UNIVERSITY OF CALIFORNIA, SAN DIEGO

Nonintrusive approaches for multiscale/multiphysics problems with random noise

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Engineering Physics)

by

Søren Henri Taverniers

Committee in charge:

Professor Daniel M. Tartakovsky, Chair
Professor Scott Baden
Professor Michael Holst
Professor David Saintillan
Professor Sutanu Sarkar

2016

The dissertation of Søren Henri Taverniers is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2016

DEDICATION

To my parents.

EPIGRAPH

Our situation is not comparable to anything in the past. It is impossible, therefore, to apply methods and measures which at an earlier age might have been sufficient. We must revolutionize our thinking, revolutionize our actions, and must have the courage to revolutionize relations among nations of the world. Clichés of yesterday will no longer do today, and will, no doubt, be hopelessly out of date tomorrow.

—Albert Einstein, in "A Message to Intellectuals" (1948)

It is in the admission of ignorance and the admission of uncertainty that there is a hope for the continuous motion of human beings in some direction that doesn't get confined, permanently blocked, as it has so many times before in various periods in the history of man.

—Richard P. Feynman, in "The Uncertainty of Values", second guest lecture presented in April 1963 at the University of Washington, Seattle

I think people get it upside down when they say the unambiguous is the reality and the ambiguous is merely uncertainty about what is really unambiguous. Let's turn it around the other way: the ambiguous is the reality and the unambiguous is merely a special case of it, where we finally manage to pin down some very special aspect.

—David J. Bohm, in "How Mathematicians Think" (2007) by William Byers

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to give immense thanks to my faculty advisor, Professor Daniel M. Tartakovsky, for supporting me throughout my PhD. Daniel's emphasis on independent, critical thinking and aim for perfection has stimulated me to continuously reassess and improve my work, profoundly increasing its quality beyond what I initially deemed possible. This proved to be key in getting my research accepted for publication in high-level scientific journals. The multidisciplinary nature of Daniel's research has also allowed me to explore a broad array of physical problems, and formulate conclusions that are relevant to a large part of the scientific community and aim to stimulate the continued improvement of techniques at the forefront of modern scientific computing. I am also grateful to Daniel for providing me with the freedom to gain professional experience at Los Alamos National Laboratory during an entire academic year. As my PhD progressed, my relation with Daniel has continued to grow stronger, making him not just an advisor, but also a true friend.

I would also like to acknowledge Dr Frank Alexander of Los Alamos National Laboratory. Frank's intimate knowledge of statistical mechanics helped me in an immeasurable way, and my successful collaboration with him on two of my five PhD projects was an important factor in getting the corresponding research published. Together with Daniel, Frank also played a key role in making my internship at Los Alamos possible and allowing me to work there with a highly talented, multidisciplinary group of scientists. But more than that, Frank was a friend and a counselor who could be called upon in the toughest of situations, weekdays or weekends, both for academic and personal matters, and would provide support without hesitation. For that, I am forever grateful to him.

Next, I would like to mention Terry Haut, Kipton Barros and Turab Lookman who guided me during my internship at Los Alamos. Their support was essential to the success of this endeavor. I would also like to give special thanks to the group leader of

T-4, Avadh Saxena, for helping me adjust to working at a National Lab, listening to my concerns in his office even on Saturdays late at night, and for improving my badminton skills after work. I can also not forget Marlene and Stacie, who helped me with all kinds of practical issues throughout my internship.

Of course, I also need to thank all past and present members of our research group, who were always ready to help me in times of need. Particular recognition should go to David Barajas-Solano, who on multiple occasions served as a seemingly infinite source of information, and Francesca Boso, who assisted me with both academic and practical matters throughout my PhD. I will also forever cherish my discussions with Yawo Ezunkpe and Joseph Bakarji, and my archery trainings with Robert Zimmerman during my internship at Los Alamos.

I would also like to acknowledge current MAE staff members Sandra de Sousa and Lydia Ramirez for helping me sort out paperwork of various natures. In addition, past member Charlotte Lauve, who was my graduate coordinator for four years, deserves special recognition. On countless occasions, Charlotte was there to help me with virtually any kind of problem, big or small, academic or personal. She was not just any graduate coordinator, she truly was the queen who ruled them all.

Last, and not by any means least, I would like to thank my parents for their unwavering support these past six years. They have done everything in their power to help me get through many difficult times, and for this I am immensely grateful to them. I hope this work makes them proud.

Chapter 2, in part, is a reprint of the material as it appears in the paper "Conservative tightly-coupled simulations of stochastic multiscale systems" by S. Taverniers, A.Y. Pigarov, and D.M. Tartakovsky, published in *Journal of Computational Physics*, 313:400–414, Elsevier Inc. (2016). The dissertation author is the primary investigator and author of this paper.

Chapter 3, in part, is a reprint of the material as it appears in the paper "A tightly-coupled domain-decomposition approach for highly nonlinear stochastic multiphysics systems" by S. Taverniers and D.M. Tartakovsky, submitted in revised form for publication in *Journal of Computational Physics*, Elsevier Inc. The dissertation author is the primary investigator and author of this paper.

Chapter 4, in part, is a reprint of the material as it appears in the paper "Impact of parametric uncertainty on estimation of the energy deposition into an irradiated brain tumor" by S. Taverniers and D.M. Tartakovsky, under review with *Journal of Computational Physics*, Elsevier Inc. The dissertation author is the primary investigator and author of this paper.

Chapter 5, in part, is a reprint of the material as it appears in the paper "Noise propagation in hybrid models of nonlinear systems: The Ginzburg-Landau equation" by S. Taverniers, F.J. Alexander, and D.M. Tartakovsky, published in *Journal of Computational Physics*, 262:313–324, Elsevier Inc. (2014). The dissertation author is the primary investigator and author of this paper.

Chapter 6, in part, is a reprint of the material as it appears in the paper "Physics-based statistical learning approach to mesoscopic model selection" by S. Taverniers, T.S. Haut, K. Barros, F.J. Alexander, and T. Lookman, published in *Physical Review E*, 92:053301, American Physical Society (2015). The dissertation author is the primary investigator and author of this paper.

VITA

2006	B. S. in Physics *cum laude*, Katholieke Universiteit Leuven, Leuven, Belgium
2011-2014	Graduate Research Assistant, University of California San Diego, San Diego, California, United States
2012	M. S. in Engineering Sciences (Engineering Physics), University of California San Diego, San Diego, California, United States
2012	C. Phil. in Engineering Sciences (Engineering Physics), University of California San Diego, San Diego, California, United States
2014-2015	Graduate Research Assistant, Los Alamos National Laboratory, Los Alamos, New Mexico, United States
2016	Ph. D. in Engineering Sciences (Engineering Physics), University of California San Diego, San Diego, California, United States

PUBLICATIONS

S. Taverniers, F.J. Alexander, and D.M. Tartakovsky. Noise propagation in hybrid models of nonlinear systems: The Ginzburg-Landau equation. *Journal of Computational Physics*, 262:313–324, 2014.

S. Taverniers, T.S. Haut, K. Barros, F.J. Alexander, and T. Lookman. Physics-based statistical learning approach to mesoscopic model selection. *Physical Review E*, 92:053301, 2015.

S. Taverniers, A.Y. Pigarov, and D.M. Tartakovsky. Conservative tightly-coupled simulations of stochastic multiscale systems. *Journal of Computational Physics*, 313:400–414, 2016.

ABSTRACT OF THE DISSERTATION

Nonintrusive approaches for multiscale/multiphysics problems with random noise

by

Søren Henri Taverniers

Doctor of Philosophy in Engineering Sciences (Engineering Physics)

University of California, San Diego, 2016

Professor Daniel M. Tartakovsky, Chair

A plethora of computational techniques have been developed for computing quantities of interest in multiscale and multiphysics problems combining processes occurring on a broad spatiotemporal range. However, a dearth exists in systematic studies of the impact of random fluctuations on the predictive ability and numerical properties of these algorithms. We consider two nonintrusive approaches for multiphysics problems with random noise: domain decomposition and stochastic collocation. A mass-conserving domain decomposition achieving tight Newton- or Picard-based coupling between linear diffusion equations, one having a Gaussian white-noise source term, reveals that Newton's iteration scales linearly with noise amplitude, while Picard's iteration may scale

superlinearly. For a given solution error, fully-converged ("implicit") coupling is more efficient than single-iteration ("explicit") coupling at low noise strength; at high noise amplitudes, this remains true provided that the time interval between two subsequent implicit coupling communications is sufficiently long. A similar strategy using Jacobian-free Newton-Krylov iteration to solve a highly nonlinear, multiscale diffusion problem forced by a truncated Gaussian boundary noise shows that ensuring path-wise continuity of the state variable and its flux, as opposed to continuity in the mean, accurately propagates random fluctuations and correctly captures system dynamics. Implicit coupling is more efficient than explicit coupling at all coefficients of variation considered, and domain decomposition with path-wise implicit coupling resolves temporally correlated boundary fluctuations when the correlation time exceeds some multiple of an appropriately defined characteristic diffusion time. Application of stochastic collocation to estimate the energy deposition into a brain tumor via X-ray irradiation with parametric uncertainty reveals that the uncertain parameters' coefficients of variation may be amplified by the problem's nonlinearity to the extent that the predictive uncertainty in the energy deposition almost equals the prediction itself. Algorithm refinement for the Ginzburg-Landau equation (GLE) demonstrates the need for adding a coarse-scale random source term to correctly propagate fine-scale Ising fluctuations throughout the computational domain. A moment-based approach with Gaussian closure enabling direct computation of the state variable's statistical moments is shown to be an accurate, and potentially more efficient, alternative to numerical time integration of the system state. A statistically learned stochastic GLE exhibits optimal predictive capacity at a complexity that may differ from that of standard models in the literature. This approach enables data-driven computation of the coarse-scale noise term's amplitude.

Chapter 1

Introduction

"Multiscale" and "multiphysics" are terms coined to describe systems consisting of respectively one or multiple physical, chemical and/or biological processes evolving on a broad spatiotemporal spectrum. The development of algorithms to simulate such problems started in earnest in the 1980s, and was driven, in large part, by the field of fluid-structure interactions (FSI) and research into modeling flows spanning a wide range of Knudsen numbers (Kn).

In FSI, the terms "monolithic" and "component partitioning" refer, respectively, to an *intrusive* approach requiring the development of novel algorithms, and a *nonintrusive* strategy involving the use of legacy solvers. Monolithic [20] methods combine the various processes into a single discrete operator; while this yields a "tight" coupling between the components, it typically necessitates the development of new codes and requires the use of a single time step. This prevents the use of readily available solvers previously developed for the individual components, and does not take into account the various time scales on which the different processes occur. Both shortcomings are addressed by component partitioning [44], also called *domain decomposition* (DD), which divides the computational domain into subregions where the processes occur and solves them

separately using legacy algorithms, each with its own time step and/or spatial grid cell size (see, e.g., Fig. 1.1). While a tight coupling comes naturally to monolithic approaches, DD also allows for a "loose" coupling which may desynchronize components by one time step of a fraction of a time step [75]. A loose coupling makes a DD method prone to instabilities and increased solution error, and is preferably transformed into a tight coupling by casting it in iterative form [43, 42, 75]. However, the accompanying increase in computational cost results in a trade-off between efficiency and accuracy.

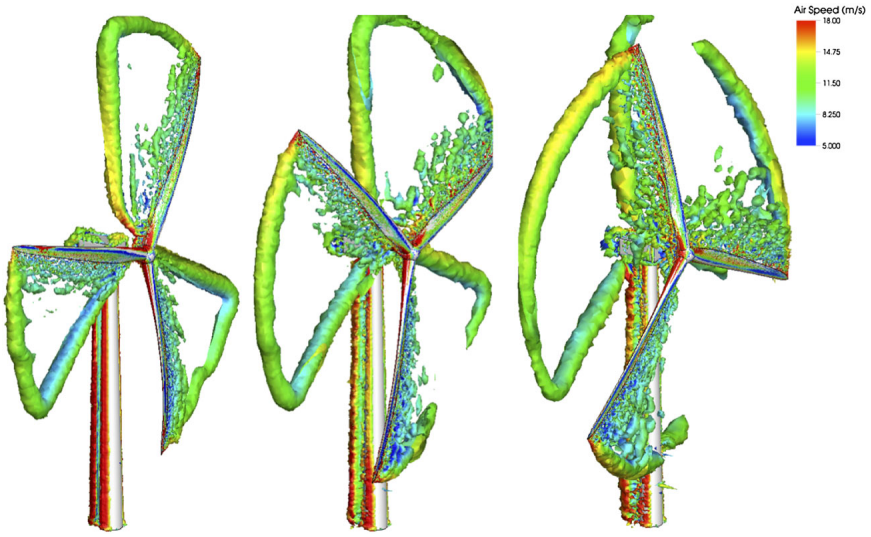

Figure 1.1: Fluid-structure interaction simulation of a wind turbine [15].

In parallel with the modeling efforts in FSI, Wadsworth et al. [119] developed the first coupling of Direct Simulation Monte Carlo [19], a fine-scale atomistic method, to its coarse-scale counterpart represented by a partial differential equation (PDE). This work was motivated by the need for efficient models of variable-Kn flows, such as those encountered around hypersonic re-entry vehicles (Fig. 1.2), and led to the concept of *algorithm refinement (AR) hybrids* [48]. These locally refine a coarse-scale algorithm to a fine-scale counterpart using a criterion based on a key system parameter: when the latter reaches a critical value, the coarse-scale method breaks down and refinement is

triggered. In the re-entry problem, this parameter is $Kn = \lambda/L$ with λ the molecular mean free path and L the characteristic length scale of the problem. In regions where $Kn \ll 1$, collisional equilibrium is achieved and a continuum description such as Navier-Stokes is valid (which is almost everywhere in the computational domain at intermediate altitudes); however, for $Kn \gg 1$, the flow is rarefied (in the wake behind the vehicle, around shocks and in boundary layers) and an atomistic method such as Direct Simulation Monte Carlo must be used. Compared to a fine-scale approach over the entire computational domain, AR hybrids may yield tremendous savings in computational cost by only using an otherwise prohibitively expensive method where absolutely required by physical constraints.

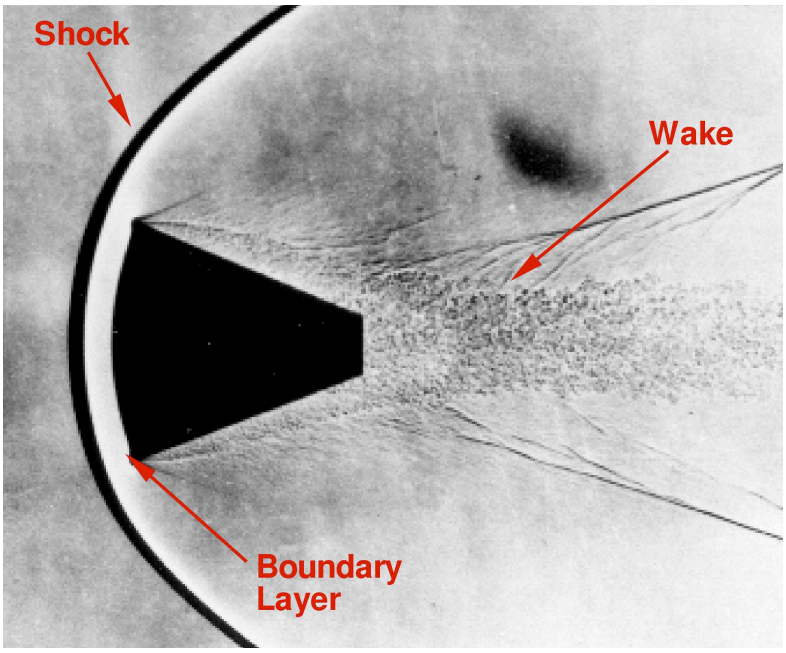

Figure 1.2: Hypersonic flow over an atmospheric re-entry vehicle [1].

DD algorithms with deterministic continuum components have been developed and analyzed extensively. Giles et al. [50] found that an otherwise unstable loose coupling in one-dimensional (1D) FSI simulations may be made stable by enforcing Neumann

boundary conditions for the structural calculation and Dirichlet boundary conditions for the fluid solver. Farhat et al. [43, 42] and Leyland et al. [75] demonstrated that standard staggered schemes for FSI simulations need to be modified by several iteratively made corrector steps to ensure conservation of energy. More recently, Errera et al. [41] investigated the stability of a coupling algorithm based on mixed interface conditions for conjugate heat transfer simulations, while Sheehan et al. [102] showed that using a finite number of iterations in a coupled linear diffusion problem may lead to conditional or unconditional stability in a nonintuitive way when using a backward Euler solver in the subdomains.

When random noise is introduced into such problems as a source term, a boundary or initial condition, an uncertain input parameter, or any combination of the above, conclusions about stability, accuracy and efficiency previously derived for fully-deterministic DD algorithms may no longer hold. In particular, the fidelity of DD simulations involving stochastic components depends on the ability of a coupling algorithm to transmit these fluctuations into adjacent subdomains, including those where a deterministic model is used. A systematic analysis of the numerical properties of DD algorithms with noise, and development of coupling techniques to accurately transport the random fluctuations across interfaces between subdomains, is still largely lacking. In Chapter 2, we contribute to this line of research by developing a mass-conserving, tightly coupled DD algorithm and applying it to the testbed problem of 1D or 2D coupled linear diffusion equations, one of which has a Gaussian space-time white-noise source term. Employing either *Picard's* or *Newton's iteration* [68], we investigate the efficiency of a fully-converged ("implicit") coupling as a function of noise strength, and compare its computational cost for a given solution error with that of single-iteration Picard's coupling for different noise strengths and time intervals between two subsequent inter-solver communications. We also analyze the stability of the DD algorithm with implicit coupling, and compare the

resulting stability conditions with those for the corresponding deterministic system in which the fluctuations have been averaged out.

When the system becomes nonlinear, the presence of noise renders the state variable's ensemble mean different from the solution to its deterministic counterpart, and accurate propagation of fluctuations is important even if one is only interested in modeling the mean behavior of the system. In Chapter 3, we develop a mass-conserving DD strategy to analyze, in 1D, the highly nonlinear testbed problem of hydrogen diffusion through a dense membrane composed of Pd and Ta layers [36, 21, 89], driven by a truncated Gaussian boundary noise. The diffusion through Pd is characterized by a diffusion coefficient [103, 4] that nonlinearly depends on the hydrogen concentration, while the diffusion through Ta is linear [4]. Tight coupling between the subdomain solvers is achieved through Jacobian-free Newton-Krylov iteration [70] with a Generalized Minimum Residual iterative linear solver [68], which only requires computation of Jacobian-vector products rather than the Jacobian itself. We analyze the accuracy of both path-wise and moment-wise (mean) exchange of concentration and flux information at interfaces between adjacent subdomains through comparison with a single full-domain ("global") algorithm, compare the efficiency of implicit with explicit coupling for various noise strengths, and perform a stability analysis of the DD algorithm with path-wise implicit coupling. We mainly consider perfectly correlated noise, but also explore the method's ability to handle fluctuations with a finite correlation time.

In Chapter 4, we increase the system's nonlinearity by considering the problem of X-ray irradiation of a brain tumor, which we model using a 2D multimaterial, equilibrium radiation-diffusion PDE [85, 99]. Here the diffusion coefficient is not only a function of the state variable (radiation energy density) as in Chapter 3, but also of its spatial gradient due to the use of a flux limiter [72]. The problem has up to three uncertain input parameters: the horizontal and vertical location of the center of an inclusion representing

the region over which an iodinated contrast agent, aimed at enhancing the energy deposition in the tumor, has spread out after being injected; and the effective atomic number within this area. We analyze how the mean and variance of the energy deposition in the tumor depend on the mean and/or variance of these input parameters, which we represent as uniformly distributed random variables. A robust, nonintrusive method traditionally used for this purpose is Monte Carlo simulation (MCS), which enables computation of statistical moments of a quantity of interest through a finite number of realizations of the time evolution of the state variable. However, the extremely slow rate of MCS often makes it prohibitively expensive. An alternative, equally nonintrusive, method is *stochastic collocation (SC)* [125, 77], which through a judicious choice of sampling points ("nodes") and their weights via a quadrature rule aims to achieve the same estimation error as MCS with fewer system realizations. We compare the computational efficiency of MCS and SC for a stochastic dimension of one, two or three.

Chapter 5 focuses on the propagation of fine-scale noise generated in atomistic-continuum AR hybrids for nonlinear systems whose macroscale dynamics is driven by microscopic fluctuations. In their traditional form, such hybrids couple the atomistic solver to a deterministic continuum counterpart (usually a PDE solver) [119, 48]. While the fine-scale algorithm inherently represents the microscale fluctuations, the deterministic coarse-scale method does not account for them. As Alexander et al. [5] first demonstrated for a linear diffusion PDE coupled to a random walk algorithm, this reduces the variance inside the atomistic region near the atomistic/continuum interface. Adding a random source term of the appropriate magnitude to the coarse-scale solver enabled the hybrid to correctly propagate the fine-scale fluctuations throughout the computational domain. Subsequent work by Bell et al. [16] and Williams et al. [121] extended this methodology to weakly nonlinear test cases based on the viscous Burgers' and Navier-Stokes equation, respectively. In systems such as those undergoing phase transitions or

instability-driven pattern formation (see, e.g., Fig. 1.3), the effect of microscale fluctuations may be amplified by the nonlinearities, drastically affecting the system's macroscale behavior. We formulate a 1D testbed AR hybrid coupling a stochastic or deterministic Ginzburg-Landau equation (GLE) [58, 24, 118, 96], containing a cubic source term, to a nearest-neighbor Ising [60] model with Glauber [52] (spin-flip) dynamics. We also derive a system of (deterministic) moment equations from the stochastic GLE using a Gaussian closure, i.e., assuming that the fluctuations of the state variable (magnetization) are Gaussian. This approach offers a direct way of calculating moments of the magnetization, such as mean and variance, rather than relying on its numerical time integration.

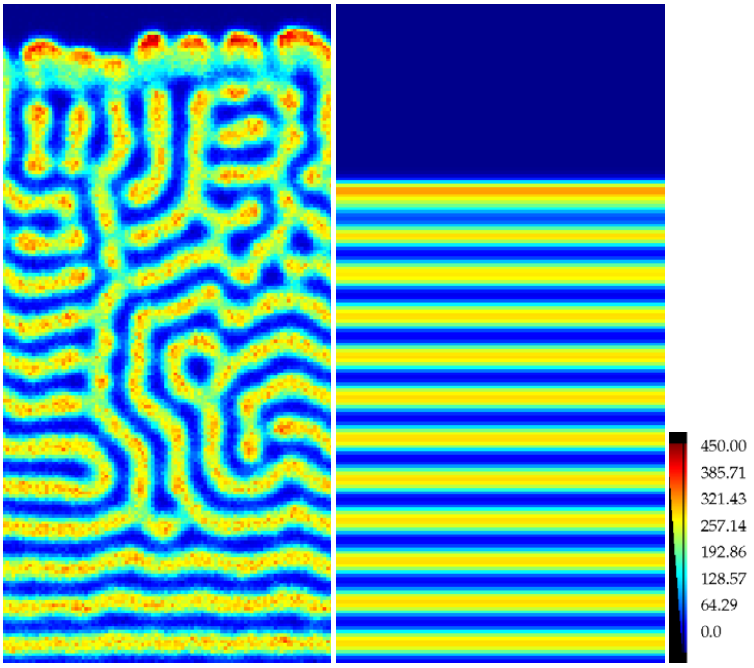

Figure 1.3: Turing-instability driven labyrinth pattern formation, modeled using fluctuating hydrodynamics with fluctuating chemistry (left), and deterministic hydrodynamics with deterministic chemistry (right) [11].

Accurate representation of the fine-scale model in an atomistic-continuum AR hybrid by a stochastic coarse-scale description requires the amplitude of the latter's random source term to match the strength of the fluctuations generated by the fine-

scale representation. As we demonstrate in Chapter 5, this may involve multiplying the noise amplitude with a "fudge factor" to correct possible mismatches stemming from approximations made in deriving the stochastic continuum model from its atomistic counterpart. To avoid such ad hoc procedures, *statistical learning* may be used to build coarse-scale models beyond those available in the literature. In Chapter 6, we employ cross-validation and regularization to build a 2D stochastic GLE from data simulated using the Ising model with Glauber dynamics ("training data"). We analyze the learned model's ability to predict Ising data independent of the training set ("test data") as a function of its complexity for different amounts of training data, using two error metrics.

Chapter 2

Conservative tightly-coupled simulations of stochastic multiscale systems

2.1 Abstract

Multiphysics problems often involve components whose macroscopic dynamics is driven by microscopic random fluctuations. The fidelity of simulations of such systems depends on their ability to propagate these random fluctuations throughout a computational domain, including subdomains represented by deterministic solvers. When the constituent processes take place in nonoverlapping subdomains, system behavior can be modeled via a domain-decomposition approach that couples separate components at the interfaces between these subdomains. Its coupling algorithm has to maintain a stable and efficient numerical time integration even at high noise strength. We propose a conservative domain-decomposition algorithm in which tight coupling is achieved by employing either Picard's or Newton's iterative method. Coupled diffusion equations,

one of which has a Gaussian white-noise source term, provide a computational testbed for analysis of these two coupling strategies. Fully-converged ("implicit") coupling with Newton's method typically outperforms its Picard counterpart, especially at high noise levels. This is because the number of Newton iterations scales linearly with the amplitude of the Gaussian noise, while the number of Picard iterations can scale superlinearly. At large time intervals between two subsequent inter-solver communications, the solution error for single-iteration ("explicit") Picard's coupling can be several orders of magnitude higher than that for implicit coupling. Increasing the explicit coupling's communication frequency reduces this difference, but the resulting increase in computational cost can make it less efficient than implicit coupling at similar levels of solution error, depending on the communication frequency of the latter and the noise strength. This trend carries over into higher dimensions, although at high noise strength explicit coupling may be the only computationally viable option.

2.2 Introduction

Many, if not most, problems of practical importance deal with complex systems that involve multiple physical (as well as chemical and biological) processes, which occur on a wide range of spatial and/or temporal scales. These processes can either spatially coexist or occur in adjacent regions of space. We focus on the latter class of multiphysics phenomena, in which different processes take place in separate spatial domains and affect each other at the interfaces between these domains. Conjugate heat transfer across a fluid-solid interface [95] is an illustrative example of such phenomena. It is central to applications as diverse as satellite cold gas propulsion systems [78] and spacecraft re-entry into Earth's atmosphere [64].

Following the terminology established in the field of fluid-structure interactions

(FSI), one can subdivide solution strategies for interfacially coupled multiphysics systems into two modeling frameworks: "monolithic" [20] and "component partitioning" [44]. The former combines all the different physics components and their interactions into a single discrete operator, which is then advanced in time. This "tight coupling" ensures temporal synchronization of all the state variables and hence possesses excellent robustness, accuracy and stability properties. However, it is computationally demanding and "intrusive", i.e., requires development of new codes. The second framework, which is also known as domain decomposition (DD), advances solutions of each physics component independently from the others, using additional solvers to exchange information at the interfaces through a coupling algorithm. It is "nonintrusive", i.e., allows for a "black-box" implementation of the physics components which can be done with existing ("legacy") codes. This operational expediency comes at a cost of reduced accuracy and stability when the physics components involved are "loosely coupled", leading to desynchronization of the state variables in the different components by one time step or a fraction of a time step [75]. Iterative coupling techniques can be used to achieve a tight coupling, which eliminates this time shift [43, 42, 75].

Despite the widespread use of DD approaches, there is a dearth of systematic studies of their numerical properties. Most studies deal with the coupling of deterministic components, which are typically represented by deterministic partial differential equations (PDEs). Representative examples include an analysis of the stability of an interfacial coupling in one-dimensional fluid-structure thermal diffusion [50], an analysis of predictor-corrector staggered schemes for simulating FSI [75], an investigation of the stability of a coupling algorithm based on mixed interface conditions for conjugate heat transfer simulations [41], and a demonstration of the effects of a non-converged iterative coupling on the stability of a coupled linear diffusion problem [102]. These and other similar studies have led to nontrivial conclusions, which are likely to be problem-specific

and demonstrate the algorithmic complexity of coupling nonlinear solvers. For example, an otherwise unstable loose coupling used in FSI simulations can be made stable by enforcing Neumann boundary conditions for the structural calculation and Dirichlet boundary conditions for the fluid solver [50]; and standard staggered schemes for FSI simulations need to be modified by several iteratively made corrector steps to ensure conservation of energy [43, 42, 75].

When random fluctuations are generated by one of the constituent solvers, conclusions drawn from numerical studies of fully-deterministic systems may need to be modified. Currently, a systematic analysis of how random noise or stochasticity of one of the constituent solvers affects the numerical performance of both the other (possibly deterministic) solvers and an algorithm used to couple them is largely missing. Such studies are needed to gain confidence in the ever-growing number of multiphysics and hybrid simulations that combine deterministic and stochastic solvers [119, 45, 32, 61]. The analysis presented below contributes to this area of research by studying the effects of random noise on numerical properties (coupling convergence, stability and accuracy) of a domain-decomposition algorithm which tightly couples a deterministic and stochastic subdomain solver. A complementary challenge, the need for adding a random source term to a (deterministic) PDE solver coupled to a stochastic solver whose microscopic fluctuations drive the macroscopic system dynamics (e.g., in highly nonlinear problems involving phase transitions), has been addressed in [5, 16, 121, 113].

In Section 2.3 we formulate a computational testbed problem, one-dimensional diffusion in a composite material one segment of which contains a Gaussian white-noise forcing. Section 2.4 contains a description of our DD approach to solving this problem, which tightly couples the deterministic (explicit Euler) and stochastic (Euler-Maruyama) diffusion solvers using Newton's or Picard's iteration. Section 2.5 presents a stability analysis of our algorithm using fully-converged Picard's iteration. In Section 2.6 we

conduct a series of numerical experiments to explore the performance of our algorithm. These findings are summarized in Section 2.7.

2.3 Problem Formulation

Consider a one-dimensional linear diffusion equation,

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left[D\frac{\partial \rho}{\partial x}\right] + f, \qquad x \in \Omega \equiv (-L/2, L/2), \quad t > 0, \tag{2.1a}$$

which describes the evolution of concentration $\rho(x,t)$ in space, x, and time, t. The diffusion coefficient $D(x)$ is piecewise constant,

$$D(x) = \begin{cases} D_1 & \text{for } x \in \Omega_1 \equiv (-L/2, 0) \\ D_2 \gg D_1 & \text{for } x \in \Omega_2 \equiv [0, L/2), \end{cases} \tag{2.1b}$$

and the source term $f(x,t)$ is defined as

$$f(x,t) = \begin{cases} 0 & \text{for } x \in \Omega_1 \\ \xi(x,t) & \text{for } x \in \Omega_2, \end{cases} \tag{2.1c}$$

where $\xi(x,t)$ is a zero-mean Gaussian space-time white noise with covariance

$$\mathbb{E}[\xi(x,t)\xi(y,\tau)] = \sigma_\xi^2 \delta(x-y)\delta(t-\tau), \quad x,y \in \Omega_2; \quad t,\tau > 0 \tag{2.1d}$$

and variance σ_ξ^2. Here $\delta(\cdot)$ denotes the Dirac delta function. Equation (2.1) is subject to Dirichlet boundary conditions

$$\rho(x = -L/2, t) = \rho_L, \qquad \rho(x = L/2, t) = \rho_R, \tag{2.2}$$

and an initial condition

$$\rho(x,0) = \rho_M \begin{cases} (\rho_L/\rho_M)^{-2x/L} & \text{for } x \in \Omega_1 \\ (\rho_R/\rho_M)^{2x/L} & \text{for } x \in \Omega_2. \end{cases} \tag{2.3}$$

The simulation domain's length, L; the concentration on the left, ρ_L, and right, ρ_R, boundaries; the initial concentration at $x = 0$, ρ_M; and the noise variance, σ_ξ^2, are given constants.

Numerical solution of this boundary-value problem (BVP) is nontrivial. First, the presence of stochastic noise $\xi(x,t)$ in the right half of the simulation domain, Ω_2, formally renders a solution $\rho(x,t)$ random over the whole domain Ω. Second, the vastly different diffusion coefficients D_1 and D_2 imply the co-existence of two disparate time scales t_1 and t_2 ($t_1 \gg t_2$). Therefore, advancing the stochastic BVP (2.1)–(2.3) in time over the whole domain Ω requires the use of a time step whose size is determined by the smallest diffusion time-scale, t_2.

Domain decomposition provides a natural alternative to solving the stochastic BVP (2.1)–(2.3) directly. It decomposes the computational domain Ω into subdomains Ω_1 and Ω_2. A different BVP is defined on each of these subdomains, such that

$$\frac{\partial \rho_1}{\partial t} = D_1 \frac{\partial^2 \rho_1}{\partial x^2}, \quad \rho_1(-L/2,t) = \rho_L, \quad \rho_1(x,0) = \rho_M \left(\frac{\rho_L}{\rho_M} \right)^{-2x/L}, \quad x \in \Omega_1 \tag{2.4}$$

and

$$\frac{\partial \rho_2}{\partial t} = D_2 \frac{\partial^2 \rho_2}{\partial x^2} + \xi, \quad \rho_2(L/2,t) = \rho_R, \quad \rho_2(x,0) = \rho_M \left(\frac{\rho_R}{\rho_M} \right)^{2x/L}, \quad x \in \Omega_2. \tag{2.5}$$

These two BVPs are coupled by enforcing the continuity of the state variable, ρ, and its

flux at the interface $x = 0$ separating the subdomains Ω_1 and Ω_2,

$$\rho_1(0,t) = \rho_2(0,t), \qquad D_1 \frac{\partial \rho_1}{\partial x}(0,t) = D_2 \frac{\partial \rho_2}{\partial x}(0,t). \tag{2.6}$$

This relatively simple computational testbed contains a number of salient features of multiphysics simulations. First, it combines deterministic and stochastic solvers used to integrate BVPs (2.4) and (2.5), respectively. Second, its constitutive solvers operate at different temporal scales defined by the diffusion coefficients D_1 and D_2. (A more complicated example of two-dimensional diffusion is presented in Section 2.6.4.)

2.4 Numerical Implementation of Domain Decomposition

Our quantity of interest is temporal snapshots of the ensemble-averaged concentration profile, $\langle \rho(x,t) \rangle$, in $\Omega = (-L/2, L/2)$ over a time interval $(0,T]$, where T is defined by the diffusion time scale of the slowest diffusion process and is set to $T = L^2/(8D_1)$. This choice of T allows $\langle \rho(x,t) \rangle$ to approach its steady-state limit for all $x \in \Omega$.

2.4.1 Spatial discretization of the computational domain

To simplify the presentation, and without any loss of generality, we discretize the computational domain Ω using a uniform mesh of cell size $\Delta x = L/N$, where N is the total number of grid cells. The solvers used to integrate BVPs (2.4) and (2.5) employ a staggered grid approach, in which diffusive fluxes, $F_i(x,t) = -D_i \partial \rho_i / \partial x$, in the ith subdomain ($i = 1,2$) are calculated at the midpoint between two subsequent concentration nodes (see Fig. 2.1). In particular, both $\rho^- = \rho_1(0,t)$ and $\rho^+ = \rho_2(0,t)$ are defined at the interfacial node $x = 0$, while the corresponding interfacial fluxes $F^- = F_1(0,t)$ and

$F^+ = F_2(0,t)$ are defined at $x = -\Delta x/2$ and $x = \Delta x/2$, respectively.

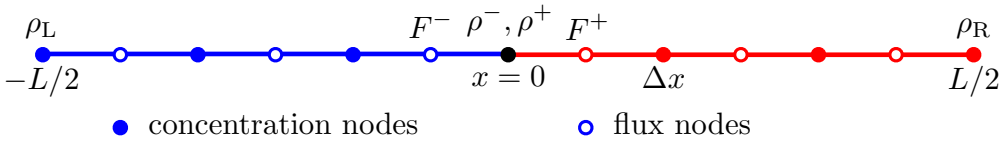

Figure 2.1: Domain $\Omega = (-L/2, L/2)$ decomposed into subdomains $\Omega_1 = (-L/2, 0)$ and $\Omega_2 = [0, L/2)$. Concentration ρ is computed at the nodes denoted by solid circles, and flux F is evaluated at the midpoint between two subsequent concentration nodes (open circles).

2.4.2 Numerical solvers for BVPs (2.4) and (2.5)

Approximation of the Laplacian with a second-order central finite difference scheme transforms (2.4) and (2.5) into systems of ordinary differential equations (ODEs)

$$\frac{\mathrm{d}\boldsymbol{\rho}_1}{\mathrm{d}t} = \mathbf{f}_1(\boldsymbol{\rho}_1), \qquad \frac{\mathrm{d}\boldsymbol{\rho}_2}{\mathrm{d}t} = \mathbf{f}_2(\boldsymbol{\rho}_2) + \boldsymbol{\xi}. \tag{2.7a}$$

Here $\boldsymbol{\rho}_i = (\rho_{i,1}, \ldots, \rho_{i,N/2-1})^\top$ are the one-dimensional arrays of size $N/2-1$ of the nodal values of the state variables $\rho_i(x,t)$ for $i = 1,2$; components $f_{i,p}$ of the one-dimensional arrays $\mathbf{f}_i(\boldsymbol{\rho}_i)$ of size $N/2-1$ are defined by

$$f_{i,p} = D_i \frac{\rho_{i,p+1} + \rho_{i,p-1} - 2\rho_{i,p}}{\Delta x^2}, \qquad p = 1, \ldots, N/2-1, \qquad i = 1,2; \tag{2.7b}$$

where $\rho_{1,0} = \rho_{\mathrm{L}}$, $\rho_{1,N/2} = \rho^-$, $\rho_{2,0} = \rho^+$ and $\rho_{2,N/2} = \rho_{\mathrm{R}}$; and components ξ_p ($p = 1, \ldots, N/2-1$) of the one-dimensional array $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{N/2-1})^\top$ are obtained by spatial discretization of the space-time white noise $\xi(x,t)$, and satisfy

$$\mathbb{E}[\xi_p(t)] = 0, \qquad \mathbb{E}[\xi_p(t)\xi_q(\tau)] = \sigma_\xi^2 \frac{\delta_{pq}}{\Delta x} \delta(t-\tau), \tag{2.7c}$$

where δ_{pq} is the Kronecker delta function.

A deterministic solver used to advance $\boldsymbol{\rho}_1$ in time employs an explicit Euler (EE) method with time step Δt_1. A stochastic solver used to advance $\boldsymbol{\rho}_2$ in time employs the Euler-Maruyama (EM) method [69, 79] with time step Δt_2. The latter advances the pth component of the random array $\boldsymbol{\rho}_2$ from $t_n = n\Delta t_2$ to $t_{n+1} = (n+1)\Delta t_2$ according to

$$\rho_{2,p}^{n+1} = \rho_{2,p}^{n} + \frac{D_2 \Delta t_2}{\Delta x^2}(\rho_{2,p+1}^{n} + \rho_{2,p-1}^{n} - 2\rho_{2,p}^{n}) + \sigma_\xi \sqrt{\frac{\Delta t_2}{\Delta x}}\, \eta_p^n, \tag{2.8}$$

where η_p^n are identically distributed standard Gaussian variables such that η_p^n and η_q^m are mutually independent for all $p \neq q$ and/or $n \neq m$.

Given a value of the interfacial ($x = 0$) concentration $\rho^-(t) = \rho^+(t)$ at a certain time t, these two solvers can operate independently from each other, yielding a deterministic solution $\boldsymbol{\rho}_1$ and a stochastic solution $\boldsymbol{\rho}_2$. These solutions will not satisfy the continuity conditions (2.6) and, hence, do not yield a solution of the original BVP (2.1). Construction of such a solution requires occasional communications between the deterministic and stochastic solvers via a coupling algorithm.

2.4.3 Interfacial coupling algorithm

Exchange of information between the two solvers can, at most, occur on the scale of the largest time step. Let Δt_{com} denote the time interval between any two successive communications between the deterministic and stochastic solvers. It is given in terms of the multiples of the inner-solver time steps Δt_1 and Δt_2, such that $\Delta t_{\text{com}} = n_1 \Delta t_1 = n_2 \Delta t_2$. In other words, the deterministic and stochastic solvers are advanced by n_1 and n_2 "micro" steps Δt_1 and Δt_2, respectively, before the inter-solver communication occurs and concentration and flux information is exchanged between them. This procedure is repeated until convergence when the coupling is iterative ("implicit"), or occurs only

once when it is noniterative ("explicit"). Either way, it advances the system state by one "macro" step Δt_{com}. In the former case, we consider successive communications during the iterations of a certain macro-step as one "overall" communication, and define "subsequent" inter-solver communications as the information exchanges associated with subsequent macro-steps.

In the coupling algorithm described below, we use Δt_{com}-averaged interfacial concentrations and fluxes. The time-averaged concentrations, $\bar{\boldsymbol{\rho}}_1$ and $\bar{\boldsymbol{\rho}}_2$, are computed as arithmetic means of $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ over their respective n_1 and n_2 micro-steps. The interfacial values of these Δt_{com}-averaged concentrations are $\bar{\rho}^-$ and $\bar{\rho}^+$. Likewise, we denote by $\bar{F}^-$ and $\bar{F}^+$ the Δt_{com}-averaged values of the interfacial fluxes F^- and F^+, respectively. Reliance on the Δt_{com}-averaged interfacial concentrations and fluxes, rather than on their counterparts computed at the last micro-step of each solver, facilitates construction of a mass-conservative coupling algorithm (see A).

To tightly couple BVPs (2.4) and (2.5), we enforce (2.6) at each inter-solver communication by means of an iterative (or "implicit") coupling algorithm based on either Picard's or Newton's method (B). In the context of the interfacial conditions (2.6), these root-finding algorithms are deployed to solve a system of coupled nonlinear algebraic equations, $\bar{\rho}^- = \bar{\rho}^+$ and $\bar{F}^- = \bar{F}^+$, at each inter-solver communication. Using notation

$$\bar{\rho}^n_{1,N/2} = \bar{\rho}^-, \quad \bar{\rho}^n_{2,0} = \bar{\rho}^+, \quad \bar{F}^n_{1,N/2-1/2} = \bar{F}^-, \quad \bar{F}^n_{2,1/2} = \bar{F}^+, \tag{2.9}$$

for the macro-step from t_n to $t_{n+1} = t_n + \Delta t_{\text{com}}$, this system is written as

$$\bar{\rho}^n_{1,N/2} = \bar{\rho}^n_{2,0}, \qquad \bar{F}^n_{1,N/2-1/2} = \bar{F}^n_{2,1/2}. \tag{2.10}$$

Picard's method recasts (2.10) into a fixed-point iteration problem

$$\bar{\rho}_{1,N/2}^{n,k+1} = [\bar{\rho}_{2,0}^{n}(\bar{F}_{2,1/2}^{n})]^{k}, \qquad \bar{F}_{2,1/2}^{n,k+1} = [\bar{F}_{1,N/2-1/2}^{n}(\bar{\rho}_{1,N/2}^{n})]^{k} \tag{2.11a}$$

where k is the iteration number, and

$$[\bar{\rho}_{2,0}^{n}(\bar{F}_{2,1/2}^{n})]^{k} \equiv \bar{\rho}_{2,1}^{n,k} + \frac{\Delta x}{D_2}\bar{F}_{2,1/2}^{n,k}, \qquad \bar{\rho}_{2,0}^{n,k} = [\bar{\rho}_{2,0}^{n}]^{k}, \quad \text{etc.}, \tag{2.11b}$$

$$[\bar{F}_{1,N/2-1/2}^{n}(\bar{\rho}_{1,N/2}^{n})]^{k} \equiv -D_1\frac{\bar{\rho}_{1,N/2}^{n,k} - \bar{\rho}_{1,N/2-1}^{n,k}}{\Delta x}. \tag{2.11c}$$

The iterations continue until

$$\max\left\{\left|\bar{\rho}_{1,N/2}^{n,k} - \bar{\rho}_{2,0}^{n,k}\right|, \left|\bar{F}_{1,N/2-1/2}^{n,k} - \bar{F}_{2,1/2}^{n,k}\right|\right\} \leq \varepsilon, \tag{2.12}$$

where ε is the prescribed tolerance.

Newton's method replaces (2.10) with an iterative system

$$\begin{pmatrix} \bar{\rho}_{1,N/2}^{n,k+1} \\ \\ \bar{F}_{2,1/2}^{n,k+1} \end{pmatrix} = \begin{pmatrix} \bar{\rho}_{1,N/2}^{n,k} \\ \\ \bar{F}_{2,1/2}^{n,k} \end{pmatrix} - \mathbf{J}^{-1}(\bar{\rho}_{1,N/2}^{n,k}, \bar{F}_{2,1/2}^{n,k}) \begin{pmatrix} g_1 \\ \\ g_2 \end{pmatrix} \tag{2.13a}$$

where $\mathbf{J}$ is the Jacobian,

$$\mathbf{J} = \begin{pmatrix} \partial g_1/\partial\bar{\rho}_{1,N/2}^{n,k} & \partial g_1/\partial\bar{F}_{2,1/2}^{n,k} \\ \\ \partial g_2/\partial\bar{\rho}_{1,N/2}^{n,k} & \partial g_2/\partial\bar{F}_{2,1/2}^{n,k} \end{pmatrix}, \tag{2.13b}$$

and

$$g_1 = \bar{\rho}_{1,N/2}^{n,k} - \bar{\rho}_{2,1}^{n,k} - \frac{\Delta x}{D_2}\bar{F}_{2,1/2}^{n,k}, \quad g_2 = -D_1\frac{\bar{\rho}_{1,N/2}^{n,k} - \bar{\rho}_{1,N/2-1}^{n,k}}{\Delta x} - \bar{F}_{2,1/2}^{n,k}. \tag{2.13c}$$

Explicit expressions for the components of the Jacobian $\mathbf{J}$ are derived in C. The iterations continue until $\max\{|g_1|, |g_2|\} \leq \varepsilon$.

2.4.4 Domain-decomposition algorithm

Let $\boldsymbol{\rho}_1^{n,l,k} \equiv \boldsymbol{\rho}_1(t_n + l\Delta t_1)$ and $\boldsymbol{\rho}_2^{n,m,k} = \boldsymbol{\rho}_2(t_n + m\Delta t_2)$ denote arrays of the nodal concentrations at inner-solver times $t_n + l\Delta t_1$ and $t_n + m\Delta t_2$ during the kth iteration of the macro-step from t_n to $t_{n+1} = t_n + \Delta t_{\text{com}}$. As before, $\bar{\rho}_{1,N/2}^{n,k}$ and $\bar{F}_{2,1/2}^{n,k}$ denote the Δt_{com}-averaged interfacial concentration and flux during the kth iteration of that macro-step. The solution is advanced from t_n to t_{n+1} as follows.

1. Initialization step. Set $\bar{\rho}_{1,N/2}^{n,0} = \rho_{1,N/2}(t_n)$ and $\bar{F}_{2,1/2}^{n,0} = F_{2,1/2}(t_n)$.

2. Evolve the state vector $\boldsymbol{\rho}_1^{n,0,k}$ of size $N/2-1$ to $\boldsymbol{\rho}_1^{n,n_1,k}$ over n_1 micro-steps, using ρ_{L} and $\bar{\rho}_{1,N/2}^{n,k}$ as the boundary conditions at $x = -L/2$ and $x = 0$, respectively.

3. Evolve the state vector $\boldsymbol{\rho}_2^{n,0,k}$ of size $N/2-1$ to $\boldsymbol{\rho}_2^{n,n_2,k}$ over n_2 micro-steps, using $\bar{\rho}_{1,N/2}^{n,k}$ and ρ_{R} as the boundary conditions at $x = 0$ and $x = L/2$, respectively.

4. Use either Picard's or Newton's coupling to calculate new iterates of the interfacial concentration, $\bar{\rho}_{1,N/2}^{n,k+1}$, and flux, $\bar{F}_{2,1/2}^{n,k+1}$.

5. Repeat steps 2 through 4 until the given tolerance ε is achieved.

6. Advance the solution by one macro-step by setting

$$\rho_{1,N/2}(t_{n+1}) = \bar{\rho}_{1,N/2}^{n,K} \quad \text{and} \quad F_{2,1/2}(t_{n+1}) = \bar{F}_{2,1/2}^{n,K},$$

where $K = K(n)$ indicates the number of iterations at convergence. By construction, $\rho_{2,0}(t_{n+1}) = \rho_{1,N/2}(t_{n+1})$ and $F_{1,N/2-1/2}(t_{n+1}) = F_{2,1/2}(t_{n+1})$.

It is worthwhile noting that the above iterative algorithms can be readily modified by using $\bar{\rho}_{2,0}^{n,k}$ and $\bar{F}_{1,N/2-1/2}^{n,k}$ as iterates.

2.5 Stability of DD Algorithm with Picard's Coupling

One micro-step of the deterministic ($l = 0, \ldots, n_1 - 1$) and stochastic ($m = 0, \ldots, n_2 - 1$) solvers, during the macro-step from t_n to t_{n+1}, is carried out, respectively, by the explicit Euler and Euler-Maruyama (2.8) methods,

$$\boldsymbol{\rho}_1^{n,l+1,k} = (\mathbf{I} + \mathbf{A}_1)\boldsymbol{\rho}_1^{n,l,k} + T_1 \boldsymbol{\rho}_{1,\mathrm{b}}^{n,k}, \qquad \boldsymbol{\rho}_{1,\mathrm{b}}^{n,k} \equiv (\rho_{\mathrm{L}}, 0, \ldots 0, \bar{\rho}_{1,N/2}^{n,k})^{\top} \tag{2.14a}$$

and

$$\boldsymbol{\rho}_2^{n,m+1,k} = (\mathbf{I} + \mathbf{A}_2)\boldsymbol{\rho}_2^{n,m,k} + T_2 \boldsymbol{\rho}_{2,\mathrm{b}}^{n,k} + \boldsymbol{\gamma}\boldsymbol{\eta}^{n,m,k}, \quad \boldsymbol{\rho}_{2,\mathrm{b}}^{n,k} \equiv (\bar{\rho}_{1,N/2}^{n,k}, 0, \ldots, 0, \rho_{\mathrm{R}})^{\top}. \tag{2.14b}$$

Here $\boldsymbol{\rho}_{1,\mathrm{b}}^{n,k}$ and $\boldsymbol{\rho}_{2,\mathrm{b}}^{n,k}$ are the vectors of size $N/2-1$ supplying the boundary conditions for the two solvers; $\mathbf{I}$ is the $(N/2-1) \times (N/2-1)$ identity matrix; the square matrices $\mathbf{A}_i$ and $\mathbf{T}_i$ ($i = 1, 2$) of size $N/2-1$ are defined by

$$\mathbf{A}_i = \frac{D_i \Delta t_i}{\Delta x^2} \mathbf{Trid}(1, -2, 1), \qquad \mathbf{T}_i = \frac{D_i \Delta t_i}{\Delta x^2} \mathbf{I}, \qquad i = 1, 2 \tag{2.14c}$$

with $\mathbf{Trid}(1, -2, 1)$ denoting a square tridiagonal matrix of size $N/2-1$, whose diagonal elements are -2 and sub- and super-diagonal elements are 1; $\gamma \equiv \sigma_\xi \sqrt{\Delta t_2 / \Delta x}$; and $\boldsymbol{\eta}^{n,m,k}$ is the vector of size $N/2-1$, whose components are independent identically distributed

standard Gaussian variables. After n_1 micro-steps, the deterministic solver (2.14a) yields

$$\boldsymbol{\rho}_1^{n,n_1,k} = \mathbf{B}_{\mathrm{L}}\boldsymbol{\rho}_1^{n,0} + \mathbf{C}_{\mathrm{L}}\boldsymbol{\rho}_{1,\mathrm{b}}^{n,k}, \tag{2.15a}$$

where $\boldsymbol{\rho}_1^{n,0} \equiv \boldsymbol{\rho}_1(t = t_n)$ and

$$\mathbf{B}_{\mathrm{L}} = (\mathbf{I} + \mathbf{A}_1)^{n_1}, \quad \mathbf{C}_{\mathrm{L}} = \sum_{l=0}^{n_1-1} (\mathbf{I} + \mathbf{A}_1)^l \, \mathbf{T}_1. \tag{2.15b}$$

After n_2 micro-steps, the stochastic solver (2.14b) gives

$$\boldsymbol{\rho}_2^{n,n_2,k} = \mathbf{B}_{\mathrm{R}}\boldsymbol{\rho}_2^{n,0} + \mathbf{C}_{\mathrm{R}}\boldsymbol{\rho}_{2,\mathrm{b}}^{n,k} + \gamma \sum_{m=0}^{n_2-1} (\mathbf{I} + \mathbf{A}_2)^{n_2-1-m} \, \boldsymbol{\eta}^{n,m,k}, \tag{2.16a}$$

where $\boldsymbol{\rho}_2^{n,0} \equiv \boldsymbol{\rho}_2(t = t_n)$ and

$$\mathbf{B}_{\mathrm{R}} = (\mathbf{I} + \mathbf{A}_2)^{n_2}, \qquad \mathbf{C}_{\mathrm{R}} = \sum_{m=0}^{n_2-1} (\mathbf{I} + \mathbf{A}_2)^m \, \mathbf{T}_2. \tag{2.16b}$$

Let us define vectors of size N

$$\mathbf{x}^{n,k} = (\boldsymbol{\rho}_1^{n,n_1,k}, \bar{\rho}_{1,N/2}^{n,k}, \bar{F}_{2,1/2}^{n,k}, \boldsymbol{\rho}_2^{n,n_2,k})^\top, \quad \mathbf{x}^{n} = (\boldsymbol{\rho}_1^{n,0}, \bar{\rho}_{1,N/2}^{n,0}, \bar{F}_{2,1/2}^{n,0}, \boldsymbol{\rho}_2^{n,0})^\top. \tag{2.17}$$

We show in C that $\mathbf{x}^{n,k}$ satisfies a recursive relation

$$\mathbf{x}^{n,k+1} = \mathbf{M}\mathbf{x}^{n,k} + \mathbf{P}\mathbf{x}^{n} + \mathbf{d}^{n,k,k+1} + \mathbf{e}, \tag{2.18a}$$

where $\mathbf{M}$ and $\mathbf{P}$ are $N \times N$ matrices,

$$\mathbf{M} = \begin{pmatrix} \mathbf{0}_{(N/2-1)\times(N/2-1)} & \mathbf{r}_{(N/2-1)\times 1} & \mathbf{s}_{(N/2-1)\times 1} & \mathbf{0}_{(N/2-1)\times(N/2-1)} \\ \mathbf{0}_{1\times(N/2-1)} & u & v & \mathbf{0}_{1\times(N/2-1)} \\ \mathbf{0}_{1\times(N/2-1)} & w & 0 & \mathbf{0}_{1\times(N/2-1)} \\ \mathbf{0}_{(N/2-1)\times(N/2-1)} & \mathbf{y}_{(N/2-1)\times 1} & \mathbf{z}_{(N/2-1)\times 1} & \mathbf{0}_{(N/2-1)\times(N/2-1)} \end{pmatrix} \tag{2.18b}$$

and

$$\mathbf{P} = \begin{pmatrix} \mathbf{B}_{\mathrm{L}} & \mathbf{0}_{(N/2-1)\times 1} & \mathbf{0}_{(N/2-1)\times 1} & \mathbf{S}_{(N/2-1)\times(N/2-1)} \\ \mathbf{0}_{1\times(N/2-1)} & 0 & 0 & \mathbf{u}_{1\times(N/2-1)} \\ \mathbf{v}_{1\times(N/2-1)} & 0 & 0 & \mathbf{0}_{1\times(N/2-1)} \\ \mathbf{0}_{(N/2-1)\times(N/2-1)} & \mathbf{0}_{(N/2-1)\times 1} & \mathbf{0}_{(N/2-1)\times 1} & \mathbf{W}_{(N/2-1)\times(N/2-1)} \end{pmatrix}, \tag{2.18c}$$

$\mathbf{d}^{n,k,k+1}$ is a vector of size N that depends on the noise, and $\mathbf{e}$ is a constant vector of size N independent of n and k. The definitions of $\mathbf{d}^{n,k,k+1}$ and $\mathbf{e}$ are provided in C. The size of sub-matrices of $\mathbf{M}$ and $\mathbf{P}$ is denoted by their subscripts, and their respective components are defined in C. At convergence, (2.18) becomes

$$\mathbf{x}^{n+1} = (\mathbf{I}-\mathbf{M})^{-1}\mathbf{P}\mathbf{x}^{n} + (\mathbf{I}-\mathbf{M})^{-1}\mathbf{d}^{n,K} + (\mathbf{I}-\mathbf{M})^{-1}\mathbf{e}, \tag{2.19}$$

where $\mathbf{d}^{n,K}$ is the value of $\mathbf{d}^{n,k,k+1}$ obtained when the iterations for the macro-step from t_n to t_{n+1} have converged. Taking the ensemble average of (2.19) yields

$$\langle \mathbf{x}^{n+1} \rangle = (\mathbf{I}-\mathbf{M})^{-1}\mathbf{P}\langle \mathbf{x}^{n} \rangle + (\mathbf{I}-\mathbf{M})^{-1}\mathbf{e}, \tag{2.20}$$

where we have used the fact that $\mathbf{d}^{n,K}$ is a zero-mean quantity (see C).

Relation (2.18) reveals that the iterations for a given macro-step (i.e., from time

t_n to time t_{n+1} for any n) converge if the spectral radius of $\mathbf{M}$, $\rho[\mathbf{M}]$, is less than one. Relation (2.20) demonstrates that the overall time advancement is numerically stable in ensemble mean if $\rho[(\mathbf{I}-\mathbf{M})^{-1}\mathbf{P}] < 1$. Both results are identical to those obtained when simulating BVP (2.1)–(2.3) with $f(x,t) \equiv 0$, and hence are independent of the presence of random noise in the stochastic solver. Therefore, if the fully-deterministic DD algorithm is stable for a certain combination of values for D_1, D_2, Δx, Δt_1, Δt_2, n_1 and n_2 (and hence $\Delta t_{\text{com}} = n_1 \Delta t_1$), then the corresponding deterministic-stochastic DD algorithm is also stable. This result is in line with the fact that stability of the Euler-Maruyama algorithm for a linear stochastic differential equation with additive noise follows from the stability of its deterministic counterpart, the explicit Euler method [87].

2.6 Simulation Results and Discussion

In the simulations reported below, the one-dimensional domain of length $L = 20.0$ is discretized into $N = 20$ intervals of length $\Delta x = 1.0$. Constant concentrations $\rho_{\text{L}} = 15.0$ and $\rho_{\text{R}} = 5.0$ are prescribed on the left ($x = -L/2$) and right ($x = L/2$) boundaries, respectively, and $\rho_{\text{M}} = 10.0$. The diffusion coefficients in the right and left halves of the computational domain are set to $D_2 = 10.0$ and either $D_1 = 1.0$ or $D_1 = 0.1$, respectively. Unless noted otherwise, the convergence tolerance for Newton and Picard iterations is $\varepsilon = 10^{-3}$. These and other values of all the physical quantities are reported in consistent units.

The presence of random noise in the stochastic solver renders a solution of BVP (2.1)–(2.3) random as well. Hence the solution is given in terms of a probability density function of the system state $\rho(x,t)$ or its ensemble moments, such as ensemble mean $\langle \rho(x,t) \rangle$ and variance $\sigma_\rho^2(x,t)$. These statistical moments are approximated by their sample counterparts computed from a finite number of independent samples. This

number has to be sufficiently large for the difference between the ensemble and sample moments not to exceed a specified tolerance.

At each discrete time t_d, to compute (2.8) we generate independent zero-mean Gaussian variables $\zeta_p = \sigma_\xi \sqrt{\Delta t_2/\Delta x}\, \eta_p$ ($p = 1, \ldots, N/2-1$) with ensemble variance $\sigma_\xi^2 \Delta t_2/\Delta x$. Hence, to obtain an estimate of the required number of samples (N_{sam}) for the sample mean and variance of $\rho(x_p, t_d)$ to approximate $\langle \rho(x_p, t_d) \rangle$ and $\sigma_\rho^2(x_p, t_d)$, respectively, within an acceptable margin of error, we use the following heuristic procedure.

1. For each p, start generating samples of ζ_p and calculate a running sample mean and sample variance as the number of samples N increases.

2. When N is such that, for all p, the N-sample mean and variance of ζ_p do not deviate more than a tolerance $\varepsilon = 10^{-2}$ from their respective ensemble counterparts 0 and $\sigma_\xi^2 \Delta t_2/\Delta x$, stop and set $N_{\text{sam}} = N$.

In order to use the same value of N_{sam} for a given σ_ξ across all of our numerical experiments, and given that in each of them $\Delta t_2/\Delta x < 1.0$ by construction, we apply the above procedure for $\Delta t_2/\Delta x = 1.0$. (Recall that the required number of samples increases with the noise's variance). We find that a sufficient number of samples for $\sigma_\xi = 0.1$, 0.2, and 0.4 (the noise strength we consider) is $N_{\text{sam}} = 500$, 1500, and 4000, respectively.

2.6.1 Temporal order of accuracy of implicitly coupled DD algorithm

We first investigate the temporal order of accuracy our DD algorithm by calculating the l^2-norm error $\mathcal{E}_{l^2}$ over the entire domain Ω between an N_{sam}-averaged, implicitly coupled EE-EM solution $\tilde{\rho}$, obtained with our time advancement scheme, and the ensemble average of the exact solution, $\langle \rho \rangle$, to the set of linear ODEs (2.7), resulting from spatially discretizing (2.4)-(2.5), using $\Delta x = 1.0$. The latter is equivalent to the exact

solution of (2.7) with $\boldsymbol{\xi} = 0$, and is approximated by an implicitly (Newton's method with $\varepsilon = 10^{-3}$) coupled EE-EE solution $\rho^{\star}$, obtained with our numerical scheme, with grid cell size $\Delta x = 1.0$, micro-steps $\Delta t_i^{\star} \equiv \Delta t^{\star} = 10^{-6}$ and $n_i = 1$ ($i = 1, 2$).

To calculate $\tilde{\rho}$, we assume equal subdomain micro-steps, $\Delta t_1 = \Delta t_2 \equiv \Delta t$, and $n_1 = n_2 = 1$, and use Newton's coupling with $\varepsilon = 10^{-3}$. (Although not shown here, similar results were obtained for Picard's coupling.) We consider the case of $D_1 = 0.1$ and $D_2 = 10.0$, with $\sigma_{\xi} = 0.4$ (i.e., $N_{\text{sam}} = 4000$).

Figure 2.2 shows that sequential reduction of Δt by a factor of two results in a near-quadratic decrease in $\mathcal{E}_{l^2}$. Repeating the experiment using the corresponding implicitly coupled EE-EE solution yields virtually identical results. These findings indicate that the implicit coupling preserves the second-order local (i.e., first-order global) order of accuracy of the subdomain solvers (for additive noise, the Euler-Maruyama method converges with strong order 1), and this irrespective of the noise strength.

2.6.2 Relative performance of Newton's and Picard's coupling

To investigate the impact of the noise strength on the computational efficiency of the Picard and Newton coupling algorithms, we conduct a series of numerical experiments for the noise amplitude $\sigma_{\xi} = 0.1$, 0.2 or 0.4. We fix $\Delta t_{\text{com}} = 5.0$, and consider two cases: $D_1 = 1.0$ and $D_2 = 10.0$ with $\Delta t_1 = 0.5$ and $\Delta t_2 = 0.05$ (which we will refer to as Test 1), and $D_1 = 0.1$ and $D_2 = 10.0$ with $\Delta t_1 = 5.0$ and $\Delta t_2 = 0.05$ (Test 2).

Table 2.1 exhibits the number of Newton and Picard iterations when approaching steady state, averaged over N_{sam} independent runs. Doubling σ_{ξ} for $D_2 = 10D_1$ nearly doubles the number of Newton iterations necessary to enforce the continuity conditions with tolerance ε. The same procedure applied to Picard's coupling almost quadruples the number of iterations. Newton's coupling outperforms its Picard counterpart, with the gain increasing with σ_{ξ}. For $D_2 = 100D_1$, doubling σ_{ξ} doubles the number of

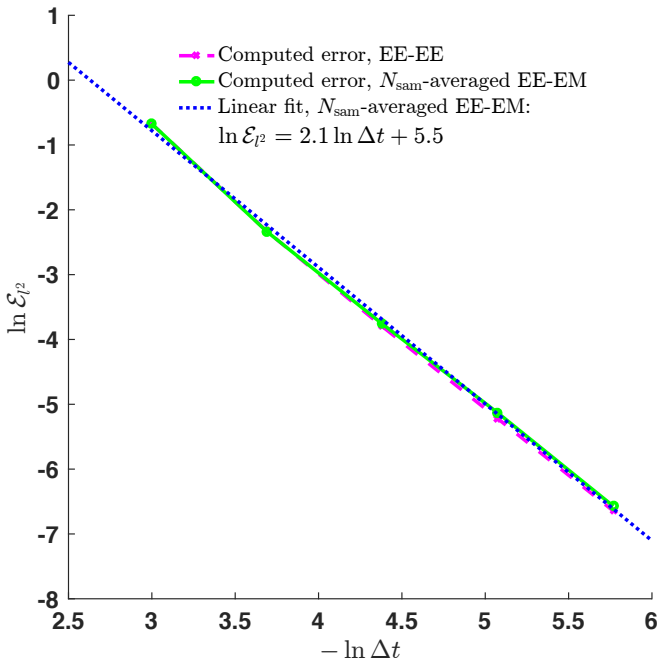

Figure 2.2: The l^2-norm error of the sample-averaged EE-EM solution and the EE-EE solution as a function of micro-step size Δt for the Newton-coupled time advancement.

Table 2.1: N_{sam}-averaged number of iterations $N_{\text{iter,av}}$ for Picard's (Tests 1.1 and 2.1) and Newton's coupling (Tests 1.2 and 2.2) with $\Delta t_{\text{com}} = 5.0$. For Tests 1.1 and 1.2, $D_1 = 1.0$ and $D_2 = 10.0$; for Tests 2.1 and 2.2, $D_1 = 0.1$ and $D_2 = 10.0$.

Test	Number of iterations, $N_{\text{iter,av}}$		
	$\sigma_\xi = 0.1$	$\sigma_\xi = 0.2$	$\sigma_\xi = 0.4$
1.1	31	114	438
1.2	13	25	49
2.1	9	20	68
2.2	12	24	48

Newton iterations at all noise strengths, and doubles the number of Picard iterations at lower strengths but increases them by a factor of about 3.5 at higher noise strengths. Consequently, Picard's coupling slightly outperforms its Newton counterpart at lower noise amplitudes but is significantly outperformed by the latter at higher noise amplitudes. This is shown in Figure 2.3, which illustrates the time evolution (in units of Δt_{com}) of the required number of iterations from $t = 0$ to $t = T/5 = 100.0$ (recall that $0 \leq t \leq T$ with $T = L^2/(8D_1)$, so that $T = 500.0$ for $D_1 = 0.1$). The initial decrease in the number of Picard iterations is absent for Newton's method and further increases the overall computation time for Picard's coupling.

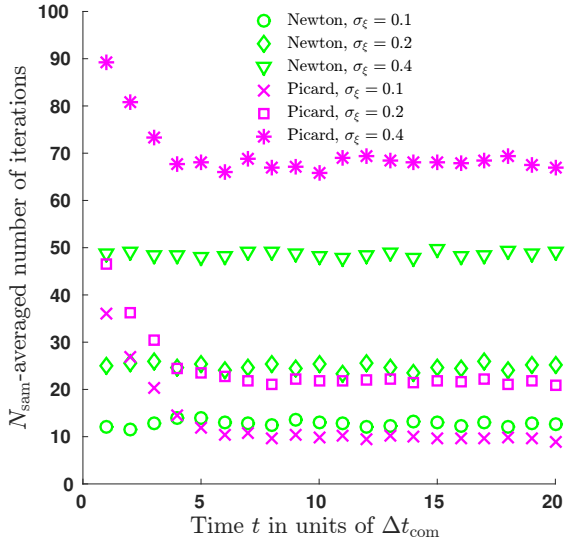

Figure 2.3: Sample-averaged number of Newton and Picard iterations as a function of time (in units of $\Delta t_{\text{com}} = 5.0$) for $D_1 = 0.1$, $D_2 = 10.0$ and several values of the noise strength σ_ξ.

2.6.3 Relative performance of implicit and explicit coupling

In a typical multiphysics simulation, the computational cost of an inter-solver communication may equal or exceed that of constitutive subdomain solvers. Reducing

the frequency of inter-solver communications (increasing Δt_{com}) or, for a given communication frequency, reducing the number of iterations by increasing the tolerance of the iterative coupling (increasing ε), may reduce the total computational cost, but, possibly, at the price of reduced accuracy. We investigate this efficiency/accuracy trade-off by comparing "implicit" coupling, in which the coupling iterations converge to within the prescribed tolerance of $\varepsilon = 10^{-3}$, with "explicit" coupling consisting of a single Picard iteration. This analysis is carried out for two communication times of the implicit coupling, $\Delta t_{\text{com}} = 5.0$ and $\Delta t_{\text{com}} = 50.0$. We focus on the case $D_1 = 0.1$ and $D_2 = 10.0$ with $\Delta t_1 = 1.0$ and $\Delta t_2 = 0.01$, and consider both $\sigma_\xi = 0.1$ and $\sigma_\xi = 0.4$.

The quantities of interest (QoIs) in these experiments are discretized ensemble-averaged concentration profiles $\langle \rho(x_i,t) \rangle$, with $x_i = -L/2 + i\Delta x$ ($i = 0, \ldots, N$), at times $t = 100.0$ and $t = 500.0$. These time points represent the time evolution of the average solution from an early stage of the simulations to the end (steady-state equilibrium). To compute the error of these QoIs, obtained with our time advancement scheme, we compare them with their "exact" counterparts $\langle \rho(x_i,t) \rangle^{\text{ex}}$ computed with implicitly (Newton's iteration with $\varepsilon = 10^{-3}$) coupled explicit Euler solvers on a fine space-time mesh of $\Delta x^{\text{ex}} = \Delta x/2^6 = 0.015625$ and $\Delta t_i^{\text{ex}} = \Delta t_2^{\text{ex}} = 10^{-5}$ (the linearity of BVP (2.1)–(2.3) suggests that the ensemble mean of its solution satisfies the deterministic version of BVP (2.1)–(2.3) in which $f(x,t) \equiv 0$) using $n_1 = n_2 = 1$. The difference between the exact and approximate solutions is reported in terms of a position-dependent relative error, $\mathcal{E}_\rho(x_i,t) = |\langle \rho(x_i,t) \rangle - \langle \rho(x_i,t) \rangle^{\text{ex}}| / \langle \rho(x_i,t) \rangle^{\text{ex}}$.

Figure 2.4 exhibits the relative errors $\mathcal{E}_\rho(x_i,t)$ for $\Delta t_{\text{com}} = 5.0$ (left) and $\Delta t_{\text{com}} = 50.0$ (right); these results were obtained with Newton's coupling algorithm for $\sigma_\xi = 0.1$. (Although not shown here, Picard's coupling was found to yield relative errors of the same magnitude as Newton's coupling.) Table 2.2 shows the N_{sam}-averaged computation times to complete a time trajectory for these and the corresponding $\sigma_\xi = 0.4$ cases. The

time it takes to evaluate the Jacobian **J** in Newton's coupling is excluded from the total, since it is time-independent and therefore computed prior to the transient simulation. The computation times are reported for an Intel Core i7 machine running at 4 GHz.

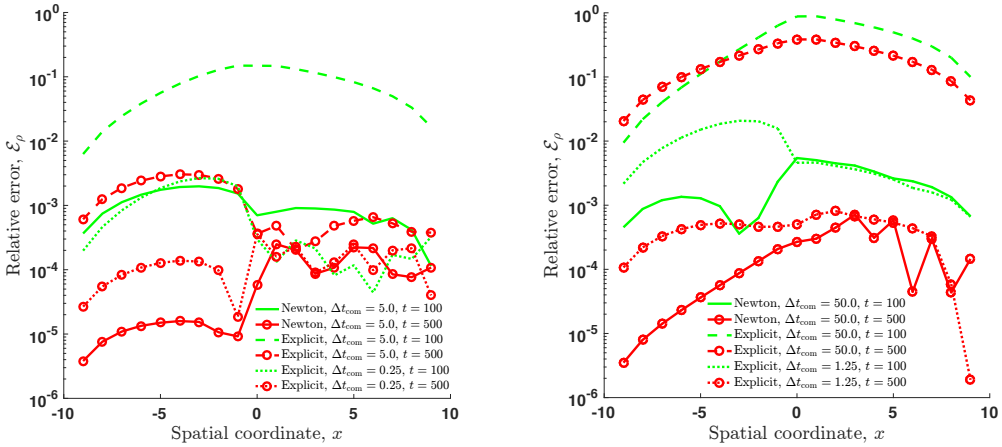

Figure 2.4: Spatial variability of the relative error $\mathcal{E}_\rho$ obtained with Newton's and explicit coupling for $\Delta t_{com} = 5.0$ (implicit, explicit) and $\Delta t_{com} = 0.25$ (explicit) on the left, and for $\Delta t_{com} = 50.0$ (implicit, explicit) and $\Delta t_{com} = 1.25$ (explicit) on the right. In both cases $\sigma_\xi = 0.1$.

The explicit coupling yields a relative error $\mathcal{E}_\rho$ that is several orders of magnitude higher than that obtained with the implicit Newton coupling, especially for $\Delta t_{com} = 50.0$ (Fig. 2.4). Decreasing Δt_{com} (by reducing the inner-solver time steps Δt_1 and Δt_2, while keeping n_1 and n_2 the same) reduces the explicit coupling's error to the level achieved by the implicit Newton coupling. This dramatically increases the explicit coupling's computation time (see Table 2.2). To gauge whether this makes the implicit coupling the more efficient choice, we compare the computation time of the explicit coupling with that of its implicit counterpart at the same solution error. For $\Delta t_{com} = 5.0$, the implicit coupling is more efficient at low noise strength ($\sigma_\xi = 0.1$), but not at high noise strength ($\sigma_\xi = 0.4$). These results are due to the fact that the number of iterations needed for the coupling to converge increases significantly with noise strength (see Section 2.6.2).

Hence, at high noise strength, the cost of the (fully-converged) implicit coupling with a lower communication frequency can outweigh that of the (single-iteration) explicit coupling with a higher communication frequency. For $\Delta t_{\text{com}} = 50.0$ however, the implicit coupling outperforms its explicit counterpart at all noise strengths. Hence, when the time between two subsequent inter-solver communications is sufficiently increased, the implicit coupling can be more efficient even at high noise strength.

Table 2.2: N_{sam}-averaged simulation time t_{sim} (in s) of the implicit and explicit coupling algorithms with several communication frequencies Δt_{com}.

Coupling	Δt_{com}	Simulation time, t_{sim}	
		$\sigma_\xi = 0.1$	$\sigma_\xi = 0.4$
implicit	5.0	3.4	13.0
explicit	0.25	5.4	5.4
implicit	50.0	1.5	4.7
explicit	1.25	10.7	10.6

2.6.4 Relative performance of implicit and explicit coupling in two dimensions

To test the generality of the previous conclusions, we consider a two-dimensional (2D) diffusion equation,

$$\frac{\partial \rho}{\partial t} = \nabla \cdot (D \nabla \rho) + f, \qquad \mathbf{x} = (x_1, x_2)^\top \in \Omega_{\text{2D}}, \quad t > 0, \tag{2.21a}$$

where $\Omega_{\text{2D}} = (-L/2, L/2) \times (-L/2, L/2)$, the diffusion coefficient $D(\mathbf{x})$ is piecewise constant,

$$D(\mathbf{x}) = \begin{cases} D_1 & \text{for } \mathbf{x} \in \Omega_{\text{2D},1} = (-L/2, 0) \times (-L/2, L/2) \\ D_2 \gg D_1 & \text{for } \mathbf{x} \in \Omega_{\text{2D},2} = [0, L/2) \times (-L/2, L/2), \end{cases} \tag{2.21b}$$

and the source term $f(\mathbf{x},t)$ is defined as

$$f(\mathbf{x},t) = \begin{cases} 0 & \text{for } \mathbf{x} \in \Omega_{2D,1} \\ \xi(\mathbf{x},t) & \text{for } \mathbf{x} \in \Omega_{2D,2}. \end{cases} \tag{2.21c}$$

Here $\xi(\mathbf{x},t)$ is a zero-mean space-time Gaussian white noise with variance σ_ξ^2 and covariance

$$\mathbb{E}[\xi(\mathbf{x},t)\xi(\mathbf{y},\tau)] = \sigma_\xi^2 \delta(\mathbf{x}-\mathbf{y})\delta(t-\tau), \quad \mathbf{x},\mathbf{y} \in \Omega_{2D,2}, \quad t,\tau > 0 \tag{2.21d}$$

Equation (2.21) is subject to boundary conditions

$$\rho(-L/2,x_2,t) = g, \qquad \rho(L/2,x_2,t) = \rho_R, \qquad \frac{\partial \rho}{\partial x_2}(x_1,\pm L/2,t) = 0, \tag{2.22}$$

where $g(x_2) = \rho_L - 2|x_2|(\rho_L - \rho_M)/L$, and an initial condition

$$\rho(\mathbf{x},0) = \rho_M \begin{cases} (g/\rho_M)^{-2x_1/L} & \text{for } \mathbf{x} \in \Omega_{2D,1} \\ (\rho_R/\rho_M)^{2x_1/L} & \text{for } \mathbf{x} \in \Omega_{2D,2}. \end{cases} \tag{2.23}$$

We employ a domain decomposition to solve the stochastic BVP (2.21)–(2.23). Solutions $\rho_1(\mathbf{x},t)$ and $\rho_2(\mathbf{x},t)$ of the BVPs defined, respectively, on the subdomains $\Omega_{2D,1}$ and $\Omega_{2D,2}$ are coupled by enforcing the continuity of the state variable and its flux at the interface $\Gamma = \{\mathbf{x} : x_1 = 0, -L/2 < x_2 < L/2\}$ separating the two subdomains,

$$\rho_1(0,x_2,t) = \rho_2(0,x_2,t), \qquad D_1 \frac{\partial \rho_1}{\partial x_1}(0,x_2,t) = D_2 \frac{\partial \rho_2}{\partial x_1}(0,x_2,t). \tag{2.24}$$

In the simulations reported below, we set $L = 20.0$, $D_1 = 0.1$, $D_2 = 10.0$, $\rho_L = 15.0$, $\rho_M = 10.0$, and $\rho_R = 5.0$. The simulation domain Ω_{2D} is discretized in both spatial

directions with a uniform mesh of cell size $\Delta x_1 = \Delta x_2 \equiv \Delta x = 1.0$ (i.e., $L/\Delta x \equiv N = 20$). We use Picard's coupling with a tolerance of $\varepsilon = 10^{-2}$, take the micro time steps $\Delta t_1 = 1.0$ and $\Delta t_2 = 0.01$, and use an inter-solver communication time $\Delta t_{\text{com}} = 5.0$. These and other values of all the physical quantities are reported in consistent units.

We compute the discretized ensemble-averaged concentration $\langle \rho(x_{1,i}, x_{2,j}, t) \rangle$ at points $x_{1,i} = -L/2 + i\Delta x$ ($i = 0, \ldots, N$) and $x_{2,j} = -L/2 + j\Delta x$ ($j = 0, \ldots, N$), and times $t = (1/5) \cdot [20.0^2/(8 \cdot 0.1)] = 100.0$ and $t = 20.0^2/(8 \cdot 0.1) = 500.0$. It is compared to its "exact" counterpart $\langle \rho(x_{1,i}, x_{2,j}, t) \rangle^{\text{ex}}$ obtained with the implicitly (Picard's iteration with $\varepsilon = 10^{-3}$) coupled explicit Euler solvers in each subdomain on a fine space-time mesh of $\Delta x_1^{\text{ex}} = \Delta x_2^{\text{ex}} \equiv \Delta x^{\text{ex}} = \Delta x/2^3 = 0.125$ and $\Delta t_1^{\text{ex}} = \Delta t_2^{\text{ex}} = 10^{-4}$ using $n_1 = n_2 = 1$. The difference between the exact and approximate solutions is reported in terms of a position-dependent relative error, $\mathcal{E}_\rho(x_{1,i}, x_{2,j}, t) = |\langle \rho(x_{1,i}, x_{2,j}, t) \rangle - \langle \rho(x_{1,i}, x_{2,j}, t) \rangle^{\text{ex}}| / \langle \rho(x_{1,i}, x_{2,j}, t) \rangle^{\text{ex}}$, at times $t = 100.0$ and $t = 500.0$.

Figure 2.5 exhibits the relative errors $\mathcal{E}_\rho(x_{1,i}, x_{2,j}, t)$ for the implicit Picard and explicit coupling methods using a noise strength $\sigma_\xi = 0.1$ (which, according to a 2D analogue of the previously described noise discretization procedure, requires $N_{\text{sam}} = 1000$). At early times ($t = 100.0$), the error of the explicit coupling can be orders of magnitude higher than that of the implicit coupling. Decreasing Δt_{com} (by reducing the inner-solver time steps Δt_1 and Δt_2, while keeping n_1 and n_2 the same) reduces the explicit coupling's error to the level achieved by its implicit counterpart; however, this makes it less efficient (see Table 2.3). A similar trend is observed for $\sigma_\xi = 0.2$, but as the noise amplitude increases to $\sigma = 0.4$, the explicit coupling becomes significantly more efficient than the implicit coupling. Although more work is needed to determine how this comparison evolves as the inter-solver communication frequency for the implicit coupling is reduced, the result for $\sigma_\xi = 0.4$ indicates that the explicit coupling may be the only computationally feasible option in 2D at high noise strength.

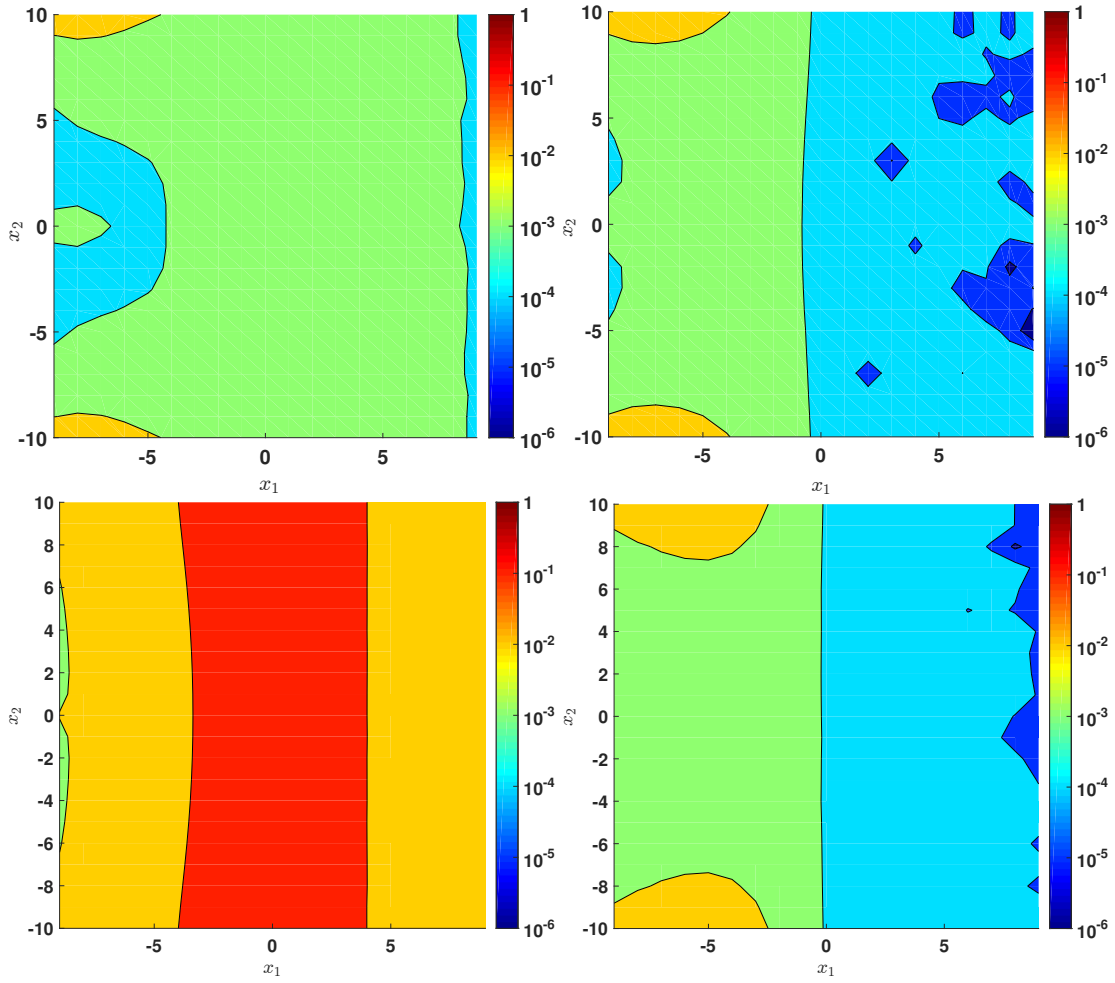

Figure 2.5: Spatial variability of the relative error $\mathcal{E}_\rho$ obtained with the implicit Picard (top) and explicit (bottom) coupling methods at $t = 100.0$ (left) and $t = 500.0$ (right). In all cases $\Delta t_{\text{com}} = 5.0$ and $\sigma_\xi = 0.1$.

Table 2.3: N_{sam}-averaged simulation time t_{sim} (in s) of the implicit Picard and explicit coupling algorithms with several communication frequencies Δt_{com}. The times listed for $\sigma_\xi = 0.2$ and 0.4 are indicative values based on averaging over only a few time trajectories.

Coupling	Δt_{com}	Simulation time, t_{sim}		
		$\sigma_\xi = 0.1$	$\sigma_\xi = 0.2$	$\sigma_\xi = 0.4$
implicit	5.0	44.1	86.7	7479.7
explicit	0.5	237.9	242.3	233.0

2.7 Summary and Conclusions

We constructed a tightly-coupled domain-decomposition approach using Picard's or Newton's method and applied it to a multiscale, interfacially-coupled linear diffusion problem driven by a Gaussian space-time white noise in one of the subdomains. We conducted a series of numerical experiments to compare the efficiency of the fully-converged ("implicit") Picard and Newton coupling methods, and to investigate the efficiency/accuracy trade-off between these implicit algorithms and a single-iteration ("explicit") Picard's coupling. These numerical properties were explored for various strengths of the Gaussian noise, and for different frequencies of communication between the constituent subdomain solvers.

Our analysis leads to the following major conclusions.

1. Implicit Newton's coupling typically outperforms its Picard counterpart, especially at high noise strength. The number of Newton iterations scales linearly with the noise amplitude, while its Picard counterpart can scale super-linearly.

2. Despite its higher cost per communication, the implicit coupling can outperform its explicit counterpart because the latter requires a higher inter-solver communication frequency to achieve a given solution error. At low noise strength, this holds true even if the implicit coupling's communication frequency is high. At high noise strength however, it requires increasing the time between two subsequent implicit coupling communications to offset the increased cost of the fully converged iterations.

3. The implicit coupling preserves the order of accuracy of the constituent solvers, even for strong random fluctuations.

4. The presence of noise does not alter the stability properties of the domain-decomposition

algorithm compared to its fully-deterministic counterpart, regardless of the strength of the fluctuations.

5. In two dimensions, we find a similar trend as in conclusion 2, although an explicit coupling may be the only computationally viable option at high noise strength.

2.8 Acknowledgements

The material contained in this Chapter is based on research supported in part by Defense Advanced Research Projects Agency under the EQUiPS program, the Air Force Office of Scientific Research under grant FA9550-12-1-0185 and by the National Science Foundation under grants EAR-1246315 and DMS-1522799. It is, in part, a reprint of the material as it appears in the paper "Conservative tightly-coupled simulations of stochastic multiscale systems" by S. Taverniers, A.Y. Pigarov, and D.M. Tartakovsky, published in *Journal of Computational Physics*, 313:400–414, Elsevier Inc. (2016) [115]. The dissertation author is the primary investigator and author of this paper.

Chapter 3

A tightly-coupled domain-decomposition approach for highly nonlinear stochastic multiphysics systems

3.1 Abstract

Multiphysics simulations often involve nonlinear components that are driven by internally generated or externally imposed random fluctuations. When used with a domain-decomposition (DD) algorithm, such components have to be coupled in a way that both accurately propagates the noise between the subdomains and lends itself to a stable and cost-effective temporal integration. We develop a conservative DD approach in which tight coupling is obtained by using a Jacobian-free Newton-Krylov (JfNK) method with a generalized minimum residual iterative linear solver. This strategy is tested on a coupled nonlinear diffusion system forced by a truncated Gaussian noise at

the boundary. Enforcement of path-wise continuity of the state variable and its flux, as opposed to continuity in the mean, at interfaces between subdomains enables the DD algorithm to correctly propagate boundary fluctuations throughout the computational domain. Reliance on a single Newton iteration (explicit coupling), rather than on the fully converged JfNK (implicit) coupling, may increase the solution error by an order of magnitude. Increase in communication frequency between the DD components reduces the explicit coupling's error, but makes it less efficient than the implicit coupling at comparable error levels for all noise strengths considered. Finally, the DD algorithm with the implicit JfNK coupling resolves temporally-correlated fluctuations of the boundary noise when the correlation time of the latter exceeds some multiple of an appropriately defined characteristic diffusion time.

3.2 Introduction and Motivation

High-performance computing facilitates the simulation of ever more complex phenomena comprising multiple physical, chemical and/or biological processes that take place on a wide range of spatiotemporal scales. Many of these problems involve constituent processes that occur in separate spatial domains and influence each other through the interfaces between these domains. One example is conjugate heat transfer across a fluid-solid interface [95], which manifests itself in applications as diverse as gas turbine cooling [82] and vehicle entry and re-entry in planetary atmospheres [30].

Construction of a single discrete operator containing the different components and their interactions yields a "tight" coupling, which guarantees temporal synchronization of state variables across inter-component interfaces. Yet, this "monolithic" [20] approach is intrusive (i.e., requires development of new software) and might become unfeasible for high-dimensional systems. The alternative strategy of "component partitioning" or

domain decomposition (DD) advances the components independently and employs a coupling method to exchange information at the interfaces. Deployment of DDs in high-performance computing facilitates an optimal distribution of the work load between the available processing cores (load balancing), while minimizing communication between cores acting on adjacent subdomains (communication scheduling) [55]. The DD approach is nonintrusive, i.e., allows for a "black-box" implementation of existing (legacy) codes, but requires an iterative coupling to avoid desynchronization of the state variables computed with the individual components[1], which may significantly increase its computational cost.

Studies of the numerical properties of DD algorithms have led to nontrivial conclusions, which might be difficult to generalize. For instance, an otherwise unstable loose coupling used in one-dimensional simulations of fluid-solid-interactions can be made stable by enforcing Neumann boundary conditions for the structural calculation and Dirichlet boundary conditions for the fluid solver [50]. And the use of a small number of iterations in a coupled linear diffusion problem leads to conditional or unconditional stability in a nonintuitive way when using a backward Euler solver in the subdomains [102]. Random fluctuations inside or on the boundary of a computational domain further affect the accuracy and performance of DD methods [115, 31].

We focus on a highly nonlinear multiscale diffusion problem driven by a temporally correlated boundary noise. A nonlinear dependence of the diffusion coefficient on the state variable (e.g., concentration) poses a host of challenges not encountered in linear [115, 31] and weakly nonlinear [31] problems. A computational testbed problem described in Section 3.3—a one-dimensional nonlinear diffusion in a composite solid forced by a truncated Gaussian noise at its left boundary—represents production of ultra-pure hydrogen gas [36]. Section 3.4 contains a description of our DD algorithm,

[1] Examples of such a desynchronization due to the use of noniterative or "loosely" coupled algorithms, and methods to iteratively correct them, can be found in [43, 42, 75].

which uses a Jacobian-free Newton-Krylov (JfNK) method to tightly couple two explicit Euler diffusion solvers. In Section 3.5 we analyze the stability of the time advancement scheme in the presence of a temporally fluctuating boundary noise. In Section 3.6 we conduct a series of computational experiments to elucidate the numerical properties of our algorithm. A summary of our findings is reported in Section 3.7.

3.3 Problem Formulation

Consider a state variable $\rho(x,t)$ whose dynamics is governed by a one-dimensional nonlinear diffusion equation,

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left[D(\rho,x)\frac{\partial \rho}{\partial x}\right], \qquad x \in \Omega \equiv (0,L), \quad t > 0, \tag{3.1a}$$

with the ρ-dependent diffusion coefficient D; this equation is defined on the simulation domain $\Omega \equiv (0,L)$ for times $t > 0$. While (3.1a) describes a large number of physical phenomena, we ground it in an application related to production of ultra-pure hydrogen gas [36]. Thus, $\rho(x,t)$ represents the concentration of atomic hydrogen (H) that diffuses through a dense composite metal membrane of thickness L. The latter is placed between streams of feed and sweep gases flowing in opposite directions in order to extract H_2 from the feed gas (for a typical configuration, see, e.g., [21] and Fig. 3.1). The membrane consists of a tantalum (Ta) layer Ω_2 sandwiched between two palladium (Pd) layers Ω_1 and Ω_3 [89]. Palladium's selective permeability to hydrogen [54] makes it suitable for use in hydrogen-separation membranes. To increase its structural stability, Pd has been alloyed with materials such as silver [59]. An alternative, and potentially superior approach, is the combination of Pd with refractory (group V) metals, such as tantalum, into layered membranes [25]. Refractory metals have even higher bulk hydrogen permeabilities than Pd or its alloys [22], and are cheaper than Pd.

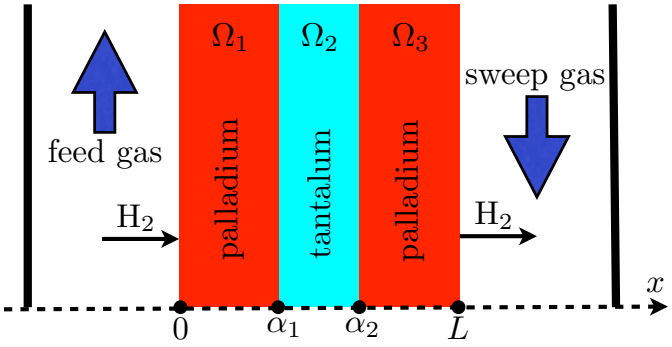

Figure 3.1: A three-layer dense membrane configuration for hydrogen separation (not to scale).

The diffusion coefficient of H in this composite is given by [103, 10]

$$D(\rho,x) = \begin{cases} D_{\mathrm{Pd}} \equiv D_{\mathrm{Pd}}^{\mathrm{int}} \left[f_1(\beta) + f_2(\rho) V_{\mathrm{Pd}} \rho \dfrac{1-\rho V_{\mathrm{Pd}}}{k_B T} \right] & \text{for } x \in \Omega_1 \cup \Omega_3 \\ D_{\mathrm{Ta}} \equiv D_{\mathrm{Ta}}^{\mathrm{int}} & \text{for } x \in \Omega_2 \end{cases} \tag{3.1b}$$

with

$$f_1 = 1 - 6\,\frac{\beta-1}{\beta}, \qquad f_2 = \frac{1}{V_{\mathrm{Pd}}}\frac{\partial \mu_e}{\partial \rho} - \frac{3\gamma}{2}\coth\left(\frac{\hbar\omega_\alpha - \gamma\,\rho V_{\mathrm{Pd}}}{2k_B T}\right). \tag{3.1c}$$

Here V_{Pd} is the Wigner-Seitz cell volume of the Pd lattice; k_B is the Boltzmann constant; T is the operating temperature of the membrane (in K); the constants γ and ω_α (energy of local vibrations of the H atoms for ρ close to 0) and variables $\beta(\rho)$ and $\mu_e(\rho)$ (electronic contribution to the chemical potential of the hydrogen subsystem) are defined in [103]; and $D_{\mathrm{Pd}}^{\mathrm{int}}$ and $D_{\mathrm{Ta}}^{\mathrm{int}}$ are the intrinsic (i.e., for ρ close to 0) diffusion coefficients[2] of H in Pd and Ta. Since the function $D_{\mathrm{Pd}}(\rho)$ in (3.1b) is discontinuous at $\rho V_{\mathrm{Pd}} = 0.656$, we replace it with a continuous approximation obtained by using quadratic splines (Fig. 3.2).

[2]Following [4], we use Arrhenius expressions $D_{\mathrm{Pd}}^{\mathrm{int}} = D_{\mathrm{Pd},0}\exp[-E_{\mathrm{Pd}}^{\mathrm{act}}/(RT)]$ and $D_{\mathrm{Ta}}^{\mathrm{int}} = D_{\mathrm{Ta},0}\exp[-E_{\mathrm{Ta}}^{\mathrm{act}}/(k_B T)]$ with R denoting the universal gas constant, $D_{\mathrm{Pd},0} = 2.9\cdot 10^{-7}$ m^2/s and $D_{\mathrm{Ta},0} = 4.4\cdot 10^{-8}$ m^2/s, and diffusion activation energies $E_{\mathrm{Pd}}^{\mathrm{act}} = 22.2$ kJ/mol and $E_{\mathrm{Ta}}^{\mathrm{act}} = 0.14$ eV/atom. We set $T = 800$ K, a temperature regime for which the expression $D_{\mathrm{Pd}}(\rho)$ in (3.1b) is valid.

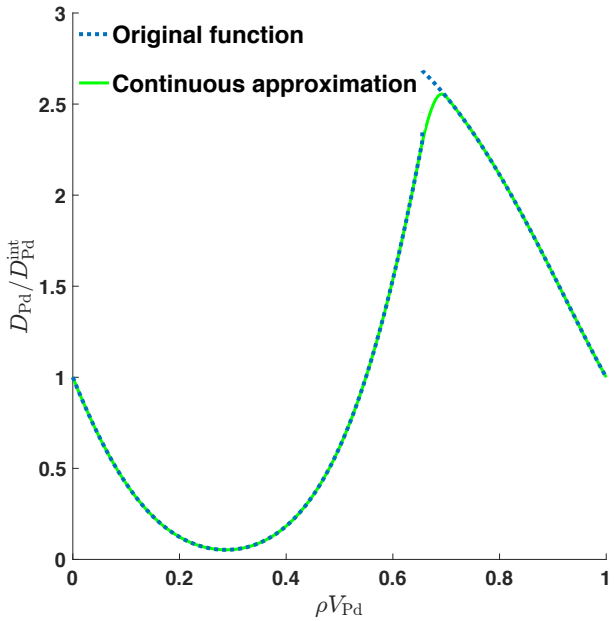

Figure 3.2: The original function $D_{\mathrm{Pd}}/D_{\mathrm{Pd}}^{\mathrm{int}}$ and its continuous approximation.

Equation (3.1) is subject to initial and Dirichlet boundary conditions

$$\rho(x,0)=\rho_L, \qquad \rho(0,t)=\rho_0(t), \qquad \rho(L,t)=\rho_L, \tag{3.2}$$

where ρ_L is a deterministic constant and $\rho_0(t)$ is the randomly fluctuating boundary function. The latter is expressed as $\rho_0(t)=\langle\rho_0\rangle+\eta(t)$ with $\langle\rho_0\rangle$ the (constant) ensemble-averaged value and $\eta(t)$ a zero-mean truncated Gaussian noise with variance σ_η^2, an exponential auto-covariance $C_\eta(t_1-t_2)=\sigma_\eta^2\exp(-|t_1-t_2|/\lambda)$ and the correlation time λ. The boundary concentrations $\rho_0(t)$ and ρ_L are related to the partial pressure of H_2 in the feed and sweep gases, respectively. At the feed gas/membrane interface, H_2 molecules are adsorbed onto the membrane surface, where they dissociate into H atoms which enter the Pd lattice; the reverse process occurs at the sweep gas/membrane interface.

The presence of the boundary noise $\eta(t)$ renders a solution $\rho(x,t)$ of the boundary-value problem (BVP) (3.1)–(3.2) random over the entire simulation domain Ω. Its

statistics, such as mean $\langle \rho(x,t) \rangle$ and variance $\sigma_\rho^2(x,t)$, can be estimated with, e.g., Monte Carlo (MC) simulations. Regardless of the noise, the disparate diffusion time scales in the Pd and Ta layers require a single full-domain algorithm to use a time step determined by the smallest diffusion time-scale (corresponding to the maximal value of D_{Pd}) in order to accurately resolve the system's dynamics. This can significantly increase the computational time of each MC realization, potentially rendering MC simulations prohibitively expensive.

A domain decomposition enables one to use different time steps in each subdomain Ω_i ($i = 1,2,3$), which are in tune with the local diffusion time-scale. Let $\rho_i(x,t)$ denote a solution of (3.1a) on the ith subdomain Ω_i ($i = 1,2,3$). These solutions are subject to the initial condition $\rho_i(x,0) = \rho_L$ for $i = 1,2,3$; additionally the external boundary conditions give rise to $\rho_1(0,t) = \rho_0(t)$ and $\rho_3(L,t) = \rho_L$. The remaining boundary conditions for these three BVPs come from enforcing the continuity of ρ and its flux at the interfaces $x = \alpha_1$ and $x = \alpha_2$ separating the three subdomains (see Fig. 3.1):

$$\begin{aligned} \rho_1(\alpha_1,t) &= \rho_2(\alpha_1,t), \qquad & D_{\mathrm{Pd}}\frac{\partial \rho_1}{\partial x}(\alpha_1,t) &= D_{\mathrm{Ta}}\frac{\partial \rho_2}{\partial x}(\alpha_1,t), \\ \rho_2(\alpha_2,t) &= \rho_3(\alpha_2,t), \qquad & D_{\mathrm{Ta}}\frac{\partial \rho_2}{\partial x}(\alpha_2,t) &= D_{\mathrm{Pd}}\frac{\partial \rho_3}{\partial x}(\alpha_2,t). \end{aligned} \tag{3.3}$$

These interfacial conditions necessitate occasional communication between the diffusion solvers in adjacent subdomains.

3.4 Numerical Implementation of Domain Decomposition

3.4.1 Spatial discretization of the computational domain

We discretize the computational domain Ω using a uniform mesh of cell size $\Delta x = L/N$, where N is the total number of grid cells. For the sake of simplicity, and without loss of generality, we assume that the interfaces $x = \alpha_1$ and $x = \alpha_2$ coincide with the nodes of this uniform grid, so that each of the three subdomains Ω_i is discretized with N_i grid cells. The solvers used to integrate the three BVPs employ a staggered grid approach, in which diffusive fluxes, $F_i(x,t) = -D_i \partial \rho_i / \partial x$, in the ith subdomain Ω_i ($i = 1,2,3$) are calculated at the midpoint between two subsequent concentration nodes (Fig. 3.3). In particular, both $\rho_l^- \equiv \rho_1(\alpha_1,t)$ and $\rho_l^+ \equiv \rho_2(\alpha_1,t)$ are defined at the interfacial node $x = \alpha_1$, while the corresponding interfacial fluxes $F_l^- \equiv F_1(\alpha_1,t)$ and $F_l^+ \equiv F_2(\alpha_1,t)$ are defined at $x = \alpha_1 - \Delta x/2$ and $x = \alpha_1 + \Delta x/2$, respectively. Likewise, $\rho_r^- \equiv \rho_2(\alpha_2,t)$ and $\rho_r^+ \equiv \rho_3(\alpha_2,t)$ are defined at $x = \alpha_2$, while the corresponding interfacial fluxes $F_r^- \equiv F_2(\alpha_2,t)$ and $F_r^+ = F_3(\alpha_2,t)$ are defined at $x = \alpha_2 - \Delta x/2$ and $x = \alpha_2 + \Delta x/2$, respectively.

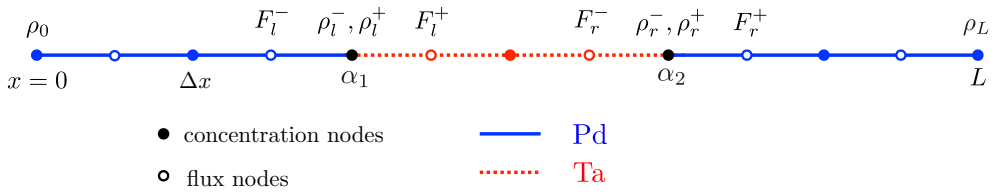

Figure 3.3: Domain $\Omega = (0,L)$ decomposed into subdomains $\Omega_1 = (0,\alpha_1)$, $\Omega_2 = [\alpha_1,\alpha_2)$ and $\Omega_3 = [\alpha_2,L)$. Concentration ρ is computed at the nodes denoted by solid circles, and flux F is evaluated at the midpoint between two subsequent concentration nodes (open circles).

3.4.2 Numerical solvers for individual BVPs

A finite-difference approximation of the spatial derivatives in (3.1a) yields a system of ordinary differential equations for each subdomain Ω_i,

$$\frac{\mathrm{d}\boldsymbol{\rho}_i}{\mathrm{d}t} = \mathbf{f}_i(\boldsymbol{\rho}_i), \qquad i = 1,2,3, \tag{3.4a}$$

where $\boldsymbol{\rho}_i = (\rho_{i,1}, \ldots, \rho_{i,N_i-1})^\top$ are the one-dimensional arrays of size $N_i - 1$ of the nodal values of the state variables $\rho_i(x,t)$; and components $f_{i,p}$ of the one-dimensional arrays $\mathbf{f}_i(\boldsymbol{\rho}_i)$ of size $N_i - 1$ are defined by

$$f_{i,p} = \frac{D_i^+ \cdot (\rho_{i,p+1} - \rho_{i,p}) - D_i^- \cdot (\rho_{i,p} - \rho_{i,p-1})}{\Delta x^2}, \tag{3.4b}$$

for $p = 1, \ldots, N_i - 1$ and $i = 1,2,3$. Here D_i^+ and D_i^- are the values of $D_i \equiv D_{\mathrm{Pd}}$ evaluated at $(\rho_{i,p+1} + \rho_{i,p})/2$ and $(\rho_{i,p} + \rho_{i,p-1})/2$, respectively, for $i = 1,3$, and $D_2^+ = D_2^- = D_2 \equiv D_{\mathrm{Ta}}$; and $\rho_{1,0} = \rho_0$, $\rho_{1,N_1} = \rho_l^-$, $\rho_{2,0} = \rho_l^+$, $\rho_{2,N_2} = \rho_r^-$, $\rho_{3,0} = \rho_r^+$ and $\rho_{3,N_3} = \rho_L$.

We use an explicit Euler method with time step Δt_i to advance $\boldsymbol{\rho}_i$ $(i = 1,2,3)$ in time. The noise enters the finite-difference scheme through advancing the $p = 1$ component of ρ_1 from $t_n = n\Delta t_1$ to $t_{n+1} = (n+1)\Delta t_1$,

$$\rho_{1,1}^{n+1} = \rho_{1,1}^n + \frac{\Delta t_1}{\Delta x^2}[D_1^+ \cdot (\rho_{1,2}^n - \rho_{1,1}^n) - D_1^- \cdot (\rho_{1,1}^n - \rho_0^n)]. \tag{3.5}$$

The random boundary term is represented as $\rho_{1,0}^n \equiv \rho_0^n = \langle \rho_0 \rangle + \eta^n$ with $\eta^n = \eta(t_n)$.

Given values of the interfacial concentrations $\rho_l^-(t) = \rho_l^+(t)$ at $x = \alpha_1$ and $\rho_r^-(t) = \rho_r^+(t)$ at $x = \alpha_2$, this spatiotemporal discretization allows one to compute, independently from each other, the three $(i = 1,2,3)$ solutions $\boldsymbol{\rho}_i(\tau)$ at any time $\tau > t$. However, these solutions will not satisfy the continuity conditions (3.3). To enforce the latter, the subdomain solvers must communicate with each other through a coupling

algorithm. The continuity conditions can be enforced either for each MC realization of the random solution (path-wise coupling) or for its statistical moments (e.g., mean) computed using a finite number of MC realizations (moment-wise coupling). The path-wise communication ensures the continuity of all the moments of concentration $\rho(x,t)$ and flux $F(x,t) = -D\partial_x\rho$ and, hence, accurately propagates the noise throughout the computational domain. The moment-wise coupling introduces an error since it does not account for higher-order moments. Yet, it reduces the computational overhead due to inter-solver communication and therefore warrants an investigation. In Section 3.6.1 we compare the spatial profiles of the mean and variance of $\rho(x,t)$ for path-wise and moment-wise (exchanging only the mean) coupling with those obtained by solving (3.1)–(3.2) with a single solver defined on the entire domain.

3.4.3 Coupling algorithm with path-wise communication

We define the communication time between solvers, Δt_{com}, as the multiples of the inner-solver time steps Δt_i, such that $\Delta t_{\text{com}} = n_i \Delta t_i$ $(i = 1,2,3)$. In other words, starting at $t = t_n$ the ith subdomain solver is advanced by n_i "micro" steps Δt_i, before communicating with adjacent solvers at $t_n + \Delta t_{\text{com}}$. Depending on whether an iterative or noniterative coupling is used, this process is repeated until convergence or carried out only once, respectively. In either case, this procedure advances the solution by one "macro" step Δt_{com}. In the iterative coupling, the set of communications during the various iterations of a particular macro-step are referred to as one overall communication.

We show in D that exchanging the Δt_{com}-averaged interfacial concentrations and fluxes (rather than their counterparts computed at the last micro-step of each solver) yields a mass-conservative coupling algorithm. The time-averaged concentrations, $\bar{\boldsymbol{\rho}}_i$ $(i = 1,2,3)$, are computed as the arithmetic means of $\boldsymbol{\rho}_i$ over their respective n_i micro-steps. The interfacial values of these Δt_{com}-averaged concentrations are $\bar{\rho}_l^-$ and $\bar{\rho}_l^+$ at

$x = \alpha_1$, and $\bar{\rho}_r^-$ and $\bar{\rho}_r^+$ at $x = \alpha_2$; the corresponding Δt_{com}-averaged interfacial fluxes are $\bar{F}_l^-$ and $\bar{F}_l^+$ at $x = \alpha_1$ and $\bar{F}_r^-$ and $\bar{F}_r^+$ at $x = \alpha_2$.

Enforcement of (3.3) provides a tight coupling of the solvers for the subdomains Ω_i ($i = 1, 2, 3$). We accomplish this by using an iterative (implicit) coupling algorithm based on the JfNK method [70] with the generalized minimum residual (GMRES) iterative linear solver [68] (see E for details). This root-finding algorithm is deployed to solve a system of coupled nonlinear algebraic equations,

$$\bar{\rho}_l^- = \bar{\rho}_l^+, \qquad \bar{F}_l^- = \bar{F}_l^+, \qquad \bar{\rho}_r^- = \bar{\rho}_r^+, \qquad \bar{F}_r^- = \bar{F}_r^+, \tag{3.6}$$

during inter-solver communication. Using notation

$$\begin{gathered} \bar{\rho}_{1,N_1}^n = \bar{\rho}_l^-, \quad \bar{\rho}_{2,0}^n = \bar{\rho}_l^+, \quad \bar{\rho}_{2,N_2}^n = \bar{\rho}_r^-, \quad \bar{\rho}_{3,0}^n = \bar{\rho}_r^+, \\ \bar{F}_{1,N_1-1/2}^n = \bar{F}_l^-, \quad \bar{F}_{2,1/2}^n = \bar{F}_l^+, \quad \bar{F}_{2,N_2-1/2}^n = \bar{F}_r^-, \quad \bar{F}_{3,1/2}^n = \bar{F}_r^+, \end{gathered} \tag{3.7}$$

for the macro-step from t_n to $t_{n+1} = t_n + \Delta t_{\text{com}}$, this system is written as

$$\bar{\rho}_{1,N_1}^n = \bar{\rho}_{2,0}^n, \quad \bar{F}_{1,N_1-1/2}^n = \bar{F}_{2,1/2}^n, \quad \bar{\rho}_{2,N_2}^n = \bar{\rho}_{3,0}^n, \quad \bar{F}_{2,N_2-1/2}^n = \bar{F}_{3,1/2}^n. \tag{3.8}$$

Newton's method, in its pure form, recasts (3.8) into an iteration problem

$$\begin{pmatrix} \bar{\rho}_{1,N_1}^{n,k+1} \\ \bar{F}_{2,1/2}^{n,k+1} \\ \bar{\rho}_{2,N_2}^{n,k+1} \\ \bar{F}_{3,1/2}^{n,k+1} \end{pmatrix} = \begin{pmatrix} \bar{\rho}_{1,N_1}^{n,k} \\ \bar{F}_{2,1/2}^{n,k} \\ \bar{\rho}_{2,N_2}^{n,k} \\ \bar{F}_{3,1/2}^{n,k} \end{pmatrix} - \mathbf{J}^{-1} \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{pmatrix} \tag{3.9a}$$

where $\mathbf{J}(\bar{\rho}_{1,N_1}^{n,k}, \bar{F}_{2,1/2}^{n,k}, \bar{\rho}_{2,N_2}^{n,k}, \bar{F}_{3,1/2}^{n,k})$ is the Jacobian,

$$\mathbf{J} = \begin{pmatrix} \partial g_1/\partial \bar{\rho}_{1,N_1}^{n,k} & \partial g_1/\partial \bar{F}_{2,1/2}^{n,k} & \partial g_1/\partial \bar{\rho}_{2,N_2}^{n,k} & \partial g_1/\partial \bar{F}_{3,1/2}^{n,k} \\ \partial g_2/\partial \bar{\rho}_{1,N_1}^{n,k} & \partial g_2/\partial \bar{F}_{2,1/2}^{n,k} & \partial g_2/\partial \bar{\rho}_{2,N_2}^{n,k} & \partial g_2/\partial \bar{F}_{3,1/2}^{n,k} \\ \partial g_3/\partial \bar{\rho}_{1,N_1}^{n,k} & \partial g_3/\partial \bar{F}_{2,1/2}^{n,k} & \partial g_3/\partial \bar{\rho}_{2,N_2}^{n,k} & \partial g_3/\partial \bar{F}_{3,1/2}^{n,k} \\ \partial g_4/\partial \bar{\rho}_{1,N_1}^{n,k} & \partial g_4/\partial \bar{F}_{2,1/2}^{n,k} & \partial g_4/\partial \bar{\rho}_{2,N_2}^{n,k} & \partial g_4/\partial \bar{F}_{3,1/2}^{n,k} \end{pmatrix}, \tag{3.9b}$$

and

$$\begin{aligned} g_1 &= \bar{\rho}_{1,N_1}^{n,k} - \bar{\rho}_{2,1}^{n,k} - \frac{\Delta x}{D_2}\bar{F}_{2,1/2}^{n,k}, \quad g_2 = -\bar{D}_1^{n,k}\frac{\bar{\rho}_{1,N_1}^{n,k} - \bar{\rho}_{1,N_1-1}^{n,k}}{\Delta x} - \bar{F}_{2,1/2}^{n,k}, \\ g_3 &= \bar{\rho}_{2,N_2}^{n,k} - \bar{\rho}_{3,1}^{n,k} - \frac{\Delta x}{\bar{D}_3^{n,k}}\bar{F}_{3,1/2}^{n,k}, \quad g_4 = -D_2\frac{\bar{\rho}_{2,N_2}^{n,k} - \bar{\rho}_{2,N_2-1}^{n,k}}{\Delta x} - \bar{F}_{3,1/2}^{n,k}. \end{aligned} \tag{3.9c}$$

Here $\bar{D}_1^{n,k}$ and $\bar{D}_3^{n,k}$ are the Δt_{com}-averaged values of the diffusion coefficients at the spatial positions $\alpha_1 - \Delta x/2$ and $\alpha_2 + \Delta x/2$, respectively.

In our numerical experiments, we do not explicitly compute the components of the Jacobian $\mathbf{J}$. Instead, we employ an inexact Newton's method, JfNK with GMRES, using a second-order finite difference expression to approximate the Jacobian-vector product. The Newton iterations continue until $\max\{|g_1|,\ldots,|g_4|\} \leq \varepsilon$, where ε is the prescribed tolerance.

3.4.4 Coupling algorithm with moment-wise communication

While the path-wise continuity of the Δt_{com}-averaged interfacial concentrations and fluxes (Section 3.4.3) guarantees the continuity of all moments of ρ and F, it is computationally intensive. To lower the computational cost, we consider an approximation involving the exchange of only the first moments (ensemble means). The modified time-advancement algorithm comprises the following steps.

1. Evolve all N_{sam} solutions, simultaneously and independently, from t_n to $t_{n+1} = t_n + \Delta t_{\text{com}}$. The value of N_{sam} is chosen according to the procedure described in Section 3.6 (and is equal to N_{sam} used in the path-wise approach);

2. Compute, at each micro-step of Ω_i ($i = 1,2,3$), the N_{sam}-average of the quantities to be exchanged between the subdomain solvers;

3. Compute the Δt_{com}-averages of the results obtained in step 2. The latter serve as inputs to the coupling algorithm, which produces new iterates for the interfacial concentration and flux (in the form of the Δt_{com}-average of their N_{sam}-averages) at $x = \alpha_1$ and $x = \alpha_2$.

3.4.5 Domain-decomposition algorithm with path-wise communication

Let $\boldsymbol{\rho}_1^{n,l,k} \equiv \boldsymbol{\rho}_1(t_n + l\Delta t_1)$, $\boldsymbol{\rho}_2^{n,m,k} = \boldsymbol{\rho}_2(t_n + m\Delta t_2)$ and $\boldsymbol{\rho}_3^{n,q,k} = \boldsymbol{\rho}_3(t_n + q\Delta t_3)$ denote arrays of the nodal concentrations at inner-solver times $t_n + l\Delta t_1$, $t_n + m\Delta t_2$ and $t_n + q\Delta t_3$ during the kth iteration of the macro-step from t_n to $t_{n+1} = t_n + \Delta t_{\text{com}}$. At all times, the arrays $\boldsymbol{\rho}_1$, $\boldsymbol{\rho}_2$ and $\boldsymbol{\rho}_3$ are of size $N_1 - 1$, $N_2 - 1$ and $N_3 - 1$, respectively. As before, $\bar{\rho}_{1,N_1}^{n,k}$ and $\bar{F}_{2,1/2}^{n,k}$ denote the Δt_{com}-averaged interfacial concentration and flux at $x = \alpha_1$, and $\bar{\rho}_{2,N_2}^{n,k}$ and $\bar{F}_{3,1/2}^{n,k}$ denote the Δt_{com}-averaged interfacial concentration and flux at $x = \alpha_2$, during the kth iteration of that macro-step. The solution is advanced from t_n to t_{n+1} as follows.

1. Initialization step. Set $\bar{\rho}_{1,N_1}^{n,0} = \rho_{1,N_1}(t_n)$, $\bar{\rho}_{2,N_2}^{n,0} = \rho_{2,N_2}(t_n)$, $\bar{F}_{2,1/2}^{n,0} = F_{2,1/2}(t_n)$ and $\bar{F}_{3,1/2}^{n,0} = F_{3,1/2}(t_n)$.

2. Evolve $\boldsymbol{\rho}_1^{n,0,k}$ to $\boldsymbol{\rho}_1^{n,n_1,k}$ over n_1 micro-steps, using $\rho_0^{n,l}$ ($l = 0, \ldots, n_1 - 1$) and $\bar{\rho}_{1,N_1}^{n,k}$ as boundary conditions at $x = 0$ and $x = \alpha_1$, respectively.

3. Evolve $\boldsymbol{\rho}_2^{n,0,k}$ to $\boldsymbol{\rho}_2^{n,n_2,k}$ over n_2 micro-steps, using $\bar{\rho}_{1,N_1}^{n,k}$ and $\bar{\rho}_{2,N_2}^{n,k}$ as boundary conditions at $x = \alpha_1$ and $x = \alpha_2$, respectively.

4. Evolve $\boldsymbol{\rho}_3^{n,0,k}$ to $\boldsymbol{\rho}_3^{n,n_3,k}$ over n_3 micro-steps, using $\bar{\rho}_{2,N_2}^{n,k}$ and ρ_L as the boundary conditions at $x = \alpha_2$ and $x = L$, respectively.

5. Use JfNK to calculate new iterates of the interfacial concentrations, $\bar{\rho}_{1,N_1}^{n,k+1}$ and $\bar{\rho}_{2,N_2}^{n,k+1}$, and fluxes, $\bar{F}_{2,1/2}^{n,k+1}$ and $\bar{F}_{3,1/2}^{n,k+1}$.

6. Repeat steps 2 through 5 until the given tolerance ε is achieved.

7. Advance the solution by one macro-step by setting

$$\rho_{1,N_1}(t_{n+1}) = \bar{\rho}^{n,K}_{1,N_1}, \quad \rho_{2,N_2}(t_{n+1}) = \bar{\rho}^{n,K}_{2,N_2},$$
$$F_{2,1/2}(t_{n+1}) = \bar{F}^{n,K}_{2,1/2}, \quad F_{3,1/2}(t_{n+1}) = \bar{F}^{n,K}_{3,1/2},$$

where $K = K(n)$ indicates the number of iterations at convergence. By construction,

$$\rho_{2,0}(t_{n+1}) = \rho_{1,N_1}(t_{n+1}), \quad \rho_{3,0}(t_{n+1}) = \rho_{2,N_2}(t_{n+1}),$$
$$F_{1,N_1-1/2}(t_{n+1}) = F_{2,1/2}(t_{n+1}), \quad F_{2,N_2-1/2}(t_{n+1}) = F_{3,1/2}(t_{n+1}).$$

Note that one could equally use $\bar{\rho}^{n,k}_{2,0}$, $\bar{\rho}^{n,k}_{3,0}$, $\bar{F}^{n,k}_{1,N_1-1/2}$ and $\bar{F}^{n,k}_{2,N_2-1/2}$ as iterates.

3.4.6 Domain-decomposition algorithm with moment-wise communication

As explained in Section 3.4.4, the moment-wise approach replaces the Δt_{com}-averages of quantities with the Δt_{com}-averages of their N_{sam}-averaged counterparts. In the time advancement algorithm of Section 3.4.5, this involves replacing, for each realization, the Dirichlet boundary conditions $\bar{\rho}^{n,k}_{1,N_1}$ at $x = \alpha_1$ and $\bar{\rho}^{n,k}_{2,N_2}$ at $x = \alpha_2$ with the Δt_{com}-average of the N_{sam}-averaged values of ρ_{1,N_1} and ρ_{2,N_2}, respectively. Thus, while each realization is characterized by a different stochastic boundary condition at $x = 0$, all of them share the same Dirichlet boundary conditions at $x = \alpha_1$ and $x = \alpha_2$ for evolving the subdomain solutions. All other quantities are modified in a similar fashion.

3.5 Stability of Implicitly Coupled DD Algorithm with Path-wise Communication

To analyze the stability of the implicitly coupled DD algorithm with path-wise communication between the subdomain solvers, we approximate our discretization of (3.4) by treating the diffusion coefficients D_1 and D_3 at respective positions $p_1\Delta x$ and $\alpha_2 + p_3\Delta x$ ($p_i = 1/2,\ldots,N_i - 1/2$) as constant over a macro-step. This effectively linearizes the algorithm around t_n for the macro-step from t_n to t_{n+1}. The most stringent constraint on the size of the macro-step for this approximation to hold comes from the time variation of D_1 at $x = \Delta x/2$ for small correlation times λ of the boundary noise $\eta(t)$; it results in the condition $\Delta t_{\mathrm{com}} \ll \lambda$.

Having chosen Δt_{com} such that the above linearization procedure yields a reasonable approximation during the macro-step from t_n to t_{n+1}, one micro-step of the left ($l = 0,\ldots,n_1 - 1$), middle ($m = 0,\ldots,n_2 - 1$) and right ($q = 0,\ldots,n_3 - 1$) subdomain solvers is given by

$$\boldsymbol{\rho}_1^{n,l+1,k} = (\mathbf{I}_{N_1-1} + \mathbf{A}_{1,n})\boldsymbol{\rho}_1^{n,l,k} + \mathbf{T}_{1,n}\,\boldsymbol{\rho}_{1,\mathrm{b}}^{n,k} + \mathbf{T}_{1,n}\,\boldsymbol{\eta}^{n,l}, \tag{3.10a}$$

$$\boldsymbol{\rho}_2^{n,m+1,k} = (\mathbf{I}_{N_2-1} + \mathbf{A}_2)\boldsymbol{\rho}_2^{n,m,k} + \mathbf{T}_2\,\boldsymbol{\rho}_{2,\mathrm{b}}^{n,k}, \tag{3.10b}$$

$$\boldsymbol{\rho}_3^{n,q+1,k} = (\mathbf{I}_{N_3-1} + \mathbf{A}_{3,n})\boldsymbol{\rho}_3^{n,q,k} + \mathbf{T}_{3,n}\,\boldsymbol{\rho}_{3,\mathrm{b}}^{n,k}, \tag{3.10c}$$

where $\boldsymbol{\rho}_{1,\mathrm{b}}^{n,k} \equiv (\langle\rho_0\rangle, 0,\ldots 0, \bar{\rho}_{1,N_1}^{n,k})^\top$, $\boldsymbol{\rho}_{2,\mathrm{b}}^{n,k} \equiv (\bar{\rho}_{1,N_1}^{n,k}, 0,\ldots,0,\bar{\rho}_{2,N_2}^{n,k})^\top$ and $\boldsymbol{\rho}_{3,\mathrm{b}}^{n,k} \equiv (\bar{\rho}_{2,N_2}^{n,k}, 0,\ldots 0,\rho_L)^\top$ are vectors of size $N_1 - 1$, $N_2 - 1$ and $N_3 - 1$, respectively, which supply the boundary conditions for the three solvers, and $\boldsymbol{\eta}^{n,l} \equiv (\eta^{n,l}, 0,\ldots,0)^\top$.[3] The identity matrices $\mathbf{I}_{N_i-1}$ ($i = 1,2,3$) are of size $N_i - 1$, and the square matrices $\mathbf{A}_{i,n}$ and $\mathbf{T}_{i,n}$ of size $N_i - 1$ ($i = 1,3$) are defined in F. Finally, the square matrices $\mathbf{A}_2$ and $\mathbf{T}_2$ of size $N_2 - 1$ are given by

[3] In equation (3.10a), $\eta^{n,l}$ is separated from the left boundary concentration to render $\boldsymbol{\rho}_{1,\mathrm{b}}^{n,k}$ independent of the micro-step, which simplifies the analysis.

$\mathbf{A}_2 = (D_2\Delta t_2/\Delta x^2)\ \mathbf{Trid}(1,-2,1)$ and $\mathbf{T}_2 = (D_2\Delta t_2/\Delta x^2)\mathbf{I}_{N_2-1}$, where $\mathbf{Trid}(1,-2,1)$ denotes the tridiagonal Toeplitz matrix of size N_2-1 whose elements on the main diagonal are -2 and those on the first sub- and super-diagonal are 1.

Lemma 3.5.1. *Given vectors, of size $N+1$,*

$$\begin{aligned}\mathbf{x}^n &= (\boldsymbol{\rho}_1^{n,0}, \bar{\rho}_{1,N_1}^{n,0}, \bar{F}_{2,1/2}^{n,0}, \boldsymbol{\rho}_2^{n,0}, \bar{\rho}_{2,N_2}^{n,0}, \bar{F}_{3,1/2}^{n,0}, \boldsymbol{\rho}_3^{n,0})^\top,\\ \mathbf{x}^{n,k} &= (\boldsymbol{\rho}_1^{n,n_1,k}, \bar{\rho}_{1,N_1}^{n,k}, \bar{F}_{2,1/2}^{n,k}, \boldsymbol{\rho}_2^{n,n_2,k}, \bar{\rho}_{2,N_2}^{n,k}, \bar{F}_{3,1/2}^{n,k}, \boldsymbol{\rho}_3^{n,n_3,k})^\top,\end{aligned} \tag{3.11}$$

representing the solution at t_n and the kth iterate of the solution at t_{n+1}, respectively, the ensemble-averaged solution at time t_{n+1}, $\langle \mathbf{x}^{n+1}\rangle$, is given by

$$\langle \mathbf{x}^{n+1}\rangle = \langle (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{P}_n\mathbf{x}^n\rangle + \langle (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{d}^n\rangle, \tag{3.12}$$

where $\mathbf{M}_n(\mathbf{x}^n, \boldsymbol{\eta}^{n,0})$ and $\mathbf{P}_n(\mathbf{x}^n, \boldsymbol{\eta}^{n,0})$ are $(N+1)\times(N+1)$ square matrices, and the vector $\mathbf{d}^n$ of size $N+1$ depends on $\mathbf{x}^n$ and all $\boldsymbol{\eta}^{n,l}$ with $l \in \{0,1,\ldots,n_1-1\}$.

Taking the 1-norm of both sides of (3.12), and using the triangle inequality, yields

$$\|\langle \mathbf{x}^{n+1}\rangle\|_1 \le \|\langle (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{P}_n\mathbf{x}^n\rangle\|_1 + \|\langle (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{d}^n\rangle\|_1. \tag{3.13}$$

For $\|\langle (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{P}_n\mathbf{x}^n\rangle\|_1$, the following result holds.

Lemma 3.5.2. *Consider $\mathbf{y}^n = (\boldsymbol{\eta}^{n,0}, \mathbf{x}^{n\top})^\top$ and $\mathbf{Q}_n = (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{P}_n$. Using Taylor's theorem to expand $Q_n(\mathbf{y}^n)$ about $\langle \mathbf{y}^n\rangle$ yields*

$$\|\langle (\mathbf{Q}_n(\mathbf{y}^n)\ \mathbf{x}^n\rangle\|_1 \le \|\mathbf{Q}_n(\langle \mathbf{y}^n\rangle)\|_1\ \|\langle \mathbf{x}^n\rangle\|_1 + V_n \tag{3.14}$$

with $V_n \in \mathbb{R}_{\ge 0}$.

Equation (3.13) and Lemma 3.5.2 yield the following necessary conditions for stability, in ensemble mean, of the time advancement.

Lemma 3.5.3. *Stability, in the mean, of the time advancement of the implicitly coupled DD algorithm with path-wise communication requires* $\|\mathbf{Q}_n(\langle\mathbf{y}^n\rangle)\|_1$ *and* V_n *in* (3.14)*, and* $\|\langle(\mathbf{I}_{N+1}-\mathbf{M}_n)^{-1}\mathbf{d}^n\rangle\|_1$ *in* (3.13)*, to be finite for every macro-step, and* $\|\mathbf{Q}_n(\langle\mathbf{y}^n\rangle)\|_1$ *to be smaller than one for all but a finite number of macro-steps. The condition on* V_n *is satisfied if* $[Q_n(\mathbf{y}^n)]_{i,j}$ *is of class* C^2 *on a open convex set containing* $\mathbf{y}^n$ *and* $\langle\mathbf{y}^n\rangle$*,* $|\partial^\alpha[Q_n(\mathbf{y}^n)]_{i,j}|$ *with* $|\alpha|=2$ *is finite, the map of the noise* η *onto* $\mathbf{y}^n$ *is monotonic,* $\nabla_{\mathbf{y}^n}\eta$ *is in* L^4*, and* $g_{i,j}(\mathbf{y}^n)$ *and* $\|\tilde{\mathbf{y}}^n\|_1^2 x_j^n$ *are in* L^2 *for all* $i,j=1,\ldots,N+1$*. Here* $\tilde{\mathbf{y}}$ *are zero-mean fluctuations in a Reynolds decomposition* $\mathbf{y}=\langle\mathbf{y}\rangle+\tilde{\mathbf{y}}$*.*

Lemmas 3.5.1–3.5.3, whose proofs are provided in F, have the following implication. The problem's nonlinearity causes the ensemble mean solution $\langle\mathbf{y}^n\rangle$ to differ from its counterpart obtained by replacing the randomly fluctuating boundary function $\rho_0(t)$ with its ensemble mean $\langle\rho_0\rangle$. Therefore, the dependence of $(\mathbf{I}_{N+1}-\mathbf{M}_n)^{-1}\mathbf{P}_n$ on $\langle\mathbf{y}^n\rangle$ suggests that the stability of the DD algorithm with random boundary conditions is different from that of the corresponding DD algorithm with the ensemble-averaged boundary conditions.[4] This illustrates the potential pitfalls of applying DD algorithms verified for deterministic nonlinear problems to nonlinear problems whose dynamics is driven by random fluctuations.

3.6 Simulation Results and Discussion

In the simulations reported below, we set $\alpha_1 = 2L/5$ and $\alpha_2 = 3L/5$. This choice of relative thicknesses ensures that diffusion is nonlinear throughout most of the

[4] This is in contrast to the case of linear diffusion, where noise does not affect the stability of DD algorithms [115].

membrane (within the Pd layers) and facilitates our computational experiments. The resulting subdomains Ω_i ($i = 1,2,3$) are discretized with $N_1 = N_3 \equiv 2N/5$ and $N_2 \equiv N/5$ nodes, respectively. All the quantities are reported in their dimensionless form,

$$\hat{x} = \frac{x}{L_{\mathrm{Ta}}}, \quad \hat{t} = \frac{tD_{\mathrm{Pd}}^{\mathrm{int}}}{L_{\mathrm{Ta}}^2}, \quad \hat{\rho} = \rho V_{\mathrm{Pd}}, \tag{3.15}$$

and the dimensionless parameters are set to $\hat{L} = 5$, $\langle \hat{\rho}_0 \rangle = 0.55$ and $\hat{\rho}_L = 0.1$. In the following, we omit the hats $\hat{\cdot}$ to simplify the notation.

Realizations of the boundary noise $\eta(t)$ with infinite correlation time ($\lambda \to \infty$), i.e., of the random variable η, were drawn from a truncated Gaussian distribution with zero mean and variance σ_η^2 using the MATLAB code by Burkardt based on [63]. Two values of the coefficient of variation, $CV_\eta \equiv \sigma_\eta / \langle \rho_0 \rangle = 0.46$ and 0.23, were considered. For the exponential correlation function with finite λ (Section 3.6.4), realizations of $\eta(t)$ were drawn from the multivariate truncated Gaussian distribution using the MATLAB codes by Benham and Luong based on [100].

The convergence tolerance for the Newton solver was set to $\varepsilon = 10^{-3}$, and that for the Krylov solver (GMRES) in the JfNK coupling to $\varepsilon_{\mathrm{K}} = 10^{-6}$. We do not precondition the GMRES algorithm as it usually converged after only a few iterations without preconditioning. The simulation time horizon, T (= 20 or 40, depending on the experiment), was chosen to allow the system to approach its steady state. The ensemble mean $\langle \rho(x,t) \rangle$ and variance $\sigma_\rho^2(x,t)$ of concentration $\rho(x,t)$ were approximated by their sample counterparts calculated from N_{sam} independent samples. This number was determined from the following conditions:

1. The difference between the prescribed ensemble average of $\rho_0(t)$ and its sample counterpart is less than $5 \cdot 10^{-3}$ at representative times $t_k = kT/5$ ($k = 1, \ldots, 5$).

2. The difference between the sample averages computed with N_{sam} and $N_{\mathrm{sam}} - 10$

realizations is less than 10^{-3} at times $t_k = kT/5$ ($k = 1,\ldots,5$).

We found these conditions to be satisfied with $N_{\text{sam}} = 1200$ for $CV_\eta = 0.23$, and $N_{\text{sam}} = 2500$ for $CV_\eta = 0.46$. The computations were performed on an Intel Core i7 machine running at 4 GHz.

3.6.1 Comparison between path-wise and moment-wise coupling

We compare the relative performance of the DD algorithms with the path-wise and moment-wise communication strategies. A single-solver ("global") method provides reference solutions $\langle\rho(x_i,t)\rangle^{\text{ref}}$ and $[\sigma_\rho^2(x_i,t)]^{\text{ref}}$ with $x_i = i\Delta x$ ($i = 0,\ldots,N$) and $t = jT/5$ ($j = 1,3,5$) for $T = 20.0$. These solutions are computed from $N_{\text{sam}} = 1200$ or 2500 (for $CV_\eta = 0.23$ or 0.46) independent runs of a single explicit Euler solver with grid cell size Δx^{ref} and time step Δt^{ref}. The latter is chosen to satisfy the stability condition $\Delta t^{\text{ref}} < (\Delta x^{\text{ref}})^2/(2D_{\max})$, where $D_{\max}$ is the maximum value of the diffusion coefficient throughout the computational domain over the entire duration of the simulation.

A position-dependent relative error $\mathcal{E}_\rho$ is defined, for $i = 1,\ldots,N-1$, as

$$\mathcal{E}_\rho(x_i,t;\Delta x^{\text{ref}},\Delta t^{\text{ref}},\Delta x,\Delta t_{\text{com}}) = \frac{|\langle\rho(x_i,t)\rangle - \langle\rho(x_i,t)\rangle^{\text{ref}}|}{\langle\rho(x_i,t)\rangle^{\text{ref}}} \tag{3.16}$$

where $\langle\rho(x_i,t)\rangle$ is computed with the DD algorithm on a grid of cell size $\Delta x \geq \Delta x^{\text{ref}}$ and with inter-solver communication time $\Delta t_{\text{com}} \geq \Delta t^{\text{ref}}$. For a fully-converged (implicit) JfNK coupling, the total relative error $\mathcal{E}_\rho^{\text{im}}$ is

$$\mathcal{E}_\rho^{\text{im}}(x_i,t;\Delta x^{\text{ref}},\Delta t^{\text{ref}},\Delta x,\Delta t_{\text{com}}) = \mathcal{E}_c^{\text{im}}(x_i,t) + \mathcal{E}_{\Delta x}^{\text{im}}(x_i,t) + \mathcal{E}_{\Delta t_{\text{com}}}^{\text{im}}(x_i,t), \tag{3.17}$$

where $\mathcal{E}_c^{\text{im}}$ is the error solely due to the use of the implicit coupling, i.e., of DD solutions with $\Delta t_{\text{com}} = \Delta t^{\text{ref}}$ and $\Delta x = \Delta x^{\text{ref}}$; $\mathcal{E}_{\Delta x}^{\text{im}}$ is the error due to the use of a coarser mesh with

$\Delta x > \Delta x^{\text{ref}}$; and $\mathcal{E}^{\text{im}}_{\Delta t_{\text{com}}}$ is the error due to the use of a $\Delta t_{\text{com}} > \Delta t^{\text{ref}}$. To compare the path-wise and moment-wise implicit coupling, we focus on $\mathcal{E}^{\text{im}}_{c}$ by using the fully-converged JfNK coupling with one micro-step per macro-step, $\Delta t_{\text{com}} = \Delta t_i = \Delta t^{\text{ref}}$ (i.e., $n_i = 1$ for $i = 1,2,3$) so that $\mathcal{E}^{\text{im}}_{\Delta t_{\text{com}}} = 0$, and a mesh with $\Delta x = \Delta x^{\text{ref}}$ so that $\mathcal{E}^{\text{im}}_{\Delta x} = 0$.

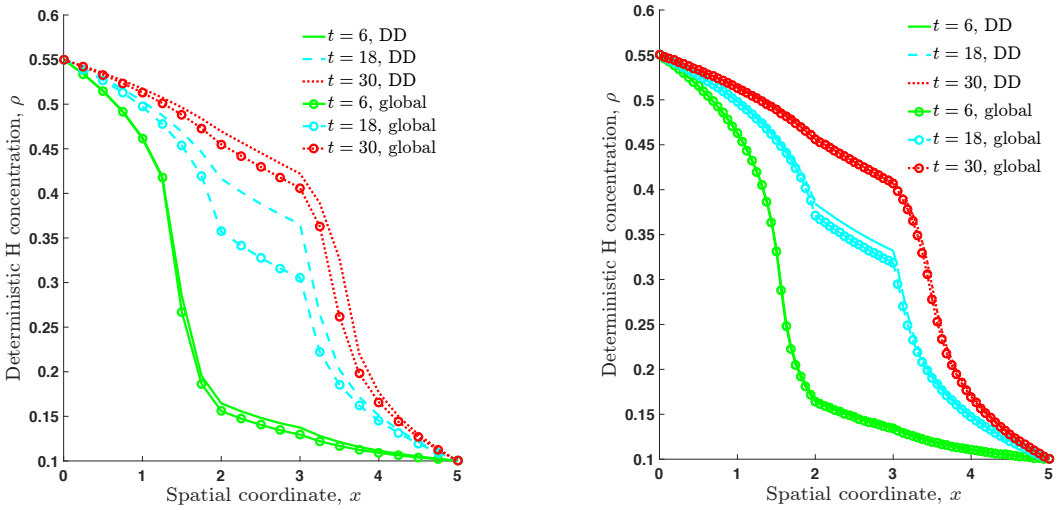

Figure 3.4: Temporal snapshots of the deterministic concentration, $\rho(x,t)$, computed with the DD and global algorithms using $\Delta x^{\text{ref}} = 0.25$ (left) and $\Delta x^{\text{ref}} = 0.0625$ (right).

First, we consider the deterministic boundary condition $\rho_0 = 0.55$, for which the DD algorithms with path-wise and moment-wise communication are equivalent, $\mathcal{E}_\rho$ is defined in terms of $\rho(x_i,t)$ rather than $\langle\rho(x_i,t)\rangle$, and $T = 30.0$ was required to approach steady state. Figure 3.4 exhibits temporal snapshots of the concentration profiles, $\rho(x,t)$, computed with the DD and global methods. The discrepancy between the two methods, i.e., $\mathcal{E}^{\text{im}}_{c}$, is maximal at intermediate times ($t = 18$) and is reduced, at all times, by refining the mesh from $\Delta x^{\text{ref}} = 0.25$ to $\Delta x^{\text{ref}} = 0.0625$. This mesh refinement necessitates the reduction of the time step, from $\Delta t^{\text{ref}} = 0.01$ to $\Delta t^{\text{ref}} = 5 \times 10^{-4}$, to satisfy the stability condition $\Delta t^{\text{ref}} < (\Delta x^{\text{ref}})^2/(2D_{\max})$.

Next, we consider the infinitely correlated boundary noise, i.e., the case $\rho_0 = \langle\rho_0\rangle + \eta_0$ with the zero-mean truncated Gaussian variable η_0. Figure 3.5 shows the

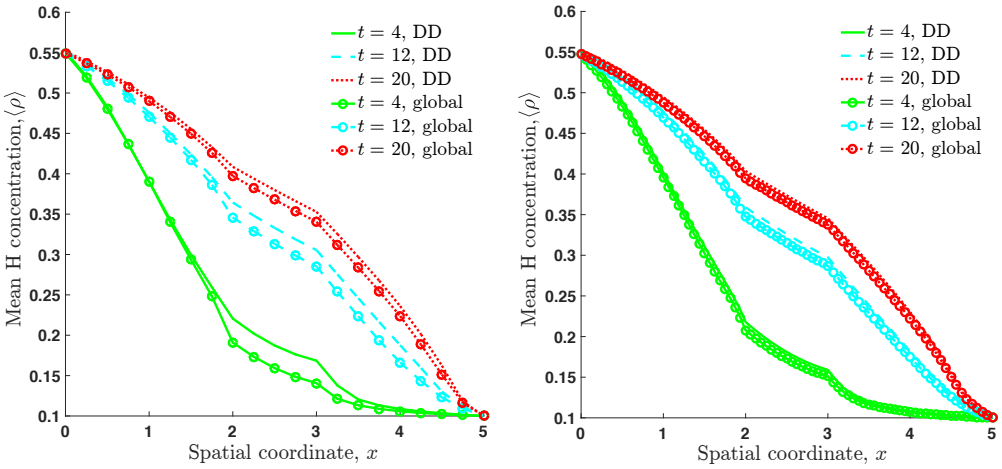

Figure 3.5: Temporal snapshots of the mean concentration, $\langle \rho(x,t) \rangle$, computed (for $CV_\eta = 0.23$) with the path-wise DD and global algorithms using $\Delta x^{\text{ref}} = 0.25$ (left) and $\Delta x^{\text{ref}} = 0.0625$ (right).

mean concentration profiles, $\langle \rho(x,t) \rangle$, computed with the path-wise DD and global algorithms, for $CV_\eta = 0.23$. The impact of time t on the discrepancy between the two solutions with $\Delta x^{\text{ref}} = 0.25$, i.e., on $\mathcal{E}_c^{\text{im}}$ is less pronounced than in the deterministic case, indicating a "smoothing effect" of the noise. As before, the mesh refinement, i.e., smaller Δx^{ref}, reduces $\mathcal{E}_c^{\text{im}}$. Figure 3.6 demonstrates that the mesh refinement also reduces the discrepancy between the DD and global solutions for the concentration variance $\sigma_\rho^2(x,t)$. Although not shown here, the case of $CV_\eta = 0.46$ yielded similar results for both the mean and variance.

Finally, we consider the performance of the DD algorithm with the moment-wise communication. Figure 3.7 reveals that this approach yields an inaccurate mean solution $\langle \rho(x,t) \rangle$ even for a moderate noise strength ($CV_\eta = 0.23$), and this solution becomes unphysical at later times ($t = 20$) as the noise strength increases ($CV_\eta = 0.46$). Although not shown here, reducing Δx^{ref} from 0.25 to 0.0625 does not yield any improvement. This finding eliminates the possibility of using the moment-wise coupling for our DD algorithm. In the remainder of this paper, we therefore exclusively use the path-wise

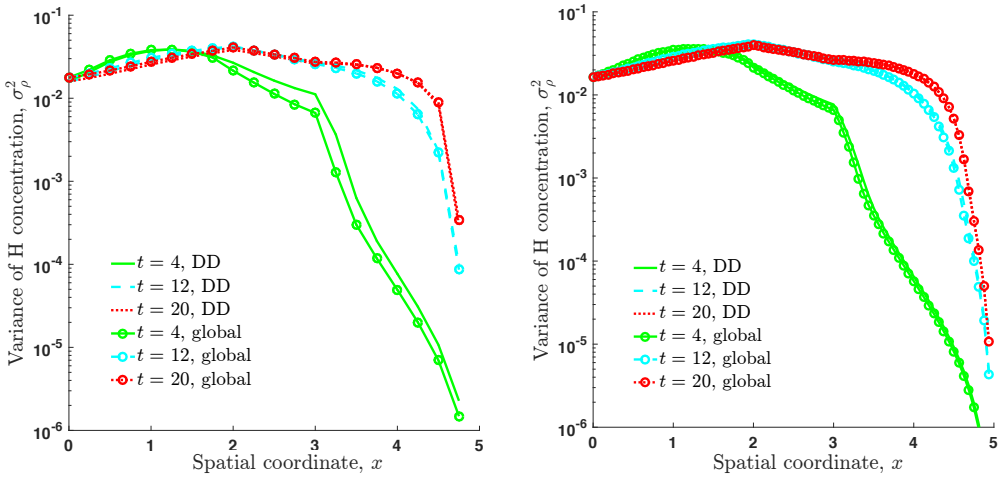

Figure 3.6: Temporal snapshots of the concentration variance, $\sigma_\rho(x,t)^2$, computed (for $CV_\eta = 0.23$) with the path-wise DD and global algorithms using $\Delta x^{\text{ref}} = 0.25$ (left) and $\Delta x^{\text{ref}} = 0.0625$ (right).

communication.

3.6.2 Relative performance of implicit and explicit coupling

Inter-solver communications can represent a large fraction of the overall computational cost of a multiphysics simulation. This cost can be mitigated by using an incomplete (i.e., partially converged) iteration or reducing the inter-solver communication frequency, both of which may lead to a higher solution error and might introduce instabilities. To explore this efficiency/accuracy trade-off, we consider the fully-converged (implicit) coupling with two communication frequencies, $\Delta t_{\text{com}} = 4.0$ and $\Delta t_{\text{com}} = 8.0$, and in each case compare its efficiency with that of a single-iteration (explicit) coupling. In these simulations, we set $\Delta x = 0.25$, $\Delta t_1 = \Delta t_3 = 0.01$, $\Delta t_2 = 0.05$, and $T = 40.0$ (to enable testing of bigger Δt_{com} values, even though steady state is approached around $t = 20$). Similar to (3.16) and (3.17), we define relative errors of the implicit ($\mathcal{E}_\rho^{\text{im}}$) and

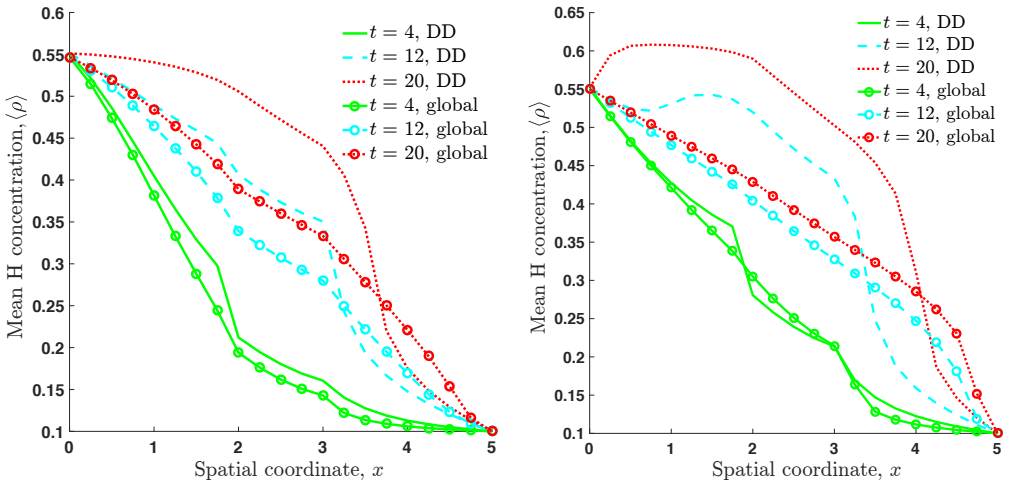

Figure 3.7: Temporal snapshots of the mean concentration, $\langle \rho(x,t) \rangle$, for $CV_\eta = 0.23$ (left) and $CV_\eta = 0.46$ (right), computed with the moment-wise DD using $\Delta x^{\text{ref}} = 0.25$.

explicit ($\mathcal{E}_\rho^{\text{ex}}$) path-wise coupling strategies,

$$\mathcal{E}_\rho^{\text{im}} = \mathcal{E}_c^{\text{im}}(x_i,t) + \mathcal{E}_{\Delta x}^{\text{im}}(x_i,t) + \mathcal{E}_{\Delta t_{\text{com}}}^{\text{im}}(x_i,t), \tag{3.18a}$$

$$\mathcal{E}_\rho^{\text{ex}} = \mathcal{E}_c^{\text{ex}}(x_i,t) + \mathcal{E}_{\Delta x}^{\text{ex}}(x_i,t) + \mathcal{E}_{\Delta t_{\text{com}}}^{\text{ex}}(x_i,t), \tag{3.18b}$$

wherein the reference solutions $\langle \rho(x_i,t) \rangle^{\text{ref}}$ are the "exact" profiles obtained with a single explicit Euler solver on a fine space-time mesh of $\Delta x^{\text{ref}} = 0.0625$ and $\Delta t^{\text{ref}} = 5 \times 10^{-4}$.

Figure 3.8 shows the errors $\mathcal{E}_\rho^{\text{im}}$ and $\mathcal{E}_\rho^{\text{ex}}$ when the implicit coupling has communication time $\Delta t_{\text{com}} = 4.0$ or 8.0, and the noise strength $CV_\eta = 0.23$. Using an identical communication time as its implicit counterpart, the explicit coupling yields a relative error $\mathcal{E}_\rho^{\text{ex}}$ that can be more than an order of magnitude higher than that of the implicit coupling, $\mathcal{E}_\rho^{\text{im}}$. This difference in solution error can be negated by reduction of Δt_{com}, which, however, increases simulation time t_{sim} (Table 3.1). This causes the explicit coupling to become slower than its implicit counterpart. As Table 3.1 shows, for the higher noise strength $CV_\eta = 0.46$ we find a similar behavior. When the implicit cou-

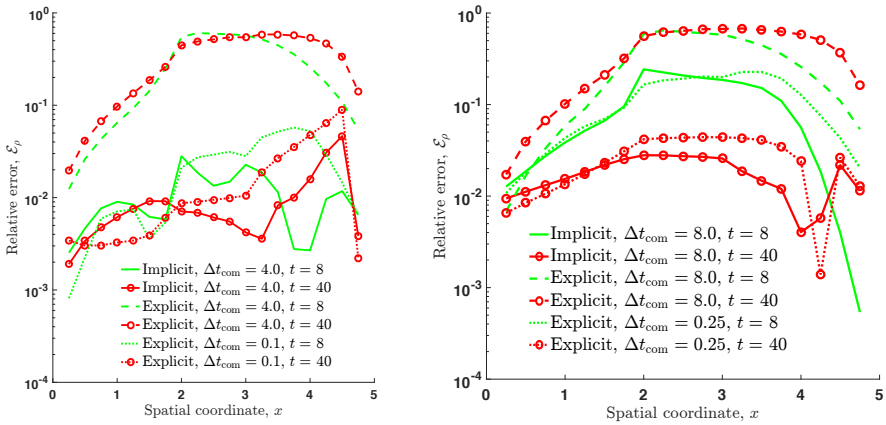

Figure 3.8: Spatial variability of the relative errors $\mathcal{E}_\rho^{\text{im}}$ and $\mathcal{E}_\rho^{\text{ex}}$ for $\Delta t_{\text{com}} = 4.0$ (implicit, explicit) and $\Delta t_{\text{com}} = 0.1$ (explicit) on the left, and for $\Delta t_{\text{com}} = 8.0$ (implicit, explicit) and $\Delta t_{\text{com}} = 0.25$ (explicit) on the right. In both cases $CV_\eta = 0.23$.

pling has a communication time $\Delta t_{\text{com}} = 8.0$, it again outperforms the explicit coupling at both coefficients of variation considered. Finally, the simulation time when using implicit coupling is smaller than its counterpart for $CV_\eta = 0.23$ for both $\Delta t_{\text{com}} = 4.0$ and $\Delta t_{\text{com}} = 8.0$. As shown in Figure 3.9 for $\Delta t_{\text{com}} = 8.0$, this can be attributed to the fact that a higher noise strength decreases the N_{sam}-averaged number of iterations per communication throughout the entire simulation, indicating again the smoothing effect of the noise (see also Section 3.6.1).

Table 3.1: N_{sam}-averaged simulation time t_{sim} (in s) with implicit and explicit coupling for several communication frequencies Δt_{com}.

Coupling	Δt_{com}	Simulation time, t_{sim}	
		$CV_\eta = 0.23$	$CV_\eta = 0.46$
implicit	4.0	4.5	3.5
explicit	0.1	11.2	10.7
implicit	8.0	7.6	6.5
explicit	0.25	8.6	8.8

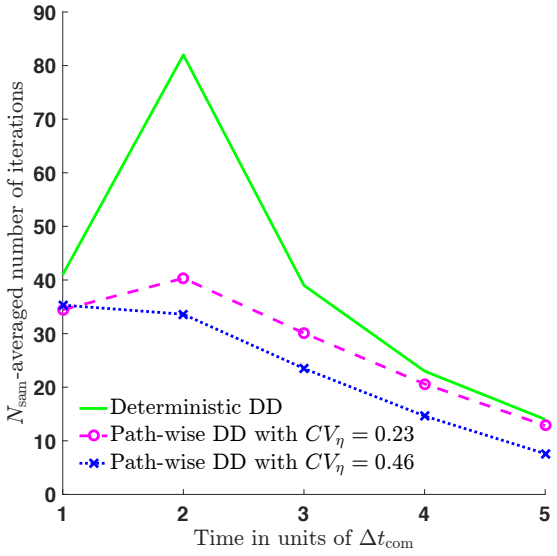

Figure 3.9: Sample-averaged number of iterations as a function of time (in units of $\Delta t_{\text{com}} = 8.0$) for a deterministic DD coupling and a path-wise DD coupling with several values of CV_η.

3.6.3 Temporal order of accuracy of implicitly coupled DD algorithm

We express the temporal order of accuracy of our DD algorithm in terms of the l^2-norm error $\mathcal{E}_{l^2} = \|\langle\rho\rangle - \langle\tilde{\rho}\rangle\|_{l^2}$ over the entire simulation domain Ω. Here $\langle\rho\rangle$ is the N_{sam}-average of the DD solution ρ obtained using JfNK coupling with $\varepsilon = 10^{-3}$, $\Delta x = 0.25$, $\Delta t_i \equiv \Delta t$ and $n_i = 1$ $(i = 1,2,3)$; and $\langle\tilde{\rho}\rangle$ is the ensemble average of the exact solution to the set of nonlinear ODEs (3.4) with $\Delta x = 0.25$, and is approximated by the N_{sam}-average of the implicitly (JfNK coupling with $\varepsilon = 10^{-3}$) coupled DD solution $\tilde{\rho}$ with spatial mesh size $\Delta\tilde{x} = 0.25$, micro-steps $\Delta\tilde{t}_i = 10^{-5}$ and $\tilde{n}_i = 1$ $(i = 1,2,3)$. We only consider perfectly correlated noise.

Figure 3.10 demonstrates, for $CV_\eta = 0.46$, that sequential reduction of Δt by a factor of two results in a near-quadratic decrease in $\mathcal{E}_{l^2}$. A similar result was obtained for $CV_\eta = 0.23$. It follows that the implicit coupling preserves the second-order local (i.e.,

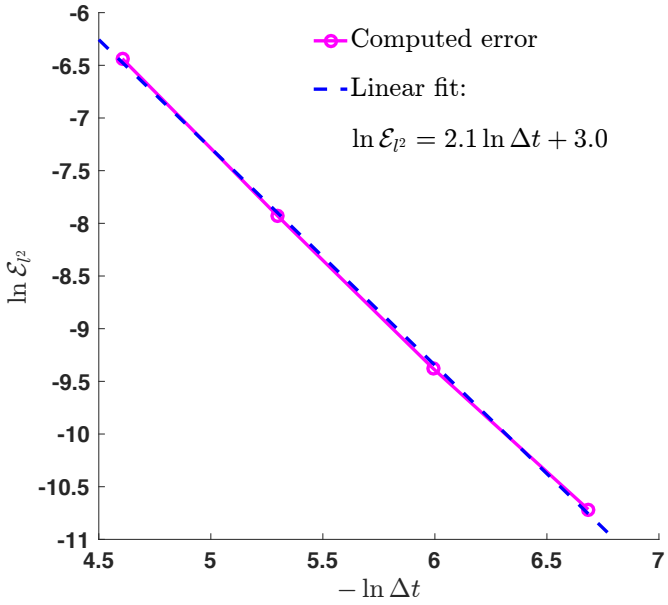

Figure 3.10: The l^2-norm error of the N_{sam}-averaged solution, $\langle \rho(x,t) \rangle$, as a function of the micro-step size Δt.

first-order global) order of accuracy of the subdomain solvers for all the boundary noise strengths considered.

3.6.4 Effect of finite noise correlation time

To generate the temporally fluctuating truncated Gaussian boundary noise $\eta(t)$ with (dimensionless) correlation time λ, we first consider a discrete version of the auto-covariance of the parent multivariate-Gaussian field $p(t)$,

$$C_p(|t_i - t_j|) = \sigma_p^2 \exp(-|i-j|\Delta t_1/\lambda), \tag{3.19}$$

where $t_i - t_j = (i-j)\Delta t_1$ with $0 < t_i, t_j \leq T$. It is used to build an $N_t \times N_t$ covariance matrix Σ_p, where N_t is the total number of discrete time steps in the subdomain Ω_1 over the simulation horizon T. The latter is then transformed into Σ_η, an $N_t \times N_t$ covariance

matrix of the truncated multivariate-Gaussian field $\eta(t)$. Finally, this matrix is used to generate N_{sam} realization arrays $\{\eta_1,\ldots,\eta_{N_t}\}$ with $\eta_i = \eta(t_i)$ at discrete times t_i for $i = 1,\ldots,N_t$. Each of the corresponding realizations of $\rho(x,t)$ was computed with discretization parameters $\Delta x = 0.25$ and $\Delta t_{\text{com}} = \Delta t_i = 10^{-2}$ $(i = 1,2,3)$; the simulation horizon was set to $T = 40.0$, and a value of 4000 was used for N_{sam}.

Figure 3.11 shows the resulting mean concentration $\langle\rho(x,t)\rangle$ computed with the implicitly coupled DD algorithm for $CV_\eta = 0.23$ and several values of λ. For $\lambda < 8.0$ the mean concentration profile becomes unphysical, indicating that the DD algorithm is not able to resolve boundary fluctuations with a correlation time smaller than eight times the characteristic diffusion time-scale $L_{\text{Ta}}^2/D_{\text{Pd}}^{\text{int}}$.

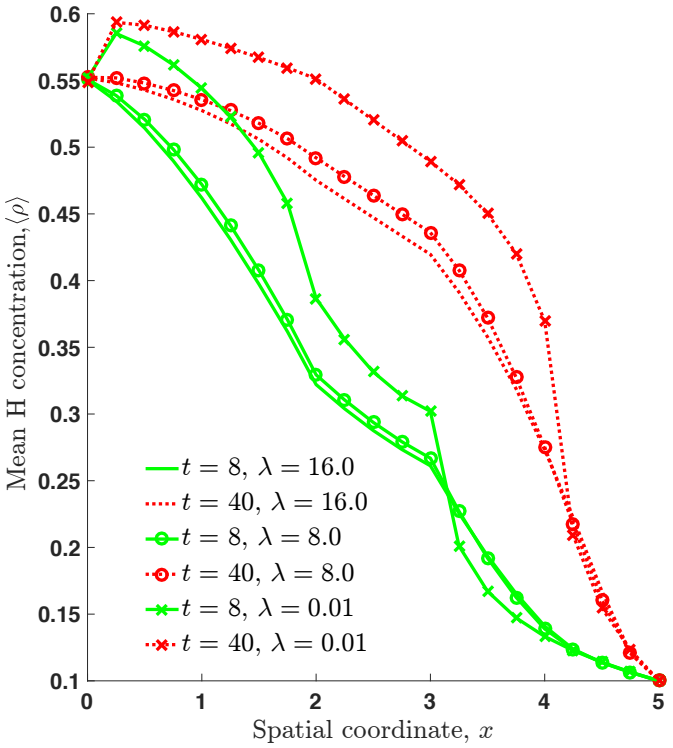

Figure 3.11: Effect of a finite noise correlation time on the mean concentration, $\langle\rho(x,t)\rangle$, for $CV_\eta = 0.23$.

3.7 Summary and Conclusions

We developed a domain-decomposition (DD) algorithm with a tight coupling based on a Jacobian-free Newton-Krylov (JfNK) method with generalized minimum residual. The DD algorithm was applied to a multiscale nonlinear diffusion problem driven by a truncated Gaussian noise at the boundary. For this problem, the DD components are coupled by enforcing the continuity of state variables, concentration ρ and flux $F = -D(\rho)\partial_x\rho$, at the interfaces between the DD subdomains. This may be done either path-wise (i.e., in each realization of the ensemble) or moment-wise (i.e., only for the ensemble average). The former strategy is exact but computationally intensive, while the latter is approximate but might be significantly faster. We explored the efficiency/accuracy trade-off between the fully-converged (implicit) JfNK coupling and its single-iteration (explicit) counterpart for different frequencies of communication between the DD components and for different strengths of the boundary noise, and analyzed the stability and order of accuracy of the implicit path-wise coupling.

Our analysis leads to the following major conclusions.

1. The DD approach with path-wise continuity provides accurate approximations of the mean and variance of $\rho(x,t)$ because it correctly propagates the boundary noise across the entire computational domain. As the spatial grid size and time step become smaller, the DD solutions for both moments converge to the reference solutions computed with a single-solver method.

2. The DD approach with moment-wise continuity fails to propagate the boundary noise into adjacent subdomains, resulting in erroneous solutions for both the mean and variance of $\rho(x,t)$.

3. The implicit coupling with path-wise continuity preserves the order of accuracy

of the constituent solvers, even for relatively high coefficients of variation of the boundary noise.

4. Despite a higher cost per communication, the fully-converged (implicit) JfNK coupling strategy outperforms its explicit counterpart at similar levels of solution error for all noise strengths considered. This is because the explicit coupling requires a higher inter-solver communication frequency to achieve the same error.

5. When the boundary fluctuations are correlated over a finite time, our path-wise DD approach correctly captures the time evolution of the mean concentration profile if the correlation time is larger than eight times the characteristic diffusion time-scale.

Future extensions of the presented analysis may include the application of our DD algorithm to higher-dimensional problems, and the development of moment-wise DD approaches in which continuity of not just the mean, but also higher moments, is enforced. As demonstrated in [31], enforcing continuity of mean and variance enables DD approaches for linear and weakly nonlinear systems to accurately propagate random fluctuations across interfaces between subdomains. Since highly nonlinear systems, such as the one considered here, are described by highly non-Gaussian state variables, their solutions likely require DD algorithms to enforce continuity of moments beyond the variance in order to adequately capture noise propagation throughout the entire computational domain.

3.8 Acknowledgements

The dissertation author thanks A.Y. Pigarov for bringing the hydrogen separation problem to his attention and for helpful discussions. The material contained in this Chapter is based on research supported in part by Defense Advanced Research Projects

Agency under the EQUiPS program and by the National Science Foundation under grant DMS-1522799. It is, in part, a reprint of the material as it appears in the paper "A tightly-coupled domain-decomposition approach for highly nonlinear stochastic multi-physics systems" by S. Taverniers and D.M. Tartakovsky, submitted in revised form for publication in *Journal of Computational Physics*, Elsevier Inc. [116]. The dissertation author is the primary investigator and author of this paper.

Chapter 4

Impact of parametric uncertainty on estimation of the energy deposition into an irradiated brain tumor

4.1 Abstract

We analyze the effect of parametric uncertainty on the total energy deposited in a brain tumor through X-ray irradiation. Both the location of the region over which a dose-enhancing, iodinated contrast agent spreads out after injection into the tumor, and the agent's concentration, are allowed to be uncertain. We model this problem via a probabilistic approach in which the coordinates of the center of the contrast agent region, as well as the effective atomic number in this area (which depends on the agent's concentration), are represented as mutually independent, uniformly distributed random variables. Employing the nonintrusive stochastic collocation (SC) method, we estimate statistical moments of the deposited energy as a function of the mean and/or variance of the random inputs. We find that in most cases, the coefficient of variation of the

uncertain parameters is amplified by the nonlinearity of the problem, yielding a larger coefficient of variation for the energy deposition. As the stochastic dimension increases, the magnitude of the predictive uncertainty in the energy deposition, as measured by its standard deviation, approaches that of the prediction (mean energy deposition) itself. This demonstrates that accurate prediction of the energy deposition requires a proper treatment of even small parametric uncertainty. Our analysis also reveals that SC outperforms standard Monte Carlo, with the largest difference in efficiency occurring for the case of a single uncertain parameter.

4.2 Introduction and Motivation

In addition to surgery and chemotherapy, radiotherapy has become one of the main treatment methods for brain tumors. It is aimed at either destroying the latter or preventing it from developing further, and may serve as the only treatment (e.g., for inoperable tumors) or in combination with surgery (to kill any remaining microscopic tumor cells) [9]. X-rays are the primary type of radiation involved in radiotherapy; their interactions with the medium in which they propagate produces energetic electrons which, in turn, lose their energy as they are slowed down through collisions. To enhance the energy dose absorbed by the tumor, a contrast agent based on a substance with high atomic number Z, such as iodine, may be injected [91]. This increases the photo-electric absorption of the X-rays (the photo-electric mass attenuation coefficient increases strongly with Z [80]), one of the main radiation-matter interactions at photon energies in the *keV* range. The photo-electric effect may also result in the emission of Auger electrons that may equally contribute to the overall energy deposition. Within the *keV* energy range, the X-rays are also likely to undergo Compton scattering and transfer their energy to existing free electrons; however, this process has a mass attenuation coefficient

that is nearly independent of Z [80].

To quantify the contrast agent's concentration within the brain/tumor, temporal and K-edge subtraction methods [39] have been employed, which inevitably introduce measurement errors preventing the agent's concentration upon injection into the tumor to be known exactly. Moreover, while the agent preferentially accumulates in the tumor interstitium due to the increased permeability of the broken blood-brain barrier, it may diffuse into the surrounding healthy brain tissue, preventing the agent's spatial extent throughout the irradiation procedure to be known with certainty. Along with other parameters whose values are not known exactly, such incomplete characterization of the contrast agent's presence within the brain/tumor renders predictions of the amount and detailed distribution of the energy deposited in the tumor, obtained through numerical simulations, uncertain.

The effect of iodine concentration on dose enhancement [107] and the uniformity of the dose distribution within the tumor [83] has been analyzed using the Monte Carlo N-particle (MCNP) [28, 123] method, which simulates trajectories of small photon packets created with their energy and propagation direction statistically selected. This allows modeling of the interaction of X-ray photons with the brain/tumor matter, and of the transport of electrons released as a result of those interactions. While such MCNP simulations may achieve high accuracy, this typically requires a large number of runs due to the slow convergence of Monte Carlo (the estimation error for the mean value of the quantity of interest decays as the inverse square root of the number of realizations [65]). Hence, as a means to determine bounds on, e.g., the dose absorbed by the tumor given certain bounds on the contrast agent's concentration and spatial extent within the brain, such methods may become prohibitively expensive.

To achieve a computationally more efficient approach to the estimation of energy deposition in X-ray irradiated brain tumors in the presence of the above mentioned para-

metric uncertainty, we combine a continuum-level approach involving a two-dimensional, flux-limited radiation-diffusion equation [74, 85] with the stochastic collocation (SC) technique [125, 77]). Similar to other nonintrusive methods such as multilevel Monte Carlo [51] as well as intrusive alternatives such as perturbation methods [76, 111, 62, 122] and polynomial chaos with stochastic Galerkin [120, 49, 126], SC aims to outperform standard Monte Carlo by a "smart" selection of sampling points and their weights. It uses a quadrature rule to approximate the weighted integrals appearing in the statistical moments of a quantity of interest. The choice of this rule is based on the distribution of the random parameter; e.g., Gauss-Hermite quadrature is employed for a normally distributed parameter, while Gauss-Legendre or Clenshaw-Curtis quadrature is used when the parameter is uniformly distributed [124, 40]. We assume that the X-rays are monochromatic and in the *keV* range [3] (such X-rays may be generated, e.g., in a synchrotron [29]), and represent the region over which the contrast agent spreads out as a square inset in the domain. We consider the coordinates of the center of this inset, as well as the effective atomic number in this area (which depends on the agent's concentration), to be uncertain. It follows that we cannot solve the radiation-diffusion equation in a deterministic setting, but need to recast it into a probabilistic framework wherein the uncertain parameters are treated as random variables. We apply SC to the resulting stochastic problem and analyze the dependence of statistical moments of a quantity of interest, the energy deposited in the tumor, on the mean and/or variance of the random inputs. We also compare the performance of SC with standard Monte Carlo for one, two or three random parameters.

The paper is organized as follows. In Section 4.3 we formulate the problem including its statistical parametrization. Section 4.4 details the spatial discretization method and time advancement algorithm used to obtain a numerical solution to the equilibrium radiation-diffusion equation, as well as the statistical moment estimators of

the deposited energy in the SC and standard Monte Carlo frameworks. In Section 4.5, we present the results of our numerical experiments. Section 4.6 formulates the conclusions of our numerical tests.

4.3 Problem Formulation

Under the conditions of local thermal equilibrium and isotropic, elastic scattering, radiation transport may be described by the radiative transfer equation [85]

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \mathbf{\Omega}\cdot\nabla I_\nu = -\kappa_{t,\nu} I_\nu + \kappa_{a,\nu} I_{\nu B}(T_m) + \frac{\kappa_{s,\nu}}{4\pi}\int_{4\pi} I_\nu \, \mathrm{d}\Omega', \tag{4.1}$$

where I_ν is the spectral radiation intensity; $I_{\nu B}$ is the Planck function; $\mathbf{\Omega}$ is the unit solid angle vector describing the direction of photon travel; $\kappa_{s,\nu}$, $\kappa_{a,\nu}$ and $\kappa_{t,\nu}$ are the spectral scattering, absorption and extinction coefficients, respectively; and T_m is the temperature of the medium through which the radiation propagates. Assuming that the collisional mean free path of the photons is small compared to the system (i.e., brain) dimensions, and that the radiation field is isotropic, (4.1) may be approximated by a diffusion equation. If we also consider the radiation field to be in thermal equilibrium with the medium, we do not need a separate equation for the latter's energy density. Under these assumptions, the transport of monochromatic[1] X-rays through a (human or animal) brain in a simplified two-dimensional, square geometry may be described through the equilibrium radiation-diffusion equation (eRDE)

$$\frac{\partial E}{\partial t} = \nabla\cdot[D(E,\nabla E,\mathbf{x})\nabla E], \qquad \mathbf{x}\in\mathcal{D}\equiv(0,L)\times(0,L), \quad t\in(0,T], \tag{4.2a}$$

[1]The use of monochromatic radiation (i.e., all photons have an identical energy) removes the need for frequency-averaging of the spectral radiation energy density, i.e., a gray approximation.

with E the radiation energy density and D the diffusion coefficient, which depends on $\mathbf{x} = (x_1, x_2)^\top$, E, and its spatial gradient, ∇E.

The domain $\mathcal{D}$, representing the brain, contains a subregion $\mathcal{T} = [L/2 - w/2, L/2 + w/2] \times [L/2 - w/2, L/2 + w/2]$, with $w = L/5$, representing the tumor (Fig. 4.1). Injection of an iodinated contrast agent causes the latter to spread out over a subdomain $\mathcal{D}_2 = [c_1 - w/2, c_1 + w/2] \times [c_2 - w/2, c_2 + w/2]$, which may fully or partially overlap with $\mathcal{T}$ depending on the values of c_1 and c_2 (Fig. 4.1). The medium in $\mathcal{D}_2$ has an "effective"[2] atomic number Z_2, which lies between the atomic number of the contrast agent, Z_c (which we take to be the atomic number of iodine, 53), and the effective atomic number of either healthy brain tissue (within $\mathcal{D}_2 \setminus \mathcal{T}$) or tumor matter (within $\mathcal{D}_2 \cap \mathcal{T}$). As healthy and tumurous brain matter have effective atomic numbers that are both small compared to Z_c, we assume they are equal and given by that of healthy brain tissue. The latter is taken to be 5, which is an average value for grey/white human brain matter over the *keV* range of photon energies [97]. It follows that Z_1, the effective atomic number within $\mathcal{D}_1 = \mathcal{D} \setminus \mathcal{D}_2$, is equal to 5.

Following [99], the diffusion coefficient D is given by

$$D(E, \nabla E, \mathbf{x}) = \frac{cE}{\{\gamma E / D_Z(Z(\mathbf{x}), E)\} + |\nabla E|}, \tag{4.2b}$$

where γ is defined in G, c is the speed of light, and $D_Z = Z^{-3} E^{3/4}$ is proportional to the diffusion coefficient without flux-limiting. The atomic number Z satisfies

$$Z(\mathbf{x}) = \begin{cases} Z_1 & \text{for } \mathbf{x} \in \mathcal{D}_1 \\ Z_2 & \text{for } \mathbf{x} \in \mathcal{D}_2. \end{cases} \tag{4.2c}$$

[2] A composite medium may be characterized by an effective atomic number, which depends on its composition and the energy of the X-ray photons.

Important steps in the derivation of (4.2) from (4.1) are given in G. Equation (4.2) is subject to the initial condition

$$E(\mathbf{x},0) = E_{\text{in}}(\mathbf{x}), \qquad \mathbf{x} \in \mathcal{D} \tag{4.3a}$$

and boundary conditions

$$E - 2D(E,\nabla E,\mathbf{x})\partial_{x_1} E = 4F_{\text{in}}, \qquad \mathbf{x} \in \Gamma_{\text{R,left}}, \tag{4.3b}$$

$$E + 2D(E,\nabla E,\mathbf{x})\partial_{x_1} E = 4F_{\text{out}}, \qquad \mathbf{x} \in \Gamma_{\text{R,right}}, \tag{4.3c}$$

$$\partial_{x_2} E = 0, \qquad \mathbf{x} \in \Gamma_{\text{N}}, \tag{4.3d}$$

with F_{in} the half-range flux in the direction of positive x_1 on $\Gamma_{\text{R,left}}$, and F_{out} the half-range flux in the direction of negative x_1 on $\Gamma_{\text{R,right}}$, respectively. Here $\Gamma_{\text{R}} = \Gamma_{\text{R,left}} \cup \Gamma_{\text{R,right}}$ with $\Gamma_{\text{R,left}} = \{\mathbf{x} : x_1 = 0, 0 < x_2 < L\}$ and $\Gamma_{\text{R,right}} = \{\mathbf{x} : x_1 = L, 0 < x_2 < L\}$, and $\Gamma_{\text{N}} = \Gamma_{\text{N,top}} \cup \Gamma_{\text{N,bottom}}$ with $\Gamma_{\text{N,top}} = \{\mathbf{x} : x_2 = L, 0 < x_1 < L\}$ and $\Gamma_{\text{N,bottom}} = \{\mathbf{x} : x_2 = 0, 0 < x_1 < L\}$, denote the Robin and homogeneous Neumann segments of the domain boundary $\partial\mathcal{D}$, respectively.

Predictions of the behavior of complex systems such as the one considered here are typically uncertain due to insufficient characterization of the input parameters (parametric uncertainty) and/or errors in the model representation of the true physical system (structural uncertainty). While the latter is introduced through use of a radiation-diffusion equation (which has a much lower fidelity than the radiative transfer equation) and of a simplified geometry for the brain/tumor/contrast agent configuration, we only consider the parametric uncertainty discussed in Section 4.2. We represent the latter in our model through uncertainty in the location of the center of $\mathcal{D}_2$, as well as in the effective atomic number Z_2 within $\mathcal{D}_2$. The latter accounts for the uncertainty in the

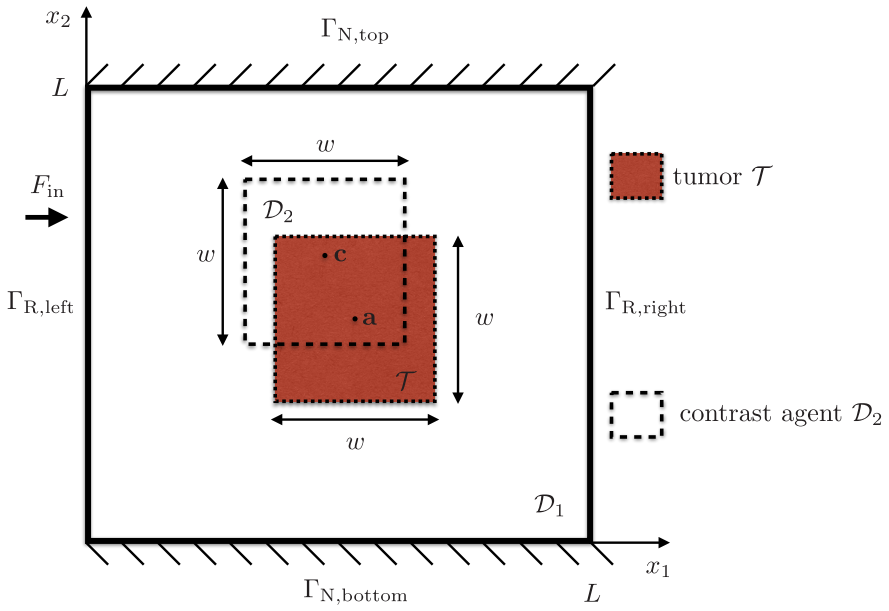

Figure 4.1: Computational domain $\mathcal{D}$ for the radiation-diffusion problem. Here $\mathbf{a} = (L/2, L/2)^\top$ and $\mathbf{c} = (c_1, c_2)^\top$, and $\mathcal{D}_1 = \mathcal{D} \setminus \mathcal{D}_2$.

contrast agent's concentration, which renders the composition of the medium within $\mathcal{D}_2$ uncertain. To simplify our representation, we assume that the concentration of the contrast agent is uniform over $\mathcal{D}_2$ and constant over time (in reality, the agent may be injected gradually over time during irradiation); we also ignore any spatial variations in the composition of the brain and tumorous tissue. Finally, we treat the brain tumor's location $\mathcal{T}$ (which may be determined via X-ray-based Computed Tomography [57], or through alternative techniques such as Magnetic Resonance Imaging [17]), and the healthy brain/tumorous matter's effective atomic number Z_1, as known quantities.

We represent the uncertain coordinates c_1 and c_2 of the center of $\mathcal{D}_2$, and the uncertain effective atomic number Z_2, as random variables. Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability triplet with Ω the event space, $\mathcal{F} \subseteq 2^\Omega$ the σ-field of Ω, and $\mathcal{P} : \mathcal{F} \to [0,1]$ the probability measure. Defining c_1, c_2 and Z_2 on Ω, i.e., $c_1 : \Omega \to \mathbb{R}$, $c_2 : \Omega \to \mathbb{R}$ and

$Z_2 : \Omega \to \mathbb{R}$ recasts (4.2) into the following stochastic boundary-value problem:

$$\frac{\partial E}{\partial t} = \nabla \cdot [D(E, \nabla E, \mathbf{x}, \omega)\nabla E], \qquad \mathbf{x} \in \mathcal{D}, \quad \omega \in \Omega, \quad t \in (0, T], \tag{4.4a}$$

where $D = cE[\ \{\gamma E / D_Z(Z(\mathbf{x}, \omega), E)\} + |\nabla E|\ ]^{-1}$, and $D_Z = Z^{-3}E^{3/4}$ with Z given by

$$Z(\mathbf{x}, \omega) = \begin{cases} Z_1 & \text{for } \mathbf{x} \in \mathcal{D}_1 \\ Z_2(\omega) & \text{for } \mathbf{x} \in \mathcal{D}_2. \end{cases} \tag{4.4b}$$

Initial and boundary conditions are provided by (4.3), provided that we take into account the dependence on ω. The state variable $E \equiv E(\mathbf{x}, t, \omega) : \mathcal{D} \times (0, T] \times \Omega \to \mathbb{R}$ is a stochastic measurable function that satisfies (4.4)–(4.3) almost surely.

To ensure $\mathcal{D}_2 \subset \mathcal{D}$, c_1 and c_2 need to lie in the interval $[w/2, L - w/2]$; to obtain a more reasonable scenario where $\mathcal{D}_2$ overlaps with at least one quarter of the tumor region $\mathcal{T}$, we require $c_1 \in [L/2 - c_{\mathrm{H}}, L/2 + c_{\mathrm{H}}]$ and $c_2 \in [L/2 - c_{\mathrm{V}}, L/2 + c_{\mathrm{V}}]$ with $0 < c_{\mathrm{H}} \leq w/2$ and $0 < c_{\mathrm{V}} \leq w/2$. The minimum and maximum theoretical bounds on the effective atomic number within $\mathcal{D}_2$ are Z_1 (no contrast agent) and Z_c (only contrast agent). However, to achieve a more realistic composition of the medium within $\mathcal{D}_2$, we require Z_2 to lie within $[Z_{\min}, Z_{\max}]$ where $Z_{\min}$ and $Z_{\max}$ are, respectively, 10 and 40. We assume that all three parameters c_1, c_2 and Z_2 are uniformly distributed on their respective intervals, and are mutually independent. The vector $\boldsymbol{\xi} = (c_1, c_2, Z_2)^\top$ is then characterized by a joint pdf

$$\rho(\mathbf{s}) = \prod_{i=1}^{3} \rho_i(s_i), \tag{4.5}$$

with support $\Lambda = [L/2 - c_{\mathrm{H}}, L/2 + c_{\mathrm{H}}] \times [L/2 - c_{\mathrm{V}}, L/2 + c_{\mathrm{V}}] \times [Z_{\min}, Z_{\max}]$, where ρ_1, ρ_2 and ρ_3 are the marginal probability density functions (PDFs) representing the univariate

uniform distributions of c_1, c_2 and Z_2, respectively. The size of $\boldsymbol{\xi}$ is called the *stochastic dimension*, and we will denote it by M. According to the Doob-Dynkin lemma, we may then substitute $E(\mathbf{x},t,\boldsymbol{\xi}(\omega)) : \mathcal{D}\times(0,T]\times\Lambda\to\mathbb{R}$ for $E(\mathbf{x},t,\omega)$. This effectively replaces the original probability triplet $(\Omega,\mathcal{F},\mathcal{P})$ with $(\Lambda,\mathcal{B},\rho(\boldsymbol{\xi})d\boldsymbol{\xi})$, where $\mathcal{B}$ is the Borel σ-algebra formed by all the open subsets of Λ. We call $E(\mathbf{x},t,\boldsymbol{\xi})$ the *stochastic response surface*, which satisfies

$$\frac{\partial E}{\partial t} = \nabla\cdot[D(E,\nabla E,\mathbf{x},\boldsymbol{\xi})\nabla E], \qquad \mathbf{x}\in\mathcal{D}, \quad \boldsymbol{\xi}\in\Lambda, \quad t\in(0,T] \tag{4.6}$$

subject to the initial and boundary conditions (4.3), modified to account for the dependence on $\boldsymbol{\xi}$.

Through some straightforward algebra, and omitting $\boldsymbol{\xi}$ for ease of notation, we may recast (4.6) into

$$\frac{\partial E^\dagger}{\partial t} = \nabla\cdot\left[\frac{cE^\dagger}{\{\gamma E^\dagger/D^\dagger_{\tilde{Z}}(\tilde{Z}(\mathbf{x}),E^\dagger)\}+|\nabla E^\dagger|}\nabla E^\dagger\right], \tag{4.7a}$$

where $D^\dagger_{\tilde{Z}} = \tilde{Z}^{-3}(E^\dagger)^{3/4}$ with $\tilde{Z}$ defined by

$$\tilde{Z}(\mathbf{x}) \equiv Z/Z_1 = \begin{cases} 1 & \text{for } \mathbf{x}\in\mathcal{D}_1 \\ Z_2/Z_1 & \text{for } \mathbf{x}\in\mathcal{D}_2, \end{cases} \tag{4.7b}$$

and $E^\dagger \equiv Z_1^{-4}E$. Next, using the characteristic length $w = L/5$ and the speed of light c, we define a dimensionless time $\tilde{t} = tc/w$. After some more algebra we find the dimensionless equation

$$\frac{\partial \tilde{E}}{\partial \tilde{t}} = \tilde{\nabla}\cdot\left[\frac{\tilde{E}}{\{\tilde{E}/\tilde{D}_{\tilde{Z}}(\tilde{Z}(\tilde{\mathbf{x}}),\tilde{E})\}+|\tilde{\nabla}\tilde{E}|}\tilde{\nabla}\tilde{E}\right] \tag{4.8a}$$

with

$$\tilde{D}_{\tilde{Z}} = \tilde{Z}^{-3}\tilde{E}^{3/4}, \quad \tilde{E} \equiv E^{\dagger}/E_0, \quad \tilde{\nabla} = w\nabla, \quad \tilde{\mathbf{x}} = \mathbf{x}/w. \tag{4.8b}$$

Here E_0 is a reference energy density satisfying $\gamma w/E_0^{3/4} = 1.0$, allowing us to eliminate γ in the denominator of the diffusion coefficient. In the following, all quantities are assumed to be in their dimensionless form, but we omit the tildes $\tilde{\cdot}$ for notational convenience.

The quantity of interest (QoI) in our problem is the total energy deposition in $\mathcal{T}$ over the time horizon T. The stochastic nature of E renders $E_{\text{abs,tot}}$ random as well. Hence, it may be characterized through its statistical moments, such as mean and variance. We compute these via the method of Stochastic Collocation, as discussed in Section 4.4.3.

4.4 Numerical Algorithm

4.4.1 Spatial discretization of the computational domain

We discretize the domain $\mathcal{D}$ by employing a cell-centered finite volume approach (CCFV) in which the numerical fluxes F are defined on the cell boundaries, and the value of E for a cell approximates the average value of the radiation energy density over that cell. We use a nonuniform mesh, finest within $\mathcal{D}_2 \cup \mathcal{T}$, with N grid cells in each spatial direction. This transforms (4.8) into the set of ordinary differential equations

$$\frac{\mathrm{d}E_{i,j}(t)}{\mathrm{d}t} = -\frac{F_{i+1/2,j}(t) - F_{i-1/2,j}(t)}{\Delta x_i} - \frac{F_{i,j+1/2}(t) - F_{i,j-1/2}(t)}{\Delta x_j}, \tag{4.9}$$

where the subscripts $i = 1,\ldots,N$ and $j = 1,\ldots,N$ indicate the position of a cell in the horizontal and vertical directions, respectively, and Δx_i and Δx_j are its dimensions in

these respective directions. We approximate the numerical flux $F_{i+1/2,j}$ using

$$F_{i+1/2,j} = -2D^{i+1/2,j}\frac{E_{i+1,j}-E_{i,j}}{\Delta x_{i+1}+\Delta x_i}, \tag{4.10}$$

where $D^{i+1/2,j}$ is defined as

$$D^{i+1/2,j} = \frac{E_{i+1,j}+E_{i,j}}{\{(E_{i+1,j}+E_{i,j})/D_{\mathrm{H}}^{i+1/2,j}\}+\{4|E_{i+1,j}-E_{i,j}|/(\Delta x_{i+1}+\Delta x_i)\}}. \tag{4.11}$$

Here $D_{\mathrm{H}}^{i+1/2,j}$ is given by the harmonic mean

$$D_{\mathrm{H}}^{i+1/2,j} = \frac{\Delta x_i \Delta x_{i+1} D_Z(Z_{i,j};E_{i,j}) D_Z(Z_{i+1,j};E_{i+1,j})}{\Delta x_{i+1} D_Z(Z_{i,j};E_{i,j})+\Delta x_i D_Z(Z_{i+1,j};E_{i+1,j})} \tag{4.12}$$

with $D_Z(Z_{i,j};E_{i,j}) = Z_{i,j}^{-3}E_{i,j}^{3/4}$. The other fluxes $F_{i-1/2,j}$, $F_{i,j+1/2}$ and $F_{i,j-1/2}$ are defined similarly. For the numerical flux at cell boundaries coinciding with $\partial\mathcal{D}$, we introduce ghost cells (denoted with subscript 0 or $N+1$ in one of the spatial dimensions) lying just outside $\mathcal{D}$ with the same dimensions (and the same effective atomic number) as the adjacent cells within $\mathcal{D}$. The boundary conditions (4.3b)–(4.3d) are then discretized according to

$$\frac{E_{0,j}+E_{1,j}}{2} - 2D^{0,j}\left(\frac{E_{1,j}-E_{0,j}}{\Delta x}\right) - 4F_{\mathrm{in}} = 0, \tag{4.13a}$$

$$\frac{E_{N+1,j}+E_{N,j}}{2} + 2D^{N+1,j}\left(\frac{E_{N+1,j}-E_{N,j}}{\Delta x}\right) - 4F_{\mathrm{out}} = 0, \tag{4.13b}$$

$$E_{i,0} = E_{i,1}, \qquad E_{i,N+1} = E_{i,N} \tag{4.13c}$$

with $i, j = 1, \ldots, N$. Here $D^{0,j}$ and $D^{N+1,j}$ are given, respectively, by

$$D^{0,j} = \frac{E_{0,j}}{\{E_{0,j}/D_Z(Z_{1,j}; E_{0,j})\} + \{|E_{1,j} - E_{0,j}|/\Delta x_{1,j}\}}, \tag{4.13d}$$

$$D^{N+1,j} = \frac{E_{N+1,j}}{\{E_{N+1,j}/D_Z(Z_{N,j}; E_{N+1,j})\} + \{|E_{N+1,j} - E_{N,j}|/\Delta x_{N,j}\}}. \tag{4.13e}$$

The Robin boundary conditions are nonlinear equations in $E_{0,j}$ or $E_{N+1,j}$ ($j = 1, \ldots, N$), which need to be solved iteratively during each time step. Figure 4.2 illustrates the relevant quantities involved in building the CCFV discretization of $\mathcal{D}$.

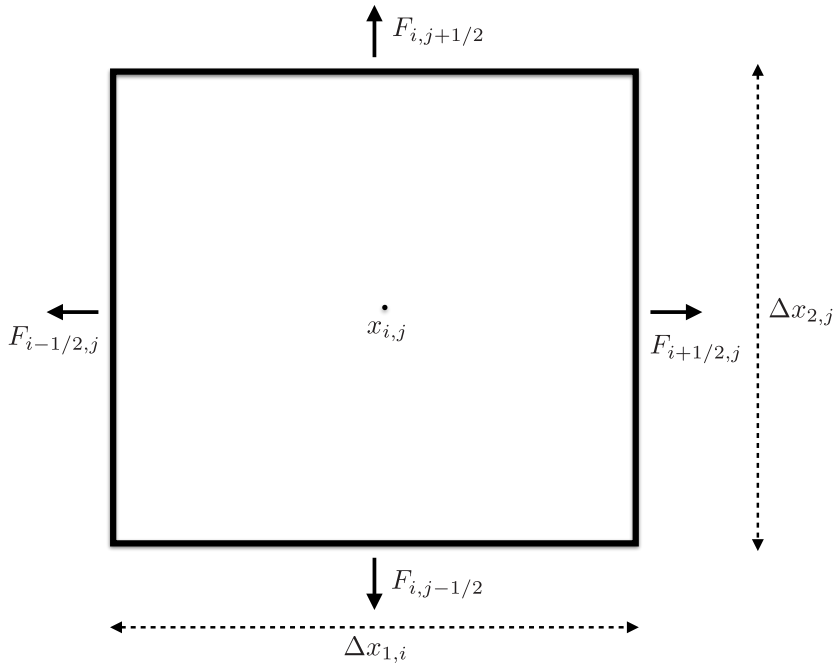

Figure 4.2: Cell-centered finite volume discretization of the computational domain $\mathcal{D}$.

4.4.2 Numerical time integration

We employ a first-order implicit Euler method which recasts (4.9) into

$$\frac{E_{i,j}^{n+1}-E_{i,j}^{n}}{\Delta t}=-\frac{F_{i+1/2,j}^{n+1}-F_{i-1/2,j}^{n+1}}{\Delta x_i}-\frac{F_{i,j+1/2}^{n+1}-F_{i,j-1/2}^{n+1}}{\Delta x_j}, \tag{4.14}$$

where the superscript n indicates the discrete time $t_n = n\Delta t$, with Δt the time step. As the right-hand-side of (4.14) involves implicitly treated nonlinear terms, we need to use an iterative algorithm to advance the solution from t_n to t_{n+1}. We employ a Jacobian-free Newton-Krylov (JfNK) solver [70] for this purpose, which avoids the explicit formation of the Jacobian while still benefiting from fast Newton-like convergence.

The time advancement of the solution from t_n to t_{n+1} consists of the following steps.

1. Initialize the Newton iterate by setting $E_{i,j,k=0}^{n+1} = E_{i,j}^{n}$ for all $i, j = 1, \ldots, N$, where k is the iteration number.

2. Inexactly solve the linear system $\mathbf{J}(\mathbf{E}_k^{n+1})\delta\mathbf{E}_k = -\mathbf{f}(\mathbf{E}_k^{n+1})$ for the Newton correction $\delta\mathbf{E}_k$ using the Krylov solver GMRES with tolerance ε_{K}. Here the components of $\mathbf{f}$ are given by

$$f_{j\cdot N+i}(\mathbf{E}_k^{n+1}) = \frac{E_{i,j,k}^{n+1}-E_{i,j}^{n}}{\Delta t}+\frac{F_{i+1/2,j,k}^{n+1}-F_{i-1/2,j,k}^{n+1}}{\Delta x_i} \tag{4.15}$$

$$+\frac{F_{i,j+1/2,k}^{n+1}-F_{i,j-1/2,k}^{n+1}}{\Delta x_j} \tag{4.16}$$

   with $i, j = 1, \ldots, N$.

3. Perform the Newton step $\mathbf{E}_{k+1}^{n+1} = \mathbf{E}_k^{n+1} + \delta\mathbf{E}_k$ where $\delta\mathbf{E}_k$ is the converged value of the kth Newton correction.

4. Perform steps 2 and 3 until a given tolerance ε is achieved.

5. Advance the solution to t_{n+1} by setting $\mathbf{E}^{n+1} = \mathbf{E}_K^{n+1}$ where K is the number of Newton iterations at convergence.

Details on our implementation of step 2 are provided in I.

4.4.3 Monte Carlo simulation and stochastic collocation

Statistical moments of $E(\mathbf{x},t,\boldsymbol{\xi})$, or a QoI derived from E, are its weighted integrals over the support Λ of $\boldsymbol{\xi}$ with respect to the latter's PDF defined in (4.5). For example, the ensemble mean and variance of the energy deposition in the tumor over time T, $E_{\text{abs,tot}}$, are defined, respectively, by

$$\langle E_{\text{abs,tot}} \rangle = \int_\Lambda E_{\text{abs,tot}}(\mathbf{s})\ \rho(\mathbf{s})\ \mathrm{d}\mathbf{s} \tag{4.17}$$

and

$$\sigma^2_{E_{\text{abs,tot}}} = \int_\Lambda E^2_{\text{abs,tot}}(\mathbf{s})\ \rho(\mathbf{s})\ \mathrm{d}\mathbf{s} - \langle E_{\text{abs,tot}} \rangle^2. \tag{4.18}$$

Monte Carlo simulation (MCS) approximates $\langle E_{\text{abs,tot}} \rangle$ and $\sigma^2_{E_{\text{abs,tot}}}$ through

$$\hat{E}^{\text{MC}}_{\text{abs,tot}} = \frac{1}{N_{\text{sam}}} \sum_{i=1}^{N_{\text{sam}}} E_{\text{abs,tot}}(\boldsymbol{\eta}_i) \tag{4.19}$$

and

$$\hat{\sigma}^{2,\text{MC}}_{E_{\text{abs,tot}}} = \frac{1}{N_{\text{sam}}} \sum_{i=1}^{N_{\text{sam}}} [E_{\text{abs,tot}}(\boldsymbol{\eta}_i)]^2 - [\hat{E}^{\text{MC}}_{\text{abs,tot}}]^2, \tag{4.20}$$

respectively. Here $\{\boldsymbol{\eta}_i\}_{i=1}^{N_{\text{sam}}}$ is a set of N_{sam} realizations of $\boldsymbol{\xi}$ sampled from (4.5).

The MCS estimation error for the mean, $\mathcal{E}^{\text{MC}}_{\text{est}} \equiv |\langle E_{\text{abs,tot}} \rangle - \hat{E}^{\text{MC}}_{\text{abs,tot}}|$, decays as $\sigma_{E_{\text{abs,tot}}}/\sqrt{N_{\text{sam}}}$ with $\sigma_{E_{\text{abs,tot}}}$ the standard deviation of $E_{\text{abs,tot}}$. While $\mathcal{E}^{\text{MC}}_{\text{est}}$ is independent

of the stochastic dimension M, the rate $1/\sqrt{N_{\text{sam}}}$ is very slow. Stochastic collocation (SC) aims to achieve faster convergence through the use of quadrature rules to approximate the weighted integrals in (4.17) and (4.18). An M-dimensional weighted integral of an integrable function $f(\mathbf{s})$ over its support Λ

$$\mathcal{I}_M[f] = \int_\Lambda f(\mathbf{s})\ \rho(\mathbf{s})\ \mathrm{d}\mathbf{s} \tag{4.21}$$

may be approximated by a quadrature formula $\mathcal{Q}_M$

$$\mathcal{I}_M[f] \approx \mathcal{Q}_M[f] = \sum_{i=1}^{P} w_i\ f(\mathbf{s}_i), \tag{4.22}$$

where $\mathbf{s}_i$ and w_i are the nodes and weights of the quadrature formula, respectively, and P is the number of nodes. The respective SC estimators for $\langle E_{\text{abs,tot}} \rangle$ and $\sigma^2_{E_{\text{abs,tot}}}$ are then given by

$$\hat{E}^{\text{SC}}_{\text{abs,tot}} = \sum_{i=1}^{P} w_i\ E_{\text{abs,tot}}(\mathbf{s}_i) \tag{4.23}$$

and

$$\hat{\sigma}^{2,\text{SC}}_{E_{\text{abs,tot}}} = \sum_{i=1}^{P} w_i\ [E_{\text{abs,tot}}(\mathbf{s}_i)]^2 - [\hat{E}^{\text{SC}}_{\text{abs,tot}}]^2. \tag{4.24}$$

By solving (4.8) at P judiciously chosen values $\mathbf{s}_i$ of $\boldsymbol{\xi}$, and using appropriate weights w_i, SC aims to achieve the same estimation error as MCS by using $P < N_{\text{sam}}$ collocation nodes.

For stochastic dimension $M > 1$, *product* or *tensor grid* rules may be constructed through the tensor product of M univariate rules [92]. However, the number of nodes in such grids increases exponentially with M, i.e., for $k+1$ nodes in each dimension it

behaves like $(k+1)^M$ (the so-called "curse of dimensionality"). Instead, *sparse grid* rules may be built from univariate rules using the Smolyak algorithm [104], whose number of nodes only increases as approximately $(2M)^k/k!$, i.e., polynomially with M.

4.5 Simulation Results and Discussion

As demonstrated in H, the dimensionless total energy deposition in the tumorous region $\mathcal{T}$ during $(0,T]$ is given by

$$E_{\text{abs,tot}} = \frac{\Delta t}{3} \sum_{i=1}^{I} \sum_{j} \sum_{k} Z_{j,k}^3 (E_{j,k}^{i+1})^{1/4} \Delta x_{1,j} \Delta x_{2,k}, \tag{4.25}$$

where we sum over the indices j and k corresponding to the grid cells located within $\mathcal{T}$. Since the nonuniform grid resulting from the CCFV discretization does not necessarily align with the boundary $\partial\mathcal{T}$ of $\mathcal{T}$, we define a uniform interpolation grid with $\Delta x_1^u = \Delta x_2^u = L/N$ that exactly lines up with $\partial\mathcal{T}$, and use the interpolated values of E and Z on this new grid to calculate $E_{\text{abs,tot}}$. Employing cubic spline interpolation for E and linear interpolation for Z was found to be optimal. The total energy deposition in $\mathcal{T}$ is then redefined as

$$E_{\text{abs,tot}} = \frac{\Delta t \Delta x_1^u \Delta x_2^u}{3} \sum_{i=1}^{I} \sum_{j=N_1^u+1}^{N_1^u+N_2^u} \sum_{k=N_1^u+1}^{N_1^u+N_2^u} Z_{j,k}^3 (E_{j,k}^{i+1})^{1/4}, \tag{4.26}$$

where $N_1^u = 2N/5$ and $N_2^u = N/5$ (the square region $\mathcal{T}$ is centered at $(L/2, L/2)^\top$ and has side $w = L/5$).

In order to quantify the error of the SC approximations of $\langle E_{\text{abs,tot}} \rangle$ and $\sigma^2_{E_{\text{abs,tot}}}$, we define the estimation errors

$$\mathcal{E}_{\text{est}}^{\text{SC}}(\langle E_{\text{abs,tot}} \rangle) = \left| \hat{E}_{\text{abs,tot}}^{\text{SC}} - \hat{E}_{\text{abs,tot}}^{\text{MC,ref}} \right| \tag{4.27}$$

and

$$\mathcal{E}_{\text{est}}^{\text{SC}}(\sigma^2_{E_{\text{abs,tot}}}) = \left|\hat{\sigma}^{2,\text{SC}}_{E_{\text{abs,tot}}} - \hat{\sigma}^{2,\text{MC,ref}}_{E_{\text{abs,tot}}}\right|. \tag{4.28}$$

Here $\hat{E}^{\text{MC,ref}}_{\text{abs,tot}}$ and $\hat{\sigma}^{2,\text{MC,ref}}_{E_{\text{abs,tot}}}$ are the "reference" MCS estimators of $\langle E_{\text{abs,tot}}\rangle$ and $\sigma^2_{E_{\text{abs,tot}}}$, respectively, which serve as surrogates for these quantities as the latter cannot be calculated analytically. We found that $N^{\text{ref}}_{\text{sam}} = 30000$ was an appropriate number of samples for converging the reference Monte Carlo estimators in all our numerical experiments except that for two uncertain parameters (c_1 and Z_2), where 35000 samples were needed. The estimation errors for MCS, $\mathcal{E}^{\text{MCS}}_{\text{est}}(\langle E_{\text{abs,tot}}\rangle)$ and $\mathcal{E}^{\text{MCS}}_{\text{est}}(\sigma^2_{E_{\text{abs,tot}}})$, are defined similarly. Figure 4.3 demonstrates, for the case where c_1, c_2 and Z_2 are all uncertain, that $\mathcal{E}^{\text{MCS}}_{\text{est}}(\langle E_{\text{abs,tot}}\rangle) \sim N^{-0.5}_{\text{sam}}$ as expected from theory [65].

Our goal is twofold: investigate the dependence of $\hat{E}^{\text{SC}}_{\text{abs,tot}}$ and $\hat{\sigma}^{2,\text{SC}}_{E_{\text{abs,tot}}}$ on the mean and variance of the uncertain parameters, and compare the efficiency of SC and MCS by analyzing how fast their estimation errors decay with the number of collocation nodes and Monte Carlo samples, respectively. For the latter exercise, we set the variance of each uncertain parameter at its maximum value established in Section 4.3, and consider SC to outperform MCS if its relative estimation error for the mean, $\mathcal{E}^{\text{SC}}_{\text{est}}(\langle E_{\text{abs,tot}}\rangle)/\hat{E}^{\text{MC,ref}}_{\text{abs,tot}}$, becomes of order 10^{-3} for P_{conv} nodes such that $P_{\text{conv}} < N_{\text{sam,conv}}$; here $N_{\text{sam,conv}}$ is the number of samples needed by MCS to reach the same relative error level.

We consider scenarios where either only c_1 or only Z_2 is uncertain (i.e., stochastic dimension $M = 1$), where both c_1 and Z_2 are uncertain ($M = 2$), and where all three input parameters c_1, c_2 and Z_2 are uncertain ($M = 3$). We model all uncertain parameters as uniformly distributed random variables over their respective ranges; an appropriate quadrature rule for the univariate uniform distribution is Clenshaw-Curtis [124, 40], whose nodes and weights are rescaled from the standard interval $[-1, 1]$ to the parameter

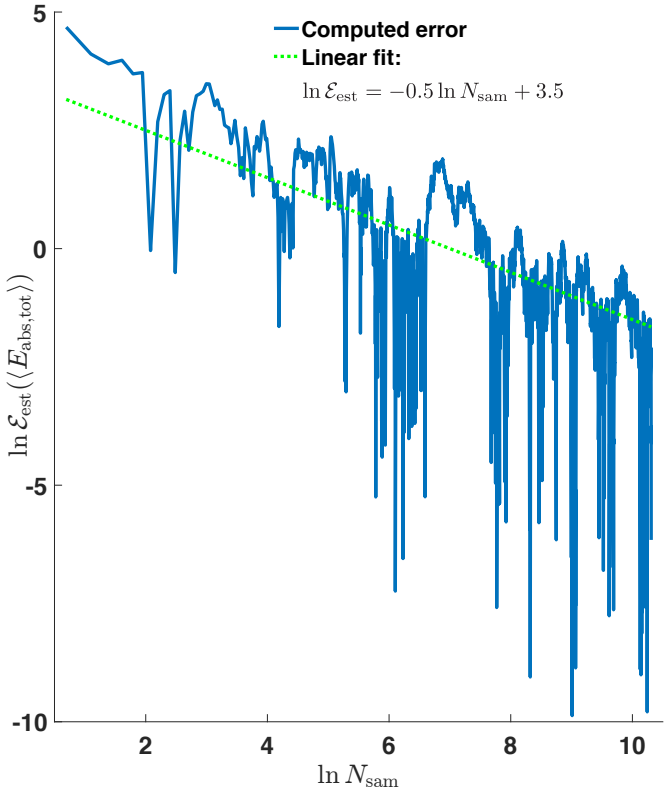

Figure 4.3: Estimation error of the mean energy deposition in the tumor for MCS as a function of the number of Monte Carlo samples for $M = 3$ with $c_{\text{H}} = c_{\text{V}} = 0.5$ and $[Z_{\text{min}}, Z_{\text{max}}] = [2, 8]$.

ranges we consider. For $M > 1$, we employ a sparse grid approach based on the Smolyak algorithm [104] with univariate Clenshaw-Curtis quadrature rules. This is implemented using MATLAB codes by Burkardt based on [90].

We set $Z_1 = 1$ and $N = 20$, and simulate the dynamics on a domain $(0,5) \times (0,5)$ over a time horizon[3] $T = 1.0$ with $\Delta t = 5 \cdot 10^{-3}$ (these and all following quantities are expressed in dimensionless values). We use $E_{\text{in}} = 5.0$, $F_{\text{in}} = 20.0$ and $F_{\text{out}} = 0$; the latter condition corresponds to a vacuum boundary, and ignores any black-body radiation emitted by the medium surrounding the brain near the exit point of the radiation. In the JfNK coupling, we employ a convergence tolerance of $\varepsilon = 10^{-3}$ for the Newton iteration and $\varepsilon_{\text{K}} = 10^{-3}$ for the GMRES algorithm. We compute the SC and MCS estimators of $\langle E_{\text{abs,tot}} \rangle$ and $\sigma^2_{E_{\text{abs,tot}}}$ using (4.26) without the prefactor of $1/3$; the latter can be easily accounted for when transforming the dimensionless results back into their dimensional counterparts. The computations were performed on an Intel Core i7 machine running at 4 GHz.

4.5.1 Case 1: Only c_1 is random

We first consider the case where $c_2 = L/2$ and $Z_2 = 6$ are fixed, and $c_1 \in [L/2 - c_{\text{H}}, L/2 + c_{\text{H}}]$ with $c_{\text{H}} = 0.1$, 0.3 or $w/2 = 0.5$. Figure 4.4 (left) reveals that SC outperforms MCS by at least one order of magnitude, and that $P_{\text{conv}} = 65$. Using 65 nodes, we investigate the dependence of $\hat{E}^{\text{SC}}_{\text{abs,tot}}$ and the coefficient of variation (CV) of $E_{\text{abs,tot}}$, $CV_{E_{\text{abs,tot}}}$, estimated by $\hat{\sigma}^{\text{SC}}_{E_{\text{abs,tot}}} / \hat{E}^{\text{SC}}_{\text{abs,tot}}$, on $CV_{c_1} \equiv \sigma_{c_1} / \langle c_1 \rangle = 0.4\, \sigma_{c_1}$. As Figure 4.5 illustrates, increasing CV_{c_1} reduces the mean value of the total energy deposition and increases the latter's coefficient of variation. The former result follows from the fact that a larger variance in the x_1-coordinate of the center of $\mathcal{D}_2$ allows this region to

[3] The results for numerical experiments using longer time horizons yielded similar results for the performance comparison between SC and MCS, and for the influence of the statistical moments of the uncertain parameters on those of the total energy deposition in the tumor.

overlap less with the tumor region $\mathcal{T}$, thereby making the dose-enhancing effect of the contrast agent less effective and increasing the dose to the surrounding healthy brain tissue. The energy absorption at the maximum CV_{c_1} of 0.12 is 27% lower than that for the deterministic case with full overlap between $\mathcal{D}_2$ and $\mathcal{T}$, demonstrating that even a small relative parametric uncertainty has a large impact on the mean value of the quantity of interest. Figure 4.5 also shows that, apart from its smallest nonzero value considered, the relative parametric uncertainty (CV_{c_1}) gets amplified (approximately doubled) by the nonlinearity of the system, and hence yields a larger relative predictive uncertainty ($CV_{E_{\text{abs,tot}}}$). Both results demonstrate that even a small uncertainty in only one parameter should be accounted for when predicting the behavior of highly nonlinear systems.

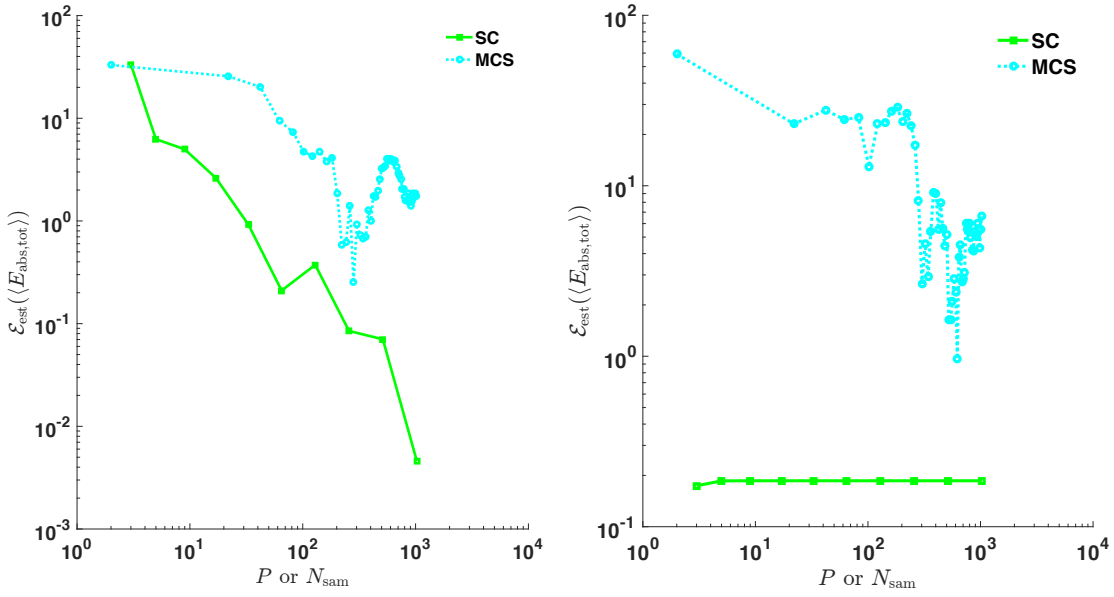

Figure 4.4: Estimation error of the mean energy deposition in the tumor for SC and MCS as a function of the number of collocation nodes/Monte Carlo samples for $M = 1$ with $c_{\text{H}} = 0.5$ (left) and $[Z_{\text{min}}, Z_{\text{max}}] = [2, 8]$ (right).

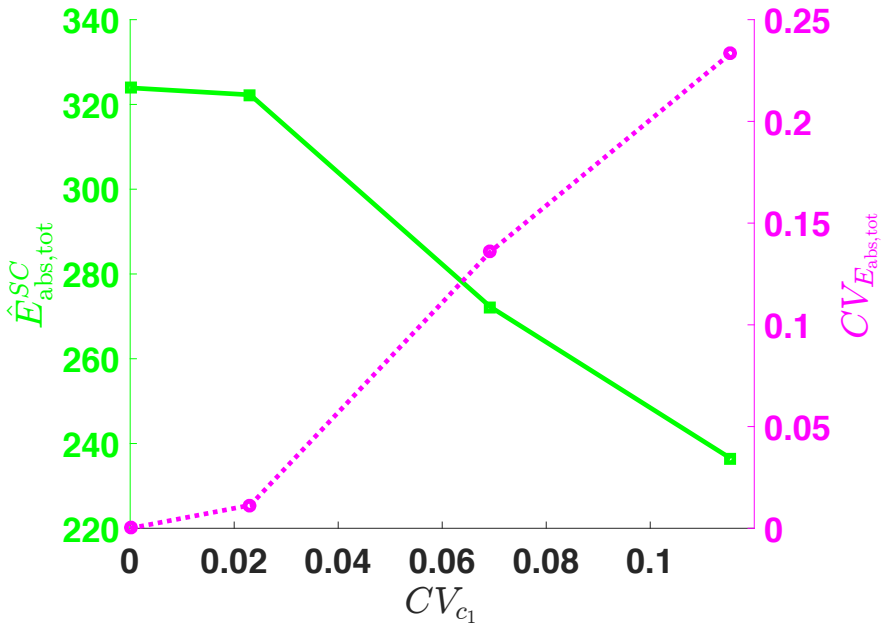

Figure 4.5: Dependence of the SC estimator for the mean and CV of the energy deposition in the tumor on the coefficient of variation of c_1 using $P_{\mathrm{conv}} = 65$ nodes for $M = 1$.

4.5.2 Case 2: Only Z_2 is random

Next, we look at a scenario where $c_1 = c_2 = L/2$ are fixed and $Z_2 \in [Z_{\min}, Z_{\max}]$, for which we consider either parameter ranges with increasing $\langle Z_2 \rangle$ and fixed $\sigma^2_{Z_2}$ ($[2,6]$, $[3,7]$ and $[4,8]$), or vice versa ($[4.6,5.4]$, $[4,6]$ and $[2,8]$). Figure 4.4 (right) shows that SC outperforms MCS by several orders of magnitude, and that $P_{\mathrm{conv}} = 9$ (to establish this value, we also looked at the estimation error for $\sigma^2_{E_{\mathrm{abs,tot}}}$). As Figure 4.6 illustrates, both increasing $\langle Z_2 \rangle$ for fixed $\sigma^2_{Z_2} = 1.33$ and increasing $\sigma^2_{Z_2}$ for fixed $\langle Z_2 \rangle = 5$ will enhance the mean energy absorption in $\mathcal{T}$. This is because higher values of Z_2 will be sampled, and the probability of photo-electric absorption of the X-rays increases with atomic number. For the case of fixed $\langle Z_2 \rangle = 5$ and increasing $\sigma^2_{Z_2}$, the maximum relative deviation between the estimated energy absorption and that for the case of $\mathcal{D}_2$ and $\mathcal{T}$ fully overlapping with $Z_2 = 5$ is 36%, even higher than the maximum relative deviation for

case 1. Figure 4.6 also reveals that increasing CV_{Z_2} for fixed $\langle Z_2 \rangle = 5$ increases $CV_{E_{\text{abs,tot}}}$, that each nonzero value of $CV_{E_{\text{abs,tot}}}$ is more than double that of the corresponding CV_{Z_2}, and that $CV_{E_{\text{abs,tot}}}$ exceeds 0.8 for the highest CV_{Z_2} considered.

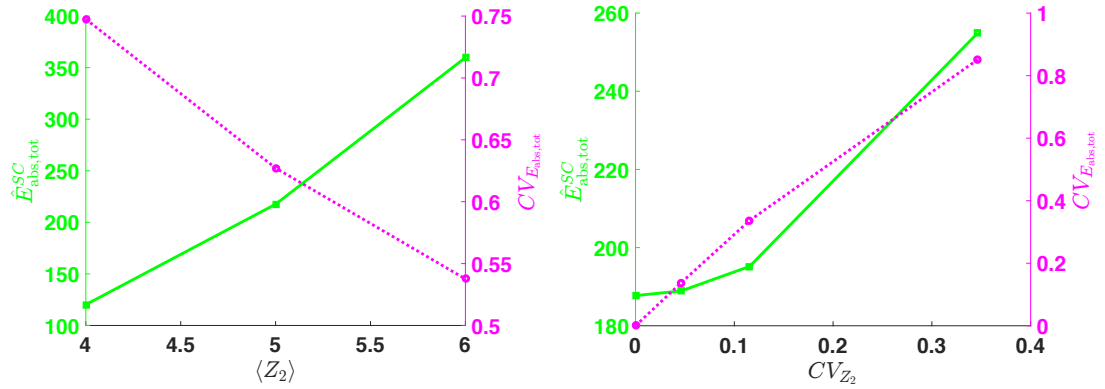

Figure 4.6: Dependence of the SC estimator for the mean and CV of the energy deposition in the tumor on the mean (left) and CV (right) of Z_2 using $P_{\text{conv}} = 9$ nodes for $M = 1$.

4.5.3 Case 3: Both c_1 and Z_2 are random

We now treat the case where $c_2 = L/2$ is fixed, $c_1 \in [L/2 - c_{\text{H}}, L/2 + c_{\text{H}}]$ with $c_{\text{H}} = 0.1$, 0.3 or 0.5, and $Z_2 \in [Z_{\min}, Z_{\max}]$ for which we again look at either $[2,6]$, $[3,7]$ and $[4,8]$, or $[4.6,5.4]$, $[4,6]$ and $[2,8]$. Figure 4.7 (left) indicates that SC outperforms MCS, and that $P_{\text{conv}} = 321$. As Figures 4.8 and 4.9 illustrate, the results for this case are mostly a combination of those obtained earlier for cases 1 and 2, except for the following:

1. For all but the smallest nonzero value of $\sigma^2_{Z_2}$ (0.05) considered, increasing $\sigma^2_{c_1}$ reduces the variance of $E_{\text{abs,tot}}$ when both c_1 and Z_2 are uncertain. This is in contrast to case 1, where increasing the uncertainty in c_1 increased the predictive uncertainty in $E_{\text{abs,tot}}$.

2. Figure 4.8 (right) shows that the energy absorption in the tumor deviates by, at

most, 35% from its value with full overlap between $\mathcal{D}_2$ and $\mathcal{T}$ and with $Z_2 = 5$. This is higher than the maximum relative deviation in case 1 but slightly lower than that in case 2 due to the counteracting effects of increasing $\sigma^2_{c_1}$ and increasing $\sigma^2_{Z_2}$ on the mean energy absorption.

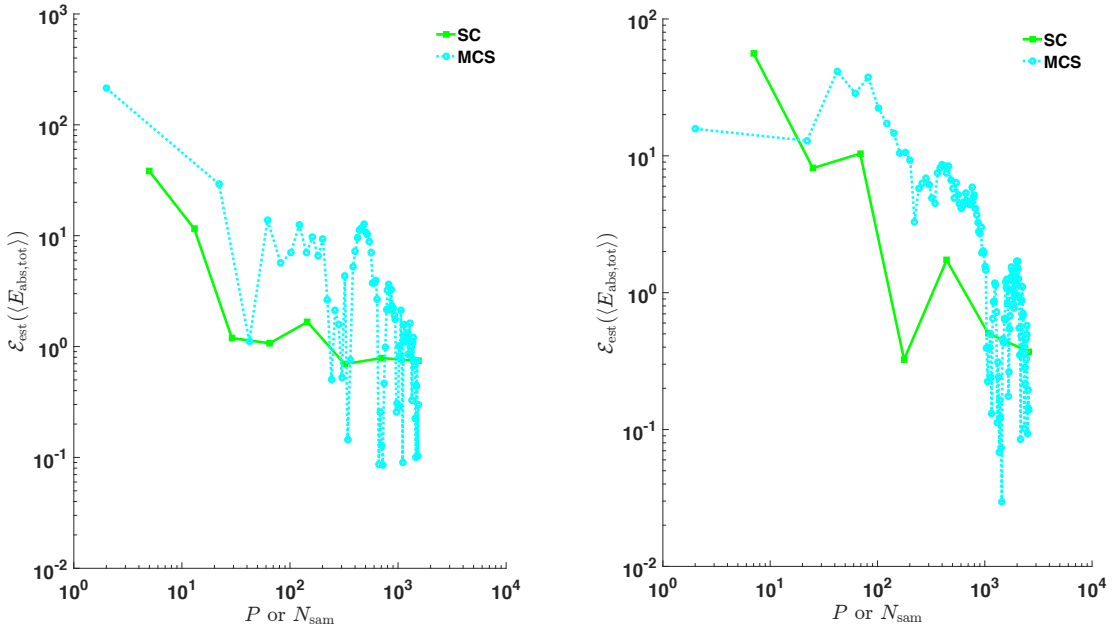

Figure 4.7: Estimation error of the mean energy deposition in the tumor for SC and MCS as a function of the number of collocation nodes/Monte Carlo samples for $M = 2$ with $c_{\text{H}} = 0.5$ and $[Z_{\text{min}}, Z_{\text{max}}] = [2, 8]$ (left), and $M = 3$ with $c_{\text{H}} = c_{\text{V}} = 0.5$ and $[Z_{\text{min}}, Z_{\text{max}}] = [2, 8]$ (right).

4.5.4 Case 4: All three parameters c_1, c_2 and Z_2 are random

Finally, we consider the scenario where $c_1 \in [L/2 - c_{\text{H}}, L/2 + c_{\text{H}}]$ with $c_{\text{H}} = 0.1$ or 0.5, $c_2 \in [L/2 - c_{\text{V}}, L/2 + c_{\text{V}}]$ with $c_{\text{V}} = 0.1$ or 0.5, and $Z_2 \in [Z_{\text{min}}, Z_{\text{max}}]$ for which we consider either $[2, 6]$ and $[4, 8]$, or $[4.6, 5.4]$ and $[2, 8]$. Figure 4.7 (right) reveals that SC again outperforms MCS, albeit only slightly, and $P_{\text{conv}} = 1073$. Although not shown here, the results for this case are mostly similar to those obtained for case 3, except for

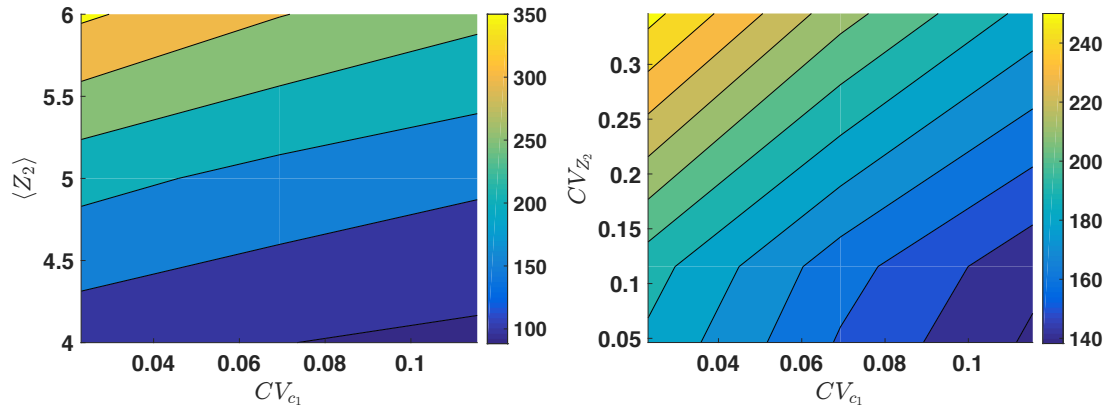

Figure 4.8: Dependence of the SC estimator for the mean energy deposition on the CV of c_1 and the mean (left) or CV (right) of Z_2 using $P_{\text{conv}} = 321$ nodes for $M = 2$.

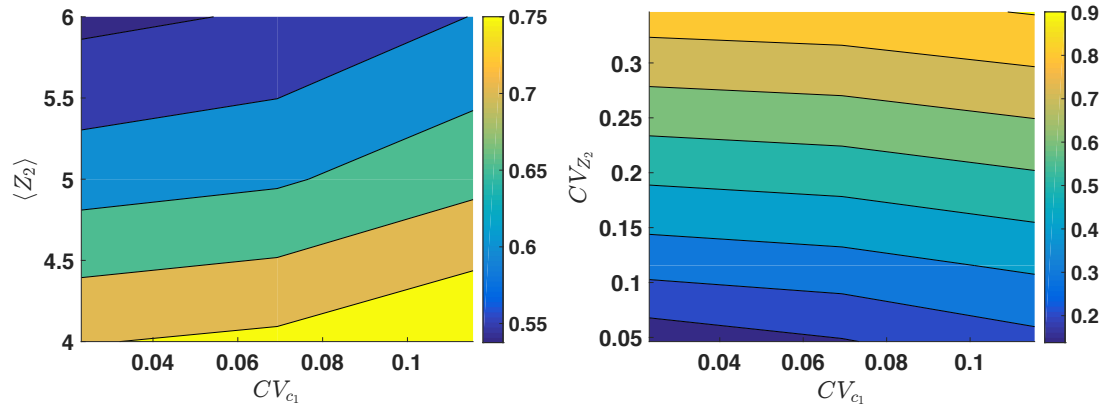

Figure 4.9: Dependence of the SC estimator for the CV of the energy deposition on the CV of c_1 and the mean (left) or CV (right) of Z_2 using $P_{\text{conv}} = 321$ nodes for $M = 2$.

the following:

1. The maximum value of $CV_{E_{\text{abs,tot}}}$ is now close to 1, i.e., the uncertainty in the prediction of the quantity of interest (as measured by the standard deviation) approximately equals the prediction (mean value) itself.

2. The maximum relative deviation of the energy absorption is now 46%, i.e., the difference in energy deposition is almost half the value of that at full overlap between $\mathcal{D}_2$ and $\mathcal{T}$ with $Z_2 = 5$. Similar to case 3, the effects of increasing $\sigma^2_{c_1}$ or $\sigma^2_{c_2}$ versus increasing $\sigma^2_{Z_2}$ will counteract each other.

4.5.5 Global sensitivity analysis

Sobol' [105] developed a global sensitivity analysis technique that decomposes the model output variance into summands of variances of the input parameters in increasing dimensionality. It computes the contribution of each input parameter and its interactions with other parameters to the overall model output variance. Based on the derivation in J, first-order sensitivity indices for our problem are defined as

$$S_{c_1} = \sigma^2_{E_{\text{abs,tot}},c_1}/\sigma^2_{E_{\text{abs,tot}}}, \quad S_{c_2} = \sigma^2_{E_{\text{abs,tot}},c_2}/\sigma^2_{E_{\text{abs,tot}}}, \quad S_{Z_2} = \sigma^2_{E_{\text{abs,tot}},Z_2}/\sigma^2_{E_{\text{abs,tot}}}, \quad (4.29)$$

where $\sigma^2_{E_{\text{abs,tot}},c_1}$ represents the contribution of c_1 to the total variance of $E_{\text{abs,tot}}$, and likewise for c_2 and Z_2. Total-order sensitivity indices are given by

$$S_{T,c_1} = \frac{1}{\sigma^2_{E_{\text{abs,tot}}}} \sum_{\alpha \in \mathcal{I}_{c_1}} \sigma^2_\alpha, \quad S_{T,c_2} = \frac{1}{\sigma^2_{E_{\text{abs,tot}}}} \sum_{\alpha \in \mathcal{I}_{c_2}} \sigma^2_\alpha, \quad S_{T,Z_2} = \frac{1}{\sigma^2_{E_{\text{abs,tot}}}} \sum_{\alpha \in \mathcal{I}_{Z_2}} \sigma^2_\alpha, \quad (4.30)$$

where $\mathcal{I}_{c_1}$ is the set of all subsets of $\{c_1, c_2, Z_2\}$ containing c_1, and likewise for $\mathcal{I}_{c_2}$ and $\mathcal{I}_{Z_2}$. To compute (4.29) and (4.30), either MCS [106], polynomial chaos [109] or stochastic collocation [26] may be used; we employ an MCS approach implemented via a

combination of an in-house numerical solver and the MATLAB code in [18]. Figure 4.10 shows that the first-order Sobol' indices for c_1 and c_2 are approximately equal, in line with the fact that a similar change in c_1 and c_2 causes a similar deviation of the contrast-agent region $\mathcal{D}_2$ from the tumor region $\mathcal{T}$; the first-order index for Z_2 is much larger, indicating that variations in this parameter will have a much larger effect on the predictive uncertainty of the total energy deposition in $\mathcal{T}$ than changes in c_1 or c_2. This may be understood from the nonlinear (cubic) dependence of the energy deposition on the effective atomic number, as illustrated by (4.25). Figure 4.10 also reveals a clear difference between the total-order Sobol' indices for c_1, c_2 and Z_2 and their first-order counterparts, indicating that there is a measurable impact of the interactions between the different parameters on the total variance in $E_{\text{abs,tot}}$.

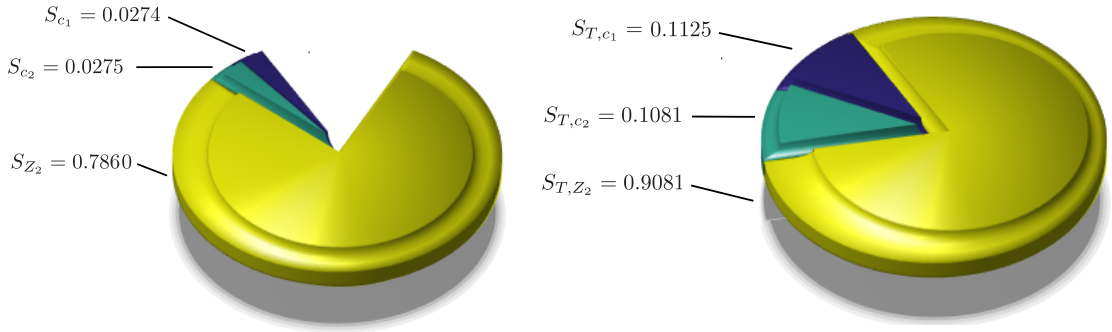

Figure 4.10: First-order (left) and total-order (right) Sobol' indices for the case of $M = 3$ with $c_{\text{H}} = c_{\text{V}} = 0.5$ and $[Z_{\text{min}}, Z_{\text{max}}] = [2, 8]$.

4.6 Summary and Conclusions

We estimated the energy deposition into a brain tumor irradiated by X-rays in the presence of parametric uncertainty using the stochastic collocation (SC) approach. We represented the uncertain input parameters, namely the coordinates of the center of the region containing an iodinated, dose-enhancing contrast agent, and the effective value of

the atomic number in this area, as mutually independent, uniformly distributed random variables. We investigated the effect of changes in their mean and/or variance on the statistical moments of the deposited energy, and compared the computational efficiency of SC to that of standard Monte Carlo simulation (MCS).

Our analysis leads to the following major conclusions.

1. SC outperforms MCS for all stochastic dimensions considered, with the biggest difference for the scenario of one random input parameter.

2. In the majority of cases, the coefficient of variation of the uncertain parameters was amplified by the nonlinearity of the problem, yielding a larger coefficient of variation for the energy deposition. Hence, even tiny parametric uncertainties may result in large predictive uncertainty in the quantity of interest.

3. As the stochastic dimension increases, the magnitude of the predictive uncertainty in the energy deposition, as measured by its standard deviation, approaches that of the prediction (mean energy deposition) itself.

4. In the presence of additional uncertain parameters, the effect of uncertainty in a specific parameter on the predictive uncertainty in the quantity of interest may differ from its effect when this parameter is the only uncertain input.

5. Global sensitivity analysis via the Sobol' method reveals that predictive uncertainty in the energy deposition is mainly influenced by variations in the effective atomic number, and is also affected by interactions between the different uncertain parameters.

The flux-limited radiation-diffusion approximation employed to model the propagation of X-rays within the brain has also been applied to a wide range of other problems including core collapse supernovae [84] and inertial confinement fusion [37]. Hence, our

findings are relevant to the quantification of predictive uncertainty across a number of research areas. Future extensions of the presented analysis may include considering a three-dimensional model with a larger number of uncertain parameters, and representing the latter by random variables with more complex probability distributions.

4.7 Acknowledgements

The dissertation author thanks Jim Morel, Michael Holst and Maxime Theillard for helpful discussions. The material contained in this Chapter is based on research supported in part by Defense Advanced Research Projects Agency under the EQUiPS program. It is, in part, a reprint of the material as it appears in the paper "Impact of parametric uncertainty on estimation of the energy deposition into an irradiated brain tumor" by S. Taverniers and D.M. Tartakovsky, under review with *Journal of Computational Physics*, Elsevier Inc. [117]. The dissertation author is the primary investigator and author of this paper.

Chapter 5

Noise propagation in hybrid models of nonlinear systems: the Ginzburg-Landau equation

5.1 Abstract

Every physical phenomenon can be described by multiple models with varying degrees of fidelity. The computational cost of higher fidelity models (e.g., molecular dynamics simulations) are invariably higher than that of their lower fidelity counterparts (e.g., a continuum model based on differential equations). While the former might not be suitable for large-scale simulations, the latter are not universally valid. Hybrid algorithms provide a compromise between the computational efficiency of a coarse-scale model and the representational accuracy of a fine-scale description. This is achieved by conducting a fine-scale computation in subdomains where it is absolutely required (e.g., due to a local breakdown of a continuum model) and coupling it with a coarse-scale computation in the rest of a computational domain. We analyze the effects of random fluctuations

generated by the fine-scale component of a nonlinear hybrid on the hybrid's overall accuracy and stability. Two variants of the time-dependent Ginzburg-Landau equation (GLE) and their discrete representations provided by a nearest-neighbor Ising model serve as a computational testbed. Our analysis shows that coupling these descriptions in a one-dimensional simulation leads to erroneous results. Adding a random source term to the GLE provides accurate prediction of the mean behavior of the quantity of interest (magnetization). It also allows the two GLE variants to correctly capture the strength of the microscale fluctuations. Our work demonstrates the importance of fine-scale noise in hybrid simulations, and suggests the need for replacing an otherwise deterministic coarse-scale component of the hybrid with its stochastic counterpart.

5.2 Introduction

Numerical modeling of complex nonlinear systems requires the development of multi-algorithm computational solvers capable of handling a wide range of spatial and/or temporal scales. While coarse-scale models are more computationally efficient than their fine-scale counterparts, they are not universally valid. For instance, continuum-scale finite element models of crack propagation break down near a crack's tip [2], and macroscopic (Darcy-scale) models of flow and transport in porous media break down for localized phenomena with high pore-scale gradients [14, 13, 23]. Standard coarse-scale models also fail to capture the effects of spontaneous microscale fluctuations on macroscopic behavior, such as spontaneous formation of ordered spatial concentration patterns in an unstirred chemical medium [56].

Fine-scale algorithms (e.g., molecular dynamics/quantum tight-binding and pore-scale simulations in the first and second examples, respectively) can model such processes, but their high computational cost renders them impractical for modeling large-scale

problems. Hybrid algorithms, which are also referred to as algorithm refinement, employ such fine-scale models only in subdomains wherein their coarse-scale counterparts break down, potentially yielding a significant reduction in computational cost [86]. With a few exceptions [110, 101], coupling of the fine- and coarse-scale components of a hybrid requires multiple iterations to ensure the continuity of state variables and their fluxes at the interface between the two components. Design and computationally efficient implementation of such coupling procedures remains a key challenge in hybrid modeling.

Stochastic fluctuations generated by a hybrid's fine-scale (particle-based) component exacerbate this task [7]. Averaging out this noise (i.e., coupling averaged quantities such as particle density and mass flux to their counterparts computed with a coarse-scale deterministic component) is adequate for computing the mean behavior of linear systems, but artificially reduces the fluctuation variance in the particle region near the particle-continuum interface [5]. In weakly nonlinear systems, such as the train model of viscous transport in gases, the averaging dampens the long-range correlations of velocity fluctuations and "can lead to a greatly altered time-dependent behavior" [6]. A nonlinear hybrid model consisting of asymmetric excluded random walk (the fine-scale component) and a viscous Burgers' equation (the coarse-scale component) revealed that the averaging tends to suppress the drift of shock location [16]. In each case, addition of a Gaussian white noise term to the hybrid's deterministic (coarse-scale) component corrected these shortcomings.

A proper treatment of noise is even more important in highly nonlinear systems, wherein even small changes in the magnitude of microscopic fluctuations can significantly affect the macroscopic dynamics. In such systems, coupling the averages of fine-scale quantities with their deterministic coarse-scale counterparts can lead to erroneous predictions of the mean system behavior. The Ginzburg-Landau theory of phase transitions [71] provides an ideal setting to study noise propagation in hybrid models,

since it establishes a rigorous relationship between fine-scale (a nearest-neighbor Ising model with spin-flip dynamics) and coarse-scale (a Ginzburg-Landau partial differential equation) representations of a highly nonlinear system. It can be used, for example, to describe the evolution of (scalar) magnetization of a uniaxial ferromagnet to thermal equilibrium [108, 58, 24].

In Section 5.3 we formulate a nearest-neighbor Ising (NNI) model and two variants of the time-dependent Ginzburg-Landau equations (GLEs). A hybrid algorithm coupling these two levels of description is discussed in Section 5.4. Simulation results reported in Section 5.5 reveal that one has to add a random source term to the Ginzburg-Landau component of the hybrid in order to correctly predict the mean and variance of the magnetization for a ferromagnet evolving to thermodynamic equilibrium. This finding facilitates the analysis of noise propagation in the NNI-GLE hybrid by allowing one to replace its NNI component with a stochastic Ginzburg-Landau equation (sGLE). A solution of the latter is presented in terms of moment equations (deterministic equations describing the evolution of the mean and covariance of magnetization). The main conclusions of our analysis are summarized in Section 5.6.

5.3 Two Modeling Scales in the Ginzburg-Landau Theory

Phase transitions in ferromagnetic systems can be described either microscopically with Ising models [60] or macroscopically with the Ginzburg-Landau theory [58, 24, 118]. Both levels of description are formulated below in the context of the magnetization of a one-dimensional (1D) ferromagnet.

5.3.1 Nearest-neighbor Ising models with spin-flip dynamics (NNIs)

Consider a ferromagnet whose atoms are arranged on a 1D lattice with sites $i = 1, \ldots, N$. A microscopic representation of this system is given by an Ising model with nearest-neighbor interactions [60]. It assumes that the spin s_i of the atom at site i can be in one of the two states designated by $s_i = \pm 1$, and interacts only with its two adjacent spins. The N-spin configuration $\mathbf{s} = \{s_1, \ldots, s_N\}$ defines the ferromagnet's state at time t; the joint probability of finding the ferromagnet in state $\mathbf{s}$ at time t is denoted by $P(\mathbf{s};t)$. Let $\mathbf{s}'$ denote an N-spin configuration that differs from configuration $\mathbf{s}$ by the value of a single spin s_j. The kinetic nearest-neighbor Ising model with spin-flip dynamics [52] (NNI) defines the evolution of $P(\mathbf{s};t)$ as a solution of the master equation

$$\frac{\mathrm{d}P(\mathbf{s};t)}{\mathrm{d}t} = \sum_{\mathbf{s}'} [w(\mathbf{s}' \to \mathbf{s})P(\mathbf{s}';t) - w(\mathbf{s} \to \mathbf{s}')P(\mathbf{s};t)] \tag{5.1}$$

where $w(\mathbf{s} \to \mathbf{s}')$ is the transition rate from state $\mathbf{s}$ to state $\mathbf{s}'$, and the summation is over all possible transitions $\mathbf{s}'$. Among the plethora of suggested functional forms for the transition rate w we consider two. The first is the Suzuki-Kubo rate

$$w_{\mathrm{SK}}(\mathbf{s} \to \mathbf{s}') = \frac{\lambda}{2}\left[1 - s_j \tanh\left(\beta J \sum_{L_j} s_{L_j}\right)\right], \tag{5.2a}$$

where λ^{-1} is the time scale of the spin-flip process that can depend both on the system temperature T and the spins other than s_j, $\beta = 1/(k_B T)$ with k_B denoting the Boltzmann constant, J is the spatially uniform exchange coupling energy associated with the interaction between neighboring spins, and L_j indicates summation over the nearest neighbors of s_j.

The second alternative is the heat-bath rate

$$w_{\mathrm{hb}}(\mathbf{s}\to\mathbf{s}') = \kappa\left[1+\mathrm{e}^{\beta\mathcal{H}(\mathbf{s}')-\beta\mathcal{H}(\mathbf{s})}\right]^{-1}, \tag{5.2b}$$

where κ^{-1} sets the time scale of the spin-flip process and can depend both on T and the spins other than s_j, and $\mathcal{H}(\mathbf{s})$ is the Hamiltonian of configuration $\mathbf{s}(t)$. The latter is given by [27]

$$\mathcal{H}(\mathbf{s}) = -\sum_{\langle i,j\rangle} J_{ij}s_is_j - \mu H\sum_{i=1}^{N} s_i, \tag{5.3}$$

where J_{ij} is the exchange coupling energy associated with the interaction between neighboring spins; μ is the spin magnetic moment; H is the external magnetic field; and $\langle i,j\rangle$ indicates the summation over pairs of adjacent spins, with each pair counted only once. In our simulations, we set $H=0$ and take $J_{ij}\equiv J>0$. The latter implies "ferromagnetic behavior", which favors (in thermal equilibrium) neighboring parallel spins over neighboring anti-parallel spins.

We refer to the master equation (5.1) with transition rate w_{SK} in (5.2a) as NNIa, and to (5.1) with transition rate w_{hb} in (5.2b) as NNIb. Both versions of NNI are implemented via a Monte Carlo (MC) algorithm, in which one step in the Ising model consists of repeating N times the following procedure:

1. Pick a random site j where $j\in\{1,\ldots,N\}$;

2. Draw a number $s^\star$ from a uniform distribution on $[0,1]$;

3. Flip the spin s_j if $s^\star < w(\mathbf{s}\to\mathbf{s}')$, otherwise leave it in its original state.

The spin lattice is initialized as follows. Each site i ($i=1,\ldots,N$) is assigned $P_+(i,t=0)$, the singlet probability of finding the spin in the "up" state. Then a number $\sigma^\star$ is drawn

from a uniform distribution on $[0,1]$. If $\sigma^\star < P_+(i,0)$ then $s_i = 1$, otherwise $s_i = -1$. We assume that the ferromagnet is initially in a state of uniform magnetization m_{in}, so that $P_+(i,0) = (1+m_{\text{in}})/2$ for all i [96]. Magnetization is a macroscopic quantity whose dynamics is governed by the Ginzburg-Landau equations.

5.3.2 Ginzburg-Landau equations (GLEs)

Magnetization (or "order parameter") is defined in terms of ensemble-averaged Ising spin states as [96]

$$m(x,t) = \frac{1}{N_b} \sum_{i \in \Lambda_b(x)} \bar{s}_i(t), \tag{5.4}$$

where $\Lambda_b(x)$ is a set of N_b spins inside the interval Δx (a continuum grid's spacing) centered around point x. Magnetization $m(x,t)$ at point x is related to the probability $P_+(i,t)$ or $P_-(i,t)$ of finding the corresponding i-th spin in the "up" or "down" state by [96]

$$P_\pm(i,t) = \frac{1 \pm m(x,t)}{2}. \tag{5.5}$$

Ginzburg-Landau equations (GLEs) provide a mean-field approximation of the NNI models. Specifically, the NNIa model gives rise to the GLE [96]

$$\frac{\partial m}{\partial t} = \Gamma_a \left(c^2 J \frac{\partial^2 m}{\partial x^2} - A\, m - B\, m^3 \right) \tag{5.6a}$$

where c is the spin lattice constant,

$$\Gamma_a = \frac{\lambda}{k_B T}, \qquad A = qJ\left(\frac{T}{T_c} - 1\right), \qquad B = \frac{qJ}{3}\left(\frac{T_c}{T}\right)^2, \tag{5.6b}$$

q is the number of nearest neighbors of any spin (for the 1D lattice, $q = 2$), and $T_c = qJ/k_B$ is the mean-field critical temperature; while the NNIb model corresponds to the GLE [73]

$$\frac{\partial m}{\partial t} = \Gamma_b \left(c^2 \frac{\partial^2 m}{\partial x^2} - r\, m \right) \tag{5.7a}$$

where

$$\Gamma_b = \kappa \left(\frac{1}{2} - W_4 \right), \qquad r = \frac{2\kappa W_4}{\Gamma_b}, \qquad W_4 = \left(1 + \mathrm{e}^{4\beta J} \right)^{-1}. \tag{5.7b}$$

We refer to the GLE (5.6) as GLEa and to the GLE (5.7) as GLEb.

When the assumptions underpinning the GLEa and GLEb models are violated, these coarse-scale models break down. For example, fine-scale fluctuations in the one-dimensional NNIa model preclude a phase transition above $T = 0$ [71]. This is at odds with the predictions of the GLEa model, which include the occurrence of a phase transition at a finite critical temperature $T_c = qJ/k_B$ from an unordered state at $T \geq T_c$ to a state with long-range order at $T < T_c$. One way of dealing with the effects of fine-scale fluctuations is to introduce a random noise term into a coarse-scale model [58]. Introducing a random source $\eta_a(x,t)$ into the GLE (5.6) yields a stochastic Ginzburg-Landau equation (sGLE)

$$\frac{\partial m}{\partial t} = \Gamma_a \left(c^2 J \frac{\partial^2 m}{\partial x^2} - A\, m - B\, m^3 \right) + \eta_a(x,t). \tag{5.8a}$$

Here $\eta_a(x,t)$ is a zero-mean Gaussian space-time white noise with an auto-covariance function

$$C^a_\eta(x,t;y,\tau) \equiv \overline{\eta_a(x,t)\eta_a(y,\tau)} = 2c\Gamma_a k_B T \delta(x-y)\delta(t-\tau). \tag{5.8b}$$

The GLEb model yields a minimum of the free energy at $T = 0$ and hence predicts a phase transition at the correct temperature. Yet it fails to capture the magnitude of the microscopic fluctuations. This failure is remedied by introducing a random source $\eta_b(x,t)$ into the GLE (5.7), which gives rise to a sGLE

$$\frac{\partial m}{\partial t} = \Gamma_b\left(c^2\frac{\partial^2 m}{\partial x^2} - r\,m\right) + \eta_b(x,t). \tag{5.9a}$$

Here $\eta_b(x,t)$ is a zero-mean Gaussian space-time white noise with an auto-covariance function

$$C^b_\eta(x,t;y,\tau) = 2D\delta(x-y)\delta(t-\tau), \qquad D = c\kappa\left[\frac{1}{2} + 2W_4(1-W_4)\right]. \tag{5.9b}$$

We refer to the sGLE (5.8) as sGLEa and to the sGLE (5.9) as sGLEb.

Numerical solutions of the stochastic GLEs (5.8) and (5.9) use a second-order central finite difference method in space, and a first-order forward Euler-Maruyama stochastic integrator [69] in time. This discretizes (5.8) into

$$\begin{aligned} m_j^{n+1} &= m_j^n + \Delta t'_a R\left[\frac{m^n_{j+1} + m^n_{j-1} - 2m^n_j}{q(\Delta x')^2} - \frac{1-R}{R}m^n_j - \frac{R^2}{3}(m^n_j)^3\right] \\ &\quad + \sqrt{\frac{2\Delta t'_a}{\Delta x'}}\xi^n_j, \qquad R \equiv \frac{T_c}{T}, \qquad \Delta x' \equiv \frac{\Delta x}{c}, \qquad \Delta t'_a \equiv \lambda\Delta t^a_c \end{aligned} \tag{5.10}$$

and (5.9) into

$$\begin{aligned} m_j^{n+1} &= m_j^n + \Gamma'_b\Delta t'_b\left[\frac{m^n_{j+1} + m^n_{j-1} - 2m^n_j}{(\Delta x')^2} - r\,m^n_j\right] + \sqrt{\frac{2D'\Delta t'_b}{\Delta x'}}\xi^n_j, \\ \Delta x' &\equiv \frac{\Delta x}{c}, \qquad \Delta t'_b \equiv \kappa\Delta t^b_c, \qquad D' \equiv D/(c\kappa). \end{aligned} \tag{5.11}$$

Here $m^n_j \equiv m(x_j,t_n)$; $x_j = j\Delta x$; $t_n = n\Delta t^{a,b}_c$; $\Delta t^{a,b}_c$ are the time steps for the GLEa and GLEb models, respectively; and ξ^n_j are independent, identically distributed standard nor-

mal random variables for each space-time point (x_j, t_n). The discretized equations (5.10) and (5.11) with the noise terms set to 0 are used to solve the deterministic GLEs (5.6) and (5.7), respectively.

5.4 Hybrid Modeling of Ginzburg-Landau Systems

Hybrid models simulate a system's dynamics with a computationally less expensive (continuum) method over the entire computational domain except in subdomains where a more accurate and expensive (discrete, e.g., atomistic) method is deployed. Our implementation of the following four hybrids is discussed below.

5.4.1 Computational domain

Consider a one-dimensional domain $[0, L)$ on which the continuum GLE or sGLE is solved everywhere except for a patch $[L/4, 3L/4)$ where the discrete NNI model is used (see Fig. 5.1). A lattice of $N_0 = L/(2c)$ spins with uniform spacing c spans the discrete subdomain $[L/4, 3L/4)$. The rest of the computational domain, $[0, L/4)$ and $[3L/4, L)$, is discretized by a mesh of size Δx. The latter is a multiple of the lattice spacing c, such that $\Delta x = N_b c$. Periodic boundary conditions are assumed for the magnetization, i.e., $m(x = 0, t) = m(x = L, t)$.

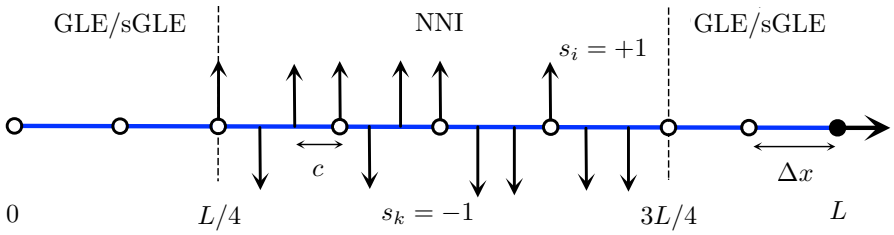

Figure 5.1: Computational domain for the NNI-GLE/sGLE hybrid.

5.4.2 Time advancement and algorithm coupling

Time integration of the hybrid is a two-step process. First, the NNI algorithm is advanced by N_0 MC steps using the algorithm described in Section 5.3.1. Second, either the sGLEa or sGLEb (or their deterministic counterparts) is advanced over one continuum time step Δt_c^a or Δt_c^b in accordance with (5.10) or (5.11), respectively. These are related to the time step in the MC algorithm Δt_d by $\Delta t_c^{a,b} = N_0 \Delta t_d$.

The NNI and GLE components of the hybrid are coupled at the interfaces $x = L/4$ and $x = 3L/4$ using the following procedure.

1. Define a left neighbor of the spin at $x = L/4$ and a right neighbor of the spin at $x = 3L/4 - c$ by extending the atomistic region into the GLE subdomain by one lattice point at $x = L/4 - c$ and $x = 3L/4$, respectively. These "virtual" spins are assigned "up" probability $P_+ = (1+m)/2$, where $m = m(L/4 - \Delta x, t)$ and $m = m(3L/4, t)$ for the left and right spins, respectively.

2. Use these virtual spins in either (5.2a) or (5.2b) to decide whether the spins at $x = L/4$ and $x = 3L/4 - c$ are to be flipped. This step follows the procedure described in Section 5.3.1.

3. Define the magnetization $m(x,t)$ at $x = L/4$ and $x = 3L/4 - \Delta x$ in order to calculate the discretized Laplacian in either (5.10) or (5.11) for $x_j = L/4 - \Delta x$ and $x_j = 3L/4$, respectively. This is done by computing the average of the spins at $x = L/4, \ldots, L/4 + \Delta x - c$ and the average of the spins at $x = 3L/4 - \Delta x, \ldots, 3L/4 - c$, respectively.

The use of a "handshake" region in step 1 is conceptually similar to that in [5], where the particle region was extended by one cell into the continuum region.

5.5 Simulation Results and Discussion

To analyze the accuracy of our hybrids and to investigate the importance of including random fluctuations into the GLEs, we compute the steady-state mean and variance of the magnetization of a one-dimensional ferromagnet evolving to thermodynamic equilibrium at temperature T. The mean $\overline{m}(x_i)$ and variance $\sigma_m^2(x_i)$ are estimated from M independent samples according to

$$\overline{m}(x_i) = \frac{1}{M}\sum_{k=1}^{M}\left(\frac{1}{N_s}\sum_{n=N_r+1}^{N_t} m_{i,k}^n\right) \tag{5.12}$$

$$\sigma_m^2(x_i) = \frac{1}{M}\sum_{k=1}^{M}\left[\frac{1}{N_s}\sum_{n=N_r+1}^{N_t} (m_{i,k}^n)^2 - \left(\frac{1}{N_s}\sum_{n=N_r+1}^{N_t} m_{i,k}^n\right)^2\right]. \tag{5.13}$$

Here N_t is the total number of continuum time steps $\Delta t_c^{a,b}$ taken, N_r is the number of steps $\Delta t_c^{a,b}$ taken before the sampling begins, and $N_s = N_t - N_r$ is the number of steps $\Delta t_c^{a,b}$ at which m_i is sampled. The value of N_r is chosen such that $N_r \Delta t_c^{a,b}$ exceeds the "relaxation time" $\tau_r^{a,b}$, which is a characteristic time needed for the system to reach equilibrium. This ensures that the magnetization is sampled at steady-state. (While the hybrid simulations evolve the system to its steady-state, the true thermal equilibrium is not attained since the detailed balance condition is not satisfied exactly by the mesoscale component.) Following [73], we obtain the relaxation times $\tau_r^a = [\lambda(R-1)]^{-1}$ and $\tau_r^b = [\kappa(1-\tanh(2R/q))]^{-1}$ for the sGLEa and sGLEb, respectively. Inside the NNI region, the magnetization m is computed by using the coarse-graining procedure (5.4) with $\bar{s}_i = s_i$, i.e., we do not ensemble average the spins at each Ising step. This preserves the fluctuations inherent to particle-based simulations [5, 6], in our case the Ising model.

The parameter values used in the simulations reported below are summarized in Table 5.5. This choice of the parameter values results in $R = 2/1.8$, $\Delta t_a' = 0.1$, $\Delta t_b' = 0.1$ and $\Delta x' = 1$. These values guarantee the stability of the linear diffusion component by

satisfying the criteria

$$\frac{R\Delta t'_a}{q(\Delta x')^2} \leq \frac{1}{2} \qquad \text{and} \qquad \frac{\Gamma'_b \Delta t'_b}{(\Delta x')^2} \leq \frac{1}{2} \tag{5.14}$$

for the the sGLEa and sGLEb, respectively. The choice of $\Delta x' = 1$ in (5.4) implies the lack of coarse-graining. The coarse-graining can be achieved by taking Δx to be an integral multiple of the spin lattice spacing c.

The results computed with the NNI models (the fine-scale descriptions) in the whole computational domain are treated as the "ground truth". The GLE and sGLE models, as well as the NNI-GLE and NNI-sGLE hybrids, provide their approximations.

Table 5.1: Parameter values. The units of λ and κ are s^{-1}. The units of $\Delta t_c^{a,b}$ are s.

Parameter	N	T	q	λ	κ	m_{in}	N_r	N_t	M	Δx	$\Delta t_c^{a,b}$
Value	40	$0.9T_c$	2	1.0	1.0	0.5	10^5	10^6	10	c	0.1

5.5.1 NNIa-GLEa and NNIa-sGLEa hybrids

The steady-state mean and variance of the magnetization computed with the NNIa model and the NNIa-GLEa hybrid are shown in Figure 5.2. The GLEa produces $\overline{m} = 0.4930$ away from the NNIa patch, which is in agreement with (K.3) in K. This agreement serves to verify our computation. Yet, the value $\overline{m} = 0.4930$ does not represent the correct system behavior, $\overline{m}(x) \equiv 0$, predicted with the NNIa (see the relevant discussion in Section 5.3.2). The NNIa-GLEa hybrid also fails to capture the magnetization variance $\sigma_m^2(x)$.

The NNIa-sGLEa hybrid is obtained by replacing the GLEa model with its stochastic counterpart, the sGLEa. The steady-state mean and variance of the magnetization computed with the NNIa-sGLEa hybrid are exhibited in Figure 5.2 alongside their

NNIa-GLEa counterparts. The NNIa-sGLEa hybrid correctly predicts the mean magnetization throughout the whole computational domain ($\overline{m} = 0$). Combined with similar behavior observed in previous studies of other (linear and nonlinear) systems [5, 6, 16], this points to a general feature of hybrid modeling: the noise generated in the fine-scale (particle-based) component of a hybrid is best handled by introducing a random source term into the hybrid's continuum (and otherwise deterministic) component. Figure 5.2 also reveals that the NNIa-sGLEa hybrid qualitatively captures the Ising fluctuations, but underestimates their strength. The exact match is obtained by multiplying the noise term η_a by $\zeta_a = 1.285$.

5.5.2 NNIb-GLEb and NNIb-sGLEb hybrids

The steady-state mean and variance of the magnetization computed with the NNIb model and the NNIb-GLEb hybrid are shown in Figure 5.3. In contrast to the NNIa-GLEa hybrid, the NNIb-GLEb hybrid yields $\overline{m} = 0$ throughout the computational domain, in agreement with the NNIb result. Similar to the NNIa-GLEa hybrid, the NNIb-GLEb hybrid fails to capture the higher moments (e.g., variance) of the magnetization.

The NNIb-sGLEb hybrid is obtained by replacing the GLEb model with its stochastic counterpart, sGLEb. The steady-state mean and variance of the magnetization computed with the NNIb-sGLEb hybrid are plotted in Figure 5.3. The NNIb-sGLEb predicts the correct mean magnetization and qualitatively reproduces the magnetization variance, overestimating its magnitude by a factor of $\zeta_b = 0.92$.

5.5.3 Computational efficiency of NNI-sGLE hybrids

Hybrid methods' *raison d'être* is their ability to improve the computational efficiency of a fine-scale algorithm without sacrificing its accuracy. We quantify the

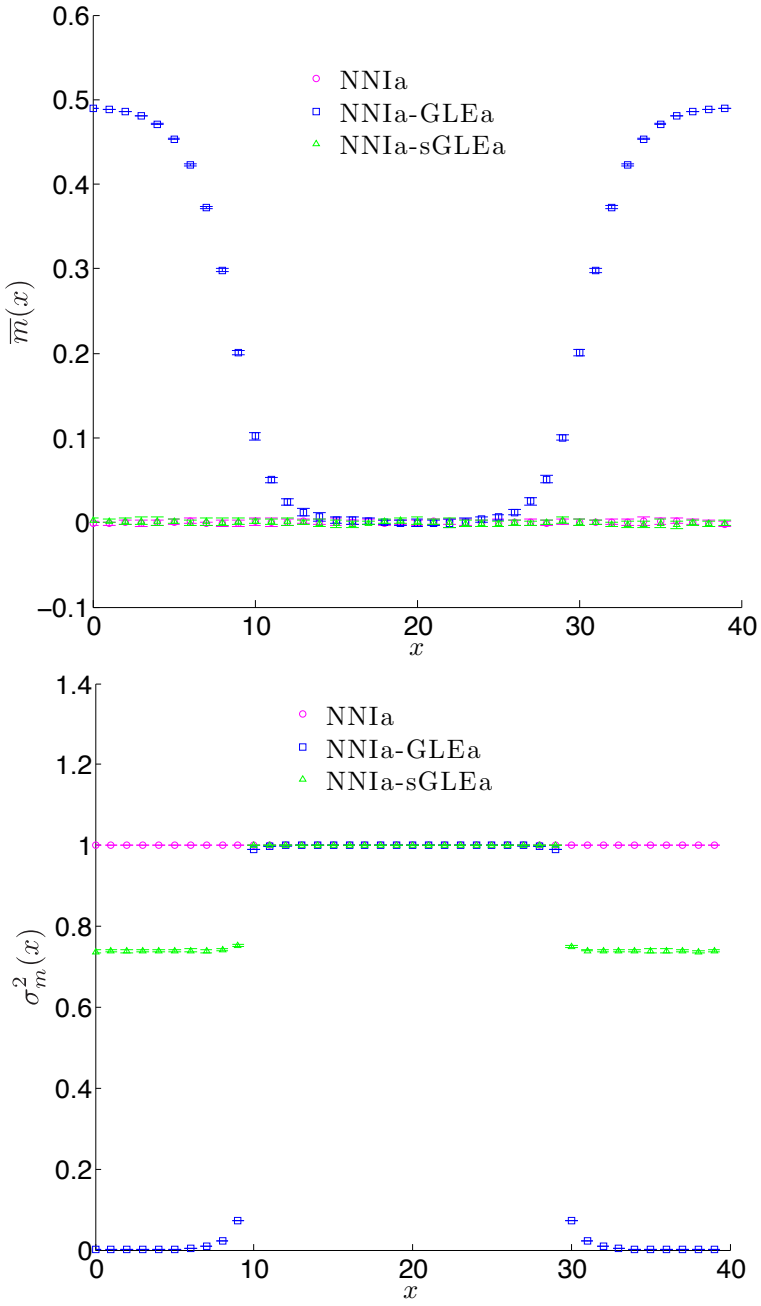

Figure 5.2: Steady-state mean (top) and variance (bottom) computed with the NNIa model and the NNIa-GLEa and NNIa-sGLEa hybrids. The NNIa solution is treated as the "ground truth".

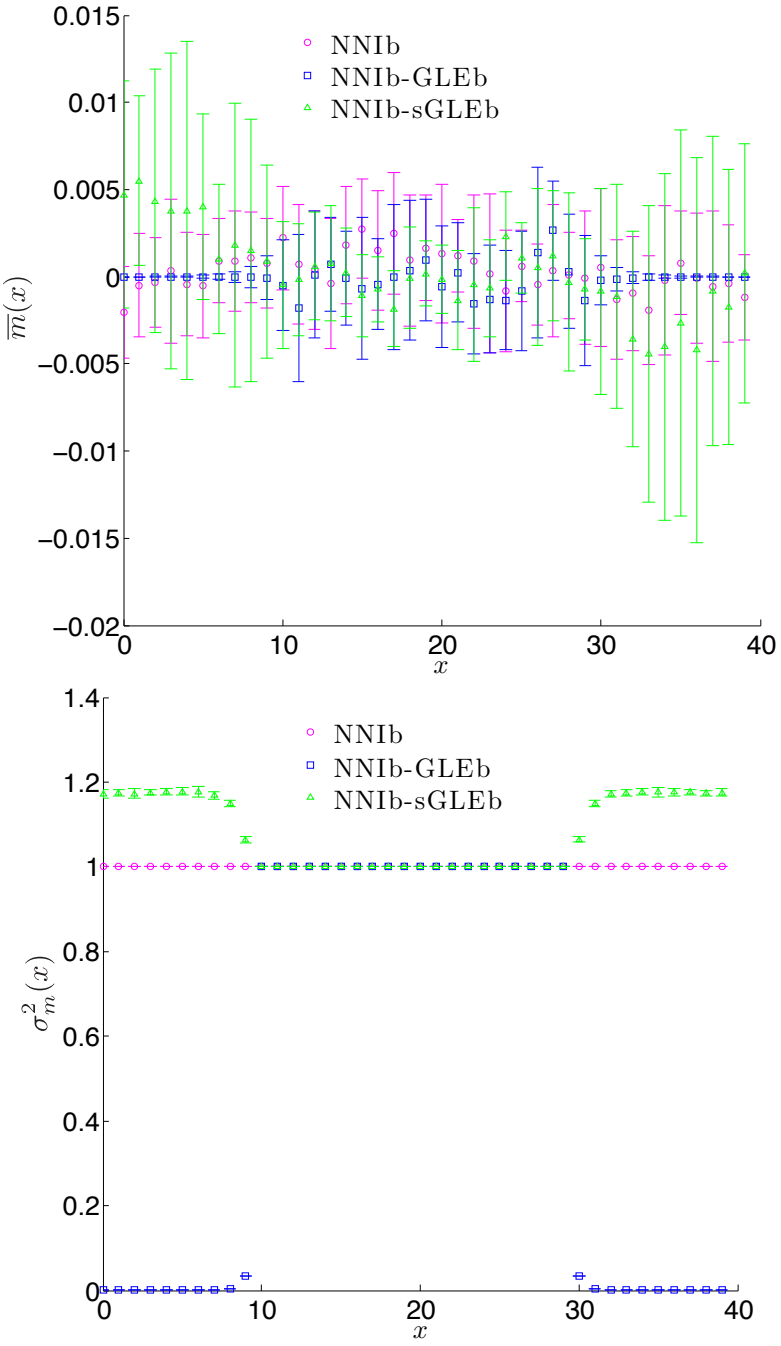

Figure 5.3: Steady-state mean (top) and variance (bottom) computed with the NNIb model and the NNIb-GLEb and NNIb-sGLEb hybrids. The NNIb solution is treated as the "ground truth".

Table 5.2: Computational gain factor $G_N = t_{\text{sim,NNI}}/t_{\text{sim,hyb}}$ with $N =$ number of spins.

Hybrid model	G_{40}	G_{160}	G_{640}
NNIa-sGLEa	1.33	2.16	3.35
NNIb-sGLEb	1.17	1.86	3.50

hybrid's gain in terms of the ratio $G = t_{\text{sim,NNI}}/t_{\text{sim,hyb}}$ of the average computation time necessary to advance the full-domain NNI algorithm by one Monte Carlo step per spin ($t_{\text{sim,NNI}}$) to the average computation time required to evolve the NNI-sGLE hybrid over one time step ($t_{\text{sim,hyb}}$). Table 5.2 summarizes the computational gain G of both the NNIa-sGLEa and NNIb-sGLEb hybrid for the systems consisting of $N = 40$, 160 and 640 spins. (These results were obtained on a quad-core 2011 MacBook Pro.) While the computational gain for the 40-spin system is negligible, it increases with N. This suggests that an NNI-sGLE hybrid might be the only computationally viable option for realistic systems (e.g., one gram of Fe contains about 10^{22} atoms).

5.5.4 Moment analysis of stochastic Ginzburg-Landau equations

The analyses of Sections 5.5.1 and 5.5.2 demonstrate that the sGLEs provide the estimates of the magnetization's mean and variance that are consistent with those computed with the corresponding NNI models. Also, the computational cost of solving the sGLEs is much smaller than that of the NNI models. These considerations lead us to replace the NNI component of the hybrid simulations with its sGLE counterpart. This procedure is fairly generic: one can reduce the computational cost of a hybrid by replacing its microscale component with the corresponding stochastic coarse-scale description, as was done in [5, 6].

In the context of the NNI-sGLE hybrids, using the sGLE instead of the NNI component replaces the hybrid with a sGLE defined on the whole computational domain. The strength of the noise term in this equation exhibits a jump discontinuity at the hybrid's

interfaces, e.g.,

$$C_\eta = 2c\Gamma_a k_B T \delta(x-y)\delta(t-\tau) \begin{cases} \zeta_a^2 & \text{for} \quad x \in [L/4, 3L/4) \\ A_2 & \text{for} \quad x \in [0, L/4) \cup [3L/4, L] \end{cases} \tag{5.15}$$

where ζ_a^2 is the noise multiplier first introduced in Section 5.5.1, and $A_2 \geq 0$ can be equal to or different from ζ_a^2. This formulation facilitates the analysis of noise propagation by employing the tools from the theory of stochastic differential equations. First, it allows us to derive (deterministic) moment equations satisfied by the statistical moments (e.g., mean and covariance) of the magnetization. Second, it enables us to compute the steady-state statistics directly, without resorting to transient simulations.

Due to the nonlinearity of the the sGLE, the derivation of the moment equations requires a closure approximation. While other closures (e.g., the Eyink-Levermore closure [8] or a closure by perturbation [112]) can be adopted for this purpose, in L we employ the Gaussian approximation to derive the moment equations governing the dynamics of the mean magnetization $\overline{m}(x,t)$, the cross-covariance $C_{\eta m}(y,\tau;x,t)$ between $m(x,t)$ and $\eta(y,\tau)$, and the auto-covariance $C_m(y,\tau;x,t)$ between $m(x,t)$ and $m(y,\tau)$,

$$\frac{\partial \overline{m}}{\partial t} = \Gamma_a \left\{ c^2 J \frac{\partial^2 \overline{m}}{\partial x^2} - \left[(A + 3B\sigma_m^2)\overline{m} + B\overline{m}^3 \right] \right\} \tag{5.16}$$

$$\frac{\partial C_{\eta m}}{\partial t} = \Gamma_a \left\{ c^2 J \frac{\partial^2 C_{\eta m}}{\partial x^2} - \left[A + 3B(\overline{m}^2 + \sigma_m^2) \right] C_{\eta m} \right\} + C_\eta \tag{5.17}$$

$$\frac{\partial C_m}{\partial t} = \Gamma_a \left\{ c^2 J \frac{\partial^2 C_m}{\partial x^2} - \left[A + 3B(\overline{m}^2 + \sigma_m^2) \right] C_m \right\} + C_{\eta m}, \tag{5.18}$$

where the coefficients A and B are defined in (5.6b). Since $m(x=0,t) = m(x=L,t)$ and $\eta(x=0,t) = \eta(x=L,t)$, their moments $\overline{m}$, C_m and $C_{\eta m}$ are periodic as well. L describes the numerical scheme used to solve the steady-state versions of these equations.

For the sGLE (5.8) with the white-noise source term $\eta_a(x,t)$ whose covariance

function is given by (5.15) to act as a valid proxy for the NNIa-sGLEa hybrid, the noise has to be statistically homogeneous, so that $A_2 = \zeta_a^2 = 1.285^2$. Both the sGLE and the corresponding moment equations accurately predict the mean steady-state magnetization to be $\overline{m}(x) \equiv 0$; this stems from the homogeneity of the mean equation (5.16). Figure 5.4 demonstrates that they also accurately reproduce the variance of the steady-state magnetization, $\sigma_m^2(x) \equiv 1$, with the relative error between the exact and the moment equations' solution equal to 0.5%.

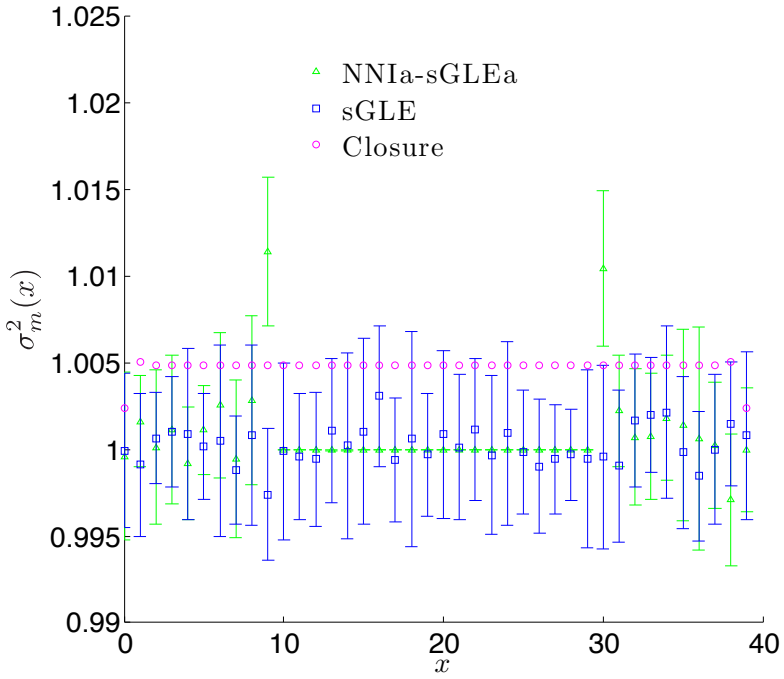

Figure 5.4: The magnetization variance computed with the closure and the sGLE (5.8) with statistically homogeneous white-noise source term $\eta_a(x,t)$ whose covariance function is given by (5.15) with $A_2 = \zeta_a^2$.

5.6 Summary and Conclusions

We investigated the performance of several hybrid models of phase transitions in one-dimensional ferromagnetic systems. These hybrids couple fine-scale (the 1D

nearest-neighbor Ising model with spin-flip dynamics, or NNI) and coarse-scale (the 1D time-dependent Ginzburg-Landau equation, or GLE) models. We considered two versions of the NNI models, one using the Suzuki-Kubo transition rate (the NNIa model) and one employing the heat-bath rate (the NNIb model). Both versions were coupled with their deterministic (the GLEa and GLEb models) and stochastic (the sGLEa and sGLEb models) coarse-scale counterparts. We used the steady-state statistics (mean and variance) of the magnetization of a ferromagnet evolving to thermodynamic equilibrium to analyze the effect of noise propagation in hybrid simulations.

Our analysis leads to the following major conclusions.

1. The NNIa-GLEa hybrid fails to predict the mean magnetization of a one-dimensional ferromagnet, while the NNIb-GLEb hybrid yields the accurate predictions. This is due to the theoretical limitations of the GLEa model in 1D. In higher spatial dimensions, both hybrids are likely to produce accurate estimates of the mean magnetization.

2. Both versions of the NNI-GLE hybrid fail to correctly estimate the magnetization variance outside the region where the NNI method is deployed. In other words, such deterministic hybrids fail to propagate the noise generated in the region with the fine-scale simulations (the NNI models) into the regions with the coarse-scale simulations (the GLE models).

3. Adding random fluctuations (a Gaussian white-noise source term) to the GLE models leads to the NNI-sGLE hybrids (both the NNIa-sGLEa and NNIb-sGLEb versions) that accurately capture the mean and variance of the magnetization. This is a general feature of hybrid modeling: the noise generated in a hybrid's fine-scale component is best handled by introducing a random source term into the hybrid's coarse-scale (and otherwise deterministic) component.

4. The NNI-sGLE hybrids provide a significant gain over the NNI simulations in terms of computational time. The gain increases exponentially with the lattice size (the number of spins) used to represent a ferromagnet in the NNI simulations.

5. Replacing the NNI component of the hybrid simulations with its sGLE counterpart further increases the computational efficiency of the hybrid models. This procedure is fairly generic: one can reduce the computational cost of a hybrid by replacing its fine-scale component with the corresponding stochastic coarse-scale description.

6. The use of the (deterministic) moment equations that govern the dynamics of the magnetization's mean and covariance in place of the sGLE provides an additional boost to the hybrid's efficiency. The derivation of the moment equations requires a closure approximation, which can reduce the hybrid's accuracy. The Gaussian approximation proved to be accurate in the setting considered here.

Future extensions of the presented analysis include the studies of the performance of the NNI-GLE hybrids in higher spatial dimensions and in phenomena where temporal dynamics is of primary importance (e.g., the identification of first-passage times). Another topic of interest is the derivation of a master equation for NNI-sGLE hybrids, which would account for the spatiotemporal discretization of the sGLE and the coupling of Ising spins to magnetization field.

5.7 Acknowledgements

The dissertation author thanks S. Puri and K. Leung for helpful discussions. The material contained in this Chapter is based on research supported in part by the Computational Mathematics Program of the Air Force Office of Scientific Research and by the National Science Foundation award EAR-1246315, and carried out as work LA-

UR-12-24412 under the auspices of the Department of Energy at Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396. It is, in part, a reprint of the material as it appears in the paper “Noise propagation in hybrid models of nonlinear systems: The Ginzburg-Landau equation” by S. Taverniers, F.J. Alexander, and D.M. Tartakovsky, published in *Journal of Computational Physics*, 262:313–324, Elsevier Inc. (2014) [113]. The dissertation author is the primary investigator and author of this paper.

Chapter 6

Physics-based statistical learning approach to mesoscopic model selection

6.1 Abstract

In materials science and many other research areas, models are frequently inferred without considering their generalization to unseen data. We apply statistical learning using cross-validation to obtain an optimally predictive coarse-grained description of a two-dimensional kinetic nearest-neighbor Ising model with Glauber dynamics (GD) based on the stochastic Ginzburg-Landau equation (sGLE). The latter is learned from GD "training" data using a log-likelihood analysis, and its predictive ability for various complexities of the model is tested on GD "test" data independent of the data used to train the model on. Using two different error metrics, we perform a detailed analysis of the error between magnetization time trajectories simulated using the learned sGLE coarse-grained description and those obtained using the GD model. We show that both for equilibrium and out-of-equilibrium GD training trajectories, the standard phenomenological description using a quartic free energy does not always yield the most predictive

coarse-grained model. Moreover, increasing the amount of training data can shift the optimal model complexity to higher values. Our results are promising in that they pave the way for the use of statistical learning as a general tool for materials modeling and discovery.

6.2 Introduction

Due to limitations in computational resources, the behavior of complex systems (e.g. climate, turbulent flow, materials under shock loading) often needs to be modeled using a coarse-grained description that captures the phenomenon of interest. Coarse-grained models cannot be perfect of course, since many microscopic degrees of freedom are absent. The Mori-Zwanzig formalism [88, 128] tells us that the relevant coarse-grained description should contain both noise [16, 113] and memory kernels to represent the "integrated out" fine scale dynamics. Deriving an appropriate coarse-graining analytically is therefore extremely difficult. Statistical learning provides a tractable way of finding a coarse-grained description that is able to predict the results of new experiments or simulations beyond those used in the model construction, which serves as a true objective test of the model. In fact, unlike traditional approaches, statistical learning can also serve as a coarse-graining strategy in cases where there is no clear separation of spatial and/or temporal scales. For example, it may be applied to problems that involve inhomogeneous flows (e.g. multicomponent fluids, complex fluids) and those in materials science where the coarse-grained description needs to account for inhomogeneities at a finer scale, e.g., microstructural defects. In such cases, techniques such as the heterogeneous multiscale method [38] that demand a clean separation in time scales are not applicable.

In the statistical approach to coarse-graining discussed here, the goal is to search over a certain class of coarse-grained models and find the complexity for which the

description is “optimally predictive”. This technique is called regularization, and to estimate the generalization error we use cross-validation. The latter involves randomly dividing data (either from experiments or simulations) into “training” and “test” samples. Ideally one would like both groups of data to be infinite, but in practice one only has a limited amount of data to work with. Experimentalists can only synthesize a small number of material samples, and in molecular dynamics simulations one is also limited to a finite number of samples. For the purposes of the current analysis, we will assume that the amount of training data is limited but that we can test our learned model on an infinite amount of data independent of the training samples. An extension to cases where both training *and* test data are finite will be the topic of future research.

Selecting the appropriate model regularization is of paramount importance because it can minimize both underfitting and overfitting. An underfitted model is too simplistic and therefore fails to capture much of the useful information available in the training data; hence, it will perform sub-optimally on data independent of the training set. In contrast, overfitting refers to the case where an overly complex model describes the many irrelevant details that appear in the training data by chance. An overfitted model will therefore be also less successful in generalizing to new data from simulations or experiments that are outside the class of the training data. The model developed in this study avoids common issues associated with overfitting by using an effectively infinite amount of test data independent of the samples on which it was trained, and selecting the complexity that makes it most predictive of this test data.

As an error estimator, cross-validation has been used for a number of years. When the amount of data is very limited though, there can be a significant difference between the cross-validation error and the actual error [33, 34]. Moreover, a detailed error analysis is often lacking in physics modeling applications. Our motivation is to learn mesoscale models from microstructural data incorporating prior domain knowledge and physical

symmetries. We will focus on the time-dependent stochastic Ginzburg-Landau equation (sGLE) which provides a coarse-grained description of a kinetic nearest-neighbor Ising model with Glauber dynamics (GD) and of which the dynamics are expected to be particularly straightforward to learn. While our approach is related to that in [35], we do not assume a prior distribution for the learned parameters and do not include a penalty for overfitting or complexity in the Bayesian Information Criterion. Moreover, instead of simply fitting a regular quartic free energy to a single or joint magnetization distribution function as in [66, 67], we consider higher order terms and find the parameters that optimally predict GD data independent of the samples on which the model was trained. The inclusion of terms beyond fourth order in the free energy accounts for the fact that we are in a regime of finite coarse-graining block sizes, and hence not at a fixed point in the renormalization group theory [24, 58]. Our current approach does not account for higher order spatial gradients, which can play an important role out of equilibrium; we plan to include these terms in future versions of the model.

Section 6.3 describes the microscopic GD model and its mesoscale description provided by the sGLE. Section 6.4 details our design loop used to select the optimal complexity of the sGLE for each amount of training data considered, after which Sec. 6.5 discusses the results of the error analysis we performed in order to arrive at an optimally predictive model. Section 6.6 summarizes our conclusions and discusses possibilities for future work.

6.3 The Kinetic Ising Model and its Coarse-scale Description by a Stochastic Ginzburg-Landau Equation

6.3.1 Kinetic Ising model with Glauber dynamics

The Ising model with nearest-neighbor interactions [60] is a simple, yet very rich, model in statistical mechanics for describing ferromagnetic behavior. Consider a two-dimensional (2D) ferromagnet with atoms arranged on an $N_1 \times N_2$ square lattice. The spin $s_{i,j}$ (where $i = 0,\ldots,N_1-1$ and $j = 0,\ldots,N_2-1$) of each atom can be in one of two states, $s_{i,j} = \pm 1$, and can only interact with its four adjacent spins. We can add dynamics to this system by flipping spins with a certain transition rate w, and the result is a kinetic nearest-neighbor Ising model with spin-flip (Glauber) [52] dynamics (which we will refer to as GD). This allows us to express the time evolution of the spin system through a master equation given by

$$\frac{d}{dt}\mathbb{P}(\sigma;t) = \sum_{\sigma'}[w(\sigma' \to \sigma)\mathbb{P}(\sigma';t) - w(\sigma \to \sigma')\mathbb{P}(\sigma;t)] \tag{6.1}$$

where $\mathbb{P}(\sigma;t)$ is the joint probability of finding the system in spin configuration σ at time t, and the w's are the transition rates between two $N_1 \times N_2$-spin configurations differing only in the value of one spin, $s_{i,j}$. For w we choose the heat bath rate, given by

$$w_{HB}(\sigma \to \sigma') = \kappa\left(1 + e^{-\beta[\mathcal{H}(\sigma)-\mathcal{H}(\sigma')]}\right)^{-1}. \tag{6.2}$$

Here $\mathcal{H}(\sigma)$ represents the Hamiltonian of the spin system with configuration σ, $\beta = 1/(k_B T)$ with T the system temperature, and κ^{-1} sets the time scale of the spin-flip process and can depend both on T and the spins other than $s_{i,j}$. We simulate this kinetic

Ising model via a Monte Carlo (MC) algorithm with one MC step per spin; i.e., to complete one step in our Ising run, we perform $N_1 \times N_2$ times the following procedure:

1. Pick a random site (i, j) where $i = 0, \ldots, N_1 - 1$ and $j = 0, \ldots, N_2 - 1$.

2. Draw a number r_1 from a uniform distribution on $[0, 1]$.

3. Flip the spin $s_{i,j}$ if $r_1 < w(\sigma \to \sigma')$, or leave it in its original state if $r_1 \geq w(\sigma \to \sigma')$.

In all of our work here, we initialize the lattice by selecting spins to be +1 or -1 randomly with equal probability.

6.3.2 Ginzburg-Landau equation

By invoking a phenomenological coarse-graining approach, it is possible to obtain a mesoscopic model of GD given by a time-dependent Ginzburg-Landau equation (*GLE*) [1]. The latter will describe the spatiotemporal evolution of an "order parameter", ϕ, a field variable which represents the instantaneous average of Ising spin values over some portion of a ferromagnetic material (also called "magnetization"). At finite temperatures, one needs to account for fluctuations, which can be added via a white noise term to obtain an overdamped stochastic relaxation equation

$$\frac{\partial \phi(\mathbf{x},t)}{\partial t} = -M \frac{\delta F[\phi(\mathbf{x},t)]}{\delta \phi(\mathbf{x},t)} + \eta(\mathbf{x},t), \tag{6.3}$$

with M the mobility which sets the time scale of the dynamics. Here $\eta(\mathbf{x},t)$ is a zero-mean Gaussian space-time white noise with an auto-covariance which, according to the fluctuation-dissipation theorem, scales linearly with M and the system temperature. Moreover, $F[\phi]$ is an effective free energy for the system, and may be developed as a

[1] A Ginzburg-Landau model for the free energy can also be used for other systems; see, e.g., [53].

power series in ϕ and its derivatives

$$F \sim \sum_{k \text{ even}} a_k \phi^k + b\, \phi \nabla^2 \phi + \ldots \tag{6.4}$$

Only even powers of ϕ are allowed in (6.4) due to the symmetry of the Ising Hamiltonian. In the context of late-stage domain growth, renormalization group arguments indicate that the Ising model with Glauber dynamics is in the universality class of Model A dynamics [58], and hence well represented at the coarse scale by having only a_2 and a_4 different from zero. However, since we are focusing on intermediate time and length scales, we relax this assumption and instead consider a model class for the free energy consisting of even-term polynomials with degree two or greater. The complexity that we eventually select is the one for which the model is optimally predictive of GD data independent of the samples from which it was learned.

6.4 Numerical Algorithm

Our goal is to learn the parameters of a discrete version of the stochastic Ginzburg-Landau equation (sGLE) which evolves the magnetization ϕ from time t_n to time t_{n+1} according to

$$\begin{aligned} \phi_{n+1,i,j} = \phi_{n,i,j} &+ \alpha_0(\phi_{n,i+1,j} + \phi_{n,i-1,j} + \phi_{n,i,j+1} \\ &+ \phi_{n,i,j-1} - 4\phi_{n,i,j}) + \sum_{k=0}^{\frac{C-1}{2}} \alpha_{k+1} \phi_{n,i,j}^{2k+1} \\ &+ \alpha_{(C+3)/2}\, \xi_{n,i,j}, \end{aligned} \tag{6.5}$$

where C is the model complexity [2] and the $\xi_{n,i,j}$ are independent, identically distributed standard normal random variables. Here n and $n+1$ refer to times t_n and t_{n+1}, while $i = 0,\ldots,\bar{N}_1 - 1$ and $j = 0,\ldots,\bar{N}_2 - 1$ indicate the spatial position of the spin blocks resulting from the coarse-graining procedure, with $\bar{N}_1$ and $\bar{N}_2$ the number of blocks in both spatial directions. Given block-averaged training data $S_{n,i,j}$ with $n = 0,\ldots,n_{\mathrm{eq}}$, where n_{eq} will be specified for each of our numerical experiments in Sec. 6.5, we would like to find the set of parameters $\boldsymbol{\alpha}_{\mathbf{opt}} \equiv \{\alpha_0, \alpha_1, \ldots, \alpha_{(C+3)/2}\}$ that maximizes the likelihood of observing this data using the sGLE model. For notational convenience, the $\bar{N}_1 \times \bar{N}_2$ block-averaged spin configuration after n steps will be denoted by S_n; the notation $S_{n,i,j}$ will refer to its (i,j)th matrix element. The same convention will be used for ϕ. It turns out $\boldsymbol{\alpha}_{\mathbf{opt}}$ is the solution to a linear system

$$A\boldsymbol{\alpha}_{\mathbf{opt}} = \mathbf{b}, \tag{6.6}$$

where the components of A and the elements of $\mathbf{b}$ involve products of S_n, its powers and its discrete Laplacian. The dimension of this system is given by the number of free parameters that make up the model (6.5), which is equal to $(C-1)/2+3$. For more details, we refer the reader to Appendix M where we have worked out the case of $C = 3$. We then test our learned sGLE model against independent GD data ("test" data) to ascertain how well learned models of different complexities perform on unseen data. By calculating the root-mean-square (RMS) error between GD test trajectories and those simulated using the learned sGLE, we then determine for which complexity C the latter is optimally predictive of the GD test data.

We will consider two error metrics in our analysis, which we will refer to as the "type 1" and "type 2" error. For the type 1 error, we calculate the sGLE grid at time

[2] We only consider odd complexities in our analysis.

t_{n+1} through (6.5) but replace ϕ_n with the block-averaged GD test data at time t_n. This error is of the same type as the error that we want to minimize when calculating $\boldsymbol{\alpha}_{\mathbf{opt}}$ from the GD training data, where we search for the set of parameters that maximizes the likelihood of observing the training data at t_{n+1} given the sGLE model and the training data at t_n, for every n (see Appendix M). For the type 2 error, we evolve the sGLE grid in time through (6.5) directly, i.e. we do not keep referring back to the GD test data at each time t_n.

A flowchart of the operational algorithm is shown in Fig. 6.1. At a high level, our approach for computing one data point in the error probability density function (pdf) for a learned sGLE model of complexity C, given a number of training samples N_{train}, can be described as follows (see Appendix N for more details).

1. We simulate N_{test} independent GD test sample trajectories. For each trajectory, we let the Ising system evolve over n_{mc} (specified in Sec. 6.5) steps, after which we take another n_{eq} steps during which we record the block-averaged Ising configuration. Each of these steps represents one MC step per spin as detailed in Sec. 6.3.1.

2. We simulate N_{train} independent GD training sample trajectories. We let the spins evolve over n_{mc} steps, and then record their block-averaged configuration over the next n_{eq} steps.

3. Using the data gathered during the last n_{eq} steps of each training trajectory, we compute the coefficients of the learned sGLE polynomial using a log-likelihood analysis (see Appendix M).

4. With the parameters calculated in step 3, we now simulate N_{test} sGLE trajectories. Each trajectory consists of n_{eq} steps, with each step involving the advancement of the discrete sGLE (6.5) from one discrete point in time to the next.

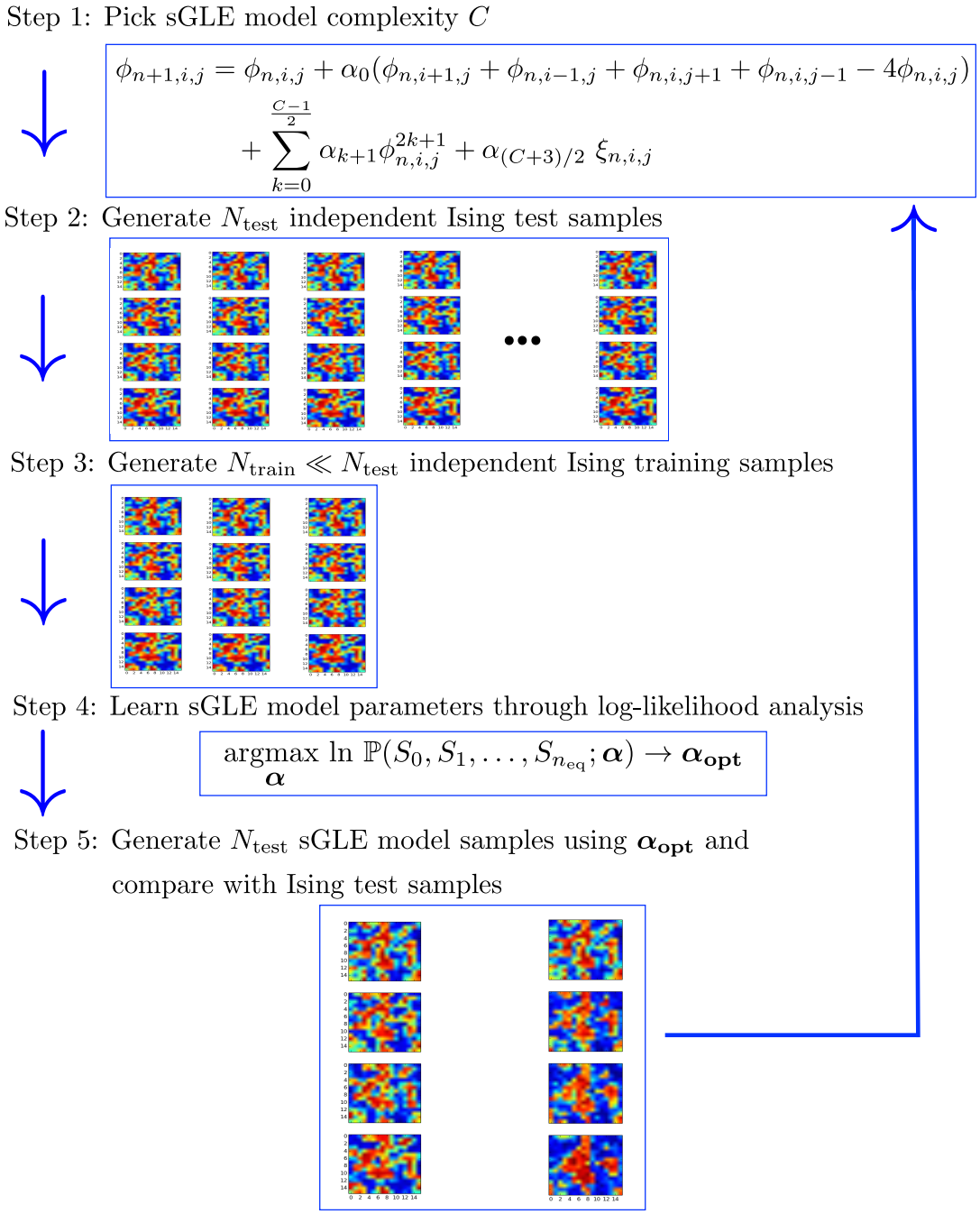

Figure 6.1: Flowchart for the model complexity selection algorithm.

5. For the kth sGLE trajectory, we calculate the RMS error ε_k between this trajectory and the kth block-averaged GD test trajectory.

6. Finally, we compute the test-averaged error

$$\varepsilon = \frac{1}{N_{\text{test}}} \sum_{k=1}^{N_{\text{test}}} \varepsilon_k, \tag{6.7}$$

which we will call the "type 1" or "type 2" test error depending on how the sGLE trajectory has been calculated (see our above definition of these errors).

The quantity ε represents one point in the error pdf for the considered complexity C and number of training samples N_{train}. The entire pdf is then obtained by repeating the above procedure except for step 1 (we use the same GD test trajectories for each point in the pdf) $N_{\text{real}} \gg 1$ number of times. We will denote the sample (i.e., realization) mean and variance of this pdf by $\bar{\varepsilon}$ and s_ε^2, respectively.

6.5 Error Analysis and Main Results

We now present the pdfs of the type 1 and type 2 errors, defined in Sec. 6.4, for different complexities $C = 3, 5, 7$ or 9, given a finite number of GD training samples N_{train}. We will use $N_{\text{test}} = 1000$ GD test samples, which provides an accurate generalization error [3]. Moreover, we will build up the error histograms using $N_{\text{real}} = 5000$ independent realizations, and consider both training data in equilibrium and out of equilibrium. For the equilibrium case, we measure the energy of the spin system and choose n_{mc} as the number of steps after which thermal equilibrium has been reached. Out of equilibrium, we choose n_{mc} such that after n_{mc} steps domains have started to form. Learning the sGLE model parameters from GD data for which the block size is smaller than the size of a

[3] In future work, we will carry out an error analysis when only a small number of test samples is available [33, 34].

typical domain allows the Laplacian in the sGLE to better capture gradients in the GD data, and should hence yield a more accurate coarse-grained description. For both cases, we determine an appropriate value for n_{eq} through trial and error, and choose values that provide a sufficient amount of input data to our log-likelihood solver. We will use $n_{\text{mc}}+1=2000$ and $n_{\text{eq}}+1=100$ for the equilibrium case, and $n_{\text{mc}}+1=n_{\text{eq}}+1=500$ out of equilibrium (these values include the number of steps plus the starting condition at $t=0$ or after n_{mc} steps). In all cases, we simulate the Glauber dynamics on a 256×256 spin lattice with periodic boundary conditions, and coarse grain using 16×16 blocks. We set the parameter κ in (6.2) equal to 1. Finally, next to each error distribution, we show a plot of the free energy constructed using the learned model parameters. These plots are meant to serve as a check of the physical soundness of our approach in the sense that, consistent with common knowledge, a single-well potential should be observed above the phase transition and a double-well potential below the phase transition. However, they do not convey any information regarding the predictiveness of our model, which follows entirely from the error distributions.

6.5.1 Type 1 test error pdfs for training data in equilibrium

We first consider the case where we train on $N_{\text{train}}=8$ GD trajectories in equilibrium, obtained by quenching the spin lattice from infinite temperature to either $T=1.6\ J/k_B$ (below $T_c=2.269\ J/k_B$ [4]) or $T=2.8\ J/k_B$ (above T_c). As Fig. 6.2 shows, at $T=1.6\ J/k_B$ the error pdf's sample mean $\bar{\varepsilon}$ clearly decreases with complexity C. Hence, the most predictive sGLE model is that with the highest complexity considered, $C=9$. At $T=2.8\ J/k_B$, however, the error pdfs for all the complexities overlap almost completely (see Fig.6.3), indicating that the regular third-order sGLE polynomial is

[4] T_c refers to the critical temperature for a second order phase transition following Onsager's [93] solution.

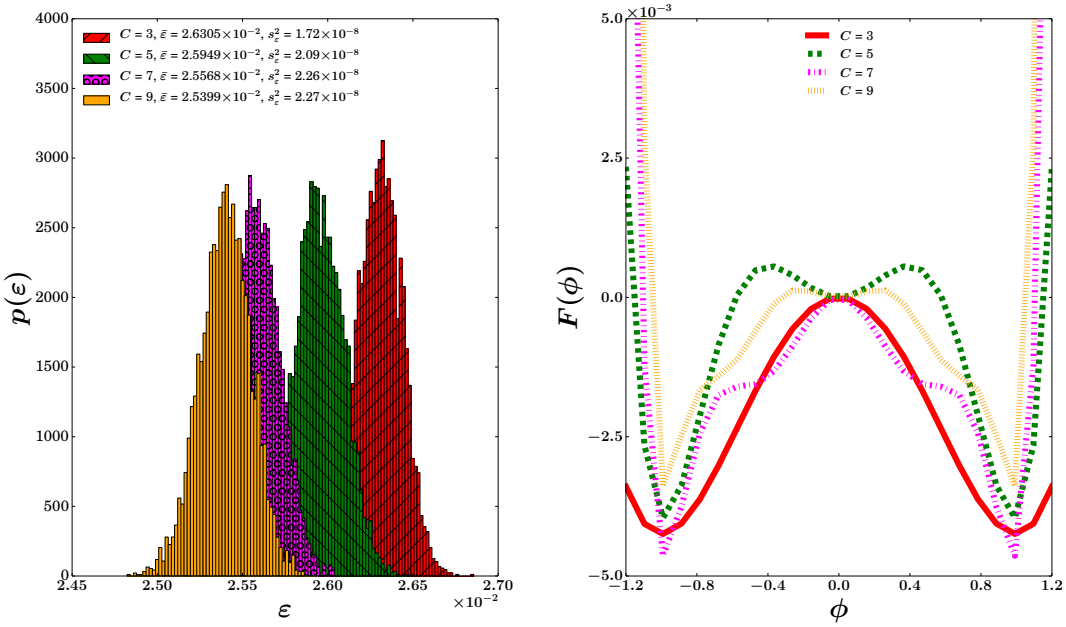

Figure 6.2: Type 1 error pdfs (left) and learned Ginzburg-Landau free energy (right) for different complexities of an sGLE learned from eight equilibrium GD training samples at $T = 1.6\, J/k_B$ (below the phase transition). The sGLE with $C = 9$ predicts the GD test data best.

adequate to predict the coarse-grained Glauber dynamics.

6.5.2 Type 1 test error pdfs for training data out of equilibrium

Next, we look at the case where we train on various amounts of GD trajectories out of equilibrium, obtained by quenching the spin lattice from infinite temperature to $T = 2.2\, J/k_B$ (just below T_c). As Figs. 6.4-6.7 show, regardless of the amount of training data the regular ϕ^4 form of the free energy ($C = 3$) is not optimally predictive of the GD test data. The complexity for which the sGLE model best predicts the GD test trajectories varies with the amount of training data. For small amounts of training data (i.e. $N_{\text{train}} = 1$ or 2), the histograms for the different complexities largely overlap. As the number of training samples is increased to 16, the pdfs for $C = 5$ and higher can be more clearly distinguished from that for $C = 3$. When one further increases the amount

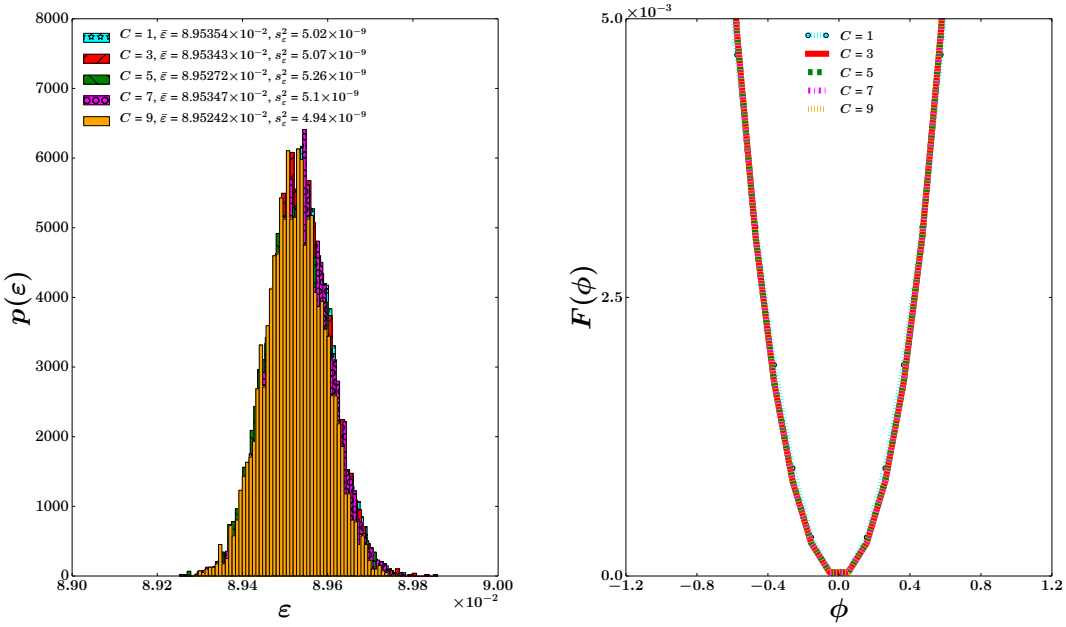

Figure 6.3: Type 1 error pdfs for different complexities of an sGLE learned from eight equilibrium GD training samples at $T = 2.8\ J/k_B$ (above the phase transition). All considered model complexities are equally predictive of the GD test data.

of training samples to 128, the pdf for $C = 3$ becomes fully distinct from those for higher complexities, and the pdf for $C = 5$ is becoming more distinct from those for $C = 7$ and $C = 9$. In sum, increasing the amount of GD training data causes the complexity at which the sGLE is optimally predictive to shift toward higher values.

6.5.3 Type 2 test error pdfs for training data out of equilibrium

Finally, we repeat the simulations in Sec. 6.5.2 for the type 2 error. We find that for this error type the regular ϕ^4 free energy *does* optimally predict the GD test data, regardless of the amount of training data. Figure 6.8 shows this for the case of one GD training sample. We note here that it is to be expected that $\bar{\varepsilon}$ is bigger for the type 2 error than for the type 1 error, since the latter is calculated in the same way as the optimization error for obtaining the $\boldsymbol{\alpha}_{\mathbf{opt}}$ is calculated from the GD training data, while the former is not.

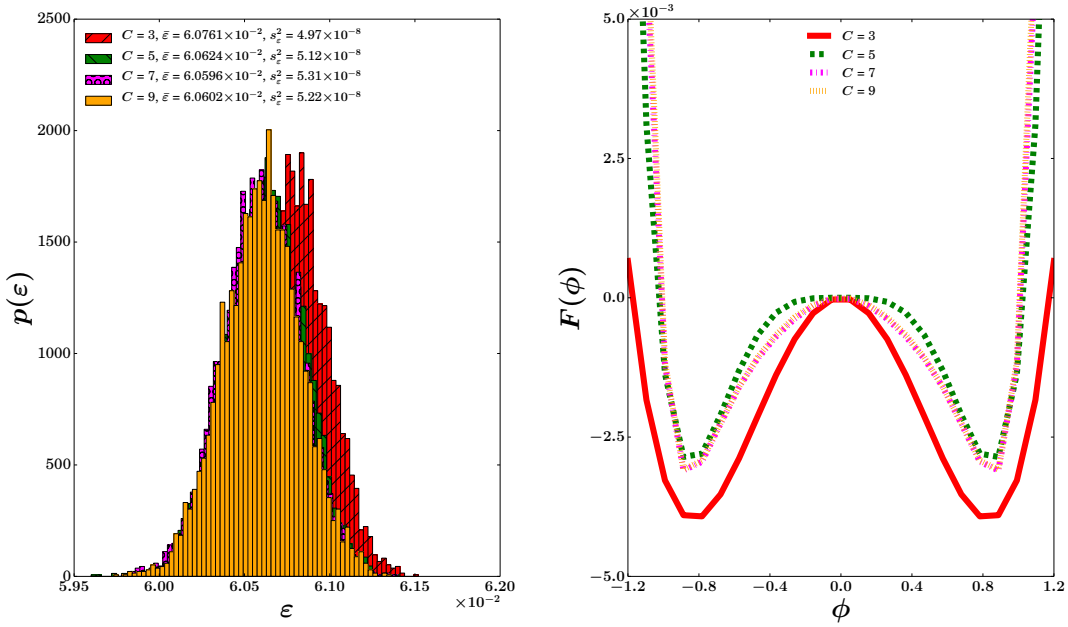

Figure 6.4: Type 1 error pdfs for different complexities of an sGLE learned from one out-of-equilibrium GD training sample at $T = 2.2\ J/k_B$ (just below the phase transition). Complexities 5 and higher are optimally predictive of the GD test data, but the corresponding error pdfs still largely overlap with that for $C = 3$.

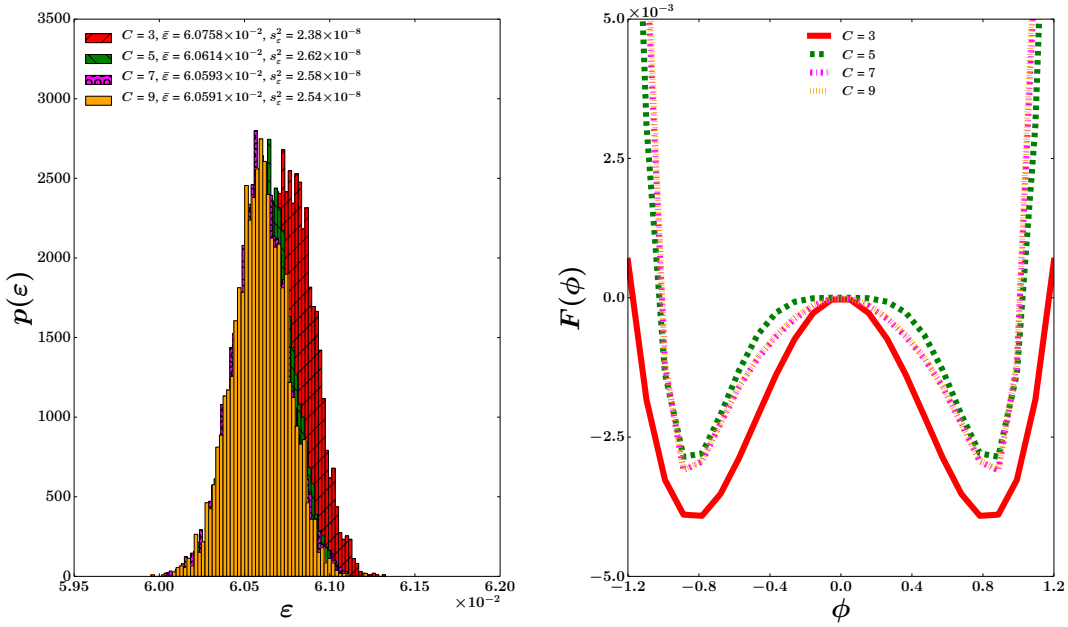

Figure 6.5: Type 1 error pdfs for different complexities of an sGLE learned from two out-of-equilibrium GD training samples at $T = 2.2\ J/k_B$.

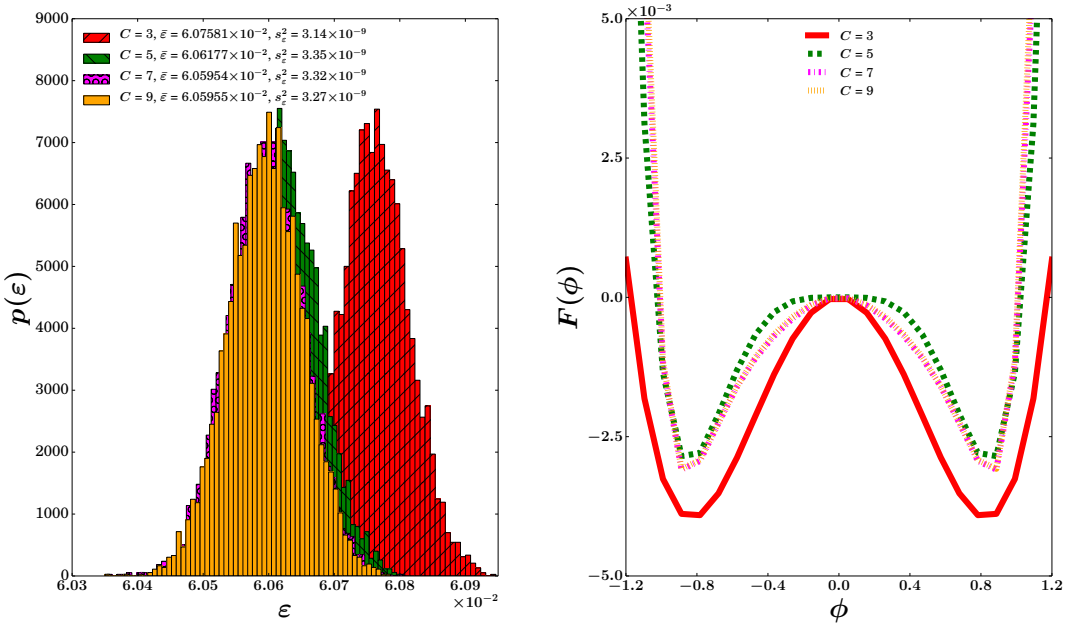

Figure 6.6: Type 1 error pdfs for different complexities of an sGLE learned from 16 out-of-equilibrium GD training samples at $T = 2.2\ J/k_B$. The error pdfs for $C = 5$ and higher are now clearly distinct from that for $C = 3$.

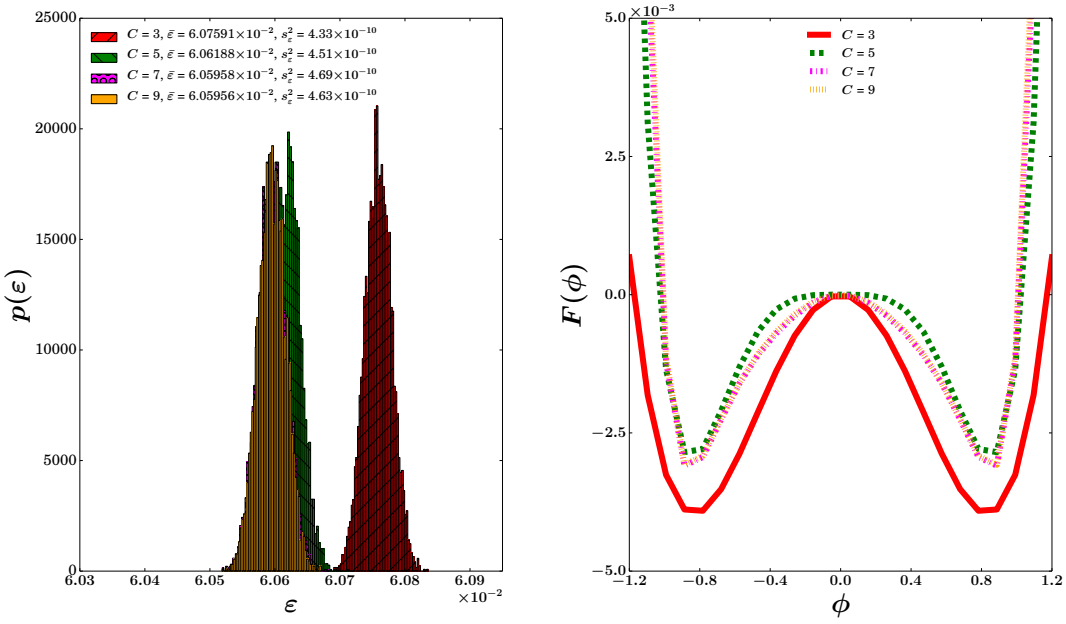

Figure 6.7: Type 1 error pdfs for different complexities of an sGLE learned from 128 out-of-equilibrium GD training samples at $T = 2.2\ J/k_B$. The error pdf for $C = 5$ and those for $C = 7$ and higher are becoming more distinct, shifting the optimally predictive model complexity to $C = 7$.

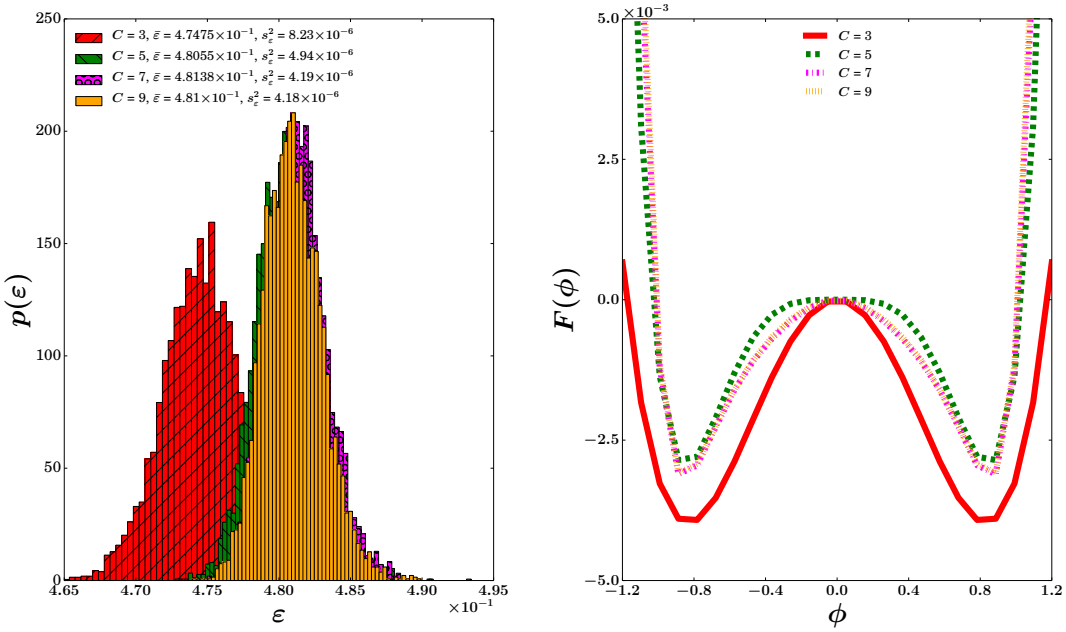

Figure 6.8: Type 2 error pdfs for different complexities of an sGLE learned from one out-of-equilibrium GD training sample at $T = 2.2\ J/k_B$. The sGLE with $C = 3$ is most predictive of the GD test data.

6.6 Summary and Conclusions

By performing a detailed error analysis in the context of a statistical learning approach using cross-validation, we derive an optimally predictive coarse-grained description of a two-dimensional kinetic nearest-neighbor Ising model with Glauber dynamics (GD) based on the stochastic Ginzburg-Landau equation (sGLE). The latter is learned from microscopic GD "training" data through a log-likelihood analysis, and its capacity to predict GD "test" data independent of the training data is analyzed for various model complexities using two error metrics and varying amounts of training data.

Our analysis yields the following major conclusions:

1. For the type 1 error, a complexity of 3 in the sGLE force equation does *not* yield an optimally predictive model for any amount of training data that we investigated. Moreover, the model complexity yielding the most predictive coarse-grained

description increases with the amount of GD training data.

2. For the type 2 error, the regular Ginzburg-Landau description using a ϕ^4 mean-field free energy *does* yield the most predictive model irrespective of the amount of GD training data.

The principled methodology developed here for simple Model A dynamics can be applied to more complicated problems such as Model H dynamics [58]. A particular application which might benefit from this work is the use of data generated from experiments, e.g. ultrafast X-ray diffraction patterns of structural phase transitions in semiconductor crystals to generate models describing crystal disordering [47]. In general, our approach can be utilized in any application using a Ginzburg-Landau functional, e.g. in phase field simulations of materials.

Directions for future work include studying the effects of the coarse-graining block size (in both space and time), performing a rigorous analysis of the stability and discretization error of our numerical scheme, and expanding the model class by including operator terms that account for higher order spatial gradients. Moreover, it is desirable to complement the current computational analysis with a rigorous theoretical derivation of an expression for the error probability density function in terms of model complexity, number of training samples and coarse-graining block size.

6.7 Acknowledgements

The material in this Chapter is based on research carried out as work LA-UR-15-23538 under the auspices of the U.S. Department of Energy by Los Alamos National Laboratory (LANL) under Contract No. DE-AC52-06NA25396, and supported by Laboratory Directed Research and Development (LDRD) Project No. 20140013DR. It is, in part, a reprint of the material as it appears in the paper “Physics-based statistical

learning approach to mesoscopic model selection" by S. Taverniers, T.S. Haut, K. Barros, F.J. Alexander, and T. Lookman, published in *Physical Review E*, 92:053301, American Physical Society (2015) [114]. Numerical simulations were performed using a hybrid C++/Python code run on LANL computing facilities. The dissertation author is the primary investigator and author of this paper.

Chapter 7

Conclusions

We considered two nonintrusive approaches for multiscale/multiphysics problems in the presence of stochastic fluctuations: domain decomposition (DD) and stochastic collocation (SC). Noise was introduced into the system as a random source term (Chapters 2, 5 and 6), a random boundary condition (Chapter 3), or as uncertain input parameters (Chapter 4). We developed and analyzed tightly-coupled DD algorithms for two testbed problems: one- and two-dimensional coupled linear diffusion equations, one of which was forced by a Gaussian space-time white noise, and one-dimensional multiscale hydrogen diffusion in a multilayer Pd-Ta dense membrane, driven by a truncated Gaussian noise at the domain boundary. In the former, tight coupling was achieved through Picard's or Newton's iteration; in the latter, Jacobian-free Newton-Krylov with Generalized Minimum Residual was employed. Applying SC to a two-dimensional multimaterial, equilibrium radiation-diffusion equation with one, two or three uncertain input parameters, we estimated the mean and variance of the energy, deposited in a brain tumor through X-ray irradiation, as a function of the statistical moments of the random inputs. We constructed an algorithm refinement (AR) hybrid for the one-dimensional real, cubic Ginzburg-Landau equation (GLE), and analyzed its ability to propagate fine-scale noise

throughout the entire computational domain; in addition, we investigated the viability of a moment approach with a Gaussian closure. Finally, using statistical learning involving a log-likelihood analysis, we derived a two-dimensional stochastic GLE, and analyzed its ability to predict microscale data outside of the data set on which the model was trained.

In Chapters 2 and 3, we demonstrated that the presence of stochastic noise alters the computational efficiency of an iterative ("implicit") coupling algorithm in a way that depends on both the noise type and the iteration technique. For a space-time white noise inside the domain, the analysis in Chapter 2 revealed that increasing the noise strength resulted, on average, in more iterations per communication. While Newton's method maintained a linear scaling with noise amplitude, Picard's coupling could scale nonlinearly. On the other hand, a truncated boundary noise was found to have a "smoothing" effect and reduced the number of Newton iterations (in a Jacobian-free Newton-Krylov coupling) per communication as its coefficient of variation was increased. The impact of noise on the performance of the iterative coupling also affected the latter's relative efficiency compared to its single-iteration ("explicit") counterpart. For a given solution error and at low noise strength, the simulations for the space-time white noise that used implicit coupling completed, on average, faster than those with explicit coupling; this was true for both the longer and shorter time interval between two subsequent implicit coupling communications that we considered. The reason is that for explicit coupling to achieve the same solution error as its implicit counterpart, its communication frequency needs to be increased, thereby increasing simulation time. At high noise strength however, the larger number of iterations per communication for the implicit coupling caused the latter to become less efficient than the explicit coupling for sufficiently short time intervals between two subsequent implicit coupling communications. Increasing the latter interval enabled the implicit coupling to again outperform its explicit counterpart. For the boundary noise however, the implicit coupling

was shown to have a higher efficiency than its explicit counterpart for all noise strengths and communication times considered.

We also investigated the effect of noise on the stability of the above DD algorithms. In Chapter 2, we found that for a linear multiscale problem the stability of the DD approach with implicit coupling was not affected by the presence of noise. This is in stark contrast to the nonlinear multiscale problem in Chapter 3, for which we showed that the concentration dependence of the diffusion coefficient in Pd causes the stability conditions to depend on the ensemble mean of the solution, which is different from its counterpart solving the corresponding deterministic problem with the fluctuations averaged out. It is therefore incorrect to base predictions about the stability of a DD algorithm for a nonlinear problem with noise on results obtained prior to the introduction of this noise into the system, or with the fluctuations averaged out.

For both the linear system in Chapter 2 and nonlinear system in Chapter 3, we found that the DD algorithm with implicit coupling preserves the temporal order of accuracy of the subdomain solvers.

The analysis in Chapter 3 revealed that exchanging sample-averaged values of concentration and flux (moment-wise communication), rather than their values for a particular realization of the system dynamics (path-wise communication), across interfaces between subdomains leads to incorrect propagation of the boundary fluctuations and erroneous spatial profiles for both the mean and variance of the state variable, regardless of the level of spatial refinement. A coupling enforcing continuity of mean and variance similar to the approach followed in [31] might approach the accuracy of our path-wise coupling while being computationally less expensive. However, unlike the solution to the weakly nonlinear problems considered in [31], the solution to highly nonlinear problems such as the hydrogen diffusion in Chapter 3 will be highly non-Gaussian, and higher moments of the state variable might need to be exchanged, in addition to the mean

and variance, for a moment-wise coupling to achieve a similar solution error as path-wise communication.

The numerical experiments in Chapter 4 demonstrated that SC outperforms standard Monte Carlo simulation for all stochastic dimensions considered, with the biggest difference for the scenario of one random input parameter. Increasing the stochastic dimension enhanced the amplification of parametric uncertainty by the system's nonlinearity to the point where the magnitude of the predictive uncertainty in the energy deposition approached that of its predicted (mean) value itself. This again demonstrates the dramatic effect of random noise, even of low strength, on the behavior of nonlinear multiphysics systems, and hence the need to incorporate an accurate model of random fluctuations in a global algorithm development strategy.

In Chapter 5, we demonstrated, using the GLE as a testbed, that for AR hybrids modeling nonlinear problems whose macroscale dynamics is driven by microscale noise, the ability of the coarse-scale model to correctly capture the atomistic algorithm's fluctuations is of paramount importance to correctly predict both the mean and higher moments of the quantity of interest. Our results also confirmed that significant savings in computational cost may be achieved by applying an AR method versus a full-domain atomistic approach, and increasingly so for bigger systems. Finally, we established that a (deterministic) moment approach with Gaussian closure was able to provide, somewhat surprisingly, a reasonably accurate estimate of the magnetization variance, offering a direct way of computing such moments as opposed to employing numerical time integration of the state variable.

In Chapter 6, we found that the optimal complexity of the learned stochastic GLE for predicting microscale Ising data outside of the data set on which the model was trained may differ from that of standard models available in the literature. Moreover, this approach established a data-driven way of computing the magnitude of the model's

random source term, which may eliminate reliance on ad hoc “fudge factors” that plagued both stochastic GLE models used in the AR hybrid discussed in Chapter 5. Hence, statistical learning may prove to be a powerful tool in building atomistic-continuum AR hybrids that accurately predict system behavior both in the fine- and coarse-scale regions.

Appendix A

Conservative versus non-conservative coupling (Chapter 2)

To understand the need for a mass-conserving coupling algorithm, we consider the total mass M inside the region $[-\Delta x/2, \Delta x/2]$ (see Fig. 2.1)

$$M(t) = \int_{-\Delta x/2}^{\Delta x/2} \rho(x,t) \mathrm{d}x. \tag{A.1}$$

Its derivative yields

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \int_{-\Delta x/2}^{\Delta x/2} \frac{\partial \rho}{\partial t} \mathrm{d}x = \int_{-\Delta x/2}^{\Delta x/2} \frac{\partial}{\partial x}\left[D\frac{\partial \rho}{\partial x}\right] \mathrm{d}x = F_{1,N/2-1/2} - F_{2,1/2}, \tag{A.2}$$

where the flux $F = -D\partial\rho/\partial x$ is given by Fick's law. The interfacial fluxes $F_{1,N/2-1/2}$ and $F_{2,1/2}$ represent the amount of mass leaving the left subdomain per unit time and the amount of mass entering the right subdomain per unit time, respectively.

Integrating (A.2) between t_n and $t_{n+1} \equiv t_n + \Delta t_{\mathrm{com}}$ gives

$$\frac{\Delta M}{\Delta t_{\mathrm{com}}} = \bar{F}^n_{1,N/2-1/2} - \bar{F}^n_{2,1/2}, \tag{A.3}$$

where ΔM denotes the change in total mass inside $[-\Delta x/2, \Delta x/2]$ between t_n and t_{n+1}, and $\bar{F}^n_{1,N/2-1/2}$ and $\bar{F}^n_{2,1/2}$ are the Δt_{com}-averaged values of $F_{1,N/2-1/2}$ and $F_{2,1/2}$, respectively. Looking at the system dynamics between two subsequent inter-solver communications, any mass leaving the left subdomain should enter the right subdomain and cannot be "trapped" inside the interface region $[-\Delta x/2, \Delta x/2]$. Hence, the total mass inside $[-\Delta x/2, \Delta x/2]$ needs to remain constant, which requires equality of the Δt_{com}-averaged interface fluxes, i.e.

$$\bar{F}^n_{1,N/2-1/2} = \bar{F}^n_{2,1/2}. \tag{A.4}$$

Hence, merely ensuring $F_{1,N/2-1/2}(t = t_{n+1}) = F_{2,1/2}(t = t_{n+1})$ would not allow one to keep the global solution in a consistent state despite the use of an iterative method.

Appendix B

Picard and Newton iterations

Consider a root-finding problem $\mathbf{f}(\mathbf{u}) = \mathbf{0}$, where $\mathbf{u}$ is a vector containing the unknowns. For Picard's method, the latter is rewritten as a fixed-point problem $\mathbf{u} = \mathbf{g}(\mathbf{u})$ where $\mathbf{g}(\mathbf{u}) \equiv \mathbf{u} - \beta\mathbf{f}(\mathbf{u})$, with $\beta > 0$ a fixed-point damping parameter, typically less than 1. A fixed-point iteration proceeds according to the algorithm in Table B.1. When $\mathbf{g}(\mathbf{u})$ is a contraction, i.e., if there exists a $\lambda \in (0,1)$ such that

$$\|\mathbf{g}(\mathbf{u}) - \mathbf{g}(\mathbf{v})\| \leq \lambda\|\mathbf{u} - \mathbf{v}\| \tag{B.1}$$

for all $\mathbf{u}, \mathbf{v}$ in a closed set containing the fixed-point solution $\mathbf{u}^*$, then the Picard iteration is guaranteed to converge based on Banach's fixed-point theorem [12]. However, even with a good initial guess, Picard iterations converge slowly, namely q-linearly in the norm [68]

$$\|\mathbf{u}^{(k+1)} - \mathbf{u}_{\text{ex}}\| \leq A\|\mathbf{u}^{(k)} - \mathbf{u}_{\text{ex}}\|, \tag{B.2}$$

where $\mathbf{u}_{\text{ex}}$ is the exact solution, the q-factor A lies in $(0,1)$, and the iteration number k is sufficiently large.

A faster alternative to Picard iterations is Newton's method (see Table B.1), which converges q-quadratically in the norm [68]

$$\|\mathbf{u}^{(k+1)} - \mathbf{u}_{\text{ex}}\| \leq A\|\mathbf{u}^{(k)} - \mathbf{u}_{\text{ex}}\|^2, \tag{B.3}$$

for $A > 0$ and sufficiently large k. Drawbacks of Newton's method include only local convergence (i.e., an initial guess for starting the iterations needs to be sufficiently good), and the cost of computing the full Jacobian $\mathbf{J}$. Although not considered in this paper, the latter can be addressed by using a Jacobian-free Newton-Krylov method which requires only the calculation of Jacobian-vector products and avoids having to explicitly form the Jacobian itself [70].

Table B.1: Algorithms for the Picard (left) and Newton (right) iterations (adapted from [94]). In the Newton method, $\mathbf{J} = \partial\mathbf{f}/\partial\mathbf{u}$ is the Jacobian matrix.

Require: Initial guess $\mathbf{u}^{(0)}$
  $k = 0$
  while not converged **do**
    $\mathbf{u}^{(k+1)} = \mathbf{g}(\mathbf{u}^{(k)})$
    $k = k + 1$
  end while

Require: Initial guess $\mathbf{u}^{(0)}$
  $k = 0$
  while not converged **do**
    $\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} - \mathbf{J}^{-1}(\mathbf{u}^{(k)})\mathbf{f}(\mathbf{u}^{(k)})$
    $k = k + 1$
  end while

In the problem considered here, during the macro-step from t_n to t_{n+1} we have $\mathbf{u}^n = (\bar{\rho}^{\,n}_{1,N/2}, \bar{F}^{n}_{2,1/2})^\top$ and need to solve the root-finding problem $\mathbf{f}(\mathbf{u}^n) = 0$ with $\mathbf{f}(\mathbf{u}^n)$ given by

$$\mathbf{f}(\mathbf{u}^n) = \begin{pmatrix} \bar{\rho}^{\,n,k}_{1,N/2} - \bar{\rho}^{\,n,k}_{2,0} \\ \\ \bar{F}^{n,k}_{1,N/2-1/2} - \bar{F}^{n,k}_{2,1/2} \end{pmatrix}. \tag{B.4}$$

Appendix C

Stability matrices and Jacobian

It follows from (2.14a) and (2.14b) that the time-averaged solution over the n_1 (n_2) micro-steps in the left (right) subdomain for the macro-step going from t_n to t_{n+1} is given by

$$\bar{\boldsymbol{\rho}}_1^{n,k} = \bar{\mathbf{B}}_{\mathrm{L}} \boldsymbol{\rho}_1^{n,0} + \bar{\mathbf{C}}_{\mathrm{L}} \boldsymbol{\rho}_{1,b}^{n,k} \tag{C.1a}$$

$$\bar{\boldsymbol{\rho}}_2^{n,k} = \bar{\mathbf{B}}_{\mathrm{R}} \boldsymbol{\rho}_2^{n,0} + \bar{\mathbf{C}}_{\mathrm{R}} \boldsymbol{\rho}_{2,b}^{n,k} + \frac{\gamma}{n_2} \sum_{m=1}^{n_2} \sum_{j=0}^{m-1} (\mathbf{I} + \mathbf{A}_2)^{m-1-j} \boldsymbol{\eta}^{n,j,k}, \tag{C.1b}$$

where

$$\bar{\mathbf{B}}_{\mathrm{L}} = \frac{1}{n_1} \sum_{l=1}^{n_1} (\mathbf{I} + \mathbf{A}_1)^l, \quad \bar{\mathbf{C}}_{\mathrm{L}} = \frac{1}{n_1} \sum_{l=1}^{n_1} \sum_{j=0}^{l-1} (\mathbf{I} + \mathbf{A}_1)^j \mathbf{T}_1, \tag{C.1c}$$

$$\bar{\mathbf{B}}_{\mathrm{R}} = \frac{1}{n_2} \sum_{m=1}^{n_2} (\mathbf{I} + \mathbf{A}_2)^m, \quad \bar{\mathbf{C}}_{\mathrm{R}} = \frac{1}{n_2} \sum_{m=1}^{n_2} \sum_{j=0}^{m-1} (\mathbf{I} + \mathbf{A}_2)^j \mathbf{T}_2. \tag{C.1d}$$

Defining $\tilde{N} = N/2$ for notational convenience, for Picard's method,

$$\bar{\rho}_{1,\tilde{N}}^{n,k+1} = \bar{\rho}_{2,1}^{n,k} + \frac{\Delta x}{D_2} \bar{F}_{2,1/2}^{n,k}, \qquad \bar{F}_{2,1/2}^{n,k+1} = -\frac{D_1}{\Delta x} (\bar{\rho}_{1,\tilde{N}}^{n,k} - \bar{\rho}_{1,\tilde{N}-1}^{n,k}), \tag{C.2}$$

this gives

$$\bar{\rho}_{1,\tilde{N}}^{n,k+1} = \left[\bar{\mathbf{B}}_{\mathrm{R}}\boldsymbol{\rho}_2^{n,0} + \bar{\mathbf{C}}_{\mathrm{R}}\boldsymbol{\rho}_{2,b}^{n,k} + \frac{\gamma}{n_2}\sum_{m=1}^{n_2}\sum_{j=0}^{m-1}(\mathbf{I}+\mathbf{A}_2)^{m-1-j}\boldsymbol{\eta}^{n,j,k}\right]_1 + \frac{\Delta x}{D_2}\bar{F}_{2,1/2}^{n,k}$$
$$\bar{F}_{2,1/2}^{n,k+1} = -\frac{D_1}{\Delta x}\left(\bar{\rho}_{1,\tilde{N}}^{n,k} - \left[\bar{\mathbf{B}}_{\mathrm{L}}\boldsymbol{\rho}_1^{n,0} + \bar{\mathbf{C}}_{\mathrm{L}}\boldsymbol{\rho}_{1,b}^{n,k}\right]_{\tilde{N}-1}\right). \tag{C.3}$$

Rewriting (2.15a) and (2.16a) as

$$\boldsymbol{\rho}_1^{n,n_1,k} = \mathbf{b}_n^1 + \mathbf{C}_{\mathrm{L}}\boldsymbol{\rho}_{1,b}^{n,k}, \quad \boldsymbol{\rho}_2^{n,n_2,k} = \mathbf{b}_n^2 + \mathbf{C}_{\mathrm{R}}\boldsymbol{\rho}_{2,b}^{n,k} + \gamma\sum_{m=0}^{n_2-1}(\mathbf{I}+\mathbf{A}_2)^{n_2-1-m}\boldsymbol{\eta}^{n,m,k} \tag{C.4}$$

we have, in component form,

$$\begin{aligned}
\rho_{1,p}^{n,n_1,k} &= b_{n,p}^1 + [C_{\mathrm{L}}]_{p,\tilde{N}-1}\sum_{q=1}^{\tilde{N}-1}[\bar{B}_{\mathrm{R}}]_{1,q}\,\rho_{2,q}^{n,0} \qquad \text{(C.5a)}\\
&\quad + [C_{\mathrm{L}}]_{p,\tilde{N}-1}[\bar{C}_{\mathrm{R}}]_{1,1}\bar{\rho}_{1,\tilde{N}}^{n,k-1} + [C_{\mathrm{L}}]_{p,\tilde{N}-1}\frac{\Delta x}{D_2}\bar{F}_{2,1/2}^{n,k-1}\\
&\quad + [C_{\mathrm{L}}]_{p,1}\rho_{\mathrm{L}} + [C_{\mathrm{L}}]_{p,\tilde{N}-1}[\bar{C}_{\mathrm{R}}]_{1,\tilde{N}-1}\rho_{\mathrm{R}}\\
&\quad + [C_{\mathrm{L}}]_{p,\tilde{N}-1}\frac{\gamma}{n_2}\sum_{m=1}^{n_2}\sum_{j=0}^{m-1}\sum_{q=1}^{\tilde{N}-1}[Z_{jm}]_{1,q}\eta_q^{n,j,k-1}\\
\rho_{2,p}^{n,n_2,k} &= b_{n,p}^2 + [C_{\mathrm{R}}]_{p,1}\sum_{q=1}^{\tilde{N}-1}[\bar{B}_{\mathrm{R}}]_{1,q}\,\rho_{2,q}^{n,0} \qquad \text{(C.5b)}\\
&\quad + [C_{\mathrm{R}}]_{p,1}[\bar{C}_{\mathrm{R}}]_{1,1}\bar{\rho}_{1,\tilde{N}}^{n,k-1} + [C_{\mathrm{R}}]_{p,1}\frac{\Delta x}{D_2}\bar{F}_{2,1/2}^{n,k-1}\\
&\quad + \left([C_{\mathrm{R}}]_{p,\tilde{N}-1} + [C_{\mathrm{R}}]_{p,1}[\bar{C}_{\mathrm{R}}]_{1,\tilde{N}-1}\right)\rho_{\mathrm{R}}\\
&\quad + [C_{\mathrm{R}}]_{p,1}\frac{\gamma}{n_2}\sum_{m=1}^{n_2}\sum_{j=0}^{m-1}\sum_{q=1}^{\tilde{N}-1}[Z_{jm}]_{1,q}\eta_q^{n,j,k-1} + \gamma\sum_{j=0}^{n_2-1}\sum_{q=1}^{\tilde{N}-1}[Z_{jn_2}]_{p,q}\eta_q^{n,j,k}
\end{aligned}$$

where $\mathbf{Z}_{jm} \equiv (\mathbf{I}+\mathbf{A}_2)^{m-1-j}$ and $p = 1,\ldots,\tilde{N}-1$. The terms in these two expressions define the matrix $\mathbf{M}$ in (2.18b) and a vector $\mathbf{P}\mathbf{x}^n = (\boldsymbol{c}_1, c_2, c_3, \boldsymbol{c}_4)^\top$ with

$$c_{1,p} = \sum_{q=1}^{\tilde{N}-1} [B_{\mathrm{L}}]_{p,q}\rho_{1,q}^{n,0} + [C_{\mathrm{L}}]_{p,\tilde{N}-1} \sum_{q=1}^{\tilde{N}-1} [\bar{B}_{\mathrm{R}}]_{1,q}\rho_{2,q}^{n,0}, \quad c_2 = \sum_{q=1}^{\tilde{N}-1} [\bar{B}_{\mathrm{R}}]_{1,q}\rho_{2,q}^{n,0}$$
$$c_3 = \frac{D_1}{\Delta x} \sum_{q=1}^{\tilde{N}-1} [\bar{B}_{\mathrm{L}}]_{\tilde{N}-1,q}\rho_{1,q}^{n,0}, \quad c_{4,p} = \sum_{q=1}^{\tilde{N}-1} [B_{\mathrm{R}}]_{p,q}\rho_{2,q}^{n,0} + [C_{\mathrm{R}}]_{p,1} \sum_{q=1}^{\tilde{N}-1} [\bar{B}_{\mathrm{R}}]_{1,q}\rho_{2,q}^{n,0}. \quad \text{(C.6)}$$

Relations (C.6) define the matrix $\mathbf{P}$ in (2.18c).

The nonzero sub-matrices of $\mathbf{M}$ and $\mathbf{P}$ in (2.18b) and (2.18c) are defined by

$$r_p = [C_{\mathrm{L}}]_{p,\tilde{N}-1}[\bar{C}_{\mathrm{R}}]_{1,1}, \quad s_p = \frac{\Delta x}{D_2}[C_{\mathrm{L}}]_{p,\tilde{N}-1}, \quad u = [\bar{C}_{\mathrm{R}}]_{1,1}, \quad v = \frac{\Delta x}{D_2},$$
$$w = -\frac{D_1}{\Delta x}(1 - [\bar{C}_{\mathrm{L}}]_{\tilde{N}-1,\tilde{N}-1}), \quad y_p = [C_{\mathrm{R}}]_{p,1}[\bar{C}_{\mathrm{R}}]_{1,1}, \quad z_p = \frac{\Delta x}{D_2}[C_{\mathrm{R}}]_{p,1} \quad \text{(C.7a)}$$

and

$$S_{p,q} = [C_{\mathrm{L}}]_{p,\tilde{N}-1}[\bar{B}_{\mathrm{R}}]_{1,q}, \quad u_q = [\bar{B}_{\mathrm{R}}]_{1,q}, \quad v_q = \frac{D_1}{\Delta x}[\bar{B}_{\mathrm{L}}]_{\tilde{N}-1,q},$$
$$W_{p,q} = [B_{\mathrm{R}}]_{p,q} + [C_{\mathrm{R}}]_{p,1}[\bar{B}_{\mathrm{R}}]_{1,q}, \quad \text{(C.7b)}$$

respectively. Equations (2.18) and (C.5) also define a vector $\mathbf{d}^{n,k,k+1}$ with components

$$d_p^{n,k,k+1} = \frac{\gamma}{n_2}\begin{cases} [C_{\mathrm{L}}]_{p,\tilde{N}-1}\sum_{m=1}^{n_2}\sum_{j=0}^{m-1}\sum_{q=1}^{\tilde{N}-1}[Z_{jm}]_{1,q}\eta_q^{n,j,k} & p=1,\ldots,\tilde{N}-1 \\ \sum_{m=1}^{n_2}\sum_{j=0}^{m-1}\sum_{q=1}^{\tilde{N}-1}[Z_{jm}]_{1,q}\eta_q^{n,j,k} & p=\tilde{N} \\ 0 & p=\tilde{N}+1 \\ [C_{\mathrm{R}}]_{p,1}\sum_{m=1}^{n_2}\sum_{j=0}^{m-1}\sum_{q=1}^{\tilde{N}-1}[Z_{jm}]_{1,q}\eta_q^{n,j,k} & \\ \quad +n_2\sum_{j=0}^{n_2-1}\sum_{q=1}^{\tilde{N}-1}[Z_{jn_2}]_{p,q}\eta_q^{n,j,k+1} & p=\tilde{N}+2,\ldots,N, \end{cases} \tag{C.8}$$

and a vector $\mathbf{e}$ with components

$$e_p = \begin{cases} [C_{\mathrm{L}}]_{p,1}\rho_{\mathrm{L}} + [C_{\mathrm{L}}]_{p,\tilde{N}-1}[\bar{C}_{\mathrm{R}}]_{1,\tilde{N}-1}\rho_{\mathrm{R}} & p=1,\ldots,\tilde{N}-1 \\ [\bar{C}_{\mathrm{R}}]_{1,\tilde{N}-1}\rho_{\mathrm{R}} & p=\tilde{N} \\ \dfrac{D_1}{\Delta x}[\bar{C}_{\mathrm{L}}]_{\tilde{N}-1,1}\rho_{\mathrm{L}} & p=\tilde{N}+1 \\ \left([C_{\mathrm{R}}]_{p,\tilde{N}-1} + [C_{\mathrm{R}}]_{p,1}[\bar{C}_{\mathrm{R}}]_{1,\tilde{N}-1}\right)\rho_{\mathrm{R}} & p=\tilde{N}+2,\ldots,N. \end{cases} \tag{C.9}$$

Finally, we derive explicit expressions for elements of the Jacobian $\mathbf{J}$ for Newton's coupling. It follows from (2.13c), (C.1a), and (C.1b) that

$$\begin{aligned} g_1 &= \bar{\rho}_{1,\tilde{N}}^{n,k} - \left[\bar{\mathbf{B}}_{\mathrm{R}}\boldsymbol{\rho}_2^{n,0} + \bar{\mathbf{C}}_{\mathrm{R}}\boldsymbol{\rho}_{2,b}^{n,k} + \frac{\gamma}{n_2}\sum_{m=1}^{n_2}\sum_{j=0}^{m-1}(\mathbf{I}+\mathbf{A}_2)^{m-1-j}\boldsymbol{\eta}^{\,n,j,k}\right]_1 - \frac{\Delta x}{D_2}\bar{F}_{2,1/2}^{n,k} \\ g_2 &= -\frac{D_1}{\Delta x}\left(\bar{\rho}_{1,\tilde{N}}^{n,k} - \left[\bar{\mathbf{B}}_{\mathrm{L}}\boldsymbol{\rho}_1^{n,0} + \bar{\mathbf{C}}_{\mathrm{L}}\boldsymbol{\rho}_{1,b}^{n,k}\right]_{\tilde{N}-1}\right) - \bar{F}_{2,1/2}^{n,k}. \end{aligned} \tag{C.10}$$

Combining this with (2.13b) leads to

$$\mathbf{J} = \begin{pmatrix} 1 - [\bar{C}_{\mathrm{R}}]_{1,1} & -\dfrac{\Delta x}{D_2} \\ \\ -\dfrac{D_1}{\Delta x}\left(1 - [\bar{C}_{\mathrm{L}}]_{\tilde{N}-1,\tilde{N}-1}\right) & -1 \end{pmatrix}. \tag{C.11}$$

Appendix D

Conservative versus non-conservative coupling (Chapter 3)

Consider the total mass M inside $[\alpha_1 - \Delta x/2, \alpha_1 + \Delta x/2]$,

$$M(t) = \int_{\alpha_1-\Delta x/2}^{\alpha_1+\Delta x/2} \rho(x,t)\, \mathrm{d}x, \tag{D.1}$$

with ρ the mass concentration. The temporal derivative of (D.1) yields

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \int_{\alpha_1-\Delta x/2}^{\alpha_1+\Delta x/2} \frac{\partial \rho}{\partial t}\mathrm{d}x = \int_{\alpha_1-\Delta x/2}^{\alpha_1+\Delta x/2} \frac{\partial}{\partial x}\left[D\frac{\partial \rho}{\partial x}\right]\mathrm{d}x = F_{1,N_1-1/2} - F_{2,1/2}. \tag{D.2}$$

Here the flux $F = -D(\rho)\partial_x\rho$ obeys Fick's law, and the interfacial fluxes $F_{1,N_1-1/2}$ and $F_{2,1/2}$ constitute the amount of mass leaving the left subdomain per unit time and the amount of mass entering the middle subdomain per unit time, respectively.

Integrating (D.2) between t_n and $t_{n+1} = t_n + \Delta t_{\mathrm{com}}$ yields

$$\frac{\Delta M}{\Delta t_{\mathrm{com}}} = \bar{F}^n_{1,N_1-1/2} - \bar{F}^n_{2,1/2}, \tag{D.3}$$

where ΔM is the change in total mass inside $[\alpha_1 - \Delta x/2, \alpha_1 + \Delta x/2]$ between t_n and t_{n+1}, and $\bar{F}^n_{1,N_1-1/2}$ and $\bar{F}^n_{2,1/2}$ are the Δt_{com}-averaged values of $F_{1,N_1-1/2}$ and $F_{2,1/2}$, respectively. Since any mass leaving the left subdomain should be transported into the middle subdomain and cannot be trapped inside the interface region $[\alpha_1 - \Delta x/2, \alpha_1 + \Delta x/2]$, the total mass inside this region must remain constant. This means that the Δt_{com}-averaged interface fluxes are equal,

$$\bar{F}^n_{1,N_1-1/2} = \bar{F}^n_{2,1/2}. \tag{D.4}$$

Identical reasoning applies to the interface $x = \alpha_2$. Therefore, only enforcing $F_{1,N_1-1/2}(t_{n+1}) = F_{2,1/2}(t_{n+1})$ and $F_{2,N_2-1/2}(t_{n+1}) = F_{3,1/2}(t_{n+1})$ cannot yield a consistent solution over the entire domain, regardless of whether or not the coupling is iterative.

Appendix E

Jacobian-free Newton-Krylov algorithm for Chapter 3

The root-finding problem $\mathbf{f}(\mathbf{u}) = \mathbf{0}$, where $\mathbf{u}$ is an n-dimensional vector containing the unknowns, can be solved iteratively using Newton's method (see Table E.1) which converges q-quadratically in the norm [68]

$$\|\mathbf{u}_{k+1} - \mathbf{u}_{\text{ex}}\| \leq A\|\mathbf{u}_k - \mathbf{u}_{\text{ex}}\|^2, \tag{E.1}$$

where $\mathbf{u}_{\text{ex}}$ is the exact solution, the q-factor $A > 0$, and the iteration number k is sufficiently large. Newton's method converges faster than, e.g., fixed-point iteration, but it is only locally convergent (i.e., requires a "good" initial guess) and requires computing the full Jacobian $\mathbf{J}$.

Table E.1: Algorithm for pure Newton's coupling (adapted from [94]).

Require: Initial guess $\mathbf{u}^{(0)}$
  $k = 0$
  while not converged **do**
    $\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} - \mathbf{J}^{-1}(\mathbf{u}^{(k)})\mathbf{f}(\mathbf{u}^{(k)})$
    $k = k + 1$
  end while

The former issue can be addressed through globalization strategies, while the latter can be overcome by using *inexact* Newton algorithms such as JfNK methods [70]. A JfNK algorithm solves a linear system $\mathbf{J}(\mathbf{u}_k)\delta\mathbf{u}_k = -\mathbf{f}(\mathbf{u}_k)$ at the kth Newton iteration inexactly using an iterative Krylov scheme such as the Generalized Minimal RESidual (GMRES) or BiConjugate Gradient STABilized (BiCGSTAB) method. The Krylov solver only requires the action of the Jacobian in the form of a matrix-vector product $\mathbf{J}(\mathbf{u}_k)\mathbf{v}$, which may be approximated by a first-order accurate finite difference expression [70]

$$\mathbf{J}(\mathbf{u}_k)\mathbf{v} \approx \frac{\mathbf{f}(\mathbf{u}_k + \varepsilon\mathbf{v}) - \mathbf{f}(\mathbf{u}_k)}{\varepsilon}, \tag{E.2}$$

or second-order accurate approximation [70]

$$\mathbf{J}(\mathbf{u}_k)\mathbf{v} \approx \frac{\mathbf{f}(\mathbf{u}_k + \varepsilon\mathbf{v}) - \mathbf{f}(\mathbf{u}_k - \varepsilon\mathbf{v})}{2\varepsilon}. \tag{E.3}$$

Here ε is a small perturbation parameter, which has to be neither too large (resulting in a poor approximation of the derivative) nor too small (leading to a big floating-point roundoff error). We define it as

$$\varepsilon = \begin{cases} \dfrac{1}{n\|\mathbf{v}\|_2} \displaystyle\sum_{i=1}^{n} \sqrt{\varepsilon}(1 + |u_i|) & \text{if } \|\mathbf{v}\|_2 > \varepsilon \\ \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \sqrt{\varepsilon}(1 + |u_i|) & \text{if } \|\mathbf{v}\|_2 \leq \varepsilon, \end{cases} \tag{E.4}$$

where n is the size of $\mathbf{u}_k$ ($n = 4$ in our case) and $\varepsilon = 2.2204 \cdot 10^{-16}$ (machine roundoff for 64-bit double precision). In our simulations we use formula (E.3).

JfNK methods provide Newton-like convergence without the cost of forming or storing the true Jacobian. Yet, their error stems from both the inexact convergence of the iterative linear solves and, more importantly, from approximating the action of the Jacobian. The latter error is directly related to the selection of a value for ε. In addition,

unless the condition number of $\mathbf{J}$ is small or its eigenvalues are clustered together, preconditioning is needed to converge the Krylov solver with reasonable effort. This causes the matrix-free appeal of (E.2) or (E.3) to yield, to some extent, to the construction and use of a preconditioning matrix $\mathbf{P}_k$ (hence, we use the term "Jacobian-free" and not "matrix-free"). Right preconditioning, which does not change the norm of the linear residual, is often used in a Newton-Krylov method. The Newton equation with right preconditioning is

$$(\mathbf{J}(\mathbf{u}_k)\mathbf{P}_k^{-1})(\mathbf{P}_k\ \delta\mathbf{u}_k) = -\mathbf{f}(\mathbf{u}_k), \tag{E.5}$$

where $\mathbf{P}_k$ is the preconditioning matrix, which should be easy to evaluate, while retaining as much of the properties of the Jacobian as possible. In practice, only the matrix elements that are needed for the action of $\mathbf{P}_k^{-1}$ are formed, which can be done in a variety of ways. Using right preconditioning, the kth Newton iteration is realized via a two-step process:

1. Solve $(\mathbf{J}(\mathbf{u}_k)\mathbf{P}_k^{-1})\mathbf{w} = -\mathbf{f}(\mathbf{u}_k)$ for $\mathbf{w}$.

2. Solve $\delta\mathbf{u}_k = \mathbf{P}_k^{-1}\mathbf{w}$ for $\delta\mathbf{u}_k$.

To execute step 1, right-preconditioned versions of (E.2) and (E.3) are defined as

$$\mathbf{J}(\mathbf{u}_k)\mathbf{P}_k^{-1}\mathbf{w} \approx \frac{\mathbf{f}(\mathbf{u}_k + \varepsilon\mathbf{P}_k^{-1}\mathbf{w}) - \mathbf{f}(\mathbf{u}_k)}{\varepsilon} \tag{E.6}$$

and

$$\mathbf{J}(\mathbf{u}_k)\mathbf{P}_k^{-1}\mathbf{w} \approx \frac{\mathbf{f}(\mathbf{u}_k + \varepsilon\mathbf{P}_k^{-1}\mathbf{w}) - \mathbf{f}(\mathbf{u}_k - \varepsilon\mathbf{P}_k^{-1}\mathbf{w})}{2\varepsilon}, \tag{E.7}$$

respectively. Next, $\mathbf{P}_k\mathbf{y} = \mathbf{w}$ is solved (inexactly) for $\mathbf{y}$, the matrix-free product $\mathbf{J}(\mathbf{u}_k)\mathbf{y} \approx [\mathbf{f}(\mathbf{u}_k + \varepsilon\mathbf{y}) - \mathbf{f}(\mathbf{u}_k)]/\varepsilon$ or $\mathbf{J}(\mathbf{u}_k)\mathbf{y} \approx [\mathbf{f}(\mathbf{u}_k + \varepsilon\mathbf{y}) - \mathbf{f}(\mathbf{u}_k - \varepsilon\mathbf{y})]/(2\varepsilon)$ is computed, and the

Krylov iteration commences. In our simulations, the Krylov solver converged sufficiently fast without the use of a preconditioner, and hence we decided against using such a matrix.

In the problem under consideration, during the macro-step from t_n to t_{n+1} we have $\mathbf{u} \equiv \mathbf{u}^n = (\bar{\rho}^n_{1,N_1}, \bar{F}^n_{2,1/2}, \bar{\rho}^n_{2,N_2}, \bar{F}^n_{3,1/2})^\top$, and employ JfNK to solve the root-finding problem $\mathbf{f}(\mathbf{u}^n) = 0$ with $\mathbf{f}(\mathbf{u}^n)$ given by

$$\mathbf{f}(\mathbf{u}^n) = \begin{pmatrix} \bar{\rho}^n_{1,N_1} - \bar{\rho}^n_{2,0} \\ \bar{F}^n_{1,N_1-1/2} - \bar{F}^n_{2,1/2} \\ \bar{\rho}^n_{2,N_2} - \bar{\rho}^n_{3,0} \\ \bar{F}^n_{2,N_2-1/2} - \bar{F}^n_{3,1/2} \end{pmatrix}. \tag{E.8}$$

Appendix F

Proofs of the stability lemmas

F.0.1 Proof of Lemma 3.5.1

After n_1 micro-steps, the left subdomain solver (3.10a) yields

$$\boldsymbol{\rho}_1^{n,n_1,k} = \mathbf{B}_{\mathrm{L}}\boldsymbol{\rho}_1^{n,0} + \mathbf{C}_{\mathrm{L}}\boldsymbol{\rho}_{1,\mathrm{b}}^{n,k} + \sum_{l=0}^{n_1-1} (\mathbf{I}_{N_1-1} + \mathbf{A}_{1,n})^{n_1-1-l}\, \mathbf{T}_{1,n}\boldsymbol{\eta}^{n,l}, \tag{F.1}$$

where $\boldsymbol{\rho}_1^{n,0} \equiv \boldsymbol{\rho}_1(t=t_n)$, $\mathbf{B}_{\mathrm{L}} = (\mathbf{I}_{N_1-1} + \mathbf{A}_{1,n})^{n_1}$ and $\mathbf{C}_{\mathrm{L}} = \sum_{l=0}^{n_1-1}(\mathbf{I}_{N_1-1} + \mathbf{A}_{1,n})^{l}\,\mathbf{T}_{1,n}$. After n_2 micro-steps, the middle subdomain solver (3.10b) gives

$$\boldsymbol{\rho}_2^{n,n_2,k} = \mathbf{B}_{\mathrm{M}}\boldsymbol{\rho}_2^{n,0} + \mathbf{C}_{\mathrm{M}}\boldsymbol{\rho}_{2,\mathrm{b}}^{n,k}, \tag{F.2}$$

where $\boldsymbol{\rho}_2^{n,0} \equiv \boldsymbol{\rho}_2(t=t_n)$, $\mathbf{B}_{\mathrm{M}} = (\mathbf{I}_{N_2-1} + \mathbf{A}_2)^{n_2}$ and $\mathbf{C}_{\mathrm{M}} = \sum_{m=0}^{n_2-1}(\mathbf{I}_{N_2-1} + \mathbf{A}_2)^{m}\,\mathbf{T}_2$. After n_3 micro-steps, the right subdomain solver (3.10c) gives

$$\boldsymbol{\rho}_3^{n,n_3,k} = \mathbf{B}_{\mathrm{R}}\boldsymbol{\rho}_3^{n,0} + \mathbf{C}_{\mathrm{R}}\boldsymbol{\rho}_{3,\mathrm{b}}^{n,k}, \tag{F.3}$$

where $\boldsymbol{\rho}_3^{n,0} \equiv \boldsymbol{\rho}_3(t=t_n)$, $\mathbf{B}_{\mathrm{R}} = (\mathbf{I}_{N_3-1} + \mathbf{A}_{3,n})^{n_3}$ and $\mathbf{C}_{\mathrm{R}} = \sum_{q=0}^{n_3-1}(\mathbf{I}_{N_3-1} + \mathbf{A}_{3,n})^{q}\,\mathbf{T}_{3,n}$. The time-averaged solutions over the n_1, n_2 and n_3 micro-steps in the left, middle and

right subdomains are given by

$$\bar{\boldsymbol{\rho}}_1^{n,k} = \bar{\mathbf{B}}_{\mathrm{L}} \boldsymbol{\rho}_1^{n,0} + \bar{\mathbf{C}}_{\mathrm{L}} \boldsymbol{\rho}_{1,b}^{n,k} + \frac{1}{n_1} \sum_{l=1}^{n_1} \sum_{j=0}^{l-1} (\mathbf{I}_{N_1-1} + \mathbf{A}_{1,n})^{l-1-j} \mathbf{T}_{1,n}\, \boldsymbol{\eta}^{n,j}, \tag{F.4a}$$

$$\bar{\boldsymbol{\rho}}_2^{n,k} = \bar{\mathbf{B}}_{\mathrm{M}} \boldsymbol{\rho}_2^{n,0} + \bar{\mathbf{C}}_{\mathrm{M}} \boldsymbol{\rho}_{2,b}^{n,k}, \qquad \bar{\boldsymbol{\rho}}_3^{n,k} = \bar{\mathbf{B}}_{\mathrm{R}} \boldsymbol{\rho}_3^{n,0} + \bar{\mathbf{C}}_{\mathrm{R}} \boldsymbol{\rho}_{3,b}^{n,k}, \tag{F.4b}$$

where

$$\bar{\mathbf{B}}_{\mathrm{L}} = \frac{1}{n_1} \sum_{l=1}^{n_1} (\mathbf{I}_{N_1-1} + \mathbf{A}_{1,n})^l, \quad \bar{\mathbf{C}}_{\mathrm{L}} = \frac{1}{n_1} \sum_{l=1}^{n_1} \sum_{j=0}^{l-1} (\mathbf{I}_{N_1-1} + \mathbf{A}_{1,n})^j \mathbf{T}_{1,n}, \tag{F.5}$$

$$\bar{\mathbf{B}}_{\mathrm{M}} = \frac{1}{n_2} \sum_{m=1}^{n_2} (\mathbf{I}_{N_2-1} + \mathbf{A}_2)^m, \quad \bar{\mathbf{C}}_{\mathrm{M}} = \frac{1}{n_2} \sum_{m=1}^{n_2} \sum_{j=0}^{m-1} (\mathbf{I}_{N_2-1} + \mathbf{A}_2)^j \mathbf{T}_2, \tag{F.6}$$

$$\bar{\mathbf{B}}_{\mathrm{R}} = \frac{1}{n_3} \sum_{q=1}^{n_3} (\mathbf{I}_{N_3-1} + \mathbf{A}_{3,n})^q, \quad \bar{\mathbf{C}}_{\mathrm{R}} = \frac{1}{n_3} \sum_{q=1}^{n_3} \sum_{j=0}^{q-1} (\mathbf{I}_{N_3-1} + \mathbf{A}_{3,n})^j \mathbf{T}_{3,n}. \tag{F.7}$$

The definitions of $A_{i,n}$ and $T_{i,n}$ $(i = 1,3)$ are given by

$$A_{i,n} = \frac{\Delta t_i}{\Delta x^2} Z_{i,n}, \qquad T_{i,n} = \frac{\Delta t_i}{\Delta x^2} W_{i,n}, \qquad i = 1,3, \tag{F.8a}$$

with

$$[Z_{i,n}]_{r,s} = \begin{cases} -D_i(\rho_{i,r+1}^{n,0};\rho_{i,r}^{n,0}) - D_i(\rho_{i,r}^{n,0};\rho_{i,r-1}^{n,0}) & \text{for } s = r = 1,\ldots,N_i-1 \\ D_i(\rho_{i,r}^{n,0};\rho_{i,r-1}^{n,0}) & \text{for } s = r-1 = 1,\ldots,N_i-2 \\ D_i(\rho_{i,r+1}^{n,0};\rho_{i,r}^{n,0}) & \text{for } s = r+1 = 2,\ldots,N_i-1 \\ 0 & \text{otherwise,} \end{cases} \tag{F.8b}$$

$$[W_{i,n}]_{r,s} = \begin{cases} 1 & \text{for } s = r = 2,\ldots,N_i-2 \\ D_i(\rho_{i,r}^{n,0};\rho_{i,r-1}^{n,0}) & s = r = 1 \\ D_i(\rho_{i,r+1}^{n,0};\rho_{i,r}^{n,0}) & s = r = N_i-1, \\ 0 & \text{otherwise,} \end{cases} \tag{F.8c}$$

where $D_i(a;b)$ stands for D_i evaluated at $(a+b)/2$ (for $i = 1$ or 3), $\rho_{1,0}^{n,0} = \rho_0^n$, $\rho_{1,N_1}^{n,0} = \rho(\alpha_1,t_n)$, $\rho_{3,0}^{n,0} = \rho(\alpha_2,t_n)$ and $\rho_{3,N_3}^{n,0} = \rho_L$. Combining (F.1)–(F.8), the expressions for A_2 and T_2 in Section 2.5, and Newton's iteration[1] (3.9), leads to a recurrence relation

$$\mathbf{x}^{n,k+1} = \mathbf{M}_n(\eta^{n,0},\mathbf{x}^n)\,\mathbf{x}^{n,k} + \mathbf{P}_n(\eta^{n,0},\mathbf{x}^n)\,\mathbf{x}^n + \mathbf{d}^n(\eta^{n,l},\mathbf{x}^n), \tag{F.9}$$

where $\mathbf{d}^n$ is a vector of size $N+1$ and $l \in \{0,1,\ldots,n_1-1\}$. Taking (F.9) to convergence, and ensemble-averaging the result, yields

$$\langle \mathbf{x}^{n+1} \rangle = \langle (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{P}_n\mathbf{x}^n \rangle + \langle (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{d}^n \rangle. \tag{F.10}$$

□

[1] Since the problem is linearized around t_n for the macro-step from t_n to t_{n+1}, we consider the pure Newton method.

F.0.2 Proof of Lemma 3.5.2

We employ the Reynolds decomposition $\mathbf{y}^n = \langle \mathbf{y}^n \rangle + \tilde{\mathbf{y}}^n$, where $\tilde{\mathbf{y}}^n$ is the zero-mean fluctuation of $\mathbf{y}^n$ about the mean $\langle \mathbf{y}^n \rangle$, and use Taylor's theorem to expand $(\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{P}_n$ around $\langle \mathbf{y}^n \rangle$. Let $\mathbf{Q}_n \equiv (\mathbf{I}_{N+1} - \mathbf{M}_n)^{-1}\mathbf{P}_n$. The ith component of $\mathbf{Q}_n(\mathbf{y}^n)\ \mathbf{x}^n$ is given by

$$[\mathbf{Q}_n(\mathbf{y}^n)\ \mathbf{x}^n]_i = \sum_{j=1}^{N+1} [Q_n(\mathbf{y}^n)]_{i,j}\ x_j^n, \tag{F.11}$$

with $i = 1, \ldots, N+1$. Taking the ensemble average of (F.11) and expanding $[\mathbf{Q}_n(\mathbf{y}^n)]_{i,j}$ using Taylor's theorem yields

$$\langle [\mathbf{Q}_n(\mathbf{y}^n)\ \mathbf{x}^n]_i \rangle = \sum_{j=1}^{N+1} \left\{ [Q_n(\langle \mathbf{y}^n \rangle)]_{i,j} \langle x_j^n \rangle + \langle (\tilde{\mathbf{y}}^n)^\top \nabla [Q_n(\langle \mathbf{y}^n \rangle)]_{i,j}\ x_j^n \rangle + \langle R_1\ x_j^n \rangle \right\} \tag{F.12}$$

where $R_1(\tilde{\mathbf{y}}^n)$ is the remainder in Taylor's formula. Taking the modulus of (F.12) and using the triangle inequality leads to

$$\left| \langle [\mathbf{Q}_n(\mathbf{y}^n)\ \mathbf{x}^n]_i \rangle \right| \leq \left| \sum_{j=1}^{N+1} \{ [Q_n(\langle \mathbf{y}^n \rangle)]_{i,j}\ \langle x_j^n \rangle \right| + V_{n,i}\ , \qquad i = 1, \ldots, N+1, \tag{F.13}$$

where $V_{n,i}$ is given by

$$V_{n,i} = \sum_{j=1}^{N+1} \left\{ |\langle (\tilde{\mathbf{y}}^n)^\top \nabla [Q_n(\langle \mathbf{y}^n \rangle)]_{i,j}\ x_j^n \rangle| + |\langle R_1(\tilde{\mathbf{y}}^n)\ x_j^n \rangle| \right\}. \tag{F.14}$$

Summing both sides of (F.13) over all values of i, we find that

$$\| \langle (\mathbf{Q}_n(\mathbf{y}^n)\ \mathbf{x}^n \rangle \|_1 \leq \| \mathbf{Q}_n(\langle \mathbf{y}^n \rangle) \langle \mathbf{x}^n \rangle \|_1 + V_n \tag{F.15}$$

with $V_n \equiv \sum_{i=1}^{N+1} V_{n,i}$ and $\|\cdot\|_1$ the l^1-norm. Using the inequality $\|\mathbf{Av}\|_1 \leq \|\mathbf{A}\|_1\|\mathbf{v}\|_1$, with $\mathbf{A} \equiv \mathbf{Q}_n(\langle \mathbf{y}^n \rangle)$ and $\mathbf{v} \equiv \langle \mathbf{x}^n \rangle$, yields (3.14). □

F.0.3 Proof of Lemma 3.5.3

A finite bound on $\|\langle(\mathbf{Q}_n(\mathbf{y}^n)\ \mathbf{x}^n\rangle\|_1$ requires the existence of a finite V_n. To find the conditions under which this is guaranteed, we proceed as follows.

Fact 1. *If $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ are two real-valued functions, then*

$$\left|\int_S fg\mathrm{d}\mu\right| \leq \int_S |fg|\mathrm{d}\mu = \|fg\|_1 \leq \|f\|_2\|g\|_2. \tag{F.16}$$

with respect to a measure μ.

The left inequality follows from the monotonicity of integral, while the second one constitutes a specific case of Hölder's inequality.

Fact 2. *For any random variable X with a probability density function (PDF) $f_X(x)$, if $Y = g(X)$ with PDF $f_Y(y)$ and g is monotonic then*

$$f_Y(y) = \left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| f_X(x(y)). \tag{F.17}$$

Let $f_{\mathbf{y}^n}$ denote the PDF of $\mathbf{y}^n$ and $g_{i,j}(\mathbf{y}^n) \equiv (\tilde{\mathbf{y}}^n)^\top \nabla[Q_n(\langle \mathbf{y}^n \rangle)]_{i,j}\ x_j^n$. Then, for each $i, j = 1, \ldots, N+1$, $|\langle g_{i,j}(\mathbf{y}^n)\rangle|$ is bounded if

$$\int_S f_{\mathbf{y}^n}^2(\mathbf{s})\mathrm{d}\mathbf{s} < \infty, \qquad \int_S |g_{i,j}(\mathbf{s})|^2\mathrm{d}\mathbf{s} < \infty. \tag{F.18}$$

where S is the support of $f_{\mathbf{y}^n}$. Let h_1 denote a function, which maps the random boundary noise η onto $\mathbf{y}^n$. Then the monotonicity of h_1 (which is expected to be the case) implies,

according to Fact 2,

$$f_{\mathbf{y}^n}(\mathbf{s}) = \|\nabla_{\mathbf{s}} v\|_2 \; f_{\eta}[v(\mathbf{s})]. \tag{F.19}$$

Hence,

$$\int_S f^2_{\mathbf{y}^n}(\mathbf{s})\mathrm{d}\mathbf{s} = \int_S \|\nabla_{\mathbf{s}} v\|_2^2 \; f^2_{\eta}[v(\mathbf{s})] \; \mathrm{d}\mathbf{s}. \tag{F.20}$$

According to Fact 1, $f(\mathbf{y}^n)$ is in L^2 if

$$\int_S \|\nabla_{\mathbf{s}} v\|_2^4 \mathrm{d}\mathbf{s} < \infty, \qquad \int_S f^4_{\eta}[v(\mathbf{s})] \; \mathrm{d}\mathbf{s} < \infty. \tag{F.21}$$

Since f_η is in L^4, then provided that $\nabla_{\mathbf{y}^n}\eta$ is in L^4, the PDF $f_{\mathbf{y}^n}$ is in L^2 and the first inequality in (F.18) holds. The second inequality in (F.18) is satisfied if $g_{i,j}(\mathbf{y}^n)$ is in L^2.

Lemma. *If $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^{k+1} on an open convex set S and $|\partial^\alpha f(\mathbf{x})| \leq W$ for $\mathbf{x} \in S$ with $|\alpha| = k+1$, then a bound for the remainder $R_{\mathbf{a},k}$ of Taylor's theorem for f about $\mathbf{a} \equiv \mathbf{x} - \mathbf{h} \in S$ is given by*

$$|R_{\mathbf{a},k}(\mathbf{h})| \leq \frac{W}{(k+1)!}\|\mathbf{h}\|_1^{k+1}, \tag{F.22}$$

where α is the multi-index $(\alpha_1, \ldots, \alpha_n)$ with $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and $\partial^\alpha f(\mathbf{x}) \equiv \partial^{|\alpha|} f/(\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n})$.

Proof. It follows from either the Lagrange or integral form of $R_{\mathbf{a},k}(\mathbf{h})$ that

$$|R_{\mathbf{a},k}(\mathbf{h})| \leq W \sum_{|\alpha|=k+1} \frac{|\mathbf{h}^\alpha|}{\alpha!}, \tag{F.23}$$

where $\mathbf{h}^\alpha \equiv h_1^{\alpha_1} \ldots h_n^{\alpha_n}$. According to the multinomial theorem, $\sum_{|\alpha|=k+1} |\mathbf{h}^\alpha| \; /\alpha! = \|\mathbf{h}\|_1^{k+1}/(k+1)!$. □

If $[Q_n(\mathbf{y}^n)]_{i,j}$ is of class C^2 on a open convex set containing $\mathbf{y}^n$ and $\langle \mathbf{y}^n \rangle$, and $|\partial^\alpha [Q_n(\mathbf{y}^n)]_{i,j}| \leq W$ with $|\alpha| = 2$, then the above Lemma yields

$$|R_1(\tilde{\mathbf{y}}^n)| \leq \frac{W}{2} \|\tilde{\mathbf{y}}^n\|_1^2, \tag{F.24}$$

and $|\langle R_1(\tilde{\mathbf{y}}^n)\ x_j^n \rangle|$ in (F.14) is bounded if $\|\tilde{\mathbf{y}}^n\|_1^2 x_j^n$ is in L^2 and $\nabla_{\mathbf{y}^n} \eta$ is in L^4. □

Appendix G

Derivation of a flux-limited, equilibrium radiation-diffusion equation from the radiative transfer and material temperature equations

Treating photons as particles, and ignoring interactions between the photons themselves, transport of radiation at the fundamental level may be described by the Boltzmann equation [98]

$$\frac{1}{c}\frac{\partial f}{\partial t} + \mathbf{\Omega}\cdot\nabla f = \mathbf{C}(f), \tag{G.1}$$

where $f(\mathbf{x},t,\mathbf{\Omega},\nu)$ is the photon phase space density; $\mathbf{\Omega}$ is the unit solid angle vector indicating the direction of photon travel; $\mathbf{C}$ is the collision operator; ν is the photon frequency (each photon has energy $h\nu$ with h the Planck constant); and c is the speed of

light. Multiplying (G.1) by $h\nu$ and expanding $\mathbf{C}$ yields the radiative transfer equation [85]

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \mathbf{\Omega}\cdot\nabla I_\nu + \kappa_{t,\nu} I_\nu = J_\nu + \frac{\kappa_{s,\nu}}{4\pi}\int_{4\pi} I_\nu(\mathbf{\Omega}')\; p_\nu(\mathbf{\Omega}'\to\mathbf{\Omega})\; \mathrm{d}\Omega'. \tag{G.2}$$

Here $I_\nu(\mathbf{x},t,\mathbf{\Omega},\nu)$ is the spectral radiation intensity, defined as the radiation energy per unit area, per unit time, per unit solid angle about the photon propagation direction, and per unit interval of photon frequency; $\kappa_{t,\nu}(\mathbf{x}) \equiv \kappa_{s,\nu}(\mathbf{x}) + \kappa_{a,\nu}(\mathbf{x})$ is the spectral extinction coefficient, consisting of the spectral scattering coefficient $\kappa_{s,\nu}$ and spectral absorption coefficient $\kappa_{a,\nu}$; J_ν is the radiation emitted by a unit volume of the medium per unit time, per unit solid angle about the photon propagation direction, and per unit interval of photon frequency; and $p_\nu(\mathbf{\Omega}'\to\mathbf{\Omega})$ is the scattering indicatrix normalized such that

$$\frac{1}{4\pi}\int_{4\pi} p_\nu(\mathbf{\Omega}'\to\mathbf{\Omega})\; d\Omega = 1. \tag{G.3}$$

Equation (G.2) essentially conveys that as a beam of radiation travels through a medium, it loses energy through absorption of radiation by the medium, gains energy through emission of radiation by the medium, and redistributes energy through scattering. Under the assumptions of local thermal equilibrium and isotropic, elastic scattering, (G.2) may be approximated by [85]

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \mathbf{\Omega}\cdot\nabla I_\nu = -\kappa_{t,\nu} I_\nu + \kappa_{a,\nu} I_{\nu B}(T_m) + \frac{\kappa_{s,\nu}}{4\pi}\int_{4\pi} I_\nu\; \mathrm{d}\Omega', \tag{G.4}$$

where $I_{\nu B}$ is the Planck function and T_m is the temperature of the medium through which the radiation propagates. Equation (G.4) may be simplified by calculating moments of the spectral radiation intensity. The zeroth moment (spectral radiation energy density E_ν)

and first moment (spectral radiation energy flux $\mathbf{F}_\nu$) are given by [85]

$$E_\nu(\mathbf{x},t,\nu) = \int_{4\pi} I_\nu(\mathbf{x},t,\mathbf{\Omega},\nu)\,\mathrm{d}\Omega, \quad \mathbf{F}_\nu(\mathbf{x},t,\nu) = \int_{4\pi} \mathbf{\Omega}\, I_\nu(\mathbf{x},t,\mathbf{\Omega},\nu)\,\mathrm{d}\Omega. \tag{G.5}$$

Integrating (G.4) over solid angle yields

$$\frac{1}{c}\frac{\partial E_\nu}{\partial t} + \nabla\cdot\mathbf{F}_\nu = \kappa_{a,\nu}(4\pi I_{\nu B}(T_m) - E_\nu). \tag{G.6}$$

Since each moment of I_ν involves the next higher moment, we need to make an approximation for the radiation pressure in order to close the otherwise infinite set of moment equations. This may be done through the *ansatz* that the radiation field is isotropic. Together with the assumption that the flux $\mathbf{F}_\nu$ varies slowly with time compared to the spatial gradient in E_ν, i.e., $(1/c)\partial\mathbf{F}_\nu/\partial t \ll (1/3)\nabla E_\nu$, we obtain, after averaging over all photon frequencies (i.e., a gray approximation), the following diffusion equation for the frequency-averaged radiation energy density

$$\frac{\partial E}{\partial t} = \nabla\cdot(D\nabla E) + c\;\kappa_\mathrm{P}(aT_m^4 - E). \tag{G.7}$$

Here a is the radiation constant; the diffusion coefficient $D = c/(3\kappa_\mathrm{R})$ with $\kappa_\mathrm{R} = \rho\sigma_\mathrm{R}$, where ρ is the material's density and σ_R is the Rosseland mean opacity; and $\kappa_\mathrm{P} = \rho\sigma_\mathrm{P}$ with σ_P the Planck mean opacity. Equation (G.7) needs to be supplemented by an equation for the material energy density u_m [85]

$$\frac{\partial u_m}{\partial t} = \int_0^\infty hc\;\kappa_{a,\nu}(E_\nu - 4\pi I_{\nu B}(T_m))\,\mathrm{d}\nu = c\;\kappa_P(E - aT_m^4). \tag{G.8}$$

In order to prevent transport of energy at speeds faster than light, we need to introduce a flux limiter λ in the expression for D. We choose the form by LeBlanc and Wilson [72]

given by $\lambda(R) = 1/(3+R)$, where $R = |\nabla E|/(\kappa_R E)$. This yields $D = c\lambda/\kappa_R$, and transforms (G.7) into

$$\frac{\partial E}{\partial t} = \nabla \cdot \left(\frac{c}{3\kappa_R + |\nabla E|/E} \nabla E \right) + c\ \kappa_P\ (aT_m^4 - E), \tag{G.9}$$

As in [99], we assume that $\kappa_R \propto Z^3 T_m^{-3}$ where Z is the material's effective atomic number. If we make the additional approximation that the radiation field is in thermal equilibrium with the material, we have $E = aT_m^4$ (i.e., the radiation has a Planckian distribution at the material temperature) and the model reduces to a single equation for the radiation energy density

$$\frac{\partial E}{\partial t} = \nabla \cdot \left(\frac{cE}{\gamma E/D_Z + |\nabla E|} \nabla E \right), \tag{G.10}$$

where $D_Z \equiv Z^{-3} E^{3/4}$ and γ combines a number of physical constants and proportionality factors and ensures dimensional correctness.

Appendix H

Derivation of the total energy deposition in the tumor

For a particular realization of the vector $\boldsymbol{\xi}$ containing the random parameters, the energy absorption per unit time and per unit surface area at position $\mathbf{x}$ and time t, $\dot{E}_{\text{abs}}$, is given by

$$\dot{E}_{\text{abs}}(\mathbf{x},t) = \kappa_{\text{R}}(\mathbf{x},t)\; cE(\mathbf{x},t), \tag{H.1}$$

with κ_{R} given by [99]

$$\kappa_{\text{R}} = \alpha Z^3 T^{-3} = \alpha\, a^{3/4} Z^3 E^{-3/4}, \tag{H.2}$$

where the proportionality constant α is assumed to be independent of the material properties. The last equality in (H.2) is based on the *ansatz* that the radiation field is in thermal equilibrium with the medium through which it propagates. From (G.10), (H.1)

and (H.2), it follows that

$$\dot{E}_{\text{abs}} = \frac{1}{3}\gamma\, cZ^3 E^{1/4}. \tag{H.3}$$

From (H.3), we conclude that energy absorption increases strongly with the material's atomic number Z, and more weakly with the energy density of the radiation field. A dimensionless form of (H.3) is given by

$$\dot{\tilde{E}}_{\text{abs}} = \frac{w}{E_0 c Z_1^4}\dot{E}_{\text{abs}} = \frac{\gamma w}{3E_0^{3/4}}\tilde{Z}^3\tilde{E}^{1/4} = \frac{1}{3}\tilde{Z}^3\tilde{E}^{1/4}. \tag{H.4}$$

where $E_0 = (\gamma w)^{4/3}$, $\tilde{Z}$ and $\tilde{E}$ are defined in Section 4.3. To obtain the energy absorption $E_{\text{abs}}(i,j,k)$ in $\mathbf{\Delta}x_{j,k} \equiv [x_{1,j} - \Delta x_{1,j}/2, x_{1,j} + \Delta x_{1,j}/2] \times [x_{2,k} - \Delta x_{2,k}/2, x_{2,k} + \Delta x_{2,k}/2]$ during the time step Δt from t_i to $t_i + \Delta t$ ($i = 0, \ldots, I-1$ with $I = T/\Delta t$, and $t_0 = 0$), where $\Delta x_{1,j}$ and $\Delta x_{2,k}$ ($j,k = 1, \ldots, N$) are the dimensions in spatial directions 1 and 2, respectively, of the grid cell centered around $(x_{1,j}, x_{2,k})^\top$, we multiply $\dot{E}_{\text{abs}}(i,j,k)$ with $\Delta t \Delta x_{1,j} \Delta x_{2,k}$. In analogy with (H.4), we define a corresponding dimensionless quantity

$$\tilde{E}_{\text{abs}}(i,j,k) = \frac{1}{E_0 Z_1^4 w^2}\dot{E}_{\text{abs}}(i,j,k)\ \Delta t \Delta x_{1,j} \Delta x_{2,k} = \frac{1}{3}\tilde{Z}_{j,k}^3 (\tilde{E}_{j,k}^{i+1})^{1/4} \Delta\tilde{t}\Delta\tilde{x}_{1,j}\Delta\tilde{x}_{2,k}, \tag{H.5}$$

with $\Delta\tilde{t} = c\Delta t/w$, $\Delta\tilde{x}_{1,j} = \Delta x_{1,j}/w$ and $\Delta\tilde{x}_{2,k} = \Delta x_{2,k}/w$. Here $\tilde{Z}_{j,k}$ is the value of $\tilde{Z}$ in $\mathbf{\Delta}\tilde{x}_{j,k}$ (we assume that $\tilde{Z}$ is constant over a finite volume cell) and $\tilde{E}_{j,k}^{i+1}$ approximates the average value of the dimensionless radiation energy density over $\mathbf{\Delta}\tilde{x}_{j,k}$ during the dimensionless time step from $\tilde{t}_i$ to $\tilde{t}_i + \Delta\tilde{t}$ (we evaluate it at $\tilde{t}_i + \Delta\tilde{t}$). The dimensionless total energy absorption $\tilde{E}_{\text{abs,tot}}$ by the medium in the entire region $\tilde{\mathcal{T}}$ over time $\tilde{T} = I\Delta\tilde{t}$

is then given by

$$\tilde{E}_{\text{abs,tot}} = \sum_{i=1}^{I} \sum_{j} \sum_{k} \tilde{E}_{\text{abs}}(i,j,k), \tag{H.6}$$

where we sum over the indices j and k corresponding to the grid cells within $\tilde{\mathcal{T}}$. Together with (H.5), and omitting $\tilde{}$, this yields

$$E_{\text{abs,tot}} = \frac{\Delta t}{3} \sum_{i=1}^{I} \sum_{j} \sum_{k} Z_{j,k}^{3} (E_{j,k}^{i+1})^{1/4} \Delta x_{1,j} \Delta x_{2,k}. \tag{H.7}$$

Appendix I

Jacobian-free Newton-Krylov algorithm for Chapter 4

Despite its quadratic convergence rate, a standard Newton method requires computation of the full Jacobian $\mathbf{J}$. For our radiation-diffusion problem, the derivatives in $\mathbf{J}$ cannot be obtained analytically and instead would need to be approximated numerically (e.g., using Fréchet derivatives). Rather than pursuing this approach, which is prone to errors and also time-consuming, we solve the linear system $\mathbf{J}(\mathbf{E}_k^{n+1})\delta\mathbf{E}_{k+1} = -\mathbf{f}(\mathbf{E}_k^{n+1})$, where $J_{i,j} = \partial f_i/\partial E_j$, at the kth Newton iteration inexactly using the iterative Krylov algorithm *Generalized Minimum RESidual (GMRES)*. (Here we have assumed that the two-dimensional matrix $\mathbf{E}_k^{n+1}$ has been reshaped into a one-dimensional array, and that the converged solution $\mathbf{E}_K^{n+1}$ is reshaped back into a two-dimensional matrix.) To implement this method we only need to represent the Jacobian-vector product $\mathbf{Jv}$, with $\mathbf{v}$ the Krylov vector, rather than explicitly calculate the Jacobian matrix elements. We may approximate $\mathbf{Jv}$ by finite differences with first-order accuracy [70]

$$\mathbf{J}(\mathbf{E}_k^{n+1})\mathbf{v} \approx \frac{\mathbf{f}(\mathbf{E}_k^{n+1} + \varepsilon\mathbf{v}) - \mathbf{f}(\mathbf{E}_k^{n+1})}{\varepsilon}, \tag{I.1}$$

or second-order accuracy [70]

$$\mathbf{J}(\mathbf{E}_k^{n+1})\mathbf{v} \approx \frac{\mathbf{f}(\mathbf{E}_k^{n+1} + \varepsilon\mathbf{v}) - \mathbf{f}(\mathbf{E}_k^{n+1} - \varepsilon\mathbf{v})}{2\varepsilon}. \tag{I.2}$$

Here ε is a small perturbation parameter, which cannot be too large (poor derivative approximation) or too small (large floating-point roundoff error). Omitting the subscript k and superscript $n+1$ for notational convenience, we use the formula

$$\varepsilon = \begin{cases} \frac{1}{n\|\mathbf{v}\|_2} \sum_{i=1}^{n} \sqrt{\varepsilon}(1 + |E_i|) & \text{if } \|\mathbf{v}\|_2 > \varepsilon \\ \frac{1}{n} \sum_{i=1}^{n} \sqrt{\varepsilon}(1 + |E_i|) & \text{if } \|\mathbf{v}\|_2 \leq \varepsilon, \end{cases} \tag{I.3}$$

where n is the size of $\mathbf{E}$ ($n = N^2$ for N grid cells in each spatial direction) and $\varepsilon = 2.2204 \cdot 10^{-16}$ (machine epsilon for 64-bit double precision). In our simulations we use expression (I.2).

Appendix J

Local and global sensitivity analysis

To identify the relative influence of the input parameters on a model's output, we may perform a local or global sensitivity analysis [46]. The former measures changes in the model output with respect to variations in a single parameter. Typically, parameters are changed one at a time in relatively small increments starting from some nominal value for all parameters, yielding a local sensitivity index for each parameter. This strategy only works when the model output is linearly related to the parameters near the reference point, and does not allow for the evaluation of simultaneous changes in all input parameters. These limitations may be overcome through a global sensitivity analysis, which enables parameters to be varied simultaneously and yields the relative contributions of each individual parameter, as well as the interactions between parameters, to the model output variance. Examples of global sensitivity techniques are Fourier amplitude sensitivity analysis (FAST) [81] and the method by Sobol' [105], which are based on variance decomposition techniques. We now focus on the latter approach.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_s)$, be the vector of input parameters. Each parameter is considered to range over some finite interval which may be assumed, after rescaling, to be $[0,1]$. Assuming that the x_i $(i = 1, \ldots, s)$ are mutually independent, uniformly

distributed random variables on $[0,1]$, we may decompose f into

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^{s} f_i(x_i) + \sum_{i=1}^{s}\sum_{j<i} f_{i,j}(x_i,x_j) + \cdots + f_{1,\ldots,s}(x_1,\ldots,x_s), \tag{J.1}$$

in which

$$\begin{aligned} f_0 &= \langle f \rangle = \int f(\mathbf{x})\, \mathrm{d}\mathbf{x} \\ f_i(x_i) &= \langle f|x_i\rangle - f_0 = \int f(\mathbf{x}) \prod_{k\neq i} \mathrm{d}x_k - f_0 \\ f_{i,j}(x_i,x_j) &= \langle f|x_i,x_j\rangle - f_0 - f_i(x_i) - f_j(x_j), \end{aligned} \tag{J.2}$$

etc. Here we used the fact that the joint probability density function of $\mathbf{x}$, $p_{\mathbf{x}}$, is given by

$$p(\mathbf{x}) = \prod_{i=1}^{s} p_i(x_i) \equiv 1, \tag{J.3}$$

where the last equality follows from the fact that each parameter is uniformly distributed on $[0,1]$. Using the orthogonality property

$$\int f_{i_1,\ldots,i_s}(x_{i_1},\ldots,x_{i,s})\mathrm{d}x_k = 0 \text{ for } k = i_1,\ldots,i_s, \tag{J.4}$$

we obtain for the variance σ_f^2 of f

$$\sigma_f^2 = \sum_{i=1}^{k} \sigma_i^2 + \sum_{i=1}^{s}\sum_{j<i} \sigma_{i,j}^2 + \cdots + \sigma_{1,\ldots,s}^2, \tag{J.5}$$

where

$$\sigma_{i_1,\ldots,i_s}^2 = \int f_{i_1,\ldots,i_s}^2 \mathrm{d}\mathbf{x}', \tag{J.6}$$

with $\mathbf{x}' = (x_{i_1}, x_{i_2}, \ldots, x_{i_s})^\top$, is the partial variance corresponding to the subset of parameters $\{x_{i_1}, \ldots, x_{i_s}\}$. The Sobol' sensitivity index for $\mathbf{x}'$ is then given by

$$S_{i_1,\ldots,i_s} = \frac{\sigma^2_{i_1,\ldots,i_s}}{\sigma^2_f}. \tag{J.7}$$

The first-order ("main effect") index $S_i = \sigma^2_i/\sigma^2_f$ measures the fractional contribution of the ith parameter to the output variance, while the second-order index $S_{i,j} = \sigma^2_{i,j}/\sigma^2_f$ is used to compute the fractional contribution from interaction between the ith and jth parameters, and so on. We may also define total-order sensitivity indices S_{Ti} quantifying the overall effect of the ith parameter on the model output through

$$S_{Ti} = \frac{1}{\sigma^2_f} \sum_{\alpha \in \mathcal{I}_i} \sigma^2_\alpha, \tag{J.8}$$

where $\mathcal{I}_i$ is the set of all subsets of $\{x_1, \ldots, x_s\}$ containing i. From (J.5), it is clear that

$$\sum_{i=1}^{k} S_i + \sum_{i=1}^{s} \sum_{j<i} S_{i,j} + \cdots + S_{1,\ldots,k} = 1. \tag{J.9}$$

The higher the sensitivity index for a parameter, the more effect it has on the model output. While no distinct cutoff exists, the value of 0.05 is frequently used to distinguish important from unimportant parameters, although a more stringent condition may apply for models with only a few input parameters [127].

Appendix K

Landau free energy

Equations (5.8) and (5.9) can be rewritten in the form of a Langevin-type equation

$$\frac{\partial m}{\partial t} = -\Gamma \frac{\delta F[m;T]}{\delta m} + \eta(x,t), \tag{K.1}$$

where $F[m(x,t);T]$ is the free-energy functional of the system and Γ is the inverse damping coefficient. This equation describes the relaxation of a system to its free-energy minimum. For the sGLEa, $F[m;T]$ is given by

$$F = F_0[T] + \int \left[\frac{A}{2} m^2 + \frac{B}{4} m^4 + \frac{c^2 J}{2} \left(\frac{\partial m}{\partial x} \right)^2 \right] \mathrm{d}x, \tag{K.2}$$

where $F_0[T]$ is a temperature-dependent constant, and the coefficients A and B are defined in (5.6b). In the deterministic case ($\eta = 0$), the free energy F reaches its minimum value at equilibrium,

$$F(T) = \begin{cases} \pm\sqrt{3(R-1)/R^3} & \text{for } T < T_c \\ 0 & \text{for } T \geq T_c \end{cases}, \tag{K.3}$$

where $R = T_c/T$ and T_c is the mean-field critical temperature. Figure K.1 illustrates this point by showing the shape of $F[m;T]$ at different temperatures around T_c.

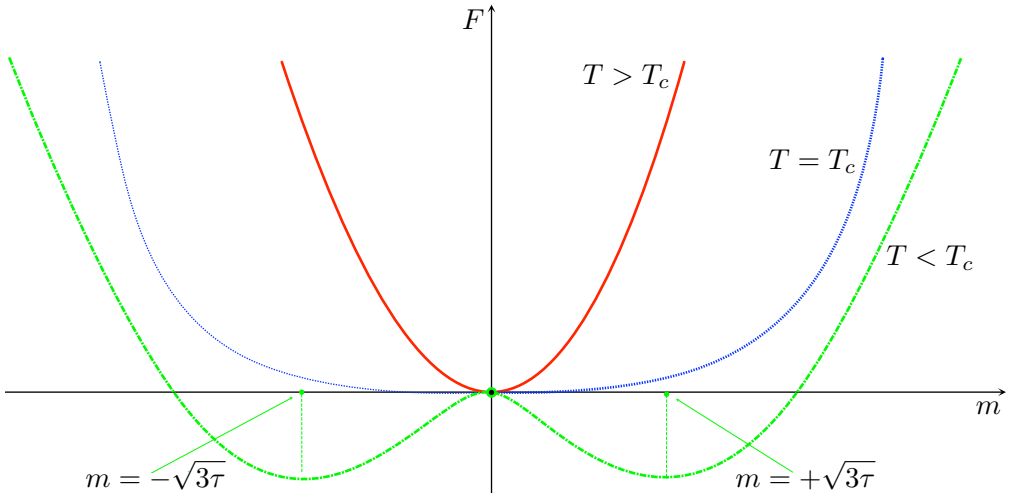

Figure K.1: Landau free energy with $\tau = 3(R-1)/R^3$.

Appendix L

Statistical moment equations for sGLEs

We use the Reynolds decomposition to represent the magnetization $m(\mathbf{x},t)$ with $\mathbf{x} \in \mathbb{R}^d$ as $m = \overline{m} + m'$, where $m'(\mathbf{x},t)$ is the zero-mean fluctuation of m about its ensemble average $\overline{m}(\mathbf{x},t)$. Expanding $f(m) = Am + Bm^3$ into a Taylor series around $\overline{m}$,

$$f(m) = \sum_{i=0}^{\infty} \frac{f^{(i)}(\overline{m})}{i!} \left(m'\right)^i, \tag{L.1}$$

and substituting the first four terms in this expansion into a d-dimensional version of (5.8), we obtain

$$\frac{\partial m}{\partial t} = \Gamma K \nabla^2 m - \Gamma \Big[A\overline{m} + B\overline{m}^3 + (A + 3B\overline{m}^2)m' + 3B\overline{m}(m')^2 + B(m')^3 \Big] + \eta. \tag{L.2}$$

where we have defined $K \equiv c^2 J$, and omitted the subscript a in Γ_a and η_a for notational convenience. Taking the ensemble average of (L.2) gives an approximate equation for

the mean dynamics,

$$\frac{\partial \overline{m}}{\partial t} = \Gamma K \nabla^2 \overline{m} - \Gamma \left[A\overline{m} + B\overline{m}^3 + 3B\overline{m}\sigma_m^2 + B\overline{(m')^3} \right], \tag{L.3}$$

where $\sigma_m^2(\mathbf{x},t) \equiv C_m(\mathbf{x},t;\mathbf{x},t)$ is the variance of $m(\mathbf{x},t)$. Assuming $m'(\mathbf{x},t)$ to be a Gaussian random field with probability density function

$$p_m(v;\mathbf{x},t) = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\left(-\frac{v^2}{2\sigma_m^2} \right), \tag{L.4}$$

we approximate (L.3) with

$$\frac{\partial \overline{m}}{\partial t} = \Gamma K \nabla^2 \overline{m} - \Gamma \left[A\overline{m} + B\overline{m}^3 + 3B\overline{m}\sigma_m^2 \right]. \tag{L.5}$$

To calculate higher moments of m, including σ_m^2, we compute the cross-covariance $C_{\eta m}(\mathbf{y},\tau;\mathbf{x},t)$. Its governing equation is obtained by multiplying (L.2) with $\eta(\mathbf{y},\tau)$ and averaging,

$$\frac{\partial C_{\eta m}}{\partial t} = \Gamma K \nabla^2 C_{\eta m} - \Gamma \Big[(A + 3B\overline{m}^2) C_{\eta m} + 3B\overline{m}\,\overline{(m')^2 \eta} + B\overline{(m')^3 \eta} \Big] + C_\eta. \tag{L.6}$$

Since $\eta(\mathbf{x},t)$ is Gaussian and $m'(\mathbf{x},t)$ is assumed to be Gaussian, their mixed moments are computed as

$$\overline{(m')^n \eta} = \frac{1}{2\pi\sqrt{\Omega}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \xi v^n \mathrm{e}^{-\Pi} \mathrm{d}\xi \mathrm{d}v, \tag{L.7}$$

where $\Omega = \sigma_\eta^2\sigma_m^2 - C_{\eta m}^2$ and $\Pi = (\xi^2\sigma_m^2 - 2\xi v C_{\eta m} + v^2\sigma_\eta^2)/(2\Omega)$. This yields

$$\begin{aligned}\overline{(m')^n\eta} &= \frac{1}{\sqrt{2\pi}}\frac{C_{\eta m}}{\sigma_m^3}\int_{-\infty}^{+\infty} v^{n+1}\exp\left(\frac{v^2C_{\eta m}^2}{2\Omega\sigma_m^2} - \frac{v^2\sigma_\eta^2}{2\Gamma}\right)\mathrm{d}v \\ &= \frac{1}{\sqrt{2\pi}}\frac{C_{\eta m}}{\sigma_m^3}\int_{-\infty}^{+\infty} v^{n+1}\exp\left(-\frac{v^2}{2\sigma_m^2}\right)\mathrm{d}v, \end{aligned} \tag{L.8}$$

so that

$$\overline{(m')^2\eta} = 0, \qquad \overline{(m')^3\eta} = \frac{1}{\sqrt{2\pi}}\frac{C_{\eta m}}{\sigma_m^3}\frac{3\sqrt{\pi}(2\sigma_m^2)^{5/2}}{4} = 3\sigma_m^2 C_{\eta m}. \tag{L.9}$$

Substituting (L.9) into (L.6) gives an equation for $C_{\eta m}(\mathbf{y},\tau;\mathbf{x},t)$,

$$\frac{\partial C_{\eta m}}{\partial t} = \Gamma K\nabla^2 C_{\eta m} - \Gamma\left[A + 3B(\overline{m}^2 + \sigma_m^2)\right]C_{\eta m} + C_\eta. \tag{L.10}$$

An equation for the two-point time-space auto-covariance $C_m(\mathbf{x},t;\mathbf{y},\tau)$ is obtained by multiplying (L.2) with $m'(\mathbf{y},\tau)$ and averaging,

$$\frac{\partial C_m}{\partial t} = \Gamma K\nabla^2 C_m - \Gamma\left[(A + 3B\overline{m}^2)C_m + B\overline{m'(y,\tau)m'(x,t)^3}\right] + C_{\eta m}. \tag{L.11}$$

Using the second expression in (L.9) and replacing η with m' yields

$$\frac{\partial C_m}{\partial t} = \Gamma K\nabla^2 C_m - \Gamma\left[A + 3B(\overline{m}^2 + \sigma_m^2)\right]C_m + C_{\eta m}. \tag{L.12}$$

One-dimensional versions of (L.5), (L.10) and (L.12) are given by (5.16)–(5.18), respectively.

To compute the steady-state solution of these equations, we set the left-hand sides of (5.16)–(5.18) to zero and use a three-point stencil for the 1D Laplacian. Discretizing the interval $[0,L]$ into I grid cells of size Δx, such that $0 \le i,j \le I-1$ and $L = I\Delta x$, this

leads to a system of dimensionless equations

$$0 = \overline{m}^{i+1} + \overline{m}^{i-1} - \Lambda^{i,i}\overline{m}^{i} - \frac{q^{\star}R^2}{3}(\overline{m}^{i})^3 \tag{L.13}$$

$$0 = \hat{C}_{\eta m}^{i+1,j} + \hat{C}_{\eta m}^{i-1,j} - \left[\Lambda^{i,i} + q^{\star}R^2(\overline{m}^{i})^2\right]\hat{C}_{\eta m}^{i,j} + \frac{2A^{i,j}q^{\star}}{R}\hat{\delta}_{\text{dis}}^{i,j} \tag{L.14}$$

$$0 = C_{m}^{i+1,j} + C_{m}^{i-1,j} - \left[\Lambda^{i,i} + q^{\star}R^2(\overline{m}^{i})^2\right]C_{m}^{i,j} + \frac{q^{\star}}{R}\hat{C}_{\eta m}^{i,j} \tag{L.15}$$

where $\hat{C}_{\eta m}^{i,j}$ is the dimensionless version of $C_{\eta m}^{i,j}$, $\hat{\delta}_{\text{dis}}^{i,j}$ is the dimensionless discretization of the delta function $\delta(x_i - y_j)$, and

$$q^{\star} = q(\Delta x')^2, \qquad \Lambda^{i,i} = 2 + q^{\star}\left(\frac{1-R}{R} + R^2 C_{m}^{i,i}\right) \tag{L.16}$$

$$A^{i,j} = \begin{cases} \zeta_a^2 & \text{for} \quad I/4 \le i < 3I/4 \\ A_2 & \text{for} \quad 0 \le i < I/4 \quad \text{or} \quad 3I/4 \le i \le I \end{cases}. \tag{L.17}$$

The subscripts indicate the covariance type, while the superscripts i and j refer to the spatial discretization of x and y, respectively. To discretization of $\delta(x_i - y_j)$, we approximate it by a Gaussian, $\delta(x_i - y_j) \approx (\pi\nu)^{-1/2}\exp[-(x_i - y_j)^2/\nu]$, which is exact in the limit of $\nu \to 0$. This results in

$$\hat{\delta}_{\text{dis}}^{i,j} = \frac{1}{\sqrt{\pi\nu'}}\exp\left[-\frac{(i-j)^2(\Delta x')^2}{\nu'}\right], \qquad \nu' = \frac{\nu}{c^2}. \tag{L.18}$$

In the simulations reported in Figure 5.4 we set $\nu' = 0.2304$.

Since $m(x=0) = m(x=L)$ and $\eta(x=0) = \eta(x=L)$, their moments $\overline{m}$, C_m and

$C_{\eta m}$ are periodic as well, i.e.

$$\begin{aligned} \overline{m}^{i=-1} &= \overline{m}^{i=I-1}, & \overline{m}^{i=I} &= \overline{m}^{i=0}, \\ C_m^{i=-1,j} &= C_m^{i=I-1,j}, & C_m^{i=I,j} &= C_m^{i=0,j}, \\ C_{\eta m}^{i=-1,j} &= C_{\eta m}^{i=I-1,j}, & C_{\eta m}^{i=I,j} &= C_{\eta m}^{i=0,j}. \end{aligned} \tag{L.19}$$

The nonlinear system of algebraic equations (L.13)-(L.19) is solved with an iterative Newton-Raphson method (with the tolerance 10^{-6}), and the resulting magnetization variance profile is compared to that obtained with the time-dependent of the stochastic GLE. In the simulations reported here, we set $I = 40$, $q = 2$ and $R = 2/1.8$, and use the following initial guess for the unknowns:

$$\overline{m}^i = 0 \qquad \text{for } i = 0, \ldots, N-1 \tag{L.20}$$

$$C_m^{i,j} = \begin{cases} 1 & \text{for } I/4 \le i < 3I/4 \text{ and } j = i \\ 1 & \text{for } 0 \le i < I/4 \text{ or } 3I/4 \le i < I \text{ and } j = i \\ 0 & \text{otherwise} \end{cases} \tag{L.21}$$

$$\hat{C}_{\eta m}^{i,j} = 1 \quad \text{for } i, j = 0, \ldots, I-1. \tag{L.22}$$

Appendix M

Computation of the parameters in the learned sGLE model

In order to calculate the coefficients of the learned discrete sGLE, we employ the widely used statistical technique of maximizing the log-likelihood that the model will predict the GD training data. The discrete sGLE we are trying to learn has the form

$$\begin{aligned}\phi_{n+1,i,j} = \phi_{n,i,j} + \alpha_0(\phi_{n,i+1,j} + \phi_{n,i-1,j} + \phi_{n,i,j+1} \\ + \phi_{n,i,j-1} - 4\phi_{n,i,j}) + \sum_{k=0}^{\frac{C-1}{2}} \alpha_{k+1}\phi_{n,i,j}^{2k+1} \\ + \alpha_{(C+3)/2}\,\xi_{n,i,j},\end{aligned} \tag{M.1}$$

where C is the model complexity (we only consider odd complexities) and the $\xi_{n,i,j}$ are independent, identically distributed standard normal random variables. Subscripts n and $n+1$ refer to times t_n and t_{n+1}, while $i = 0,\ldots,\bar{N}_1 - 1$ and $j = 0,\ldots,\bar{N}_2 - 1$ are the spatial coordinates of the blocks obtained by coarse-graining the spin lattice, with $\bar{N}_1$ and $\bar{N}_2$ the number of blocks in both spatial directions. Given the block-averaged training data $S_{n,i,j}$ with $n = 0,\ldots,n_{\mathrm{eq}}$, we would like to find the set of parameters

$\boldsymbol{\alpha}_{\mathbf{opt}} \equiv \{\alpha_0, \alpha_1, \ldots, \alpha_{(C+3)/2}\}$ that maximizes the probability of observing S_0, S_1,..., $S_{n_{\text{eq}}}$ from model (M.1). Here S_n refers to the $\bar{N}_1 \times \bar{N}_2$ block-averaged spin configuration after n steps following the convention introduced in Sec. 6.4. This probability is given by

$$\begin{aligned}
&\mathbb{P}(S_0, S_1, \ldots, S_{n_{\text{eq}}}; \boldsymbol{\alpha}) \\
&= \mathbb{P}(S_0)\mathbb{P}(S_1 \mid S_0; \boldsymbol{\alpha}) \ldots \mathbb{P}(S_{n_{\text{eq}}} \mid S_1, \ldots, S_{n_{\text{eq}}-1}; \boldsymbol{\alpha}) \\
&= \mathbb{P}(S_0)\mathbb{P}(S_1 \mid S_0; \boldsymbol{\alpha}) \ldots \mathbb{P}(S_{n_{\text{eq}}} \mid S_{n_{\text{eq}}-1}; \boldsymbol{\alpha}).
\end{aligned} \tag{M.2}$$

Given S_n, we can see from (M.1) that for each i and j, $S_{n+1,i,j}$ is normally distributed with mean

$$\begin{aligned}
Y_{n,i,j} = S_{n,i,j} + \alpha_0 (&S_{n,i+1,j} + S_{n,i-1,j} + S_{n,i,j+1} \\
&+ S_{n,i,j-1} - 4S_{n,i,j}) + \sum_{k=0}^{\frac{C-1}{2}} \alpha_{k+1} S_{n,i,j}^{2k+1}
\end{aligned} \tag{M.3}$$

and variance $\alpha^2_{(C+3)/2}$. Therefore, we have

$$\begin{aligned}
&\mathbb{P}\left(S_{n+1,i,j} = s \mid S_n; \boldsymbol{\alpha}\right) \\
&= \frac{1}{\alpha_{(C+3)/2}\sqrt{2\pi}} \exp\left[-\frac{\left(s - Y_{n,i,j}\right)^2}{2\alpha^2_{(C+3)/2}}\right],
\end{aligned} \tag{M.4}$$

and since the $\xi_{n,i,j}$ are independent,

$$\begin{aligned}
&\mathbb{P}\left(S_{n+1} \mid S_n; \boldsymbol{\alpha}\right) \\
&= \frac{1}{\left(\alpha^2_{(C+3)/2} 2\pi\right)^{(\bar{N}_1\bar{N}_2)/2}} \exp\left[-\frac{\|S_{n+1} - Y_n\|^2}{2\alpha^2_{(C+3)/2}}\right],
\end{aligned} \tag{M.5}$$

where Y_n is defined according to the notational convention in Sec. 6.4. Hence,

$$\begin{aligned}
&\ln \mathbb{P}\left(S_0, S_1, \ldots, S_{n_{\mathrm{eq}}}; \boldsymbol{\alpha}\right) \\
&= -\sum_{n=0}^{n_{\mathrm{eq}}-1} \frac{\left\|S_{n+1} - Y_n\right\|^2}{2\alpha_{(C+3)/2}^2} \\
&\quad - n_{\mathrm{eq}} \bar{N}_1 \bar{N}_2 \ln\left(\alpha_{(C+3)/2}\right) + \text{constant}.
\end{aligned} \tag{M.6}$$

For notational simplicity, let us now focus on the case of $C = 3$; the generalization to higher order nonlinearity is straightforward. From (M.6), it follows that we need to minimize

$$\begin{aligned}
\mathcal{L}\left(S_1, \ldots, S_{n_{\mathrm{eq}}}; \boldsymbol{\alpha}\right) = \frac{1}{2}\alpha_3^{-2} f\left(\alpha_0, \alpha_1, \alpha_2\right) \\
+ n_{\mathrm{eq}} \bar{N}_1 \bar{N}_2 \ln\left(\alpha_3\right),
\end{aligned} \tag{M.7}$$

where

$$\begin{aligned}
f\left(\alpha_0, \alpha_1, \alpha_2\right) = \sum_{n=0}^{n_{\mathrm{eq}}-1} \sum_{i=0}^{\bar{N}_1-1} \sum_{j=0}^{\bar{N}_2-1} [(S_{n+1,i,j} - S_{n,i,j}) \\
- \alpha_0 (S_{n,i+1,j} + S_{n,i-1,j} + S_{n,i,j+1} \\
+ S_{n,i,j-1} - 4S_{n,i,j}) - \alpha_1 S_{n,i,j} \\
- \alpha_2 S_{n,i,j}^3]^2.
\end{aligned} \tag{M.8}$$

From

$$\frac{\partial \mathcal{L}}{\partial \alpha_3} = -\alpha_3^{-3} f\left(\alpha_0, \alpha_1, \alpha_2\right) + n_{\mathrm{eq}} \bar{N}_1 \bar{N}_2 \alpha_3^{-1} = 0, \tag{M.9}$$

we can solve for α_3

$$\alpha_3 = \sqrt{\frac{f(\alpha_0, \alpha_1, \alpha_2)}{n_{\text{eq}} \bar{N}_1 \bar{N}_2}}. \tag{M.10}$$

Also, for $k \neq 3$, we have

$$\frac{\partial \mathcal{L}}{\partial \alpha_k} = \frac{1}{2} \alpha_3^{-2} \frac{\partial f}{\partial \alpha_k}. \tag{M.11}$$

If we now define

$$\begin{aligned} &D_{n,i,j} = S_{n+1,i,j} - S_{n,i,j}, \\ &A_{n,i,j} = S_{n,i+1,j} + S_{n,i-1,j} + S_{n,i,j+1} + S_{n,i,j-1} - 4S_{n,i,j}, \\ &B_{n,i,j} = S_{n,i,j},\ C_{n,i,j} = S_{n,i,j}^3, \end{aligned} \tag{M.12}$$

then

$$\begin{aligned} \frac{\partial f}{\partial \alpha_0} &= \partial_{\alpha_0} \sum_{n=0}^{n_{\text{eq}}-1} \sum_{i=0}^{\bar{N}_1-1} \sum_{j=0}^{\bar{N}_2-1} (D_{n,i,j} - \alpha_0 A_{n,i,j} \\ &\quad - \alpha_1 B_{n,i,j} - \alpha_2 C_{n,i,j})^2 \\ &= -2 \left(\sum_{n=0}^{n_{\text{eq}}-1} \sum_{i=0}^{\bar{N}_1-1} \sum_{j=0}^{\bar{N}_2-1} D_{n,i,j} A_{n,i,j} \right) \\ &\quad + 2 \left(\sum_{n=0}^{n_{\text{eq}}-1} \sum_{i=0}^{\bar{N}_1-1} \sum_{j=0}^{\bar{N}_2-1} A_{n,i,j}^2 \right) \alpha_0 \\ &\quad + 2 \left(\sum_{n=0}^{n_{\text{eq}}-1} \sum_{i=0}^{\bar{N}_1-1} \sum_{j=0}^{\bar{N}_2-1} B_{n,i,j} A_{n,i,j} \right) \alpha_1 \\ &\quad + 2 \left(\sum_{n=0}^{n_{\text{eq}}-1} \sum_{i=0}^{\bar{N}_1-1} \sum_{j=0}^{\bar{N}_2-1} C_{n,i,j} A_{n,i,j} \right) \alpha_2 \\ &= -2 (a_0 - a_{00} \alpha_0 - a_{01} \alpha_1 - a_{02} \alpha_2). \end{aligned} \tag{M.13}$$

Similarly,

$$\frac{\partial f}{\partial \alpha_1} = -2\left(a_1 - a_{10}\alpha_0 - a_{11}\alpha_1 - a_{12}\alpha_2\right), \tag{M.14}$$

where

$$\begin{aligned}
a_1 &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} D_{n,i,j}B_{n,i,j},\\
a_{10} &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} A_{n,i,j}B_{n,i,j},\\
a_{11} &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} B_{n,i,j}^2,\\
a_{12} &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} C_{n,i,j}B_{n,i,j}.
\end{aligned} \tag{M.15}$$

Finally,

$$\frac{\partial f}{\partial \alpha_2} = -2\left(a_2 - a_{20}\alpha_0 - a_{21}\alpha_1 - a_{22}\alpha_2\right), \tag{M.16}$$

where

$$\begin{aligned}
a_2 &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} D_{n,i,j}C_{n,i,j},\\
a_{20} &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} A_{n,i,j}C_{n,i,j},\\
a_{21} &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} B_{n,i,j}C_{n,i,j},\\
a_{22} &= \sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1} C_{n,i,j}^2.
\end{aligned} \tag{M.17}$$

Hence, we need to solve the linear system

$$\begin{aligned} a_{00}\alpha_0 + a_{01}\alpha_1 + a_{02}\alpha_2 &= a_0, \\ a_{10}\alpha_0 + a_{11}\alpha_1 + a_{12}\alpha_2 &= a_1, \\ a_{20}\alpha_0 + a_{21}\alpha_1 + a_{22}\alpha_2 &= a_2. \end{aligned} \tag{M.18}$$

We can see that despite the nonlinearity of the sGLE, optimizing the log-likelihood function can be reduced to solving a linear system.

Appendix N

Details of the operational procedure for calculating the error pdfs

To compute one data point in the error pdf for a learned sGLE model of complexity C given a number of training samples N_{train}, we do the following.

1. We simulate N_{test} independent GD test sample trajectories. The kth trajectory is obtained as follows:

   (a) Starting from a random initial Ising configuration, we march over n_{mc} steps.

   (b) Starting from the resulting Ising configuration, we march over n_{eq} steps and store the block-averaged time history over these steps in a 3D matrix $s^{\text{av,test}}$ with dimensions $(n_{\text{eq}}+1)\times\bar{N}_1\times\bar{N}_2$. Here $\bar{N}_1$ and $\bar{N}_2$ represent the number of spin blocks in each spatial direction.

   (c) This matrix $s^{\text{av,test}}$ will be the kth element of a 4D matrix $s^{\text{av,test,all}}$ with dimensions $N_{\text{test}}\times(n_{\text{eq}}+1)\times\bar{N}_1\times\bar{N}_2$.

   The initial spins $s_{i,j}$ ($i=0,\ldots,N_1-1$ and $j=0,\ldots,N_2-1$) are given by $s_{i,j}=1-2\,r_{i,j}$, where the $r_{i,j}$ are drawn from a discrete uniform distribution on the

half-open interval $[0,2)$. Furthermore, with "block-averaged time history", we refer to the time evolution of the block-averaged spin configuration of the Ising lattice. At each discrete point in time, we group the individual spins into blocks of a certain size, and then calculate the average spin values over the different blocks. The resulting coarsened grid is then recorded.

2. We simulate N_{train} independent GD training sample trajectories. A trajectory is calculated as follows:

   (a) Starting from a random initial Ising configuration, we march over n_{mc} steps.

   (b) Starting from the final Ising configuration, we march over n_{eq} steps and store the block-averaged time history over these steps in a 3D matrix $s^{\text{av,train}}$ with dimensions $(n_{\text{eq}}+1)\times\bar{N}_1\times\bar{N}_2$.

   (c) We concatenate $s^{\text{av,train}}$ of the current training sample with the corresponding matrices of the previous training samples along the first (time) dimension, and hence obtain a bigger matrix $s^{\text{av,train,all}}$ with dimensions $N_{\text{train}}\times(n_{\text{eq}}+1)\times\bar{N}_1\times\bar{N}_2$.

3. Using the training data stored in $s^{\text{av,train,all}}$, we compute the coefficients of the learned sGLE polynomial using a log-likelihood analysis (see Appendix D).

4. With the parameters calculated in step 3, we now simulate N_{test} sGLE trajectories. The kth trajectory is obtained as follows:

   (a) We define a 3D matrix ϕ with dimensions $(n_{\text{eq}}+1)\times\bar{N}_1\times\bar{N}_2$.

   (b) We take the block-averaged Ising test configuration stored in $s^{\text{av,test,all}}_{k,0,:,:}$ as the initial condition and define $s^{\text{av,test}}\equiv s^{\text{av,test,all}}_{k,:,:,:}$. We then define $\phi_{0,:,:}\equiv s^{\text{av,test}}_{0,:,:}$.

   (c) We march over n_{eq} steps and store the time history over these steps in ϕ.

(d) The configuration $\phi_{n+1,:,:}$ is calculated according to

$$\phi_{n+1,i,j} = x_{i,j} + \mathcal{G}(x_{:,:};\boldsymbol{\alpha}), \tag{N.1}$$

where $x_{:,:}$ is either $s^{\text{av,test}}_{n,:,:}$ or $\phi_{n,:,:}$, $i = 0,\ldots,\bar{N}_1 - 1$ and $j = 0,\ldots,\bar{N}_2 - 1$.

In (d), we have defined $\mathcal{G}(x_{:,:};\boldsymbol{\alpha})$ as

$$\begin{aligned}
&\mathcal{G}(x_{:,:};\boldsymbol{\alpha}) \\
&= \alpha_0(x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1} - 4x_{i,j}) \\
&\quad + \sum_{k=0}^{\frac{C-1}{2}} \alpha_{k+1}(x_{i,j})^{2k+1} + \alpha_{(C+3)/2}\,\xi_{n,i,j}.
\end{aligned} \tag{N.2}$$

5. For the kth sGLE trajectory, we calculate the RMS error

$$\varepsilon_k = \sqrt{\frac{1}{a}\sum_{n=0}^{n_{\text{eq}}-1}\sum_{i=0}^{\bar{N}_1-1}\sum_{j=0}^{\bar{N}_2-1}(\phi_{n+1,i,j} - s^{\text{av,test}}_{n+1,i,j})^2}, \tag{N.3}$$

where $a = n_{\text{eq}}\bar{N}_1\bar{N}_2$.

6. Finally, we compute the test-averaged error

$$\varepsilon = \frac{1}{N_{\text{test}}}\sum_{k=1}^{N_{\text{test}}}\varepsilon_k, \tag{N.4}$$

which we will call the "type 1" test error if the $\phi_{n+1,i,j}$ are calculated using $x_{:,:} = s^{\text{av,test}}_{n,:,:}$ in (N.1), or the "type 2" test error if the $\phi_{n+1,i,j}$ are calculated using $x_{:,:} = \phi_{n,:,:}$ in (N.1).

Bibliography

[1] Nonequilibrium Gas and Plasma Dynamics Laboratory website, Department of Aerospace Engineering, University of Michigan, Ann Arbor. http://ngpdlab.engin.umich.edu/hypersonics/hybrid-particle_continuum-method.

[2] F. F. Abraham, J. Q. Broughton, N. Bernstein, and E. Kaxiras. Spanning the continuum to quantum length scales in a dynamic simulation of brittle fracture. *Europhys. Lett.*, 44(6):783, 1998.

[3] J.-F. Adam, H. Elleaume, A. Joubert, M.-C. Biston, A.-M. Charvet, J. Balosso, J.-F. Le Bas, and F. Estève. Synchrotron radiation therapy of malignant brain glioma loaded with an iodinated contrast agent: first trial on rats bearing F98 gliomas. *Int. J. Radiation Oncology Biol. Phys.*, 57(5):1413–1426, 2003.

[4] G. Alefeld and J. Völkl. *Hydrogen in Metals I*. Springer-Verlag, Berlin, 1978.

[5] F. J. Alexander, A. L. Garcia, and D. M. Tartakovsky. Algorithm refinement for stochastic partial differential equations: I. Linear diffusion. *J. Comput. Phys.*, 182(1):47–66, 2002.

[6] F. J. Alexander, A. L. Garcia, and D. M. Tartakovsky. Algorithm refinement for stochastic partial differential equations: II. Correlated systems. *J. Comput. Phys.*, 663:915–922, 2003.

[7] F. J. Alexander, A. L. Garcia, and D. M. Tartakovsky. Noise in algorithm refinement methods. *Comput. Sci. Eng.*, 7(3):32–38, 2005.

[8] F. J. Alexander, G. Johnson, G. L. Eyink, and I. G. Kevrekidis. Equation-free implementation of statistical moment closures. *Phys. Rev. E*, 77:26701, 2008.

[9] American Brain Tumor Association. *Conventional Radiation Therapy*. 2014.

[10] S. Arrhenius. Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Zeit. Phys. Chem.*, 4:226–248, 1889.

[11] K. Balakrishnan, J. B. Bell, A. Donev, and A. L. Garcia. Fluctuating hydrodynamics and Direct Simulation Monte Carlo. *Proceedings of the 28th International Symposium on Rarefied Gas Dynamics*, 1501:695–704, 2012.

[12] S. Banach. Sur les opérations dans les ensembles abstraits et leur applications aux équations intégrales. *Fund. Math.*, 3:133–181, 1922.

[13] I. Battiato and D. M. Tartakovsky. Applicability regimes for macroscopic models of reactive transport in porous media. *J. Contam. Hydrol.*, 120-21:18–26, 2011.

[14] I. Battiato, D. M. Tartakovsky, A. M. Tartakovsky, and T. Scheibe. On breakdown of macroscopic models of mixing-controlled heterogeneous reactions in porous media. *Adv. Water Resour.*, 32:1664–1673, 2009.

[15] Y. Bazilevs, A. Korobenko, X. Deng, and J. Yan. Novel structural modeling and mesh moving techniques for advanced fluid-structure interaction simulation of wind turbines. *Int. J. Numer. Meth. Engng.*, 102(3), 2014.

[16] J. B. Bell, J. Foo, and A. L. Garcia. Algorithm refinement for the stochastic Burgers' equation. *J. Comput. Phys.*, 223(1):451–468, 2007.

[17] D. Bhattacharyya and T. Kim. Brain tumor detection using MRI image analysis. *Ubiquitous Computing and Multimedia Applications: Second International Conference, UCMA 2011, Daejeon, Korea, April 13-15, 2011*, 2011.

[18] N. Bilal. Implementation of Sobol's method of global sensitivity analysis to a compressor simulation model. *22nd International Compressor Engineering Conference at Purdue, July 14-17, 2014*, 2014.

[19] G. A. Bird. *Molecular gas dynamics and the direct simulation of gas flows*. Oxford University Press, New York, 1994.

[20] F. J. Blom. A monolithical fluid-structure interaction algorithm applied to the piston problem. *Comput. Methods Appl. Mech. Engrg.*, 167:369–391, 1998.

[21] J. Boon, H. Li, J. W. Dijkstra, and J. A. Z. Pieterse. 2-dimensional membrane separator modelling: Mass transfer by convection and diffusion. *Energy Procedia*, 4:699–706, 2011.

[22] A. C. Bose. *Inorganic membranes for energy and environmental applications*. Springer-Verlag, Berlin, 2009.

[23] F. Boso and I. Battiato. Homogenizability conditions for multicomponent reactive transport. *Adv. Water Resour.*, 62:254–265, 2013.

[24] A. J. Bray. Theory of phase-ordering kinetics. *Adv. Phys.*, 43(3):357–459, 1994.

[25] R. E. Buxbaum and A. B. Kinney. Hydrogen transport through tubular membranes of palladium-coated tantalum and niobium. *Ind. Eng. Chem. Res.*, 35:530–537, 1996.

[26] G. T. Buzzard and D. Xiu. Variance-based global sensitivity analysis via sparse-grid interpolation and cubature. *Commun. Comput. Phys.*, 9(3):542–567, 2010.

[27] J. Cardy. *Scaling and renormalization in statistical physics*. Cambridge Univ. Press, 1st edition, 1996.

[28] E. D. Cashwell and C. J. Everett. *A practical manual on the Monte Carlo method for random walk problems*. Pergamon Press, London, 1959.

[29] A. M. Charvet, J. F. LeBas, H. Elleuame, C. Schulze, P. Suortti, and P. Spanne. Medical applications of synchrotron radiation at the ESRF. *Proc. International School E. Fermi-CXXVIII Course-Biomedical Applications of Synchrotron Radiation, IOS Press, Amsterdam*, pages 355–377, 1995.

[30] G. C. Cheng, K. D. Neroorkar, Y. S. Chen, T. S. Wang, and E. O. Daso. Numerical study of flow augmented thermal management for entry and re-entry environments. *25th AIAA Applied Aerodynamics Conference, Miami, Florida (2007)*.

[31] H. Cho, X. Yang, D. Venturi, and G. E. Karniadakis. Algorithms for propagating uncertainty across heterogeneous domains. *SIAM J. Sci. Comput.*, 37:A3030–A3054, 2015.

[32] T.-J. Choi, M. R. Maurya, D. M. Tartakovsky, and S. Subramaniam. Stochastic hybrid modeling of intracellular calcium dynamics. *J. Chem. Phys.*, 133(16):165101, 2010.

[33] L. A. Dalton and E. R. Dougherty. Optimal mean-square-error calibration of classifier estimators under Bayesian models. *Pattern Recogn.*, 45(6):2308–2320, 2012.

[34] L. A. Dalton and E. R. Dougherty. Optimal classifiers with minimum expected error within a Bayesian framework–part I: Discrete and Gaussian models. *Pattern Recogn.*, 46(5):1301–1314, 2013.

[35] B. C. Daniels and I. Nemenman. Automated adaptive inference of phenomenological dynamical models. *Nat. Commun.*, 6:8133, 2015.

[36] M. De Falco, L. Marrelli, and G. Iaquaniello. *Membrane reactors for hydrogen production processes*. Springer-Verlag, Berlin, 2011.

[37] J. J. Duderstadt and G. A. Moses. *Inertial Confinement Fusion*. John Wiley and Sons, New York, 1982.

[38] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. Heterogeneous multiscale methods: A review. *Commun. Comput. Phys.*, 2(3):367–450, 2007.

[39] H. Elleaume, A.-M. Charvet, S. Corde, F. Estève, and J.-F. Le Bas. Performance of computed tomography for contrast agent concentration measurements with monochromatic x-ray beams: comparison of K-edge versus temporal subtraction. *Phys. Med. Biol.*, 47:3369–3385, 2002.

[40] H. Engels. *Numerical Quadrature and Cubature*. Academic Press, London, 1980.

[41] M.-P. Errera and S. Chemin. Optimal solutions of numerical interface conditions in fluid-structure thermal analysis. *J. Comput. Phys.*, 245:431–455, 2013.

[42] C. Farhat and M. Lesoinne. Two efficient staggered procedures for the serial and parallel solution of three-dimensional nonlinear transient aeroelastic problems. *Comput. Methods Appl. Mech. Engrg.*, 182:499–516, 2000.

[43] C. Farhat, M. Lesoinne, and P. Le Tallec. Load and motion transfer algorithms for fluid/structure interaction problems with non-matching discrete interfaces: Momentum and energy conservation, optimal discretization and application to aeroelasticity. *Comput. Methods Appl. Mech. Engrg.*, 157:95–114, 1998.

[44] C. A. Felippa and K. C. Park. Staggered transient analysis procedures for coupled mechanical systems: Formulation. *Comput. Methods Appl. Mech. Engrg.*, 24:61–111, 1980.

[45] G. Frantziskonis, K. Muralidharan, P. Deymier, S. Simunovic, P. Nukala, and S. Pannala. Time-parallel multiscale/multiphysics framework. *J. Comput. Phys.*, 228(21):8085–8092, 2009.

[46] H. C. Frey and S. R. Patil. Identification and review of sensitivity analysis methods. *Risk Anal.*, 22:553–578, 2002.

[47] K. J. Gaffney, A. M. Lindenberg, J. Larsson, K. Sokolowski-Tinten, C. Blome, O. Synnergren, J. Sheppard, C. Caleman, A. G. MacPhee, D. Weinstein, D. P. Lowney, T. Allison, T. Matthews, R. W. Falcone, A. L. Cavalieri, D. M. Fritz, S. H. Lee, P. H. Bucksbaum, D. A. Reis, J. Rudati, A. T. Macrander, P. H. Fuoss, C. C. Kao, D. P. Siddons, R. Pahl, K. Moffat, J. Als-Nielsen, S. Duesterer, R. Ischebeck, H. Schlarb, H. Schulte-Schrepping, J. Schneider, D. von der Linde, O. Hignette, F. Sette, H. N. Chapman, R. W. Lee, T. N. Hansen, J. S. Wark, M. Bergh, G. Huldt, D. van der Spoel, N. Timneanu, J. Hajdu, R. A. Akre, E. Bong, P. Krejcik, J. Arthur, S. Brennan, K. Luening, and J. B. Hastings. Observation of structural anisotropy and the onset of liquidlike motion during the nonthermal melting of InSb. *Phys. Rev. Lett*, 95:125701, 2005.

[48] A. L. Garcia. Adaptive mesh and algorithm refinement using Direct Simulation Monte Carlo. *J. Comput. Phys.*, 154:134–155, 1999.

[49] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag, New York, 1991.

[50] M. B. Giles. Stability analysis of numerical interface conditions in fluid-structure thermal analysis. *Int. J. Num. Methods Fluids*, 25:421–436, 1997.

[51] M. B. Giles. Multilevel Monte Carlo path simulation. *Oper. Res.*, 56(3):607–617, 2008.

[52] R. J. Glauber. Time-dependent statistics of the Ising model. *J. Math. Phys.*, 4(2):294–307, 1963.

[53] M. E. Gracheva, J. M. Rickman, and J. D. Gunton. Coarse-grained Ginzburg-Landau free energy for Lennard-Jones systems. *J. Chem. Phys.*, 113(9):3525–3529, 2000.

[54] T. Graham. On the adsorption and dialytic separation of gases by colloid septa. *Philos. Trans. R. Soc. London*, 156:399, 1866.

[55] W. Gropp. Parallel Computing and Domain Decomposition. *Proceedings of the 5th International Conference on Domain Decomposition Methods in Norfolk, Virginia (May 6-8, 1991)*.

[56] H. M. Hastings, R. J. Field, and S. G. Sobel. Microscopic fluctuations and pattern formation in a supercritical oscillatory chemical system. *J. Chem. Phys.*, 119(6):3291–3296, 2003.

[57] G. T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer-Verlag, London Limited, 2nd edition, 2009.

[58] P. C. Hohenberg and B. I. Halperin. Theory of dynamic critical phenomena. *Rev. Mod. Phys.*, 49(3):435–479, 1977.

[59] J. B. Hunter. Palladium alloy diffusion process for hydrogen purification. *Platinum Met. Rev.*, 4(4):130–131, 1960.

[60] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Z. Phys.*, 31:253–258, 1925.

[61] A. Ivanov, V. Sanchez, R. Stieglitz, and K. Ivanov. Internal multi-scale multi-physics coupled system for high fidelity simulation of light water reactors. *Ann. Nucl. Energy*, 66:104–112, 2014.

[62] K. D. Jarman and T. F. Russell. Eulerian moment equations for 2-D stochastic immiscible flow. *Multiscale Model. Simul.*, 1:598–608, 2003.

[63] N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions Volume 1*. John Wiley and Sons, New York, 2nd edition, 1994.

[64] O. Joshi and P. Leyland. Implementation of surface radiation and fluid-structure thermal coupling in atmospheric reentry. *Int. J. Aerospace Engrg.*, 2012:402653, 2012.

[65] M. H. Kalos and P. A. Whitlock. *Monte Carlo Methods*. Wiley-VCH, Weinheim, 2nd edition, 2008.

[66] K. Kaski, K. Binder, and J. D. Gunton. A study of a coarse-grained free energy functional for the three-dimensional Ising model. *J. Phys. A: Math. Gen.*, 16:623–627, 1983.

[67] K. Kaski, K. Binder, and J. D. Gunton. Study of cell distribution functions of the three-dimensional Ising model. *Phys. Rev. B*, 29(7):3996–4009, 1984.

[68] C. T. Kelley. *Iterative Methods for Linear and Nonlinear Equations*. SIAM Philadelphia, 1995.

[69] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 3rd edition, 1999.

[70] D. A. Knoll and D. E. Keyes. Jacobian-free Newton-Krylov methods: a survey of approaches and applications. *J. Comput. Phys.*, 193:357–397, 2004.

[71] L. D. Landau and E. M. Lifshitz. *Statistical physics, part 1: Course of theoretical physics (Volume 5)*. Butterworth-Heinemann, 3rd edition, 1980.

[72] J. M. LeBlanc and J. R. Wilson. A numerical example of the collapse of a rotating magnetized star. *Astrophys. J.*, 161:541, 1970.

[73] K. Leung. Heuristic derivation of continuum kinetic equations from microscopic dynamics. *Phys. Rev. E*, 63:016102, 2000.

[74] C. D. Levermore and G. C. Pomraning. A flux-limited diffusion theory. *Astrophys. J.*, 248:321–334, 1981.

[75] P. Leyland, V. Carstens, F. Blom, and T. Tefy. Fully coupled fluid-structure algorithms for aeroelasticity and forced vibration induced flutter. *Revue Européenne des Éléments*, 9(6-7):763–803, 2000.

[76] P. C. Lichtner and D. M. Tartakovsky. Upscaled effective rate constant for heterogeneous reactions. *Stoch. Environ. Res. Risk Assess.*, 17:419–429, 2003.

[77] G. Lin, A. M. Tartakovsky, and D. M. Tartakovsky. Uncertainty quantification via random domain decomposition and probabilistic collocation on sparse grids. *J. Comp. Phys.*, 229:6995–7012, 2010.

[78] J. Lindström and J. Nordström. A stable and high-order accurate conjugate heat transfer problem. *J. Comput. Phys.*, 229:14:5440–5456, 2010.

[79] G. Lythe and S. Habib. Stochastic PDEs: convergence to the continuum? *Comput. Phys. Commun.*, 142:29–35, 2001.

[80] A. Macovski. *Medical Imaging Systems*. Prentice-Hall, New Jersey, 1983.

[81] G. J. McRae, J. W. Tilden, and J. H. Seinfeld. Global sensitivity analysisa computational implementation of the Fourier Amplitude Sensitivity Test (FAST). *Comput. Chem. Eng.*, 6(1):15–25, 1982.

[82] A. Mensch and K. A. Thole. Conjugate heat transfer analysis of the effects of impingement channel height for a turbine blade endwall. *Int. J. Heat Mass Tran.*, 82:66–77, 2015.

[83] A. V. Mesa, A. Norman, T. D. Solberg, J. J. Demarco, and J. B. Smathers. Dose distributions using kilovoltage x-rays and dose enhancement from iodine contrast agents. *Phys. Med. Biol.*, 44:1955–1968, 1999.

[84] A. Mezzacappa, A. C. Calder, S. W. Bruenn, J. M. Blondin, M. W. Guidry, M. R. Strayer, and A. S. Umar. An investigation of neutrino-driven convection and the core collapse supernova mechanism using multigroup neutrino transport. *Astrophys. J.*, 495:911–926, 1998.

[85] D. Mihalas and B. Weibel-Mihalas. *Foundations of radiation hydrodynamics*. Oxford Univ. Press, New York, 1st edition, 1984.

[86] R. E. Miller and E. B. Tadmor. A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. *Modell. Simul. Mater. Sci. Eng.*, 17:053001, 2009.

[87] G. N. Milstein and N. V. Tretyakov. *Stochastic Numerics for Mathematical Physics*. Springer, 2004.

[88] H. Mori. Transport, collective motion and Brownian motion. *Prog. Theor. Phys.*, 33(3):423–455, 1965.

[89] T. S. Moss, N. M. Peachey, R. C. Snow, and R. C. Dye. Multilayer metal membranes for hydrogen separation. *Int. J. Hydrogen Energy*, 23(2):99–106, 1998.

[90] F. Nobile, R. Tempone, and C. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2309–2345, 2008.

[91] A. Norman, K. S. Iwamoto, and S. T. Cochran. Iodinated contrast agents for brain tumor localization and radiation dose enhancement. *Invest. Radiol.*, 26:Suppl. 1: S120–121; discussion S125–128, 1991.

[92] E. Novak and K. Ritter. Simple cubature formulas with high polynomial exactness. *Constr. Approx.*, 15:499–522, 1999.

[93] L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65:117–149, 1944.

[94] R. Pawlowski, R. Bartlett, N. Belcourt, R. Hooper, and R. Schmidt. A theory manual for multi-physics code coupling in LIME (Version 1.0). *Technical Report SAND2011-2195, Sandia National Laboratories*, 2011.

[95] T. L. Perelman. On conjugated problems of heat transfer. *Int. J. Heat Mass Transfer*, 3(4):293–303, 1961.

[96] S. Puri and V. Wadhawan. *Kinetics of phase transitions*. CRC Press, 1st edition, 2009.

[97] S. Ramu. Effective atomic numbers for photon energy absorption and photon attenuation of tissues from human organs. *Med. Dosim.*, 27(1):1–9, 2002.

[98] L. E. Reichl. *A Modern Course in Statistical Physics*. Wiley-VCH, Weinheim, 3rd edition, 2009.

[99] W. J. Rider, D. A. Knoll, and G. L. Olson. A multigrid Newton-Krylov method for multimaterial equilibrium radiation diffusion. *J. Comp. Phys.*, 152:164–199, 1999.

[100] C. P. Robert. Simulation of truncated normal variables. *Statistics and Computing*, 5:121–125, 1995.

[101] D. Roubinet and D. M. Tartakovsky. Hybrid modeling of heterogeneous geochemical reactions in fractured porous media. *Water Resour. Res.*, 49:doi:10.1002/2013WR013999, 2013.

[102] B. Sheehan, D. Estep, S. Tavener, J. Cary, S. Kruger, A. Hakim, A. Pletzer, J. Carlsson, and S. Vadlamani. The interaction of iteration error and stability for linear partial differential equations coupled through an interface. *Adv. Math. Phys.*, 2015:Article ID 787198, 2015.

[103] L. I. Smirnov. Effect of the concentration dependence of the diffusion coefficient on the distribution and flow of hydrogen in palladium membranes. *Inzh.-Fiz. Zh.*, 59(10):602–606, 1990.

[104] S. A. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 4:240–243, 1963.

[105] I. M. Sobol'. Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Exp.*, 1:407–414, 1993.

[106] I. M. Sobol'. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.*, 55:271–280, 2001.

[107] T. D. Solberg, K. S. Iwamoto, and A. Norman. Calculation of radiation dose enhancement factors for dose enhancement therapy of brain tumors. *Phys. Med. Biol.*, 37:439–443, 1992.

[108] H. E. Stanley. *Introduction to Phase Transitions and Critical Phenomena*. Oxford Univ. Press, 1st edition, 1971.

[109] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Syst. Safe.*, 93:964–979, 2008.

[110] A. M. Tartakovsky, D. M. Tartakovsky, T. D. Scheibe, and P. Meakin. Hybrid simulations of reaction-diffusion systems in porous media. *SIAM J. Sci. Comput.*, 30(6):2799–2816, 2008.

[111] D. M. Tartakovsky, M. Dentz, and P. C. Lichtner. Probability density functions for advective-reactive transport in porous media with uncertain reaction rates. *Water Resources Res.*, 45:W07414, 2009.

[112] D. M. Tartakovsky, A. Guadagnini, and M. Riva. Stochastic averaging of nonlinear flows in heterogeneous porous media. *J. Fluid Mech.*, 492:47–62, 2003.

[113] S. Taverniers, F. J. Alexander, and D. M. Tartakovsky. Noise propagation in hybrid models of nonlinear systems: The Ginzburg-Landau equation. *J. Comput. Phys.*, 262:313–324, 2014.

[114] S. Taverniers, T. S. Haut., K. Barros, F. J. Alexander, and T. Lookman. Physics-based statistical learning approach to mesoscopic model selection. *Phys. Rev. E*, 92:053301, 2015.

[115] S. Taverniers, A. Y. Pigarov, and D. M. Tartakovsky. Conservative tightly-coupled simulations of stochastic multiscale systems. *J. Comput. Phys.*, 313:400–414, 2016.

[116] S. Taverniers and D. M. Tartakovsky. A tightly-coupled domain-decomposition approach for highly nonlinear stochastic multiphysics systems. *J. Comput. Phys.*, submitted in revised form.

[117] S. Taverniers and D. M. Tartakovsky. Impact of parametric uncertainty on estimation of the energy deposition into an irradiated brain tumor. *J. Comput. Phys.*, under review.

[118] A. Tröster and C. Dellago. Coarse graining the ϕ^4 model: Landau-Ginzburg potentials from computer simulations. *Ferroelectrics*, 354:225–237, 2007.

[119] D. C. Wadsworth and D. A. Erwin. One-dimensional hybrid continuum/particle simulation approach for rarefied hypersonic flows. *Proc. AIAA/ASME 5th Joint Thermophysics and Heat Transfer Conference, Seattle, Washington, June 18-20 (1990)*.

[120] N. Wiener. The homogeneous chaos. *Am. J. Math.*, 60:897–936, 1938.

[121] S. A. Williams, J. B. Bell, and A. L. Garcia. Algorithm refinement for fluctuating hydrodynamics. *Multiscale Model. Simul.*, 6(4):1256–1280, 2008.

[122] C. L. Winter, D. M. Tartakovsky, and A. Guadagnini. Moment equations for flow in highly heterogeneous porous media. *Surv. Geophys.*, 24:81–106, 2003.

[123] X-5 Monte Carlo Team. MCNP - A general Monte Carlo N-Particle transport code, Version 5, Volume I: Overview and Theory. *LA-UR-03-1987, Los Alamos National Laboratory, 2008*, 2008.

[124] D. Xiu. *Numerical Methods for Stochastic Computations: A Spectral Method Approach*. Princeton University Press, Princeton, NJ, 2010.

[125] D. Xiu and J. S. Hesthaven. High order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.

[126] D. Xiu and G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.

[127] X. Y. Zhang, M. N. Trame, L. J. Lesko, and S. Schmidt. Sobol sensitivity analysis: A tool to guide the development and evaluation of systems pharmacology models. *CPT Pharmacometrics Syst. Pharmacol.*, 4:69–79, 2015.

[128] R. Zwanzig. Nonlinear generalized Langevin equations. *J. Stat. Phys.*, 9:215–220, 1973.