UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

TRANSFORMED PATH INTEGRAL BASED APPROACHES FOR STOCHASTIC DYNAMICAL SYSTEMS: PREDICTION, FILTERING, AND OPTIMAL CONTROL

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

GNANA MURUGAN SUBRAMANIAM

Norman, Oklahoma

2023

TRANSFORMED PATH INTEGRAL BASED APPROACHES FOR STOCHASTIC DYNAMICAL SYSTEMS: PREDICTION, FILTERING, AND OPTIMAL CONTROL

A DISSERTATION APPROVED FOR THE
SCHOOL OF AEROSPACE AND MECHANICAL ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Prakash Vedula, Chair

Dr. Peter Attar

Dr. Jivtesh Garg

Dr. Dimitrios Papavassiliou

Epigraph

குடிஎன்னும் குன்றா விளக்கம் மடிஎன்னும்
மாசுஊர மாய்ந்து கெடும்

Sluggishness will dim even one's bright innate inextinguishable qualities.

A **dynamic approach** is better than a fixed approach.

தொடங்கற்க எவ்வினையும் எள்ளற்க முற்றும்
இடங்கண்ட பின்அல் லது

Choose an advantageous location before engaging your opponent.

Focus resources at the **regions of importance**.

Thirukkural, Thiruvalluvar [1]

[1] Thirukkural: Navalar Urai, V.R. Nedunchezhiyan

Dedication

In loving memory of

Tmt. Gnanam Nayagam

&

Prof. C. Subramania Nayagam

Acknowledgements

I would like to extend my thanks and most sincere gratitude to my advisor Dr. Prakash Vedula. Without his support, patience, and invaluable insights this dissertation would not have been possible. I found our discussions deeply intellectually satisfying and a wellspring for ideas. May they long continue. I would also like to extend my gratitude to Dr. Peter Attar whose constant support at various stages of my growth as a graduate student and whose contributions to my research are greatly appreciated. I would like to thank the rest of my committee members both past and present: Dr. Jivtesh Garg, Dr. Dimitrios Papavassiliou, Dr. David Miller, and Dr. Andrea L'Afflitto. Their contributions to this research and my professional development are much appreciated. I am thankful for the support I received from the School of Aerospace and Mechanical Engineering, its chair Dr. Parthasarathy, former chair Dr. Siddique, Graduate Liaison Dr. Liu, and the rest of the faculty and staff members.

This work is dedicated to the memory of my *aachi* (grandmom) and my *appa* (father). My *aachi* instilled in me a deep love of science from a very young age. She also showed me that, through determination and hardwork, one can overcome any obstacle. And with my *appa*, I could not have asked for a better role model. He was the epitome of dilligence and virtue. He taught me to never compromise on the ideals and values I hold dear. He encouraged me to always trust my own judgement and live my life on my own terms. I have striven to incoporate their teachings in my life and in my work. I know they would have been proud to see me complete this work.

I am eternally grateful for the unconditional loving support of my dear *amma* (mother). It is her incredible sacrifices that have provided me with the opportunity to succeed. I am grateful for my *annan* (brother) and family (Shaila, Zen, and Wyn) whose support during

difficult times is greatly appreciated. I am thankful for the support of Meena *chithi* (aunt) and Thiru *chithappa* (uncle) who kept me motivated while I was in Chennai.

I am grateful for the good company of my friends Aseem Nevrekar and Siddharth John, who have been a source of mirth, camaraderie, and support through various parts of my journey as a graduate student in the U.S. I am thankful for the guidance provided by various teachers throughout my school years. In particular, I am thankful for my high school teacher Ravishankar who showed me the absolute joy in Physics and set me on the path to pursue a career in science. Lastly, I am most thankful for my beloved fiancée Madeline. She has been my pillar of support, a guardian angel, and a guiding light as I neared completion of this work. I eagerly look forward to us building a life together filled with joy, happines, and our annoying cats. I jest. I am thankful for our cats, Biff and Buster, and for their occasional emotional support.

Contents

List of Tables

List of Figures

Abstract

Novel path integral–based frameworks for efficient solutions to problems in prediction, nonlinear filtering, and optimal control of stochastic dynamical systems are presented in this dissertation. Namely, (1) the transformed path integral (TPI) approach for solution of the Fokker-Planck equation in stochastic dynamical systems with a full rank diffusion coefficient matrix, (2) the generalized transformed path integral (GTPI) approach—a non-trivial extension of the TPI to stochastic dynamical systems with rank deficient diffusion coefficient matrices, (3) the generalized transformed path integral filter (GTPIF) for solution of the nonlinear filtering problem, and (4) the generalized transformed path integral control (GTPIC) for solution of a large class of stochastic optimal control problems are presented. The frameworks are based on dynamic transformations of the state variables that ensure the appropriate distributions in the transformed space (state distributions in TPI and GTPI; and corresponding conditional distributions in GTPIF and GTPIC) always have zero mean and identity covariance. In systems where the dynamics are linear with respect to the state variables and initial distribution is Gaussian, the appropriate distributions in the transformed space remain invariant with a standard normal distribution. The frameworks thus allow for the underlying distributions necessary for evaluating the quantities of interest to be accurately represented and evolved in a transformed computational domain. In particular, compared to conventional fixed grid approaches and Monte Carlo simulations, the challenges in dynamical systems with large drift, diffusion, and concentration of PDF can be addressed more efficiently using the proposed frameworks. In addition, straightforward error bounds for the underlying distributions in the transformed space can be established via Chebyshev's inequality.

In each of these frameworks, novel short-time propagators for the evolution of distri-

butions in the transformed space are developed. Additionally, necessary update equations for the mean and covariance of the distribution in the original space needed for the evolution of the distribution in the transformed space are derived from the underlying stochastic models. In addition to the normalization condition to preserve the zeroth moment property, conditions to preserve the first and second moment properties of the distributions in the transformed space are also established. The set of update equations in TPI and GTPI approaches present an efficient solution for the Fokker-Planck equation. In the case of the GTPIF, they are an efficient solution of the Zakai equation. And the GTPIC is an efficient solution for the corresponding stochastic Hamilton-Jacobi-Bellman equation. The GTPI approach (which is also used in GTPIF and GTPIC) is applicable to (a) second order stochastic dynamical systems, (b) stochastic dynamical systems with zero process noise, (c) certain stochastic dynamical systems with non-white noise excitation, as well as (d) systems with a full rank diffusion coefficient matrix (where the TPI method is recovered). In the case of (b) stochastic dynamical systems with zero process noise, the GTPI approach represents a solution to the corresponding Liouville equation.

The benefits of these frameworks over those based on conventional fixed grid approaches and Monte-Carlo simulations are illustrated using representative examples in one-dimensional and two-dimensional spaces for the respective problems. Influence of the system parameters and simulation parameters on the error in the TPI and GTPI approaches are also studied. The extensions of the proposed frameworks to more general problems as well as recommendations to address their limitations are discussed.

CHAPTER 1

Introduction

1.1 Overview

The study of stochastic systems, their characterization, and control are of great importance to many fields in science and engineering [1–4]. This often involves obtaining accurate estimates of *quantities of interest* such as the system state or the expected cost in nonlinear dynamical systems subjected to random forces. The nonlinearity here refers to the nonlinear dependence of the system dynamics and the random forcing parameters on the state variables. Such systems can be encountered, for instance, in the field of structural dynamics where loads can often be thought of as random processes and we may need accurate estimates of the displacements due to these loads. In flutter analysis, the pressure fluctuations due to a turbulent boundary layer can be considered as random forces [5], while the excitation from waves can be considered as random forces in studying the effect of irregular waves on the rolling motion of a ship [6]. A system model with nonlinear dynamics is used in tracking of a falling body, i.e., the estimation of its altitude, velocity, and constant ballistic coefficient using noisy radar measurements taken at discrete instants [7]. Similarly, in the distributed control of a team of cooperative unmanned aerial vehicles, the motion of each vehicle may be described using a nonlinear dynamical system model [8].

Under the assumption of a broadband delta correlated random excitation, the response of these systems can very often be accurately described by applying the theory of Markov processes, i.e., using an Itô stochastic differential equation (SDE). The equation describes the evolution of the state of a system along a deterministic path specified by a drift vector function with deviations from the path specified by a diffusion coefficient matrix. Alterna-

tively, one can consider the corresponding Fokker-Planck equation (FPE) [1] which describes the time evolution of the state distribution and allow us to compute averages of the involved stochastic quantities in the system.

Solutions to problems in nonlinear filtering, which involve the estimation of the state of a stochastic dynamical system from noisy observations, require us to evaluate the conditional state distribution conditioned on the observation history. The evolution of this conditional distribution is governed by the Kushner equation—a stochastic partial differential equation [3]. A closed form solution to the equation exists in the case of a linear dynamical system, namely, the Kalman-Bucy filter [9]. However, the Kushner equation for the general nonlinear filtering problem is quite hard to solve in spite of recent advances [10–12]. Fortunately, optimal estimates of the conditional distribution for the general continuous-discrete nonlinear filtering problem, where the dynamical system is represented by a continuous stochastic process and measurements are taken at discrete time instants, can be obtained by solving the Fokker-Planck equation, coupled with a Bayesian-update rule [3].

Likewise, the optimal control of a stochastic dynamical system require us to solve the stochastic Hamilton-Jacobi-Bellman equation—a nonlinear partial differential equation. In a large class of stochastic optimal control problems, where the system dynamics are linear and the cost function is quadratic with respect to the control variables, the solution maybe obtained by solving an equation of the form of a backward Kolmogorov equation [13]. Equivalently, one may solve its adjoint equation which is of the form of a Fokker-Planck equation. Thus, the solution to many problems in the estimation, prediction, and control of stochastic dynamical systems may be obtained via the solution of a corresponding Fokker-Planck equation.

1.2 Solutions of the Fokker-Planck Equation

Exact analytical solutions of the FPE exist only for a few special cases [14–20]. In most other cases, approximate solutions based on either analytical methods or numerical meth-

ods need to be found. The equivalent linearization method [21, 22], stochastic averaging method [23, 24], eigenfunction expansion [25–27], and method of matrix continued fractions [28] are some of the approximate analytical methods. Among the numerical methods, a commonly used method is the numerical integration of the mathematically equivalent stochastic differential equation (SDE) using Monte Carlo (MC) techniques [29, 30]. MC simulations are particularly useful when information is needed only about certain averaged statistical properties of the system. It also finds ease of application in multidimensional problems. However, the MC approach faces challenges due to the inherent sampling errors leading to inaccurate representation of the PDF tail information [31]. The effect of sampling errors on the accuracy of estimates maybe reduced through variance reduction techniques [32].

Grid based methods, where sufficient representation of the PDF tail information can be enforced, are devoid of sampling errors and hence offer benefits over MC methods. The finite difference (FD) and finite element (FE) methods are among the commonly used grid based methods [33–35]. Although these methods have the potential for a better representation of the tail information than MC, they can be computationally expensive for applications in multidimensional problems. For systems with large drift (as compared to the effects of diffusion), traditional implementations of FD and FE methods give invalid solutions where the probability values can become negative and may develop into oscillatory behavior. Other challenges are also presented by the use of fixed grids in the conventional implementations of FD and FE methods. Often these fixed grids are insufficient in accurately representing the transient behavior of the PDF, especially in cases where the system contains large drift, large diffusion and/or highly concentrated regions of the PDF. In other words, a conventional fixed grid based computational domain may lead to inaccurate results in the case of large drift when the PDF peaks approach the domain boundaries. Similarly, it is possible that in the case of large diffusion, the growth of the PDF in the tail regions may not be accurately captured as the boundaries of the computational domain are fixed. Additionally, when the PDFs tend to get highly concentrated such as in cases with low diffusion and stable fixed

points of the corresponding deterministic dynamical system, the resolution of the fixed grid based computational domain might be insufficient to accurately represent the PDF. Adaptive methods such as moving finite elements [36] and adaptive mesh refinement [37] can address some of these issues.

The direct quadrature method of moments (DQMOM) is another adaptive method that can address several of the issues presented above [38, 39]. The method involves the dynamic representation of the state conditional PDF as a finite weighted sum of Dirac delta functions. The weights and locations, which are functions of time, are determined based on constraints imposed by the evolution of the moments. However, for nonlinear systems, there is an issue of closure where the constraints involve higher order moments that are not evolved through the moment evolution equations obtained from the underlying system. Thus, these moments are not always representative of the statistics of the underlying PDF or the propagator in the short-time limit. The path integral (PI) approach based on the short-time propagator does not suffer from such closure approximation requirements. This is owing to the fact that, in the short-time limit, the PI based propagator is an exact solution to the FPE [1]. Thus PI methods ensure that certain statistical properties of the PDF are preserved without having to explicitly impose them. Another advantage is that the approximate solutions obtained from the method are inherently centered around the deterministic path.

Path integrals, originally introduced by Weiner [40] in the context of Brownian motion and reinvented later by Feynman [41] in the reformulation of quantum mechanics, have found wide appeal as a theoretical tool in diverse areas from physics to finance [42]. In path integrals, the contributions of each path weighted by the probability density function of the initial state is integrated over all paths from the initial state to the final state. Several analytical forms for the integral were proposed and investigated for different cases [43–45]. However, it has been shown [46] that such continuous forms are meaningless without the associated discretization prescription. In spite of several attractive properties such as numerical stability and the possibility for the accurate capture of the tail information, the

path integral as a numerical tool for the solution of the FPE has been, till recently, relatively ignored.

One of the earliest numerical implementations of the PI method utilized a histogram based approximation for the PDF on a grid [47]. Similar approaches involving the numerical integration of path integrals using cubic B-splines [48–50] and Gauss-Legendre schemes [51, 52] were also proposed. These methods however inherit the issues associated with a fixed grid (FG) representation, especially those pertaining to cases with large drift or large diffusion. Another issue is the curse of dimensionality, which poses challenges to the solution of multidimensional problems. A meshless implementation based on numerical integration over sampled paths was investigated by Kappen [13]. Such an approach addresses some of the issues arising from the curse of dimensionality, however it still suffers from the shortcomings due to sampling errors. More recently, a method based on variational calculus was proposed [53–55] where the contributions from paths other than the most likely path are neglected thereby reducing the computational complexity involved in the evaluation of the path integral. The method is more efficient than other conventional methods but the inherent approximations involved in neglecting the contributions of certain paths as described above, especially for processes with large diffusion, contribute to a loss in accuracy in the solution and pose challenges to the applicability of the method to a diverse class of problems.

1.3 The Transformed Path Integral Approach

In this dissertation, a novel approach referred to as the transformed path integral method (TPI), for solving the Fokker-Planck equation is proposed. In Chapter 2, we restrict our attention to a class of stochastic dynamical systems with a full rank diffusion coefficient matrix (referred to as non-singular systems). Extensions to systems with a rank deficient diffusion coefficient matrix are discussed in Chapter 3. In the proposed TPI method for non-singular systems, a new formulation for the short-time propagator is developed that allows for the propagation of the state PDF to be performed in a transformed computational

domain where a more accurate representation of the distribution can be achieved. The novel form of the propagator, developed through a dynamic transformation of the state space involving the mean and covariance of the distribution as parameters, also inherits salient features of a conventional PI based propagator, including preservation of important properties of the underlying stochastic process. One such property is non-negativity of the distributions, which is ensured by the propagator form without it having to be explicitly specified. Additionally, in the short-time limit, the propagator is an analytical solution of the corresponding Fokker-Planck equation in the transformed computational domain. In particular, for linear drift processes with a constant diffusion coefficient and Gaussian initial distribution, the TPI method results in a PDF that remains an invariant standard normal distribution under propagation in the associated transformed computational domain.

A unique feature of the generated transformed computational domain is that a fixed grid in this space corresponds to an adaptive grid that is dynamically centered at the mean and scaled with the covariance of the state variables in the original state space. Thus the adaptive grid is non-stationary and is translated with the mean enabling the proposed method to tackle challenges arising from large drift processes such as ensuring accurate and sufficient representation without loss of efficiency (see Fig. 1.1). Similarly, the dynamic scaling of the domain bounds of the discretized state space in the TPI method accommodates growth of the PDF in the tail regions leading to better representation of the tail information especially in large diffusion processes (see Fig. 1.2). In addition, the method provides for an improved grid resolution that can address challenges in accurately representing concentrated PDFs with large peak values, which may arise in small noise processes with stable fixed points in the corresponding deterministic dynamical system. These features contribute to improved accuracy in representation of the PDF and especially that of the tail information with the TPI method. This is additionally facilitated by the possibility to estimate error bounds on the representation of probability in the transformed space through Chebyshev's inequality. As a part of the proposed method, along with the transformed propagator, we derive a set

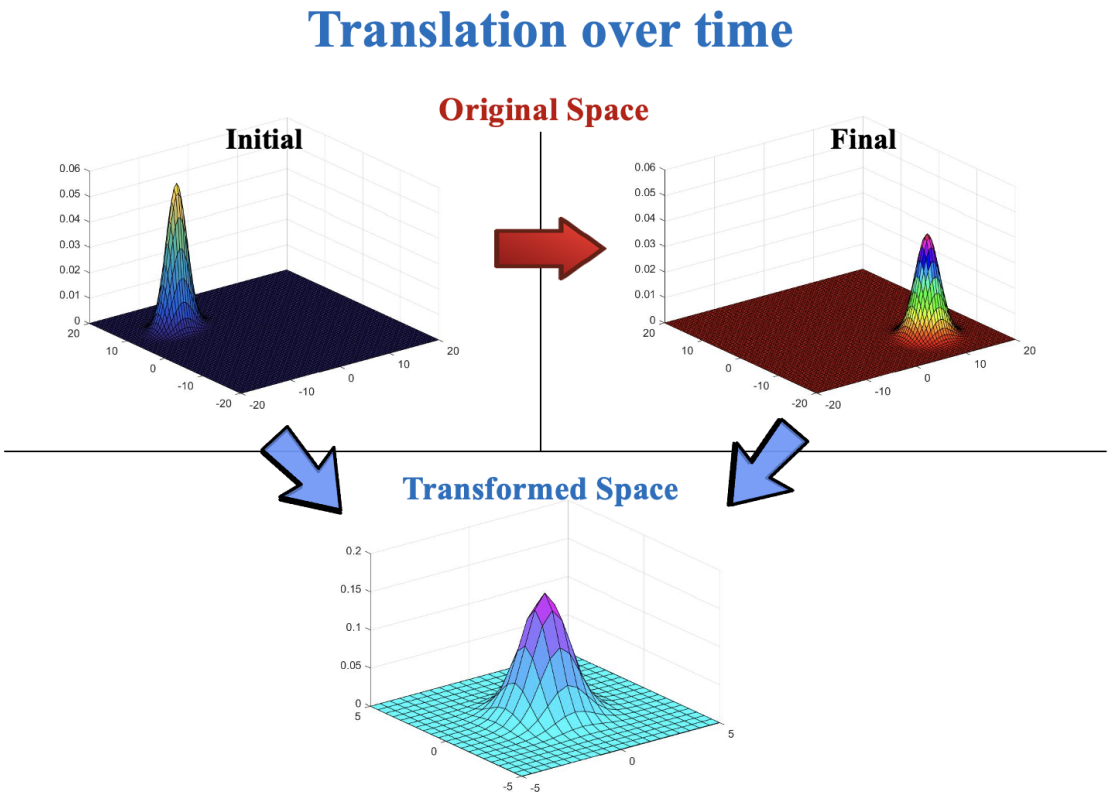

Figure 1.1: Illustration of some features and benefits of the proposed transformed computational domain in addressing challenges arising in prediction of stochastic dynamical systems with large drift.

of update equations for the mean, covariance, and the PDF of the state variables in the transformed space based on the properties of the underlying stochastic difference equation. These update equations, not too unlike those proposed in the unscented transform method (UT) [56], go further in their ability to deal with processes involving nonlinear and non-Gaussian behavior.

However, the TPI approach also inherits a limitation of conventional grid-based PI approaches—the applicability of the approach is restricted to stochastic dynamical systems with a full rank diffusion coefficient matrix. Stochastic dynamical systems with singular diffusion coefficient matrices, such as in second order dynamical systems [57, 58] and certain systems with non-white noise excitation [59–63], present challenges to the application of conventional lattice-based path integral implementations. The challenges, which are also present

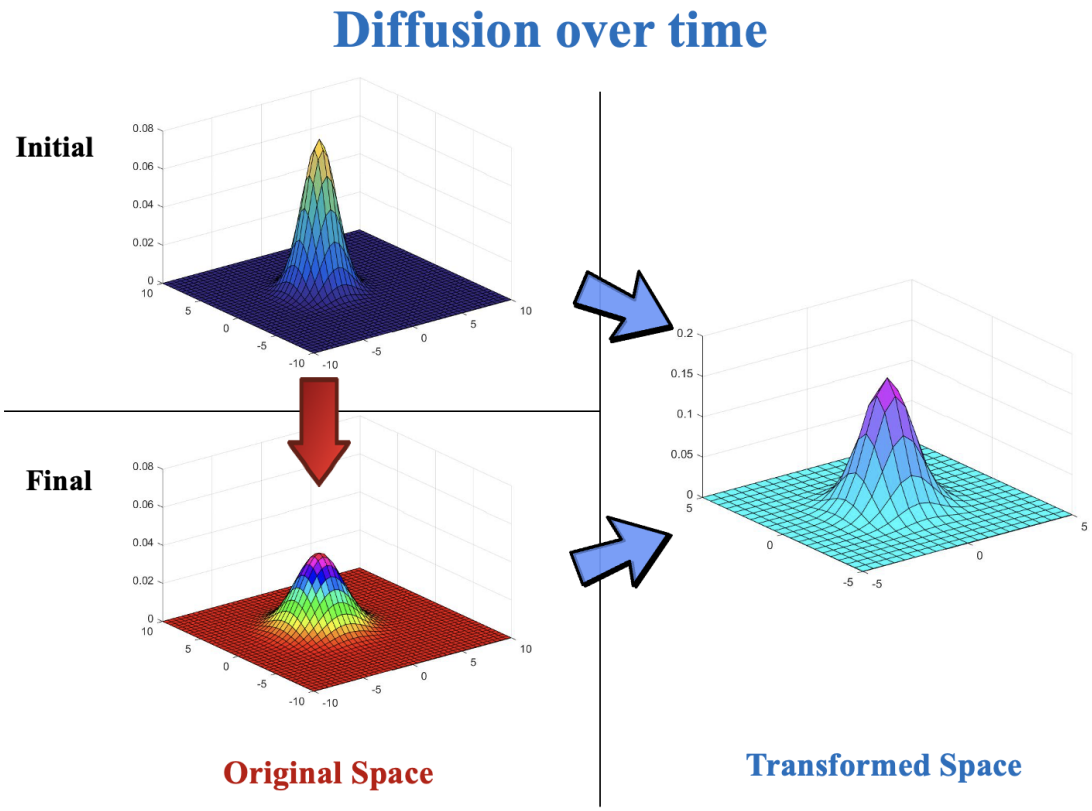

Figure 1.2: Illustration of some features and benefits of the proposed transformed computational domain in addressing challenges arising in prediction of stochastic dynamical systems with large diffusion.

in the TPI method, arise because these diffusion coefficient matrices are non-invertible. A few scholars [64–66] explicitly studied the challenges posed by these "singular systems". However, their proposed solutions inherit the limitations of conventional lattice-based implementations, such as those pertaining to large drift, diffusion, and concentration of PDF. Thus, in addition to TPI, we also propose a transformed path integral-based approach that can address the aforementioned challenges and is applicable to both singular and non-singular systems.

1.4 The Generalized Transformed Path Integral Approach

In order to generalize the TPI approach to applications for both singular and non-singular systems, the generalized transformed path integral (GTPI) approach is proposed in Chap-

ter 3. The proposed GTPI approach, which preserves the salient features of the TPI approach, utilizes the Trotter product formula to split the transformed Fokker-Planck equation, i.e., the Fokker-Planck equation in the transformed space coordinates. The transformed FPE describes the evolution of a complementary stochastic dynamical system obtained through the dynamic transformation of the state variables described earlier. We refer to this system as the standard transformed stochastic dynamical system. The state mean and covariance of the transformed system do not change with evolution, and the choice of our transformation parameters ensure that they are zero and identity respectively. In fact, in dynamical systems with a linear drift function (with respect to the state variables) and a constant diffusion coefficient, the standard normal distribution is a stationary solution for the standard transformed stochastic dynamical system; this feature is also the theoretical underpinning for the capture of invariant solutions mentioned earlier in connection with the TPI approach. Thus, the GTPI, like the TPI, allows for PDF propagation to be performed in a "well contained" transformed computational domain where a more accurate representation of evolved distributions can be achieved. The benefits of the TPI approach in addressing challenges posed by systems with large drift, diffusion, or concentration of PDF are also preserved in the GTPI approach. Accurate error bounds on the distributions via Chebyshev's inequality can also be established. In addition to the renormalization condition to preserve zeroth moment properties, numerical implementations of our approach allows us to establish conditions to preserve the first and second moment properties of the state distribution in the transformed space.

Our splitting scheme in the GTPI approach, which is designed to address the challenges of rank deficient diffusion coefficient matrices, consists of splitting the corresponding transformed FPE into (1) a Liouville operator and (2) a Fokker-Planck operator with a full rank diffusion coefficient sub-matrix. In effect, the former governs contributions from "singular dimensions" while the latter concerns the "nonsingular dimensions". We consider a solution based on method of characteristics for the Liouville equation; A TPI type short-time

propagator is considered for the Fokker-Planck equation with the sub-matrix. Additionally, update equations for mean and covariance of the state variables, which are used for PDF propagation in the transformed space, are developed from the underlying stochastic system models.

The new set of update equations allow us to consider solutions of (a) second order dynamical systems, (b) dynamical systems with zero process noise, (c) certain dynamical systems with non-white noise excitation, and (d) systems with a full rank diffusion coefficient matrix (where we recover the TPI method). Note that for the case (b) the corresponding set of update equations represent a solution for the Liouville equation. The GTPI approach presents a non-trivial extension of the applicability of TPI approach to a more general class of stochastic dynamical systems listed above. Thus, the GTPI approach allows us to consider transformed path integral-based frameworks for the problems of nonlinear filtering (Chapter 4) and stochastic optimal control (Chapter 5) in systems belonging to this general class of stochastic dynamical systems.

1.5 The Generalized Transformed Path Integral Filter

The nonlinear filtering problem, i.e., the estimation of the state of a stochastic dynamical system based on noisy observations, is encountered in a wide variety of applications from different fields—for instance, the tracking of space objects [67], navigation system design for autonomous aircraft landing [68], and data assimilation for weather forecasting applications [69]. The nonlinearity refers to the nonlinear dependence of the system dynamics and/or the observation function on the state variables. In the probabilistic approach to nonlinear filtering [70, 71], the solution requires accurate estimates of the conditional state distribution conditioned on the measurement observation history. The evolution of this conditional distribution is governed by the kushner equation—a nonlinear stochastic partial differential equation. A more tractable bilinear stochastic partial differential equation, the Zakai equation, governs the evolution of the un-normalized conditional density function [72].

As discussed earlier, the solutions to the general linear filtering problem are laid out in the pioneering works by Kalman and Bucy [9, 73]. The extended Kalman filter (EKF) is a sub-optimal filter that extends the applicability to nonlinear filtering problems via a linearization of the nonlinear dynamics and observation function around the previous state estimates [74]. Although the EKF inherits the salient features of the Kalman filter, the desirability of the approach is adversely affected by the difficulty in deriving the required Jacobian matrices as well as filter instability arising from inaccuracies in the linearized approximations [75]. The particle filter approach, where the distribution is expressed as a set of evolved weighted samples is better equipped to address the challenges arising from nonlinear dynamics and non-Gaussian distributions [76]. However, the standard particle filter approaches may need large number of samples to counter particle degeneracy where the contributions of all but one particle become negligible. The particle flow filter [77] attempts to address the issue of particle degeneracy by ensuring that there are enough particles in the regions of importance.

In contrast to particle filters and other Monte-Carlo based approaches, the unscented Kalman filter (UKF) employs deterministically chosen points that exhibit certain specific properties (e.g., have a given mean and covariance) to represent the distribution [78]. Thus, the UKF can be used to overcome limitations of EKF at lower computational costs compared to particle filters. The UKF is designed to preserve the first and second moments of the distribution with propagation; Variations of the approach to preserve higher-order moments is an area of active research. Approaches which attempt to preserve the complete statistics of the underlying distributions with propagation may be more accurate than UKF, especially in cases with a high degree of nonlinearity or non-Gaussianity of distributions. One such approach involves the use of a path integral–based short-time propagator to solve the Zakai equation [10] on a fixed computational grid. Another approach involves the solution of a Fokker-Planck equation using a finite difference model on an adaptive grid along with a Bayesian rule for the measurement update [79]. The former approach faces challenges in application to systems with large drift, diffusion, and concentration of PDF. While the

latter is better equipped to address some of these challenges, it inherits other issues of finite difference based approaches pertaining to preservation of non-negativity of the distributions and oscillatory behavior in systems with large drift.

The generalized transformed path integral filter (GTPIF) is a transformed path integral-based framework for the continuous-discrete nonlinear filtering problem (see Chapter 4). As discussed earlier, the continuous-discrete nonlinear filtering problem may be solved through a two step process—namely, (1) System update via the solution of a Fokker-Planck equation and (2) Measurement update using Bayes' theorem. The basis for the GTPIF approach is a dynamical transformation of the state variables using the mean and covariance of the conditional distribution as parameters. Consequently, the evolution of the conditional distribution of the transformed state variables are governed by the corresponding transformed Fokker-Planck equation between observations. A GTPI-like short-time propagator is used to evolve this distribution while the conditional mean and covariance are evolved using the corresponding update equations.

At an observation, a new Bayes' update rule for the conditional distribution of the transformed state variables is presented. Additionally, the necessary new measurement update equations for the conditional mean and covariance are also derived. The new set of system update and measurement update equations together constitute a solution of the Kushner equation in the transformed space variables. They are better equipped to handle the challenges from systems with large drift, diffusion, and concentration of PDF. In fact, we recover the Kalman filter in the case of a linear dynamical system and observation function with Gaussian initial condition. However, the GTPIF is also better equipped than the EKF to accommodate nonlinearities and non-Gaussian distributions.

In contrast to the UKF, the GTPIF is better equipped to preserve the higher-order moments of the distribution. The GTPIF is not affected by the issues of particle degeneracy or sample impoverishment present in particle filters nor is it affected by the need to explicitly enforce the non-negativity of evolved distributions. Unlike other grid-based filters, the grid

points in the GTPIF are located in "regions of importance" (which can be estimated via a Chebyshev inequality). This allows GTPIF to offset some of the computational costs associated with a grid-based approach for filtering problems.

1.6 The Generalized Transformed Path Integral Control

The objective of determining the best strategy to control a dynamical system in the presence of uncertainties is common to problems in a wide range of fields—control of batch distillation processes in chemical engineering [80], trajectory optimization for low-thrust space missions [81], and reinforcement learning algorithms in robotics [82], to name a few. Often, a set of controls are sought that maximize a measure of performance or minimize an expected cost function. The optimized cost function satisfies the stochastic Hamilton-Jacobi-Bellman (HJB) equation—a nonlinear partial differential equation to be solved backwards in time [4]. The stochastic HJB equation can be reduced to a more tractable form for a large group of stochastic optimal control problems, namely, those with a linear system dynamics and quadratic cost function with respect to the control variables. It is to be noted that the dynamics and cost may still be nonlinear with respect to the state variables.

For these cases, the solution can be obtained by solving an equation of the form of a backward Kolmogorov equation (BKE) [1, 83] which is solved backwards in time. Equivalently, one may solve its adjoint equation, namely, a Fokker-Planck–type equation forward in time. The equation governs the evolution of the conditional state distribution conditioned on the initial state for a diffusion process that occurs alongside a killing process. An approach to this class of control problems involves the solution of this Fokker-Planck–type equation using sample based approximations of path integrals [83]. The approach is shown to be more efficient than conventional gradient based approaches. However, the approach may inherit issues such as sampling errors. Such issues maybe addressed through a grid based approach, but the approach would need to address issues pertaining to large drift, diffusion, and concentration of PDF.

The generalized transformed path integral control (GTPIC) is a transformed path integral–based approach for the stochastic optimal control problem (see Chapter 5). The GTPIC employs a dynamic transformation of the state space with the mean and covariance of the conditional distribution as parameters. Under this transformation, a novel short-time propagator for the evolution of the conditional distribution in the transformed space is developed. Additionally, update equations for the conditional mean and covariance are also presented. The GTPIC, in contrast to conventional grid based approaches, is better equipped to address challenges from large drift, diffusion, and concentration of PDF.

1.7 Selected Examples and Outline

The benefits of the proposed frameworks over conventional approaches were showcased using illustrative examples in one-dimensional and multi-dimensional spaces. The performance of the TPI approach was shown in systems with full rank diffusion coefficient matrices—(1) linear dynamical systems such as the pure diffusion process, constant drift process, and Ornstein Uhlenbeck process; as well as (2) nonlinear dynamical systems with cubic nonlinearities. Likewise, the performance of the GTPI approach was showcased in systems with rank deficient diffusion coefficient matrices, namely, (a) the stochastic harmonic oscillator, (b) the stochastic van der Pol oscillator, (c) the stochastic Caughey oscillator, (d) the stochastic Duffing oscillator with zero process noise, and (e) bistable stochastic flow driven by non-white noise. Note that the evolution of the stochastic Duffing oscillator with zero process noise (example (d)) is governed by a Liouville equation. A bistable stochastic flow, i.e., a system with nonlinear dynamics, driven by white noise and a linear measurement function was chosen to showcase benefits of the GTPIF over the conventional extended Kalman filter. The performance of the GTPIC and its benefits over conventional approaches are illustrated using (i) the controlled pure diffusion process and (ii) the control of a bistable stochastic flow driven by white noise.

The outline for the rest of the dissertation is as follows. The transformed path inte-

gral (TPI) approach and its application to non-singular systems is discussed in Chapter 2. The generalization of the TPI approach to both singular and non-singular systems via the generalized transformed path integral (GTPI) approach along with selected examples are presented in Chapter 3. Extensions of the GTPI approach to problems in nonlinear filtering (via a novel formulation, GTPIF) and stochastic optimal control (via a novel formulation, GTPIC) are presented in Chapters 4 and 5 respectively. Concluding remarks along with directions for future work are discussed in Chapter 6.

CHAPTER 2

The Transformed Path Integral Approach for Stochastic Processes

2.1 Scope of the Chapter

In this chapter, a novel path integral (PI) based approach for the solution of the Fokker-Planck equation is presented. The proposed approach, termed the transformed path integral (TPI) approach, utilizes a new formulation for the underlying short-time propagator to perform the evolution of the probability density function (PDF) in a transformed computational domain where a more accurate representation of the PDF can be ensured. The new formulation, based on a dynamic transformation of the original state space with the statistics of the PDF as parameters, preserves the non-negativity of the PDF and incorporates short-time properties of the underlying stochastic process. New update equations for the state PDF in a transformed space and the parameters of the transformation (including mean and covariance) that better accommodate the nonlinearities in drift and non-Gaussian behavior in distributions are proposed (based on properties of the SDE). Owing to the choice of the transformation considered, the proposed approach maps a fixed grid in transformed space to a dynamically adaptive grid in the original state space. The TPI approach, in contrast to conventional approaches such as Monte Carlo simulations and fixed grid based approaches, is able to better represent the distributions (especially the tail information) and better address challenges in processes with large diffusion, large drift, and large concentration of PDF. Additionally, in the proposed TPI approach, error bounds on the probability in the computational domain can be obtained using the Chebyshev's inequality. The benefits of the TPI approach over conventional approaches are illustrated through simulations of linear and nonlinear drift processes in one-dimensional and multidimensional state spaces. The effects

of spatial and temporal grid resolutions as well as that of the diffusion coefficient on the error in the PDF are also characterized.

2.2 Mathematical Formulations

In this section the mathematical formulation behind the proposed TPI approach and its derivation are presented. Starting with a brief introduction to the Wiener path integral, the form of the accompanying short-time propagator for the FPE is provided. A new form of the short-time propagator is then developed for single degree of freedom (SDOF) systems based on a time dependent transformation involving the statistics of the distribution and is then later extended to systems in multiple dimensions (MDOF case) as well. A set of associated update equations for the PDF, the mean and covariance are also derived. The invariance of the PDF from a standard normal distribution in the transformed space for linear drift and diffusion processes with a Gaussian initial PDF is briefly discussed along with an examination of the estimation of the PDF error bounds based on the Chebyshev's inequality. A discrete representation of the proposed formalism is also provided in this section.

2.2.1 Path integral formalism

A stochastic process with the Markov assumption can be expressed as an Itô stochastic differential equation (SDE) [3] as follows

$$\mathrm{d}\mathbf{x}(t) = \mathbf{f}\Big(\mathbf{x}(t), t\Big)\,\mathrm{d}t + \mathbf{A}\Big(\mathbf{x}(t), t\Big)\,\mathrm{d}\mathbf{w}(t) \tag{2.1}$$

where $\mathbf{x}(t) \in \mathbb{R}^{N_s \times 1}$ represents the state of the system at time t. The equation describes evolution of the state of a system along a deterministic path specified by the drift vector function $\mathbf{f}\Big(\mathbf{x}(t), t\Big) \in \mathbb{R}^{N_s \times 1}$ while being subjected to random excitations modeled as Gaussian white noise of strength $\mathbf{A}\Big(\mathbf{x}(t), t\Big) \in \mathbb{R}^{N_s \times N_w}$. The increments $\mathrm{d}\mathbf{w}(t) \in \mathbb{R}^{N_w \times 1}$ are independent and identically distributed zero mean Gaussian random vectors with the auto-

correlation $\left\langle \mathrm{d}w_i(t)\,\mathrm{d}w_j(s)\right\rangle = \delta(t-s)\,\delta_{ij}\,\mathrm{d}t$ for $i,j=1,\ldots,N_w$. The stochastic process $\mathbf{w}(t)$ is the Wiener process.

The state of the system can also be charaterized by a probability density function $p(\mathbf{x},t)$ whose evolution is governed by the Fokker-Planck equation (FPE), a second order partial differential equation of the parabolic type given by

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial x_i} f_i(\mathbf{x},t) - \frac{\partial^2}{\partial x_i \partial x_j} G_{ij}(\mathbf{x},t)\right] p(\mathbf{x},t) = 0 \tag{2.2}$$

where we have used Einstein's notation convention, i.e., repeated indices imply summation. Note that the spatial variables $\mathbf{x}$ used in Eq. (2.2) is not the same as the random vector $\mathbf{x}(t)$ representing the state of the dynamical system in Eq. (2.1). Nevertheless, the Fokker-Planck equation features the drift vector function $\mathbf{f}(\mathbf{x},t)$ while the strength of the white noise excitation $\mathbf{A}(\mathbf{x},t)$ is related to the diffusion coefficient matrix $\mathbf{G}(\mathbf{x},t)$ as $\mathbf{G}(\mathbf{x},t) = \mathbf{A}(\mathbf{x},t)\mathbf{A}(\mathbf{x},t)^{\mathrm{T}}/2$. For the sake of brevity, wherever the context is clear we will represent both the state variables and their associated spatial variables with the same symbol. Although a general analytical solution for Eq. (2.2) does not currently exist, it can be shown [1] for small time differences $\Delta t = t' - t$ the transition PDF $p(\mathbf{x}', t' \,|\, \mathbf{x}, t)$ is given by

$$\begin{aligned} p(\mathbf{x}', t' \,|\, \mathbf{x}, t) = \|4\pi\Delta t\, \mathbf{G}(\mathbf{x},t)\|^{-1/2} \\ \exp\left\{-\frac{1}{4\,\Delta t}\left[\mathbf{x}_\mathbf{e}^{\mathrm{T}}\, \mathbf{G}(\mathbf{x},t)^{-1}\, \mathbf{x}_\mathbf{e}\right]\right\} \end{aligned} \tag{2.3}$$

with $\mathbf{x}_\mathbf{e} = \mathbf{x}' - \mathbf{x} - \mathbf{f}(\mathbf{x},t)\,\Delta t$. The relation in Eq. (2.3) is also known as the short-time propagator. Based on the Chapman-Kolmogorov equation, the time evolution of the PDF is performed by the equation

$$p(\mathbf{x}', t') = \int p(\mathbf{x}', t' \,|\, \mathbf{x}, t)\, p(\mathbf{x}, t)\, \mathrm{d}\mathbf{x} \tag{2.4}$$

where the integral is over the entire domain of $\mathbf{x}$. With the repeated application of Eq. (2.4)

we have

$$p(\mathbf{x},t) = \lim_{N\to\infty} \underbrace{\int\cdots\int}_{N \text{ times}} \prod_{k=0}^{N-1} \left(\mu_k \, \mathrm{d}\mathbf{x_k}\right) \exp\left\{-\Delta t \sum_{k=0}^{N-1} \mathcal{L}(\mathbf{x_{k+1}}, \mathbf{x_k}, t_k, \Delta t)\right\} p(\mathbf{x_0}, t_0) \tag{2.5}$$

where $\mu_k = \|4\pi\,\Delta t\,\mathbf{G}(\mathbf{x_k}, t_k)\|^{-1/2}$ and

$$\mathcal{L}(\mathbf{x_{k+1}}, \mathbf{x_k}, t_k, \Delta t) = \Big[\mathbf{x_{k+1}} - \mathbf{x_k} - \mathbf{f}(\mathbf{x_k}, t_k)\,\Delta t\Big]^{\mathrm{T}} \mathbf{G}(\mathbf{x_k}, t_k)^{-1} \Big[\mathbf{x_{k+1}} - \mathbf{x_k} - \mathbf{f}(\mathbf{x_k}, t_k)\,\Delta t\Big] \tag{2.6}$$

with $\mathbf{x_N} = \mathbf{x}$. Eq. (2.5) is an integral over all possible paths from $\mathbf{x_0}$ to $\mathbf{x}$ and is called the Weiner path integral. Analytical solutions to Eq. (2.5) exist only for few special cases. Alternatively, $p(\mathbf{x}, t)$ can be obtained iteratively using Eq. (2.3) and Eq. (2.4). The numerical solution for the latter involves the evaluation of a convolution integral which poses significant computational challenges especially in high dimensional state spaces. Several attempts have been made to consider a discrete representation of the PDF in order to reduce the computational complexity [47, 48]. They reduce the convolution operation to a matrix vector multiplication. However, these approaches do not account for the challenges posed by processes with large drift, diffusion, and concentration of PDF.

2.2.2 Transformed path integral approach for SDOF systems

We develop here the equations in one-dimensional state space for the proposed (TPI) approach. These equations, as we shall see later, can be easily extended to their corresponding forms in the MDOF case. Let us consider a one-dimensional stochastic process given by the Itô SDE

$$\mathrm{d}x(t) = f\Big(x(t), t\Big)\,\mathrm{d}t + a\Big(x(t), t\Big)\,\mathrm{d}w(t) \tag{2.7}$$

where $w(t)$ is a scalar Brownian motion process with $\Big\langle \mathrm{d}w(t)\,\mathrm{d}w(s)^{\mathrm{T}} \Big\rangle = \delta(t-s)\,\mathrm{d}t$. Eq. (2.7) can be equivalently written as a difference equation

$$x(t+\mathrm{d}t) = x(t) + f\Big(x(t),t\Big)\,\mathrm{d}t + a\Big(x(t),t\Big)\,\mathrm{d}w(t). \tag{2.8}$$

We propose the following time dependent transformation

$$z(t) = \mathcal{Z}\Big(x(t),t\Big) = \frac{x(t)-\mu(t)}{\sigma(t)} \tag{2.9}$$

where $\mu(t)$ and $\sigma(t)$ are respectively the mean and the standard deviation of $x(t)$ at time t. Applying this transformation to Eq. (2.8), we get

$$\begin{aligned} &\sigma(t+\mathrm{d}t)\,z(t+\mathrm{d}t) + \mu(t+\mathrm{d}t) \\ &\qquad = \sigma(t)\,z(t) + \mu(t) + f\Big(x(t),t\Big)\,\mathrm{d}t + a\Big(x(t),t\Big)\,\mathrm{d}w(t) \end{aligned} \tag{2.10}$$

After rearranging, we obtain

$$\begin{aligned} z(t+\mathrm{d}t) = z(t) &+ \left(\frac{\sigma(t)}{\sigma(t+\mathrm{d}t)} - 1\right) z(t) \\ &+ \frac{1}{\sigma(t+\mathrm{d}t)}\delta\tilde{f}\Big(z(t),t\Big)\,\mathrm{d}t + \frac{1}{\sigma(t+\mathrm{d}t)}\tilde{a}\Big(z(t),t\Big)\,\mathrm{d}w(t) \end{aligned} \tag{2.11}$$

where $\tilde{a}\Big(z(t),t\Big) = a\Big(\sigma(t)\,x(t)+\mu(t),t\Big)$, $\delta\tilde{f}\Big(z(t),t\Big) = \delta f\Big(\sigma(t)\,x(t)+\mu(t),t\Big)$ and $\delta f\Big(x(t),t\Big) = f\Big(x(t),t\Big) - \Big\langle f\Big(x(t),t\Big)\Big\rangle$. Thus, the form of the short-time propagator in the transformed space is given by

$$p(z',t' \mid z,t) = \frac{\sigma}{\sqrt{2\pi\,\tilde{a}(z,t)^2\,\Delta t}}\,\exp\left\{-\frac{\Big(\sigma' z' - \sigma z - \delta\tilde{f}(z,t)\,\Delta t\Big)^2}{2\tilde{a}(z,t)^2\,\Delta t}\right\} \tag{2.12}$$

where we have used primed variables to represent the quantities at time t'. Also, $\Delta t = t'-t$. An alternate form of the transformed short-time propagator, to the one proposed here, can

be obtained based on the Itô's lemma as shown in A. The time evolution of the PDF in the transformed space is performed by the equation

$$p(z', t') = \int p(z', t' \mid z, t)\, p(z, t)\, \mathrm{d}z \tag{2.13}$$

Eq. (2.12) along with Eq. (2.13) allow us to evaluate the PDF in the transformed space. The PDFs in the original space and the transformed space are related as follows

$$p_{x(t)}(x, t) = \frac{1}{\sigma(t)}\, p_{z(t)}\left(\mathcal{Z}(x, t),\, t\right) \tag{2.14}$$

However, in order to evaluate the integral in Eq. (2.13) that denotes an update equation for the PDF of the state variables in the transformed space, the mean $\mu(t)$ and variance $\sigma(t)^2$ need to be computed at every time step. The update equations for these quantities can be obtained by taking expectations of Eq. (2.7). We thus have respectively for the mean and variance

$$\mu' = \mu + \Delta t \left\langle \tilde{f}(z, t) \right\rangle \tag{2.15}$$

$$\sigma'^2 = \sigma^2 + \Delta t \left\langle \tilde{a}(z, t)^2 \right\rangle + \Delta t \left\langle \sigma z\, \delta\tilde{f}(z, t) \right\rangle + (\Delta t)^2 \left\langle \delta\tilde{f}(z, t)^2 \right\rangle \tag{2.16}$$

2.2.3 Features of the transformed path integral approach

In the framework of the TPI approach, we can capture certain invariant solutions as described below.

Lemma 2.2.1. *Given the transformed space created by the transformation Eq. (2.9), for an initial distribution $p(z, t) \sim \mathcal{N}(0, 1)$ the form of the PDF remains invariant for all times $t' > t$ under the influence of a stochastic process with linear dynamics and additive white noise excitation wth a constant diffusion coefficient.*

Proof. Considering a process with a linear drift given by $f(x, t) = m\, x + c$ and a constant

diffusion coefficient $g(x,t) = a(x,t)^2/2 = g$, the update equations (Eq. (2.15) and Eq. (2.16)) for the mean and variance, given a Gaussian initial PDF, $p(x,t) \sim \mathcal{N}(\mu, \sigma)$, can be reduced to

$$\mu' = \mu\,(1 + m\,\Delta t) + c\,\Delta t \tag{2.17}$$

$$\sigma'^2 = \sigma^2\,(1 + m\,\Delta t)^2 + 2\,g\,\Delta t \tag{2.18}$$

Using these in Eq. (2.12) and Eq. (2.13) we obtain

$$p(z,t) \sim \mathcal{N}(0,1) \Rightarrow p(z',t') \sim \mathcal{N}(0,1) \quad \forall\, t' > t \tag{2.19}$$

□

The importance of this result is that for processes with linear drift and constant diffusion coefficient, the loss of tail information due to the use of finite computational domains in the transformed space can be estimated.

Given a domain with lower and upper bounds $[x_l, x_u]$, an estimate of the error in the PDF $p(x,t)$ due to the loss of tail information outside of these bounds is given by

$$\varepsilon = 1 - \int_{x_l}^{x_u} p(x,t)\,\mathrm{d}x \tag{2.20}$$

For a Gaussian distribution $x \sim \mathcal{N}(\mu, \sigma^2)$, Eq. (2.20) can be written as

$$\varepsilon = 1 - \frac{1}{2}\left[\mathrm{erf}\left(\frac{x_u - \mu}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{x_l - \mu}{\sqrt{2}\sigma}\right)\right] \tag{2.21}$$

which for the standard normal distribution in a domain with bounds $[-l, l]$ reduces to $\mathrm{erfc}\left(l/\sqrt{2}\right)$. For $l = 3$, we have this error value to be 0.0027. Thus, for a linear drift and constant diffusion process, 99.73 % of the probability in the computational domain can be represented by considering a transformed computational domain with bounds $[-3, 3]$.

This can also be extended to processes where the PDF remains approximately Gaussian, i.e. near-Gaussian distributions, to obtain reliable estimates for the error bounds.

For the general case, estimates of the error bounds can be obtained through Chebyshev's inequality: for a random variable x with finite expected value μ and finite non-zero variance σ^2, for any real number $k > 0$,

$$Pr\Big(|x - \mu| \geq k\sigma\Big) \leq \frac{1}{k^2} \tag{2.22}$$

where the left hand side represents the probability that the random variable x takes a value outside of the domain with bounds k standard deviations on either side of the mean. In the transformed space, the inequality reduces to

$$Pr\Big(|z| \geq k\Big) \leq \frac{1}{k^2} \tag{2.23}$$

Thus for a fixed domain with bounds $[-k, k]$ in the transformed space, which corresponds to an adaptive domain with bounds k standard deviations on either side of the mean in the original space, the probability that the state variable will take a value outside of these bounds has the upper bound $1/k^2$. This probability is given by

$$Pr\Big(|z| \geq k\Big) = 1 - \int_{-k}^{k} p(z, t)\, \mathrm{d}z \tag{2.24}$$

which from Eq. (2.20) is $\varepsilon(-k, k)$. Thus we have from the Chebyshev's inequality

$$\varepsilon \leq \frac{1}{k^2} \tag{2.25}$$

The error in the PDF due to the loss in tail information for a fixed grid in the transformed space has an upper bound. In theory, due to the fixed grid in the transformed space, this upper bound is valid for all times under the propagation. However, in actual implementation,

the growth of discretization error over a large period of time also needs to be accounted for.

2.2.4 Transformed path integral approach for MDOF systems

The formulation presented above can be easily generalized to the MDOF case. Starting from the SDE Eq. (2.1) we consider the following transformation

$$\mathbf{z}(t) = \mathcal{Z}\Big(\mathbf{z}(t), t\Big) = \mathbf{R}(t)^{-1}\Big(\mathbf{x}(t) - \boldsymbol{\mu}(t)\Big) \tag{2.26}$$

where the symbols $\boldsymbol{\mu}(t)$ and $\boldsymbol{\Sigma}(t) \equiv \mathbf{R}(t)\mathbf{R}(t)^{\mathrm{T}}$ denote the mean and covariance matrix of $\mathbf{x}(t)$ at time t respectively. Following the steps laid out earlier, we can arrive at the expression for the transformed short-time propagator

$$p(\mathbf{z}', t' \,|\, \mathbf{z}, t) = \|\mathbf{R}\| \, \Big\|4\pi \, \Delta t \, \tilde{\mathbf{G}}(\mathbf{z}, t)\Big\|^{-1/2} \exp\Big\{-\frac{1}{4\,\Delta t}\Big[\mathbf{z}_{\mathbf{e}}^{\mathrm{T}} \tilde{\mathbf{G}}(\mathbf{z}, t)^{-1} \mathbf{z}_{\mathbf{e}}\Big]\Big\} \tag{2.27}$$

where $\mathbf{z}_{\mathbf{e}} = \mathbf{R}'\mathbf{z}' - \mathbf{R}\mathbf{z} - \mathrm{d}t \, \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)$ and $\boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t) = \tilde{\mathbf{f}}(\mathbf{z}, t) - \Big\langle \tilde{\mathbf{f}}(\mathbf{z}, t) \Big\rangle$. Here $\tilde{\mathbf{f}}$ and $\tilde{\mathbf{G}}$ are functions in the transformed space given by $\tilde{\mathbf{f}}(\mathbf{z}, t) = \mathbf{f}(\mathbf{R}\mathbf{z} + \boldsymbol{\mu}, t)$ and $\tilde{\mathbf{G}}(\mathbf{z}, t) = \mathbf{G}(\mathbf{R}\mathbf{z} + \boldsymbol{\mu}, t)$. The time evolution of the PDF is performed by

$$p(\mathbf{z}, t) = \int p(\mathbf{z}', t' \,|\, \mathbf{z}, t) \, p(\mathbf{z}, t) \, \mathrm{d}\mathbf{z} \tag{2.28}$$

The update equations for the mean and covariance are as follows

$$\boldsymbol{\mu}' = \boldsymbol{\mu} + \Delta t \, \Big\langle \tilde{\mathbf{f}}(\mathbf{z}, t) \Big\rangle \,, \tag{2.29}$$

$$\begin{aligned} \boldsymbol{\Sigma}' = \boldsymbol{\Sigma} &+ 2\,\Delta t \, \Big\langle \tilde{\mathbf{G}}(\mathbf{z}, t) \Big\rangle + \Delta t \, \Big\langle \mathbf{R}\mathbf{z} \, \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)^{\mathrm{T}} \Big\rangle \\ &+ \Delta t \, \Big\langle \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t) \, \mathbf{z}^{\mathrm{T}} \mathbf{R}^{\mathrm{T}} \Big\rangle + (\Delta t)^2 \Big\langle \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t) \, \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)^{\mathrm{T}} \Big\rangle \,. \end{aligned} \tag{2.30}$$

The Chebyshev's inequality expressed in Eq. (2.22) for a random variable can also be easily extended for the MDOF case where for a random vector $\mathbf{x} \in \mathbb{R}^{N_s \times 1}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, for any real number $k > 0$, we have

$$Pr\Big((\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \geq k^2\Big) \leq \frac{N_s}{k^2} \tag{2.31}$$

which in the transformed space reduces to

$$Pr\Big(\mathbf{z}^{\mathrm{T}} \mathbf{z} \geq k^2\Big) \leq \frac{N_s}{k^2} \tag{2.32}$$

2.2.5 Numerical implementation of the transformed path integral approach

The set of Eqs. (2.12), (2.13), (2.15) and (2.16) form the basis of the proposed approach for solving SDOF systems. Similarly, we have Eqs. (2.27) to (2.30) for MDOF systems. The long-time evolution of the PDF, based on the update equations presented above, accounts for the contributions of all possible paths in the transformed space via the transformed short-time propagator. This is analogous to traditional path integral based formulations of Weiner, Feynman and Kac. We thus refer to the novel approach presented here as the *transformed path integral* (TPI) approach.

Lattice-based implementations of path integrals often involve discrete representations of distributions on a grid—for instance, $p_i = p(\mathbf{z_i}, t)$ for $i = 1\ldots, N$. The index i here represents the discretization of the computational domain into, N grid points and should not be confused with the vector indices. Hence, p_i here is the value of $p(\mathbf{z}, t)$ evaluated at the i-th grid point $\mathbf{z_i}$. Under this representation, the update equations for PDF propagation, namely Eq. (2.13) and Eq. (2.28) can be expressed in the form

$$p_i' = \sum_{j=1}^{N} p_{ij}\, p_j\, \Delta z_j \tag{2.33}$$

where Δz_j is the grid spacing in the transformed domain while p_{ij} represents the short-time

propagator in the transformed space whose individual terms are obtained from Eq. (2.12) and Eq. (2.27) for SDOF and MDOF systems respectively. Thus, the convolution operation of the path integral is reduced to a matrix vector multiplication. Similarly, the expectations in the update equations for the mean and covariance (Eqs. (2.15), (2.16), (2.29) and (2.30)) become weighted averages over the grid points of the respective quantities.

The form of the short-time propagator ensures that the terms of the matrix p_{ij} are non-negative. Hence, starting with a non-negative set of values for p_j, we will obtain a non-negative set of values for p'_i under the propagation. This feature ensures that the non-negativity property of the PDF can also be preserved in the discrete implementation of the TPI approach for all times without it having to be explicitly specified. Additionally, the propagator matrix must obey the properties of a Markov transition matrix, namely the following zeroth moment condition.

$$\int p(\mathbf{z}', t' \,|\, \mathbf{z}, t)\, \mathrm{d}\mathbf{z}' = 1 \tag{2.34}$$

This maybe achieved by making the transformation

$$p_{ij}^{(\mathrm{new})} = \frac{p_{ij}^{i(\mathrm{old})}}{\sum_i p_{ij}\, \Delta z_i} \tag{2.35}$$

Similarly, the zeroth moment condition for p_i is enforced by

$$p_i^{(\mathrm{new})} = \frac{p_i^{(\mathrm{old})}}{\sum_i p_i\, \Delta z_i} \tag{2.36}$$

The enforcement of these zeroth moment properties further increases and maintains the high levels of accuracy in the numerical implementations.

2.3 Performance of the Transformed Path Integral Approach

In this section we consider several numerical examples wherein the key aspects that contribute to the superior performance of the TPI approach are highlighted to illustrate the application and benefit of the approach over the conventional fixed grid approaches and MC simulations. In particular, the ability of the TPI approach to overcome the limitations of conventional fixed grid representations such as those due to fixed grid bounds, stationary domains, and a finite fixed grid resolution are showcased through the simulations of specific linear drift processes (in 1D) with known analytical solutions. A zero drift and constant diffusion process is considered in Example 1 to showcase the ability of the TPI approach to overcome the limitations of finite fixed grid bounds. The challenges posed by a static computational domain and the capacity of the TPI approach to tackle them are investigated in Example 2 through simulations of a constant drift and diffusion process. The better performance of the TPI approach for the accurate representation of the PDF in cases where there is large concentration of the PDF in a small region of state space is illustrated by the simulation of the Ornstein-Uhlenbeck (OU) process in Example 3.

The TPI approach, with a set of update equations for the mean and covariance as well as the PDF of the state variable in the transformed space, can better accommodate nonlinearities in the drift without any linearization approximations and more accurately represent the non-Gaussian behavior of PDFs. A typical nonlinear process (in 1D) is provided in Example 4 and the performance of the TPI approach is examined. The benefits of the TPI approach over MC simulations, especially in the case of low dimensional problems, are revealed in by comparing the error growth over time for the various approaches. The influence of the diffusion coefficient as well as that of the spatial and temporal grid resolutions on the error in the solutions obtained using the TPI approach are further examined through simulations of the pure diffusion process. The existence of a potential similarity parameter, that can account for the influences of the above simulation parameters, is also studied. Application

of the TPI approach to MDOF systems is presented in Examples 5 to 8. These examples cover linear and nonlinear stochastic dynamical systems (including those with anisotropic diffusion tensors). Through these examples, the advantage of the TPI approach over conventional approaches (based on fixed grid and MC) in terms of performance (relevant to accuracy and computational time) is demonstrated.

2.3.1 SDOF stochastic dynamical systems

We begin by studying processes in one dimensional state spaces. As we will see later in the examples in multidimensional state spaces, the benefits highlighted here in SDOF systems are also observed in MDOF systems.

Example 1: A pure diffusion process in 1D

A salient feature of path integral based approaches, especially compared to MC simulations, is their ability to accurately represent the PDF tail information. However, numerical implementations based on a fixed grid (FG) approach due to fixed and finite grid domain bounds, may not be able to fully utilize this ability. Hence, they face challenges in obtaining accurate estimates of the long time evolution of the distribution for processes where the PDF is unbound and diffuses over time. One such process is the pure diffusion process where the PDF information beyond the bounds of the domain, in a conventional fixed grid implementation, grows in time with the evolution of the PDF. This constitutes a loss in tail information that significantly contributes to the error in the PDF representation.

The pure diffusion process is considered here to showcase the benefits afforded by the TPI approach in addressing challenges for the accurate representation of the tail information. The process is governed by the following equation

$$\mathrm{d}x = a\,\mathrm{d}w \tag{2.37}$$

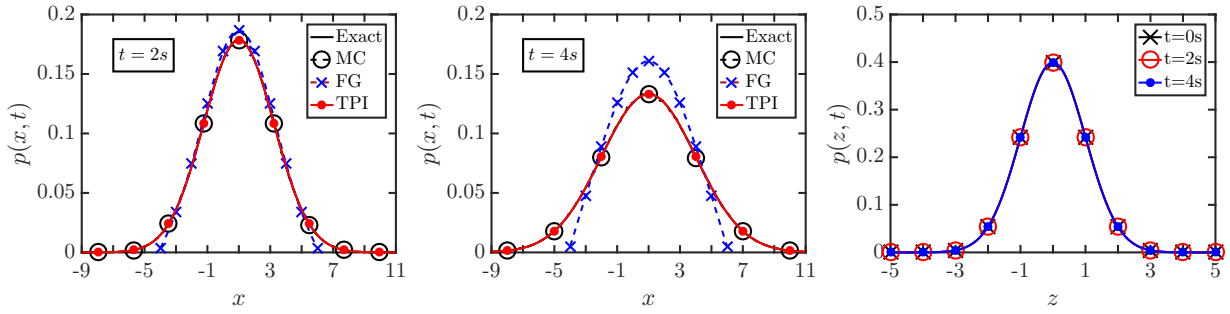

Figure 2.1: Comparison of PDFs in the original space at $t = 2\,\mathrm{s}$ (left) and $t = 4\,\mathrm{s}$ (center) as well as in the transformed space (right) for the pure diffusion process Eq. (2.37).

where the strength of the white noise process is chosen to be $a = \sqrt{2}$. The simulation is run from $0\,\mathrm{s}$ to $4\,\mathrm{s}$ with a time step of $\Delta t = 0.01\,\mathrm{s}$. The initial distribution is assumed to be a unit mean and unit variance Gaussian distribution i.e $p(x, t_0) = \mathcal{N}(\mu_0, \sigma_0^2)$ where $\mu_0 = \sigma_0 = 1$.

The simulation is performed on a uniform grid with 201 grid points between the upper and lower bounds $[x_l, x_u] = [-4, 6]$ in the original space and on a uniform grid with 201 grid points between the bounds $[z_l, z_u] = [-5, 5]$ in the transformed space. In order to facilitate comparisons, the number of grid points as well as the density of these points in their respective domains have been chosen to be the same. However, it is to be noted that the chosen fixed grid in the transformed space is equivalent to an adaptive grid in the original space with bounds that are five standard deviations on either side of the mean i.e. $[\mu - 5\sigma, \mu + 5\sigma]$, where μ and σ represent the mean and variance respectively at time t.

Better performance of the TPI approach over the conventional fixed grid approach can be seen from Fig. 2.1. The plots in the left and center of Fig. 2.1 show the time evolution of the state PDF where the solutions obtained from the TPI, FG and MC approaches are shown along with the exact analytical solution. It can be seen that the distribution for this problem remains Gaussian centered about $\mu_0 = 1$ with a variance $\sigma(t)^2$ that grows linearly with time. Besides, the state variables in the transformed computational domain is found to have a standard normal distribution that remains invariant in time as suggested by Lemma. (2.2.1) and demonstrated by the plots in the right of Fig. 2.1. Thus from Eq. (2.21), the error in the PDF due to the loss of tail information in the conventional fixed grid formulation for

this problem can be evaluated to be erfc $\left(5/\sqrt{2}\sigma_t\right)$. As the standard deviation grows in time for this process, the PDF area outside of the bounds in FG increases leading to the growth in error. The value of this error at time $t = 4\,\mathrm{s}$ is approximately 0.096 i.e. close to 10 % of the probability in the transformed space is lost contributing to the poor performance seen in (center) in Fig. 2.1.

In contrast, the error from loss of tail information in the TPI approach for this process is found to be erfc $\left(5/\sqrt{2}\right)$. This error is a constant under the propagation and approximately equal to 5.73×10^{-7} i.e. less than 0.0001% of the probability in the transformed space is lost. The constant value of the error is due to the invariance of the PDF (from a standard normal distribution) in the transformed space for this process which can be seen (right) in Fig. 2.1. The results from the simulation of the pure diffusion process show that the transformed path integral approach, for the same number of grid points, performs significantly better than the conventional fixed grid approach.

Example 2: A constant drift and diffusion process

The constant drift and diffusion process is another simple stochastic process with a well known analytical solution. Processes with large drift pose significant challenges to the traditional grid based approaches, which are limited by a stationary grid. Over time, the PDF may drift out of the bounds of the original computational grid space. The benefits offered by the TPI approach to address these challenges are illustrated in this example. The process considered is given by the following equation

$$\mathrm{d}x = \kappa\,\mathrm{d}t + a\,\mathrm{d}w \tag{2.38}$$

where the constant drift is given by $\kappa = 3$ and the strength of the white noise process is chosen to be $a = 1/3$. The simulation is run from 0 s to 4 s with a time step of $\Delta t = 0.01\,\mathrm{s}$. The initial distribution is assumed to be a Gaussian distribution given by $p(x, t_0) = \mathcal{N}(1, 2)$.

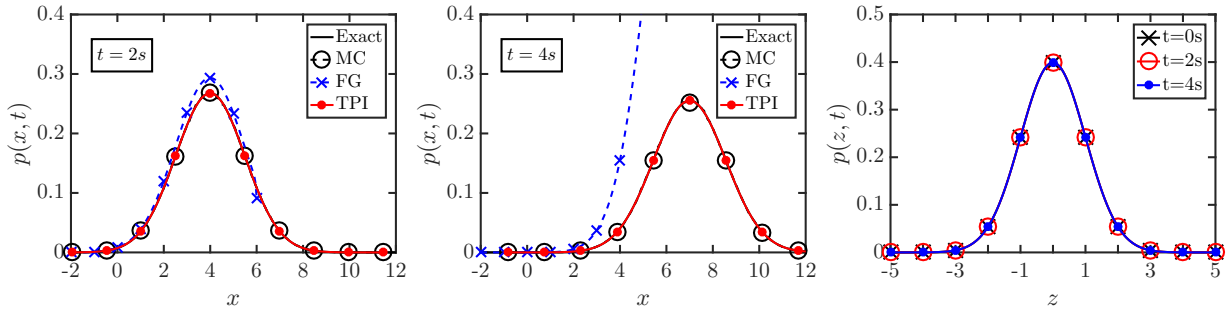

Figure 2.2: Comparison of PDFs in the original space at $t = 2\,\mathrm{s}$ (left) and $t = 4\,\mathrm{s}$ (center) as well as in the transformed space (right) for the constant drift process Eq. (2.38).

The bounds on the grid and the number of grid points used for this problem are the same as those used in the previous example for the simulation of the pure diffusion process.

In Fig. 2.2, the PDF plots in the original state space at times $t = 2\,\mathrm{s}$ (left) and $t = 4\,\mathrm{s}$ (center) show the failure of the FG approach, as the PDF drifts out of the bounds $[-4, 6]$ in the original computational domain. In spite of the mean and variance growing linearly with time, it can be seen that the form of the distribution remains Gaussian. Thus, following the steps laid out in the pure diffusion case, the failure of the FG approach can be similarly quantified. The error resulting from the inaccurate representation of the PDF in the tail regions at $t = 4s$ for FG is 0.7388 i.e. a loss of almost 74 % of the probability in the transformed space. This high loss is reflected in the poor performance of FG seen (center) in Fig. 2.2. The corresponding error obtained from the TPI approach remains less than 0.0001 %. This example illustrates the unique property of the TPI approach where the PDF exhibiting large translation in the original space is still well contained within the fixed bounds of the computational domain in the transformed space. Additionally, as an evidence to Lemma. (2.2.1), the invariance of the PDF in the transformed space for the linear drift and constant diffusion process considered here can be seen (right) in Fig. 2.2.

Example 3: OU process

Processes with stable fixed points in the associated deterministic dynamical system, i.e. in the absence of noise, may exhibit concentration of the probability density functions near

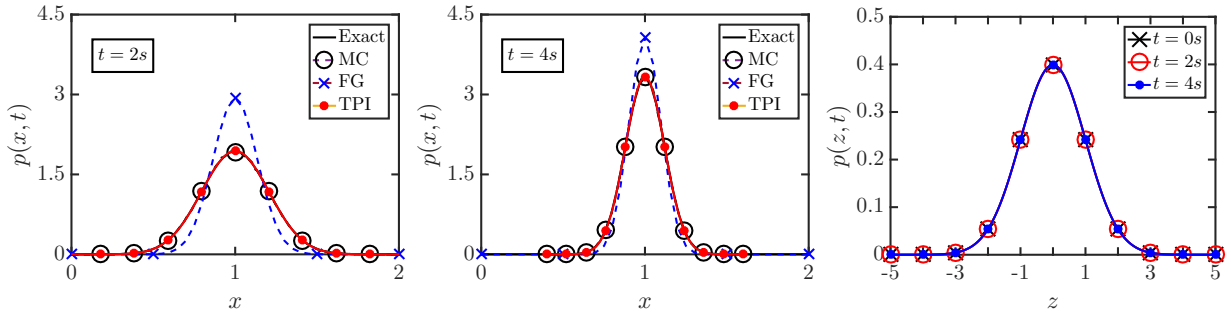

Figure 2.3: Comparison of PDFs in the original space at $t = 2\,\mathrm{s}$ (left) and $t = 4\,\mathrm{s}$ (center) as well as in the transformed space (right) for the OU process Eq. (2.39).

these points, especially in the case of processes with small diffusion coefficients corresponding to low noise processes. These distributions are often characterized by large peak values at the fixed points and with very thin tails. Conventional fixed grid based approaches face challenges, due to a fixed and finite grid resolution, in efficiently and accurately representing the transient behavior of distributions near these stationary points. In contrast, the TPI approach offers the possibility to ensure a better and sufficient resolution. This better performance is showcased here with the help of a linear drift process with a constant diffusion coefficient. The process, known as the Ornstein-Uhlenbeck process, is a process with a stationary point and a known analytical solution [1]. It is given by the following equation

$$\mathrm{d}x = -\beta(x - \theta)\,\mathrm{d}t + a\,\mathrm{d}w. \tag{2.39}$$

where $\beta = 1/2$, $\theta = 1$ and the strength of the white noise process is $a = 0.1$. The initial distribution, given by $p(x, t_0) = \mathcal{N}(\mu_0, \sigma_0^2)$ where $\mu_0 = 1$ and $\sigma_0 = 0.5$, is evolved from $0\,\mathrm{s}$ to $4\,\mathrm{s}$ in time steps of $\Delta t = 0.01\,\mathrm{s}$. The dimensions of the grid are the same as those in the pure diffusion example.

The process described here has a stationary point at $x = \theta = 1$. As we evolve in time, the PDF will have a peak value at this point. This can be seen from the PDF plots in the original space at $t = 2\,\mathrm{s}$ (left) and $t = 4\,\mathrm{s}$ (center) in Fig. 2.3. The failure of the FG approach near the stationary point can also be seen. As the variance σ^2 for this process shrinks and

approaches the steady state value of $a^2/2 = 0.01$, the PDF peak value at the stationary point increases and the PDF near this point becomes more narrow. The FG approach may not be able to ensure sufficient resolution near the stationary point to accurately represent this behavior over time as shown (left and center) in Fig. 2.3. In contrast, sufficient resolution around the stationary point can be achieved with the TPI approach leading to the more accurate representation. This is due to the fact that a uniform grid spacing of Δz in the transformed space is equivalent to an adaptive grid with a spacing of $\Delta x = \sigma \Delta z$ in the original state space. Thus a fine resolution when the variance is small, in order to ensure a better representation of the PDF near the stationary point, is automatically accounted for by the TPI approach.

Example 4: A stochastic process with nonlinear drift

The TPI approach is able to sufficiently accommodate the nonlinearities in the drift that may be present in a typical real world system. Additionally, the approach can also accommodate the non-Gaussian behavior of PDFs which arise in these nonlinear drift processes. We consider the following process to illustrate the same.

$$\mathrm{d}x = -\gamma x(x^2 - \delta^2)\,\mathrm{d}t + a\,\mathrm{d}w \tag{2.40}$$

where $\gamma = 1/16$, $\delta = 4$ and the strength of the white noise process is chosen to be $a = 3$. The simulation is run from 0 to 16s with a time step of $\Delta t = 0.01$s. The initial distribution is assumed to be a Gaussian distribution given by $p(x, t_0) \sim \mathcal{N}(1, 2)$.

The problem considered is of a nonlinear drift process for which the steady state solution is known. The transient solutions and the steady state solution for the problem are inherently non-Gaussian. The steady state solution is given by

$$p(x) = \frac{1}{\eta} \exp\left\{-\gamma \frac{x^2}{a^2}\left(\frac{x^2}{2} - \delta\right)\right\} \tag{2.41}$$

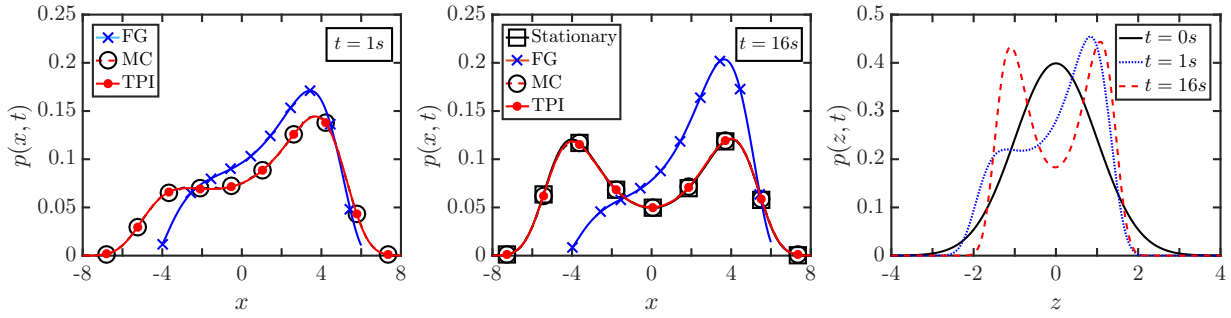

Figure 2.4: Comparison of PDFs in the original space at $t = 1\,\mathrm{s}$ (left) and $t = 16\,\mathrm{s}$ (center) as well as in the transformed space (right) for the 1D non-linear process Eq. (2.40).

where η is a normalization constant. Clearly, the steady state solution is a bimodal PDF with the peaks located at $x = \pm\delta = \pm 4$. It can be seen from Fig. 2.4 that the particular transformation considered is able to sufficiently capture the non-Gaussian PDF behavior as well. The limitations of a fixed grid can be seen here for this particular problem as well. It is to be noted that in the transformed space we do not observe the invariant solution behavior exhibited by the linear drift processes.

Error analysis

In the preceding examples, we have seen clear benefits for the proposed transformed path integral (TPI) approach over the conventional fixed grid (FG) approach. The results also indicate good agreement between the TPI approach and Monte Carlo (MC) simulations, especially for the low dimensional problems considered. Further, the performance of these approaches can be better quantified via an RMS error measure in the PDF. A discrete representation of this error measure on an uniform grid is given by

$$\varepsilon_{RMS} = \sqrt{\frac{1}{N_{pts}} \sum_{i=1}^{N_{pts}} \left[p(x_i) - \hat{p}(x_i)\right]^2} \tag{2.42}$$

where $\hat{p}(x_i)$ is the numerical solution obtained using MC, FG or the TPI approach for the pure diffusion process (Eq. (2.37)) and the OU process (Eq. (2.39)) at the i-th grid point. The summation in the definition is taken over the entire set of grid points (or samples). The

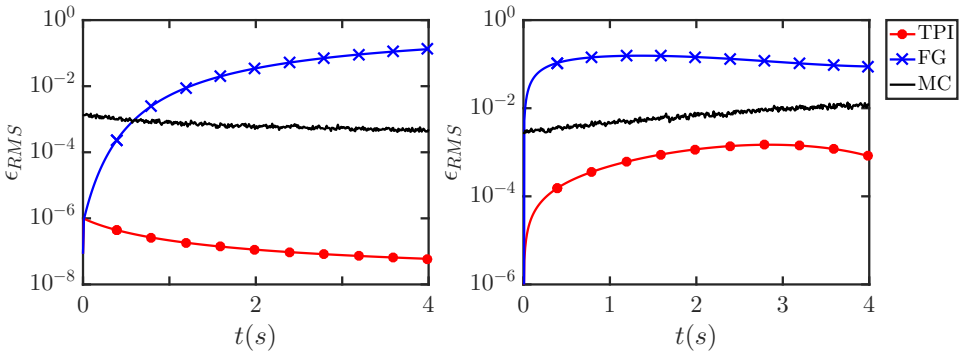

Figure 2.5: Comparison of the time evolution of the RMS error in the PDF obtained by fixed grid (FG), Monte Carlo (MC) and transformed path integral (TPI) approaches for the pure diffusion process Eq. (2.37) (left) and the OU process Eq. (2.39) (right).

		Linear Stochastic Dynamical Systems in 2D					
		Pure Diffusion		OU Process		Nonlinear Dynamics	
	N_{pts}	T_{comp}(s)	ε_{RMS}	T_{comp}(s)	ε_{RMS}	T_{comp}(s)	ε_{RMS}
FG	201	2.4	1.33×10^{-1}	2.9	8.85×10^{-2}	9.3	5.45×10^{-2}
MC	10^6	34	4.94×10^{-4}	33.9	1.23×10^{-2}	132	7.93×10^{-4}
TPI	201	2.9	5.95×10^{-8}	3.7	8.11×10^{-4}	10.8	6.46×10^{-4}

Table 2.1: Comparison of RMS error in the PDF obtained by fixed grid (FG), Monte Carlo (MC) and transformed path integral (TPI) approaches at $t = 4s$ for the pure diffusion process Eq. (2.37) and the OU process Eq. (2.39).

analytical solution $p(x)$ is known at all times for the problems considered here. Based on this error measure, we can clearly see that the TPI approach performs significantly better than the conventional fixed grid approach and MC simulations. For instance, at $t = 4s$ the error for the TPI approach is approximately four orders of magnitude lower than that for MC and six orders of magnitude lower than that for FG as also shown in Table. 2.1. It can be seen that the error plots for the OU process show similar benefits for the TPI approach over the conventional fixed grid approach and Monte Carlo simulations. Additionally, from the comparison of the RMS error in the PDF at $t = 4s$ (for the case of pure diffusion and OU process) and the corresponding computational costs involved, as shown in Table. 2.1, we see that the reduction in the error for the TPI approach is obtained at a significantly lower computational cost (as compared to MC) or comparable computational cost (as compared

to FG) in terms of the computation time (T_{comp}). In particular, this benefit is also seen for the nonlinear drift process Eq. (2.40). Note that in this case, the RMS error (at a large time, $t = 16s$) is evaluated by considering the stationary PDF as a reference to the exact solution. Such a restriction (on comparison with stationary PDF) may often be necessary when exact solutions corresponding to transient behavior of stochastic systems are not readily available.

Further reduction in the PDF error in the TPI approach can be achieved by varying the parameters of the numerical simulation. Intuitively, the expectation is that with increase in the number of grid points N_{pts} there is a decrease in the error ε_{RMS}. While this is true for large number of grid points, as shown in Fig. 2.6, the opposite behavior is observed when a small number of grid points are used. In addition, we see that beyond a point there is no further reduction in the error with increase in the number of grid points. This is because of the error arising from the loss of tail information. Increase in the domain size of the grid points ensure better representation of the tail information, but for the same number of grid points this would increase the discretization error. Thus, it can be seen that having a larger domain along with an increase in the number of grid points produces a greater reduction in error.

The influence of the size of the time step Δt on the error ε_{RMS} is shown in Fig. 2.6, where the existence of regions corresponding to a loss of accuracy in ε_{RMS} can be seen. It can also be seen that this loss of accuracy can be reduced and shifted to a smaller time step with increase in the number of grid points. Thus the effects of the time step Δt and that of the number of grid points N_{pts} on the error ε_{RMS} are not independent of each other. The plot also indicates that there might exist a similarity parameter that could collapse the different curves into one.

It can be seen from Fig. 2.7, that this similarity parameter may also depend on the diffusion constant a of the process. We propose a possible candidate given by

$$\alpha = \frac{\sigma_{ref}^2 \, (\Delta z)^2}{2a^2 \Delta t} \tag{2.43}$$

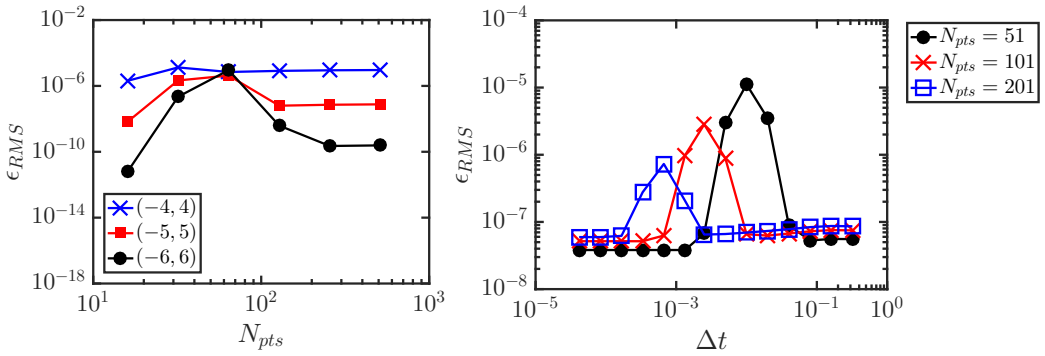

Figure 2.6: Influence of the number of grid points (left) and the size of the time step (right) on the RMS error in the PDF obtained through transformed path integral approach (TPI) for the pure diffusion process Eq. (2.37). The RMS error was calculated for each point at $t = 4s$.

where $\Delta z \propto 1/N_{pts}$ is the grid spacing in the transformed space and σ_{ref} is a scaling parameter to ensure the non-dimensionality of α. In our analysis, the scaling parameter σ_{ref} was chosen to be equal to unity. The analysis is analogous to the von Neumann stability analysis as performed for the heat equation. Previous studies in the literature based on fixed grid implementations of the path integral [47] have noted that in order to obtain correct results, the grid spacing in the original space Δx, the time step Δt, and the diffusion coefficient a need to satisfy the relation $\Delta x = [a^2 \Delta t]^{1/2}$. Our analysis extends the previous study further by analyzing the error ε_{RMS} for different α values as shown (right) in Fig. 2.7.

The plot of the variation of the error with α shows a partial collapse of the curves. The regions, where there is a loss of accuracy, overlap on the α axis (≈ 1.5). However, the peak values themselves are different and also the collapse seems to be better in certain regions of the α space ($1.5 \leq \alpha \leq 10$). From the plot, we can infer that the best choices for Δt to minimize the error are values that correspond to $\alpha > 10$ or $\alpha < 0.5$. But further analysis may be required in order to better understand the similarity parameters associated with the TPI and their influence on the error.

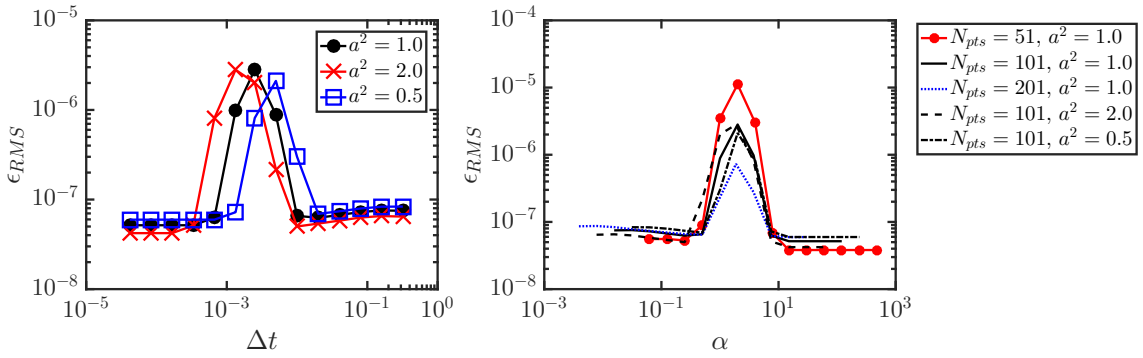

Figure 2.7: Influence of the size of the time step (left) and the proposed similarity parameter (right) on the RMS error in the PDF obtained through transformed path integral approach (TPI) and calculated at $t = 4s$ for the pure diffusion process Eq. (2.37).

2.3.2 MDOF stochastic dynamical systems

The ability of the TPI approach to better address challenges arising from large diffusion, large drift and concentration in the PDF also extends to MDOF systems. In this context, performance of the TPI approach is evaluated by consideration of selected problems in two-dimensional state space, as given by (i) Example 5: A pure diffusion process in 2D (including isotropic and anisotropic diffusion tensor with non-zero cross-covariances), (ii) Example 6: A coupled linear system in 2D (analogous to 1D OU process), (iii) Example 7: An uncoupled nonlinear system in 2D and (iv) Example 8: A coupled nonlinear system in 2D. Note that the latter two examples involve cubic nonlinearities (analogous to Example 4). Through analysis of problems considered in examples 5 to 8, benefits of the TPI approach in comparison to FG approach and MC simulations are demonstrated for two-dimensional stochastic dynamical systems.

Example 5: A pure diffusion process in 2D

The conventional fixed grid (FG) approaches face challenges in accurately representing the PDF in systems with large diffusion. The error, arising from a loss of tail information in the PDF, can be reduced with the TPI approach. We consider the two dimensional pure

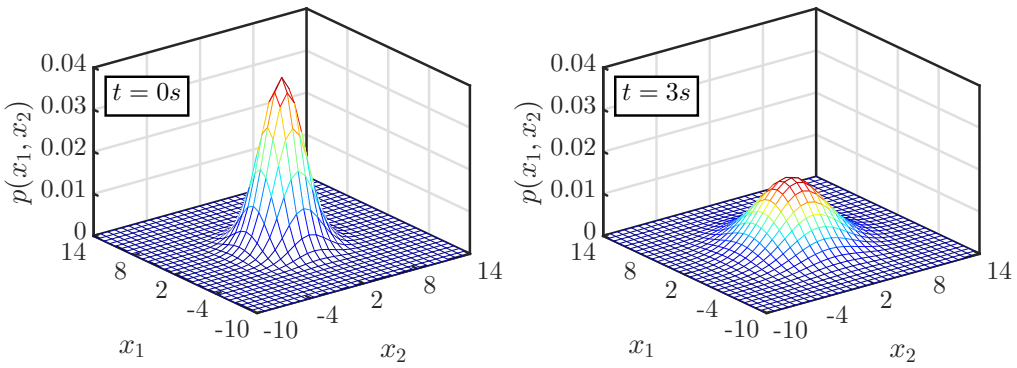

Figure 2.8: The plots above show the exact analytical solution for the time evolution of the joint PDF of the state for the 2D pure diffusion process Eq. (2.44) from $t = 0\,\mathrm{s}$ to $t = 3\,\mathrm{s}$.

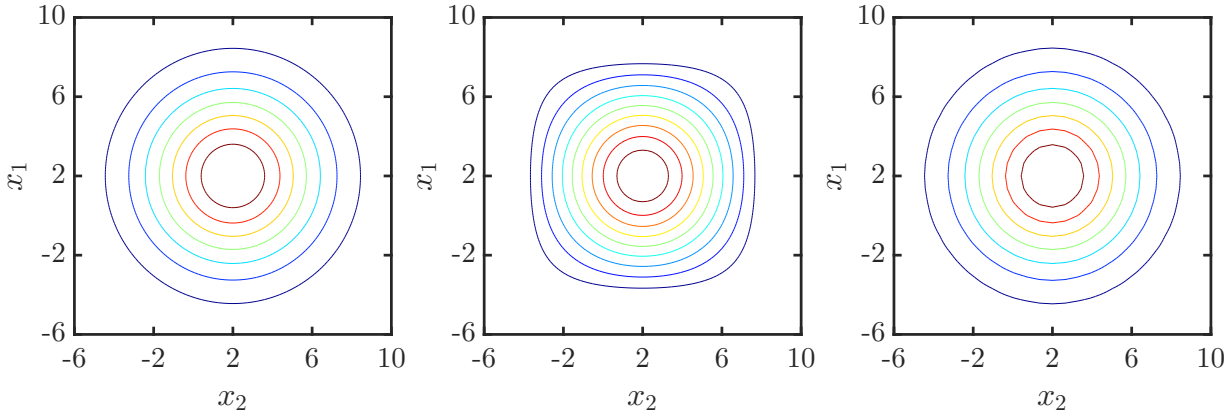

Figure 2.9: Contour plots of the joint PDF at $t = 3\,\mathrm{s}$ for the 2D pure diffusion process Eq. (2.44) obtained by the analytical solution (left), the fixed grid approach (center) and the TPI approach (right).

diffusion process to show this benefit. The process is given by the equation

$$\mathrm{d}\mathbf{x} = \mathbf{A}\mathrm{d}\mathbf{w} \tag{2.44}$$

where the diffusion coefficient matrix $\mathbf{G} \equiv \mathbf{D}^{(2)} = \mathbf{A}\mathbf{A}^T/2$ is chosen to be the identity matrix of size 2. Note that the diffusion tensor considered in this case (labeled Case A) is isotropic. The simulation is run from 0 s to 3 s in steps of 0.01 s with a Gaussian initial distribution of mean $\begin{bmatrix} 2 & 2 \end{bmatrix}^T$ and a (2×2) diagonal covariance matrix with the diagonal elements $\begin{bmatrix} 4 & 4 \end{bmatrix}$. The simulations were performed using a 51×51 grid resolution in both the original and transformed space. The domain bounds were taken to be $[x_l, x_u] = [-4, 8]$

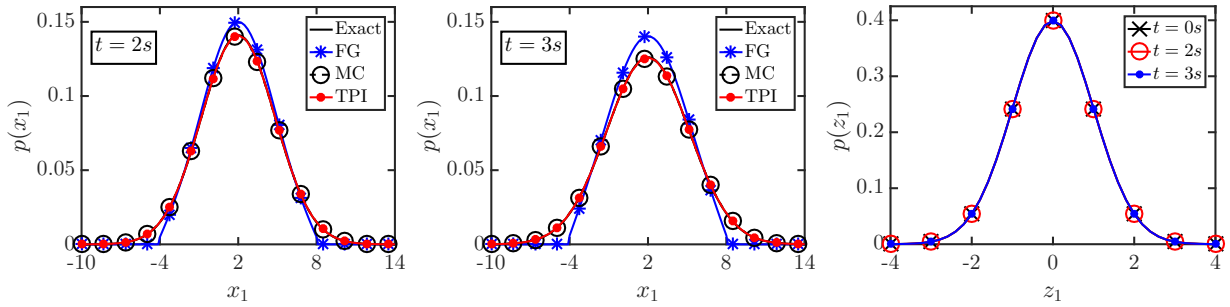

Figure 2.10: Comparison of marginal PDFs in the original space at $t = 2\,\mathrm{s}$ (left) and $t = 3\,\mathrm{s}$ (center) as well as in the transformed space (right) for the 2D pure diffusion process Eq. (2.44).

in either dimensions of the original space for FG simulations and $[z_l, z_u] = [-5, 5]$ in either dimensions of the transformed space for the TPI simulations.

The analytical solution for this problem is known and the evolution of the PDF is shown in Fig. 2.8. The superior performance of the TPI over the conventional fixed grid approaches can be seen from the contour plots at $t = 3\,\mathrm{s}$ shown in Fig. 2.9. This can also be seen from the plots of the marginal PDFs shown in Fig. 2.10. Invariance of the TPI solution in the transformed space can also be noted in Fig. 2.10 (right), as expected via generalization of Lemma. (2.2.1) for multidimensional (linear) problems. Performance of the TPI approach in the context of accuracy and computational time can also be assessed using data shown in Table. 2.2. Higher accuracy (or lower RMS error) of the TPI approach in comparison to FG and MC approaches can be observed (in Table. 2.2) for comparable computational times.

The performance of the TPI approach is further assessed for the case of a pure diffusion process with an ansiotropic diffusion tensor (including non-zero, off-diagonal terms). For this case (labeled Case B), the elements of the diffusion tensor are chosen to be $G_{11} = 3$, $G_{22} = 1$ and $G_{12} = G_{21} = \sqrt{3}/2$. The simulation is run from $0\,\mathrm{s}$ to $2\,\mathrm{s}$ in steps of $0.1\,\mathrm{s}$ with a multivariate standard normal distribution as the initial distribution.

It can be seen from the analytical solution shown in Fig. 2.11 for the evolution of the PDF and the contour plots at $t = 2\,\mathrm{s}$ shown in Fig. 2.12, that the PDF undergoes stretching and rotation. Based on Fig. 2.12, a clearly visible benefit of the TPI approach over the FG

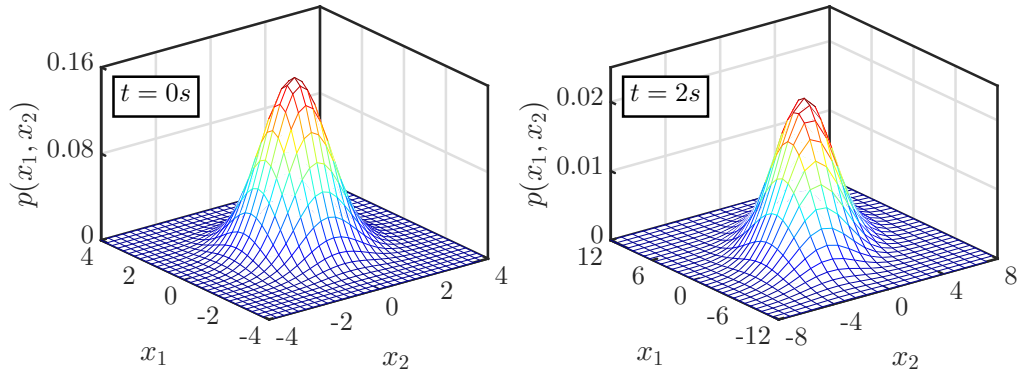

Figure 2.11: The plots above show the exact analytical solution for the time evolution of the joint PDF of the state for the 2D pure diffusion process Eq. (2.44) from $t = 0\,\mathrm{s}$ to $t = 2\,\mathrm{s}$.

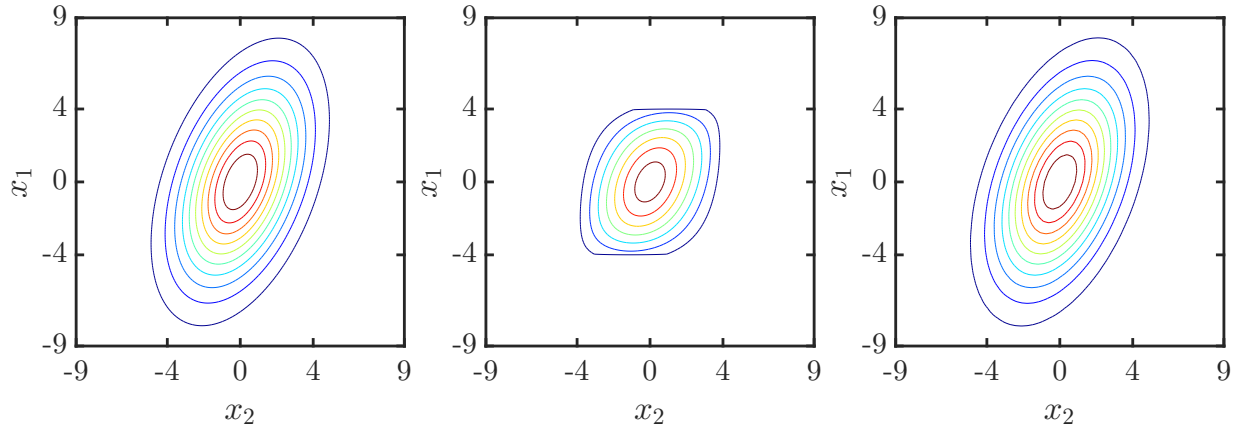

Figure 2.12: Contour plots of the joint PDF at $t = 2\,\mathrm{s}$ for the 2D pure diffusion process Eq. (2.44) obtained by the analytical solution (left), the fixed grid approach (center) and the TPI approach (right).

approach can be noted. Clear visible benefit can be seen in Fig. 2.12 for the TPI approach over the FG approach. Additionally, the marginal PDF plots in the original space shown in Fig. 2.13, illustrate the large growth of error with time for results obtained using the FG approach. The superior performance of the TPI approach is further underscored by the predicted invariant solution behavior exhibited by PDF in the transformed space as shown in Fig. 2.13 (right).

Better performance of the TPI approach for this case (with anisotropic diffusion tensor) can also be observed from Table. 2.2. In comparison to the FG approach, it can be noted that the TPI approach has significantly greater accuracy for comparable computational times. The TPI approach is also observed to outperform the MC approach, by virtue of lower RMS

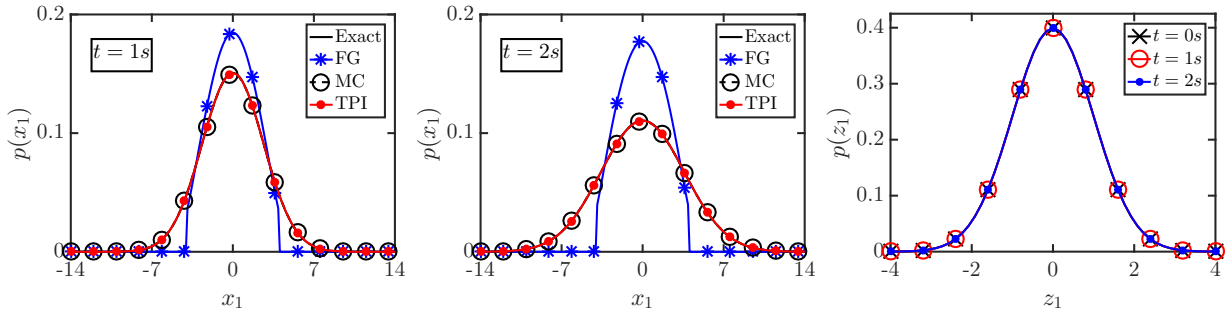

Figure 2.13: Comparison of marginal PDFs in the original space at $t = 1\,\mathrm{s}$ (left) and $t = 2\,\mathrm{s}$ (center) as well as in the transformed space (right) for the 2D pure diffusion process Eq. (2.44).

error and lower computational time.

Example 6: A coupled linear process in 2D

Concentration of the PDF, like in SDOF systems (Example 3), poses challenges to the conventional FG approach in MDOF systems. We consider simulation of a coupled system of Ornstein-Uhlenbeck processes in a two dimensional space, given by the equation

$$\mathrm{d}\mathbf{x} = \mathbf{K}(\mathbf{c} - \mathbf{x})\,\Delta t + \mathbf{A}\mathrm{d}\mathbf{w} \tag{2.45}$$

where $\mathbf{K}$ is a lower triangular matrix with elements $K_{11} = 0.5$ and $K_{22} = K_{21} = 0.25$, while $\mathbf{c}$ is the column vector $\begin{bmatrix}1 & 1\end{bmatrix}^{\mathrm{T}}$. This particular coupled linear stochastic system has a stable stationary point (corresponding to the underlying deterministic system) at $\mathbf{x} = \mathbf{c}$ and hence a concentration of the PDF around this stationary point can be expected for large times (for relatively small process noise levels). An anisotropic diffusion matrix is considered here, with component terms given by $G_{11} = G_{12} = G_{21} = 1/32$ and $G_{22} = 1/16$, to illustrate the limitations of the conventional fixed grid approaches in accurately representing this PDF behavior. Starting with a multivariate standard normal distribution (as our initial condition), the simulation is run from $0\,\mathrm{s}$ to $2\,\mathrm{s}$ in steps of $0.001\,\mathrm{s}$, using a 51×51 grid resolution for TPI and FG approaches. The bounds of the finite computational domain were taken to be $[\xi_l, \xi_u] = [-4, 4]$ in either dimensions of both, the original and the transformed spaces.

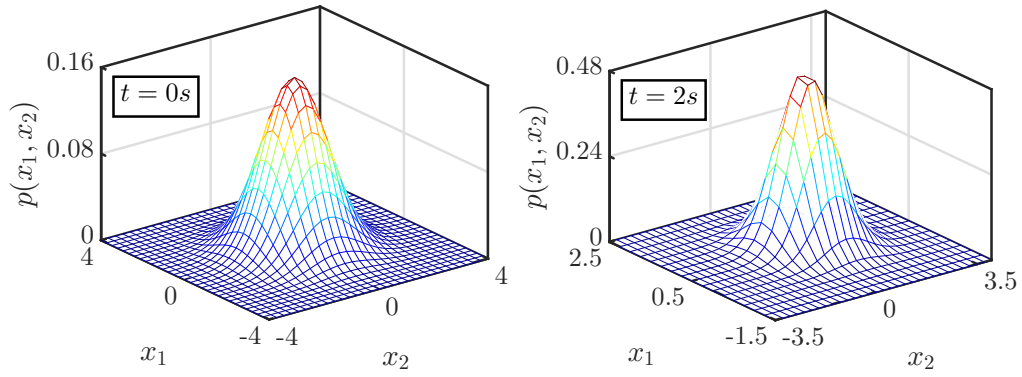

Figure 2.14: The plots above show the exact analytical solution for the time evolution of the joint PDF of the state for the 2D OU process Eq. (2.45) from $t = 0\,\mathrm{s}$ to $t = 2\,\mathrm{s}$.

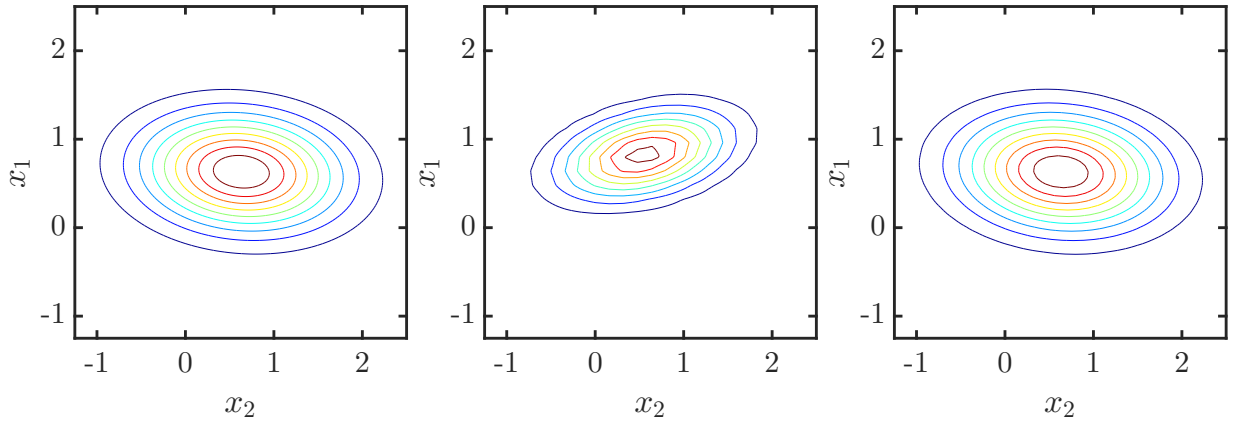

Figure 2.15: Contour plots of the joint PDF at $t = 2\,\mathrm{s}$ for the 2D OU process Eq. (2.45) in 2D obtained by the analytical solution (left), the fixed grid approach (center) and the TPI approach (right).

In addition to concentration of the PDF near stationary points, the coupled linear drift process considered here can cause the PDF to be stretched and rotated. This behavior can be seen in the analytical solution shown in Fig. 2.14. High concentration of the PDF, similar to the 1D OU process considered earlier in Example 3, may lead to errors with the FG approach due to an insufficient resolution in the fixed grid, as shown in Fig. 2.15 and Fig. 2.16. These figures also illustrate the relative benefits of the TPI approach. The ability of the TPI approach to achieve a sufficiently fine resolution in regions of large concentration in the PDF, as noted for the SDOF case given in Example 3, is also observed in this two-dimensional case. The TPI solution in the transformed space, shown in Fig. 2.16 (right), is also found to be invariant (as expected). Better performance of the TPI approach (for

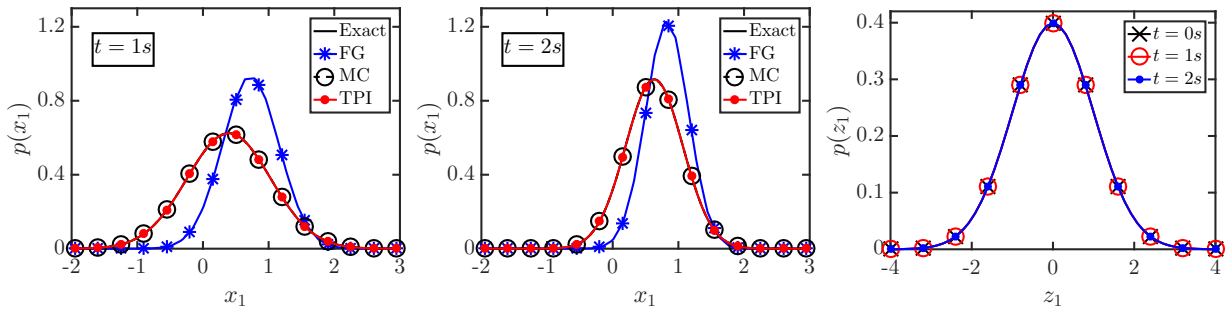

Figure 2.16: Comparison of marginal PDFs in the original space at $t = 1\,\mathrm{s}$ (left) and $t = 2\,\mathrm{s}$ (center) as well as in the transformed space (right) for the 2D OU process Eq. (2.45).

		Linear Stochastic Dynamical Systems in 2D					
		Isotropic Diffusion (Case A)		Anisotropic Diffusion (Case B)		Coupled System (OU Process)	
	N_{pts}	$T_{\mathrm{comp}}(\mathrm{s})$	$\varepsilon_{\mathrm{RMS}}$	$T_{\mathrm{comp}}(\mathrm{s})$	$\varepsilon_{\mathrm{RMS}}$	$T_{\mathrm{comp}}(\mathrm{s})$	$\varepsilon_{\mathrm{RMS}}$
FG	51×51	114	6.64×10^{-4}	9.5	2.05×10^{-3}	601	3.94×10^{-2}
MC	10^7	235	1.98×10^{-5}	93.9	3.03×10^{-5}	1660	4.36×10^{-4}
TPI	51×51	128	1.40×10^{-5}	10.0	1.04×10^{-5}	790	2.29×10^{-4}

Table 2.2: Comparison of computational times and RMS errors in PDFs obtained by fixed grid (FG), Monte Carlo (MC) and transformed path integral (TPI) approaches for the following two-dimensional linear stochastic dynamical systems: the 2D pure diffusion process Eq. (2.44) (with isotropic and anisotropic diffusion tensors, considered in Example 5) and the coupled system of OU process in 2D Eq. (2.45)

this problem) in the context of accuracy and computational time can also be noted from Table. 2.2 and the benefits over FG and MC approaches are found to be qualitatively similar to those found in anisotropic diffusion case in Example 5.

Example 7: An uncoupled nonlinear process in 2D

Like in the scalar case, the application of the TPI approach can be easily extended to nonlinear drift processes. The following process is considered to illustrate the same.

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\,\mathrm{d}t + \mathbf{A}\,\mathrm{d}\mathbf{w} \tag{2.46}$$

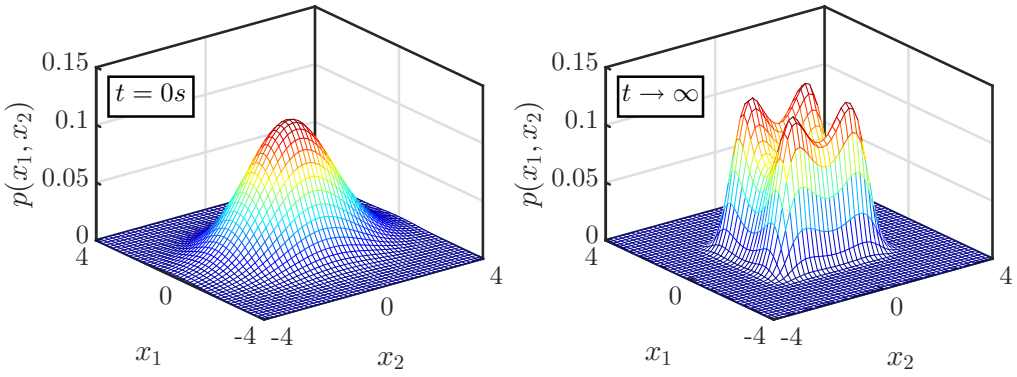

Figure 2.17: The plots above show the initial joint PDF (left) and the stationary joint PDF (right) of the state for the nonlinear drift process Eq. (2.46)

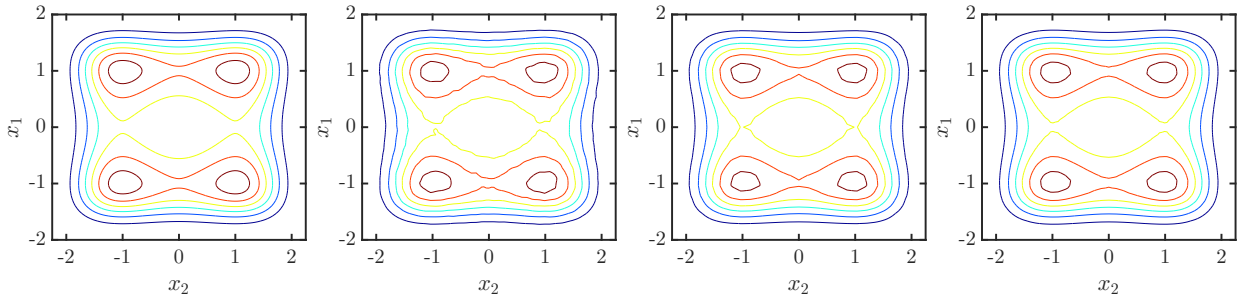

Figure 2.18: Contour plots of the stationary joint PDF (far left) and the joint PDF at $t = 2\,\mathrm{s}$ obtained by MC simulations (center left), the FG approach (center right) and the TPI approach (far right) for the 2D uncoupled nonlinear drift process Eq. (2.46)

where the nonlinear drift is given by $\mathbf{f}(\mathbf{x}, t) = \begin{bmatrix} x_1 - x_1^3 & x_2 - x_2^3 \end{bmatrix}^{\mathrm{T}}$ and the diffusion coefficient matrix for the process $\mathbf{G} \equiv \mathbf{A}\mathbf{A}^{\mathrm{T}}/2$, is a (2×2) diagonal matrix with diagonal elements $[0.5, 1]$. For this case (labeled Case C), the initial distribution is a multivariate Gaussian distribution given by $p(\mathbf{x}, t_0) = \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_0})$, where $\mathbf{\Sigma_0}$ is a (2×2) diagonal matrix with diagonal elements $[1, 2]$. The MC solution for the evolution of the PDF is shown in Fig. 2.17. The PDF, as expected, does not remain Gaussian. The TPI approach is able to sufficiently capture the multidimensional non-Gaussian behavior as can be seen from the comparison of the contour plots at $t = 2\,\mathrm{s}$ shown in Fig. 2.18. The marginal plots shown in Fig. 2.19 further illustrate the same. The plots of the marginal PDF in the transformed space for this problem, as expected, do not exhibit the invariant solution behavior seen in the pure diffusion case.

Based on the data shown in Table. 2.3, the performance of the TPI approach appears to

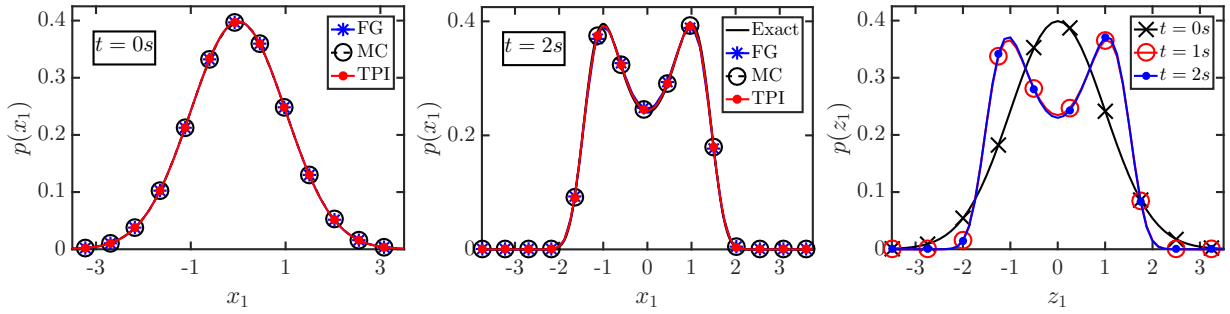

Figure 2.19: Comparison of marginal PDFs at $t = 0\,\mathrm{s}$ (left) and $t = 2\,\mathrm{s}$ (center) in the original space and transformed space (right), for the 2D uncoupled nonlinear drift process Eq. (2.46).

be marginally better (in terms of accuracy) than the fixed grid approach, albeit for slightly longer computational time. When compared with MC approach, comparable accuracy can be achieved by the TPI approach at a slightly lower computational time. Note that the bounds for the FG simulations and TPI simulations are chosen to be $[x_l, x_u] = [-6, 6]$ and $[z_l, z_u] = [-7/2, 7/2]$ respectively in each dimension. In this case, the bounds of the fixed grid computational domain are chosen such that inital and transient behaviors of the PDF are well represented (for the duration of the simulation time). Although it is possible to choose the bounds of the fixed grid such that the initial PDF is well represented, it might not always be possible to choose the bounds such that the transient behavior of the PDF is adequately represented for the duration of the simulation time. While the FG approach can be significantly affected by this limitation, the TPI approach can handle this limitation better as fixed bounds in the transformed space (of the TPI approach) correspond to dynamically adaptive bounds in the original space.

To illustrate a situation where the fixed grid might be significantly affected by the choice of bounds, we select a problem with the nonlinear drift considered in Eq. (2.46) with a different set of parameters. In particular, the diffusion coefficient matrix for the process is now chosen to be a (2×2) diagonal matrix with diagonal elements $G_{11} = G_{22} = 2$. Note that an isotropic diffusion tensor is considered in this case (labeled Case D), in contrast to the anisotropic diffusion tensor considered in the earlier case. The initial distribution for this

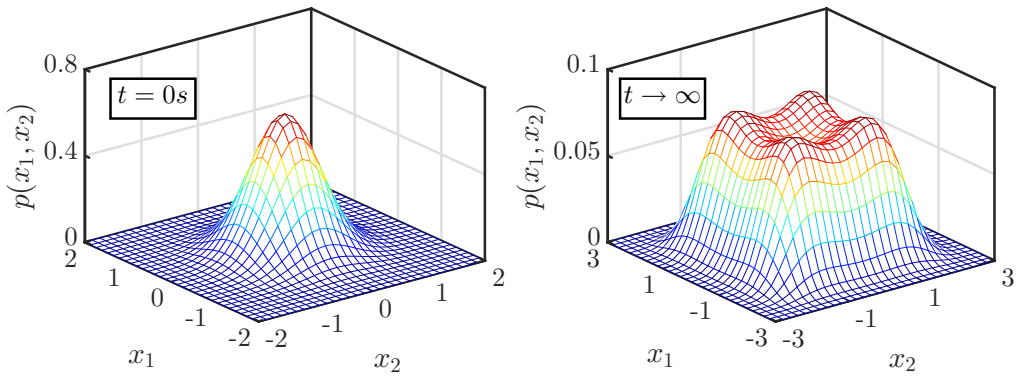

Figure 2.20: The plots above show the initial joint PDF (left) and the stationary joint PDF (right) of the state for the uncoupled nonlinear drift process Eq. (2.46).

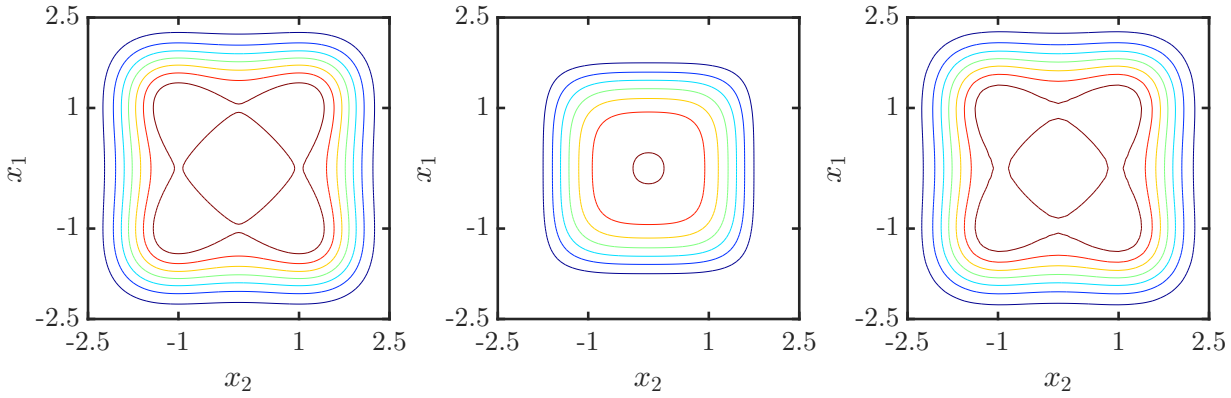

Figure 2.21: Contour plots of the joint PDF at $t = 2\,\mathrm{s}$ for the 2D uncoupled nonlinear drift process Eq. (2.46) obtained by the analytical solution (left), the fixed grid approach (center) and the TPI approach (right).

case is again a multivariate Gaussian distribution with the mean at the origin and a (2×2) diagonal covariance matrix, but with the diagonal elements $[0.25, 0.25]$.

The benefits of the TPI approach over the FG approach for this problem can be seen from the comparison of the contour plots at $t = 2\,\mathrm{s}$ shown in Fig. 2.21 as well as from the marginal PDF plots shown in Fig. 2.22. The chosen fixed grid for the problem is sufficient to adequately represent the initial PDF as shown in Fig. 2.22 (far left). However, unlike the previous case, the FG approach fails in accurately representing the transient behavior of the PDF as seen in Fig. 2.22 (center left and center right). The TPI approach, on the other hand, is able to represent the PDF more accurately than the FG approach.

Considerably better performance (in terms of accuracy) of the TPI approach over the FG approach (for comparable computational time) can also be seen from the data shown in

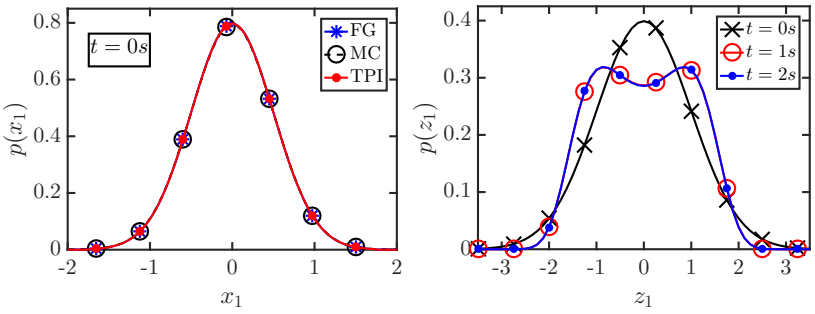

Figure 2.22: Comparison of marginal PDFs at $t = 0\,\mathrm{s}$ (left) and $t = 2\,\mathrm{s}$ (center) in the original space and the transformed space (right), for the 2D uncoupled nonlinear drift process Eq. (2.46).

Table. 2.3. Besides, marginally better performance of the TPI approach over MC (in terms of the accuracy and computational time) is also observed in Table. 2.3.

Example 8: A coupled nonlinear process in 2D

Additionally, we consider a coupled nonlinear process in 2D given by $\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\,\mathrm{d}t + \mathbf{A}\,\mathrm{d}\mathbf{w}$ where $\mathbf{f}(\mathbf{x}, t) = \begin{bmatrix} -\gamma x_1(x_1^2 + x_2^2 - c^2) & -\gamma x_2(x_1^2 + x_2^2 - c^2) \end{bmatrix}^{\mathrm{T}}$ with $\gamma = c = 1$ and the diffusion coefficient matrix $\mathbf{G} \equiv \mathbf{A}\mathbf{A}^{\mathrm{T}}/2$, is a (2×2) diagonal matrix with diagonal elements $G_{11} = G_{22} = g = 0.5$. The initial distribution is a zero mean multivariate Gaussian distribution with a (2×2) diagonal matrix and diagonal elements $[4/49, 4/49]$. A 57×57 grid resolution is chosen for both the fixed grid and the TPI grid, with bounds $[x_l, x_u] = [-1, 1]$ and $[z_l, z_u] = [-7/2, 7/2]$, respectively, for each dimension. The simulation is run from $0\,\mathrm{s}$ to $2\,\mathrm{s}$ in steps of $\Delta t = 0.01\,\mathrm{s}$.

The underlying deterministic system in this example has a stable limit cycle on a circle of radius c. The stationary solution for the PDF is given by

$$p(x_1, x_2) = \frac{1}{\eta} \exp\left[-\frac{\gamma}{4g} \left(x_1^2 + x_2^2\right) \left(x_1^2 + x_2^2 - 2c^2\right) \right] \tag{2.47}$$

where η is a normalization constant. The stationary PDF is shown in Fig. 2.23 (right) where the presence of a limit cycle, characterized by the concentration of the PDF around it, can be seen.

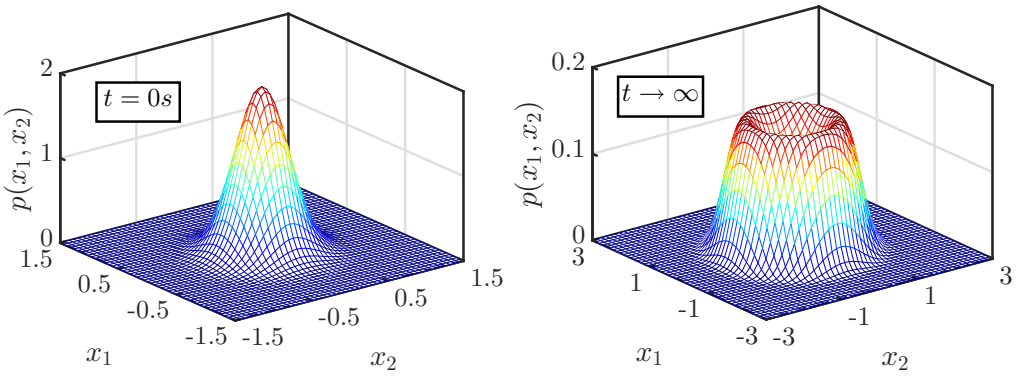

Figure 2.23: The plots above show the initial joint PDF (left) and the stationary joint PDF (right) of the state for the coupled nonlinear process considered in Example 8 (section 3.2.4).

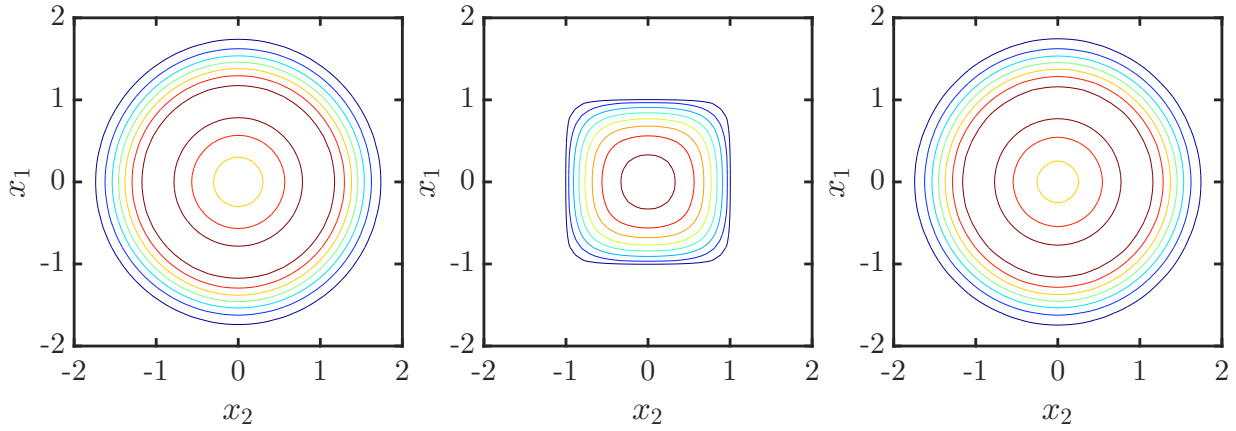

Figure 2.24: Contour plot of the stationary joint PDF (left) and the joint PDF at $t = 2\,\mathrm{s}$ obtained by the fixed grid approach (center) and the TPI approach (right) for the 2D coupled nonlinear process considered in Example 8 (section 3.2.4).

Insufficient resolution in the original space to accurately represent PDF near the limit cycle can contribute to errors in the FG approach. This is observed from the contour plots at $t = 2\,\mathrm{s}$ shown in Fig. 2.24 and the marginal plots shown in Fig. 2.25. The figures also illustrate the ability of the TPI approach to accurately represent the multidimensional non-Gaussian behavior in a coupled nonlinear drift process. The evolution of the marginal PDF in the transformed space is shown in Fig. 2.25 (right).

Accuracy of the TPI approach can also be assessed from the data in table 2.3, where significant benefit over FG approach is observed. The table also shows that for this problem, the TPI approach is faster in comparison to the MC approach for comparable accuracy. Overall, based on Examples 1-8, the TPI approach was found to be more accurate than the

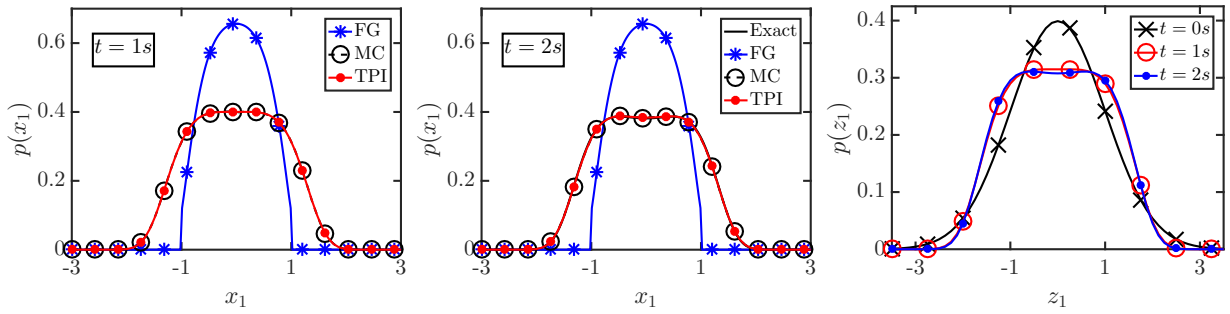

Figure 2.25: Comparison of marginal PDFs in the original space at $t = 1\,\mathrm{s}$ (left) and $t = 2\,\mathrm{s}$ (center) as well as in the transformed space (right) for the 2D coupled nonlinear process considered in Example 8 (section 3.2.4).

		Nonlinear Stochastic Dynamical Systems in 2D					
		Uncoupled (Case C)		Uncoupled (Case D)		Coupled	
	N_{pts}	$T_{\text{comp}}(\text{s})$	ε_{RMS}	$T_{\text{comp}}(\text{s})$	ε_{RMS}	$T_{\text{comp}}(\text{s})$	ε_{RMS}
FG	57×57	114	3.36×10^{-3}	132	2.51×10^{-2}	133	5.06×10^{-2}
MC	10^7	199	8.81×10^{-4}	195	7.68×10^{-4}	205	6.67×10^{-4}
TPI	57×57	172	8.42×10^{-4}	163	7.17×10^{-4}	180	5.99×10^{-4}

Table 2.3: Comparison of computational times and RMS errors in PDFs obtained by fixed grid (FG), Monte Carlo (MC) and transformed path integral (TPI) approaches at $t = 2s$ for the following two-dimensional nonlinear stochastic dynamical systems: the uncoupled nonlinear system with anisotropic diffusion tensor (Case C), the uncoupled nonlinear system with isotropic diffusion tensor (Case D) Eq. (2.46) considered in Example 7 and the coupled nonlinear system considered in Example 8 (section 3.2.4).

FG approach and MC simulations for prediction of linear and nonlinear stochastic dynamical systems in one-dimensional and two-dimensional spaces. This observation is also supported by the data in shown in Tables. 2.1 to 2.3.

2.4 Concluding Remarks

In this chapter, an accurate and efficient path integral based approach has been developed for the solution of the Fokker-Planck equation. The proposed approach, termed the transformed path integral (TPI) approach, has been shown to preserve the non-negativity of the involved distributions as well as the short-time properties of the transition PDF and the underlying

stochastic process. As a part of the approach, a new form of the short time propagator was developed based on a dynamic transformation of the state space with the mean and covariance of the state variables as parameters. The new form of the propagator allows for the propagation to be performed in a transformed computational domain where a more accurate representation of the PDF can be ensured. The dynamic transformation in the TPI approach is such that a fixed grid in the transformed space corresponds to an adaptive grid in the original state space. Hence the establishment of error bounds based on the Chebyshev's inequality, for solutions obtained from the TPI approach, are more straightforward. A new set of update equations was derived for the PDF in the transformed space as well as the mean and covariance of the state variables which enables the TPI approach to better accommodate nonlinear drift processes and non-Gaussian distributions.

The proposed TPI approach is able to better address challenges arising from finite fixed grid bounds, a static computational domain, and a finite fixed grid resolution, which maybe encountered in processes with large diffusion coefficients, large drift vectors, or large concentrations of the PDF. The benefits of the TPI approach and limitations of conventional approaches such as Monte Carlo simulations and fixed grid based approaches, in tackling these challenges were clearly illustrated through the pure diffusion, the constant drift and, the Ornstein-Uhlenbeck processes (in 1D). Additionally, the TPI approach solutions show much better agreement with the analytical solutions than those obtained from conventional approaches. Simulations of a nonlinear drift process showed similar benefits and good agreement with the stationary solution for the TPI approach over conventional approaches. Analysis of the error in the PDF as compared with the analytical solutions in the linear case and the stationary solution in the nonlinear case show that the TPI approach is more efficient than conventional approaches in terms of the computational costs involved to obtain a desired level of accuracy, especially for problems in one dimensional space. Our analysis also indicated the possibility of the existence of a similarity parameter. Our proposed candidate for the parameter given by $\alpha = \sigma_{ref}^2 \, (\Delta z)^2/(2a^2\Delta t)$ yielded a partial collapse of the various

error curves. It was shown that there is further reduction of the error for $\alpha > 10$ or $\alpha < 0.5$. The extension of the TPI approach to problems in multidimensional space was illustrated through simulations of linear and nonlinear problems in a two dimensional space.

Although, the TPI approach addresses the domain issues associated with conventional fixed grid approaches, issues relevant to the curse of dimensionality need to be investigated further. Another challenge for the conventional PI approaches and inherited by the TPI approach is handling of stochastic processes with a singular diffusion matrix. These issues will be addressed in the next chapter.

CHAPTER 3

The Generalized Transformed Path Integral Approach for Stochastic Processes

3.1 Scope of the Chapter

In this chapter, we present the generalized transformed path integral approach (GTPI): a grid-based path integral approach for PDF evolution in a large class of dynamical systems. We showcase the application of our proposed approach to nonsingular systems, second order (and higher order) dynamical systems, dynamical systems with zero process noise, and certain dynamical systems with non-white noise excitation. As a part of the approach, we present a reformulation of the problem in terms of the solution of a complementary system—the standard transformed stochastic dynamical system—obtained through a dynamic transformation of the state variables. The state mean and covariance of the transformed system do not change with evolution and the choice of our transformation parameters ensure that they are zero and identity respectively. This allows us to consider PDF propagation in a "well contained" transformed space where greater numerical accuracy for the distribution can be ensured. A fixed grid in the transformed space corresponds to an adaptive grid in the original space coordinates which allow our approach to better address the challenges of large drift, diffusion, and concentration of PDF. In addition, error bounds for distributions in the transformed space can be easily obtained using Chebyshev's inequality. Using a splitting solution of the Fokker-Planck equation in the transformed space, we derive a set of update equations for the state PDF, its mean, and covariance. Illustrative examples were considered to showcase the benefits of GTPI over conventional fixed grid (FG) approaches in the afore-

mentioned stochastic dynamical systems. In all the cases, the GTPI results show excellent agreement with results from MC simulation and available analytical (and stationary) solutions, while FG results show large errors. The effect of simulation parameters and system parameters on the numerical error in our approach were also studied.

3.2 The Transformed Path Integral Approach: An Overview

Dynamical systems subjected to Gaussian white noise excitation can be described by the Itô stochastic differential equation (SDE)

$$\mathrm{d}\mathbf{x}(t) = \mathbf{f}\Big(\mathbf{x}(t), t\Big)\,\mathrm{d}t + \mathbf{A}\Big(\mathbf{x}(t), t\Big)\,\mathrm{d}\mathbf{w}(t) \tag{3.1}$$

where $\mathbf{x}(t) \in \mathbb{R}^{N_s \times 1}$ represents the state of the system at time t. The equation describes evolution of the state of a system along a deterministic path specified by the drift vector function $\mathbf{f}\Big(\mathbf{x}(t), t\Big) \in \mathbb{R}^{N_s \times 1}$ while being subjected to random excitations modeled as Gaussian white noise of strength $\mathbf{A}\Big(\mathbf{x}(t), t\Big) \in \mathbb{R}^{N_s \times N_w}$. The increments $\mathrm{d}\mathbf{w}(t) \in \mathbb{R}^{N_w \times 1}$ are independent and identically distributed zero mean Gaussian random vectors with the autocorrelation $\Big\langle \mathrm{d}w_i(t)\,\mathrm{d}w_j(s) \Big\rangle = \delta(t-s)\,\delta_{ij}\,\mathrm{d}t$ for $i, j = 1, \ldots, N_w$. The stochastic process $\mathbf{w}(t)$ is the Wiener process.

The state of the system can also be characterized by a probability density function $p(\mathbf{x}, t)$ whose evolution is governed by the Fokker-Planck equation (FPE), a second order partial differential equation of the parabolic type given by

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial x_i} f_i(\mathbf{x}, t) - \frac{\partial^2}{\partial x_i \partial x_j} G_{ij}(\mathbf{x}, t)\right] p(\mathbf{x}, t) = 0 \tag{3.2}$$

where we have used Einstein's notation convention, i.e., repeated indices imply summation. Note that the spatial variables $\mathbf{x}$ used in Eq. (3.2) is not the same as the random vector $\mathbf{x}(t)$ representing the state of the dynamical system in Eq. (3.1). Nevertheless, the Fokker-

Planck equation features the drift vector function $\mathbf{f}(\mathbf{x}, t)$ while the strength of the white noise excitation $\mathbf{A}(\mathbf{x}, t)$ is related to the diffusion coefficient matrix $\mathbf{G}(\mathbf{x}, t)$ as $\mathbf{G}(\cdot) = \mathbf{A}(\cdot)\mathbf{A}(\cdot)^{\mathrm{T}}/2$. For the sake of brevity, wherever the context is clear we will represent both the state variables and their associated spatial variables with the same symbol. Exact analytical solutions to the Fokker-Planck equation are known only for certain cases [14, 16, 18–20]. In most other cases, approximate solutions based on either analytical [22, 23, 28] or numerical methods [30, 33, 34] are sought.

Several authors [45, 84–86] pursued an alternative approach, namely to seek a solution to the Fokker-Planck equation in the form of a path integral:

$$p(\mathbf{x}, t|\mathbf{x_0}, t_0) = \int_{\mathbf{x}(0)=\mathbf{x_0}}^{\mathbf{x}(t)=\mathbf{x}} \mathcal{D}\mu\Big(\mathbf{x}(\tau)\Big)\, \exp\Big\{-\int_{t_0}^{t} \mathrm{d}\tau\, \mathcal{L}(\mathbf{x}(\tau), \dot{\mathbf{x}}(\tau), \tau)\Big\}. \tag{3.3}$$

It represents a functional integral with integration measure $\mathcal{D}\mu\Big(\mathbf{x}(\tau)\Big)$ over paths $\mathbf{x}(\tau)$ between states $\mathbf{x}(0) = \mathbf{x_0}$ and $\mathbf{x}(t) = \mathbf{x}$. The path integral approach presents another, albeit formal, mechanism to describe the behavior of dynamical systems; and the Onsager-Machlup functional $\mathcal{L}$, if it can be specified, is similar to the Lagrangian of a dynamical system [44]. However, a general form of the functional, especially for systems with multiple degrees of freedom, has not been uniquely specified.

Stratonovich [84], for instance, derived an expression for the functional by evaluating probability along the paths in a small neighborhood of a smooth path:

$$\begin{aligned}
\mathcal{L}\Big(\mathbf{x}(t), \dot{\mathbf{x}}(t), t\Big) &= \frac{1}{2}\left[\frac{1}{2}\,\dot{x}_i\Big(\mathbf{G}^{-1}\Big)_{ij}\dot{x}_j - 2\,m_i\Big(\mathbf{A}^{-1}\Big)_{ij}\dot{x}_j + m_i\,m_j + A_{ij}\frac{\partial m_i}{\partial x_j}\right]\mathrm{d}t \\
m_i &= \Big(\mathbf{A}^{-1}\Big)_{ij}\left(f_j - \frac{1}{2}\frac{\partial A_{kl}}{\partial x_r}\,A_{kr}\right)
\end{aligned}$$

where for the sake of brevity we have used index notations x_i, $\dot{x}_i$, and m_i for $\mathbf{x}(t)$, $\dot{\mathbf{x}}(t)$, and $\mathbf{m}\Big(\mathbf{x}(t), t\Big)$. Also, $\Big(\mathbf{G}^{-1}\Big)_{ij}$ and $\Big(\mathbf{A}^{-1}\Big)_{ij}$ represent the (i, j) term of the matrix inverse of the corresponding quantities. More expressions for the functional were derived using operator

calculus [45], Fourier sum representations of stochastic paths [85], and from the asymptotic behavior of transition probability density for a short time interval [86].

In contrast to the formal continuous forms, other scholars [43, 87] investigated more mathematically rigorous definitions of the path integral such as the limit

$$p(\mathbf{x},t) = \lim_{N\to\infty} \int\!\cdots\!\int \prod_{k=0}^{N-1} (\mu_k \, \mathrm{d}\mathbf{x_k}) \exp\left\{-\Delta t \sum_{k=0}^{N-1} \mathcal{L}(\mathbf{x_{k+1}}, \mathbf{x_k}, \Delta t)\right\} p(\mathbf{x_0}, t_0) \tag{3.4}$$

along with an expression for $\mathcal{L}(\mathbf{x_{k+1}}, \mathbf{x_k}, \Delta t)$ and a discretization prescription to go from Eq. (3.3) to Eq. (3.4). In this representation, the time axis is discretized on an equidistant lattice $[t_0, t_1, \ldots, t_k, t_{k+1}, \ldots, t_{N-1}, t]$ with time step $\Delta t = t_{k+1} - t_k = (t - t_0)/N$ while $\mathbf{x_{k+1}}$ and $\mathbf{x_k}$ represent the state of the system at times t_{k+1} and t_k respectively. Wissel showed [46] that a whole family of expressions for the path integral of the form of Eq. (3.4) can be derived based on the different ways of representing the δ-function. Irrespective of the choice of the Onsager-Machlup functional, the evaluation of the path integral presents a daunting task.

Numerical methods proposed for evaluating the path integral Eq. (3.4) have largely fallen under two broad categories: Monte Carlo based simulations [88–90] and lattice based approaches employing discretizations of the Onsager-Machlup functional [47, 49, 91, 92]. The latter involves repeated application of the Chapman-Kolmogorov equation

$$p(\mathbf{x_{k+1}}, t_k + \Delta t) = \int p(\mathbf{x_{k+1}}, t_k + \Delta t | \mathbf{x_k}, t_k) \, p(\mathbf{x_k}, t_k) \, \mathrm{d}\mathbf{x_k} \tag{3.5}$$

where the transition probability density function

$$p(\mathbf{x_{k+1}}, t_k + \Delta t | \mathbf{x_k}, t_k) = \mu_k \exp\Big\{ - \Delta t \mathcal{L}(\mathbf{x_{k+1}}, \mathbf{x_k}, t_k, \Delta t)\Big\} \tag{3.6}$$

is often referred to as the short-time propagator of the system. Wehner and Wolfer [47,91,92]

developed a numerical procedure utilizing the following form of the short-time propagator

$$p(\mathbf{x_{k+1}}, t_k + \Delta t \,|\, \mathbf{x_k}, t_k) = (4\pi\Delta t)^{-N_s/2} \left\|\mathbf{G}(\mathbf{x_k}, t_k)\right\|^{-1/2} \exp\left\{-\frac{1}{4\,\Delta t}\left[\mathbf{x_e}^{\mathrm{T}}\,\mathbf{G}(\mathbf{x_k}, t_k)^{-1}\,\mathbf{x_e}\right]\right\} \tag{3.7}$$

with $\mathbf{x_e} = \mathbf{x_{k+1}} - \mathbf{x_k} - \mathbf{f}(\mathbf{x_k}, t_k)\,\Delta t$. Their procedure involved a histogram-based discretization of the state space to reduce the convolution integral in Eq. (3.5) to a matrix-vector multiplication; while Naess and Johnsen [49] used cubic B-splines which offer greater flexibility and accuracy especially for dynamical systems with cubic nonlinearities. Such conventional "fixed-grid" implementations of the path integral approach, however, are unable to efficiently handle challenges posed by processes with large drift, large diffusion, or large concentration in the probability density function.

In an earlier work [93], we explored the limitations of fixed-grid approaches and proposed the transformed path integral (TPI) approach to address those limitations. The approach allows for the propagation of the probability density function to be performed in a transformed computational domain using a novel short-time propagator given by

$$p(\mathbf{z}', t + \Delta t \,|\, \mathbf{z}, t) = (4\,\pi\,\Delta t)^{-N_s/2}\,\|\mathbf{R}(t)\|\,\left\|\tilde{\mathbf{G}}(\mathbf{z}, t)\right\|^{-1/2} \exp\left\{-\frac{1}{4\,\Delta t}\left[\mathbf{z_e}^{\mathrm{T}}\,\tilde{\mathbf{G}}(\mathbf{z}, t)^{-1}\,\mathbf{z_e}\right]\right\} \tag{3.8}$$

where $\mathbf{z_e} = \Delta\,(\mathbf{R}\,\mathbf{z}) - \Delta t\,\boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)$ and $\boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t) = \tilde{\mathbf{f}}(\mathbf{z}, t) - \left\langle\tilde{\mathbf{f}}(\mathbf{z}, t)\right\rangle$. We have used $\left\langle\,\cdot\,\right\rangle$ to denote expectation. In addition, here, $\tilde{\mathbf{f}}(\mathbf{z}, t) = \mathbf{f}(\mathbf{R}\mathbf{z} + \boldsymbol{\mu}, t)$ and $\tilde{\mathbf{G}}(\mathbf{z}, t) = \mathbf{G}(\mathbf{R}\mathbf{z} + \boldsymbol{\mu}, t)$ are functions in the transformed space generated by the dynamic transformation

$$\mathbf{z} = \mathcal{Z}\Big(\mathbf{x}, t\Big) = \mathbf{R}(t)^{-1}\Big[\mathbf{x} - \boldsymbol{\mu}(t)\Big] \tag{3.9}$$

with $\boldsymbol{\mu}(t)$ and $\boldsymbol{\Sigma}(t) \equiv \mathbf{R}(t)\,\mathbf{R}(t)^{\mathrm{T}}$ denoting, respectively, the mean and covariance of the state $\mathbf{x}(t)$ at time t. Under this transformation, a fixed grid in the transformed space ($\mathbf{z}$-space)

corresponds to a dynamically adaptive grid in the original space ($\mathbf{x}$-space). The dynamic grid translates with the state mean and scales with the state covariance. Thus, our proposed approach is better equipped to address the challenges arising from large drift, diffusion, and concentration of PDF. We used $\Delta\Big(\mathbf{Rz}\Big) = \mathbf{R}'\mathbf{z}' - \mathbf{Rz}$ where the primed variables denote the corresponding quantities at time $t + \Delta t$. Other equivalent versions of the transformed path integral short-time propagator up to order $\mathcal{O}\left((\Delta t)^{3/2}\right)$ can be obtained for different discretization prescriptions of $\Delta\left(\mathbf{Rz}\right)$. The state distribution in the transformed space is evolved via the corresponding Chapman-Kolmogorov equation:

$$p(\mathbf{z}', t + \Delta t) = \int p(\mathbf{z}', t + \Delta t \,|\, \mathbf{z}, t)\, p(\mathbf{z}, t)\, \mathrm{d}\mathbf{z}. \tag{3.10}$$

In addition, the necessary update equations for the state mean and covariance were obtained by taking expectations of Eq. (3.1):

$$\boldsymbol{\mu}' = \boldsymbol{\mu} + \Delta t \left\langle \tilde{\mathbf{f}}(\mathbf{z}, t) \right\rangle, \tag{3.11}$$

$$\begin{aligned} \boldsymbol{\Sigma}' = \boldsymbol{\Sigma} + 2\,\Delta t \left\langle \tilde{\mathbf{G}}(\mathbf{z}, t) \right\rangle + \Delta t \left\langle \mathbf{Rz}\, \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)^{\mathrm{T}} + \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)\, \mathbf{z}^{\mathrm{T}}\, \mathbf{R}^{\mathrm{T}} \right\rangle \\ + (\Delta t)^2 \left\langle \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)\, \boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{z}, t)^{\mathrm{T}} \right\rangle. \end{aligned} \tag{3.12}$$

Eqs. (3.8) to (3.12) form the basis of the transformed path integral approach. It has several desirable qualities: potential for accurate representation of PDFs—especially the tail information—in processes with large drift, diffusion, or concentration of PDF; accurate estimates of error bounds on the numerical representations of evolved PDFs via the Chebyshev's inequality; and preservation of stochastic properties of the underlying dynamical system. A limitation of the TPI approach is the restriction of its applicability to stochastic dynamical systems with a nonsingular, i.e., full rank diffusion coefficient matrix.

3.3 The Generalized Transformed Path Integral Approach

A large category of stochastic dynamical systems is characterized by the presence of a rank deficient diffusion coefficient matrix. Second order dynamical systems where the random excitation acts along only some of the dimensions fall within this category [57,58]. As would systems excited by colored noise where the excitation can be represented as a filtered white noise process [59–63]. These "singular systems" present challenges to the application of conventional lattice-based path integral implementations. The limitation is also present in the TPI approach because the singularity is preserved under the transformation Eq. (3.9). Since singular matrices are non-invertible, the short-time propagators defined in Eqs. (3.7) and (3.8) cannot be used to obtain solutions of these dynamical systems.

Some of the solutions proposed to address this challenge involve Taylor series expansion to arrive at an SDE, and hence a short time propagator, with a nonsingular diffusion matrix; power series expansion of the cumulant generating function [64]; and splitting of the Fokker-Planck equation into individual contributions using the Trotter product formula [65, 66]. These approaches, however, do not address the issues associated with conventional lattice-based implementations in the way the transformed path integral approach does for nonsingular systems. Motivated by this unaddressed challenge, we propose a more general transformed path integral-based approach applicable to both singular and nonsingular systems.

3.3.1 Mathematical formulations

The starting point for our proposed approach is the dynamic transformation of the state variables

$$\mathbf{z}(t) = \mathcal{Z}\Big(\mathbf{x}(t), t\Big) = \mathbf{R}(t)^{-1}\Big[\mathbf{x}(t) - \boldsymbol{\mu}(t)\Big] \tag{3.13}$$

with yet to be specified functions $\boldsymbol{\mu}(t) \in \mathbb{R}^{N_s \times 1}$ and $\mathbf{R}(t) \in \mathbb{R}^{N_s \times N_s}$. The distribution of the transformed state variables is related to that of the original state variables as

$$p_{\mathbf{x}(t)}(\mathbf{x}, t) = \frac{1}{\|\mathbf{R}(t)\|} \, p_{\mathbf{z}(t)}\Big(\mathbf{R}(t)^{-1}\Big[\mathbf{x} - \boldsymbol{\mu}(t)\Big], t\Big). \tag{3.14}$$

In further discussions, we will drop the subscripts in our representations of the distribution functions and rely on the reader to distinguish between them based on the context. Note that,

$$\boldsymbol{\mu}(t) = \langle \mathbf{x} \rangle - \mathbf{R}(t) \, \langle \mathbf{z} \rangle \,, \tag{3.15}$$

$$\mathbf{R}(t)\Big[\Big\langle \mathbf{z}\mathbf{z}^{\mathrm{T}} \Big\rangle - \langle \mathbf{z} \rangle \langle \mathbf{z} \rangle^{\mathrm{T}}\Big]\mathbf{R}(t)^{\mathrm{T}} = \Big\langle \mathbf{x} - \langle \mathbf{x} \rangle \Big\rangle \Big\langle \mathbf{x} - \langle \mathbf{x} \rangle \Big\rangle^{\mathrm{T}}. \tag{3.16}$$

Thus, defining our transformed space, i.e., specifying $\boldsymbol{\mu}(t)$ and $\mathbf{R}(t)$, is equivalent to specifying the desired state mean and covariance of the transformed state variables at time t.

Given the transformation in Eq. (3.13), we know from Itô's lemma that $\mathbf{z}(t)$ is also a stochastic process of the form of Eq. (3.1). Specifically, we now have a complementary stochastic dynamical system whose evolution is described by the Itô SDE

$$\mathrm{d}\mathbf{z}(t) = \boldsymbol{\Phi}\Big(\mathbf{z}(t), t\Big) \, \mathrm{d}t + \boldsymbol{\Lambda}\Big(\mathbf{z}(t), t\Big) \, \mathrm{d}\mathbf{w}(t). \tag{3.17}$$

Its drift vector function $\boldsymbol{\Phi}$ and diffusion coefficient matrix $\boldsymbol{\Gamma} \equiv \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{T}}$ are given by

$$\boldsymbol{\Phi}(\mathbf{z}, t) = \mathbf{R}(t)^{-1}\Big(\tilde{\mathbf{f}}(\mathbf{z}, t) - \dot{\mathbf{R}}(t) \, \mathbf{z} - \dot{\boldsymbol{\mu}}(t)\Big) \tag{3.18}$$

$$\boldsymbol{\Gamma}(\mathbf{z}, t) = \mathbf{R}(t)^{-1} \, \tilde{\mathbf{G}}(\mathbf{z}, t) \, \mathbf{R}(t)^{-1^{\mathrm{T}}} \tag{3.19}$$

where $\tilde{\mathbf{f}}(\mathbf{z}, t) = \mathbf{f}\Big(\mathbf{R}(t) \, \mathbf{z} + \boldsymbol{\mu}(t), t\Big)$ and $\tilde{\mathbf{G}}(\mathbf{z}, t) = \mathbf{G}\Big(\mathbf{R}(t) \, \mathbf{z} + \boldsymbol{\mu}(t), t\Big)$. Evolution of the state distribution for this dynamical system is governed by the corresponding Fokker-Planck

equation, henceforth referred to as the transformed Fokker-Planck equation:

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial z_i}\Phi_i(\mathbf{z},t) - \frac{\partial^2}{\partial z_i \partial z_j}\Gamma_{ij}(\mathbf{z},t)\right] p(\mathbf{z},t) = 0. \tag{3.20}$$

Eqs. (3.17) to (3.20) allow us to describe the evolution of a stochastic dynamical system by considering the evolution of a complementary system obtained through the transformation in Eq. (3.13) for some specified $\boldsymbol{\mu}(t)$ and $\mathbf{R}(t)$. We will refer to this complementary system as the transformed stochastic dynamical system. Let us seek a transformation that will preserve the mean and covariance of the transformed state variables with propagation. In other words, we seek $\boldsymbol{\mu}(t)$ and $\mathbf{R}(t)$ such that

$$\frac{\partial}{\partial t}\langle \mathbf{z} \rangle = \left\langle \boldsymbol{\Phi}(\mathbf{z},t) \right\rangle = \mathbf{0}, \tag{3.21}$$

$$\frac{\partial}{\partial t}\left\langle \mathbf{z}\mathbf{z}^{\mathrm{T}} \right\rangle = \left\langle \mathbf{z}\,\boldsymbol{\Phi}(\mathbf{z},t)^{\mathrm{T}} \right\rangle + \left\langle \boldsymbol{\Phi}(\mathbf{z},t)\,\mathbf{z}^{\mathrm{T}} \right\rangle + 2\left\langle \boldsymbol{\Gamma}(\mathbf{z},t) \right\rangle = \mathbf{0}. \tag{3.22}$$

Substituting Eqs. (3.18) and (3.19) in Eqs. (3.21) and (3.22) we obtain the coupled ordinary differential equations

$$\dot{\boldsymbol{\mu}}(t) = \left\langle \tilde{\mathbf{f}}(\mathbf{z},t) \right\rangle - \dot{\mathbf{R}}(t)\,\langle \mathbf{z} \rangle\,, \tag{3.23}$$

$$\begin{aligned}\mathbf{R}(t)\left[\left\langle \mathbf{z}\mathbf{z}^{\mathrm{T}} \right\rangle - \langle \mathbf{z} \rangle \langle \mathbf{z} \rangle^{\mathrm{T}}\right]\dot{\mathbf{R}}(t)^{\mathrm{T}} + \dot{\mathbf{R}}(t)\left[\left\langle \mathbf{z}\mathbf{z}^{\mathrm{T}} \right\rangle - \langle \mathbf{z} \rangle \langle \mathbf{z} \rangle^{\mathrm{T}}\right]\mathbf{R}(t)^{\mathrm{T}} \\ = \mathbf{R}(t)\left\langle \mathbf{z}\,\delta\tilde{\mathbf{f}}(\mathbf{z},t)^{\mathrm{T}} \right\rangle + \left\langle \delta\tilde{\mathbf{f}}(\mathbf{z},t)\,\mathbf{z}^{\mathrm{T}} \right\rangle \mathbf{R}(t)^{\mathrm{T}} + 2\left\langle \tilde{\mathbf{G}}(\mathbf{z},t) \right\rangle\end{aligned} \tag{3.24}$$

where $\delta\tilde{\mathbf{f}}(\mathbf{z},t) = \tilde{\mathbf{f}}(\mathbf{z},t) - \left\langle \tilde{\mathbf{f}}(\mathbf{z},t) \right\rangle$. Thus, given a stochastic dynamical system of the form of Eq. (3.1), we can construct a transformed stochastic dynamical system where the state mean and covariance do not change with the evolution of the system. In particular, we consider the transformed system with zero mean and identity covariance; we will refer to this system as the "standard transformed stochastic dynamical system". Clearly, then $\boldsymbol{\mu}(t)$ is the state mean in the original space and $\boldsymbol{\Sigma} \equiv \mathbf{R}(t)\mathbf{R}(t)^{\mathrm{T}}$ the corresponding state covariance. Making these substitutions in Eqs. (3.23) and (3.24) while noting that $\dot{\boldsymbol{\Sigma}}(t) = \dot{\mathbf{R}}(t)\mathbf{R}(t)^{\mathrm{T}} + \mathbf{R}(t)\dot{\mathbf{R}}(t)^{\mathrm{T}}$

we obtain

$$\dot{\boldsymbol{\mu}}(t) = \left\langle \tilde{\mathbf{f}}(\mathbf{z}, t) \right\rangle, \tag{3.25}$$

$$\dot{\boldsymbol{\Sigma}}(t) = \mathbf{R}(t) \left\langle \mathbf{z}\, \delta\tilde{\mathbf{f}}(\mathbf{z}, t)^{\mathrm{T}} \right\rangle + \left\langle \delta\tilde{\mathbf{f}}(\mathbf{z}, t)\, \mathbf{z}^{\mathrm{T}} \right\rangle \mathbf{R}(t)^{\mathrm{T}} + 2 \left\langle \tilde{\mathbf{G}}(\mathbf{z}, t) \right\rangle. \tag{3.26}$$

Eqs. (3.20), (3.25) and (3.26) represent another way of describing the evolution of stochastic dynamical systems; one where the evolution of the corresponding system in a "well contained" transformed space can be used to describe the state of the system in the original space. The aforementioned equations are the governing equations for such an approach.

3.3.2 Salient features

Since the state mean and covariance of the transformed state variables are constant in our approach, we are able to consider more accurate computational domains for propagating the state distribution of the transformed variables. In fact, we are able to capture certain invariant solutions of the standard transformed stochastic dynamical system as shown below.

Lemma 3.3.1. *For a stochastic dynamical system of the form of Eq. (3.1) with a linear drift vector function (with respect to the state variables) and a constant diffusion coefficient matrix, the standard normal distribution is a stationary solution for the corresponding standard transformed stochastic dynamical system.*

Proof. Consider a process with a linear drift vector function given by $\mathbf{f}(\mathbf{x}, t) = \mathbf{M}\mathbf{x} + \mathbf{c}$. The drift vector function of the standard transformed stochastic dynamical system for this case is given by $\boldsymbol{\Phi}(\mathbf{z}, t) = \boldsymbol{\alpha}(t)\, \mathbf{z}$ where $\boldsymbol{\alpha}(t) = \mathbf{R}(t)^{-1} \left[\mathbf{M}\mathbf{R}(t) - \dot{\mathbf{R}}(t) \right]$. A stationary solution of this system needs to satisfy

$$\left[\alpha_{ii} + \alpha_{ij} z_j \frac{\partial}{\partial z_i} - \Gamma_{ij} \frac{\partial^2}{\partial z_i \partial z_j} \right] p(\mathbf{z}, t) = 0 \tag{3.27}$$

which is just the stationary version of the transformed Fokker-Planck equation Eq. (3.20) for

this system. For a standard normal distribution, $p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$, the gradient and Hessian matrix are given by

$$\frac{\partial p}{\partial z_i} = -z_i \, p(\mathbf{z}) \tag{3.28}$$

$$\frac{\partial^2 p}{\partial z_i \partial z_j} = \left[z_i z_j - \delta_{ij}\right] \, p(\mathbf{z}). \tag{3.29}$$

Consequently, the left hand side of Eq. (3.27) becomes $\left[\mathrm{Tr}\left(\boldsymbol{\alpha} + \boldsymbol{\Gamma}\right) - \mathbf{z}^{\mathrm{T}} \left(\boldsymbol{\alpha} + \boldsymbol{\Gamma}\right) \mathbf{z}\right] p(\mathbf{z})$ for the standard normal distribution. From the covariance update in Eq. (3.26) for this system (or equivalently from Eq. (3.22)) we get the relation $\left[\boldsymbol{\alpha} + \boldsymbol{\Gamma}\right] + \left[\boldsymbol{\alpha} + \boldsymbol{\Gamma}\right]^{\mathrm{T}} = 0$. In other words, $\boldsymbol{\alpha} + \boldsymbol{\Gamma}$ is a skew-symmetric matrix. We know that the trace of a skew-symmetric matrix is zero and so also is the matrix quadratic form $\mathbf{z}^{\mathrm{T}} \left(\boldsymbol{\alpha} + \boldsymbol{\Gamma}\right) \mathbf{z}$ for all $\mathbf{z} \in \mathbb{R}^{N_s \times 1}$. Therefore, the standard normal distribution is a solution of Eq. (3.27) and consequently is a stationary solution of the system. □

In our approach, we are able to easily establish PDF error bounds based on domain bounds. Let us consider a computational domain $\mathbf{S}$ on which the distribution is represented. An estimate of the error in PDF due to loss of information outside the bounds of this computational domain is given by

$$\varepsilon = 1 - \int\limits_{\boldsymbol{\xi} \in \mathbf{S}} p(\boldsymbol{\xi}, t) \, \mathrm{d}\boldsymbol{\xi} \tag{3.30}$$

For a multivariate standard normal distribution on a rectangular computational domain with domain bounds $\begin{bmatrix} -k & k \end{bmatrix}$ in each dimension, we have $\varepsilon = 1 - \left[\mathrm{erf}(k/\sqrt{2})\right]^{N_s}$. This means we can represent 99.46 % of the distribution by considering a transformed computational domain with bounds $\begin{bmatrix} -3 & 3 \end{bmatrix}$ in either dimensions for a two-dimensional system with a linear drift and constant diffusion coefficient. Such an analysis involving error bounds can also be extended to systems where the distribution remains approximately Gaussian or near Gaussian distributions.

For the general case, we may obtain error bounds via the multidimensional Chebyshev's inequality which for a random vector $\mathbf{x} \in \mathbb{R}^{N_s \times 1}$ with mean $\boldsymbol{\mu}$, covariance $\boldsymbol{\Sigma}$, and any real number $k > 0$ is given by

$$Pr\Big((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \geq k^2\Big) \leq \frac{N_s}{k^2} \tag{3.31}$$

In the transformed space, we have a more succinct expression given by

$$Pr\left(\mathbf{z}^T \mathbf{z} \geq k^2\right) \leq \frac{N_s}{k^2} \tag{3.32}$$

In two-dimensional systems, $\mathbf{z}^T \mathbf{z} \geq k^2$ represents the region outside a circle of radius k in the transformed space. Let us consider a square region $\mathbf{S}$ with side length $2k$ circumscribing the circular region. It represents a rectangular computational domain in the transformed space with bounds $\begin{bmatrix} -k & k \end{bmatrix}$ in either dimensions. This can be extended to multi-dimensional systems, for instance, a cubical region of side length $2k$ enclosing spherical region of radius k. Thus, for the general case we have the relation (inequality) for the error bounds given by

$$\varepsilon = 1 - Pr\Big(\mathbf{z} \in \mathbf{S}\Big) \leq \frac{N_s}{k^2}. \tag{3.33}$$

3.3.3 Update equations

Having established some of the salient features of the standard transformed stochastic dynamical system, we seek a path integral based solution for the system. Path integral based solutions have several desirable qualities, for instance, preserving the non-negativity of the distributions. However, if $\mathbf{G}$ is singular, and hence $\boldsymbol{\Gamma}$, the short time propagator in the transformed space given by Eq. (3.8) cannot be used to solve Eq. (3.20). For such a system, through appropriate transformation of the state variables where necessary, $\boldsymbol{\Gamma}$ can be

expressed in the block matrix form

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma^{(r)}} \end{bmatrix} \tag{3.34}$$

where $\mathbf{0}$ represents appropriate zero-matrices. Note that in this representation $\mathbf{\Gamma^{(r)}} \in \mathbb{R}^{N_r \times N_r}$ is nonsingular and $N_r = \text{rank}\,(\mathbf{\Gamma})$. Let $N_\varepsilon = N_s - N_r$. Without loss of generality, we may assume the dynamical system and $\mathbf{\Gamma}$ to already be of this form.

In our approach, informed by the form of $\mathbf{\Gamma}$, we consider grouping the transformed state variables $\mathbf{z} = [z_1, z_2, \ldots, z_{N_s}]^{\mathrm{T}}$ into singular variables $\mathbf{q}$ and nonsingular variables $\mathbf{v}$ as

$$\mathbf{q} = \begin{bmatrix} z_1 \\ \vdots \\ z_{N_\varepsilon} \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} z_{N_\varepsilon+1} \\ \vdots \\ z_{N_s} \end{bmatrix} \tag{3.35}$$

Similarly, we define $\mathbf{\Phi^{(\varepsilon)}}(\cdot) = \left[\Phi_1^{(\varepsilon)}(\cdot), \ldots, \Phi_{N_\varepsilon}^{(\varepsilon)}(\cdot)\right]^{\mathrm{T}}$ and $\mathbf{\Phi^{(r)}}(\cdot) = \left[\Phi_{N_\varepsilon+1}^{(r)}(\cdot), \ldots, \Phi_{N_s}^{(r)}(\cdot)\right]^{\mathrm{T}}$ allowing us to write Eq. (3.20) in the form

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial u_i}\Phi_i^{(\varepsilon)}(\mathbf{q}, \mathbf{v}, t) + \frac{\partial}{\partial v_i}\Phi_i^{(r)}(\mathbf{q}, \mathbf{v}, t) - \frac{\partial^2}{\partial v_i \partial v_j}\Gamma_{ij}^{(r)}(\mathbf{q}, \mathbf{v}, t)\right] p(\mathbf{q}, \mathbf{v}, t) = 0 \tag{3.36}$$

or more succinctly as $\partial p/\partial t = \left[\mathcal{L}_{\mathrm{I}} + \mathcal{L}_{\mathrm{II}}\right] p$ with the operators

$$\mathcal{L}_{\mathrm{I}} = -\frac{\partial}{\partial u_i}\Phi_i^{(\varepsilon)}(\mathbf{q}, \mathbf{v}, t) \tag{3.37}$$

$$\mathcal{L}_{\mathrm{II}} = -\frac{\partial}{\partial v_i}\Phi_i^{(r)}(\mathbf{q}, \mathbf{v}, t) + \frac{\partial^2}{\partial v_i \partial v_j}\Gamma_{ij}^{(r)}(\mathbf{q}, \mathbf{v}, t) \tag{3.38}$$

operating on $p = p(\mathbf{q}, \mathbf{v}, t)$. A solution based on operator splitting [94] is given by

$$p(\mathbf{q}', \mathbf{v}', t + \Delta t) = e^{[\Delta t \mathcal{L}_{\mathrm{II}}]}\, e^{[\Delta t \mathcal{L}_{\mathrm{I}}]}\, p(\mathbf{q}, \mathbf{v}, t) + \mathcal{O}\left(\Delta t\right). \tag{3.39}$$

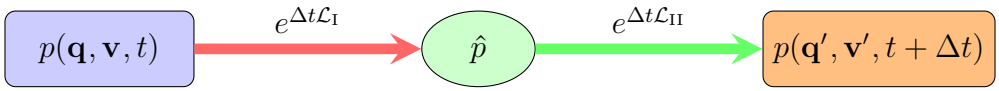

Figure 3.1: A splitting solution for PDF propagation in the transformed space

where, for notational simplicity, we use primed variables to denote corresponding quantities at time $t+\Delta t$. Equivalently, Eq. (3.39) may be interpreted as the two step iterative sequence (see also Fig. 3.1): One may also employ symmetric Trotter splitting or other high-order splitting schemes [95] in place of Eq. (3.39). The $\mathcal{L}_{\mathrm{I}}$ operator corresponds to a Liouville equation, while $\mathcal{L}_{\mathrm{II}}$ corresponds to a Fokker-Planck equation with a nonsingular diffusion coefficient sub-matrix. Equivalently, Eq. (3.39) may be interpreted as the two step iterative sequence (see also Fig. 3.1):

$$\hat{p} = e^{[\Delta t \mathcal{L}_{\mathrm{I}}]}\, p \tag{3.39a}$$

$$p' = e^{[\Delta t \mathcal{L}_{\mathrm{II}}]}\, \hat{p} \tag{3.39b}$$

where $p' = p(\mathbf{q}', \mathbf{v}', t + \Delta t)$, $p = p(\mathbf{q}, \mathbf{v}, t)$, and $\hat{p}$ is just an intermediate variable to indicate that the Fokker-Planck equation operates on the solution of the Liouville equation. The numerical implementation involves (Step 1) solution of the Liouville equation

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial u_i}\Phi_i^{(\varepsilon)}(\mathbf{q}, \mathbf{v}, t)\right] p = 0 \tag{3.39c}$$

for $\hat{p}$ at $t + \Delta t$ with initial condition $p(\mathbf{q}, \mathbf{v}, t)$, followed by (Step 2) solution of the Fokker-Planck equation

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial v_i}\Phi_i^{(r)}(\mathbf{q}, \mathbf{v}, t) - \frac{\partial^2}{\partial v_i \partial v_j}\Gamma_{ij}^{(r)}(\mathbf{q}, \mathbf{v}, t)\right] p = 0 \tag{3.39d}$$

for $p(\mathbf{q}', \mathbf{v}', t + \Delta t)$ with initial condition $\hat{p}$ at t.

Step 1: Evolve PDF via solution of a Liouville equation

The contribution of the singular dimensions to the evolution of the state distribution is evaluated via Eq. (3.39c). A solution based on the method of characteristics for this Liouville equation is given by

$$p(\mathbf{q}', \mathbf{v}, t') = p(\mathbf{q}, \mathbf{v}, t) \, \exp\left\{ - \int_{t}^{t'} \mathrm{d}\tau \, \left. \frac{\partial \Phi_i^{(\varepsilon)}}{\partial u_i} \right|_{\mathbf{q},\mathbf{v},\tau} \right\} \tag{3.40}$$

where $\mathbf{q}'$ and $\mathbf{q}$ are points along the characteristic curve generated by $\dot{\mathbf{u}} = \mathbf{\Phi}^{(\varepsilon)}(\mathbf{q}, \mathbf{v}, \tau)$. Here, $\mathbf{v}$ is a constant. Our solution is implemented as a semi-Lagrangian scheme:

$$\hat{p} = p(\mathbf{q}, \mathbf{v}, t) \, \exp\left\{ -\Delta t \, \left. \frac{\partial \Phi_i^{(\varepsilon)}}{\partial u_i} \right|_{\mathbf{q},\mathbf{v},t} \right\}. \tag{3.41}$$

where $\hat{p}$ is computed at $\begin{bmatrix} \mathbf{q}' & \mathbf{v} \end{bmatrix}^{\mathrm{T}}$ with $\mathbf{q}' = \mathbf{q} + \Delta t \, \mathbf{\Phi}^{(\varepsilon)}(\mathbf{q}, \mathbf{v}, t)$, i.e., a forward Euler scheme. In grid-based approaches, this could represent off-grid points and might require interpolation of the distributions evaluated at those points back on to the points of our computational grid. Other integration schemes can also be used, for instance, fourth order Runge–Kutta method, backward Euler method, or Gauss-Legendre methods.

Step 2: Update evolved PDF from Step 1 through solution of a Fokker-Planck equation

The solution of the Liouville equation, represented by $\hat{p}$, is then updated with the contribution from the nonsingular dimensions using Eq. (3.39d). We employ a transformed path integral based solution for this Fokker-Planck equation. Namely, we evolve the distribution via the Chapman-Kolmogorov equation

$$p(\mathbf{q}', \mathbf{v}', t + \Delta t) = \int \rho(\mathbf{v}', t + \Delta t \,|\, \mathbf{v}, t) \, \hat{p} \, \mathrm{d}\mathbf{v} \tag{3.42}$$

using the short-time propagator

$$\rho(\mathbf{v}', t+\Delta t \,|\, \mathbf{v}, t) = (4\,\pi\,\Delta t)^{-N_r/2} \left\| \mathbf{\Gamma}^{(\mathbf{r})}(\mathbf{q}, \mathbf{v}, t) \right\|^{-1/2} \exp\left\{ -\frac{1}{4\,\Delta t} \left[\mathbf{v_e}^{\mathrm{T}}\, \mathbf{\Gamma}^{(\mathbf{r})}(\mathbf{q}, \mathbf{v}, t)^{-1}\, \mathbf{v_e} \right] \right\} \tag{3.43}$$

where $\mathbf{v_e} = \mathbf{v}' - \mathbf{v} - \Delta t\, \mathbf{\Phi}^{(r)}(\mathbf{q}, \mathbf{v}, t)$. Here, complementary to Step 1, $\mathbf{q}$ is a constant.

Computation of $\mathbf{\Phi}(\mathbf{z}, t)$ via Eq. (3.18) requires us to evaluate $\dot{\boldsymbol{\mu}}(t)$ and $\dot{\mathbf{R}}(t)$. While $\dot{\boldsymbol{\mu}}(t)$ may be easily evaluated from Eq. (3.25), $\dot{\mathbf{R}}(t)$ from Eq. (3.26) is not straightforward in general. Moreover, updated values of $\boldsymbol{\mu}(t)$ and $\mathbf{R}(t)$ are needed to compute the distribution of the original state variables from that of the transformed state variables via Eq. (3.14). In our implementation we address these challenges by considering the first order approximation $\Delta\mathbf{R} \equiv \mathbf{R}(t+\Delta t) - \mathbf{R}(t)$ for $\dot{\mathbf{R}}(t)\,\Delta t$ and forward Euler schemes for Eqs. (3.25) and (3.26). Thus, we have the update relations:

$$\boldsymbol{\mu}(t+\Delta t) = \boldsymbol{\mu}(t) + \Delta t \left\langle \tilde{\mathbf{f}}(\mathbf{z}, t) \right\rangle, \tag{3.44}$$

$$\mathbf{\Sigma}(t+\Delta t) = \mathbf{\Sigma}(t) + \Delta t \left[\mathbf{R}(t) \left\langle \mathbf{z}\, \tilde{\mathbf{f}}(\mathbf{z}, t)^{\mathrm{T}} \right\rangle + \left\langle \tilde{\mathbf{f}}(\mathbf{z}, t)\, \mathbf{z}^{\mathrm{T}} \right\rangle \mathbf{R}(t)^{\mathrm{T}} + 2 \left\langle \tilde{\mathbf{G}}(\mathbf{z}, t) \right\rangle \right]. \tag{3.45}$$

We obtain $\mathbf{R}(t+\Delta t)$ by performing a Cholesky factorization of $\mathbf{\Sigma}(t+\Delta t)$. The distribution of the transformed state variables are updated using Eqs. (3.41) to (3.43) with $\mathbf{\Phi}(\mathbf{z}, t)\Delta t = \mathbf{R}(t)^{-1} \left[\delta\tilde{\mathbf{f}}(\mathbf{z}, t)\Delta t - \Delta\mathbf{R}\, \mathbf{z} \right]$ and $\mathbf{\Gamma}(\mathbf{z}, t) = \mathbf{R}(t)^{-1} \tilde{\mathbf{G}}(\mathbf{z}, t)\, \mathbf{R}(t)^{-1^{\mathrm{T}}}$.

3.3.4 Numerical implementation

Lattice-based implementations of path integrals often involve discrete representations of distributions on a grid—for instance, $p_i = p(\mathbf{z_i}, t)$ for $i = 1 \ldots, N$. The index i here represents the discretization of the computational domain into, N grid points and should not be confused with the vector indices. Hence, p_i here is the value of $p(\mathbf{z}, t)$ evaluated at the i-th grid point $\mathbf{z_i}$. Under this representation, the update equations for PDF propagation,

namely Eqs. (3.41) and (3.42) can be expressed in the form

$$\hat{p}_i = p_i \, J_i \tag{3.46}$$

$$p'_i = \sum_{j=1}^{N} \rho_{ij} \, \hat{p}_j \tag{3.47}$$

where $J(\cdot) = \exp\left\{-\Delta t \; \partial \Phi_k^{(\varepsilon)}(\cdot)/\partial u_k\right\}$ and ρ_{ij} represents the short-time propagator in the transformed space whose individual terms are obtained from Eq. (3.43). Similarly, in the discrete versions of the update equations for the mean and covariance, the expectations are represented as weighted summations. Eqs. (3.41) to (3.45) and their discrete versions form the basis of our proposed approach. The long-time evolution of the PDF, based on these update equations, accounts for the contributions of all possible paths in the transformed space. In fact, for the case of nonsingular systems, i.e., systems where $N_r = N_s$, our approach becomes equivalent to the TPI approach described in [93].

The formulations are also applicable for the case when $N_r = 0$, i.e., a system without stochastic forcing. These systems may be encountered, for instance, in the area of uncertainty quantification [96]. The evolution of the probability density function in such systems are governed by the Liouville equation [97], which in the transformed space is given by (see Chapter B for details of the derivation)

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial z_i}\Phi_i(\mathbf{z}, t)\right] p(\mathbf{z}, t) = 0. \tag{3.48}$$

Note that Eq. (3.48) is of the form of Eq. (3.39c) from Step 1 of our proposed approach. Thus our approach when applied to this case presents a numerical solution for the Liouville equation. Also of note are certain dynamical systems driven by non–white noise stochastic processes where the excitations can be modeled as a filtered white noise process [59–63].

These systems may be described by the set of equations

$$\mathrm{d}\boldsymbol{\theta}(t) = \mathbf{h}\Big(\boldsymbol{\theta}(t), t\Big)\,\mathrm{d}t + \boldsymbol{\xi}(t)\,\mathrm{d}t \tag{3.49}$$

$$\mathrm{d}\boldsymbol{\xi}(t) = \mathbf{m}\Big(\boldsymbol{\xi}(t), t\Big)\,\mathrm{d}t + \mathbf{B}\Big(\boldsymbol{\xi}(t), t\Big)\,\mathrm{d}\mathbf{w}(t) \tag{3.50}$$

where the auxiliary variable $\boldsymbol{\xi}$ represents the filtered white noise process. We can express the system as an Itô SDE, i.e., of the form of Eq. (3.1), by considering the augmented state $\mathbf{x} = \begin{bmatrix}\boldsymbol{\theta} & \boldsymbol{\xi}\end{bmatrix}^{\mathrm{T}}$. In this representation, the system dynamics takes the form $\mathbf{f} = \begin{bmatrix}\mathbf{h} + \boldsymbol{\xi} & \mathbf{m}\end{bmatrix}^{\mathrm{T}}$ and the strength of white noise excitation is given by

$$\mathbf{A} = \begin{bmatrix}\mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}\end{bmatrix} \tag{3.51}$$

As a consequence, we obtain a singular diffusion coefficient matrix $\mathbf{G} = \mathbf{A}\mathbf{A}^{\mathrm{T}}/2$. Similarly, second order dynamical systems given by

$$\ddot{\boldsymbol{\theta}} + \mathbf{h}\Big(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, t\Big) = \mathbf{B}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, t)\,\dot{\mathbf{w}}(t) \tag{3.52}$$

where $\dot{\mathbf{w}}(t)$ represents the formal derivative of the Wiener process may be expressed as an Itô SDE of the form of Eq. (3.1) with the augmented state $\mathbf{x} = \begin{bmatrix}\boldsymbol{\theta} & \dot{\boldsymbol{\theta}}\end{bmatrix}^{\mathrm{T}}$ and the system dynamics $\mathbf{f} = \begin{bmatrix}\dot{\boldsymbol{\theta}} & -\mathbf{h}\end{bmatrix}^{\mathrm{T}}$. The strength of white noise excitation once again takes the form of Eq. (3.51).

Essentially, in our formulation we have expressed an n-dimensional second order system as a first order system in $2n$ dimensions. In general, an m-th order system in n dimensions can be expressed as an $m \times n$ dimensional first order system. We can also have a combination of these systems such as a second order system with zero process noise or non-white noise expressed as the output of a second order filter. Nevertheless, they may be expressed in the form of Eq. (3.1) with a singular diffusion coefficient matrix and thereby within the

framework of our proposed approach. Thus our proposed approach can be applied to a diverse set of dynamical systems: nonsingular dynamical systems (reduces to TPI approach), second order (and higher order) dynamical systems, those with zero process noise, and certain non-white noise dynamical systems. We thus refer to our proposed approach as the generalized transformed path integral approach.

In order to maintain high levels of accuracy in numerical implementations of path integral approaches [47, 49], the zeroth moment renormalization condition

$$p_i^{(new)} = \frac{p_i^{(old)}}{\sum_i p_i^{(old)} \Delta z_i} \tag{3.53}$$

is often imposed for both the transition probability density and state distribution at every time step. The generalized transformed path integral approach allows us to easily enforce conditions to also conserve the first and second moments of the state distribution. At each time step we apply the transformation

$$p_i^{(new)}(\mathbf{z_i}) = p_i^{(old)}(\mathbf{S_z z_i} + \mathbf{M_z}) \tag{3.54}$$

where $\mathbf{M_z} = \sum_i \mathbf{z_i}\, p_i^{(old)}\, \Delta z_i$ and $\mathbf{S_z} = \sum_i \mathbf{z_i z_i}^T\, p_i^{(old)}\, \Delta z_i$. To our knowledge, this conservation of the first and second moment properties in path integral based approaches is unique to our work. The enforcement of Eqs. (3.53) and (3.54) contributes greatly to the numerical stability and accuracy of the generalized transformed path integral approach.

3.4 Performance of the Generalized Transformed Path Integral Approach

In this section, we showcase the benefits of the generalized transformed path integral (GTPI) approach. Conventional fixed grid approaches (FG), where solutions of stochastic systems are obtained on a grid with fixed resolution and domain bounds in the original state space,

are unable to efficiently address the challenges arising from large drift, diffusion, and concentration of PDF. These were higlighted in our earlier work [93]. There, we proposed the transformed path integral (TPI) approach for systems with nonsingular diffusion coefficient matrices. We illustrated its advantages over conventional approaches in addressing the aforementioned challenges. Here, we showcase performance of the GTPI approach and demonstrate its applicability to a more general class of stochastic systems. Using representative examples, we demonstrate the benefits over a fixed grid approach, specifically, over a path integral based approach using operator splitting to address singular diffusion but performed in the original space.

We consider three broad categories: second order systems, systems with zero process noise, and those with non-white noise excitation. In the category of second order systems we present solutions of (a) the stochastic harmonic oscillator, (b) the stochastic van der Pol oscillator, and (c) the stochastic Caughey oscillator. We then study the solution of (d) the stochastic Duffing oscillator with zero process noise. And finally, we present the solution of (e) bistable stochastic flow driven by non–white noise. These systems and cases were chosen to highlight specific benefits and features of the GTPI approach.

3.4.1 Stochastic harmonic oscillator

The stochastic harmonic oscillator is a linear second order dynamical system with a known analytical solution. We consider an oscillator with $h = \gamma\dot{q} + \omega^2 q$. The linearity of the system allows us to isolate and study, through three different cases, effects of large drift, diffusion, and concentration of PDF. The system has only one stationary point at the origin. Therefore, in all our cases, the state distribution will evolve to a Gaussian stationary distribution centered at the origin as $t \to \infty$. In addition, if the initial distribution is Gaussian, then the state distribution will remain Gaussian. Consequently, the distribution in the transformed space will remain standard multivariate normal in such a scenario.

In Case A, we consider a system with $\gamma = 0.5$, $\omega = 1$, and $a = 0.5$. The state variable is

initially normally distributed about $\begin{bmatrix} 4 & 4 \end{bmatrix}^{\mathrm{T}}$ with identity covariance. The evolution of the state distribution from 0 s to 6 s is shown in Fig. 3.2. The simulations were performed with a timestep of 0.02 s. The system considered is an example of a large drift system where there is large translation of the state distribution. Conventional fixed grid approaches may face challenges in accurately describing the transient behavior of the system.

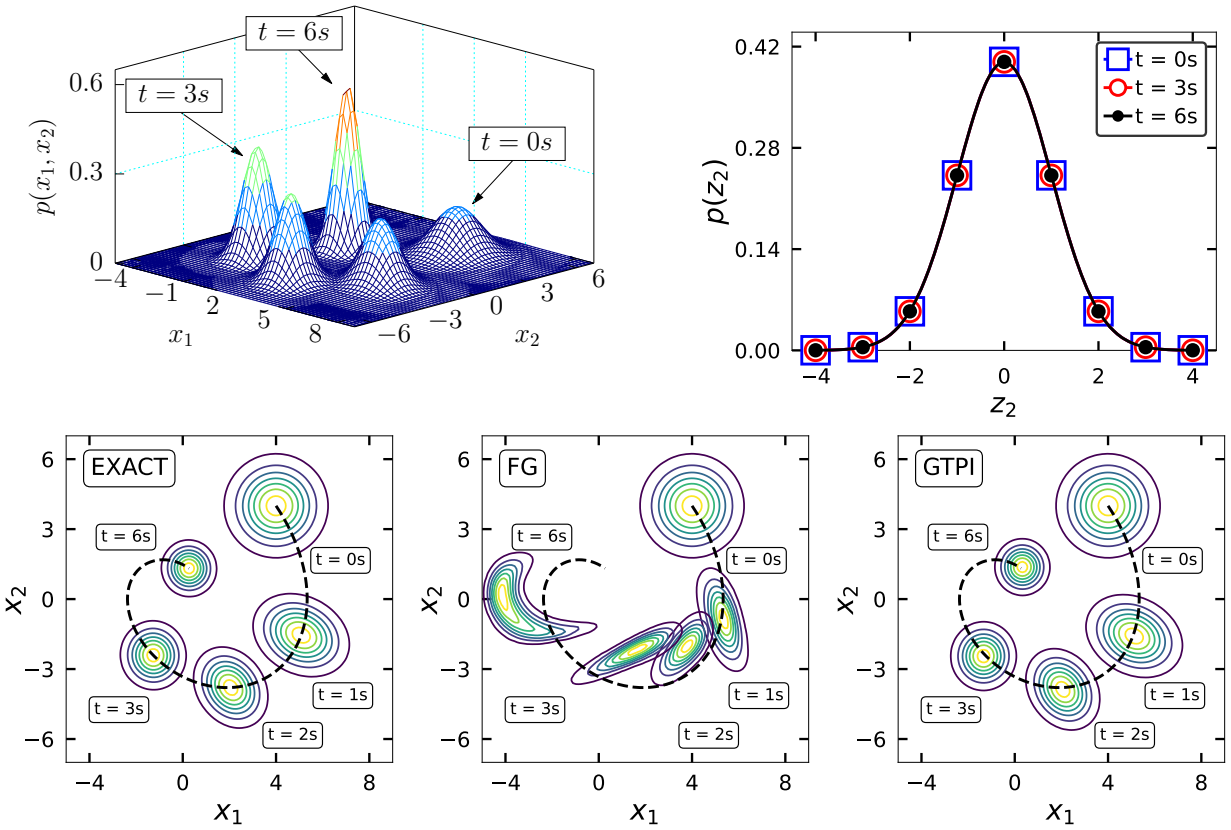

Figure 3.2: Plots of the state distribution at different times for a stochastic harmonic oscillator exhibiting large drift (Case A). Surface plot of the analytical solution is shown on the top left. The transformed space marginal plots from the GTPI approach are on the top right. Contour plots of the joint distribution obtained via Monte Carlo simulation, the FG approach, and the GTPI approach are shown on the bottom. The dotted lines show the deterministic path.

With a fixed grid in the original space, grid bounds that are sufficient to accurately represent the initial distribution might be insufficient at later times. In the case considered, a grid with domain bounds $\begin{bmatrix} -1 & 9 \end{bmatrix}$, i.e., five standard deviations about the initial mean,

cannot be used to represent the distribution at $t = 3\,\mathrm{s}$. However, larger domain bounds with the same number of points might introduce errors due to losses in resolution. This can be seen in the bottom center plot of Fig. 3.2 which shows a fixed grid simulation performed on a uniform grid with bounds $\begin{bmatrix} -8 & 12 \end{bmatrix}$ in either dimension. In contrast, the simulation using our proposed GTPI approach shown in the bottom right plot of Fig. 3.2 is more accurate than the fixed grid approach. This can also be seen from Table. 3.1 showing the comparison of error estimates in the numerical solutions.

We define the RMS error measure in the PDF at time t as

$$\varepsilon_{\mathrm{RMS}}(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\Big[p(x_i,t) - \hat{p}(x_i,t)\Big]^2} \tag{3.55}$$

where $\hat{p}(x_i,t)$ represents the numerical solution obtained from the FG or GTPI approach evaluated at the i-th grid point and at time t. The analytical solution $p(x_i,t)$ for the stochastic harmonic oscillator is known at all times. In cases where the analytical solution or the stationary solution is not known, the solution obtained from MC simulations may be used instead for $p(x_i,t)$. The error is evaluated by averaging the square of the difference in these PDF values over all points in the computational grid and then taking its square root. The solution obtained from the GTPI approach has an error that is an order of magnitude lower than that from the FG approach. Although the computation time is higher for the same number of grid points—GTPI approach has more operations—the computation time and error of GTPI on a (81×81) grid are lower than that of FG on a (121×121) grid. In other words, our proposed approach is more computationally efficient than the conventional fixed grid based approaches for comparable error.

In the generalized transformed path integral approach the propagation of the state distribution is performed in the transformed space. Note that a fixed grid in the transformed space corresponds to an adaptive grid in the original space translated and scaled according to the mean and covariance of the state distribution. The bounds in the transformed space,

		Stochastic Harmonic Oscillator					
		Case A (at $t = 6\,\mathrm{s}$)		Case B (at $t = 2\,\mathrm{s}$)		Case C (at $t = 2\,\mathrm{s}$)	
	N_{pts}	$T_{\mathrm{comp}}(\mathrm{s})$	$\varepsilon_{\mathrm{RMS}}$	$T_{\mathrm{comp}}(\mathrm{s})$	$\varepsilon_{\mathrm{RMS}}$	$T_{\mathrm{comp}}(\mathrm{s})$	$\varepsilon_{\mathrm{RMS}}$
FG	81×81	106.2	6.42×10^{-2}	30.0	8.27×10^{-4}	31.5	1.28×10^{-1}
	121×121	218.1	3.93×10^{-2}	67.0	8.31×10^{-4}	76.9	8.34×10^{-2}
GTPI	81×81	156.2	5.39×10^{-3}	48.5	1.40×10^{-5}	49.6	1.25×10^{-3}

Table 3.1: Comparison of RMS error in the PDF obtained by fixed grid (FG) and generalized transformed path integral (TPI) approaches at $t = 6\,\mathrm{s}$ for Case A (large drift) of the stochastic harmonic oscillator, and at $t = 2\,\mathrm{s}$ for Case B (large diffusion) and Case C (large concentration of PDF).

$\begin{bmatrix} -5 & 5 \end{bmatrix}$ in either dimension for this case, corresponds to an adaptive grid with grid bounds approximately five standard deviations from the mean of the state distribution. As a consequence, our grid points are located in the regions of "importance" contributing to better performance of our proposed approach for this large drift second order system. Further evidence for the accuracy of our proposed approach can be seen in Fig. 3.2 (top right) showing plots of the marginal distributions in the transformed space; they remain standard normal in this case.

For Case B, we increase the strength of the white noise excitation to $a = 4$ while using the same system dynamics parameters as in Case A, i.e., $\gamma = 0.5$ and $\omega = 1$. In addition, we consider a zero-mean Gaussian initial distribution with identity covariance. Because the stationary distribution is also zero-mean, the challenges from large drift seen in Case A are minimized here. However, in this large diffusion system, the initial distribution will dissipate to one with with larger covariance, smaller peak value, and more tail information. Conventional fixed grid approaches may face challenges in accurately describing the long time behavior of large diffusion systems. Plots of the marginal distributions in the original space shown in Fig. 3.3 (left and middle) demonstrate this shortcoming. They also show the better performance of our proposed GTPI approach. Consequently, the greater accuracy and higher computational efficiency of the GTPI approach is also seen in Table. 3.1.

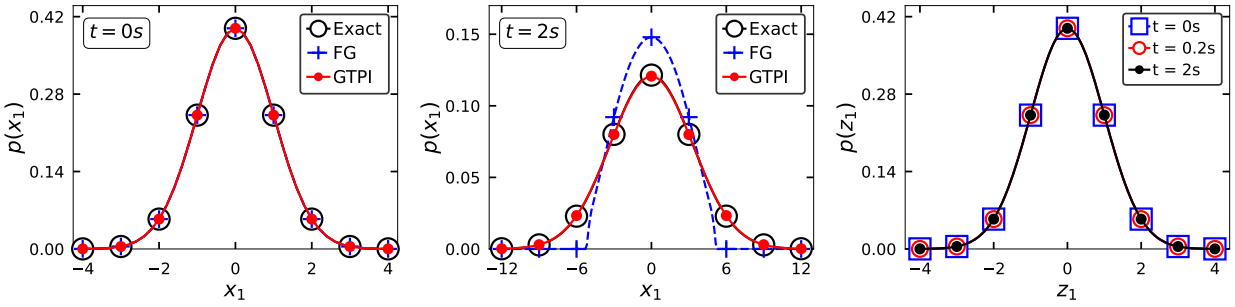

Figure 3.3: Comparison of marginal PDFs in the original space at $t = 0\,\mathrm{s}$ (left) and $t = 2\,\mathrm{s}$ (center) as well as in the transformed space (right) for a stochastic harmonic oscillator with large diffusion (Case B).

The domain bounds in the original space, $\begin{bmatrix} -5 & 5 \end{bmatrix}$ in either dimensions, seem sufficient to accurately represent the initial distribution. However, these fixed bounds lead to significant loss of tail information as the distribution dissipates. This loss of information can cause large errors as seen in Fig. 3.3 (middle). In contrast, the GTPI approach, with $\begin{bmatrix} -5 & 5 \end{bmatrix}$ bounds in the transformed space, does not exhibit any significant loss of tail information. It is able to accurately describe the long time behavior in this large diffusion second order system. Additionally, the transformed space marginal distributions in Fig. 3.3 (right) remain standard normal showcasing the accuracy of our proposed approach. Simulations of the evolution were carried out in increments of $\Delta t = 0.02\,\mathrm{s}$ between $0\,\mathrm{s}$ and $2\,\mathrm{s}$; they were performed on uniform grids with (81×81) grid resolution in both the original and transformed space.

In Case C, we consider a system with $\gamma = 2.5$, $\omega = 1$, and $a = 0.2$. In contrast to Case B, here an initial multivariate standard normal distribution will evolve to a Gaussian distribution with smaller covariance and larger peak value at $t = 2\,\mathrm{s}$. In other words, the distribution becomes more concentrated at the origin over time. Regions of large concentration in the PDF might pose challenges to conventional fixed grid approaches—namely, the distributions in these regions may become under-resolved with PDF evolution. For our simulations, we considered grid resolutions and domain bounds identical to those in Case B, i.e., uniform grids with (81×81) grid resolution and $\begin{bmatrix} -5 & 5 \end{bmatrix}$ domain bounds in either dimensions. Once again, the simulations were carried out in increments of $\Delta t = 0.02\,\mathrm{s}$ between $0\,\mathrm{s}$ and $2\,\mathrm{s}$.

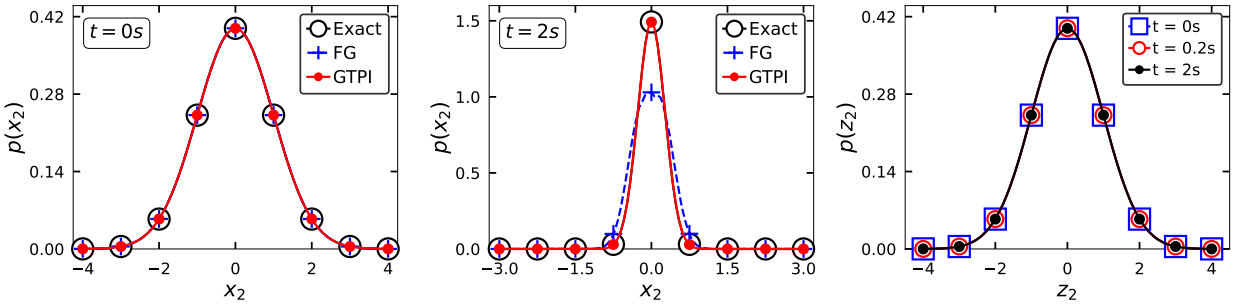

Figure 3.4: Comparison of marginal PDFs in the original state space at $t = 0\,\mathrm{s}$ (left) and $t = 2\,\mathrm{s}$ (center) as well as in the transformed space (right) for a stochastic harmonic oscillator with large concentration in PDF (Case C).

The grid resolution in the original space is sufficient to accurately represent the initial distribution. But, at later times the distribution becomes under-resolved as seen in Fig. 3.4 (middle). The GTPI approach does not face such issues of resolution here. It is able to accurately describe the long time behavior in this second order system with large concentration of PDF. The distribution in the transformed space remains standard normal for this case as well as seen in Fig. 3.4 (right). Table. 3.1 illustrates the greater accuracy and computational efficiency of the GTPI approach.

In summary, the three cases considered here highlight some of the challenges faced by fixed grid based approaches in obtaining accurate numerical solutions to stochastic dynamical systems. Namely, we want large domain bounds to minimize errors due to large drift and large diffusion; we want higher grid resolution to minimize errors due to large concentration of PDF; and we would like to keep computation times small. In fixed grid based approaches, when we have large domain bounds and high grid resolution we have larger computation times. When we compromise on the domain bounds or the grid resolution, we have lower accuracy. In our proposed generalized path integral approach, we can have higher accuracy with lower computation time when compared with the fixed grid based approaches.

The accuracy of the GTPI approach can be further improved by controlling the simulation parameters. The effects of varying the grid resolution, i.e., N_z of a $(N_z \times N_z)$ grid, and the time step Δt on the accuracy of our approach can be seen in Fig. 3.5. For our analysis,

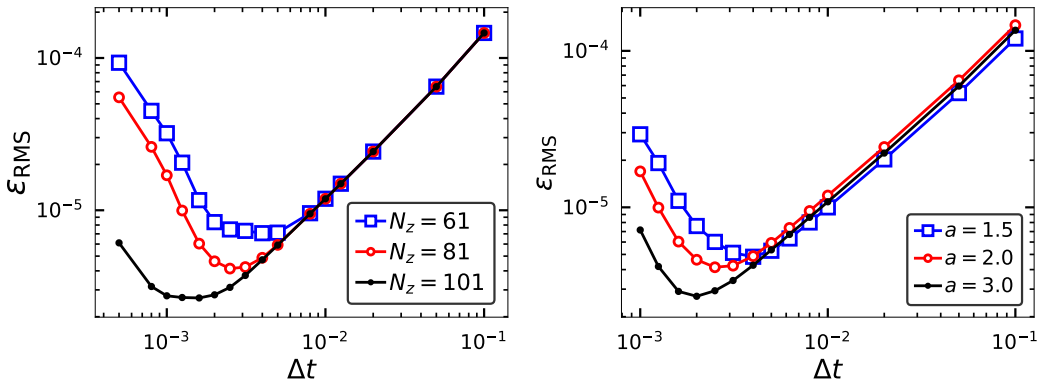

Figure 3.5: The influence of varying simulation parameters on the numerical error in the GTPI approach for the stochastic harmonic oscillator with $a = 2$, $\gamma = 0.5$, and $\omega^2 = 1$. The RMS error in the PDF ($\varepsilon_{\mathrm{RMS}}$) is evaluated using Eq. (3.55) and it is evaluated at $t = 2\,\mathrm{s}$. The variation of $\varepsilon_{\mathrm{RMS}}$ with the time step Δt is shown on the left and its variation with the grid resolution, specifically, N_z of a $(N_z \times N_z)$ uniform grid is shown on the right.

we used the stochastic harmonic oscillator with strength of white noise excitation $a = 2$, damping coefficient $\gamma = 0.5$ and stiffness coefficient $\omega^2 = 1$. Three immediate insights can be gleaned from these figures. Firstly, for time steps larger than $\Delta t = 1 \times 10^{-2}\,\mathrm{s}$, increasing the number of grid points does not result in a reduction of error. However, reducing the time step does reduce the error; halving the time step cuts the error approximately by half. In other words, the numerical error from the simulations is approximately of $\mathcal{O}(\Delta t)$ matching the expectations from our theoretical formulations—the GTPI approach uses a first order splitting solution Eq. (3.39) and the short-time propagator Eq. (3.43) is a first order accurate solution of the corresponding Fokker-Planck equation.

Secondly, for a given grid resolution there is an optimal time step $\Delta t^\star$ below which further reduction of the time step produces larger error. And the error grows more rapidly. Halving the time step almost quadruples the error. Lastly, for time steps equal to or smaller than the optimal value, increasing the number of grid points reduces the error. Note that the optimal value as well as the error at the optimal value $\varepsilon^\star_{\mathrm{RMS}} \equiv \varepsilon_{\mathrm{RMS}}(\Delta t^\star)$ become smaller for finer grid resolutions, i.e, there is more room to vary Δt so as to reduce the error. However, having a finer grid resolution for small time steps will increase the computational cost of the approach. In summary, for a given grid resolution one may reduce error in the GTPI

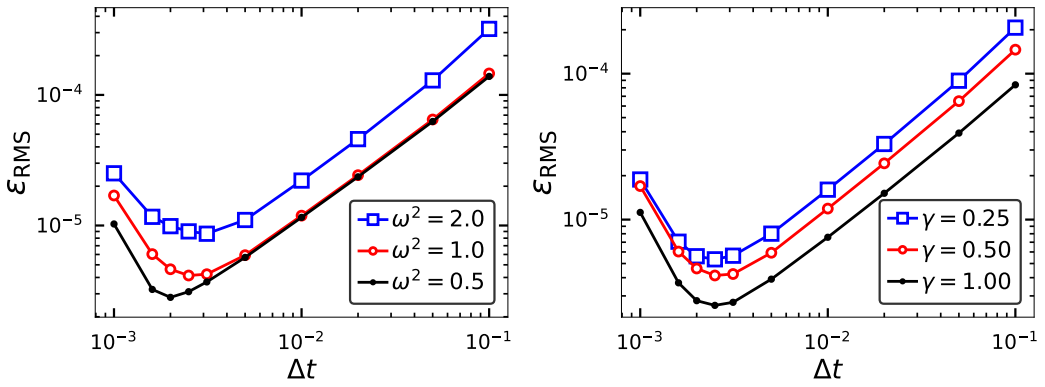

Figure 3.6: The influence of system parameters on the numerical error in the GTPI approach for the stochastic harmonic oscillator. The RMS error in the PDF ($\varepsilon_{\mathrm{RMS}}$) is evaluated using Eq. (3.55) and it is evaluated at $t = 2\,\mathrm{s}$. The influence of white noise strength a on $\varepsilon_{\mathrm{RMS}}$ is shown on the left. The influence of the stiffness coefficient ω^2 is shown in the middle and that of the damping coefficient γ is shown on the right.

approach by reducing the time step up to an optimal value $\Delta t^\star$. Further reduction in the error would require a finer grid resolution.

The influence of the system parameters, namely, the strength of white noise excitation a, the stiffness coefficient ω^2, and the damping coefficient γ on the optimal time step $\Delta t^\star$ can be seen in Fig. 3.6. A system with more process noise has a lower $\Delta t^\star$ as well as a lower $\varepsilon^\star_{\mathrm{RMS}}$ as seen in Fig. 3.6 (left). Likewise, so does a system with a lower stiffness coefficient seen in Fig. 3.6 (middle). Also, the RMS error at any time step seems to be lower for a system with a lower stiffness coefficient. In a system with more damping, both the RMS error and its minimum value $\varepsilon^\star_{\mathrm{RMS}}$ are lower as seen in Fig. 3.6 (right). Interestingly, the optimal time step $\Delta t^\star$ seems to remain constant. Thus, the benefits of our proposed generalized transformed path integral approach may be more pronounced in cases where the damping in the system is large.

3.4.2 Stochastic van der Pol oscillator

The van der Pol oscillator, a non-conservative oscillator with nonlinear damping, is a second order dynamical system. This nonlinear dynamical system is widely used as a standard model for studying oscillatory processes in diverse fields [98–100]. Challenges due to large

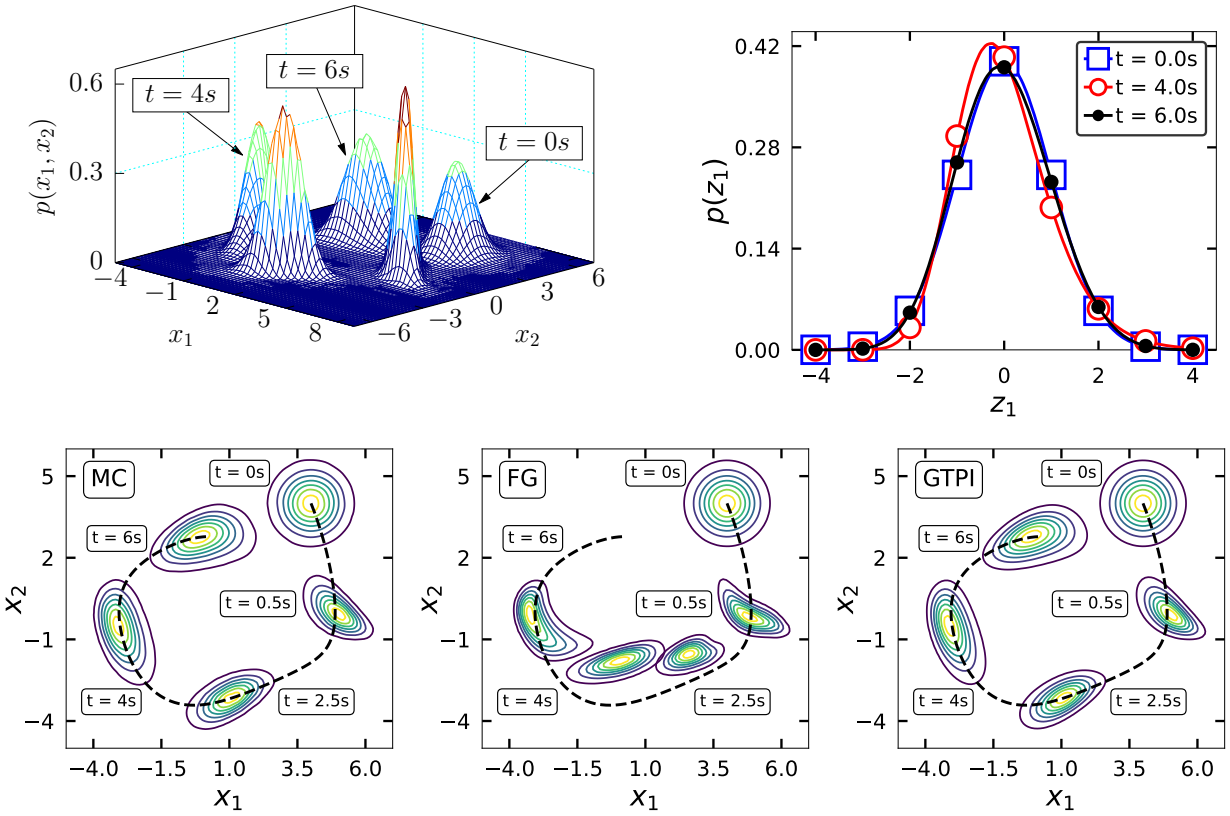

Figure 3.7: Plots of the state distribution at different times for the van der Pol oscillator. Surface plot of Monte Carlo (MC) solution is shown on the top left. Transformed space marginal plots from the GTPI approach are on the top right. Contour plots of the joint distribution obtained via MC simulation, FG approach, and GTPI approach are shown on the bottom. The dotted lines show the deterministic path.

drift, diffusion, and concentration of PDF, that were highlighted in the linear cases, may also be encountered in nonlinear dynamical systems with the nonlinearity adding further complexity for prediction of state evolution. Our proposed GTPI approach can be used to predict the evolution while addressing the challenges faced by conventional approaches.

Here, we consider a van der Pol oscillator with $h = 0.1\,\dot{q}\,(1 - q^2) - q$ and system noise $a = 0.5$. The initial distribution is Gaussian with mean $\begin{bmatrix}4 & 4\end{bmatrix}$ and a diagonal covariance matrix. The diagonal elements are $\begin{bmatrix}0.5 & 0.5\end{bmatrix}$. Simulations of the system evolution between 0 s and 6 s were carried out via MC, FG, and GTPI approaches in increments of $\Delta t = 0.02$ s. These are shown in Fig. 3.7. Clearly, large drift, diffusion, and concentration of PDF are all

		Nonlinear Second Order Stochastic Dynamical Systems			
		Stochastic van der Pol Oscillator		Stochastic Caughey Oscillator	
	N_{pts}	T_{comp}(s)	ε_{RMS}	T_{comp}(s)	ε_{RMS}
FG	81×81	109.4	6.63×10^{-2}	476.0	4.37×10^{-2}
	121×121	226.8	2.17×10^{-2}	865.3	4.38×10^{-2}
TPI	81×81	151.3	3.28×10^{-3}	679.7	6.81×10^{-4}

Table 3.2: Comparison of RMS error in the PDF obtained by fixed grid (FG) and generalized transformed path integral (TPI) approaches at $t = 6\,\text{s}$ for stochastic dynamical systems with nonlinear (with respect to the state variables) drift functions—the stochastic van der Pol oscillator and the stochastic Caughey oscillator.

present in this system.

The contour plots in Fig. 3.7 (bottom) show good agreement between the solutions obtained from Monte-Carlo simulations and the generalized transformed path integral approach. However, solutions from the fixed grid approach show large errors. The greater accuracy of our approach can also be seen in Table. 3.2 showing comparisons of the RMS error in the probability density functions. Also evident is the greater computational efficiency of our approach. Since an exact analytical solution for this case is not known, the error here is evaluated by considering the PDF obtained from Monte-Carlo simulations as a reference for the exact solution. The error in our approach on a (81×81) uniform grid is about an order of magnitude lower than that from the fixed grid approach on a (121×121) uniform grid. The computation cost of the fixed grid approach is also greater—about 1.5 times that of the generalized transformed path integral approach.

As discussed earlier, the fixed grid approach faces challenges in addressing the effects of large drift, diffusion, and concentration of PDF. Here, the fixed grid simulations were performed on a grid in the original space with domain bounds $\begin{bmatrix} -8 & 10 \end{bmatrix}$ in either dimensions. On the other hand, the generalized transformed path integral approach was performed on a grid in the transformed space with domain bounds $\begin{bmatrix} -5 & 5 \end{bmatrix}$ in either dimensions. This

corresponds to an adaptive grid in the original space that is spread about five standard deviations from the state mean. The grid points in our approach are located in the regions of "importance" thereby contributing to the better performance even in this nonlinear dynamical system. These regions lie along the deterministic path shown in dotted lines on the contour plots of Fig. 3.7 (bottom). Note that the distribution in the transformed space shown in Fig. 3.7 (top right) does not remain standard normal for this nonlinear dynamical system.

3.4.3 Stochastic Caughey oscillator

The Caughey oscillator is a nonlinear dynamical system with a known analytical solution for the stationary distribution of the state [14, 20, 49]. Challenges due to large drift, diffusion, and concentration of PDF may also be present in the prediction of long term behavior of dynamical systems. We study the stationary behavior in a system with $h = 2\,\dot{q}\,(1-q^2-\dot{q}^2)-q$ and system noise $a = 1$. In this system, a multivariate standard normal initial distribution will evolve to a non-Gaussian distribution as shown in Fig. 3.8 (top). The origin is an unstable stationary point and the unit circle is a stable limit cycle for this system.

Simulations of the system evolution between 0 s and 6 s were carried out via MC, FG, and GTPI approaches in increments of $\Delta t = 0.005\,\text{s}$. The domain bounds, $\begin{bmatrix} -1 & 1 \end{bmatrix}$ in either dimension, and grid resolution (81×81) in the original space are sufficient to accurately represent the initial distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu} = \begin{bmatrix} 0 & 0 \end{bmatrix}^{\mathrm{T}}$ and $\boldsymbol{\Sigma} = \text{diag}\left(\begin{bmatrix} 1/25 & 1/25 \end{bmatrix}^{\mathrm{T}}\right)$. However, as the distribution evolves and gets more concentrated around the unit circle, the grid is insufficient to accurately describe the distribution. The generalized transformed path integral approach on a (81×81) grid in the transformed space with domain bounds $\begin{bmatrix} -5 & 5 \end{bmatrix}$ in either dimension is able to more accurately describe the distribution.

Fig. 3.8 (bottom middle) shows good agreement between solutions obtained from Monte Carlo simulations, GTPI approach, and stationary distribution while there are large errors in the conventional fixed grid approach. Evolution of the marginal distribution in the trans-

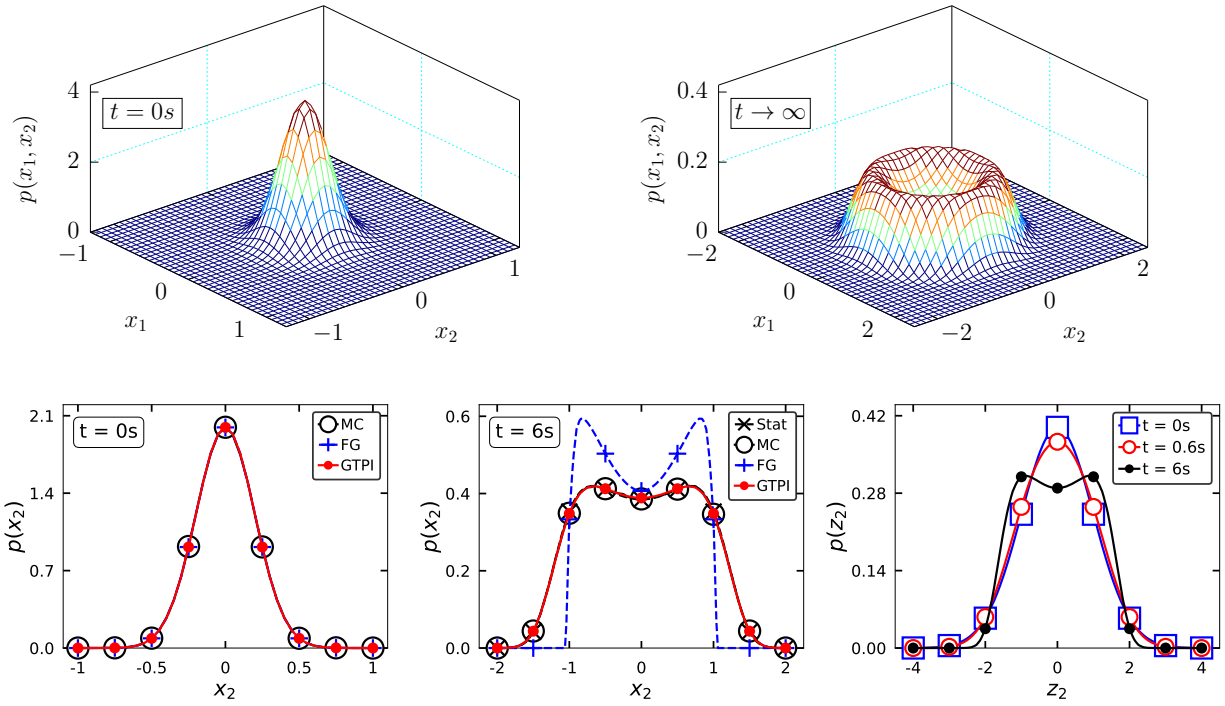

Figure 3.8: Plots of the state distribution at different times for the Caughey oscillator. Surface plots of the initial and stationary distribution are shown on the top. Marginal PDFs in the original space (left and middle) as well as in the transformed space (right) are shown on the bottom.

formed space is shown in Fig. 3.8 (bottom right). It does not remain standard normal in the transformed space. Similar insights may also be gleaned from Table. 3.2 showing comparisons of RMS error in the PDF. The error is computed using the stationary solution as a reference for the exact solution. The error from the GTPI approach is about two orders of magnitude lower than that from the conventional fixed grid approach.

3.4.4 Stochastic Duffing oscillator with zero process noise

The Duffing oscillator is a second order dynamical system with a cubic stiffness term. Originally used to study the behavior of a pendulum, the system has since become a standard model to describe several phenomena in science and engineering [101]. Depending on the choice of parameters, the system exhibits qualitatively different phenomena, including local and global bifurcations as well as chaos. Here we consider a Duffing oscillator with

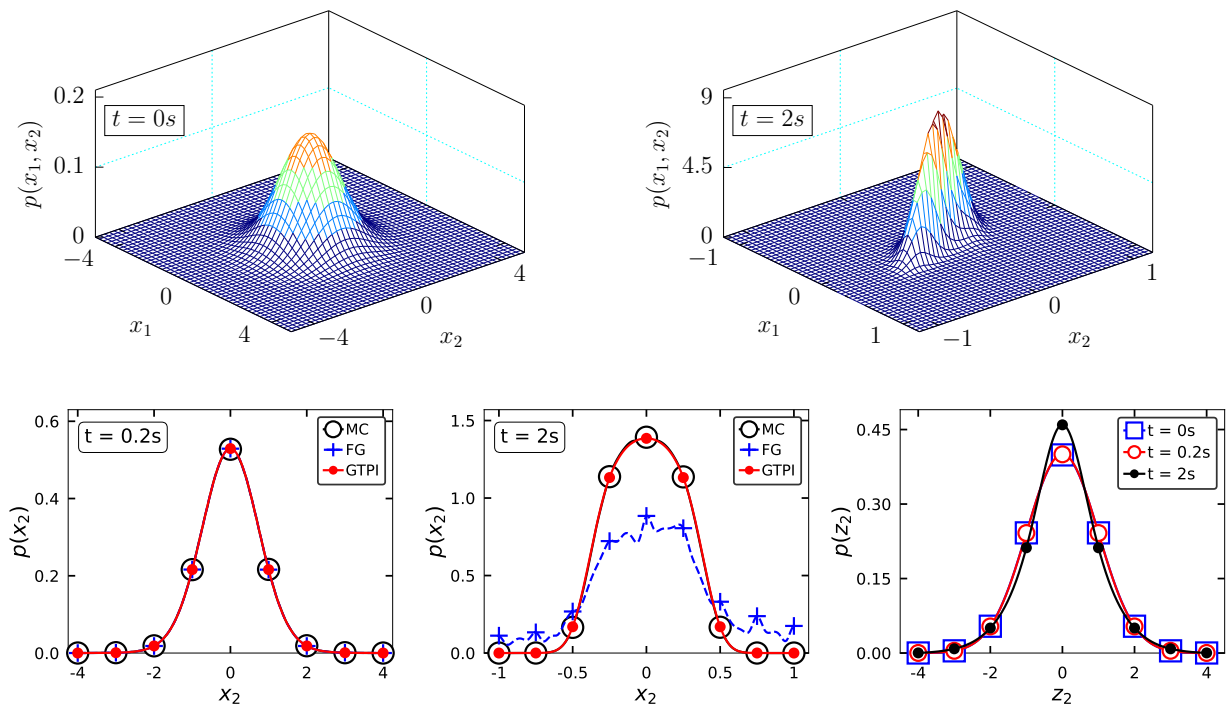

Figure 3.9: Plots of the state distribution at different times for the Duffing oscillator with zero process noise. Surface plots at $t = 0\,\mathrm{s}$ and $t = 2\,\mathrm{s}$ are shown on the top. Marginal PDFs in the original space (left and middle) as well as in the transformed space (right) are shown on the bottom.

$h = 2\,\dot{q} + 2\,q + q^3/4$ but with no process noise, i.e., $a = 0$. However, there is uncertainty in the initial conditions given by $p(q, \dot{q}) \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

The evolution of state distribution for such a system (with zero process noise) is described by the Liouville equation (LE). Note that for the case considered we have a stable stationary point at the origin. As the system evolves there will be large concentration of PDF at the origin. This can be seen in Fig. 3.9 (top) showing the evolution from a multivariate standard normal distribution at $t = 0\,\mathrm{s}$ to a concentrated non-Gaussian distribution at $t = 2\,\mathrm{s}$. Any method of solution for the LE will also have to address the challenge of large concentration of PDF in this case.

Simulations of the PDF evolution were performed in increments of $\Delta t = 0.001\,\mathrm{s}$. The fixed grid approach (FG) and the generalized transformed path integral approach (GTPI) were performed on uniform 81×81 grids with bounds $\begin{bmatrix} -5 & 5 \end{bmatrix}$ in either dimensions. While

		Zero Process Noise and Non-White Noise Excited Systems			
		Stochastic Duffing Oscillator (with zero process noise)		Bistable Stochastic Flow (with non-white noise)	
	N_{pts}	T_{comp}(s)	ε_{RMS}	T_{comp}(s)	ε_{RMS}
FG	81×81	279.7	1.99×10^{-1}	175.6	1.78×10^{-3}
	121×121	781.0	1.13×10^{-1}	325.2	1.26×10^{-3}
GTPI	81×81	558.7	8.78×10^{-2}	232.3	5.81×10^{-4}

Table 3.3: Comparison of RMS error in the PDF obtained by fixed grid (FG) and generalized transformed path integral (GTPI) approaches at $t = 2\,\text{s}$ for a stochastic Duffing oscillator with zero process noise and a bistable stochastic flow driven by colored noise.

the former is unable to accurately represent the distribution at $t = 2\,\text{s}$, the latter shows good agreement with results obtained from Monte Carlo simulations Fig. 3.9 (bottom center). The better performance of the GTPI approach in this case can also be seen from comparisons of the RMS error in Table. 3.3. Although FG grid resolution in the original space is sufficient to accurately describe the initial distribution, it becomes insufficient as the distribution becomes more concentrated at later times. In contrast, because its grid points are located at the regions of importance, GTPI is able to better manage the challenge of large concentration in PDF. Note that as the distribution becomes more non-Gaussian in the original space, so does the distribution in the transformed space Fig. 3.9 (bottom right).

3.4.5 Bistable stochastic flow driven by non-white noise

Brownian motion in a double well potential has been used to study different phenomena in a wide variety of fields [102–104]. In many situations however, one may have to consider excitations with a non-negligible correlation time. In this context, bistable systems driven by certain non-white noise processes have also been investigated [59–61]. We consider a system in the form of Eqs. (3.49) and (3.50) with $m(\theta, t) = \theta - \theta^3$, $b(\xi, t) = -\xi/\tau$, and $g(\xi, t) = 1/\tau$. Here, $\tau = 0.1$ represents the correlation time of the non-white noise.

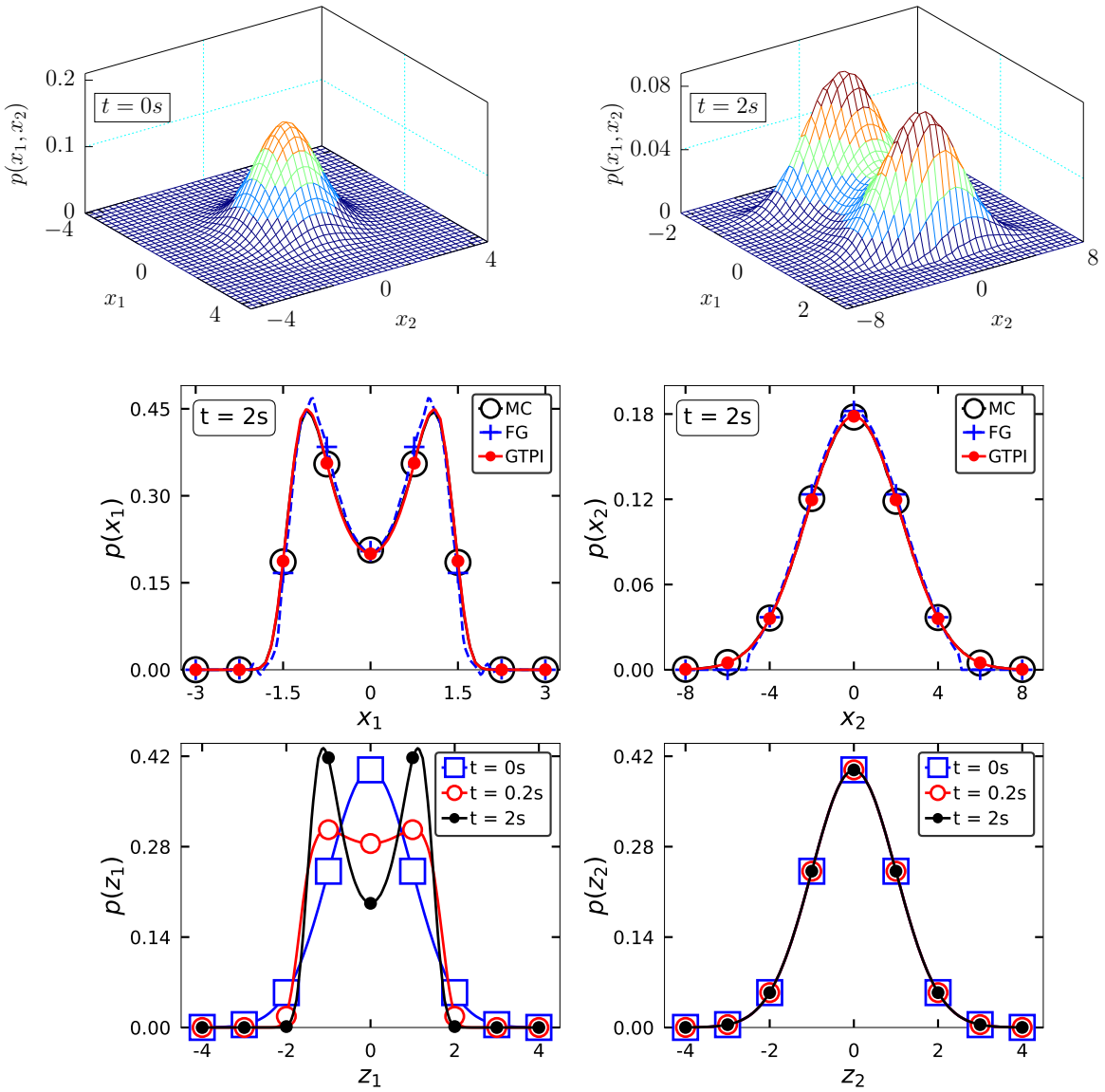

Figure 3.10: Plots of the state distribution at different times for a bistable stochastic flow driven by colored noise. Surface plots at $t = 0\,\mathrm{s}$ and $t = 2\,\mathrm{s}$ are shown on the top. Marginal PDFs in either dimensions of the original space (far left and middle left) as well as the transformed space (far right and middle right) are shown on the bottom.

The evolution of an initial multivariate standard normal distribution to a non-Gaussian distribution at $t = 2\,\mathrm{s}$ is shown in Fig. 3.10 (top). However, the dynamical system along the dimension corresponding to ξ is linear and does not depend on θ. Thus, the marginal distribution in the dimension will remain Gaussian though the joint distribution may become non-Gaussian. Consequently, the marginal distribution in the corresponding dimension of

the transformed space will remain standard normal as seen in Fig. 3.10 (bottom far right). This serves to reinforce the accuracy of our proposed generalized transformed path integral approach (GTPI) in nonlinear dynamical systems.

In this nonlinear dynamical system, there is large diffusion in one of the dimensions and large concentration of the probability density function in the other. The fixed grid approach (FG) faces issues in accurately representing this behavior while the GTPI is able to better address the challenge. This is evident from Fig. 3.10 (bottom far left and bottom center left) where the marginal plots at $t = 2\,\mathrm{s}$ in the original space show good agreement between the results obtained from Monte-Carlo simulations (MC) and our approach. In fact, as demonstrated earlier in other systems, Table. 3.3 shows that the GTPI approach on a (81×81) uniform grid in the transformed space has greater accuracy and lower computation cost than the FG approach on a (121×121) uniform grid in the original space. The simulations were performed in increments of $\Delta t = 0.005\,\mathrm{s}$ on grids with domain bounds $\begin{bmatrix} -5 & 5 \end{bmatrix}$ in either dimension.

3.5 Concluding Remarks

In this chapter, we propose the generalized transformed path integral approach (GTPI) for efficient numerical predictions of a broad class of stochastic dynamical systems (including those with singular diffusion coefficient matrices). As a part of our approach, we develop a path integral-based solution of a transformed Fokker-Planck equation (FPE). The equation describes the evolution of a complementary stochastic dynamical system whose state mean and covariance do not change with time; we refer to the system as the standard transformed stochastic dynamical system. Thus, we present an alternate approach to describe the evolution of stochastic dynamical systems in terms of the evolution of its standard transformed stochastic dynamical system. The system, which is obtained through a dynamic transformation of the original state variables with their state mean and covariance as parameters, allows for propagation of the state distribution to be performed in a transformed computational

domain where greater accuracy can be ensured. Note that a fixed grid in the transformed space corresponds to an adaptive grid in the original state space. Evolution of the underlying probability distribution(s) in the transformed computational domain can ensure greater accuracy and accurate estimation of error bounds via a Chebyshev's inequality. As a result, our approach enables efficient predictions in a broach class of stochastic dynamical systems, including those characterized by large drift, large diffusion, and highly localized PDFs.

We use operator splitting of the transformed FPE to develop appropriate short-time propagators that can handle challenges arising from rank deficient, i.e., singular diffusion coefficient matrices. The splitting scheme consists of a Liouville operator (Eq. (3.37)) and a Fokker-Planck operator (Eq. (3.38)); The latter has a full rank diffusion coefficient submatrix. We propose new update equations for mean and covariance of state variables along with evolution of the underlying state distribution in the transformed space. A novel feature of our approach is that under the dynamic transformation considered the proposed update equations ensure zero mean and identity covariance for the transformed state variables. Thus, in addition to the renormalization condition to preserve zeroth moment properties, our approach also allows us to establish conditions to preserve the first and second moment properties of the state distribution in the transformed space. In fact, the proposed framework allows us to capture invariant solutions for linear stochastic dynamical systems with a Gaussian initial distribution; The state distribution remains standard normal in the transformed space for such systems.

This work generalizes our earlier transformed path integral (TPI) formulation [93] which was restricted to stochastic dynamical systems with full rank, i.e., nonsingular diffusion coefficient matrices. Our non-trivial generalization extends the application of TPI to a broader class of stochastic dynamical systems with singular diffusion coefficient matrices. Notably, probabilistic descriptions of second and higher-order systems can be made using this framework. It can be applied to certain systems with colored noise excitation where the excitation may be expressed as a filtered white noise process. The framework can be used for

probabilistic description of systems with zero process noise but with uncertainties in initial conditions and parameters (equivalent to a solution of the Liouville equation). Likewise, the GTPI approach is also applicable to systems with full rank diffusion coefficient matrices.

To evaluate the performance of our generalized transformed path integral approach we considered several examples of linear and nonlinear stochastic dynamical systems with singular diffusion coefficient matrices. These examples include (a) the stochastic harmonic oscillator, (b) the stochastic van der Pol oscillator, (c) the stochastic Caughey oscillator, (d) the stochastic Duffing oscillator with zero process noise, and (e) bistable stochastic flow driven by non-white noise. Note that the evolution of the stochastic Duffing oscillator with zero process noise (example(d)) is governed by a Liouville equation. The examples and parameters were also chosen to highlight challenges faced by conventional fixed grid approaches to accurately describe evolution of distributions in dynamical systems with large drift, diffusion, and concentration of PDF. The results obtained from our proposed GTPI approach in these examples show good agreement with those from MC simulations and analytical solutions (where available). Comparisons of numerical error in the PDF show that the GTPI approach is also more efficient than conventional fixed grid approaches in terms of the computational costs at achieving a desired level of accuracy. The analysis of numerical error in the GTPI approach for the stochastic harmonic oscillator also indicate the presence of an optimal value ($\Delta t^\star$) for the time step below which the numerical simulations produce larger numerical error in the PDF. Both $\Delta t^\star$ and $\varepsilon^\star_{\mathrm{RMS}}$, the error in the PDF at this time step, are lower for a finer grid resolution. We also investigated the influence of the grid resolution and system parameters on $\Delta t^\star$ and $\varepsilon^\star_{\mathrm{RMS}}$ for the stochastic harmonic oscillator. In spite of the documented benefits of our proposed approach over conventional approaches, issues pertaining to the curse of dimensionality which may arise for instance in second order (and higher order) multiple degree of freedom systems need to be investigated further.

CHAPTER 4

Generalized Transformed Path Integral Based Approaches for Nonlinear Filtering

4.1 Scope of the Chapter

In this chapter the generalized transformed path integral filter (GTPIF) is presented. The framework is a path-integral–based approach for continuous-discrete nonlinear filtering problems. In the continuous-discrete filtering problem, the system dynamics is continuous in time and the measurements are taken at discrete time instants. A brief overview of the mathematical formulations for the problem are presented, namely, the probabilistic description of the problem.

The formulations for the proposed GTPIF approach are derived. The update equation for the evolution of the conditional distribution in the transformed space between observations are presented. Additionally, update equations for the conditional mean and covariance are also presented. At an observation, novel update equations for the conditional distribution, the conditional mean, and covariance are derived based on Bayes' theorem. The salient features of the approach are also presented.

4.2 Nonlinear Filtering

In the continuous-discrete nonlinear filtering problem, the evolution of the dynamical system is described by the Itô stochastic differential equation

$$\mathrm{d}\mathbf{x}(t) = \mathbf{f}\Big(\mathbf{x}(t), t\Big)\,\mathrm{d}t + \mathbf{A}\Big(\mathbf{x}(t), t\Big)\,\mathrm{d}\mathbf{w}(t) \tag{4.1}$$

where $\mathbf{x}(t)$ represents the state of the system at time t. The system evolves along a deterministic path determined by the drift vector function $\mathbf{f}\Big(\mathbf{x}(t), t\Big)$ while being subjected to Gaussian white noise excitations of strength $\mathbf{A}\Big(\mathbf{x}(t), t\Big)$. The increments $\mathrm{d}\mathbf{w}(t)$ represent independent identically distributed zero mean Gaussian random variables with the autocorrelation $\Big\langle \mathrm{d}w_i(t)\, \mathrm{d}w_j(s)^{\mathrm{T}} \Big\rangle = \delta(t-s)\, \delta_{ij}$ for $i, j = 1, \ldots, N_w$. The system is observed through the process

$$\mathbf{y_k} = \mathbf{h}\Big(\mathbf{x}(t_k), t_k\Big) + \boldsymbol{\nu_k} \tag{4.2}$$

where $\boldsymbol{\nu_k}$ represents zero mean Gaussian white noise with covariance matrix $\mathbf{M_k}$. The goal of the filtering problem is to evaluate the conditional distribution $p(\mathbf{x}, t \,|\, \mathcal{Y}_k)$ where $\mathcal{Y}_k = \Big\{\mathbf{y_0}, \ldots, \mathbf{y_k}\Big\}$ represents the set of observations obtained from Eq. (4.2). The evolution of this distribution between observations is governed by the Fokker-Planck equation

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial x_i} f_i(\mathbf{x}, t) - \frac{\partial^2}{\partial x_i \partial x_j} G_{ij}(\mathbf{x}, t)\right] p(\mathbf{x}, t \,|\mathcal{Y}_k) = 0, \qquad\qquad t_k \le t < t_{k+1} \tag{4.3}$$

where $\mathbf{G}(\mathbf{x}, t) = \mathbf{A}(\mathbf{x}, t)\, \mathbf{A}(\mathbf{x}, t)^{\mathrm{T}}/2$. At an observation, i.e., at t_k the conditional distribution can be evolved using the Bayes' update relation

$$p(\mathbf{x}, t_k \,|\, \mathcal{Y}_k) = \frac{p(\mathbf{y_k} \,|\, \mathbf{x})\, p(\mathbf{x}, t_k \,|\, \mathcal{Y}_{k-1})}{p(\mathbf{y_k} \,|\, \mathcal{Y}_{k-1})} \tag{4.4}$$

where the likelihood $p(\mathbf{y_k} \,|\, \mathbf{x})$ is given by

$$p(\mathbf{y_k} \,|\, \mathbf{x}) = \|2\pi\, \mathbf{M_k}\|^{-1/2} \exp\left\{-\frac{1}{2}\Big[\mathbf{y_k} - \mathbf{h}(\mathbf{x}, t_k)\Big]^{\mathrm{T}} \mathbf{M_k}^{-1} \Big[\mathbf{y_k} - \mathbf{h}(\mathbf{x}, t_k)\Big]\right\} \tag{4.5}$$

4.3 The Generalized Transformed Path Integral Filter

In our proposed solution for the continuous-discrete nonlinear filtering problem, we consider a dynamic transformation of the state variables:

$$\mathbf{z}(t) = \mathbf{R}_{t|k}^{-1}\Big[\mathbf{x}(t) - \boldsymbol{\mu}_{t|k}\Big] \tag{4.6}$$

where $\boldsymbol{\mu}_{t|k}$ is the conditional state mean and $\boldsymbol{\Sigma}_{t|k} = \mathbf{R}_{t|k}\,\mathbf{R}_{t|k}^{\mathrm{T}}$ the conditional state covariance, i.e.,

$$\begin{aligned}
\boldsymbol{\mu}_{t|k} &= \langle \mathbf{x} \,|\, \mathcal{Y}_k \rangle \\
&= \int \mathbf{x}\, p_{\mathbf{x}(t)}(\mathbf{x}, t \,|\, \mathcal{Y}_k)\, \mathrm{d}\mathbf{x} \qquad (4.7) \\
\boldsymbol{\Sigma}_{t|k} &= \left\langle \Big[\mathbf{x} - \boldsymbol{\mu}_{t|k}\Big]\Big[\mathbf{x} - \boldsymbol{\mu}_{t|k}\Big]^{\mathrm{T}} \,\middle|\, \mathcal{Y}_k \right\rangle \\
&= \int \Big[\mathbf{x} - \boldsymbol{\mu}_{t|k}\Big]\Big[\mathbf{x} - \boldsymbol{\mu}_{t|k}\Big]^{\mathrm{T}} p_{\mathbf{x}(t)}(\mathbf{x}, t \,|\, \mathcal{Y}_k)\, \mathrm{d}\mathbf{x}. \qquad (4.8)
\end{aligned}$$

Given this transformation, we may re-formulate our filtering problem to one of evaluating the conditional distribution $p_{\mathbf{z}(t)}(\mathbf{z}, t \,|\, \mathcal{Y}_k)$. This conditional distribution in the transformed state variables is related to that of the original state variables as

$$p_{\mathbf{z}(t)}(\mathbf{z}, t \,|\, \mathcal{Y}_k) = \Big\|\mathbf{R}_{t|k}\Big\|\; p_{\mathbf{x}(t)}\Big(\mathbf{R}_{t|k}\,\mathbf{z} + \boldsymbol{\mu}_{t|k}, t \,\Big|\, \mathcal{Y}_k\Big) \tag{4.9}$$

In further discussions, we will drop the subscripts in our representations of the distribution functions and rely on the reader to distinguish between them based on the context.

System Update

The evolution of $p(\mathbf{z}, t \,|\, \mathcal{Y}_k)$ between observations is governed by the transformed Fokker-Planck equation

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial z_i}\Phi_i(\mathbf{z}, t) - \frac{\partial^2}{\partial z_i \partial z_j}\Gamma_{ij}(\mathbf{z}, t)\right] p(\mathbf{z}, t \,|\, \mathcal{Y}_k) = 0, \qquad t_k \le t < t_{k+1} \tag{4.10}$$

where the drift vector function $\mathbf{\Phi} \equiv [\Phi_i]$ and diffusion coefficient matrix $\mathbf{\Gamma} \equiv [\Gamma_{ij}]$ are given by

$$\mathbf{\Phi}(\mathbf{z}, t) = \mathbf{R}_{t|k}^{-1}\Big(\tilde{\mathbf{f}}(\mathbf{z}, t) - \dot{\mathbf{R}}_{t|k}\,\mathbf{z} - \dot{\boldsymbol{\mu}}_{t|k}\Big) \tag{4.11}$$

$$\mathbf{\Gamma}(\mathbf{z}, t) = \mathbf{R}_{t|k}^{-1}\,\tilde{\mathbf{G}}(\mathbf{z}, t)\,\mathbf{R}_{t|k}^{-1^{\mathrm{T}}} \tag{4.12}$$

with $\tilde{\mathbf{f}}(\mathbf{z}, t) = \mathbf{f}\Big(\mathbf{R}_{t|k}\,\mathbf{z} + \boldsymbol{\mu}_{t|k}, t\Big)$ and $\tilde{\mathbf{G}}(\mathbf{z}, t) = \mathbf{G}\Big(\mathbf{R}_{t|k}\,\mathbf{z} + \boldsymbol{\mu}_{t|k}, t\Big)$. If $\mathbf{G}$ is singular, and hence $\mathbf{\Gamma}$, through appropriate transformation of the state variables where necessary $\mathbf{\Gamma}$ can be expressed in the block matrix form

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}^{(\mathbf{r})} \end{bmatrix} \tag{4.13}$$

where $\mathbf{0}$ represents appropriate zero-matrices. Note that in this representation $\mathbf{\Gamma}^{(\mathbf{r})} \in \mathbb{R}^{N_r \times N_r}$ is nonsingular and $N_r = \operatorname{rank}(\mathbf{\Gamma})$. Let $N_\varepsilon = N_s - N_r$. Without loss of generality, we may assume the dynamical system and $\mathbf{\Gamma}$ to already be of this form.

In our approach, informed by the form of $\mathbf{\Gamma}$, we consider grouping the transformed state variables $\mathbf{z} = \Big[z_1, z_2, \ldots, z_{N_s}\Big]^{\mathrm{T}}$ into singular variables $\mathbf{q} = \Big[z_1, \ldots, z_{N_\varepsilon}\Big]^{\mathrm{T}}$ and nonsingular variables $\mathbf{v} = \Big[z_{N_\varepsilon+1}, \ldots, z_{N_s}\Big]^{\mathrm{T}}$. Similarly, we define $\mathbf{\Phi}^{(\boldsymbol{\varepsilon})} = \Big[\Phi_1^{(\varepsilon)}, \ldots, \Phi_{N_\varepsilon}^{(\varepsilon)}\Big]^{\mathrm{T}}$ and $\mathbf{\Phi}^{(\boldsymbol{r})} =$

$\left[\Phi^{(r)}_{N_\varepsilon+1}, \ldots, \Phi^{(r)}_{N_s}\right]^{\mathrm{T}}$ allowing us to write Eq. (4.10) in the form

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial q_i}\Phi^{(\varepsilon)}_i(\mathbf{q},\mathbf{v},t) + \frac{\partial}{\partial v_i}\Phi^{(r)}_i(\mathbf{q},\mathbf{v},t) - \frac{\partial^2}{\partial v_i \partial v_j}\Gamma^r_{ij}(\mathbf{q},\mathbf{v},t)\right] p(\mathbf{q},\mathbf{v},t\,|\,\mathcal{Y}_k) = 0, \tag{4.14}$$

for $t_k \leq t < t_{k+1}$. The solution to Eq. (4.14) via the two-step generalized transformed path integral (GTPI) approach is given by

$$\hat{p}(\mathbf{q}',\mathbf{v}\,|\,\mathcal{Y}_k) = p(\mathbf{q},\mathbf{v},t\,|\,\mathcal{Y}_k)\, \exp\left\{-\Delta t\, \left.\frac{\partial \Phi^{(\varepsilon)}_i}{\partial q_i}\right|_{\mathbf{q},\mathbf{v},t}\right\}, \tag{4.15}$$

$$p(\mathbf{q}',\mathbf{v}',t+\Delta t\,|\,\mathcal{Y}_k) = \int p(\mathbf{v}'\,|\,\mathbf{v})\, \hat{p}(\mathbf{q}',\mathbf{v}\,|\,\mathcal{Y}_k)\, \mathrm{d}\mathbf{v} \tag{4.16}$$

Note that $\hat{p}$ in Eq. (4.15) is computed at $\begin{bmatrix}\mathbf{q}' & \mathbf{v}\end{bmatrix}^{\mathrm{T}}$ with $\mathbf{q}' = \mathbf{q} + \Delta t\, \boldsymbol{\Phi}^{(\boldsymbol{\varepsilon})}(\mathbf{q},\mathbf{v},t)$. In grid-based approaches, this could represent off-grid points and might require interpolation of the distributions evaluated at those points back on to the points of our computational grid. The short-time propagator $p(\mathbf{v}'\,|\,\mathbf{v})$ in Eq. (4.16) is given by

$$p(\mathbf{v}'\,|\,\mathbf{v}) = \left\|4\,\pi\,\Delta t\,\boldsymbol{\Gamma}^{(\mathbf{r})}(\mathbf{q},\mathbf{v},t)\right\|^{-1/2} \exp\left\{-\frac{1}{4\,\Delta t}\left[\mathbf{v}_\mathbf{e}^{\mathrm{T}}\,\boldsymbol{\Gamma}^{(\mathbf{r})}(\mathbf{q},\mathbf{v},t)^{-1}\,\mathbf{v}_\mathbf{e}\right]\right\} \tag{4.17}$$

with $\mathbf{v_e} = \mathbf{v}' - \mathbf{v} - \Phi^{(r)}(\mathbf{q},\mathbf{v},t)\,\Delta t$. A first order approximation $\Delta\mathbf{R} \equiv \mathbf{R}_{t+\Delta t|k} - \mathbf{R}_{t|k}$ is employed in evaluating $\boldsymbol{\Phi}(\mathbf{q},\mathbf{v},t)\,\Delta t = \mathbf{R}^{-1}_{t|k}\left[\boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{q},\mathbf{v},t)\,\Delta t - \Delta\mathbf{R}\,\mathbf{z}\right]$ with $\boldsymbol{\delta}\tilde{\mathbf{f}}(\mathbf{q},\mathbf{v},t) = \tilde{\mathbf{f}}(\mathbf{q},\mathbf{v},t) - \left\langle\tilde{\mathbf{f}}(\mathbf{q},\mathbf{v},t)\right\rangle$. In addition, we have the update relations

$$\boldsymbol{\mu}_{t+\Delta t|k} = \boldsymbol{\mu}_{t|k} + \Delta t\left\langle\tilde{\mathbf{f}}(\mathbf{z},t)\,|\,\mathcal{Y}_k\right\rangle, \tag{4.18}$$

$$\boldsymbol{\Sigma}_{t+\Delta t|k} = \boldsymbol{\Sigma}_{t|k} + \Delta t\Big[\mathbf{R}_{t|k}\left\langle\mathbf{z}\,\tilde{\mathbf{f}}(\mathbf{z},t)^{\mathrm{T}}\,|\,\mathcal{Y}_k\right\rangle + \left\langle\tilde{\mathbf{f}}(\mathbf{z},t)\,\mathbf{z}^{\mathrm{T}}\,|\,\mathcal{Y}_k\right\rangle\mathbf{R}^{\mathrm{T}}_{t|k} + 2\left\langle\tilde{\mathbf{G}}(\mathbf{z},t)\,|\,\mathcal{Y}_k\right\rangle\Big]. \tag{4.19}$$

We obtain $\mathbf{R}_{t+\Delta t|k}$ by performing a Cholesky factorization of $\mathbf{\Sigma}_{t+\Delta t|k}$.

Measurement Update

We consider the transformations

$$\mathbf{z}^* = \mathbf{R}_{k|k}^{-1}\left[\mathbf{x}(t_k) - \boldsymbol{\mu}_{k|k}\right] \qquad \text{and} \qquad \mathbf{z}' = \mathbf{R}_{k|k-1}^{-1}\left[\mathbf{x}(t_k) - \boldsymbol{\mu}_{k|k-1}\right] \tag{4.20}$$

where $\boldsymbol{\mu}_{k|k}$ is the posterior conditional state mean and $\mathbf{\Sigma}_{k|k} = \mathbf{R}_{k|k}\mathbf{R}_{k|k}^{\mathrm{T}}$ is the posterior conditional state covariance obtained from $p(\mathbf{x}, t_k \,|\, \mathcal{Y}_k)$. Similarly, $\boldsymbol{\mu}_{k|k-1}$ and $\mathbf{R}_{k|k-1}$ are the corresponding quantities associated with the prior distribution $p(\mathbf{x}, t_k \,|\, \mathcal{Y}_{k-1})$. At an observation, i.e., at t_k the prior and posterior distributions satisfy the update relation in Eq. (4.4). Applying the transformations from Eq. (4.20), we obtain the relation

$$p(\mathbf{z}^*, t \,|\, \mathcal{Y}_k) = \frac{\left\|\mathbf{R}_{t|k}\right\|}{\left\|\mathbf{R}_{t|k-1}\right\|} \frac{1}{N}\, p(\mathbf{y_k} \,|\, \mathbf{z}')\, p(\mathbf{z}', t \,|\, \mathcal{Y}_{k-1}) \tag{4.21}$$

where $N = \int p(\mathbf{y_k} \,|\, \mathbf{z}')\, p(\mathbf{z}', t \,|\, \mathcal{Y}_{k-1})\, \mathrm{d}\mathbf{z}'$ and $\mathbf{z}' = \mathbf{R}_{k|k-1}^{-1}\Big[\mathbf{R}_{k|k}\, \mathbf{z}^* + \boldsymbol{\mu}_{k|k} - \boldsymbol{\mu}_{k|k-1}\Big]$. The likelihood $p(\mathbf{y_k} \,|\, \mathbf{z}')$ is given by

$$p(\mathbf{y_k} \,|\, \mathbf{z}') = \|2\pi\, \mathbf{M_k}\|^{-1/2} \exp\left\{-\frac{1}{2}\left[\mathbf{y_k} - \tilde{\mathbf{h}}(\mathbf{z}', t_k)\right]^{\mathrm{T}} \mathbf{M}_\mathbf{k}^{-1}\Big[\mathbf{y_k} - \tilde{\mathbf{h}}(\mathbf{z}', t_k)\Big]\right\} \tag{4.22}$$

with $\tilde{\mathbf{h}}(\mathbf{z}', t_k) = \mathbf{h}\Big(\mathbf{R}_{k|k-1}^{-1}\left[\mathbf{x} - \boldsymbol{\mu}_{k|k-1}\right], t_k\Big)$. Eq. (4.21) is the measurement update relation for the conditional distribution of the transformed state variables. However, we also need to specify update relations for the conditional mean and covariance, i.e, relations to obtain $\boldsymbol{\mu}_{k|k}$ and $\mathbf{R}_{k|k}$ from $\boldsymbol{\mu}_{k|k-1}$ and $\mathbf{R}_{k|k-1}$. Let us define

$$\boldsymbol{\psi_1} = \frac{1}{N}\int_{-\infty}^{\infty} \mathbf{z}'\, p(\mathbf{y_k} \,|\, \mathbf{z}')\, p(\mathbf{z}', t \,|\, \mathcal{Y}_{k-1}))\, \mathrm{d}\mathbf{z}' \tag{4.23}$$

$$\boldsymbol{\psi_2} = \frac{1}{N}\int \Big[\mathbf{z}' - \boldsymbol{\psi_1}\Big]\Big[\mathbf{z}' - \boldsymbol{\psi_1}\Big]^{\mathrm{T}}\, p(\mathbf{y_k} \,|\, \mathbf{z}')\, p(\mathbf{z}', t \,|\, \mathcal{Y}_{k-1})\, \mathrm{d}\mathbf{z}' \tag{4.24}$$

Taking appropriate expectations of Eq. (4.21), we can then show that

$$\langle \mathbf{z}^* \,|\, \mathcal{Y}_k \rangle = \mathbf{R}_{t|k}^{-1} \left[\mathbf{R}_{t|k-1} \boldsymbol{\psi_1} + \boldsymbol{\mu}_{t|k-1} - \boldsymbol{\mu}_{t|k} \right] \tag{4.25}$$

$$\left\langle \mathbf{z}^* \mathbf{z}^{*\mathrm{T}} \,|\, \mathcal{Y}_k \right\rangle = \mathbf{R}_{t|k}^{-1} \, \mathbf{R}_{t|k-1} \, \boldsymbol{\psi_2} \, \mathbf{R}_{t|k-1}^{\mathrm{T}} \, \mathbf{R}_{t|k}^{-1^{\mathrm{T}}} \tag{4.26}$$

We seek update relations that ensure $\langle \mathbf{z}^* \,|\, \mathcal{Y}_k \rangle = \mathbf{0}$ and $\left\langle \mathbf{z}^* \mathbf{z}^{*\mathrm{T}} \,|\, \mathcal{Y}_k \right\rangle = \mathbf{I}$. Thus, we get

$$\boldsymbol{\mu}_{t|k} = \boldsymbol{\mu}_{t|k-1} + \mathbf{R}_{t|k-1} \, \boldsymbol{\psi_1} \tag{4.27}$$

$$\boldsymbol{\Sigma}_{t|k} = \mathbf{R}_{t|k-1} \, \boldsymbol{\psi_2} \, \mathbf{R}_{t|k-1}^{\mathrm{T}} \tag{4.28}$$

We obtain $\mathbf{R}_{k|k}$ by taking the Cholesky factorization of $\boldsymbol{\Sigma}_{k|k}$.

4.4 Performance of the Generalized Transformed Path Integral Filter

In this section, the benefits of the GTPI-based filter (GTPIF) for nonlinear filtering are illustrated. The conventional extended Kalman filter (EKF) is used extensively to obtain solutions to several nonlinear filtering problems. However, the approach is not well-equipped to obtain accurate estimates in dynamical systems with large nonlinearities. Such nonlinearities may arise from the system dynamics and/or the measurement function. In contrast, the GTPIF is better equipped to address challenges from such large nonlinearities.

A stochastic dynamical system with a cubic nonlinearity is considered. The evolution of the system is described the equation

$$\mathrm{d}x = -kx(x^2 - c)\,\mathrm{d}t + a\mathrm{d}w \tag{4.29}$$

with $k = 4$ and $c = 1$. Clearly, the system has two stable stationary points at $x = \pm c = \pm 1$. Intuitively, we expect the system state to evolve towards one of these points. The strength

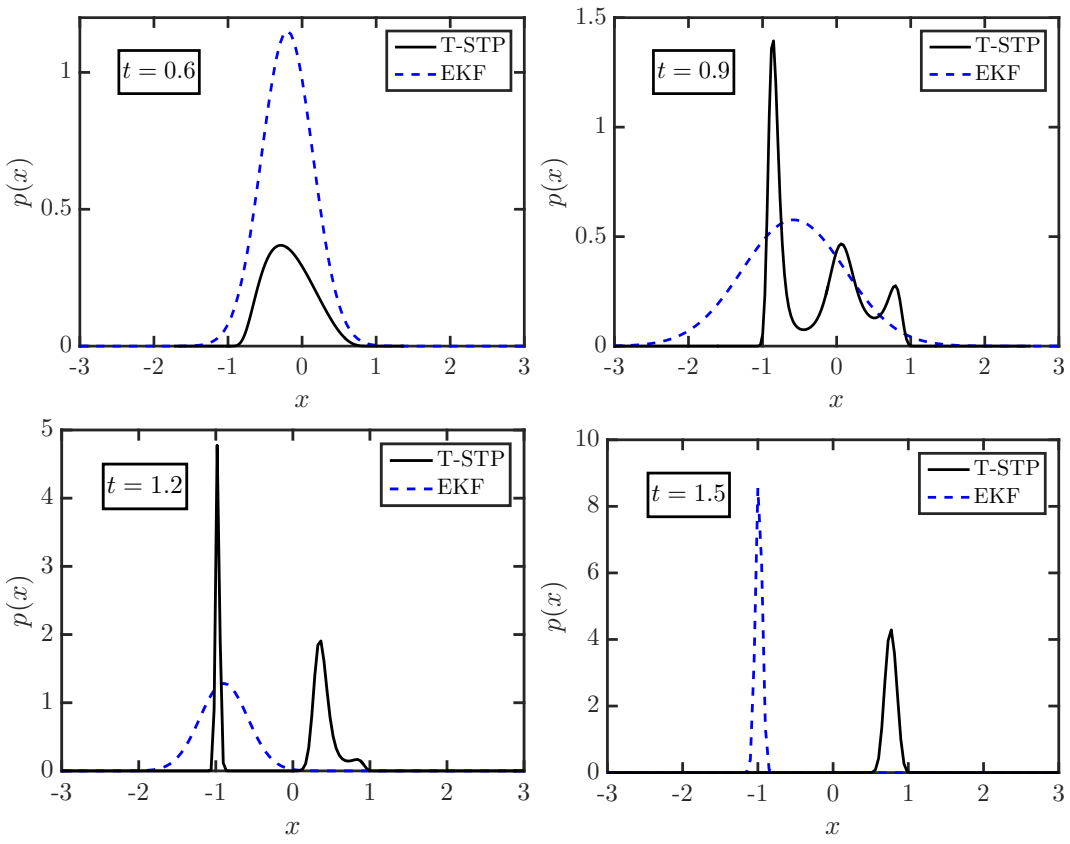

Figure 4.1: Evolution of the state distribution conditioned on the observation history for the nonlinear filtering problem (Eqs. (4.29) and (4.30)) at different times.

of the white noise excitation is $a = 0.1$. The system is observed through the measurement process

$$y_k = x + \nu_k, \tag{4.30}$$

with $\nu_k = 0.2$. The initial distribution for the state was chosen to be $\mathcal{N}(-0.02, 10^{-9})$. The evolution of the state distribution given a set of observation values is shown in Fig. 4.1.

The simulations were performed on a grid with 151 grid points between bounds $[-5, 5]$ in the transformed space. The simulations were performed with a time step $\Delta t = 0.01$ from $t = 0\,\mathrm{s}$ to $t = 2\,\mathrm{s}$. The GTPI based filter is able to better represent the non-Gaussian transient behavior of the distribution for this system. Moreover, owing to the presence of grid points at the regions of importance in the GTPI based approach, it is able to better describe the evolution of the distribution when there is large drift. Also, as the system evolves and

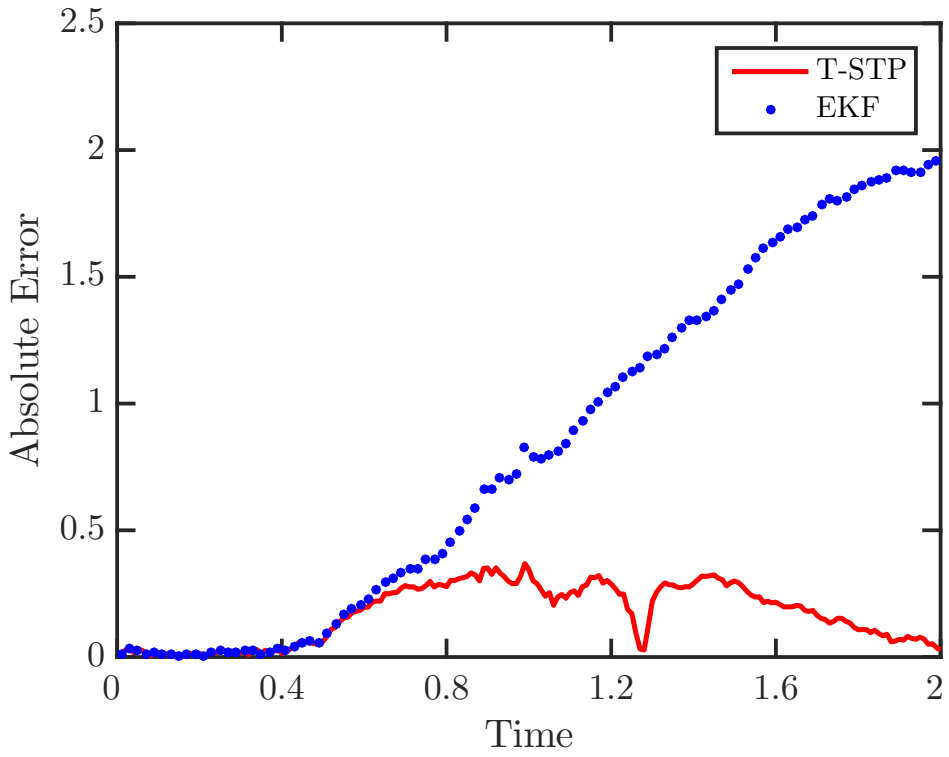

Figure 4.2: Comparison of the error in the state estimate obtained from the Extended Kalman Filter and the GTPI based filter for the nonlinear filtering problem Eqs. (4.29) and (4.30).

distribution is updated with the measurement information, there is large concentration of the PDF. The GTPI based approach is able to better represent the distribution in regions where there is large concentration of PDF. While the EKF estimate is initially comparable to that of the GTPI based filter, over time the error for the EKF estimate can grow rapidly as shown in Fig. 4.2.

4.5 Concluding Remarks

The generalized transformed path integral filter (GTPIF)—a novel path integral-based filter was presented. New approaches for updating the prior probabilities as well as posterior probabilities (consistent with Bayes' rule) in the transformed space were proposed. In particular, a set of update equations for the evolution of the conditional distribution in the transformed space along with update equations for the conditional mean and covariance of the original

state variables were derived. The proposed approach is able to better accommodate nonlinearities in the system dynamics and measurement function than the conventional extended Kalman filter (EKF). The approach is better equipped than the Unscented Kalman filter (UKF) at preserving higher order moments of the conditional distribution. GTPIF does not encounter the sampling errors present in Monte-Carlo (MC) based approaches and have lower errors than conventional grid based approaches for comparable computational costs. The set of update equations for GTPIF are better equipped to address the challenges arising from fixed computational domain bounds and grid resolution in conventional fixed grid based approaches. The approach allows for straightforward evaluation of error bounds on evolved distributions via Chebyshev's inequality.

CHAPTER 5

Generalized Transformed Path Integral Based Approaches for Stochastic Optimal Control

5.1 Scope of the Chapter

In this chapter, the generalized transformed path integral control (GTPIC) is presented. The proposed framework is a transformed path-based approach for the control of a large class of stochastic dynamical systems where the system dynamics is linear and the cost function is quadratic in terms of the control variables. However, the dynamics and cost may still contain nonlinearities with respect to the state variables. The governing equation for problems in stochastic optimal control, the stochastic Hamilton-Jacobi-Bellman (HJB) equation, is reduced to a more tractable form for these systems.

The mathematical formulations for the stochastic optimal control problem are presented in this chapter. A brief derivation of the stochastic HJB is also provided. Likewise, the formulations for the class of problems considered here are shown where the governing equation can be reduced to a backward Kolmogorov type equation. Its adjoint equation, a Fokker-Planck type equation, governs the evolution of the distribution for a diffusion process that occurs alongside a killing process. The equation forms the basis for the use of path integrals to solve this class of stochastic optimal control problems.

The framework for the GTPIC is presented. A set of update equations for solving the Fokker-Planck type equation are presented. A short-time propagator based on the GTPI approach that allows for the evolution of distributions in the transformed space is presented. Necessary update equations for the conditional mean and variance are also presented.

5.2 Stochastic Optimal Control

Let us consider a controlled stochastic dynamical system described by the Itô stochastic differential equation

$$\mathrm{d}\mathbf{x}(t) = \mathbf{f}\Big(\mathbf{x}(t), \mathbf{u}(t), t\Big)\,\mathrm{d}t + \mathbf{A}\Big(\mathbf{x}(t), \mathbf{u}(t), t\Big)\,\mathrm{d}\mathbf{w}(t) \tag{5.1}$$

where the random vector $\mathbf{x}(t) \in \mathbb{R}^{N_s \times 1}$ represents the state of the system at time t and $\mathbf{u}(t) \in \mathbb{R}^{N_u \times 1}$ is the set of control inputs to the system at time t. The drift vector function $\mathbf{f}\Big(\mathbf{x}(t), \mathbf{u}(t), t\Big)$ specifies the deterministic path of evolution for the system and $\mathbf{A}\Big(\mathbf{x}(t), \mathbf{u}(t), t\Big)$ is the strength of the Gaussian white noise excitations. The increments $\mathrm{d}\mathbf{w}(t)$ represent independent identically distributed zero mean Gaussian random variables with the autocorrelation $\Big\langle \mathrm{d}w_i(t)\,\mathrm{d}w_j(s)^{\mathrm{T}} \Big\rangle = \delta(t-s)\,\delta_{ij}$ for $i, j = 1, \ldots, N_w$. We consider a cost functional of the Bolza type given by

$$C\Big(\mathbf{x_0}, t_0, \mathbf{u}(\cdot)\Big) = \left\langle \phi\Big(\mathbf{x}(T)\Big) + \int_{t_0}^{T} L\Big(\mathbf{x}(t), \mathbf{u}(t), t\Big)\,\mathrm{d}t \,\middle|\, \mathbf{x}(t_0) \right\rangle \tag{5.2}$$

which represents an expected cost conditioned on the initial state of the system $\mathbf{x}(t_0)$. It consists of the terminal cost $\phi\Big(\mathbf{x}(T)\Big)$ and the path cost (or running cost) $L\Big(\mathbf{x}(t), \mathbf{u}(t), t\Big)$. The objective of the stochastic optimal control problem is to find the control function $\mathbf{u}^\star(t)$ for $t_0 \leq t < T$ such that the cost $C\Big(\mathbf{x_0}, t_0, \mathbf{u}(\cdot)\Big)$ is minimized. In other words,

$$\mathbf{u}^\star(\cdot) = \underset{\mathbf{u}(\cdot)}{\operatorname{argmin}}\, C\Big(\mathbf{x_0}, t_0, \mathbf{u}(\cdot)\Big) \tag{5.3}$$

subject to the constraint given by Eq. (5.1).

The Stochastic Hamilton-Jacobi-Bellman Equation

Let us define the value function (or cost-to-go) as

$$V(\mathbf{x},t) = \min_{\mathbf{u}(\cdot)} \left\langle \phi\Big(\mathbf{x}(T)\Big) + \int_t^T L\Big(\mathbf{x}(s),\mathbf{u}(s),s\Big)\,\mathrm{d}s \,\middle|\, \mathbf{x}(t)=\mathbf{x} \right\rangle \tag{5.4}$$

which is the optimal cost to evolve the system in Eq. (5.1) starting from state $\mathbf{x}$ at time t. Applying Bellman's principle of optimality, we can write Eq. (5.4) as the recursive relation

$$V(\mathbf{x},t) = \min_{\mathbf{u}(\cdot)} \left\langle V(\mathbf{x}',t') + \int_t^{t'} L\Big(\mathbf{x}(s),\mathbf{u}(s),s\Big)\,\mathrm{d}s \,\middle|\, \mathbf{x} \right\rangle. \tag{5.5}$$

where the minimization is with respect to $\mathbf{u}(s)$ for $t \le s < t'$. Let us consider $t' = t + \Delta t$. Applying Dynkin's formula, we obtain

$$\min_{\mathbf{u}(\cdot)} \left\langle \int_t^{t+\Delta t} \Big[\mathcal{A}\,V\Big(\mathbf{x}(s),s\Big) + L\Big(\mathbf{x}(s),\mathbf{u}(s),s\Big)\Big]\mathrm{d}s \,\middle|\, \mathbf{x} \right\rangle = 0 \tag{5.6}$$

where $\mathcal{A}$ is the infinitesimal generator of the stochastic process $\begin{bmatrix}\mathrm{d}t & \mathrm{d}\mathbf{x}(t)\end{bmatrix}^{\mathrm{T}}$. It is given by

$$\mathcal{A} = \frac{\partial}{\partial t} + f_i(\mathbf{x},\mathbf{u},t)\frac{\partial}{\partial x_i} + G_{ij}(\mathbf{x},\mathbf{u},t)\frac{\partial^2}{\partial x_i \partial x_j}. \tag{5.7}$$

Thus, in the limit $\Delta t \to 0$ we have

$$-\frac{\partial V(\mathbf{x},t)}{\partial t} = \min_{\mathbf{u}} \left[L(\mathbf{x},\mathbf{u},t) + f_i(\mathbf{x},\mathbf{u},t)\frac{\partial V(\mathbf{x},t)}{\partial x_i} + G_{ij}(\mathbf{x},\mathbf{u},t)\frac{\partial^2 V(\mathbf{x},t)}{\partial x_i \partial x_j} \right] \tag{5.8}$$

which is solved backwards in time with the boundary condition $V(\mathbf{x},T) = \phi(\mathbf{x})$. The equation is nonlinear because of the pointwise minimization with respect to $\mathbf{u}$. It is known as the stochastic Hamilton-Jacobi-Bellman (HJB) equation.

Path Integral Control

Let us consider a stochastic dynamical system of the form

$$\mathrm{d}\mathbf{x}(t) = \Big[\mathbf{f}\Big(\mathbf{x}(t),t\Big) + \mathbf{B}\Big(\mathbf{x}(t),t\Big)\,\mathbf{u}(t)\Big]\mathrm{d}t + \mathbf{A}\Big(\mathbf{x}(t),t\Big)\,\mathrm{d}\mathbf{w}(t). \tag{5.9}$$

and a cost functional of the form

$$C\Big(\mathbf{x_0},t_0,\mathbf{u}(\cdot)\Big) = \Big\langle \phi\Big(\mathbf{x}(T)\Big) + \int_{t_0}^{T}\Big[L\Big(\mathbf{x}(t),t\Big) + \frac{1}{2}\mathbf{u}(t)^{\mathrm{T}}\mathbf{Q}\Big(\mathbf{x}(t),t\Big)\,\mathbf{u}(t)\Big]\,\mathrm{d}t \,\Big|\, \mathbf{x_0}\Big\rangle. \tag{5.10}$$

In other words, we consider system dynamics linear in $\mathbf{u}$ and a path cost function quadratic in $\mathbf{u}$. For the case considered, the stochastic HJB becomes

$$\begin{aligned}-\frac{\partial V(\mathbf{x},t)}{\partial t} = \min_{\mathbf{u}}\Big[L(\mathbf{x},t) + \frac{1}{2}u_i u_j Q_{ij}(\mathbf{x},t) \\ + \Big[f_i(\mathbf{x},t) + u_j B_{ij}(\mathbf{x},t)\Big]\frac{\partial V(\mathbf{x},t)}{\partial x_i} + G_{ij}\Big(\mathbf{x},t\Big)\frac{\partial^2 V(\mathbf{x},t)}{\partial x_i \partial x_j}\Big].\end{aligned} \tag{5.11}$$

Minimization with respect to $\mathbf{u}$ gives us

$$\mathbf{u}^{\star}(\mathbf{x},t) = -\mathbf{Q}(\mathbf{x},t)^{-1}\mathbf{B}(\mathbf{x},t)^{\mathrm{T}}\,\nabla_{\mathbf{x}}V(\mathbf{x},t). \tag{5.12}$$

Substituting Eq. (5.12) in Eq. (5.11) we get

$$\begin{aligned}-\frac{\partial V(\mathbf{x},t)}{\partial t} = L(\mathbf{x},t) - \frac{1}{2}\frac{\partial V(\mathbf{x},t)}{\partial x_i}\frac{\partial V(\mathbf{x},t)}{\partial x_j}M_{ij}(\mathbf{x},t) \\ + f_i(\mathbf{x},t)\frac{\partial V(\mathbf{x},t)}{\partial x_i} + G_{ij}(\mathbf{x},t)\frac{\partial^2 V(\mathbf{x},t)}{\partial x_i \partial x_j}\end{aligned} \tag{5.13}$$

with $\mathbf{M}(\mathbf{x},t) = \mathbf{B}(\mathbf{x},t)\,\mathbf{Q}(\mathbf{x},t)^{-1}\mathbf{B}(\mathbf{x},t)^{\mathrm{T}}$. In order to make Eq. (5.13) more tractable, we consider the Cole-Hopf transformation

$$V(\mathbf{x},t) = -\lambda \log \psi(\mathbf{x},t) \tag{5.14}$$

with a yet to be defined constant λ. Under this transformation, we see that

$$\begin{aligned}
&-\frac{1}{2}\frac{\partial V(\mathbf{x},t)}{\partial x_i}\frac{\partial V(\mathbf{x},t)}{\partial x_j}M_{ij}(\mathbf{x},t)+G_{ij}(\mathbf{x},t)\frac{\partial^2 V(\mathbf{x},t)}{\partial x_i \partial x_j} \\
&\qquad = -\frac{\lambda}{\psi(\mathbf{x},t)}G_{ij}(\mathbf{x},t)\frac{\partial^2 \psi(\mathbf{x},t)}{\partial x_i \partial x_j} \\
&\qquad\quad + \frac{\lambda}{\psi(\mathbf{x},t)^2}\left[G_{ij}(\mathbf{x},t)-\frac{\lambda}{2}M_{ij}(\mathbf{x},t)\right]\frac{\partial \psi(\mathbf{x},t)}{\partial x_i}\frac{\partial \psi(\mathbf{x},t)}{\partial x_j}
\end{aligned} \tag{5.15}$$

If we choose λ such that

$$\mathbf{G}(\mathbf{x},t)=\frac{1}{2}\lambda\,\mathbf{M}(\mathbf{x},t) \tag{5.16}$$

the nonlinear terms in Eq. (5.15) will vanish. Such a λ is always possible in the one dimensional case. However, in multivariate dynamical systems the condition Eq. (5.16) restricts the possible values for $\mathbf{G}(\mathbf{x},t)$ and $\mathbf{Q}(\mathbf{x},t)$. In cases where Eq. (5.16) holds, the relation implies that the controls are less expensive along dimensions with higher process noise strength. This insight makes sense from a control theoretic standpoint; under a large disturbance, a significant control authority is required to bring the system back to a desirable state. Thus, with the transformation in Eq. (5.14) and the relation in Eq. (5.16), we may reduce Eq. (5.13) to

$$\begin{aligned}
\frac{\partial \psi(\mathbf{x},t)}{\partial t} &= -\left[f_i(\mathbf{x},t)\frac{\partial}{\partial x_i}+G_{ij}(\mathbf{x},t)\frac{\partial^2}{\partial x_i \partial x_j}-\frac{L(\mathbf{x},t)}{\lambda}\right]\psi(\mathbf{x},t) \\
&= -\mathcal{H}\,\psi(\mathbf{x},t)
\end{aligned} \tag{5.17}$$

where $\mathcal{H}$ is a linear operator operating on $\psi(\mathbf{x},t)$. Eq. (5.17) is of the form of a backward Kolmogorov equation with the boundary condition $\psi(\mathbf{x},T)=\exp\{-\phi(\mathbf{x})/\lambda\}$. We can consider a stochastic process $\boldsymbol{\xi}(t)$ whose infinitesimal generator is $\mathcal{H}$. It is to be noted that this process is different from the stochastic process in Eq. (5.9). The evolution of the conditional

probability density function for this process is governed by the equation

$$\begin{aligned}\frac{\partial \rho(\mathbf{x}', t' \,|\, \mathbf{x}, t)}{\partial t'} &= \left[-\frac{\partial}{\partial x_i'} f_i(\mathbf{x}', t') + \frac{\partial^2}{\partial x_i' \partial x_j'} G_{ij}(\mathbf{x}', t') - \frac{L(\mathbf{x}', t')}{\lambda} \right] \rho(\mathbf{x}', t' \,|\, \mathbf{x}, t) \\ &= \mathcal{H}^\star \rho(\mathbf{x}', t' \,|\, \mathbf{x}, t) \end{aligned} \tag{5.18}$$

with initial condition $\rho(\mathbf{x}', t \,|\, \mathbf{x}, t) = \delta(\mathbf{x}' - \mathbf{x})$ where $\boldsymbol{\xi}(t') = \mathbf{x}'$ and $\boldsymbol{\xi}(t) = \mathbf{x}$. Here, $\mathcal{H}^\star$ is the Hermitian adjoint of $\mathcal{H}$. Thus, the quantity $\langle \rho \,|\, \psi \rangle = \int_{-\infty}^{\infty} \rho(\mathbf{x}', t' \,|\, \mathbf{x}, t)\, \psi(\mathbf{x}', t')\, \mathrm{d}\mathbf{x}'$ is independent of t'. This can be seen by taking the derivative of $\langle \rho \,|\, \psi \rangle$ with respect to t' and applying the property $\langle \rho \,|\, \mathcal{H}\, \psi \rangle = \langle \mathcal{H}^\star \rho \,|\, \psi \rangle$. Evaluating $\langle \rho \,|\, \psi \rangle$ at $t' = t$ and at $t' = T$, we arrive at the useful relation

$$\psi(\mathbf{x}, t) = \int_{-\infty}^{\infty} \rho(\mathbf{x}', T \,|\, \mathbf{x}, t)\, \psi(\mathbf{x}', T)\, \mathrm{d}\mathbf{x}'. \tag{5.19}$$

Thus, we can compute $\psi(\mathbf{x}, t)$ by solving Eq. (5.18) to obtain $\rho(\mathbf{x}', T \,|\, \mathbf{x}, t)$ and then applying Eq. (5.19). An iterative path integral–based solution of Eq. (5.18) involves the evolution of conditional distribution by repeated application of the Chapman-Kolmogorov equation

$$\rho(\mathbf{x_{i+1}}, t_i + \Delta t \,|\, \mathbf{x}, t) = \int_{-\infty}^{\infty} \rho(\mathbf{x_{i+1}}, t_i + \Delta t \,|\, \mathbf{x_i}, t_i)\, \rho(\mathbf{x_i}, t_i \,|\, \mathbf{x}, t)\, \mathrm{d}\mathbf{x_i} \tag{5.20}$$

for $i = 1, \ldots, N$ using the short-time propagator

$$\begin{aligned}\rho(\mathbf{x_{i+1}}, t_i + \Delta t \,|\, \mathbf{x_i}, t_i) = &\left\| 4\pi \Delta t\, \mathbf{G}\Big(\mathbf{x_i}, t_i\Big) \right\|^{-1/2} \\ &\exp\left\{ -\frac{1}{4\Delta t}\, \mathbf{x_e}^{\mathrm{T}}\, \mathbf{G}(\mathbf{x_i}, t_i)^{-1} \mathbf{x_e} - \frac{L(\mathbf{x_i}, t_i)\, \Delta t}{\lambda} \right\} \end{aligned} \tag{5.21}$$

with $\mathbf{x_e} = \mathbf{x_{i+1}} - \mathbf{x_i} - \mathbf{f}(\mathbf{x_i}, t_i)\, \Delta t$.

5.3 The Generalized Transformed Path Integral Control

In the generalized transformed path integral control we consider a dynamical system of the form in Eq. (5.9) and cost functional of the form in Eq. (5.10). Our starting point is the stochastic process $\boldsymbol{\xi}(t)$. Evolution of the associated conditional distribution $\rho(\mathbf{x}, t \,|\, \mathbf{x_0}, t_0)$ is governed by Eq. (5.18). We consider a dynamic transformation of $\boldsymbol{\xi}(t)$ given by

$$\mathbf{z}(t) = \mathcal{Z}\Big(\boldsymbol{\xi}(t), t\Big) = \mathbf{R}^{-1}_{t|\mathbf{x_0}}\Big[\boldsymbol{\xi}(t) - \boldsymbol{\mu}_{t|\mathbf{x_0}}\Big] \tag{5.22}$$

where $\boldsymbol{\mu}_{t|\mathbf{x_0}}$ and $\boldsymbol{\Sigma}_{t|\mathbf{x_0}} \equiv \mathbf{R}_{t|\mathbf{x_0}}\mathbf{R}^{\mathrm{T}}_{t|\mathbf{x_0}}$ are respectively the conditional mean and conditional covariance of $\boldsymbol{\xi}(t)$ conditioned on the initial state of the system $\boldsymbol{\xi}(t_0) = \mathbf{x_0}$. In other words,

$$\boldsymbol{\mu}_{t|\mathbf{x_0}} = \int_{-\infty}^{\infty} \mathbf{x}\, \rho(\mathbf{x}, t \,|\, \mathbf{x_0}, t_0)\, \mathrm{d}\mathbf{x} \tag{5.23}$$

$$\boldsymbol{\Sigma}_{t|\mathbf{x_0}} = \int_{-\infty}^{\infty} \Big[\mathbf{x} - \boldsymbol{\mu}_{t|\mathbf{x_0}}\Big]\Big[\mathbf{x} - \boldsymbol{\mu}_{t|\mathbf{x_0}}\Big]^{\mathrm{T}} \rho(\mathbf{x}, t \,|\, \mathbf{x_0}, t_0)\, \mathrm{d}\mathbf{x}. \tag{5.24}$$

The conditional distribution of the original state variables are related to that of the transformed state variables as

$$\rho_{\boldsymbol{\xi}(t)}(\mathbf{x}, t \,|\mathbf{x_0}, t_0) = \Big\|\mathbf{R}^{-1}_{t|\mathbf{x_0}}\Big\|\; \rho_{\mathbf{z}(t)}\Big(\mathcal{Z}(\mathbf{x}, t), t \,\Big|\, \mathbf{z_0}, t_0\Big). \tag{5.25}$$

In further discussions, we will drop the subscripts in our representations of the distribution functions and rely on the reader to distinguish between them based on the context. Since $\rho(\mathbf{x}, t \,|\, \mathbf{x_0}, t_0) = \rho(\mathbf{x}, t \,|\, \mathbf{z_0}, t_0)$, we have $\Big[\boldsymbol{\mu}_{t|\mathbf{x_0}}, \mathbf{R}_{t|\mathbf{x_0}}, \boldsymbol{\Sigma}_{t|\mathbf{x_0}}\Big] = \Big[\boldsymbol{\mu}_{t|\mathbf{z_0}}, \mathbf{R}_{t|\mathbf{z_0}}, \boldsymbol{\Sigma}_{t|\mathbf{z_0}}\Big]$. The evolution equation for the conditional distribution of the transformed state variables $\rho(\mathbf{z}, t \,|\, \mathbf{z_0}, t_0)$ is given by

$$\frac{\partial \rho(\mathbf{z}, t \,|\, \mathbf{z_0}, t_0)}{\partial t} = \left[-\frac{\partial}{\partial z_i}\Phi_i(\mathbf{z}, t) + \frac{\partial^2}{\partial z_i \partial z_j}\Gamma_{ij}(\mathbf{z}, t) - \frac{\tilde{L}(\mathbf{z}, t)}{\lambda}\right] \rho(\mathbf{z}, t \,|\, \mathbf{z_0}, t_0). \tag{5.26}$$

The drift vector function $\mathbf{\Phi}$ and diffusion coefficient matrix $\mathbf{\Gamma}$ are given by

$$\mathbf{\Phi}(\mathbf{z},t) = \mathbf{R}_{t|\mathbf{z_0}}^{-1}\Big(\tilde{\mathbf{f}}(\mathbf{z},t) - \dot{\mathbf{R}}_{t|\mathbf{z_0}}\,\mathbf{z} - \dot{\boldsymbol{\mu}}_{t|\mathbf{z_0}}\Big) \tag{5.27}$$

$$\mathbf{\Gamma}(\mathbf{z},t) = \mathbf{R}_{t|\mathbf{z_0}}^{-1}\,\tilde{\mathbf{G}}(\mathbf{z},t)\,\mathbf{R}_{t|\mathbf{z_0}}^{-1\ \mathrm{T}}. \tag{5.28}$$

with $\tilde{\mathbf{f}}(\mathbf{z},t) = \mathbf{f}\Big(\mathcal{Z}^{-1}(\mathbf{z},t),t\Big)$, $\tilde{\mathbf{G}}(\mathbf{z},t) = \mathbf{G}\Big(\mathcal{Z}^{-1}(\mathbf{z},t),t\Big)$, and $\tilde{L}(\mathbf{z},t) = L\Big(\mathcal{Z}^{-1}(\mathbf{z},t),t\Big)$ where $\mathcal{Z}^{-1}$ is the inverse transformation of Eq. (5.22). Along with Eq. (5.26) for the evolution of the conditional distribution, we have update equations for the conditional mean and conditional covariance given by

$$\dot{\boldsymbol{\mu}}_{t|\mathbf{z_0}} = \Big\langle \tilde{\mathbf{f}}(\mathbf{z},t)\,\Big|\,\mathbf{z_0}\Big\rangle - \frac{1}{\lambda}\mathbf{R}_{t|\mathbf{z_0}}\Big\langle \mathbf{z}\tilde{L}(\mathbf{z},t)\,\Big|\,\mathbf{z_0}\Big\rangle, \tag{5.29}$$

$$\begin{aligned}\dot{\mathbf{\Sigma}}_{t|\mathbf{z_0}} = \mathbf{R}_{t|\mathbf{z_0}}\Big\langle \mathbf{z}\,\delta\tilde{\mathbf{f}}(\mathbf{z},t)^{\mathrm{T}}\Big\rangle + \Big\langle \delta\tilde{\mathbf{f}}(\mathbf{z},t)\,\mathbf{z}^{\mathrm{T}}\Big\rangle\,\mathbf{R}_{t|\mathbf{z_0}}^{\mathrm{T}} + 2\Big\langle\tilde{\mathbf{G}}(\mathbf{z},t)\Big\rangle \\ - \frac{1}{\lambda}\mathbf{R}_{t|\mathbf{z_0}}\Big\langle \mathbf{z}\mathbf{z}^{\mathrm{T}}\tilde{L}(\mathbf{z},t)\,\Big|\,\mathbf{z_0}\Big\rangle\mathbf{R}_{t|\mathbf{z_0}}^{\mathrm{T}}.\end{aligned} \tag{5.30}$$

The transformation in Eq. (5.22) ensures that the conditional mean $\Big\langle \mathbf{z}\,\Big|\,\mathbf{z_0}\Big\rangle$ and the conditional covariance $\Big\langle \mathbf{z}\mathbf{z}^{\mathrm{T}}\,\Big|\,\mathbf{z_0}\Big\rangle$ remain constant with propagation. These constants are zero and identity respectively. We consider the iterative path-integral based solution for Eq. (5.26) given by

$$\rho(\mathbf{z_{i+1}},t_i+\Delta t\,|\,\mathbf{z_0},t) = \int_{-\infty}^{\infty}\rho(\mathbf{z_{i+1}},t_i+\Delta t\,|\,\mathbf{z_i},t)\,\rho(\mathbf{z_i},t_i\,|\,\mathbf{z_0},t_0)\,\mathrm{d}\mathbf{z_0} \tag{5.31}$$

for $i = 1,\ldots,N$ with the short-time propagator

$$\rho(\mathbf{z_{i+1}},t_i+\Delta t\,|\,\mathbf{z_i},t_i) = \|4\pi\,\Delta t\,\mathbf{\Gamma}(\mathbf{z_i},t_i)\|^{-1/2}\exp\left\{-\frac{1}{4\Delta t}\mathbf{z_e}^{\mathrm{T}}\,\mathbf{\Gamma}(\mathbf{z},t)^{-1}\mathbf{z_e}\right\} \tag{5.32}$$

and $\mathbf{z_e} = \mathbf{z_{i+1}} - \mathbf{z_i} - \mathbf{\Phi}(\mathbf{z_i}, t_i)\,\Delta t$. Thus, we obtain the conditional distribution $\rho(\mathbf{z}', t' \,|\, \mathbf{z}, t)$. Using Eq. (5.25) and Eq. (5.19) we get

$$\psi(\mathbf{x}, t) = \int_{-\infty}^{\infty} \rho(\mathbf{z}', T \,|\, \mathbf{z}, t)\, \psi(\mathbf{x}', t)\, \mathrm{d}\mathbf{x}' \tag{5.33}$$

5.4 Performance of the Generalized Transformed Path Integral Control

The benefits of the generalized path integral control (GTPIC) over conventional fixed grid (FG) based approaches are illustrated in this section. The conventional FG based approaches to path integral control face limitations for their applicability to the control of stochastic dynamical systems with large drift, diffusion, and concentration of PDF. On the other hand, Monte-Carlo based implementations of path integral control encounter sampling errors which present challenges for the accurate representations of the underlying distributions and the set of optimal controls. The control of a pure diffusion process is first considered to showcase the benefits in a system where the dynamics are linear with respect to the state variables. Then the control of a nonlinear dynamical system with a cubic nonlinearity is presented.

Recall that in path integral control, the dynamical system is of the form

$$\mathrm{d}x = \Big[f(x,t) + bu\Big]\,\mathrm{d}t + a\mathrm{d}w \tag{5.34}$$

and the cost functional to be optimized is of the form

$$C[u] = \left\langle \phi(x) + \int \Big[L(x) + \frac{1}{2}Ru^2\Big]\mathrm{d}x \,\middle|\, x_0 \right\rangle \tag{5.35}$$

where $\phi(x)$ is the end cost and $L(x)$ is the state dependent path cost. In the case of controlled pure diffusion we consider $f(x,t) = 0$. Thus the system is completely driven by the random excitation and the control inputs. The strength of the white noise excitation is $a = 1$.

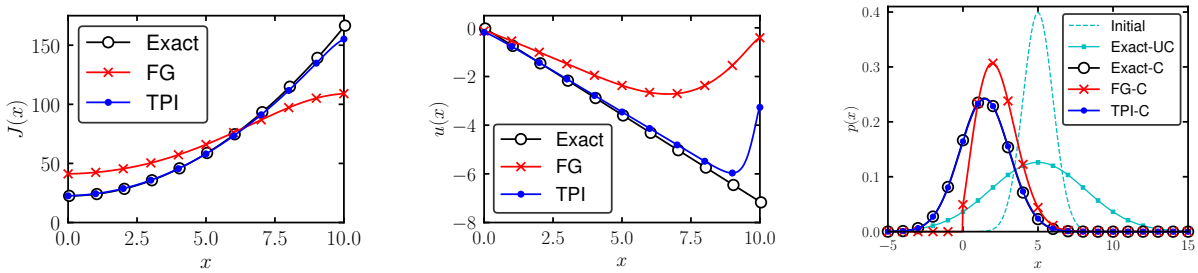

Figure 5.1: Plot of the optimal cost function (left), the optimal controls (center), and the evolution of the state distributions (right) obtained via the conventional fixed grid approach (FG), the generalized transformed path integral control (TPI), and the exact analytical solution for the controlled 1D pure diffusion process.

Additionally, we consider $b = 1$, $R = 1$ and a constant path cost $L(x) = 1$.

In this system, the covariance of the distribution grows linearly with time. Conventional fixed grid (FG) based approaches will not be able to accurately represent this behavior over large times. Consequently, large errors are encountered in the estimates of the optimal cost function and the optimal controls as seen in Fig. 5.1 (left and middle). The GTPIC estimates match the exact results much better than those from FG. Thus, the estimates of the controlled distribution obtained from the GTPIC are more accurate than those from FG as illustrated in Fig. 5.1 (right). An initial distribution of $\mathcal{N}(4, 1)$ was considered for the simulation. They were performed on a grid with 201 grid points between bounds $[-6, 6]$ in both the original space as well as the transformed computational domain. The simulations were performed from $t = 0\,\mathrm{s}$ to $t = 1\,\mathrm{s}$ in increments of $\Delta t = 0.01\,\mathrm{s}$.

Next, the control of a dynamical system with nonlinear dynamics $f(x) = x/4 - x^3$ is considered. The other system parameters are the same as those considered in the controlled pure diffusion case. The initial distribution was chosen to be $\mathcal{N}(1, 1)$. The simulation was performed on a grid with 201 grid points between bounds $[-4, 4]$ in the original space and bounds $[-5, 5]$ in the transformed space. It is worth pointing out that the conventional FG approach would be sufficient for the system parameters and simulation parameters considered for this case. Clearly, the estimates obtained from the GTPIC match those from FG for this case as seen in Fig. 5.2. The simulations were performed from $t = 0\,\mathrm{s}$ to $t = 1\,\mathrm{s}$ in increments

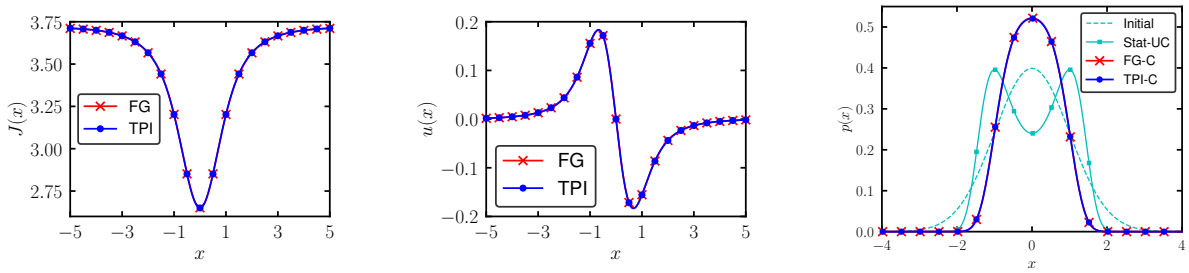

Figure 5.2: Plot of the optimal cost function (left), the optimal controls (center), and the evolution of the state distributions (right) obtained via the conventional fixed grid approach (FG) and the generalized transformed path integral control (TPI) for the control of a 1D nonlinear dynamical system with a cubic nonlinearity.

of $\Delta t = 0.01\,\mathrm{s}$.

5.5 Concluding Remarks

The generalized transformed path integral control (GTPIC), a path-integral based framework for the efficient control of a large class of stochastic dynamical systems is presented. These systems have linear dynamics and quadratic cost functions with respect to the control variables. The optimal controls for these systems maybe obtained through the solution of a backward Kolmogorov-type equation or equivalently through the solution of a Fokker-Planck-type equation. The Fokker-Planck type equation governs the evolution of distributions for a stochastic diffusion process that occurs alongside a growth/decay process.

As a part of the proposed framework, a novel short-time propagator for the evolution of the conditional distribution conditioned on the initial state of the system is presented. The framework allows for the evolution in a transformed space where a more accurate representation of the distribution can be ensured. As a consequence, more accurate estimates of the optimal controls than in conventional grid based approaches maybe obtained. The framework also does not face issues arising from sampling errors that are present in Monte-Carlo based approaches. In addition, necessary update equations for the conditional mean and covariance of the distribution are also presented. Error bounds for the evolved distributions,

and for the optimal controls, can be easily established using Chebyshev's inequality.

CHAPTER 6

Summary and Future Work

In this dissertation, novel path integral–based frameworks to obtain efficient solutions to problems in prediction, nonlinear filtering, and optimal control of stochastic dynamical systems have been presented. The foundation for these frameworks is the transformed path integral (TPI) method—a novel path integral–based solution of the Fokker-Planck equation (FPE). The equation governs the evolution of the state distribution for the underlying stochastic dynamical system. Many quantities of interest in the aforementioned problems, such as the a posteriori state estimate or expected cost, require accurate estimates of the state distribution. The TPI method allows us to obtain more accurate estimates of the state distribution compared to conventional fixed grid-based methods.

The basic TPI method is however not applicable to stochastic dynamical systems with a rank deficient (singular) diffusion coefficient matrix. The generalized transformed path integral (GTPI) method—a non-trivial generalization of the TPI method—extends the applicability to a larger class of systems. Specifically, it can be applied to systems with a singular diffusion coefficient matrix. As a consequence, the GTPI method is applicable to second order stochastic dynamical systems, systems with zero process noise, certain stochastic dynamical systems with non-white noise excitation, as well as systems with a full rank diffusion coefficient matrix (where we recover the TPI method). The framework of the GTPI method was also applied to the problem of nonlinear filtering and stochastic optimal control.

A transformed path integral-based framework for the continuous-discrete nonlinear filtering problem, termed the generalized transformed path integral filter (GTPIF) was presented. The GTPIF consists of a GTPI-like method for the system update (or prediction) and a novel

set of update equations for the conditional state distribution, state mean, and covariance in order to perform the Bayesian measurement update in the transformed computational domain. Likewise, a transformed path integral–based framework for a broad class of stochastic optimal control problems, termed the generalized transformed path integral control (GTPIC) was also presented. In the GTPIC, the corresponding stochastic Hamilton-Jacobi-Bellman equation is solved by solving an associated Fokker-Planck type equation using a GTPI-based method. The benefits of the proposed frameworks in their respective areas were shown through simulations of corresponding example problems in one-dimensional and multi-dimensional spaces. Ideas to address limitations of the proposed frameworks along with extensions to future work are also discussed.

6.1 The Transformed Path Integral Method

The theoretical formulations for the transformed path integral (TPI) method are presented in Chapter 2. As a part of the method, a new form of the short-time propagator was developed based on a dynamic transformation of the state space with the state mean and state covariance as parameters. The new propagator allows for the propagation of distributions to be performed in a transformed computational domain where a more accurate representation of distributions can be realized. In addition to the novel short-time propagator for PDF propagation, update equations for the state mean and state covariance were also derived.

Under the dynamic transformation in the TPI method a fixed grid in the transformed space corresponds to an adaptive grid in the original space that translates with the state mean and scales with the state covariance. Hence, straightforward error bounds based on Chebyshev's inequality were established for solutions obtained by the TPI method. A feature of the TPI method is that the propagated PDF in the transformed space always has a mean of zero and identity covariance. In fact, for linear dynamical systems with a Gaussian initial distribution, the distribution in the transformed space remains invariant with a standard normal distribution as shown in Lemma. (2.2.1).

Consequently, the TPI method is able to better address challenges arising from finite fixed grids, a static computational domain, and finite fixed grid resolution which may be encountered in processes with large diffusion coefficients, large drift vectors, or large concentrations of the PDF. The benefits of the TPI method and limitations of conventional methods such as the Monte Carlo simulations and fixed grid approaches, in tackling these challenges were clearly illustrated in the one-dimensional systems: (a) one-dimensional pure diffusion process, (b) one-dimensional constant drift process, (c) one-dimensional Ornstein-Uhlenbeck process, and (d) white noise driven bistable stochastic flow in one dimension. They were also illustrated in multi-dimensional systems: (e) multi-dimensional pure diffusion process, (g) multi-dimensional Ornstein-Uhlenbeck process, (h) multi-dimensional uncoupled nonlinear dynamical system, and (i) multi-dimensional coupled nonlinear dynamical system. Solutions obtained from the TPI method show much better agreement with analytical solutions where available than those obtained from conventional methods.

Analysis of the error in the PDF as compared with the analytical solutions in the linear case and the stationary solution in the nonlinear case show that the TPI method is more efficient than conventional methods in terms of the computational costs involved to obtain a desired level of accuracy, especially for problems in one dimensional space. The analysis also indicated the possibility of the existence of a similarity parameter. A candidate for the parameter, namely $\alpha = \sigma_{ref}^2 (\Delta z)^2/(2a^2\Delta t)$, yielded a partial collapse of the various error curves. It was shown that there is further reduction of the error for $\alpha > 10$ or $\alpha < 0.5$. A challenge for the conventional path integral methods and inherited by the TPI method is handling of stochastic processes with a rank deficient diffusion coefficient matrix. An extension to the TPI method is needed to extend the applicability to these processes.

6.2 The Generalized Transformed Path Integral Approach

The generalized transformed path integral method (GTPI), a generalization of the TPI method that extends the applicability to a larger class of systems (including those with

a rank deficient diffusion coefficient matrix) was presented in Chapter 3. This non-trivial generalization is based on the insight that the dynamic transformation employed in the transformed path integral method allows for an alternate description of the evolution of stochastic dynamical systems in terms of the evolution of a complementary system. We termed this system as the standard transformed stochastic dynamical system and we will refer to it here simply as the transformed system. A feature of this system is that its state mean and state covariance remain constant with evolution of the system. The constants are zero and the identity matrix respectively.

The evolution of the state distribution for the transformed system is governed by the corresponding (transformed) Fokker-Planck equation. In the GTPI method, appropriate short-time propagators that can handle challenges arising from singular diffusion coefficient matrices are developed using operator splitting of the transformed FPE. The splitting scheme consists of a Liouville operator and a Fokker-Planck operator; the latter has a full rank diffusion coefficient sub-matrix. Additionally, update equations for mean and covariance of the state variables, which are used for PDF propagation in the transformed space, are developed from the underlying stochastic system models. The new set of update equations allow for consideration of solutions to (i) second order dynamical systems, (ii) dynamical systems with zero process noise, (iii) certain dynamical systems with non-white noise excitation, and (iv) systems with a full rank diffusion coefficient matrix (where we recover the TPI method).

Unsurprisingly, the generalized transformed path integral method inherits several of the salient features of the TPI method, including the benefits in addressing challenges arising from large drift, diffusion, and concentration of PDF. It also inherits the feature allowing for straightforward establishment of error bounds based on Chebyshev's inequality. Finally, like in TPI, the transformed space distribution in the GTPI framework also remains invariant with a standard normal distribution for linear dynamical systems with a Gaussian initial condition Lemma. (3.3.1). In addition to the renormalization condition to preserve zeroth moment properties, conditions to preserve the first and second moment properties of the

transformed space state distribution were also established.

Several examples of linear and nonlinear stochastic dynamical systems with a singular diffusion coefficient matrix were considered to evaluate the performance of the generalized transformed path integral approach. These examples include (a) the stochastic harmonic oscillator, (b) the stochastic van der Pol oscillator, (c) the stochastic Caughey oscillator, (d) the stochastic Duffing oscillator with zero process noise, and (e) bistable stochastic flow driven by non-white noise. Of note is the stochastic Duffing oscillator with zero process noise (example (d)) which is governed by a Liouville equation. The results obtained from the GTPI method in these examples show good agreement with those from MC simulations and analytical solutions (where available). Comparisons of numerical error in the PDF show that the GTPI method is also more efficient than conventional fixed grid approaches in terms of the computational costs at achieving a desired level of accuracy.

The analysis of numerical error in the GTPI approach for the stochastic harmonic oscillator also indicated the presence of an optimal value ($\Delta t^\star$) for the time step below which the numerical simulations produce larger numerical error in the PDF. Both $\Delta t^\star$ and $\varepsilon^\star_{\text{RMS}}$, the error in the PDF at this time step, were lower for a finer grid resolution. The influence of the grid resolution and system parameters on $\Delta t^\star$ and $\varepsilon^\star_{\text{RMS}}$ for the stochastic harmonic oscillator was also investigated. Thus, the GTPI method addresses the domain issues associated with conventional fixd grid approaches and also extends the applicability to a larger class of systems (that includes processes with singular diffusion coefficient matrices). In spite of these benefits, issues relevant to the curse of dimensionality need to be investigated further. Nevertheless, the GTPI method presents an efficient framework for obtaining solutions to problems in stochastic dynamical systems. In particular, the framework can be applied to the nonlinear filtering problem as well as the stochastic optimal control problem.

6.3 The Generalized Transformed Path Integral Filter

The framework for the generalized transformed path integral filter (GTPIF) was presented in Chapter 4. The GTPIF is a path integral–based solution for the continuous-discrete nonlinear filtering problem. Recall that the objective of the nonlinear filtering problem is to obtain accurate estimates of the state of a stochastic dynamical system based on noisy measurements. In the continuous-discrete case, the dynamical system is represented by a continuous stochastic process and measurements are taken at discrete time instants. Optimal estimates for this system can be obtained via a two step process involving (A) the solution of a Fokker-Planck equation (i.e., the system update or prediction step) coupled with (B) the Bayesian update rule (i.e., the measurement update).

The GTPIF employs a dynamic transformation of the state space using the conditional state mean and conditional state covariance conditioned on the set of all prior measurements (i.e., filtration). Thereby, accurate estimates of the posterior conditional distribution conditioned on the filtration for the system update can be efficiently obtained by solving the corresponding transformed Fokker-Planck equation. Accordingly, a novel short-time propagator for the evolution of the conditional distribution in the transformed space was presented. Also, update equations for the conditional mean and covariance were derived for the prediction step. Likewise, new update equations for the measurement update step for the conditional distribution, state mean, and covariance were obtained based on the Bayes' update rule. The performance of the generalized path integral filter (GTPIF) was evaluated in a bistable stochastic flow excited by white noise that is observed through a linear measurement function with Gaussian white noise. The results obtained from the GTPIF were more accurate than those from the conventional extended Kalman filter.

6.4 The Generalized Transformed Path Integral Control

The framework for the generalized transformed path integral control (GTPIC) was presented in Chapter 5. The GTPIC is applicable to a large class of stochastic optimal control problems where the system dynamics are linear and the cost functional is quadratic in terms of the control variables. Recall that the objective of the stochastic optimal control problem is to obtain a set of control inputs to a stochastic dynamical system that optimizes a performance criterion (or minimizes a cost). The optimal controls may be obtained by solving the stochastic Hamilton-Jacobi-Bellman equation (HJB)—a nonlinear partial differential equation that governs the evolution of the value function (i.e., the cost-to-go function). In the class of problems considered for GTPIC, the stochastic HJB may be solved by solving an associated Fokker-Planck–type equation for a conditional distribution conditioned on the initial state of the system.

The GTPIC applies the GTPI framework to obtain the solution to the Fokker-Planck–type equation. It employs a dynamic transformation of the state space with the conditional mean and conditional covariance conditioned on the initial state of the system. A novel short-time propagator was presented that allows for the evolution of the conditional distribution in the transformed space generated by the transformation. Additionally, novel update equations for the conditional mean and conditional covariance of the original state variables necessary for the evolution were presented. The distribution in the transformed space coordinates will always have a zero mean and identity covariance. Thus, accurate and efficient estimates of the evolved conditional distribution and optimal controls maybe obtained with the GTPIC.

The performance of the generalized transformed path integral control was evaluated in a controlled pure diffusion process. The results obtained from the proposed approach showed good agreement with those from the analytical solution. Additionally, the GTPIC results were more accurate than the results from the conventional FG approach. The performance was also evaluated in the control of a bistable stochastic flow with white noise excitation.

The results from the GTPIC showed excellent agreement with expected behavior.

6.5 Future Work

The efficiency of the proposed frameworks can be increased further by using matrix free approaches for the propagation of the distributions. Since the numerical implementation of the propagation involves a matrix-vector multiplication, parallel computation algorithms can be employed to reduce computation times without reduction in the accuracy. Approaches based on fast multipole methods such as the fast Gauss transform (FGT) and symmetric fast Gauss transform (SFGT) can be incorporated to reduce computations times further for marginal reductions in the accuracy. Another approach involves developing the propagator for larger time steps as the product of the propagator for smaller time steps. These techniques can increase the applicability of the proposed approaches to problems in multi-dimensional systems with a large number of dimensions.

The framework can be applied to first passage problems where the governing equation is the backward Kolmogorov equation. Another area for future work is in Langrangian PDF methods for turbulent flows where stochastic models are used to describe evolution of the flow properties. The framework can also be extended to general stochastic control problems via new formulations in the transformed space based on the Bellman recursive relation. Likewise, the distributed stochastic optimal control of multiple agents is an area of active research where the benefits of the GTPI frameworks can be extended. The GTPI framework can also be extended to problems in population growth where the governing equation is the Fisher's equation (also known as the Kolmogorov-Petrovsky-Piskunov equation).

Bibliography

[1] H. Risken. *The Fokker-Planck Equation : Methods of Solution and Applications*. Springer Series in Synergetics. Springer-Verlag, Berlin Heidelberg New York, 2nd edition, 1989.

[2] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer-Verlag, Berlin Heidelberg New York, 6th edition, 2003.

[3] Andrew H. Jazwinski. *Stochastic Processes and Filtering Theory*. Mathematics in science and engineering. Academic Press, New York, 1970.

[4] Robert F Stengel. *Optimal control and estimation*. Courier Corporation, 1994.

[5] R. A. Ibrahim, P. O. Orono, and S. R. Madaboosi. Stochastic flutter of a panel subjected to random in-plane forces part 1: Two mode interaction. *AIAA Journal*, 28(4):694–702, 1990.

[6] J. B. Roberts and M. Vasta. Markov modelling and stochastic identification for nonlinear ship rolling in random waves. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 358(1771):1917–1941, 2000.

[7] M. Athans, R. Wishner, and A. Bertolini. Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements. *IEEE Transactions on Automatic Control*, 13(5):504–514, 1968.

[8] Neng Wan, Aditya Gahlawat, Naira Hovakimyan, Evangelos A Theodorou, and Petros G Voulgaris. Cooperative path integral control for stochastic multi-agent systems. In *2021 American Control Conference (ACC)*, pages 1262–1267. IEEE, 2021.

[9] R. E. Kalman and R. S. Bucy. New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, 83(1):95–108, 03 1961.

[10] Bhashyam Balaji. Universal nonlinear filtering using Feynman path integrals ii: the continuous-continuous model with additive noise. *PMC Physics A*, 3:1–28, 2009.

[11] Kazufumi Ito and Boris Rozovskii. Approximation of the kushner equation for nonlinear filtering. *SIAM Journal on Control and Optimization*, 38(3):893–23, 2000.

[12] Dan Crisan and Terry Lyons. A particle approximation of the solution of the kushner–stratonovitch equation. *Probability Theory and Related Fields*, 115:549–578, 1999.

[13] H. J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(11), 2005.

[14] Thomas K Caughey. Nonlinear theory of random vibrations. *Advances in applied mechanics*, 11:209–253, 1971.

[15] T. K. Caughey and J. K. Dienes. Analysis of a nonlinear first-order system with a white noise input. *Journal of Applied Physics*, 32(11):2476–2479, 1961.

[16] Crispin W Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer Series in Synergetics. Springer-Verlag, Berlin New York, 3rd edition, 2004.

[17] M. F. Dimentberg. An exact solution to a certain non-linear random vibration problem. *International Journal of Non-Linear Mechanics*, 17(4):231–236, 1982.

[18] M. Vasta. Exact stationary solution for a class of non-linear systems driven by a non-normal delta-correlated process. *International Journal of Non-Linear Mechanics*, 30(4):407–418, 1995.

[19] C. Proppe. Exact stationary probability density functions for non-linear systems under Poisson white noise excitation. *International Journal of Non-Linear Mechanics*, 38(4):557–564, 2003.

[20] G. Q. Cai and Y. K. Lin. Exact and approximate solutions for randomly excited MDOF non-linear systems. *International Journal of Non-Linear Mechanics*, 31(5):647–655, 1996.

[21] J. B. Roberts and P. D. Spanos. *Random vibration and statistical linearization*. Dover Publications, Mineola, NY, 2003.

[22] L Socha. Linearization in analysis of nonlinear stochastic systems: recent results-part i: theory. *Applied Mechanics Reviews*, 58(3):178–205, 2005.

[23] J. B. Roberts and P. D. Spanos. Stochastic averaging: an approximate method of solving random vibration problems. *International Journal of Non-Linear Mechanics*, 21(2):111–134, 1986.

[24] WQ Zhu. Recent developments and applications of the stochastic averaging method in random vibration. *Applied Mechanics Reviews*, 49(10S):S72–S80, 1996.

[25] J. D. Atkinson. Eigenfunction expansions for randomly excited non-linear systems. *Journal of Sound and Vibration*, 30(2):153–172, 1973.

[26] J. P. Johnson and R. A. Scott. Extension of eigenfunction-expansion solutions of a Fokker-Planck equation-ii. second order system. *International Journal of Non-Linear Mechanics*, 15(1):41–56, 1980.

[27] M. Yar and J. K. Hammond. Approximate eigenfunction analysis of first order nonlinear systems with application to a cubic system. *Journal of Sound and Vibration*, 111(3):457–466, 1986.

[28] K. Voigtlaender and H. Risken. Solutions of the Fokker-Planck equation for a double-well potential in terms of matrix continued fractions. *Journal of Statistical Physics*, 40(3-4):397–429, 1985.

[29] D. L. Ermak and H. Buckholz. Numerical-integration of the Langevin equation - Monte-Carlo simulation. *Journal of Computational Physics*, 35(2):169–182, 1980.

[30] C. Proppe, H. J. Pradlwarter, and G. I. Schueller. Equivalent linearization and Monte Carlo simulation in stochastic dynamics. *Probabilistic Engineering Mechanics*, 18(1):1–15, 2003.

[31] E. A. Johnson, S. F. Wojtkiewicz, L. A. Bergman, and B. F. Spencer. Observations with regard to massively parallel computation for Monte Carlo simulation of stochastic dynamical systems. *International Journal of Non-Linear Mechanics*, 32(4):721–734, 1997.

[32] L. Tierney. Markov chains for exploring posterior distributions. *Annals of Statistics*, 22(4):1701–1728, 1994.

[33] P. Kumar and S. Narayanan. Solution of Fokker-Planck equation by finite element and finite difference methods for nonlinear systems. *Sadhana*, 31(4):445–461, 2006.

[34] BF Spencer Jr and LA Bergman. On the numerical solution of the Fokker-Planck equation for nonlinear stochastic systems. *Nonlinear Dynamics*, 4(4):357–372, 1993.

[35] A. N. Drozdov and M. Morillo. Solution of nonlinear Fokker-Planck equations. *Physical Review E*, 54(1):931–937, 1996.

[36] Gray W Harrison. Numerical solution of the Fokker Planck equation using moving finite elements. *Numerical methods for Partial differential Equations*, 4(3):219–232, 1988.

[37] S. L. Cotter, T. Vejchodsky, and R. Erban. Adaptive finite element method assisted by stochastic simulation of chemical systems. *SIAM Journal on Scientific Computing*, 35(1):B107–B131, 2013.

[38] P. J. Attar and P. Vedula. Direct quadrature method of moments solution of the Fokker-Planck equation. *Journal of Sound and Vibration*, 317(1-2):265–272, 2008.

[39] Y. J. Xu and P. Vedula. A quadrature-based method of moments for nonlinear filtering. *Automatica*, 45(5):1291–1298, 2009.

[40] N. Wiener. Generalized harmonic analysis. *Acta Mathematica*, 55(1):117–258, 1930.

[41] R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Reviews of Modern Physics*, 20(2):367–387, 1948.

[42] Hagen Kleinert. *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*. World Scientific, Hackensack, N.J., 5th edition, 2010.

[43] H. Haken. Generalized Onsager-Machlup function and classes of path integral solutions of Fokker-Planck equation and master equation. *Zeitschrift für Physik B Condensed Matter*, 24(3):321–326, 1976.

[44] L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Physical Review*, 91(6):1505–1512, 1953.

[45] R. Graham. Path integral formulation of general diffusion processes. *Zeitschrift für Physik B Condensed Matter*, 26(3):281–290, 1977.

[46] C. Wissel. Manifolds of equivalent path integral solutions of the Fokker-Planck equation. *Zeitschrift für Physik B Condensed Matter*, 35(2):185–191, 1979.

[47] M. F. Wehner and W. G. Wolfer. Numerical evaluation of path-integral solutions to Fokker-Planck equations. *Physical Review A*, 27(5):2663–2670, 1983.

[48] Arvid Naess and John M. Johnsen. Response statistics of nonlinear dynamic systems by path integration. In *Nonlinear Stochastic Mechanics*, pages 401–414. Springer, 1992.

[49] A. Naess and J.M. Johnsen. Response statistics of nonlinear, compliant offshore structures by the path integral solution method. *Probabilistic Engineering Mechanics*, 8(2):91–106, 1993.

[50] A. Naess and V. Moe. Efficient path integration methods for nonlinear dynamic systems. *Probabilistic Engineering Mechanics*, 15(2):221–231, 2000.

[51] J. S. Yu, G. Q. Cai, and Y. K. Lin. A new path integration procedure based on Gauss-Legendre scheme. *International Journal of Non-Linear Mechanics*, 32(4):759–768, 1997.

[52] J. S. Yu and Y. K. Lin. Numerical path integration of a non-homogeneous Markov process. *International Journal of Non-Linear Mechanics*, 39(9):1493–1500, 2004.

[53] I. A. Kougioumtzoglou and P. D. Spanos. An analytical Wiener path integral technique for non-stationary response determination of nonlinear oscillators. *Probabilistic Engineering Mechanics*, 28:125–131, 2012.

[54] I. A. Kougioumtzoglou and P. D. Spanos. Nonstationary stochastic response determination of nonlinear systems: A Wiener path integral formalism. *Journal of Engineering Mechanics*, 140(9), 2014.

[55] I. A. Kougioumtzoglou, A. Di Matteo, P. D. Spanos, A. Pirrotta, and M. Di Paola. An efficient Wiener path integral technique formulation for stochastic response determination of nonlinear mdof systems. *Journal of Applied Mechanics*, 82(10), 2015.

[56] Simon J Julier and Jeffrey K Uhlmann. New extension of the Kalman filter to nonlinear systems. In *AeroSense'97*, pages 182–193. International Society for Optics and Photonics, 1997.

[57] Yu Kweng Lin and Guo Qiang Cai. *Probabilistic structural dynamics: advanced theory and applications*. McGraw-Hill New York, 1995.

[58] Mircea Grigoriu. *Stochastic calculus: applications in science and engineering*. Springer Science & Business Media, 2013.

[59] Peter Hänggi. Path integral solutions for non-markovian processes. *Zeitschrift für Physik B Condensed Matter*, 75(2):275–281, 1989.

[60] Peter Häunggi and Peter Jung. Colored noise in dynamical systems. *Advances in chemical physics*, 89:239–326, 1994.

[61] SJB Einchcomb and AJ McKane. Use of Hamiltonian mechanics in systems driven by colored noise. *Physical Review E*, 51(4):2974, 1995.

[62] Wei Chai, Arvid Naess, and Bernt J Leira. Filter models for prediction of stochastic ship roll response. *Probabilistic Engineering Mechanics*, 41:104–114, 2015.

[63] Apostolos F Psaros, Olga Brudastova, Giovanni Malara, and Ioannis A Kougioumtzoglou. Wiener path integral based response determination of nonlinear systems subject to non-white, non-Gaussian, and non-stationary stochastic excitation. *Journal of Sound and Vibration*, 433:314–333, 2018.

[64] Alexander N Drozdov. High-accuracy discrete path integral solutions for stochastic processes with noninvertible diffusion matrices. *Physical Review E*, 55(3):2496, 1997.

[65] Alexander N Drozdov. High-accuracy discrete path integral solutions for stochastic processes with noninvertible diffusion matrices. II. Numerical evaluation. *The Journal of chemical physics*, 107(9):3505–3520, 1997.

[66] Horacio S Wio. On the solution of Kramers' equation by Trotter's formula. *The Journal of Chemical Physics*, 88(8):5251–5252, 1988.

[67] Alinda Mashiku, James Garrison, and J. Russell Carpenter. *Statistical Orbit Determination using the Particle Filter for incorporating Non-Gaussian Uncertainties.* 2012.

[68] I. Kaminer, Wei Kang, O. Yakimenko, and A. Pascoal. Application of nonlinear filtering to navigation system design using passive sensors. *IEEE Transactions on Aerospace and Electronic Systems*, 37(1):158–172, 2001.

[69] Peter Jan van Leeuwen. *Nonlinear Data Assimilation for high-dimensional systems*, pages 1–73. Springer International Publishing, Cham, 2015.

[70] R. L. Stratonovich. Conditional markov processes. *Theory of Probability & Its Applications*, 5(2):156–178, 1960.

[71] Harold J. Kushner. On the differential equations satisfied by conditional probablitity densities of markov processes, with applications. *Journal of the Society for Industrial and Applied Mathematics Series A Control*, 2(1):106–119, 1964.

[72] Moshe Zakai. On the optimal filtering of diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 11(3):230–243, 1969.

[73] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960.

[74] Maria Isabel Ribeiro. Kalman and extended Kalman filters: Concept, derivation and properties. *Institute for Systems and Robotics*, 43(46):3736–3741, 2004.

[75] Simon J Julier. The scaled unscented transformation. In *Proceedings of the 2002 American Control Conference*, volume 6, pages 4555–4559. IEEE, 2002.

[76] M Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on signal processing*, 50(2):174–188, 2002.

[77] Fred Daum, Jim Huang, and Arjang Noushin. Exact particle flow for nonlinear filters. In *Signal processing, sensor fusion, and target recognition XIX*, volume 7697, pages 92–110. SPIE, 2010.

[78] S.J. Julier and J.K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

[79] S. Challa and Y. Bar-Shalom. Nonlinear filter design using Fokker-Planck-Kolmogorov probability density evolutions. *IEEE Transactions on Aerospace and Electronic Systems*, 36(1):309–315, 2000.

[80] Vicente Rico-Ramirez, Urmila M. Diwekar, and Benoit Morel. Real option theory from finance to batch distillation. *Computers & Chemical Engineering*, 27(12):1867–1882, 2003.

[81] Naoya Ozaki, Stefano Campagnola, and Ryu Funase. Tube stochastic optimal control for nonlinear constrained trajectory optimization problems. *Journal of Guidance, Control, and Dynamics*, 43(4):645–655, 2020.

[82] Evangelos Theodorou, Jonas Buchli, and Stefan Schaal. A generalized path integral control approach to reinforcement learning. *The Journal of Machine Learning Research*, 11:3137–3181, 2010.

[83] Hilbert J Kappen. An introduction to stochastic control theory, path integrals and reinforcement learning. In *AIP conference proceedings*, volume 887, pages 149–181. American Institute of Physics, 2007.

[84] RL Stratonovich. On the probability functional of diffusion processes. *Selected Trans. in Math. Stat. Prob*, 10:273–286, 1971.

[85] H Dekker. Functional integration and the Onsager-Machlup Lagrangian for continuous Markov processes in Riemannian geometries. *Physical Review A*, 19(5):2102, 1979.

[86] Hidemi Ito. Probabilistic construction of lagrangean of diffusion process and its application. *Progress of Theoretical Physics*, 59(3):725–741, 1978.

[87] F Langouche, D Roekaerts, and E Tirapegui. Functional integrals and the Fokker-Planck equation. *Il Nuovo Cimento B (1971-1996)*, 53(1):135–159, 1979.

[88] JA Barker. A quantum-statistical Monte Carlo method; path integrals with boundary conditions. *The Journal of Chemical Physics*, 70(6):2914–2918, 1979.

[89] David M Ceperley. Path integrals in the theory of condensed helium. *Reviews of Modern Physics*, 67(2):279, 1995.

[90] H. J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(11), 2005.

[91] MF Wehner and WG Wolfer. Numerical evaluation of path-integral solutions to fokker-planck equations. ii. restricted stochastic processes. *Physical Review A*, 28(5):3003, 1983.

[92] MF Wehner and WG Wolfer. Numerical evaluation of path-integral solutions to Fokker-Planck equations. iii. time and functionally dependent coefficients. *Physical Review A*, 35(4):1795, 1987.

[93] Gnana M Subramaniam and Prakash Vedula. A transformed path integral approach for solution of the Fokker–Planck equation. *Journal of Computational Physics*, 346:49–70, 2017.

[94] Hale F Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, 10(4):545–551, 1959.

[95] Alexander N Drozdov and Peter Talkner. Path integrals for Fokker–Planck dynamics with singular diffusion: Accurate factorization for the time evolution operator. *The Journal of chemical physics*, 109(6):2080–2091, 1998.

[96] Mani Razi, Peter J Attar, and Prakash Vedula. Adaptive finite difference solutions of Liouville equations in computational uncertainty quantification. *Reliability Engineering & System Safety*, 142:267–278, 2015.

[97] Tsu T Soong. *Random differential equations in science and engineering*. Mathematics in Science and Engineering. Academic Press, New York, 1973.

[98] Julyan HE Cartwright, Víctor M Eguíluz, Emilio Hernández-García, and Oreste Piro. Dynamics of elastic excitable media. *International Journal of Bifurcation and Chaos*, 9(11):2197–2202, 1999.

[99] Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal*, 1(6):445–466, 1961.

[100] Jinichi Nagumo, Suguru Arimoto, and Shuji Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.

[101] Ivana Kovacic and Michael J Brennan. *The Duffing equation: nonlinear oscillators and their behaviour*. John Wiley & Sons, 2011.

[102] Rolf Landauer. Fluctuations in bistable tunnel diode circuits. *Journal of Applied Physics*, 33(7):2209–2216, 1962.

[103] Klaus Schulten, Zan Schulten, and Attila Szabo. Dynamics of reactions involving diffusive barrier crossing. *The Journal of Chemical Physics*, 74(8):4426–4432, 1981.

[104] Benny Carmeli and Abraham Nitzan. Non-Markovian theory of activated rate processes. III. bridging between the Kramers limits. *Physical Review A*, 29(3):1481, 1984.

APPENDIX A

TPI formulation from Itô's lemma

In this appendix, an alternate formulation for the TPI method based on Itô's lemma is presented. We will show later that the two forms are equivalent the short time limit. We consider the 1D stochastic process given in Eq. (2.7) and the time dependent transformation in Eq. (2.9). We know from Itô's lemma [2] that z_t is also an Itô process given by

$$\mathrm{d}z_t = \left(\frac{\partial \mathcal{Z}}{\partial t} + f\,\frac{\partial \mathcal{Z}}{\partial x} + \frac{1}{2} g^2\,\frac{\partial^2 \mathcal{Z}}{\partial x^2} \right) \mathrm{d}t + g\,\frac{\partial \mathcal{Z}}{\partial x}\,\mathrm{d}\beta_t$$

which, for the particular transformation considered, reduces to

$$\mathrm{d}z_t = \frac{1}{\sigma_t}\left[-\frac{\mathrm{d}\mu_t}{\mathrm{d}t} - z_t\,\frac{\mathrm{d}\sigma_t}{\mathrm{d}t} + f(x_t,t) \right] \mathrm{d}t + \left[\frac{g(x_t,t)}{\sigma_t} \right] \mathrm{d}\beta_t.$$

Multiplying throughout by σ_t and with some rearranging we have

$$\sigma_t\,\mathrm{d}z_t + z_t\,\mathrm{d}\sigma_t = \delta\tilde{f}(z_t,t)\,\mathrm{d}t + \tilde{g}(z_t,t)\,\mathrm{d}\beta_t \tag{A.1}$$

Thus we have the corresponding form of the short time propagator in the transformed space

$$p(z,t|z',t') = \frac{\sigma'}{\sqrt{2\,\pi\,\tilde{g}^2(z',t')\mathrm{d}t}} \times \exp\left\{ -\frac{\left[\sigma'\,(z-z') + z'\,(\sigma-\sigma') - \delta\tilde{f}(z',t')\,\mathrm{d}t\right]^2}{2\,\tilde{g}^2(z',t')\,\mathrm{d}t} \right\}. \tag{A.2}$$

The form presented above is equivalent to the form proposed in Eq. (2.12) up to an order of $\mathcal{O}(\mathrm{d}t^2)$. This can be seen by rewriting Eq. (A.1) as below

$$\begin{aligned} z_{t+\mathrm{d}t} &= z_t \left(1 - \frac{\mathrm{d}\sigma_t}{\sigma_t}\right) + \frac{\delta \tilde{f}(z_t, t)}{\sigma_t}\mathrm{d}t + \frac{\tilde{g}(z_t, t)}{\sigma_t}\mathrm{d}\beta_t \\ &= z_t \left[2 - \frac{1}{\left(1 - \frac{\mathrm{d}\sigma_t}{\sigma_{t+\mathrm{d}t}}\right)}\right] + \frac{\delta \tilde{f}(z_t, t)\,\mathrm{d}t}{\sigma_{t+\mathrm{d}t}\left(1 - \frac{\mathrm{d}\sigma_t}{\sigma_{t+\mathrm{d}t}}\right)} + \frac{\tilde{g}(z_t, t)\,\mathrm{d}\beta_t}{\sigma_{t+\mathrm{d}t}\left(1 - \frac{\mathrm{d}\sigma_t}{\sigma_{t+\mathrm{d}t}}\right)} \end{aligned}$$

Making use of the expansion $1/(1-x) = 1 + x + x^2 + x^3 + \ldots$ and after simplification we have

$$\begin{aligned} z_{t+\mathrm{d}t} = z_t + \left(\frac{\sigma_t}{\sigma_{t+\mathrm{d}t}} - 1\right) z_t + \frac{\delta \tilde{f}(z_t, t)}{\sigma_{t+\mathrm{d}t}}\mathrm{d}t + \frac{\tilde{g}(z_t, t)}{\sigma_{t+\mathrm{d}t}}\mathrm{d}\beta_t \\ + \mathcal{O}\left(\left(\frac{\dot{\sigma}_t}{\sigma_{t+\mathrm{d}t}}\right)^2 (\mathrm{d}t)^2\right) + \mathcal{O}\left(\left(\frac{\dot{\sigma}_t}{\sigma^2_{t+\mathrm{d}t}}\right)(\mathrm{d}t)^2\right) + \mathcal{O}\left(\left(\frac{\dot{\sigma}_t}{\sigma^2_{t+\mathrm{d}t}}\right)(\mathrm{d}t\,\mathrm{d}\beta_t)\right) \end{aligned} \tag{A.3}$$

For small time steps the contributions from the higher order terms of $\mathrm{d}t$ are negligible. Thus we see that the two forms of the transformed short time propagator Eq. (A.2) and Eq. (2.12) are equivalent in the limit $\mathrm{d}t \to 0$.

APPENDIX B

Liouville Equation in Transformed Space Coordinates

Let us consider the Liouville equation given by

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial x_i} f_i(\mathbf{x}, t)\right] p(\mathbf{x}, t) = 0. \tag{B.1}$$

We seek to express this equation in a new set of coordinates related to the old coordinates via the transformation $\mathbf{z} = \mathcal{Z}(\mathbf{x}, t)$. Let this transformation be invertible and the inverse transformation be given by $\mathbf{x} = \mathcal{Z}^{-1}(\mathbf{z}, t)$. The probability densities with respect to the old and new coordinates, i.e., $p_{\mathbf{x}}$ and $p_{\mathbf{z}}$ are related by

$$p_{\mathbf{z}}(\mathbf{z}, t) = \|\mathbf{J}\| \; p_{\mathbf{x}}(\mathbf{x}, t) \tag{B.2}$$

where $\mathbf{J} = \partial\mathbf{x}/\partial\mathbf{z} \equiv \left[\partial x_i/\partial z_j\right]$ for $i, j = 1, \ldots, N_s$ is the Jacobian matrix and $\|\mathbf{J}\|$ is its determinant. Consequently, $\mathbf{J}^{-1} = \partial\mathbf{z}/\partial\mathbf{x} \equiv \left[\partial z_i/\partial x_j\right]$. The spatial derivatives in the old coordinates can be expressed in the new coordinates as [1]

$$\frac{\partial}{\partial x_i} = \frac{1}{\|\mathbf{J}\|} \frac{\partial}{\partial z_k} \frac{\partial z_k}{\partial x_i} \; \|\mathbf{J}\| \tag{B.3}$$

Similarly, the expression for the time derivative in the new coordinates is given by

$$\left(\frac{\partial}{\partial t}\right)_{\mathbf{x}} = \frac{1}{\|\mathbf{J}\|} \left(\frac{\partial}{\partial t}\right)_{\mathbf{z}} \|\mathbf{J}\| + \frac{1}{\|\mathbf{J}\|} \frac{\partial}{\partial z_k} \left(\frac{\partial z_k}{\partial t}\right)_{\mathbf{x}} \|\mathbf{J}\| \tag{B.4}$$

where $\left(\partial/\partial t\right)_{\mathbf{x}}$ denotes derivative with respect to t while keeping $\mathbf{x}$ constant and $\left(\partial/\partial t\right)_{\mathbf{z}}$ denotes the derivative while keeping $\mathbf{z}$ constant. We may drop these subscripts for brevity

where the context is clear. Thus, substituting Eqs. (B.3) and (B.4) in Eq. (B.1) while noting the relationship in Eq. (B.2) we get

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial z_k}\Phi_k(\mathbf{z}, t)\right] p(\mathbf{z}, t) = 0 \tag{B.5}$$

where

$$\Phi_k(\mathbf{z}, t) = \frac{\partial z_k}{\partial t} + \frac{\partial z_k}{\partial x_i}\tilde{f}_i(\mathbf{z}, t) \tag{B.6}$$

$$\tilde{f}_i(\mathbf{z}, t) = f_i\Big(\mathcal{Z}^{-1}(\mathbf{z}, t), t\Big) \tag{B.7}$$

For the transformation considered in Eq. (3.13) we have

$$\mathbf{\Phi}(\mathbf{z}, t) = \mathbf{R}(t)^{-1}\Big[\tilde{f}(\mathbf{z}, t) - \dot{\mathbf{R}}(t)\,\mathbf{z} - \dot{\boldsymbol{\mu}}(t)\Big] \tag{B.8}$$

$$\tilde{f}(\mathbf{z}, t) = \mathbf{f}\Big(\mathbf{R}(t)\,\mathbf{z} + \boldsymbol{\mu}(t)\Big). \tag{B.9}$$

Fokker-Planck Equation in Transformed Space Coordinates

Applying Eq. (B.3) twice we obtain an expression for the second derivative in terms of the new coordinates:

$$\frac{\partial^2}{\partial x_i \partial x_j} = \frac{1}{\|\mathbf{J}\|}\frac{\partial^2}{\partial z_k \partial z_r}\frac{\partial z_k}{\partial x_i}\frac{\partial z_r}{\partial x_j}\|\mathbf{J}\| - \frac{1}{\|\mathbf{J}\|}\frac{\partial}{\partial z_k}\frac{\partial^2 z_k}{\partial x_i \partial x_j}\|\mathbf{J}\|. \tag{B.10}$$

Substituting Eqs. (B.3) and (B.10) in the Fokker-Planck equation Eq. (3.2) we obtain

$$\left[\frac{\partial}{\partial t} + \frac{\partial}{\partial z_k}\Phi_k(\mathbf{z}, t) - \frac{\partial^2}{\partial z_k \partial z_r}\Gamma_{kr}(\mathbf{z}, t)\right] p(\mathbf{z}, t) = 0 \tag{B.11}$$

with $\Phi_k = \left(\partial z_k/\partial t\right) + \left(\partial z_k/\partial x_i\right) \tilde{f}_i(\mathbf{z},t) + \left(\partial^2 z_k/\partial x_i \partial x_j\right) \tilde{G}_{ij}(\mathbf{z},t), \Gamma_{kr} = \left(\partial z_k/\partial x_i\right)\left(\partial z_r/\partial x_j\right) \tilde{G}_{ij}(\mathbf{z},t)$

Note that for the transformation considered in Eq. (3.13), we obtain

$$\mathbf{\Phi}(\mathbf{z},t) = \mathbf{R}(t)^{-1}\left[\tilde{f}(\mathbf{z},t) - \dot{\mathbf{R}}(t)\,\mathbf{z} - \dot{\boldsymbol{\mu}}(t)\right] \tag{B.12}$$

$$\mathbf{\Gamma}(\mathbf{z},t) = \mathbf{R}(t)^{-1}\,\tilde{\mathbf{G}}(\mathbf{z},t)\,\mathbf{R}(t)^{-1^{\mathrm{T}}}. \tag{B.13}$$

APPENDIX C

Derivation of update equations for mean and covariance

In this appendix, we detail the steps involved in the derivation of the update equations for mean and covariance presented in Eqs. (3.25) and (3.26). We want to obtain relations such that the first and second moments of the state variables in the transformed space are preserved with propagation under our proposed approach. By definition, the transformation in Eq. (3.13) ensures that state variables in the transformed space are initially distributed with zero mean and identity covariance before propagation. Thus, our relations should ensure that the state variables after propagation are also distributed with zero mean and identity covariance, i.e.,

$$\langle \mathbf{z} \rangle = \langle \mathbf{z}' \rangle = \mathbf{0} \tag{C.1}$$

$$\left\langle \mathbf{z}\mathbf{z}^{\mathrm{T}} \right\rangle = \left\langle \mathbf{z}'\mathbf{z}'^{\mathrm{T}} \right\rangle = \mathbb{I} \tag{C.2}$$

where the primed variables denote the quantities after propagation. Note that in our approach, we consider a partition $\mathbf{z}' \equiv \begin{bmatrix} \mathbf{q}' & \mathbf{v}' \end{bmatrix}^{\mathrm{T}}$ of the state space in the transformed domain into "singular" and "nonsingular" variables as detailed in Eq. (3.35). Also, since $\mathbf{q}' = \mathbf{q} + \Delta t\, \mathbf{\Phi}^{(\varepsilon)}(\mathbf{q}, \mathbf{v}, t)$ we have for any real-valued, compactly supported, continuous function f with support contained in $\mathbf{q}'(\mathbf{q})$

$$\int f(\mathbf{q}')\, \mathrm{d}\mathbf{q}' = \int f\big(\mathbf{q}'(\mathbf{q})\big) \left\| \mathbb{I} + \Delta t\, \mathrm{D}\mathbf{\Phi}^{(\varepsilon)} \right\| \mathrm{d}\mathbf{q}$$

where $\mathrm{D}\mathbf{\Phi}^{(\varepsilon)} \equiv \left[\partial \Phi_i^{(\varepsilon)} / \partial q_j\right]$. For small time steps, i.e., $\Delta t \to 0$

$$\int f(\mathbf{q}')\,\mathrm{d}\mathbf{q}' \approx \int f\Big(\mathbf{q}'(\mathbf{q})\Big)\,\exp\left\{\Delta t\,\frac{\partial \Phi_i^{(\varepsilon)}}{\partial u_i}\bigg|_{\mathbf{q},\mathbf{v},t}\right\}\mathrm{d}\mathbf{q}$$

Thus, using Eqs. (3.41) and (3.42) we have

$$\int f(\mathbf{q}')\,p(\mathbf{q}',\mathbf{v}',t+\Delta t)\,\mathrm{d}\mathbf{q}' \approx \int\int f\Big(\mathbf{q}'(\mathbf{q})\Big)\,\rho(\mathbf{v}',t+\Delta t|\mathbf{v},t)\,p(\mathbf{q},\mathbf{v},t)\,\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{v} \tag{C.3}$$

where $\rho(\mathbf{v}',t+\Delta t|\mathbf{v},t)$ is given by Eq. (3.43). Let us consider

$$\langle \mathbf{q}' \rangle = \int\int \mathbf{q}'\,p(\mathbf{q}',\mathbf{v}',t+\Delta t)\,\mathrm{d}\mathbf{q}'\,\mathrm{d}\mathbf{v}'$$
$$\langle \mathbf{v}' \rangle = \int\int \mathbf{v}'\,p(\mathbf{q}',\mathbf{v}',t+\Delta t)\,\mathrm{d}\mathbf{q}'\,\mathrm{d}\mathbf{v}'$$

Applying Eq. (C.3) and changing the order of integration we have

$$\langle \mathbf{q}' \rangle = \langle \mathbf{q} \rangle + \Delta t \left\langle \mathbf{\Phi}^{(\varepsilon)}(\mathbf{q},\mathbf{v},t) \right\rangle \tag{C.4}$$
$$\langle \mathbf{v}' \rangle = \langle \mathbf{v} \rangle + \Delta t \left\langle \mathbf{\Phi}^{(r)}(\mathbf{q},\mathbf{v},t) \right\rangle \tag{C.5}$$

Combining Eqs. (C.4) and (C.5) we have

$$\langle \mathbf{z}' \rangle = \langle \mathbf{z} \rangle + \Delta t \left\langle \mathbf{\Phi}(\mathbf{z},t) \right\rangle \tag{C.6}$$

Applying Eq. (C.1), we obtain

$$\left\langle \mathbf{\Phi}(\mathbf{z},t) \right\rangle = 0 \tag{C.7}$$

Proceeding similarly and dropping terms of $\mathcal{O}\left((\Delta t)^2\right)$ we can show that

$$\left\langle \mathbf{z}'\mathbf{z}'^{\mathrm{T}} \right\rangle = \left\langle \mathbf{z}\,\mathbf{z}^{\mathrm{T}} \right\rangle + \Delta t \left\langle \mathbf{z}\,\mathbf{\Phi}(\mathbf{z},t)^{\mathrm{T}} \right\rangle + \Delta t \left\langle \mathbf{\Phi}(\mathbf{z},t)\,\mathbf{z}^{\mathrm{T}} \right\rangle + 2\,\Delta t\,\langle \mathbf{\Gamma}(\mathbf{z},t) \rangle \tag{C.8}$$

Applying Eq. (C.2), we obtain

$$\Big\langle \mathbf{z}\, \boldsymbol{\Phi}(\mathbf{z},t)^{\mathrm{T}} \Big\rangle + \Big\langle \boldsymbol{\Phi}(\mathbf{z},t)\, \mathbf{z}^{\mathrm{T}} \Big\rangle + 2\Big\langle \boldsymbol{\Gamma}(\mathbf{z},t) \Big\rangle = 0 \tag{C.9}$$

The relations in Eqs. (C.7) and (C.9) represent the conditions to preserve the state mean and covariance with propagation using our proposed approach. Substituting Eqs. (3.18) and (3.19) in these conditions we can simplify them to Eqs. (3.25) and (3.26).