_t\}$ only, i.e. $\theta_{s_j} \neq \theta_{s_j-}$ and $x_{s_j} = x_{s_j-}$. Then the process $\{\theta_t, x_t\}$ is a hybrid switching diffusion [Yin and Zhu, 2010], and common practice is to work with transition matrix $\left[\pi_{\theta\eta}(x)\right]$ with $\pi_{\theta\eta}(x) = \lambda_{\theta\eta}(x)$, for $\eta \neq \theta$, and $\pi_{\theta\theta}(x) = \sum_{\eta \neq \theta} -\lambda_{\theta\eta}(x)$.

Remark 5.2.5: If $\lambda_{\theta\eta}(x)$ is x-invariant for each θ, η, then a hybrid switching diffusion is a Markov switching diffusion

5.2.4 Rare event estimation

The problem is to estimate the probability γ that $\{z_t\} = \{\theta_t, x_t\}$ reaches a closed subset $D \subset \{\theta \times \mathbb{R}^{d(\theta)}; \theta \in \Theta\}$ within finite period $[0,T]$, i.e.

$$\gamma = P(\tau < T) \tag{5.2.4}$$

with τ being the first hitting time of D by $\{\theta_t, x_t\}$:

$$\tau = \inf\{t > 0, (\theta_t, x_t) \in D\} \tag{5.2.5}$$

To analyse rare event estimation for GSHS, without loss of generality, we adopt the following condition:

C0. The discrete-valued process $\{\theta_t\}$ embeds the rare event indicator process $\kappa_t = 1\{t \geq \tau\}$, which implies $\partial D \subset \partial \Xi$, and the process $\{\theta_t, x_t\}$ stops evolving upon hitting D, i.e. $\lambda(\theta, x) = f(\theta, x) = g(\theta, x) = 0$, if $\kappa_t = \theta_t^{\kappa} = 1$.

Together with assumptions A1-A4, condition C0 is assumed to hold true throughout the remainder of this chapter.

5.3 Importance Sampling of GSHS

In a GSHS, there are four candidate components for Importance Sampling (IS): the diffusion components *f* and *g*, the transition rate function λ, and the transition measure Q. In this chapter we restrict our attention to IS modification of the latter two only.

5.3.1 Importance sampling modification of λ and Q

Let λ and Q be IS modified to λ^* and Q^*, in such a way that the modified process $\{z_t^*\}$ has a significant higher rare event probability $\gamma^* = P(\tau^* < T)$, with τ^* being the first hitting time of D by $\{z_t^*\}$.

Between successive stopping times s_j^* and s_{j+1}^*, the process $\{z_t^*\}$ evolves according to a W-adapted flow: $z_t^* = \Phi^W_{s_j^*, z^*_{s_j}}(t - s_j^*)$, where $\Phi^W_{s_j^*, z^*_{s_j}}(t - s_j^*)$ is the solution of SDE (5.2.1), starting at $(s_j^*, z^*_{s_j^*})$, and given W. Thanks to Assumption A2, this solution exists and is pathwise unique.

Hence, the evolution of $\{z_t^*\}$ on $[0,T]$ is embedded in $\{W, z^*_{s_0^*}, (s_1^* - s_0^*), z^*_{s_1^*}, \ldots.., (s_n^* - s_{n-1}^*), z^*_{s_n^*}\}$, with s_n^* the last switching time prior to T. Because $\{z_t^*\}$ does not evolve when $\kappa_t^* = 1$, we know $\kappa^*_{s_j^*} = 0$ for each $j < n$, and during the period $(s_n^*, T]$ no jump happens.

The above embedding is subsequently used to characterize the likelihood ratio to compensate a biased reach probability estimate that is obtained by conducting Monte Carlo simulation of a GSHS with the modified λ^* and Q^*.

Proposition 5.3.1

Let λ and Q be modified to λ^* and Q^* respectively. Conducting a Monte Carlo simulation of N_{runs} for this modified process yields the unbiased estimator of γ

$$E\{\gamma\} = \frac{1}{N_{run}} \sum_{i=1}^{N_{run}} 1\{z_{n,i}^{\kappa} = 1\} L(Z_i, W_i)\big|_{z_{n,i}^{\kappa}=1} \tag{5.3.1}$$

with likelihood ratio:

$$L(Z,W)_{|z_n^\kappa=1} = L(t_1, z_1, \ldots\ldots, t_n, z_n |W)_{|z_n^\kappa=1}$$

$$= \frac{\exp\{-\Lambda^W_{s_{n-1},z_{n-1}}(t_n)\}}{\exp\{-\Lambda^{*W}_{s_{n-1},z_{n-1}}(t_n)\}} \cdot \prod_{j=1}^{n-1} \left[\frac{\lambda^W_{s_{j-1},z_{j-1}}(t_j)^{1\{t_j < t^W_{s_{j-1},z_{j-1}}\}} \exp\{-\Lambda^W_{s_{j-1},z_{j-1}}(t_j)\}}{\lambda^{*W}_{s_{j-1},z_{j-1}}(t_j)^{1\{t_j < t^W_{s_{j-1},z_{j-1}}\}} \exp\{-\Lambda^{*W}_{s_{j-1},z_{j-1}}(t_j)\}} \cdot \frac{Q_{\Phi^W_{s_{j-1},z_{j-1}}(t_j)}(z_j)}{Q^*_{\Phi^W_{s_{j-1},z_{j-1}}(t_j)}(z_j)} \right] \tag{5.3.2}$$

where:

$$\lambda^W_{s_j,z_j}(t) \triangleq \lambda(\Phi^W_{s_j,z_j}(t)),\ \lambda^{*W}_{s_j,z_j}(t) \triangleq \lambda^*(\Phi^W_{s_j,z_j}(t)) \tag{5.3.3}$$

$$\Lambda^W_{s_j,z_j}(t) \triangleq \int_0^t \lambda^W_{s_j,z_j}(s)ds,\ \ \Lambda^{*W}_{s_j,z_j}(t) \triangleq \int_0^t \lambda^{*W}_{s_j,z_j}(s)ds \tag{5.3.4}$$

$$t^W_{s_j,z_j} \triangleq \inf\{t > 0;\ \Phi^W_{s_j,z_j}(t) \in \partial\Xi\}. \tag{5.3.5}$$

Proof:

Because the IS modification influences GSHS elements λ and Q only, the process $\{z_t^*\}$ has the following similarities with $\{z_t\}$: i) $p_{z_0^*}(.) = p_{z_0}(.)$; ii) $f^*(.) = f(.)$; and iii) $g^*(.) = g(.)$. Hence, the conditional joint probability density of $\{z_t^*, t \in [0, s_n^*]\}$, given W, satisfies:

$$p_{z^*_{s_0^*},(s_1^*-s_0^*),z^*_{s_1^*},\ldots\ldots,(s_n^*-s_{n-1}^*),z^*_{s_n^*}}(z_0,t_1,z_1,..,t_n,z_n|W) = p_{z^*_{s_0^*}}(z_0) \prod_{j=1}^{n} \left[p_{(s_j^*-s_{j-1}^*)|z^*_{s_{j-1}^*}}(t_j \mid z_{j-1},W)\, p_{z^*_{s_j^*}|(s_j^*-s_{j-1}^*),z^*_{s_{j-1}^*}}(z_j \mid t_j, z_{j-1},W) \right] \tag{5.3.6}$$

A similar joint probability density holds for $\{z_t\}$. Hence, for $Z = \{z_0, t_1, z_1, ..., t_n, z_n\}$, the Radon Nikodym derivative $L(Z,W)$ equals a quotient between the conditional probability densities of $\{z_t\}$ and $\{z_t^*\}$, given W:

$$L(Z,W) = L(Z|W)\frac{p_W(W)}{p_W(W)} = L(Z|W) = L(t_1, z_1, \ldots\ldots, t_n, z_n |W)$$

$$= \prod_{j=1}^{n} \left[\frac{p_{(s_j-s_{j-1})|z_{s_{j-1}}}(t_j \mid z_{j-1},W)}{p_{(s_j^*-s_{j-1}^*)|z^*_{s_{j-1}^*}}(t_j \mid z_{j-1},W)} \cdot \frac{p_{z_{s_j}|(s_j-s_{j-1}),z_{s_{j-1}}}(z_j \mid t_j, z_{j-1},W)}{p_{z^*_{s_j^*}|(s_j^*-s_{j-1}^*),z^*_{s_{j-1}^*}}(z_j \mid t_j, z_{j-1},W)} \right]$$

By conducting a Monte Carlo simulation of N_{run} runs for the process $\{z_t^*\}$ yields realizations $\{W_i, Z_i,\ i = 1,.., N_{run}\}$. Hence, an unbiased estimator of γ then is:

$$E\{\gamma\} = \frac{1}{N_{run}} \sum_{i=1}^{N_{run}} 1\{z^\kappa_{n,i} = 1\} L(Z_i, W_i)$$

Because $L(Z_i, W_i)$ only plays a role in $E\{\gamma\}$ if $z^\kappa_{n,i} = 1$, this yields eq. (5.3.1) with:

$$L(Z,W)_{|z_n^\kappa=1} = L(t_1, z_1, \ldots\ldots, t_n, z_n |W)_{|z_n^\kappa=1} = \frac{p_{(s_n-s_{n-1})|z_{s_{n-1}}}(t_n \mid z_{n-1},W)}{p_{(s_n^*-s_{n-1}^*)|z^*_{s_{n-1}^*}}(t_n \mid z_{n-1},W)} \cdot$$

$$\cdot \prod_{j=1}^{n-1} \left[\frac{p_{(s_j-s_{j-1})|z_{s_{j-1}}}(t_j \mid z_{j-1},W)}{p_{(s_j^*-s_{j-1}^*)|z^*_{s_{j-1}^*}}(t_j \mid z_{j-1},W)} \cdot \frac{p_{z_{s_j}|(s_j-s_{j-1}),z_{s_{j-1}}}(z_j \mid t_j, z_{j-1},W)}{p_{z^*_{s_j^*}|(s_j^*-s_{j-1}^*),z^*_{s_{j-1}^*}}(z_j \mid t_j, z_{j-1},W)} \right] \tag{5.3.7}$$

The terms in (5.3.7) are characterized as follows:

$$p_{z_{s_{j+1}}|z_{s_j},(s_{j+1}-s_j)}(z_{j+1}\,|\,z_j,t_j,W)=Q_{\Phi^W_{s_j,z_j}(t_j)}(z_{j+1})$$

$$p_{(s_{j+1}-s_j)|z_{s_j}}(t_j\,|\,z_j,W)=\begin{cases}\lambda^W_{s_j,z_j}(t_j)\exp\{-\Lambda^W_{s_j,z_j}(t_j)\}, & \text{for } t_j<t^W_{s_j,z_j}\\ \delta_{t^W_{s_j,z_j}}(t_j)\exp\{-\Lambda^W_{s_j,z_j}(t_j)\}, & \text{for } t_j\geq t^W_{s_j,z_j}\end{cases}$$

Substituting these characterizations in (5.3.7), and subsequent evaluation, yields eq. (5.3.2).

Q.E.D.

5.3.2 Optimal IS for a GSHS

This section develops an optimal IS strategy for λ and Q of a GSHS, under the condition that coefficients *f* and *g* are not modified. This extends the optimal IS development by Chraibi et al. [2019] for a PDMP to a GSHS, i.e. for the case $g=0$ in (5.2.1) to $g\neq 0$.

Definition 5.3.2: IS strategy that modifies λ and Q.

$$Q^*_\xi(s_j,z)=\frac{U^{*|W}_{s_j}(z)}{U^{\neq|W}_{s_j}(\xi)}Q_\xi(z)\ \text{, all}\,(\xi,z)\in\bar{\bar{\Xi}}\times\bar{\bar{\Xi}}. \tag{5.3.8}$$

$$p_{(s^*_{s_{j+1}}-s^*_{s_j})|z^*_{s^*_j}}(u|z,W)=\frac{U^{\neq|W}_{s_j+u}(\Phi^W_{s_j,z}(u))}{c^W(s_j,z)}p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W) \tag{5.3.9}$$

with:

$$U^{*|W}_s(z)\triangleq P\{\tau<T\,|\,z_s=z,W\} \tag{5.3.10}$$

$$U^{\neq|W}_s(\xi)\triangleq\sum_{\eta\in\Theta^\kappa}\int_{\bar{X}^\eta}U^{*|W}_s(\eta,y)Q_\xi(\eta,y)d\varpi(y) \tag{5.3.11}$$

$$c^W(s_j,z)\triangleq\int_0^\infty U^{\neq|W}_{s_j+u}(\Phi^W_{s_j,z}(u))p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)d\varpi(u), \tag{5.3.12}$$

where $\Phi^W_{s_j,z}(u)$ is the solution of SDE (5.2.1) at moment s_j+u given W and $(\theta_{s_j},x_{s_j})=z$.

Theorem 5.3.3

If GSHS coefficients *f* and *g* are not modified, then the IS strategy of Definition 5.3.2 is optimal.

To prepare for the proof of Theorem 5.3.3, we first derive Lemmas 5.3.4 and 5.3.5 below.

Lemma 5.3.4: $U^{\neq|W}_s(\xi),\xi\in\bar{\bar{\Xi}},$ is the conditional hit probability given $z_{s-}=\xi$, $z_s\neq\xi$, and W, i.e.

$$U^{\neq|W}_s(\xi)=P\{\tau<T\,|\,z_s\neq z_{s-}=\xi,W\}. \tag{5.3.13}$$

Proof: Evaluation of $U^{\neq|W}_s(\xi)$ in (5.3.11) yields:

$$
\begin{aligned}
U_s^{\neq|W}(\xi) &= \sum_{\eta\in\Theta} \int_{\bar{X}^\eta} U_s^{*|W}(\eta,y) Q_\xi(\eta,y) d\varpi(y) \\
&= \sum_{\eta\in\Theta^\kappa} \int_{\bar{X}^\eta} \left[P\{\tau<T \,|\, z_s=(\eta,y),W\} Q_\xi(\eta,y) \right] d\varpi(y) \\
&= \sum_{\eta\in\Theta^\kappa} \int_{\bar{X}^\eta} \left[P\{\tau<T,(\theta_s,x_s)\in\{\eta\}\times dy \,|\, z_s\neq z_{s-}=\xi,W\} \right] \\
&= P\{\tau<T \,|\, z_s\neq z_{s-}=\xi,W\}
\end{aligned}
$$

Q.E.D.

Lemma 5.3.5: For all $z\in\Xi$:

$$
p_{(s^*_{s_{j+1}}-s^*_{s_j})|z^*_{s^*_j}}(u|z,W) = \frac{U_{s_j+u}^{\neq|W}(\Phi^W_{s_j,z}(u))}{U_{s_j}^{*|W}(z)} p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W) \tag{5.3.14}
$$

Proof: From Lemma 5.3.4 we get:

$$
U_{s_j+u}^{\neq|W}(\Phi^W_{s_j,z}(u)) = P\{\tau<T \,|\, z_{s_j+u}\neq z_{s_j+u-}=\Phi^W_{s_j,z}(u),W\}
$$

Substituting this in eq. (5.3.12) yields:

$$
\begin{aligned}
c^W(s_j,z) &= \int_0^\infty P\{\tau<T \,|\, z_{s_j+u}\neq z_{s_j+u-}=\Phi^W_{s_j,z}(u),W\} p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W) d\varpi(u) \\
&= \int_0^\infty P\{\tau<T \,|\, z_{s_{j+1}}\neq z_{s_{j+1}-}=\Phi^W_{s_j,z}(s_{j+1}-s_j),W\} p_{s_{j+1}-s_j|z_{s_j}}(u|z,W) d\varpi(u) \\
&= P\{\tau<T \,|\, z_{s_{j+1}}\neq z_{s_{j+1}-}=\Phi^W_{s_j,z}(s_{j+1}-s_j),W\} \\
&= P\{\tau<T \,|\, z_{s_{j+1}-}=\Phi^W_{s_j,z}(s_{j+1}-s_j),W\} \\
&= P\{\tau<T \,|\, z_{s_j}=z,W\} = U_{s_j}^{*|W}(z)
\end{aligned}
$$

Substituting this in eq. (5.3.9) yields (5.3.14). **Q.E.D.**

Proof of Theorem 5.3.3: The IS strategy of Definition 5.3.2 yields the process $\{z^*_t\}$ with transition density:

$$
p_{z^*_{s_{j+1}}|z^*_{s_j},(s^*_{j+1}-s^*_j)}(z_{j+1}\,|\,z_j,t_{j+1},W) = Q^*_{\Phi^W_{s_j,z_j}(t_{j+1})}(s_j+t_{j+1},z_{j+1}) = \frac{U_{s_{j+1}}^{*|W}(z_{j+1})}{U_{s_j+t_{j+1}}^{\neq|W}(\Phi^W_{s_j,z_j}(t_{j+1}))} Q_{\Phi^W_{s_j,z_j}(t_{j+1})}(z_{j+1})
$$

Substituting this together with eq. (5.3.14) in eq. (5.3.6), yields:

$$p_{\{z_t^*\}}(Z|W)=$$

$$=p_{z_{s_0}^*}(z_0)\prod_{j=1}^{n}\left[p_{(s_j^*-s_{j-1}^*)|z_{s_{j-1}}^*}(t_j\mid z_{j-1},W)\,p_{z_{s_j}^*|(s_j^*-s_{j-1}^*)z_{s_{j-1}}^*}(z_j\mid t_j,z_{j-1},W)\right]$$

$$=p_{z_{s_0}^*}(z_0)\prod_{j=1}^{n}\left[\frac{U_{s_{j-1}+t_j}^{\neq|W}(\Phi_{s_{j-1},z_{j-1}}^{W}(t_j))}{U_{s_{j-1}}^{*|W}(z_{j-1})}p_{(s_j-s_{j-1})|z_{s_{j-1}}}(t_j\mid z_{j-1},W)\right]$$

$$\cdot\prod_{j=1}^{n}\left[\frac{U_{s_j}^{*|W}(z_j)}{U_{s_{j-1}+t_j}^{\neq|W}(\Phi_{s_{j-1},z_{j-1}}^{W}(t_j))}p_{z_{s_j}|(s_j-s_{j-1}),z_{s_{j-1}}}(z_j\mid t_j,z_{j-1},W)\right]$$

$$=\prod_{j=1}^{n}\left[\frac{U_{s_{j-1}+t_j}^{\neq|W}(\Phi_{s_{j-1},z_{j-1}}^{W}(t_j))}{U_{s_{j-1}}^{*|W}(z_{j-1})}\frac{U_{s_j}^{*|W}(z_j)}{U_{s_{j-1}+t_j}^{\neq|W}(\Phi_{s_{j-1},z_{j-1}}^{W}(t_j))}\right].p_{\{z_t\}}(Z|W)$$

Subsequent evaluation yields:

$$p_{\{z_t^*\}}(Z|W)=\prod_{j=1}^{n}\left[\frac{U_{s_j}^{*|W}(z_j)}{U_{s_{j-1}}^{*|W}(z_{j-1})}\right].p_{\{z_t\}}(Z|W)=\frac{U_{s_n}^{*|W}(z_n)}{U_{s_0}^{*|W}(z_0)}.p_{\{z_t\}}(Z|W)$$

$$=\frac{P\{\tau<T|z_{s_n}=z_n,W\}}{P\{\tau<T|z_{s_0}=z_0,W\}}p_{\{z_t\}}(Z|W) \tag{5.3.15}$$

$$=\frac{P\{z_n^{\kappa}=1|z_{s_n}=z_n,W\}}{P\{\tau<T|z_{s_0}=z_0,W\}}p_{\{z_t\}}(Z|W)$$

where z_n^{κ} denote the κ component of z_n.

Eq. (5.3.15) implies that IS of Definition 5.3.2 is optimal. **Q.E.D.**

5.3.3 Characterizing the optimal IS for a GSHS

This Section develops a characterization of the optimal IS strategy of Theorem 5.3.3. Of this IS strategy, $Q_{\xi}^*(s_j,z)$ satisfies eqs. (5.3.8) and (5.3.10), from Definition 5.3.2, while $p_{(s_{s_{j+1}}^*-s_{s_j}^*)|z_{s_j^*}^*}(u|z,W)$ satisfies eqs. (5.3.14) and (5.3.13) from Lemmas 5.3.5 and 5.3.4 respectively.

Theorem 5.3.6: For $t<t_z^*$, given *W*:

$$\lambda_{s_j,z}^{*|W}(t)=\frac{U_{s_j+t}^{\neq|W}(\Phi_z^W(t))}{U_{s_j+t}^{*|W}(\Phi_z^W(t))}\lambda_z^W(t) \tag{5.3.16}$$

Proof: The jump rate $\lambda_{s_j^*,z}^{*|W}(t)$, given *W*, satisfies for $t<t_z^*$:

$$\lambda_{s_j^*,z}^{*|W}(t)=\frac{p_{(s_{j+1}^*-s_j^*)|z_{s_j^*}^*}(t|z,W)}{1-\int_0^t p_{(s_{j+1}^*-s_j^*)|z_{s_j^*}^*}(u|z,W)du}=\frac{p_{(s_{j+1}^*-s_j^*)|z_{s_j^*}^*}(t|z,W)}{\int_t^{\infty} p_{(s_{j+1}^*-s_j^*)|z_{s_j^*}^*}(u|z,W)du}$$

Substitution of eq. (5.3.14) yields:

$$\lambda^{*|W}_{s_j,z}(t)=\frac{\dfrac{U^{\neq|W}_{s_j+t}(\Phi^W_z(t))}{U^{*|W}_{s_j}(z)}\,p_{(s_{j+1}-s_j)|z_{s_j}}(t|z,W)}{\int\limits_t^\infty \dfrac{U^{\neq|W}_{s_j+u}(\Phi^W_z(u))}{U^{*|W}_{s_j}(z)}\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du}=\frac{U^{\neq|W}_{s_j+t}(\Phi^W_z(t))\,p_{(s_{j+1}-s_j)|z_{s_j}}(t|z,W)}{\int\limits_t^\infty U^{\neq|W}_{s_j+u}(\Phi^W_z(u))\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du}$$

Using $p_{(s_{j+1}-s_j)|z_{s_j}}(t\,|\,z,W)=\lambda^W_z(t)\exp\{-\Lambda^W_z(t)\}$ yields:

$$\lambda^{*|W}_{s_j,z}(t)=\frac{U^{\neq|W}_{s_j+t}(\Phi^W_z(t))\lambda^W_z(t)\exp\{-\Lambda^W_z(t)\}}{\int\limits_t^\infty U^{\neq|W}_{s_j+u}(\Phi^W_z(u))\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du}$$

Hence, it remains to be proven that, given *W*, for $t<t^*_z$:

$$\int\limits_t^\infty U^{\neq|W}_{s_j+u}(\Phi^W_z(u))\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du=U^{*|W}_{s_j+t}(\Phi^W_z(t))\exp\{-\Lambda^W_z(t)\}$$

This remaining proof is accomplished as follows.

From eq. (5.3.13), we get:

$$U^{\neq|W}_{s_j+u}(\Phi^W_z(u))=P\{\tau<T\,\big|\,z_{s_j+u}\neq z_{s_j+u-}=\Phi^W_z(u),W\}$$

Substitution and subsequent evaluation, yields for $t<t^*_z$:

$$\begin{aligned}
&\int\limits_t^\infty U^{\neq|W}_{s_j+u}(\Phi^W_z(u))\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du\\
&=\int\limits_t^\infty P\Big\{\tau<T\,\Big|\,z_{s_j+u}\neq z_{s_j+u-}=\Phi^W_z(u),W\Big\}\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du\\
&=\int\limits_t^\infty P\Big\{\tau<T\,\Big|\,z_{s_{j+1}}\neq z_{s_{j+1}-}=\Phi^W_z(u),W\Big\}\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du\\
&=\int\limits_t^\infty P\Big\{\tau<T\,\Big|\,z_{s_{j+1}}\neq z_{s_{j+1}-},z_{s_j+t}=\Phi^W_z(t),W\Big\}\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du\\
&=\int\limits_t^\infty U^{*|W}_{s_j+t}(\Phi^W_z(t))\,p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du\\
&=U^{*|W}_{s_j+t}(\Phi^W_z(t))\int\limits_t^\infty p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W)du\\
&=U^{*|W}_{s_j+t}(\Phi^W_z(t))\exp\{-\Lambda^W_z(t)\} \qquad\qquad \text{Q.E.D.}
\end{aligned}$$

Corollary 5.3.7

$$U^{*|W}_{s_j+t^*_z-}(\Phi^W_z(t^*_z))=U^{\neq|W}_{s_j+t^*_z}(\Phi^W_z(t^*_z)) \tag{5.3.17}$$

Proof: From eqs. (5.3.10) and (5.3.13) we get:

$$U^{*|W}_{s_j+t^*_z-}(\Phi^W_z(t^*_z)) = P\{\tau < T \big| z_{s_j+t^*_z-} = \Phi^W_z(t^*_z), W\}$$

$$= P\{\tau < T \big| z_{s_j+t^*_z} \neq z_{s_j+t^*_z-} = \Phi^W_z(t^*_z), W\} = U^{\neq|W}_{s_j+t^*_z}(\Phi^W_z(t^*_z))$$ Q.E.D.

Corollary 5.3.8: Given *W*, for $t < t^*_z$:

$$\frac{\partial U^{*|W}_{s_j+t}(\Phi^W_z(t))}{\partial t} = \left[U^{*|W}_{s_j+t}(\Phi^W_z(t)) - U^{\neq|W}_{s_j+t}(\Phi^W_z(t)) \right] \lambda^W_z(t) \tag{5.3.18}$$

Proof: From the proof of Theorem 5.3.6, given *W,* for $t < t^*_z$:

$$U^{*|W}_{s_j+t}(\Phi^W_z(t)) \exp\{-\Lambda^W_z(t)\} = \int_t^\infty U^{\neq|W}_{s_j+u}(\Phi^W_z(u)) p_{(s_{j+1}-s_j)|z_{s_j}}(u|z,W) du$$

Partial derivation w.r.t. $t < t^*_z$ yields the result. **Q.E.D.**

Remark 5.3.9: [Chraibi et al., 2019, Theorems 4.4-4.6] derived PDMP versions of Theorem 5.3.6 and Corollaries 5.3.7 and 5.3.8. The next Theorem provides additional insight.

Theorem 5.3.10: Given *W*, for $t_j < t^*_z$:

$$\lambda^{*|W}_{s_j,z_j}(t_j) Q^*_{\Phi_{z_j}(t_j)}(s_{j+1}, z_{j+1}) = \frac{U^{*|W}_{s_j+t_j}(z_{j+1})}{U^{*|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))} \lambda^W_{z_j}(t_j) Q_{\Phi^W_{z_j}(t_j)}(z_{j+1}) \tag{5.3.19}$$

Proof: We start from eq. (5.3.16), for $(z,t) = (z_j, t_j),\ t_j < t^*_z$:

$$\lambda^{*|W}_{s_j,z_j}(t_j) = \frac{U^{\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))}{U^{*|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))} \lambda^W_{z_j}(t_j)$$

Multiplication of the left and right hand sides by $Q^*_{\Phi^W_{z_j}(t_j)}(s_j + t_j, z_{j+1})$ yields:

$$\lambda^{*|W}_{s_j,z_j}(t_j) Q^*_{\Phi^W_{z_j}(t_j)}(s_j + t_j, z_{j+1}) = \frac{U^{\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))}{U^{*|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))} \lambda^W_{z_j}(t_j) Q^*_{\Phi^W_{z_j}(t_j)}(s_j + t_j, z_{j+1})$$

From (5.3.8) we get:

$$Q^*_{\Phi^W_{z_j}(t_j)}(s_j + t_j, z_{j+1}) = \frac{U^{*|W}_{s_j+t_j}(z_{j+1})}{U^{\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j)} Q_{\Phi^W_{z_j}(t_j)}(z_{j+1})$$

Substituting this in the preceding eq. yields:

$$\lambda^{*|W}_{s_j,z_j}(t_j) Q^*_{\Phi^W_{z_j}(t_j)}(s_j + t_j, z_{j+1})$$

$$= \frac{U^{\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))}{U^{*|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))} \lambda^W_{z_j}(t_j) \frac{U^{*|W}_{s_j+t_j}(z_{j+1})}{U^{\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j)} Q_{\Phi^W_{z_j}(t_j)}(z_{j+1})$$

$$= \frac{U^{*|W}_{s_j+t_j}(z_{j+1})}{U^{*|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))} \lambda^W_{z_j}(t_j) Q_{\Phi^W_{z_j}(t_j)}(z_{j+1})$$ Q.E.D.

5.4. Approximation of Optimal IS

This section studies an approximation of the Optimal IS for a GSHS which has failing subsystems in parallel redundancy, and the Euclidean valued $\{x_t\}$ is pathwise continuous, i.e.:

C1. GSHS consists of a number of subsystems in parallel redundancy that are subject to failure and repair;

C2. $Q_{(\theta,x)}(\eta,y) = Q^{\Theta}_{(\theta,x)}(\eta)\delta_x(y)$, for all $(\theta,x),(\eta,y)\in\Xi$.

First, subsection 5.4.1, follows the idea of [Chraibi et al., 2019] in adopting a family of approximated $U_s^{*|W}(z)$. Next, subsection 5.4.2 evaluates this family for a GSHS. Subsection 5.4.3 shows what this means for a CTMC.

5.4.1 Family of approximated $U_s^{*|W}(z)$ for a GSHS

We propose to approximate the optimal $U_s^{*|W}(z)$ by a family of functions $U_s^{a|W}(z)$, $a\in S_a$. The best setting for parameter a can be determined through minimizing the Kullback-Leibler distance (or cross-entropy) between the laws of the GSHS under the optimal IS of $U_s^{*|W}(z)$ and the approximate IS of $U_s^{a|W}(z)$.

Similar to the characterization of the optimal IS in Section 5.3.3, an approximate $U_s^{a|W}(z)$ yields the following IS strategy:

$$\lambda^{a|W}_{s_j,z_j}(t_j) = \frac{U^{a\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))}{U^{a|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))}\lambda^W_{z_j}(t_j),\ \text{for } t_j < t_z^* \tag{5.4.1}$$

$$Q^a_{\Phi^W_{z_j}(t_j)}(s_j+t_j,\theta_{j+1}) = \frac{U^{a|W}_{s_j+t_j}(\theta_{j+1},\phi^W_{z_j}(t_j))}{U^{a\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))}Q^{\Theta}_{\Phi^W_{z_j}(t_j)}(\theta_{j+1}) \tag{5.4.2}$$

where $U_s^{a\neq|W}(\theta,x) = \sum_{\eta\in\Theta} U_s^{a|W}(\eta,x)Q^{\Theta}_{\theta,x}(\eta)$.

For a PDMP that satisfies C1 and C2, the $U_s^a(\theta,x)$ family proposed by [Chraibi et al., 2019] is:

$$U_s^a(\theta,x) = H\left(a_i,b(\theta)\right)F(s,x) \tag{5.4.3}$$

$$H\left(a_i,b(\theta)\right) = \exp\{a_i.b(\theta^2)\} \tag{5.4.4}$$

where index $i=1$ if $(\theta,x)\in\Xi$ and $i=2$ if $(\theta,x)\in\partial\Xi$, $b(\theta)$ is the number of failing sub-systems under mode θ, and $F(s,x)$ is a function of time s and Euclidean-valued state x.

To take the W-dependence into account for a GSHS, the $U_s^{a|W}(\theta,x)$ family we adopt is:

$$U_s^{a|W}(\theta,x) = H\left(a_i,b(\theta)\right)F^W(s,x) \tag{5.4.5}$$

$$= \exp\{a_i.b(\theta^2)\} \tag{5.4.9}$$

where $H\left(a_i, b(\theta)\right)$ satisfies (5.4.4) and $F^W(\cdot)$ is a $W-$ conditional function of time s and Euclidean-valued state x.

5.4.2 Approximated IS strategies for GSHS

The family proposal (5.4.4)-(5.4.5) is characterized in Theorems 4.4 and 4.5.

Theorem 5.4.4: Under **C1** and **C2**, for $y_j \in X^{\theta_j}$:

$$\lambda^a(\theta_j, y_j) Q^a_{\theta_j, y_j}(s_{j+1}, \theta_{j+1}) = \\ = \begin{cases} \exp\left\{a_1\left(1+2b(\theta_j)\right)\right\} \cdot \lambda(\theta_j, y_j) Q^{\Theta}_{\theta_j, y_j}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j)+1 \\ \lambda(\theta_j, y_j) Q^{\Theta}_{\theta_j, y_j}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j) \\ \exp\left\{a_1\left(1-2b(\theta_j)\right)\right\} \cdot \lambda(\theta_j, y_j) Q^{\Theta}_{\theta_j, y_j}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j)-1 \end{cases} \tag{5.4.6}$$

where $\lambda^a(z_j) \triangleq \lambda^{a|W}_{s_j, z_j}(0)$.

Proof: By multiplication of (5.4.1) and (5.4.2), and subsequent cancelling of $U^{a\neq|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))$ in nominator and denominator:

$$\lambda^{a|W}_{s_j, z_j}(t_j) Q^a_{\Phi^W_{z_j}(t_j)}(s_j+t_j, \theta_{j+1}) = \frac{U^{a|W}_{s_j+t_j}(\theta_{j+1}, \phi^W_{z_j}(t_j))}{U^{a|W}_{s_j+t_j}(\Phi^W_{z_j}(t_j))} \lambda^W_{z_j}(t_j) Q^{\Theta}_{\Phi^W_{z_j}(t_j)}(\theta_{j+1}) \tag{5.4.7}$$

Substitution of (5.4.5) and (5.4.4) in (5.4.7) for $(s,x) \in \Xi$ yields:

$$\begin{aligned} &\lambda^{a|W}_{s_j, \theta_j, x_j}(t_j) Q^a_{\Phi^W_{\theta_j, x_j}(t_j)}(s_{j+1}, \theta_{j+1}) \\ &= \frac{H(a_1, b(\theta_{j+1})) F^W(s_j+t_j, \phi^W_{\theta_j, x_j}(t_j))}{H(a_1, b(\theta_j)) F^W(s_j+t_j, \phi^W_{\theta_j, x_j}(t_j))} \lambda^W_{z_j}(t_j) Q^{\Theta}_{\Phi^W_{\theta_j, x_j}(t_j)}(\theta_{j+1}) \\ &= \frac{H(a_1, b(\theta_{j+1}))}{H(a_1, b(\theta_j))} \lambda^W_{z_j}(t_j) Q^{\Theta}_{\Phi^W_{\theta_j, x_j}(t_j)}(\theta_{j+1}) \\ &= \frac{\exp\{a_1 b(\theta_{j+1})^2\}}{\exp\{a_1 b(\theta_j)^2\}} \lambda^W_{z_j}(t_j) Q^{\Theta}_{\Phi^W_{\theta_j, x_j}(t_j)}(\theta_{j+1}) \end{aligned}$$

Straightforward elaboration yields for $(s,x) \in \Xi$:

$$\lambda^{a|W}_{s_j, \theta_j, x_j}(t_j) Q^a_{\Phi^W_{\theta_j, x_j}(t_j)}(s_{j+1}, \theta_{j+1}) = \\ = \begin{cases} \exp\left\{a_1\left(1+2b(\theta_j)\right)\right\} \cdot \lambda^W_{\theta_j, x_j}(t_j) Q^{\Theta}_{\Phi^W_{\theta_j, x_j}(t_j)}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j)+1 \\ \lambda^W_{\theta_j, x_j}(t_j) Q^{\Theta}_{\Phi^W_{\theta_j, x_j}(t_j)}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j) \\ \exp\left\{a_1\left(1-2b(\theta_j)\right)\right\} \cdot \lambda^W_{\theta_j, x_j}(t_j) Q^{\Theta}_{\Phi^W_{\theta_j, x_j}(t_j)}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j)-1 \end{cases} \tag{5.4.8}$$

This means that the product λQ is multiplied by factors that are *W*-invariant. To make this explicit we define:

$$\lambda^a(\Phi^W_{s_j, z_j}(t_j)) \triangleq \lambda^{a|W}_{s_j, z_j}(t_j)$$

By definition in subsection 5.3.1, we also know: $\lambda^W_{s_j,z_j}(t_j) \triangleq \lambda(\Phi^W_{s_j,z_j}(t_j))$.

Substituting both in (5.4.12) yields for $(s,x) \in \Xi$:

$$\lambda^a(\Phi^W_{s_j,\theta_j,x_j}(t_j))Q^a_{\Phi^W_{\theta_j,x_j}(t_j)}(s_{j+1},\theta_{j+1}) = \tag{5.4.9}$$
$$= \begin{cases} \exp\left\{a_1\left(1+2b(\theta_j)\right)\right\}\cdot\lambda(\Phi^W_{s_j,\theta_j,x_j}(t_j))Q^\Theta_{\Phi^W_{\theta_j,x_j}(t_j)}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j)+1 \\ \lambda(\Phi^W_{s_j,\theta_j,x_j}(t_j))Q^\Theta_{\Phi^W_{\theta_j,x_j}(t_j)}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j) \\ \exp\left\{a_1\left(1-2b(\theta_j)\right)\right\}\cdot\lambda(\Phi^W_{s_j,\theta_j,x_j}(t_j))Q^\Theta_{\Phi^W_{\theta_j,x_j}(t_j)}(\theta_{j+1}), & \text{if } b(\theta_{j+1}) = b(\theta_j)-1 \end{cases}$$

Now we define $y_j \triangleq \phi^W_{s_j,\theta_j,x_j}(t_j) \in X^{\theta_j}$; substituting this in (5.4.9) yields (5.4.6). Q.E.D.

Theorem 5.4.5: Under **C1** and **C2**, for each $y \in \bar{X}^{\theta_j}$:

$$Q^a_{(\theta_j,y)}(s_j+t^*_z,\theta_{j+1}) = \frac{Q^\Theta_{(\theta_j,y)}(\theta_{j+1})}{\sum_{\eta\in\Theta}\left[\exp\{a_2[b(\eta)^2-b(\theta_{j+1})^2]\}Q^\Theta_{(\theta_j,y)}(\eta)\right]} \tag{5.4.10}$$

Proof: For $t_j = t^*_z$, only the Kernel transition applies.

From (5.4.2) we get:

$$Q^a_{\Phi^W_{z_j}(t^*_z)}(s_j+t^*_z,\theta_{j+1}) = \frac{U^{a|W}_{s_j+t^*_z}(\theta_{j+1},\phi^W_{z_j}(t^*_z))Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\theta_{j+1})}{\sum_{\eta\in\Theta}U^{a|W}_{s_j+t^*_z}(\eta,\phi^W_{z_j}(t^*_z))Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\eta)}$$

Substitution of (5.4.5) and (5.4.4) yields:

$$\begin{aligned} &Q^a_{\Phi^W_{z_j}(t^*_z)}(s_j+t^*_z,\theta_{j+1}) \\ &= \frac{H\left(a_2,b(\theta_{j+1})\right)F^W(s_j+t^*_z,\phi^W_{z_j}(t^*_z))Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\theta_{j+1})}{\sum_{\eta\in\Theta}H\left(a_2,b(\eta)\right)F^W(s_j+t^*_z,\phi^W_{z_j}(t^*_z))Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\eta)} \\ &= \frac{H\left(a_2,b(\theta_{j+1})\right)Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\theta_{j+1})}{\sum_{\eta\in\Theta}H\left(a_2,b(\eta)\right)Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\eta)} = \frac{\exp\{a_2b(\theta_{j+1})^2\}Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\theta_{j+1})}{\sum_{\eta\in\Theta}\exp\{a_2b(\eta)^2\}Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\eta)} \end{aligned}$$

Division by $\exp\{a_2b(\theta_{j+1})^2\}$, and elaboration yields:

$$Q^a_{\Phi^W_{z_j}(t^*_z)}(s_j+t^*_z,\theta_{j+1}) = \frac{Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\theta_{j+1})}{\sum_{\eta\in\Theta}\left[\exp\{a_2[b(\eta)^2-b(\theta_{j+1})^2]\}Q^\Theta_{\Phi^W_{z_j}(t^*_z)}(\eta)\right]}$$

Finally, we substitute $\Phi^W_{z_j}(t^*_z) = (\theta_j, y)$, where $y = \phi^W_{z_j}(t^*_z) \in \bar{X}^{\theta_j}$. This yields (4.10) for $y \in \bar{X}^{\theta_j}$.

Q.E.D.

Theorems 5.4.4 and 5.4.5 imply that under conditions C1 and C2, the Brownian motion does not play a role anymore in the approximated IS strategy. Hence, if $g = 0$,then Theorems 5.4.4 and 5.4.5 coincide with the characterizations by [Chrabi et al., 2019, Subsection 5.1].

5.4.3 Approximated IS strategy for CTMC

In earlier proposed IS strategies for CTMC's, e.g. [Reijsbergen et al., 2012], the IS factors apply to transition probabilities of Q. From Theorem 5.4.4, we get an IS strategy for CTMC in which IS factors apply to the transition rates of λQ.

Corollary 5.4.6: For a CTMC, i.e. $d=0$, satisfying **C1**:

$$\lambda^{a}(\theta_j)Q^{a}_{\theta_j}(\theta_{j+1})$$
$$=\begin{cases} \exp\left\{a_1\left(1+2b(\theta_j)\right)\right\}\cdot\lambda(\theta_j)Q^{\ominus}_{\theta_j}(\theta_{j+1}), & \text{if } b(\theta_{j+1})=b(\theta_j)+1 \\ \lambda(\theta_j)Q^{\ominus}_{\theta_j}(\theta_{j+1}), & \text{if } b(\theta_{j+1})=b(\theta_j) \\ \exp\left\{a_1\left(1-2b(\theta_j)\right)\right\}\cdot\lambda(\theta_j)Q^{\ominus}_{\theta_j}(\theta_{j+1}), & \text{if } b(\theta_{j+1})=b(\theta_j)-1 \end{cases}$$

5.5. Application to GSHS Example

In this section we apply IS for the reach probability estimation of a modified version of the heated room example of Charibi et al. (2019, pp. 901-902). In the original example, the room is losing heat to the environment, which has a constant temperature. The extra complication considered is that the external temperature is no longer fixed, though evolves according to a Geometric Brownian motion (GBM). The latter changes the original PDMP into a GSHS example.

5.5.1. Heated room system example

The heated room example is about a room whose temperature $x_{R,t}$ is influenced by three identical heaters, and the energy flow to the exterior of the room. The challenge is to estimate $\mathbb{P}\{\tau<T\}$ with $\tau\triangleq\inf\left\{t>0:x_{R,t}\leq 0\right\}$, for $T=100$ hour.

The Euclidean valued state includes the room temperature $x_{R,t}$ at time *t*, and the exterior temperature $x_{E,t}$ at time *t*. Hence, the Euclidean valued state space is 2-dimensional.The room temperature $x_{R,t}$ evolves according to the differential:

$$dx_{R,t}=[-\beta_1\left(x_{R,t}-x_{E,t}\right)+\beta_2\sum_{i=1}^{3}1_{\theta_{i,t}=ON}]dt \tag{5.5.1}$$

with β_1 the rate of heat transition to the exterior, β_2 the heating rate capacity of a heater that is ON, and $x_{E,t}$ is the exterior temperature. $x_{E,t}<0$ evolves according to SDE:

$$d\left(-x_{E,t}\right)=-x_{E,t}\sigma_E dW_t \tag{5.5.2}$$

where W_t is standard Brownian motion, $x_{E,0}<0$, and σ_E is a volatility parameter. Given initial value $x_{E,0}<0$, eq. (5.5.2) has the following solution:

$$x_{E,t}=\left(x_{E,0}\right)\exp\left\{-\tfrac{\sigma_E{}^2}{2}t+\sigma_E W_t\right\} \tag{5.5.3}$$

The discrete-valued state θ_t consists of a product of the discrete-valued states of the three heaters, i.e. $\theta_t=\left(\theta_{1,t},\theta_{2,t},\theta_{3,t}\right)$. Each of the three heater may switch between three mode values: ON, OFF, and Failed (F). Each Failed heater may be repaired at a spontaneous rate λ_R. Each non-failed heater may switch to mode F at spontaneous rate:

$$\lambda_{F,t}=c_O+c_T x_{R,t}\,. \tag{5.5.4}$$

Non-failed heaters make forced switches if the temperature of the room, $x_{R,t}$, becomes too high or too low. The forced switching laws of the heaters are:

- Switching law 1: If $x_{R,t} \leq x_{\min}$, the second heater activates only if the first one is failed, and the third one activates only if the two other heaters are failed.

- Switching law 2: If $x_{R,t} \leq x_{\min}$, and repair of heater *i* occurs, then the heater status is set to ON only if all other heaters are failed, otherwise the repaired heater is set to OFF.

- Switching law 3: If switching law 1 selects heater *i* as the heater to be turned on, then this heater may fail on demand with probability P_F.

- Switching law 4: If $x_{R,t} \geq x_{\max}$, then all non-failed heaters are switched to mode OFF with probability 1.

Table 5.1 lists the parameter values for this example. Parameter values are from (Charibi et al., 2019), except those for the novel parameter σ_E in eq. (5.5.2).

Table 5.1. Parameter values

Parameter	Value	Unit
β_1	0.1	h^{-1}
β_2	5	ºC/h
c_O	0.0021	h^{-1}
c_T	0.00015	$(h\cdot{}^{o}C)^{-1}$
σ_E	0. ... 0.5	-
D	$(-\infty,0]$	ºC
P_F	0.01	-
T	100	h
$x_{E,0}$	-1.5	ºC
$x_{\max}$	5.5	ºC
$x_{\min}$	0.5	ºC
$x_{R,0}$	7.5	ºC
$\theta_{1,0}$	OFF	-
$\theta_{2,0}$	OFF	-
$\theta_{3,0}$	OFF	-
λ_R	0.2	h^{-1}

5.5.2. IS strategy for Heated room example

The IS approach to be demonstrated for the Heated room example is the approximated IS strategy that has been developed in section 5.4.2 for a GSHS. In the heated room example, spontaneous switchings are: i) spontaneous failure of a heater which is in the ON or OFF mode; and ii) spontaneous repair of a failed heaters. The forced switchings are defined by the switching laws of the heaters. Next we explain how application of Theorems 5.4.4 and 5.4.5 to the spontaneous and forced switchings, yields the IS factors in Table 5.2.

Application of Theorem 5.4.4 to spontaneous failure of a heater yields the IS factor $\exp\{a_1[1+2b(\theta)]\}$. Application of Theorem 5.4.4 to spontaneous repairs yields the IS factor $\exp\{a_1[1-2b(\theta)]\}$. Following law 3, if heater *i* is selected to be switched to ON, then it switches to ON with probability $1-P_F$, and it switches to Failure (F) with probability P_F. Applying this to Theorem 5.4.5 yields as denominators $P_F + (1-P_F)\exp\{-a_2(1+2b(\theta))\}$ and $P_F \exp\{a_2(1+2b(\theta))\} + (1-P_F)$ respectively.

Table 5.2. IS Factors for spontaneous failure/repair rates, and for failure/non-failure probabilities of forced switchings

Transition	**Rate or Probability**	**IS Factor, given joint mode θ**
Spontaneous Failure	$\lambda_{F,t}$	$e^{a_1\{1+2b(\theta)\}}$
Spontaneous Repair	λ_R	$e^{a_1\{1-2b(\theta_j)\}}$
Failure of forced switching	P_F	$\dfrac{1}{P_F + (1-P_F)\exp\{-a_2(1+2b(\theta))\}}$
Non-Failure of forced switching	$1-P_F$	$\dfrac{1}{P_F \exp\{a_2(1+2b(\theta))\} + (1-P_F)}$

As is shown in Section 5.4, the same approximated IS strategy applies if $g=0$, i.e. $\sigma_E = 0$. Hence the minimization of the Kullback-Leibler divergence also coincides. Charibi et al. (2019) obtained optimal values by minimizing the Kullback-Leibler divergence (or cross-entropy); the resulting values are: $a_1 = 0.915$ and $a_2 = 1.197$. The minimization of the Kullback-Leibler divergence for $\sigma_E > 0$ yields the same values for the parameters a_1 and a_2.

5.5.3. Importance Sampling and MC simulation results

For the Heated room example, we estimate $\bar{\gamma}$ by conducting IS and MC simulations that consist of $N_{run} = 100{,}000$. To also estimate the standard deviation, we conduct both each IS and each MC simulation $N_{\bar{\gamma}}$ times. Hence we get $\bar{\gamma}^i, i = 1,..,N_{\bar{\gamma}}$. These results are used to assess the mean $\hat{\gamma}$, the percentage ρ_S of successful IS runs, and the normalized root-mean-square error (RMSE), i.e.

$$\hat{\gamma} = \frac{1}{N_{\bar{\gamma}}} \sum_{i=1}^{N_{\bar{\gamma}}} \bar{\gamma}^i \tag{5.5.5}$$

$$\rho_S = \frac{1}{N_{\bar{\gamma}}} \sum_{i=1}^{N_{\bar{\gamma}}} 1(\bar{\gamma}^i > 0) \tag{5.5.6}$$

$$RMSE = \sqrt{\frac{1}{N_{\bar{\gamma}}} \sum_{i=1}^{N_{\bar{\gamma}}} (\bar{\gamma}^i - \hat{\gamma})^2} \tag{5.5.7}$$

In addition, we estimate the acceleration factor that is obtained by using IS instead of MC:

$$F_{IS} = \left(\frac{RMSE_{MC}}{RMSE_{IS}} \right)^2 \frac{CPU_{MC}}{CPU_{IS}} \tag{5.5.8}$$

The results obtained are collected in Table 5.3. The results in Table 5.3 show that IS outperforms straightforward MC simulation by about a factor 10^4, for this GSHS example. Without Brownian motion, i.e. $\sigma_E = 0$, the acceleration factor is slightly lower than 10^4. The acceleration factor is slightly higher than 10^4, when Brownian motion influences the rare event probability.

Table 5.3. Simulation results for MC and IS applied to the Heated room example for $\Delta = 0.01s$, $N_{run} = 100,000$, $N_{\bar{\gamma}} = 50$.

	σ_E	$\hat{\gamma}$	ρ_S	$RMSE / \hat{\gamma}$	CPU time	F_{IS}
MC	0	1.22×10^{-5}	74%	94.07%	0.92s	1
	0.001	1.23×10^{-5}	72%	94.67%	1.41s	1
	0.01	-	0%	-	-	1
IS	0	1.29×10^{-5}	100%	0.54%	2.71 s	9,215
	0.001	1.29×10^{-5}	100%	0.51%	3.90 s	11,326
	0.01	1.24×10^{-5}	100%	0.54%	4.02 s	>>
	0.5	1.25×10^{-5}	100%	0.52%	4.24 s	>>

5.6. Conclusion

This chapter studied IS for rare event for a GSHS, i.e. a PDMP that involves Brownian motion, i.e. $g \neq 0$. As outcome of this study, the IS developments by Chraibi et al., 2019 for PDMP have been extended to GSHS. Section 5.2 has defined GSHS and its execution process. Section 5.3 has characterized the optimal IS strategy for a GSHS; Brownian motion plays a key role in these derivations. Section 5.4 has developed an approximated IS strategy for a GSHS that consists of a number of subsystems in parallel redundancy, that are subject to failure and repair. Under the additional assumption that the Euclidean valued process has no discontinuities, this approximated IS strategy has shown to be invariant to the Brownian motion in a GSHS. Thanks to the latter, the minimization of the Kullback-Leibler distance works for a GSHS from this class the same as it works for its corresponding PDMP [Chraibi et al., 2019]. For a simple GSHS example the approximated IS strategy has been demonstrated to work well.

Follow-on research is to extend the developed approximated IS strategy to a GSHS where the Euclidean-valued process involves discontinuities. A complementary extension is to study optimal and approximated IS strategies for GSHS that not only modify lambda and Q, though also the drift coefficient *f*.

ACKNOWLEDGEMENT

The first author would like to thank Thomas Galtier for helpful discussion of results in Chraibi et al. (2019).

5.7 References

Blom, H.A.P., Krystul, J., Bakker, G.J., Klompstra, M.B., Klein Obbink, B. (2007). Free flight collision risk estimation by sequential MC simulation. Eds: C.G, Cassandras and J. Lygeros, Stochastic hybrid systems, Taylor & Francis, pp.249-281.

Bujorianu, M.L. and Lygeros, J. (2006). Toward a general theory of stochastic hybrid systems. Eds: H.A.P. Blom and J. Lygeros, Stochastic hybrid systems, Springer, Berlin, pp. 3-30.

Chraibi, H., Dutfoy, A., Galtier, T. and Garnier, J. (2019). On the optimal importance process for piecewise deterministic markov process. ESAIM: Probability and Statistics, 23, pp.893-921.

Davis, M.H. (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. Journal of the Royal Statistical Society: Series B (Methodological), 46(3), pp.353-376.

Dupuis, P., Spiliopoulos, K., and Wang H., Importance Sampling for Multiscale Diffusions, SIAM Multiscale Model. Simul., Vol. 10 (2012), pp. 1–27.

Glasserman, P., Heidelberger, P., and Shahabuddin, P. (1999). Importance sampling in the Heath-Jarrow-Morton framework, J. of Derivatives, Vol. 6, pp. 32-50.

Glasserman, P. (2004). Monte Carlo methods in financial engineering Springer NY, pp. xiii-596.

Goyal, A., Shahabuddin, P., Heidelberger, P., Nicola, V. F., and Glynn, P. W. (1992). A unified framework for simulating Markovian models of highly reliable systems. IEEE Transactions on Computers, C-41, 36–51.

Heidelberger, P. (1995). Fast simulation of rare events in queueing and reliability models, ACM Transactions on Modeling and Computer Simulation, Vol. 5, pp. 43-85.

Juneja, S., and Shahabuddin, P. (2006). Rare event simulation techniques: An introduction and recent advances. In S. G. Henderson & B. L. Nelson (Eds.) Handbooks in operations research and management science (pp. 291–350). Amsterdam: Elsevier, Chap. 11.

L'Ecuyer, P., Mandjes, M., Tuffin, B. (2010) Importance sampling in rare event simulation, Eds: G. Rubino and B. Tuffin, Rare event simulation using Mont Carlo methods, Wiley, pp. 17-38.

L'Ecuyer, Pierre, Tuffin, Bruno, (2011). Approximating zero variance importance sampling in a reliability setting. Ann. Oper. Res. 189 (1), 277–297.

Nakayama, M. K. and Shahabuddin, P. (2004). Quick simulation methods for estimating the unreliability of regenerative models of large highly reliable systems. Probability in the Engineering and Information Sciences, 18, 339–368.

Papadopoulos, C. and Limnios, N. (2002) Quick Monte Carlo methods in stochastic systems and reliability, Ed: S. Osaki, Stochastic Models in Reliability and Maintenance, Springer, pp. 307-333.

Reijsbergen, D., De Boer, P.T., Scheinhardt, W., Haverkort, B. (2012) Rare event simulation for highly dependable systems with fast repairs, Performance Evaluation, Vol. 69, pp. 336-355.

Shahabuddin, P. (1994). Importance sampling for the simulation of highly reliable Markovian systems. Management Science, 40(3), pp.333-352.

Yin, G.G., and Zhu, C. (2010) Hybrid Switching Diffusions, Properties and Applications, Springer.

Zhang, W., Wang, H., Hartmann, C., Weber, M., and Schutte, C., Applications of the cross-entropy method to Importance Sampling and optimal control of diffusions, SIAM Journal on Scientic Computing, Vol. 36 (2014), 18 pages.

6

Conclusion and Future Research

This Ph.D. thesis studied rare event estimation using MC simulation. The motivation for these studies stems from the increased need to evaluate a design of a future ATM ConOps on safety and capacity. With a focus on the GSHS method and Monte Carlo acceleration techniques, the overall aim of this thesis is:

To develop significant improvement in rare event simulation for GSHS

The overall aim has been achieved through a series of interconnected works conducted in this thesis. The comprehensive findings and results will be discussed in this chapter. The structure of this chapter is as follows: Section 6.1 addresses the results from rare event simulations of GSHS, Section 6.2 discusses the contributions of these results, and Section 6.3 proposes several directions for future work.

6.1 Results from rare event simulations of GSHS

This section explains the results obtained for each objective identified in the introduction.

Objective 1: Error Analysis of Multilevel Splitting.

Objective 1 has been studied in Chapter 2 for a multi-dimensional diffusion process using the IPS framework of Cérou et al. (2006). More specifically, the IPS performances have been analysed for four splitting strategies: MR, MS, RMS, and FAS, when employing a finite number of particles. These strategies differ in how they sample the new set of particles from the set of successful particles. It has been proven that the dominance of the four splitting methods in terms of variance satisfies:

$$V_{FAS} \leq V_{RMS} \leq V_{MS} \leq V_{MR}$$

where $V_{FAS}, V_{RMS}, V_{MS}, V_{MR}$ the variance used by IPS-MR, IPS-MS, IPS-RMS and IPS-FAS, respectively. The proof of this inequality has been extended to GSHS in Chapter 3.

These proofs have been realized in a step-wise approach. During the first step, a novel characterizations have been derived for the conditional variances at splitting level *k,* given the set of survived particles at the beginning of the *k*-th IPS cycle; this is done for all four splitting strategies. Subsequently, these conditional variances are compared for the different splitting strategies, and their relative dominance is proven. Then this dominance proof is extended to the variance given that only during the *k*-th IPS cycle different splitting strategies are used. The final step has used the latter result in an inductive way to complete the proof of variance dominance.

The difference in IPS performance under the four splitting strategies has been illustrated for a one-dimensional geometric Brownian motion example for which the reach probabilities are analytically known.

Objective 2: Understanding Effect of Transforming Spontaneous Jumps to Forced Jumps.

Chapter 3 has studied the question: “Can a spontaneous jump in a GSHS model be transformed to a forced jump without changing process behavior that is relevant for IPS based rare event estimation?” The transformation proposed by Lygeros and Prandini (2010) produces a GSHS version that includes an auxiliary Euclidean-valued state component. Chapter 3 first reveals that the execution of such transformed GSHS version no longer satisfies the strong Markov Property. Subsequently it is proven that the effect of the above transformation has a negative effect on the performance of IPS for a GSHS.

To maintain the strong Markov property, the extra state component should be treated as being unobservable for other process(es) than the GSHS execution considered. To formalize this in applying IPS, prior to applying the transformation by Lygeros and Prandini (2010), the original GSHS should be enriched with the first hitting times of the IPS subsets, and the extra state component should be refreshed at these hitting times. The latter refreshment induces a significant improvement in particle diversity at the start of each IPS cycle. As a result of this improved particle diversity, IPS performance in reach probability estimation is expected to significantly improve when reach probability estimation becomes a challenge. For purpose of comparison, in Chapter 3, an algorithm for the direct simulation of a GSHS execution within IPS cycles is specified. Based on theory, use of this algorithm in IPS for GSHS will yield similar good performance as applying IPS to the original GSHS model, though significantly better than applying IPS in combination with the original transformation of Lygeros and Prandini (2010).

The expected IPS performance for the three IPS versions have been illustrated for a GSHS example.

The findings of Chapter 3 mean that for IPS based reach probability estimation for an arbitrary GSHS model, there are two equally well working approaches, The first approach is to apply IPS to the original GSHS model with spontaneous jumps. The second approach is to apply IPS to a GSHS that is obtained through the following two steps: i) To enrich the original GSHS with the first hitting times of the IPS subsets, without affecting the pathwise behavior of the GSHS execution; and ii) To apply the transformation by Lygeros and Prandini (2010) to this enriched GSHS.

Objective 3: Error Analysis of sampling per mode within IPS

Objective 3 has been studied in Chapter 4 for sampling per mode strategies within IPS based rare event simulation for a GSHS. The resulting IPSmode algorithm has been combined with four mode-directed splitting strategies: MRmode, MSmode, RMSmode, and FASmode, each employing a finite number of particles. These mode-directed splitting strategies differ in how they sample a new set of particles from the set of successful particles. It has been proven that under certain conditions the following inequalities satisfy:

$$V_{\text{IPSmode-RMSmode}} \leq V_{\text{IPSmode-MSmode}} \leq V_{\text{IPSmode-FASmode}} \approx V_{\text{IPSmode-MRmode}} \leq V_{\text{IPS-MR}}$$

where $V_{\text{IPSmode-}Z\text{mode}}$ denotes the variance for the algorithm IPSmode-*Z*mode, for $Z \in \{\text{RMS, MS, FAS, MR}\}$, and $V_{\text{IPS-MR}}$ denotes normal IPS with multinomial resampling.

The crucial part in these proofs was to capture the effect of particle weights in each sampling per mode strategy such that the estimated reach probability remains unbiased. In contrast to normal IPS (in Chapters 3), the mode-directed splitting strategies employing RMS and MS outperform the FAS approach. The explanation is that for the FASmode splitting approach it is more demanding to take proper account of the effect of particle weights in mode-dependent splitting.

The various IPSmode versions and IPS-MR have been simulated for a simple GSHS example; the simulation results obtained show similar relative performance of the different IPS versions.

Objective 4: Extending Charibi's IS results for PDMP to GSHS.

Chapter 5 has extended the IS developments by Chraibi et al., (2019) for a PDMP to a GSHS, i.e. a PDMP that involves Brownian motion. First, an optimal IS strategy has been characterized for a GSHS; this showed that Brownian motion plays a key role in the optimal IS strategy for a GSHS. Next, the approximated IS strategy of Chraibi et al. (2019) for a PDMP has been evaluated for application to a GSHS; in both cases the Euclidean valued state is assumed to have no discontinuities. This revealed that this approximated IS strategy does not depend on the Brownian motion in a GSHS, and is therefore equal to the approximated IS strategy for a PDMP. The latter equality implies that the minimization of the Kullback-Leibler divergence of the approximated IS strategy, yields the same parameter values for a GSHS as it yields for its corresponding PDMP.

For a simple GSHS extension of the PDMP example of Chraibi et al. (2019), the approximated IS strategy has been demonstrated to work well.

6.2 Contributions to safety and capacity assessment of Future ATM ConOps Designs

This section explains the contributions to safety and capacity assessment of future ATM ConOps designs for each chapter.

Objective 1. Error Analysis of Multilevel Splitting

For the IPS based rare event simulation of a GSHS model for an advanced ATM ConOps, [Blom et al., 2007] used the MS splitting strategy as a heuristic improvement of the MR splitting strategy. In the MS strategy, the set of successful particles is copied once, and then the remaining particles are selected according to an MR strategy.

Thanks to the IPS splitting strategy results derived in chapter 3, for an arbitrary GSHS, the following two improvements can be realized for the rare event simulation of a GSHS for an ATM ConOps design: 1) It has formally been proven that the use of MS instead of MR will yield a lower variance; and 2) Further reduction of variance will be obtained by using FAS instead of MS. In the practice of rare event simulation of a GSHS model of an ATM ConOps design, the use of IPS-FAS makes it possible to significantly reduce the number of particles that have to be simulated to get the desired level of accuracy in risk estimation.

Objective 2: Understanding Effect of Transforming Spontaneous Jumps to Forced Jumps.

For the IPS-based rare event simulation of an advanced ATM ConOps, common practice, e.g. Blom et al. (2007), is to model non-exponential time delays as forced jumps. The results of Chapter 3 have shown that direct application of IPS to a GSHS with these forced jumps may undermine particle diversity.

To apply the two well working IPS approaches from chapter 3, the given GSHS model has to be transformation in the opposite direction to the transformation of Lygeros and Prandini (2010); i.e., to transform each forced jump model of a random time delay to a spontaneous jump model. This is done as follows for each random delay model. First, an auxiliary state component q_t^*, representing "passed time" starts at an applicable stopping time τ at initial condition $q_\tau^* = 0$, and subsequently evolves as $dq_t^* = dt$. Second, let $p_{delay}(s)$ denote the probability density of the random time delay, then the spontaneous jump rate is increased from $\lambda(x_t, \theta_t)$ to $\lambda(x_t, \theta_t) + \lambda^*(q_t^*)$, with $\lambda^*(q) = p_{delay}(q) / \int_q^\infty p_{delay}(s) ds$. Third, the "remaining time" state component q_t should be deleted.

After this transformation in opposite direction, chapter 3 has shown that there are two equally well working approaches in applying IPS to the resulting GSHS. The first approach is to apply IPS to the resulting GSHS model with spontaneous jumps. The second approach is to apply IPS to a GSHS that is obtained through the following two steps: i) To enrich the resultng GSHS with the first hitting times of the IPS subsets, without affecting the pathwise behavior of the GSHS execution; and ii) To apply the transformation by Lygeros and Prandini (2010) to this enriched GSHS. In contrast to the first approach, the second approach depends on the adopted IPS levels. For this reason, the first approach often will be preferred in application to a GSHS model of an ATM ConOps design.

Objective 3. Error Analysis of Multilevel Splitting

For IPS based rare event simulation of a GSHS model of an advanced ATM ConOps design, [Blom et al., 2007] uses IPS-MR or IPS-MS. Thanks to the results derived in Chapter 4, it has become clear that the variance in estimated reach probability can be reduced by replacing IPS-

MR or IPS-MS by IPSmode-RMSmode. Although it is reasonable to expect that IPSmode-RMSmode will outperform IPS-FAS of chapter 3, formally this has not been proven. Therefore it is recommended to both implement IPS-FAS and IPSmode-RMSmode for the GSHS model considered, and then to compare the two methods through conducting computer simulations.

Objective 4: Extending Charibi's IS results for PDMP to GSHS.

In rare event simulation for a GSHS model of an advanced ATM ConOps, Blom et al. (2007) used IPS.

The question is how the IS results of chapter 5 can be used for rare event simulation of this GSHS model?

Because the optimal IS has no direct practical value, the advantage has to come from the approximated IS strategy for GSHS. Unfortunately, the assumption that Euclidean valued state should evolve pathwise continuous, is a condition often not satisfied by a GSHS model of an ATM Conps. For example, a mode switch from level flight mode to a climb or descent mode involves a simultaneous jump in the climb or descent rate, which invalidates the pathwise continuity assumption. The latter means that the conditions assumed in the given proofs are not satisfied for the approximated IS strategy of chapter 5. However, this does not mean that the approximated IS strategy does not work.

Therefore it is a worthwhile to give it a try, i.e. to adopt the approximated IS strategy for the PDMP version of the GSHS considered. To evaluate how well this IS based rare event simulation works in practice, a simulation based comparison with an IPS based rare event simulation should be conducted.

6.3 Follow-on research

There are several directions for follow-on research on rare event simulation for a GSHS model of an ATM ConOps design. Based on the studies conducted within this PhD thesis, the following five are of special interest.

Performance comparison of IPSmode-RMSmode versus IPS-FAS

In section 6.2, under Objective 3, it was explained that IPSmode-RMSmode will outperform IPS-FAS. To be sure in case of a specifc GSHS model of an ATM CnOps desig, it was recommend to verify this expectation by comparing the two methods through conducting simulations. To avoid the need of such verification, relevant follow-on research is to conduct for GSHS a theoretical comparison of the variances in the estimated reach probabilities under IPS-FAS and IPSmode-RMSmode.

Approximated IS strategy when Euclidean-valued component of GSHS is discontinuous

In section 6.2, under Objective 4, it was identified that the chapter 5 adopted continuity condition of the approximated IS strategy, does not satisfy for a GSHS model of a given ATM ConOps design. It was also explained that it remains worthwhile to give it a try, i.e. to adopt the approximated IS strategy for the PDMP version of the GSHS considered. To evaluate how well this IS based rare event simulation works in practice, a simulation based comparison with an IPS based rare event simulation should be conducted. Obviously, a more powerful approach would be to study the development of an approximated IS strategy for an arbitrary GSHS where the Euclidean-valued process involves discontinuities.

Sensitivity Analysis and IPS

A GSHS model of a given ATM ConOps design involves a large number of parameters, e.g. the mean and variance of the delay in a response of a pilot or air traffic controller to specific events. In safety risk assessment it is important to assess the sensitivity of the estimated reach probability to changes in these parameter values. The most simple sensitivity analysis varies one parameter value at a time, i.e. conduct two rare event simulations: one for the assumed parameter value μ, and another one for a parameter value $\mu+\Delta$. The estimated sensitivity then $[Risk(\mu+\Delta)-Risk(\mu)]/\Delta$ satisfies. The straightforward approach is to use independent random samples during each of the two assessments; this however leads to a large error in the numerator $[Risk(\mu+\Delta)-Risk(\mu)]$. As explained by Glasserman (2003, pp. 380-381), this error in the numerator can be largely reduced by using the same random numbers during both simulations. The challenge that remains to be studied is how to accomplish the use of the same randomly generated numbers in conducting IPS for a GSHS.

IS strategy that modifies the drift coefficient *f*

Chraibi's IS results have been expanded for PDMP to GSHS, i.e. a PDMP that involves Brownian motion. The addition of Brownian motion opens the opportunity to use an IS strategy that is based on modifying the drift coefficient f , e.g.[Glasserman, 2003, section 4.6]. For example, if two aircraft are fly near each other, then the normal Brownian motion model for wind may give rise to an extremely small probability of mid-air collision. By modifying f such that the two aircraft models evolve on collision course, then the mid-air collision risk may go up by multiple orders of magnitude. To get an unbiased reach probability estimate, the latter has to be compensated by the corresponding likelihood ratio. This example shows that there is high potential for developing an IS strategy for a GSHS that can not only modify (λ, Q) though also f.

IPS x IS

This thesis studied IPS and IS separately from each other. The complexity of a GSHS model of an ATM ConOps design makes it very challenging to develop an IS strategy for such given GSHS model. The advantage of the IPS approach is that is provides a means to decompose the rare event simulation problem into a sequence of rare event simulation cycles. During each IPS cycle the next set to be reached is reduced in size. This also gives the possibility to use different IS strategies during different IPS cycles. For example, to use an IS strategy for rare failures during an IPS cycle where aircraft remain well separated, and to use an IS strategy that modifies drift coefficient f during an IPS cycle where aircraft are flying at a potential unsafe distance from each other.

6.4 References

Blom, H.A.P., Krystul, J., and Bakker, G.J., Klompstra, M.B., Klein Obbink, B. (2007). Free flight collision risk estimation by sequential MC simulation. Eds: C.G, Cassandras and J. Lygeros, Stochastic hybrid systems, Taylor & Francis, pp.249-281.

Cérou, F., Del Moral, P., Legland, F., and Lezaud, P. (2006). Genetic genealogical models in rare event analysis. Latin American J. of Probability & Mathematical Statistics, 1, pp. 181-203.

Chraibi, H., Dutfoy, A., Galtier, T. and Garnier, J. (2019). On the optimal importance process for piecewise deterministic Markov process. ESAIM: Probability and Statistics, 23, pp. 893-921.

Glasserman, P. (2003). Monte Carlo Methods in Financial Engineering. Springer Science.

Lygeros, J. and Prandini, M., (2010). Stochastic hybrid systems: a powerful framework for complex, large scale applications. European Journal of Control, 16(6), pp.583-594.

Curriculum Vitae

Hao Ma, born on October 27, 1988, in Xianyang, China. From 2007 to 2011, he studied at Northwestern Polytechnical University (NPU), School of Electronics and Information, in China and received a Bachelor's degree in Electronic Engineering in 2011. During this period, he participated in the NPU Aircraft Model Team, which later won the first prize in the National Aircraft Model Competition. In addition, the pressure sensor system he designed with other students was awarded the second prize in the National Aircraft Model Competition. From 2011 to 2014, he pursued a Master's degree in Circuit Systems at NPU and completed his Master's thesis at the NPU Unmanned Aircraft Research Institute. His master thesis was on the redundancy design. After graduating in 2014, he had the opportunity to pursue a Ph.D. at the School of Aeronautics. During his Ph.D., he participated in several research projects, including the 973 project. At the end of 2016, he had the opportunity to pursue a Ph.D. at Air Transport and Operations (ATO), Delft University of Technology, where his main research was on rare event simulation. During this period, Hao participated three times in ARCH competitions (Applied Verification for Continuous and Hybrid Systems) .

In his spare time, Hao participated in the Robot deliver project in RSA (Robot Student Associate) and was mainly responsible for building the hardware equipment.

Publications

Ma, H. and Blom, H.A.P., Random Assignment Versus Fixed Assignment in Multilevel Importance Splitting for Estimating Stochastic Reach Probabilities. Methodol Comput Appl Probab 24, 2313–2338, 2022. https://doi.org/10.1007/s11009-021-09892-4

Ma, H. and Blom, H.A.P., Interacting particle system based estimation of reach probability of general stochastic hybrid systems. *Nonlinear Analysis: Hybrid Systems*, *47*, 2023, p.101303.

Blom, H.A.P, Ma, H. and Bakker, G.J.. Interacting particle system-based estimation of reach probability for a generalized stochastic hybrid system. *IFAC-PapersOnLine*, *51*(16), 2018, pp.79-84.

Abate, A., Blom, H.A.P., Cauchi, N., Degiorgio, K., Fraenzle, M., Hahn, E.M., Haesaert, S., Ma, H., Oishi, M., Pilch, C., Remke, A., Salamati, M., Soudjani, S., Huijgevoort, B., and Vinod, A. 2019. ARCH-COMP19 Category Report: Stochastic Modelling. In Frehse G., and Althoff, M., editors, 6th International Workshop on Applied Verification of Continuous and Hybrid Systems (ARCH19), vol 61 of EPiC Series in Computing, 2019, pages 62-102.

Abate, A., Blom, H.A.P., Bouissou, M., Cauchi, N., Chraibi, H., Delicaris, J., Haesaert, S., Hartmanns, A., Khaled, M., Lavaei, A., Ma, H., Mallik, K., Niehage, M., Remke, A., Schupp, S., Shmarov, F., Soudjani, S., Thorpe, A., Turcuman, V., and Zuliani, P., ARCH-COMP21 Category Report: Stochastic models. In Frehse, G. and Althoff, M., editors, 8th International Workshop on Applied Verification of Continuous and Hybrid Systems (ARCH21), vol 80 of EPiC Series in Computing, 2021, pages 55–89.

Abate, A., Blom, H.A.P., Delicaris, J., Haesaert, S., Hartmanns, A., Huijgevoort, B., Lavaei, A., Ma, H., Niehage, M., Remke, A., Schön, O., Schupp, S., Soudjani S., and Willemsen, L., ARCH-COMP22 Category Report: Stochastic Models. In: Frehse, G., Althoff, M., Schoitsch E., and Guiochet, J., editors, 9th International Workshop on Applied Verification of Continuous and Hybrid Systems (ARCH22), vol 90 of EPiC Series in Computing, 2022, pages 113-141.

Under review

Ma, H. and Blom, H.A.P., Sampling per mode strategies in rare event simulation of stochastic hybrid systems. Statistics and Computing, submitted in 2022.

Ma, H. and Blom, H.A.P., Importance Sampling in Estimation of Reach Probability of General Stochastic Hybrid Systems. *Methodol Comput Appl Probab*, submitted in 2023.