This is to certify that the

dissertation entitled

Stochastic Differential Equations
and their Numerical Approximations

presented by

Liying Huang

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Applied Mathematics

David H. Y. Yen
Major professor

Date July 28, 1995

Stochastic Differential Equations and their Numerical Approximations

By

Liying Huang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1995

ABSTRACT

Stochastic Differential Equations and their Numerical Approximations

By

Liying Huang

This thesis is on numerical methods for Fokker-Planck equations, especially those equations with degenerate diffusion coefficients. Emphasis is focused on the two-dimensional case which corresponds to second order stochastic differential equations.

First, Gauss-Galerkin/finite difference methods are proposed. To solve two dimensional Fokker-Planck equations which involve two *spatial* variables and one time variable, the idea of variable splitting is adopted. More specifically, finite difference methods are utilized in one of the spatial variables and Gauss-Galerkin methods are employed in the other. As a consequence, the Gauss-Galerkin approach is used only in one dimension at each time step.

After two-dimensional Fokker-Planck equations are discretized to semi-discrete equations by finite difference methods in one direction, the convergence of the difference approximation is established based on energy type estimates. Using theory of measures and moments, the Gauss-Galerkin approximation for the semi-discrete equations is shown to converge when one-dimensional Gauss-Galerkin approximation is used in the second direction. Combining the above results, the convergence of the Gauss-Galerkin/finite difference approximation is established.

Computer implementation and intensive numerical tests of Gauss-Galerkin/finite difference methods are then carried out. The methods appear very efficient and very accurate.

To the loving memory of my mother

ACKNOWLEDGMENTS

I would like to thank Professor *David H.Y. Yen*, my dissertation advisor, for his constant encouragement and support during my graduate study at Michigan State University. I would also like to thank him for suggesting the problem and helpful directions.

I would like to thank my other dissertation committee members: Professor *Qiang Du*, Professor *T.Y. Li*, Professor *Habib Salehi*, and Professor *ZhengFang Zhou* for their valuable suggestions and time.

Contents

List of Tables

List of Figures

1 Introduction

This thesis is on numerical methods for Fokker-Planck equations, especially those equations with degenerate diffusion coefficients.

It has three major parts:

1). Study of various settings in which Fokker-Planck equations are formulated;

2). Derivation and convergence analysis of Gauss-Galerkin/finite difference methods for two-dimensional Fokker-Planck equations;

3). Numerical implementation of Gauss-Galerkin/finite difference methods and the application of these algorithms to problems in nonlinear random oscillations governed by second order stochastic differential equations.

The first part of this thesis provides a study of Fokker-Planck equations corresponding to second order differential equations excited by white noise and various solution techniques such as the iterative methods. Fokker-Planck equations, which are defined in unbounded domains with unbounded coefficients, are reduced to equations defined in bounded domains with bounded coefficients at finite time. This reduction lays the groundwork for efficient numerical approximations.

In the second part, Gauss-Galerkin/finite difference methods for solving two-dimensional Fokker-Planck equations numerically are proposed. These generalize the results of one-dimensional Gauss-Galerkin methods originally given by Dawson [5]. The basic idea of one-dimensional Gauss-Galerkin methods is that discrete measures are constructed for the approximation of the expected values and the Gauss quadrature formula is used to find the nodes and the weights in the numerical computation. To deal with a two-dimensional problem, the idea of variable splitting is adopted. More specifically, finite difference methods are first utilized in one variable and Gauss-Galerkin methods are employed in the other. As a consequence, the Gauss-Galerkin approach is used only in one dimension at each time step.

After two-dimensional Fokker-Planck equations are discretized to semi-discrete equations by finite difference methods in one direction, the convergence of the difference approximation is established based on energy type estimates. Using theories of measures and moments, the Gauss-Galerkin approximation for semi-discrete equations is shown to converge when one-dimensional Gauss-Galerkin approximation is used in the second direction. Combining the above results, the convergence of Gauss-

Galerkin/finite difference approximation is derived. The analysis is much more rigorous and general than the corresponding results for one-dimensional Gauss-Galerkin methods previously obtained by Abrouk, Dawson, HajJafar, Salehi and Yen [1, 5, 6].

In the third part, computer implementation and intensive numerical tests of Gauss-Galerkin/finite difference methods are carried out. In order to find nodes and weights from moments in one-dimensional Gauss-Galerkin methods, a different implementation from those by Abrouk and HajJafar [1, 6] is used which appears to be very efficient and very accurate. The coupling of difference approximation with Gauss-Galerkin approach is used successfully for the two-dimensional case. Numerical results on several test problems are also presented. For the model test problem 1, we compare the numerical solution based on the Gauss-Galerkin/finite difference method with that based on the conventional finite difference method. It is demonstrated that the Gauss-Galerkin/finite difference methods seem to be superior than the two-dimensional finite difference method in achieving high accuracy. For the test problem 2, the exact solution is known analytically. This, therefore, allows us to compare the numerical solution with the exact solution. For the test problem 3, a simple second order linear random oscillation problem is formulated in terms of its Fokker-Planck equation and the partial differential equation is solved by the Gauss-Galerkin/finite difference methods. Graphics plots are provided to show various behaviors of the exact solution and the numerical solution.

We now outline the remainder of the thesis. In Chapter 2, some basic definitions related to stochastic processes are given. In Chapter 3, we discuss the first and the second order stochastic differential equations and their corresponding Fokker-Planck equations. In Chapter 4, some general approximation methods for Fokker-Planck equations are discussed. In Chapter 5, some model problems and their Gauss-Galerkin/finite difference approximation are presented. Chapter 6 is devoted to the convergence analysis of the two-dimensional Gauss-Galerkin/finite difference method. In Chapter 7, the convergence analysis of Gauss-Galerkin/finite difference method is extended to a second order nonlinear random oscillation problem. In Chapters 8 and 9, computational algorithms and their implementations are presented. Discussions on the future work and some concluding remarks are given in Chapter 10.

2 Basic definitions

In this chapter, we introduce some definitions and notations needed for our work. Some results from stochastic analysis are also presented.

2.1 Random variables and stochastic processes

Probability theory pertains to the various possible outcomes that might be obtained and possible events that might occur when an experiment is performed. The collection of all possible outcomes of an experiment is called a sample space of the experiment, denoted by Ω. An event E can be regarded as a certain subset of possible outcomes in the space. In defining a random variable we shall proceed in accordance with a probability measure, with each elementary event e, we associate with it a certain number $X = f(e)$. We say that X is a random variable if the function $f(e)$ is measurable relative to the probability. In other words, we demand that for each value of x $(-\infty < x < +\infty)$ the set Ax of those e for which $f(e) < x$ should belong to the set of random events and hence, that for it there should be defined the probability $P(X < x) = P(Ax) = F(x)$ which we have called the distribution function of X.

Thus a random variable is a variable quantity whose values depend on chance and for which a distribution function of probabilities has been defined.

After introducing the random variables, we will introduce the definition of a stochastic process. Let U be a set of elementary events and t a continuous parameter. A stochastic process is defined as the function of two arguments: $X(t) = y(e, t)$ $(e \in U)$. For every value of the parameter t, the function $y(e, t)$ is a function of e only and, consequently is a random variable. For every fixed value of the argument e (i.e. for every elementary event), $y(e, t)$ depends only on t and is thus simply a function of one real variable. Every such function is called a realization of the stochastic process $X(t)$. We may regard a stochastic process either as a collection of random variables $X(t)$ that depend on the parameter t, or as a collection of the realizations of the process $X(t)$. Thus a process is determined if the probability measure in the function space of its realization is specified. Two stochastic processes $X(t)$ and $Y(t)$ are said to be stochastically equivalent if, for every t, we have $X(t) = Y(t)$ with probability 1. Then $X(t)$ is called a version of $Y(t)$ and vice versa.

2.2 Transition probability density

Assume that the condition probability density functions exist. A stochastic process is Markovian if, for $t_1 < t_2 < \cdots < t_n < \cdots < t_{m+n}$,

$$\begin{aligned} p(x_{n+1}, t_{n+1}; \ldots; x_{n+m}, t_{n+m} \mid x_1, t_1; \ldots; x_n, t_n) \\ = p(x_{n+1}, t_{n+1}; \ldots; x_{n+m}, t_{n+m} \mid x_n, t_n) \end{aligned} \tag{2.1}$$

It means that "the past and future are statistically independent when the present is known". The following is generally true for probability density functions.

$$\begin{aligned} & p(x_1, t_1; \ldots; x_n, t_n) \\ & = p(x_2, t_2; \ldots; x_n, t_n \mid x_1, t_1) \cdot p(x_1, t_1) \\ & = p(x_3, t_3; \ldots; x_n, t_n \mid x_1, t_1;\ x_2, t_2) \cdot p(x_2, t_2 \mid x_1, t_1) \cdot p(x_1, t_1) \\ & = \cdots \\ & = p(x_n, t_n \mid x_1, t_1;\ x_2, t_2; \ldots; x_{n-1}, t_{n-1}) \cdots p(x_2, t_2 \mid x_1, t_1) \cdot p(x_1, t_1) \end{aligned}$$

If $t_1 < t_2 < \cdots < t_n$, the Markovian property of $x(t)$ simplifies the above equation:

$$\begin{aligned} & p(x_1, t_1; \ldots; x_n, t_n) \\ & = p(x_n, t_n \mid x_{n-1}, t_{n-1}) \cdots p(x_2, t_2 \mid x_1, t_1) \cdot p(x_1, t_1) \end{aligned}$$

The condition probability density function $p(x_i, t_i \mid x_{i-1}, t_{i-1})$ is, for the Markov process, called the transition probability density function. The transition probability density function gives the density of probability of a transition from one point in phase space, x_{i-1}, at time t_{i-1} to another point in phase space, x_i, at time t_i, where $t_i > t_{i-1}$.

2.3 Markov process and Chapman-Kolmogorov equations

Using the Markov property; if $t_1 < t_2 < t_3$,

$$p(x_2, t_2; x_3, t_3 \mid x_1, t_1) = p(x_3, t_3 \mid x_2, t_2) \cdot p(x_2, t_2 \mid x_1, t_1)$$

Integrating over x_2, this becomes

$$p(x_3, t_3 \mid x_1, t_1) = \int_{\mathbf{R}^m} p(x_3, t_3 \mid x_2, t_2) \cdot p(x_2, t_2 \mid x_1, t_1) dx_2. \tag{2.2}$$

This equation is known as the Chapman-Kolmogorov or Smoluchowski equation [2].

2.4 Wiener process, Gaussian distribution

A Wiener process or Brownian motion is a special stochastic process, satisfying

(i) $x(0) = 0$.

(ii) If $t_1 < t_2 < \cdots < t_n$, the differences

$$[w(t_2 - w(t_1)], [w(t_2) - w(t_2)], \cdots, [w(t_n) - w(t_{n-1})]$$

are independent.

(iii) If $0 \leq s < t$, $w(t) - w(s)$ is normally distributed with

$$E[w(t) - w(s)] = (t-s) \cdot \mu \text{ , } E((w(t) - w(s))^2) = \sigma^2 \cdot (t-s)$$

where μ is called the drift and σ^2 is called the variance.

We can express these properties in terms of differentials

(i) $dw(t_i)$, $i = 1, \ldots, n$, are independent.

(ii) $dw(t)$ has a normal distribution and

$$E(dw(t)) = \mu \, dt \text{ , } \qquad E(dw^2(t)) = \sigma^2 dt.$$

Combining (i) & (ii) together, we have

$$E[dw(t) \cdot dw(s)] = \sigma^2 \cdot \delta(t-s) dt \cdot ds$$

where $\delta(t-s)$ is the Dirac-delta function.

If $\mu = 0$, $\sigma^2 = 1$, it is called normalized Wiener process. For any Brownian motion with mean μ and variance σ^2, $(\frac{x(t)-\mu t}{\sigma})$ is a normalized Wiener process. A random variable x has a normal distribution with mean μ and variance σ^2 $(-\infty < \mu < +\infty, \sigma > 0)$ if x has a continuous distribution for which the probability density function $f(x \mid \mu, \sigma^2)$ is given as follows:

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \text{ for } -\infty < x < \infty. \tag{2.3}$$

If $\vec{x} \in \mathbf{R}^n$, $\vec{x}$ has a normal distribution with mean vector $\vec{\mu}$ and covariance matrix k if $\vec{x}$ has a continuous distribution for which the probability density function $f(\vec{x} \mid \vec{\mu}, k)$ as follows:

$$f(\vec{x} \mid \vec{\mu}, k) = \frac{1}{(2\pi)^{\frac{n}{2}} |k|^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T k^{-1} (\vec{x}-\vec{\mu})}. \tag{2.4}$$

A d-dimensional process $w(t) = (w_1(t), \ldots, w_d(t))$ is called a d-dimensional Brownian motion if each process $w_i(t)$ is a Brownian motion and if the *σ-fields* $\mathcal{F}(w_i(t),\ t \geq 0)$, $1 \leq i \leq n$, are independent.

3 Stochastic differential equations

3.1 First order stochastic differential equations

Many problems in applications can be modeled as systems of first order differential equations of the form

$$\frac{dx}{dt} = a(t,x) + \sum_{k=1}^{m} b_k(t,x)\frac{dw_k(t)}{dt} \tag{3.1}$$

$$x(t_0) = y \tag{3.2}$$

where x, a, b_k, $k = 1, \dots, m$ are m vectors and $w_k(t)$ for $k = 1, 2, \dots, m$ are independent processes of Brownian motion. The vectors $a(t,x)$, $b_k(t,x)$ are defined for $t \in [t_0, T]$, $x \in \mathbf{R}^m$ and their ranges are in $\mathbf{R}^m$. We can write (3.1) in differential form since $\frac{dx}{dt}$ may exist almost nowhere,

$$dx = a(t,x)dt + \sum_{k=1}^{m} b_k(t,x)dw_k(t), \tag{3.3}$$

$$x(t_0) = y.$$

(3.3) is equivalent to the integral equation

$$x(t) = y + \int_{t_0}^{t} a[s, x(s)]ds + \sum_{k=1}^{m} \int_{t_0}^{t} b_k[s, x(s)]dw_k(s). \tag{3.4}$$

Theorem 1 (Ito 1951) *Let $a(t,x)$ and $\{b_j(t,x), j = 1, 2, \dots, m\}$ denote Borel functions defined for $t \in [t_0, T]$ and $x \in \mathbf{R}^m$ with ranges in $\mathbf{R}^m$. If there exists a constant κ, such that*

$$|a(t,x)|^2 + \sum_{k=1}^{m} |b_k(t,x)|^2 \leq \kappa^2(1 + |x|)^2,$$

$$|a(t,x) - a(t,y)| + \sum_{k=1}^{m} |b_k(t,x) - b_k(t,y)| \leq \kappa|x - y|,$$

for every $x, y \in \mathbf{R}^m$, then (3.4) has a solution $x(t)$ which is unique up to a stochastic equivalence and continuous with probability one. This solution $x(t)$ is a Markov process whose transition probability $P(A, t|y, s)$ for $s < t$ is given by the relation $P(A, t|y, s) = P(x_{s,y}(t) \in A)$ where $x_{s,y}(t)$ is the solution of the integral equation

$$x_{s,y}(t) = y + \int_{s}^{t} a[\xi, x_{s,y}(\xi)]d\xi + \sum_{k=1}^{m} \int_{s}^{t} b_k[\xi, x_{s,y}(\xi)]dw_k(\xi).$$

If the functions $a(t,x)$ and $b_k(t,x), k = 1,\ldots,m$ are continuous with respect to t, then the process $x(t)$ is a diffusion process with transfer vector $a(t,x)$ and diffusion operator $B(t,x)$ satisfying the equation

$$[B(t,x)z,z] = \sum_{k=1}^{m}[b_k(t,x),z]^2.$$

If $a(t,x)$ and $b_k(t,x)$, $k = 1,\ldots,m$ do not depend on t, and satisfy the condition of the above theorem, then the equation $dx = a(x)dt + \sum_{k=1}^{m} b_k(x)dw_k(t)$ is transition invariant with respect to t and the solution $x(t)$ is a homogeneous Markov process, that is, the transition probability $P(A, t+\tau \mid y,t)$ is independent of t.

3.2 Diffusion process

A stochastic process $x(t)$ is a diffusion process if the following conditions are satisfied:

(a) For every y and every $\varepsilon > 0$, the transition probability density function $p(x,t \mid y,s)$ satisfies the condition

$$\int_{|x-y|>\varepsilon} p(x,t \mid y,s)dx = o(t-s) \tag{3.5}$$

uniformly over $t > s$ and $x,y \in \mathbf{R}^m$.

(b) There exist functions $a_i(t,x)$, $B_{ij}(s,y)$ such that for every $y \in \mathbf{R}^m$ and every $\varepsilon > 0$,

(i) for $1 \le i \le m$,

$$\int_{|x-y|\le\varepsilon} (x_i - y_i)p(x,t \mid y,s)dx = a_i(t,y)(t-s) + o(t-s)\ , \tag{3.6}$$

(ii) for $1 \le i,j \le m$,

$$\int_{|x-y|\le\varepsilon} (x_i - y_i)(x_j - y_j)p(x,t \mid y,s)dx = B_{ij}(s,y)(t-s) + o(t-s), \tag{3.7}$$

uniformly for $t \ge s$ and $x,y \in \mathbf{R}^m$.

Theorem 2 (Fokker-Planck-Kolmogorov)

(i) If the transition probability $P(A,t \mid y,s)$ has a density $p(x,t \mid y,s)$ so that

$$P(A,t \mid y,s) = \int_A p(x,t \mid y,s)dx; \tag{3.8}$$

(ii) If conditions (a) and (b) are satisfied uniformly with respect to y, and if there exist continuous derivatives

$$\frac{\partial}{\partial t}[p(x,t \mid y,s)], \quad \frac{\partial}{\partial x_i}[a_i(t,x)p(x,t \mid y,s)]$$

and

$$(\frac{\partial^2}{\partial x_i \partial x_j})[B_{ij}(t,x)p(x,t \mid y,s)] \quad i,j = 1,\dots,m,$$

then $p(x,t \mid y,s)$ for $x,y \in \mathbf{R}^m$ and $t \in [s,T]$ satisfies the equation

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{m} \frac{\partial}{\partial x_i}[a_i(t,x)p] + \frac{1}{2}\sum_{i,j=1}^{m} \frac{\partial^2}{\partial x_i \partial x_j}[B_{ij}(t,x)p] \tag{3.9}$$

with initial condition

$$\lim_{t \downarrow s} p(x,t \mid y,s) = \delta(x-y). \tag{3.10}$$

Proof. Without loss of generality, let us consider $m = 1$, i.e. we try to prove:

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}[a(t,x)p] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[b(t,x)p]\,.$$

Suppose that $Q(x)$ is any function with continuous first and second derivatives, and has compact support in $[a,b]$. Then by continuity $Q(a) = Q'(a) = Q''(a) = Q(b) = Q'(b) = Q''(b) = 0$.

$$\begin{aligned}
I &= \int_{-\infty}^{+\infty} Q(u)\frac{\partial p}{\partial t}du \\
&= \int_a^b Q(u)\frac{\partial}{\partial t}p(u,t \mid y,s)du \\
&= \lim_{\Delta t \to 0}\int_a^b Q(u)\frac{p(u,t+\Delta t \mid y,s) - p(u,t \mid y,s)}{\Delta t}du \\
&= \lim_{\Delta t \to 0}\frac{1}{\Delta t}\left[\int_a^b Q(u)\int_{-\infty}^{+\infty} p(x,t \mid y,s)p(u,t+\Delta t \mid x,t)dxdu \right. \\
&\qquad \left. - \int_a^b Q(u)p(u,t \mid y,s)du\right]
\end{aligned}$$

(by (2.2), $p(u,t+\Delta t \mid y,s) = \int_{-\infty}^{+\infty} p(x,t \mid y,s)\,p(u,t+\Delta t \mid x,t)dx$.)

$$\begin{aligned}
&= \lim_{\Delta t \to 0}\frac{1}{\Delta t}\left\{\int_{-\infty}^{+\infty} p(x,t \mid y,s)[\int_a^b Q(u)p(u,t+\Delta t \mid x,t)du]dx \right. \\
&\qquad \left. - \int_a^b Q(x)p(x,t \mid y,s)dx\right\} \\
&= \lim_{\Delta t \to 0}\frac{1}{\Delta t}\int_{-\infty}^{+\infty} p(x,t \mid y,s)[\int_a^b p(u,t+\Delta t \mid x,t)Q(u)du - Q(x)]dx.
\end{aligned}$$

By (3.5) and the fact that $Q(x)$ is bounded since it is continuous on $[a,b]$, we have

$$\int_{|u-x|>\epsilon} p(u,t+\Delta t \mid x,t)Q(u)du \;=\; o(\Delta t)\;.$$

So

$$\begin{aligned}
& \frac{1}{\Delta t}\left[\int_a^b p(u,t+\Delta t \mid x,t)Q(u)du - Q(x)\right] \\
= \; & \frac{1}{\Delta t}\left[\int_{|u-x|\le\epsilon} p(u,t+\Delta t \mid x,t)Q(u)du - Q(x)\right] + o(1)\;.
\end{aligned}$$

We expand $Q(u)$ about x:

$$Q(u) = Q(x) + Q'(x)(u-x) + \frac{Q''(x)}{2}(u-x)^2 + o(u-x)^2\;.$$

Then

$$\begin{aligned}
& \frac{1}{\Delta t}\left[\int_{|u-x|\le\epsilon} p(u,t+\Delta t \mid x,t)Q(u)du - Q(x)\right] \\
= \; & \frac{1}{\Delta t}\{\int_{|u-x|\le\epsilon} p(u,t+\Delta t \mid x,t)[Q(x) + Q'(x)(u-x) + \frac{Q''(x)}{2} \\
& \quad +o(u-x)^2]du \; - Q(x)\} \\
= \; & -Q(x)\frac{1}{\Delta t}\int_{|u-x|>\epsilon} p(u,t+\Delta t \mid x,t)du \\
& +Q'(x)\frac{1}{\Delta t}\int_{|u-x|\le\epsilon} p(u,t+\Delta t \mid x,t)du \\
& +\frac{Q''(x)}{2}\frac{1}{\Delta t}\int_{|u-x|\le\epsilon}(u-x)^2 p(u,t+\Delta t \mid x,t)du \\
& +\frac{1}{\Delta t}\int_{|u-x|\le\epsilon} o(u-x)^2 p(u,t+\Delta t \mid x,t)du \\
= \; & Q'(x)\,a(t,x) + \frac{Q''(x)}{2}b(t,x)\;, \text{ as } \Delta t \to 0\;, \\
& \text{(By (3.5), (3.6) and (3.7).)}
\end{aligned}$$

Therefore

$$\begin{aligned}
I \;&=\; \int_{-\infty}^{+\infty} p(x,t \mid y,s)[Q'(x)a(t,x) + \frac{Q''(x)}{2}b(t,x)]dx \\
&=\; \int_{-\infty}^{+\infty} a(t,x)p(x,t \mid y,s)Q'(x)dx \\
&\quad +\frac{1}{2}\int_{-\infty}^{+\infty} b(t,x)p(x,t \mid y,s)Q''(x)dx\;.
\end{aligned}$$

Since

$$\begin{aligned}
\int_{-\infty}^{+\infty} apQ'(x)dx &= (apQ)\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} Q\frac{\partial}{\partial x}(ap)dx \\
&= -\int_{-\infty}^{+\infty} Q\frac{\partial}{\partial x}(ap)dx\ , \\
\int_{-\infty}^{+\infty} bpQ''(x)dx &= (bpQ')\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} Q'\frac{\partial}{\partial x}(bp)dx \\
&= [-\frac{\partial}{\partial x}(bp)Q]\Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} Q\,\frac{\partial^2}{\partial x^2}(bp)dx \\
&= \int_{-\infty}^{+\infty} Q\,\frac{\partial^2}{\partial x^2}(bp)dx\ ,
\end{aligned}$$

we have

$$\begin{aligned}
I &= \int_{-\infty}^{+\infty} Q\,\frac{\partial p}{\partial t}dx \\
&= -\int_{-\infty}^{+\infty} Q\,\frac{\partial}{\partial x}(ap)dx + \frac{1}{2}\int_{-\infty}^{+\infty} Q\,\frac{\partial^2}{\partial x^2}(bp)dx \\
&= \int_{-\infty}^{+\infty} Q\,[-\frac{\partial}{\partial x}(ap) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(bp)]dx\ .
\end{aligned}$$

So

$$\int_{-\infty}^{+\infty} Q\,\frac{\partial p}{\partial t}dx = \int_{-\infty}^{+\infty} Q\,[-\frac{\partial}{\partial x}(ap) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(bp)]dx\ ,$$

and finally we get:

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(ap) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(bp)\ . \quad \square$$

This equation is called the Fokker-Planck equation (Fokker 1914) or the forward Kolmogorov equation (Kolmogorov 1931).

Theorem 3 (Kolmogorov)

(i) If the transition probability $P(A, t \mid y, s)$ has a density $p(x, t \mid y, s)$.

(ii) If conditions (a) and (b) above are satisfied, and if there exist continuous derivatives

$$\frac{\partial}{\partial s}[p(x,t \mid y,s)], \quad \frac{\partial}{\partial y_i}[p(x,t \mid y,s)]$$

and

$$\frac{\partial^2}{\partial y_i \partial y_j}[p(x,t \mid y,s)] \quad i,j = 1,\ldots,m,$$

then $p(x,t \mid y,s)$ for $y \in \mathbf{R}^m$ and $s \in [0,t]$ satisfies the equation

$$-\frac{\partial p}{\partial s} = \sum_{i=1}^{m} a_i(s,y)\frac{\partial p}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^{m} B_{ij}(s,y)\frac{\partial^2 p}{\partial y_i \partial y_j} \tag{3.11}$$

with initial condition

$$\lim_{s\uparrow t} p(x,t \mid y,s) = \delta(x-y). \tag{3.12}$$

This is the so-called backward or converse Kolmogorov equation. The proof is similar to the previous one.

3.3 Transformation of second order stochastic differential equations

Nonlinear equations of second order excited by white noise have the form

$$\ddot{x} + f(H)\dot{x} + g(x) = \dot{w}(t);\ x(0) = y,\ \dot{x}(0) = \dot{y}, \tag{3.13}$$

where

$$E(dw^2(t)) = 2Ddt\,,$$

and

$$H = \frac{1}{2}\dot{x}^2 + \int_0^x g(\eta)d\eta\,.$$

In order to transform it to a first order differential system, let

$$y_1 = x\ ,\ \ y_2 = \dot{x}\ ,$$

and

$$\vec{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\ ,\ \ \dot{\vec{W}}(t) = \begin{pmatrix} \dot{\tilde{w}} \\ \frac{\dot{w}(t)}{\sqrt{2D}} \end{pmatrix}\ ,$$

where $\dot{\tilde{w}}$, $\frac{\dot{w}(t)}{\sqrt{2D}}$ are independent and identically distributed. Then

$$\dot{\vec{Y}} = \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix}.$$

Therefore

$$\dot{\vec{Y}} = \begin{pmatrix} y_2 \\ -f(H)y_2 - g(y_1) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2D} \end{pmatrix} \dot{\vec{W}}(t).$$

We rewrite the above equation in differential form

$$d\vec{Y} = \begin{pmatrix} y_2 \\ -f(H)y_2 - g(y_1) \end{pmatrix} dt + \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2D} \end{pmatrix} d\vec{W}(t), \tag{3.14}$$

i.e.,

$$d\vec{Y} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \vec{Y} dt + \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2D} \end{pmatrix} \vec{G}(\vec{Y}) dt + \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2D} \end{pmatrix} d\vec{W}(t) \tag{3.15}$$

with initial condition

$$\vec{Y}(0) = \begin{pmatrix} y \\ \dot{y} \end{pmatrix}, \tag{3.16}$$

where

$$\vec{G}(\vec{Y}) = \begin{pmatrix} 0 \\ -\sqrt{2D}(f(H)y_2 + g(y_1)) \end{pmatrix}.$$

The above equation is a special case of the following system of stochastic differential equations:

$$d\vec{Y} = A\vec{Y}dt + B\vec{G}(\vec{Y})dt + Bd\vec{W}. \tag{3.17}$$

The existence and uniqueness of solutions of the above systems and the corresponding Fokker-Planck equations have been studied in, for example, [8, 9, 10, 11], under proper assumptions on the coefficients. The matrix B corresponding to equation (3.15) is

$$B = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2D} \end{pmatrix} \quad \text{and} \quad \tilde{B} = BB' = \begin{pmatrix} 0 & 0 \\ 0 & 2D \end{pmatrix}.$$

The corresponding Fokker-Planck equation can be written as

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\sum_{i=1}^{2} \frac{\partial}{\partial y_i}[a_i(t,\vec{Y})p] + \frac{1}{2}\sum_{i,j=1}^{2} \frac{\partial^2}{\partial y_i, \partial y_j}[\tilde{B}_{ij}(t,\vec{Y})p] \\ &= -\frac{\partial}{\partial y_1}[y_2 p] - \frac{\partial}{\partial y_2}[(-f(H)y_2 - g(y_1))p] + \frac{1}{2}\frac{\partial^2}{\partial y_2^2}[\tilde{B}_{22}(t,\vec{Y})p] \\ &= -y_2 \cdot \frac{\partial p}{\partial y_1} + \frac{\partial}{\partial y_2}[(f(H)y_2 + g(y_1))p] + \frac{2D}{2} \cdot \frac{\partial^2 p}{\partial y_2^2}, \end{aligned}$$

i.e.

$$\frac{\partial p}{\partial t} = -\dot{x}\frac{\partial p}{\partial x} + \frac{\partial}{\partial \dot{x}}[(f(H)\dot{x} + g(x))p] + D\frac{\partial^2 p}{\partial \dot{x}^2}. \tag{3.18}$$

4 Approximations of Fokker-Planck equations

4.1 Brief introduction

Since, in general, it is not possible to obtain exact statistics for the response of a nonlinear system excited by white noises, a number of techniques have been developed to obtain approximate solutions. Here, we will introduce the iterative method, which is used by Ilin and Khasminskii [8] and generalized by Kushner [11] to establish the existence and uniqueness of solutions of the Fokker-Planck-Kolmogorov equations. It can be used to construct approximate solutions of the Fokker-Planck-Kolmogorov equations.

4.2 Iterative methods

Consider the following system:

$$\begin{cases} \ddot{x} + \beta\dot{x} + x + \epsilon g(x,\dot{x}) = \dot{w}(t)\,, \quad \beta > 0\,, |\epsilon| < 1\,. \\ E(dw(t)^2) = 2\,D\,dt\,, \\ x(0) = y \;\;,\; \dot{x}(0) = \dot{y}\,. \end{cases} \tag{4.1}$$

The associated Fokker-Planck equation is:

$$\frac{\partial p}{\partial t} = -\dot{x}\frac{\partial p}{\partial x} + \frac{\partial}{\partial \dot{x}}\{[\beta\dot{x} + x + \epsilon g(x,\dot{x})]p\} + D\frac{\partial^2 p}{\partial \dot{x}^2} \tag{4.2}$$

with initial condition:

$$p(\vec{x},0 \mid \vec{y}) = \delta(x-y)\,\delta(\dot{x}-\dot{y})\,, \quad \vec{x} = (x,\dot{x}) \in \mathbf{R}^2\,.$$

Define

$$L_0 u = -\dot{x}\frac{\partial u}{\partial x} + \frac{\partial}{\partial \dot{x}}[(\beta\dot{x} + x)u] + D\frac{\partial^2 u}{\partial \dot{x}^2}\,. \tag{4.3}$$

We now rewrite (4.2) in the form

$$\frac{\partial p}{\partial t} = L_0 p + \epsilon\,\frac{\partial}{\partial \dot{x}}[g(x,\dot{x})p]\,. \tag{4.4}$$

The existence of a unique solution is guaranteed if $g(x,\dot{x})$, $g_x(x,\dot{x})$, and $g_{\dot{x}}(x,\dot{x})$ are bounded for all $\vec{x} \in \mathbf{R}^2$ and (4.2) satisfies some appropriate conditions [2].

Consider:

$$\begin{cases} \frac{\partial G}{\partial t} = L_0 G\,, \\ G(\vec{x},0 \mid \vec{y}) = \delta(x-y)\delta(\dot{x}-\dot{y})\,. \end{cases} \tag{4.5}$$

It is equivalent to:

$$\begin{cases} \ddot{x} + \beta\dot{x} + x = \dot{w}(t) \;\;,\; \beta > 0, \\ E(dw(t)^2) = 2\,D\,dt\,, \\ x(0) = y \;\;,\; \dot{x}(0) = \dot{y}\,. \end{cases} \tag{4.6}$$

We rewrite (4.4) in the form of an integral equation

$$p(\vec{x},t \mid \vec{y}) = p_0(\vec{x},t \mid \vec{y}) + \epsilon \int_0^t \int_{\mathbb{R}^2} p_0(\vec{x},t-\tau \mid \vec{\xi}) \frac{\partial}{\partial \xi_2}[g(\vec{\xi})p(\vec{\xi},\tau \mid \vec{y})]\, d\vec{\xi} d\tau. \tag{4.7}$$

Integrating by parts w.r.t ξ_2, we obtain

$$p(\vec{x},t \mid \vec{y}) = p_0(\vec{x},t \mid \vec{y}) - \epsilon \int_0^t \int_{\mathbb{R}^2} g(\vec{\xi})\, p(\vec{\xi},\tau \mid \vec{y}) \frac{\partial}{\partial \xi_2}[p_0(\vec{x},t-\tau \mid \vec{\xi})]\, d\vec{\xi} d\tau. \tag{4.8}$$

We introduce the following iterative scheme:

$$p_n(\vec{x},t \mid \vec{y}) = p_0(\vec{x},t \mid \vec{y}) - \epsilon \int_0^t \int_{\mathbb{R}^2} g(\vec{\xi})\, p_{n-1}(\vec{\xi},\tau \mid \vec{y}) \frac{\partial}{\partial \xi_2}[p_0(\vec{x},t-\tau \mid \vec{\xi})]\, d\vec{\xi} d\tau. \tag{4.9}$$

In order to start this scheme, we need to solve $p_0(\vec{x},t \mid \vec{y})$ first. we transform (4.6) to a system of first order differential equations (Section 3.3)

$$d\vec{y} = A\vec{y}dt + B\, d\vec{w}(t)$$

with

$$\vec{y} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ -1 & -\beta \end{pmatrix}, \quad \text{and } B = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2D} \end{pmatrix}.$$

Assuming that $p_0(\vec{x},t \mid \vec{y})$ is the solution of (4.5), then $p_0(\vec{x},t \mid \vec{y})$ is the transition probability density function for (4.6). Since (4.6) is a linear system and $w(t)$ is a Wiener process, $p_0(\vec{x},t \mid \vec{y})$ is Gaussian with mean value vector $m(t \mid \vec{y}) = e^{At}\vec{y}$ and covariance matrix $K(t \mid \vec{y}) = \int_0^t e^{As}BB'e^{A's}ds$.

We now establish

Lemma 1

(i) $m(t \mid \vec{y})$ is bounded.

(ii) $K(t \mid \vec{y})$ is positive-definite , $\forall t > 0$.

Proof:

(i) is clear since the eigenvalues of A have negative real parts.

(ii) Clearly $K(t \mid \vec{y})$ is nonnegative-definite.

Suppose for some $t > 0$, $\exists$ nonzero $\vec{y}$, $\ni \vec{y}'K\vec{y} = 0$, i.e.

$$\begin{aligned} \vec{y}'K\vec{y} &= \vec{y}' \left(\int_0^t e^{As}BB'e^{A's}ds \right) \vec{y} \\ &= \int_0^t \vec{y}'e^{As}BB'e^{A's}\vec{y}ds \\ &= 0\ . \end{aligned}$$

Thus $B'e^{A's}\vec{y} = 0$, $\forall s \in [0,t]$ (since $\vec{y}'e^{As}BB'e^{A's}\vec{y} = \|B'e^{A's}\vec{y}\|^2 \geq 0$).

Let

$$\vec{y} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \quad , \quad e^{A's} = \begin{pmatrix} a_{11}(s) & a_{12}(s) \\ a_{21}(s) & a_{22}(s) \end{pmatrix} .$$

We have

$$\begin{aligned} 0 &= B'e^{A's}\vec{y} \\ &= \sqrt{2D} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11}(s) & a_{12}(s) \\ a_{21}(s) & a_{22}(s) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \\ &= \sqrt{2D} \begin{pmatrix} 0 & 0 \\ a_{21}(s) & a_{22}(s) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \\ &= \sqrt{2D} \begin{pmatrix} 0 \\ a_{21}(s)\xi_1 + a_{22}(s)\xi_2 \end{pmatrix} . \end{aligned}$$

Thus

$$a_{21}(s)\xi_1 + a_{22}(s)\xi_2 = 0 , \forall s \in [0,t]. \tag{4.10}$$

Moreover

$$\frac{d}{ds}\left(e^{A's}\right) = A'e^{A's} \Longrightarrow \frac{d}{ds}\left(e^{A's}\vec{y}\right) = A'e^{A's}\vec{y} ,$$

and

$$\begin{aligned} \frac{d}{ds}\left(e^{A's}\vec{y}\right) &= \frac{d}{ds} \begin{pmatrix} a_{11}(s)\xi_1 + a_{12}(s)\xi_2 \\ a_{21}(s)\xi_1 + a_{22}(s)\xi_2 \end{pmatrix} \\ &= \frac{d}{ds} \begin{pmatrix} a_{11}(s)\xi_1 + a_{12}(s)\xi_2 \\ 0 \end{pmatrix} , \end{aligned}$$

$$\begin{aligned} A'e^{A's}\vec{y} &= \begin{pmatrix} 0 & -1 \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} a_{11}(s)\xi_1 + a_{12}(s)\xi_2 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ a_{11}(s)\xi_1 + a_{12}(s)\xi_2 \end{pmatrix} . \end{aligned}$$

Thus

$$\frac{d}{ds} \begin{pmatrix} a_{11}(s)\xi_1 + a_{12}(s)\xi_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ a_{11}(s)\xi_1 + a_{12}(s)\xi_2 \end{pmatrix} .$$

Therefore

$$a_{11}(s)\xi_1 + a_{12}(s)\xi_2 = 0 , \quad \forall s > 0 . \tag{4.11}$$

Combining (4.10) and (4.11) together, we obtain: $e^{A's}\vec{y} = 0$, which implies $\vec{y} = 0$. It is a contradiction to the assumption $\vec{y} \neq 0$! Thus $K(t \mid \vec{y})$ is positive-definite. □

The above implies that $p_0(\vec{x}, t \mid \vec{y})$ can be written explicitly,

$$p_0(\vec{x}, t \mid \vec{y}) = \frac{1}{2\pi |K|^{\frac{1}{2}}} e^{-\frac{1}{2}(\vec{x}-e^{At}\vec{y})'K^{-1}(\vec{x}-e^{At}\vec{y})} .$$

Theorem 4 *There exists some constant* $0 < \gamma < 1$ *such that*

$$|p_1(\vec{x}, t \mid \vec{y}) - p_0(\vec{x}, t \mid \vec{y})| \leq \gamma\, p_0(\alpha\vec{x}, t \mid \alpha\vec{y}) .$$

Proof: Assume that c is a generic positive constant. For $0 < \alpha < 1$,

$$\frac{|\frac{\partial}{\partial \xi_2}[p_0(\vec{x}, t \mid \vec{y})]|}{p_0(\alpha\vec{x}, t \mid \alpha\vec{y})} = |(0\ ,\ 1)e^{At}K^{-1}(\vec{x} - e^{At}\vec{y})|\; e^{-\frac{1-\alpha^2}{2}(\vec{x}-e^{At}\vec{y})'K^{-1}(\vec{x}-e^{At}\vec{y})}.$$

Define

$$u = K^{-\frac{1}{2}}(\vec{x} - e^{At}\vec{y}) \quad , \qquad v = K^{-\frac{1}{2}} e^{A't} \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

$$\begin{aligned} |(0\ ,\ 1)e^{At}K^{-1}(\vec{x} - e^{At}\vec{y})| &= |v'u| \leq |v'v|^{\frac{1}{2}}\, |u'u|^{\frac{1}{2}} \\ &= |(0\ ,\ 1)e^{At}K^{-1}e^{A't}\begin{pmatrix} 0 \\ 1 \end{pmatrix}|^{\frac{1}{2}} \cdot |(\vec{x} - e^{At}\vec{y})'K^{-1}(\vec{x} - e^{At}\vec{y})|^{\frac{1}{2}}. \end{aligned}$$

Thus

$$\begin{aligned} \frac{|\frac{\partial}{\partial \xi_2}[p_0(\vec{x}, t \mid \vec{y})]|}{p_0(\alpha\vec{x}, t \mid \alpha\vec{y})} \leq\ & \underbrace{|(0\ ,\ 1)e^{At}K^{-1}e^{A't}\begin{pmatrix} 0 \\ 1 \end{pmatrix}|^{\frac{1}{2}}}_{\text{(I)}} \\ & \cdot \underbrace{|(\vec{x} - e^{At}\vec{y})'K^{-1}(\vec{x} - e^{At}\vec{y})|^{\frac{1}{2}} e^{-\frac{1-\alpha^2}{2}(\vec{x}-e^{At}\vec{y})'K^{-1}(\vec{x}-e^{At}\vec{y})}}_{\text{(II)}}. \end{aligned}$$

(II) is bounded since $f(x) = |x|^{\frac{1}{2}} e^{-\frac{1-\alpha^2}{2}x}$ is a bounded function on $(-\infty, \infty)$ if $0 < \alpha < 1$. Now we only need to consider (I).

Case (1): $t > \delta$. Then

$$K(t) > K(\delta) \Longrightarrow K^{-1}(t) < K^{-1}(\delta) .$$

So K^{-1} is uniformly bounded.

$$|v'v|^{\frac{1}{2}} \;=\; \|v\|$$

$$
\begin{aligned}
&= \|K^{-\frac{1}{2}}e^{A't}\begin{pmatrix}0\\1\end{pmatrix}\| \\
&\leq \|K^{-\frac{1}{2}}\|\,\|e^{A't}\lambda\|\begin{pmatrix}0\\1\end{pmatrix}\| \\
&\leq c\,\|e^{A't}\| \quad \text{(since } K^{-1} \text{is uniformly bounded)}
\end{aligned}
$$

Since the eigenvalues of A have negative real parts, it follows that

$$|v'v|^{\frac{1}{2}} \sim ce^{-rt} \ . \tag{4.12}$$

Case (2): $0 < t \leq \delta << 1$.

$$
\begin{aligned}
K &= \int_0^t e^{As}BB'e^{A's}ds \\
&= 2D\int_0^t e^{As}\begin{pmatrix}0\\1\end{pmatrix}(0\ 1)e^{A's}ds \ .
\end{aligned}
$$

Lemma 2 *If* $0 < s \leq t \leq \delta << 1$, *then*

$$K^{-1} \sim \frac{c}{t^3}\begin{pmatrix}1 & -\frac{t}{2}\\ -\frac{t}{2} & \frac{t^2}{3}\end{pmatrix} \ . \tag{4.13}$$

Proof: Suppose A has eigenvalues λ_1 , λ_2 and the corresponding eigenvectors are v_1 , v_2 .

Case (i): $\lambda_1 \neq \lambda_2$, then $v_1 = \begin{pmatrix}1\\ \lambda_1\end{pmatrix}$, $v_2 = \begin{pmatrix}1\\ \lambda_2\end{pmatrix}$.

Clearly: $\begin{pmatrix}0\\1\end{pmatrix} = \frac{1}{\lambda_1-\lambda_2}\left[\begin{pmatrix}1\\ \lambda_1\end{pmatrix} - \begin{pmatrix}1\\ \lambda_2\end{pmatrix}\right]$.

Thus

$$
\begin{aligned}
e^{As}\begin{pmatrix}0\\1\end{pmatrix} &= \frac{1}{\lambda_1-\lambda_2}\left[e^{As}\begin{pmatrix}1\\ \lambda_1\end{pmatrix} - e^{As}\begin{pmatrix}1\\ \lambda_2\end{pmatrix}\right] \\
&= \frac{1}{\lambda_1-\lambda_2}\left[e^{\lambda_1 s}\begin{pmatrix}1\\ \lambda_1\end{pmatrix} - e^{\lambda_2 s}\begin{pmatrix}1\\ \lambda_2\end{pmatrix}\right]
\end{aligned}
$$

(since $e^{As}\vec{v} = e^{\lambda s}\vec{v}$ if $\vec{v}$ is an eigenvector of A with eigenvalue λ)

$$= \begin{pmatrix}\frac{e^{\lambda_1 s}-e^{\lambda_2 s}}{\lambda_1-\lambda_2}\\ \frac{\lambda_1 e^{\lambda_1 s}-\lambda_2 e^{\lambda_2 s}}{\lambda_1-\lambda_2}\end{pmatrix} \ .$$

Since $\lim_{s\to 0}\frac{e^{\lambda_1 s}-e^{\lambda_2 s}}{\lambda_1-\lambda_2} = s$ and $\lim_{s\to 0}\frac{\lambda_1 e^{\lambda_1 s}-\lambda_2 e^{\lambda_2 s}}{\lambda_1-\lambda_2} = 1$, for $0 < s \le t \le \delta <<$ 1 , there exist positive constant $c_1(\delta)$, $c_2(\delta)$, satisfying $c_1(\delta) < 1 < c_2(\delta)$ and

$$\lim_{\delta\to 0} c_1(\delta) = \lim_{\delta\to 0} c_2(\delta) = 1 \, .$$

So the following is true:

$$c_1(\delta)\begin{pmatrix} s \\ 1 \end{pmatrix} \le \begin{pmatrix} \frac{e^{\lambda_1 s}-e^{\lambda_2 s}}{\lambda_1-\lambda_2} \\ \frac{\lambda_1 e^{\lambda_1 s}-\lambda_2 e^{\lambda_2 s}}{\lambda_1-\lambda_2} \end{pmatrix} \le c_2(\delta)\begin{pmatrix} s \\ 1 \end{pmatrix} ,$$

which implies

$$2D\,c_1^2(\delta)\begin{pmatrix} s \\ 1 \end{pmatrix}(s,\ 1) \le e^{As}BB'e^{A's} \le 2D\,c_2^2(\delta)\begin{pmatrix} s \\ 1 \end{pmatrix}(s,\ 1) \, .$$

Therefore

$$c_1^2(\delta)\int_0^1 \begin{pmatrix} s^2 & s \\ s & 1 \end{pmatrix} ds \le \frac{K}{2D} \le c_2^2(\delta)\int_0^1 \begin{pmatrix} s^2 & s \\ s & 1 \end{pmatrix} ds \, .$$

Hence

$$K^{-1} \sim c\begin{pmatrix} \frac{t^3}{3} & \frac{t^2}{2} \\ \frac{t^2}{2} & t \end{pmatrix}^{-1} \sim \frac{c}{t^3}\begin{pmatrix} 1 & -\frac{t}{2} \\ -\frac{t}{2} & \frac{t^2}{3} \end{pmatrix} .$$

Case (ii): In this case , $\lambda_1 = \lambda_2 = -1$, $\beta = 2$.

Let $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, then $A = (v_1 \, , \, v_2)\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}(v_1 \, , \, v_2)^{-1}$. Thus

$$\begin{aligned} e^{As} &= (v_1 \, , \, v_2)\exp\left\{\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} s\right\}(v_1 \, , \, v_2)^{-1} \\ &= (v_1 \, , \, v_2)\exp\left\{\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} s + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} s\right\}(v_1 \, , \, v_2)^{-1} \\ &= (v_1 \, , \, v_2)\exp\left\{\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} s\right\}\exp\left\{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} s\right\}(v_1 \, , \, v_2)^{-1} \\ &= e^{-s}(v_1 \, , \, v_2)\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}(v_1 \, , \, v_2)^{-1} \, . \end{aligned}$$

Therefore

$$e^{As}(v_1 \, , \, v_2) = e^{-s}(v_1 \, , \, v_2)\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} .$$

Her

M

S

I

Hence

$$e^{As}\begin{pmatrix}0\\1\end{pmatrix} = e^{-s}(sv_1+v_2) = e^{-s}\begin{pmatrix}s\\-s+1\end{pmatrix}.$$

Moreover

$$e^{-\delta}(1-\delta)\begin{pmatrix}s\\1\end{pmatrix} \le e^{-s}\begin{pmatrix}s\\-s+1\end{pmatrix} \le \begin{pmatrix}s\\1\end{pmatrix}.$$

(since $0 < s \le \delta$)

Similar as in (i), we get:

$$2De^{-2\delta}(1-\delta)^2\int_0^t \begin{pmatrix}s^2 & s\\ s & 1\end{pmatrix} ds \le K \le 2D\int_0^t \begin{pmatrix}s^2 & s\\ s & 1\end{pmatrix} ds.$$

Hence

$$K^{-1} \sim c\begin{pmatrix}\frac{t^3}{3} & \frac{t^2}{2}\\ \frac{t^2}{2} & 1\end{pmatrix}^{-1} \sim \frac{c}{t^3}\begin{pmatrix}1 & -\frac{t}{2}\\ -\frac{t}{2} & \frac{t^2}{3}\end{pmatrix}. \quad \square$$

By (4.13) and since $0 < t \le \delta << 1$, we have

$$e^{At}K^{-1}e^{A't} \sim \frac{c}{t^3}\begin{pmatrix}1 & -\frac{t}{2}\\ -\frac{t}{2} & \frac{t^2}{3}\end{pmatrix}.$$

Therefore

$$\begin{aligned}|v'v|^{\frac{1}{2}} &= |(0\,,\,1)\,e^{At}K^{-1}e^{A't}\begin{pmatrix}0\\1\end{pmatrix}|^{\frac{1}{2}}\\ &\sim (\frac{c}{t^3})^{\frac{1}{2}}|(0\,,\,1)\begin{pmatrix}1 & -\frac{t}{2}\\ -\frac{t}{2} & \frac{t^2}{3}\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}|^{\frac{1}{2}},\end{aligned}$$

which implies

$$|v'v|^{\frac{1}{2}} \sim ct^{-\frac{1}{2}}. \tag{4.14}$$

Thus we have

$$\begin{aligned}&|p_1(\vec{x},t \mid \vec{y}) - p_0(\vec{x},t \mid \vec{y})|\\ =\;& |\epsilon|\,|\int_0^t\int_{\mathbf{R}^2} g(\vec{\xi})\,p_0(\vec{\xi},\tau \mid \vec{y})\,\frac{\partial}{\partial\xi_2}[p_0(\vec{x},t-\tau \mid \vec{\xi})]\,d\vec{\xi}d\tau|\\ &\text{(by (4.2.9))}\\ \le\;& |\epsilon|\,M\int_0^t\int_{\mathbf{R}^2} p_0(\vec{\xi},\tau \mid \vec{y})\,|\frac{\partial}{\partial\xi_2}[p_0(\vec{x},t-\tau \mid \vec{\xi})]|\,d\vec{\xi}d\tau\\ &\text{(since } g \text{ is bounded in } \mathbf{R}^2)\\ \le\;& |\epsilon|\,M\int_0^{t-\delta}\int_{\mathbf{R}^2} p_0(\vec{\xi},\tau \mid \vec{y})\,|\frac{\partial}{\partial\xi_2}[p_0(\vec{x},t-\tau \mid \vec{\xi})]|\,d\vec{\xi}d\tau\end{aligned}$$

$$
\begin{aligned}
& + |\epsilon|\, M \int_{t-\delta}^{t} \int_{\mathbb{R}^2} p_0(\vec{\xi}, \tau \mid \vec{y}) \, |\frac{\partial}{\partial \xi_2}[p_0(\vec{x}, t-\tau \mid \vec{\xi})]| \, d\vec{\xi} d\tau \\
\leq \; & |\epsilon|\, M \int_{0}^{t-\delta} \int_{\mathbb{R}^2} p_0(\vec{\xi}, \tau \mid \vec{y}) \, c\, e^{-r\tau} \, p_0(\alpha\vec{x}, t-\tau \mid \alpha\vec{\xi}) \, d\vec{\xi} d\tau \\
& + |\epsilon|\, M \int_{t-\delta}^{t} \int_{\mathbb{R}^2} p_0(\vec{\xi}, \tau \mid \vec{y}) \, c\, \tau^{-\frac{1}{2}} \, p_0(\alpha\vec{x}, t-\tau \mid \alpha\vec{\xi}) \, d\vec{\xi} d\tau \\
& (\text{ by (4.12) and (4.14) }) \\
\leq \; & |\epsilon|\, M \int_{0}^{t-\delta} c\, e^{-r\tau} \left[\int_{\mathbb{R}^2} p_0(\alpha\vec{x}, t-\tau \mid \alpha\vec{\xi}) \, p_0(\alpha\vec{\xi}, \tau \mid \alpha\vec{y}) \, d\vec{\xi} \right] d\tau \\
& + |\epsilon|\, M \int_{t-\delta}^{t} c\, \tau^{-\frac{1}{2}} \left[\int_{\mathbb{R}^2} p_0(\alpha\vec{x}, t-\tau \mid \alpha\vec{\xi}) \, p_0(\alpha\vec{\xi}, \tau \mid \alpha\vec{y}) \, d\vec{\xi} \right] d\tau \\
= \; & |\epsilon|\, c \left(\int_{0}^{t-\delta} e^{-r\tau} d\tau + \int_{t-\delta}^{t} \tau^{-\frac{1}{2}} d\tau \right) p_0(\alpha\vec{x}, t \mid \alpha\vec{y}) \\
& (\text{ by (2.2) }) \\
= \; & |\epsilon|\, c \left[\frac{1}{r} - \frac{e^{-r(t-\delta)}}{r} + 2\sqrt{t} - 2\sqrt{t-\delta} \right] p_0(\alpha\vec{x}, t \mid \alpha\vec{y}) \\
\leq \; & |\epsilon|\, c \left(\frac{1}{r} + 2\sqrt{t} - 2\sqrt{t-\delta} \right) p_0(\alpha\vec{x}, t \mid \alpha\vec{y}) \\
\leq \; & c\, |\epsilon|\, p_0(\alpha\vec{x}, t \mid \alpha\vec{y}) \\
& (\text{since } 2\sqrt{t} - 2\sqrt{t-\delta} \text{ is bounded for all} t \geq 0 \text{ when } \delta \text{ is small}).
\end{aligned}
$$

Thus $\gamma = c\,|\epsilon| < 1$ if $|\epsilon|$ is chosen small enough. □

Similarly, we can get

$$
|p_n - p_{n-1}| \leq \gamma^n \, p_0(\alpha\vec{x}, t \mid \alpha\vec{y}) \, .
$$

Therefore the series

$$
p(\vec{x}, t \mid \vec{y}) = p_0(\vec{x}, t \mid \vec{y}) + \sum_{n=0}^{\infty} [p_{n+1}(\vec{x}, t \mid \vec{y}) - p_n(\vec{x}, t \mid \vec{y})]
$$

converges uniformly w.r.t $\vec{x}$ and $\vec{y}$ for any fixed t and satisfies the integral equation (4.7).

An approximate solution of (4.7) is given by the partial sum p_n , where

$$
|p_n - p| \leq \frac{\gamma^{n+1}}{1-\gamma} \, p_0(\alpha\vec{x}, t \mid \alpha\vec{y}) \, .
$$

Since $\gamma < 1$, $|p_n - p| \to 0$ as $n \to \infty$. In addition , $\forall n$, we have

$$
\lim_{|\epsilon| \to 0} |\, p_n(\vec{x}, t \mid \vec{y}) - p(\vec{x}, t \mid \vec{y}) \,| = 0 \, .
$$

Therefore, the iterative scheme (4.9) is convergent.

5 Gauss-Galerkin methods

5.1 Fokker-Planck equations for nonlinear oscillations

Let $x_1 = x$ and $x_2 = \dot{x}$ and let us recall the Fokker-Planck equation obtained previously from the second order stochastic differential equation that models the nonlinear oscillation with random forcing term,

$$\frac{\partial p}{\partial t} = -x_2\frac{\partial p}{\partial x_1} + \frac{\partial}{\partial x_2}\{[f(H)x_2 + g(x_1)]p\} + D\frac{\partial^2 p}{\partial x_2^2} \quad , \tag{5.1}$$

with initial condition

$$p(x_1, x_2, 0) = q(x_1, x_2) \,. \tag{5.2}$$

Our interest is to solve the above equation numerically by extending the idea of Gauss-Galerkin approximation developed for the one-dimensional Fokker-Planck equations.

The most straightforward extension of the one dimensional Gauss-Galerkin method involves applying a finite difference type approximation in one of the spatial variables and the Gauss-Galerkin method in the other spatial variable. There are numerous issues to be considered such as how should the difference approximation in an infinite domain be applied.

5.2 Model equations

In order to construct and analyze the Gauss-Galerkin type approximation for equations of the form (5.1), we first study the following model equation:

$$\frac{\partial p}{\partial t} = -f(y)\frac{\partial p}{\partial y} + \varepsilon\frac{\partial^2 p}{\partial y^2} - \frac{\partial}{\partial x}(\alpha p) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(\beta^2 p) \ ,$$

for $x, y \in (-\infty, \infty)$, $t > 0$. Here, $\varepsilon > 0$, $f(y) > 0$ are assumed to be positive. Further assumptions on the coefficients $\alpha = \alpha(x, y)$ and $\beta = \beta(x, y)$ will be made later.

Let

$$\tilde{x} = \frac{1}{2}(1 + \coth x) = \frac{e^x}{e^x + e^{-x}} \quad , \quad \tilde{y} = \frac{1}{2}(1 + \coth y) = \frac{e^y}{e^y + e^{-y}} \ .$$

Then $\tilde{x}\,, \tilde{y} \in (0, 1)$ if $x, y \in (-\infty, \infty)$. Moreover, we have

$$v_x \ = \ v_{\tilde{x}}\frac{\partial \tilde{x}}{\partial x}$$

$$
\begin{aligned}
&= v_{\tilde{x}}\,\frac{e^{x}(e^{x}+e^{-x})-e^{x}(e^{x}-e^{-x})}{2(e^{x}+e^{-x})^{2}} \\
&= v_{\tilde{x}}\,\frac{2}{(e^{x}+e^{-x})^{2}} \\
&= 2\tilde{x}(1-\tilde{x})\,v_{\tilde{x}}\ , \\
v_{xx} &= (v_{\tilde{x}})_{x}\,\frac{2}{(e^{x}+e^{-x})^{2}}+v_{\tilde{x}}\left[\frac{2}{(e^{x}+e^{-x})^{2}}\right]_{x} \\
&= v_{\tilde{x}\tilde{x}}\left[\frac{2}{(e^{x}+e^{-x})^{2}}\right]^{2}+v_{\tilde{x}}\,\frac{-4\,(e^{x}-e^{-x})}{(e^{x}+e^{-x})^{3}} \\
&= 4\tilde{x}^{2}(1-\tilde{x})^{2}\,v_{\tilde{x}\tilde{x}}\;-\;4\tilde{x}(1-\tilde{x})(2\tilde{x}-1)\,v_{\tilde{x}}\ , \\
v_{y} &= v_{\tilde{y}}\,\frac{\partial\tilde{y}}{\partial y} \\
&= v_{\tilde{y}}\,\frac{e^{y}(e^{y}+e^{-y})-e^{y}(e^{y}-e^{-y})}{2(e^{y}+e^{-y})^{2}} \\
&= v_{\tilde{y}}\,\frac{2}{(e^{y}+e^{-y})^{2}} \\
&= 2\tilde{y}(1-\tilde{y})\,v_{\tilde{y}}\ , \\
v_{yy} &= (v_{\tilde{y}})_{y}\,\frac{2}{(e^{y}+e^{-y})^{2}}+v_{\tilde{y}}\left[\frac{2}{(e^{y}+e^{-y})^{2}}\right]_{y} \\
&= v_{\tilde{y}\tilde{y}}\left[\frac{2}{(e^{y}+e^{-y})^{2}}\right]^{2}+v_{\tilde{y}}\,\frac{-4\,(e^{y}-e^{-y})}{(e^{y}+e^{-y})^{3}} \\
&= 4\tilde{y}^{2}(1-\tilde{y})^{2}\,v_{\tilde{y}\tilde{y}}\;-\;4\tilde{y}(1-\tilde{y})(2\tilde{y}-1)\,v_{\tilde{y}}\ .
\end{aligned}
$$

Therefore the model equation becomes:

$$
\begin{aligned}
\frac{\partial p}{\partial t} &= -2[\tilde{y}(1-\tilde{y})\tilde{f}(\tilde{y})+2\varepsilon\tilde{y}(1-\tilde{y})(2\tilde{y}-1)]\frac{\partial p}{\partial\tilde{y}}+4\varepsilon\tilde{y}^{2}(1-\tilde{y})^{2}\frac{\partial^{2}p}{\partial\tilde{y}^{2}} \\
&\quad -2\tilde{x}(1-\tilde{x})\frac{\partial}{\partial\tilde{x}}(\alpha p)-2\tilde{x}(1-\tilde{x})(2\tilde{x}-1)\frac{\partial}{\partial\tilde{x}}(\beta^{2}p)+2\,\tilde{x}^{2}(1-\tilde{x})^{2}\frac{\partial^{2}}{\partial\tilde{x}^{2}}(\beta^{2}p)\,.
\end{aligned}
$$

Without loss of generality, replace $\tilde{x}$ by x , $\tilde{y}$ by y , and $\tilde{f}$ by f . This leads to

$$
\begin{aligned}
\frac{\partial p}{\partial t} &= -2[\,y(1-y)f(y)+2\varepsilon y(1-y)(2y-1)\,]\frac{\partial p}{\partial y}+4\varepsilon y^{2}(1-y)^{2}\frac{\partial^{2}p}{\partial y^{2}} \\
&\quad -2x(1-x)\frac{\partial}{\partial x}(\alpha p)-2x(1-x)(2x-1)\frac{\partial}{\partial x}(\beta^{2}p) \\
&\quad +2\,x^{2}(1-x)^{2}\frac{\partial^{2}}{\partial x^{2}}(\beta^{2}p). \qquad (5.3)
\end{aligned}
$$

For simplicity, we pose the boundary conditions

$$
p(x,0,t)=p(x,1,t)=0\ ,\ \ \forall\,t>0\,, \qquad (5.4)
$$

and the initial condition $p(x,y,0) = q(x,y)$.

Let us use $\mathbf{H} = \mathbf{L}^2(0,1)$ to denote the L^2 integrable function space. We denote the inner product by

$$(f,g) = \int_0^1 f(x)g(x)dx \ , \ \forall f,g \in \mathbf{H} \, .$$

We also define the Sobolev spaces by

$$\mathbf{W}^{m,2}(0,1) = \{f \in \mathbf{H} \ | \ f', f'', ..., f^m \in \mathbf{H}\} \ , \ m = 1,2,... .$$

Then the additional boundary conditions at $x = 0$ and $x = 1$ are specified in a way such that the following always holds,

$$(Lp, v) = (p, L^*v) \ , \forall v \in \mathbf{W}^{2,2}(0,1) \cap \mathbf{C}^2[0,1] \qquad (5.5)$$

where

$$Lp = -2x(1-x)\frac{\partial}{\partial x}(\alpha p) - 2x(1-x)(2x-1)\frac{\partial}{\partial x}(\beta^2 p) + 2\,x^2(1-x)^2\frac{\partial^2}{\partial x^2}(\beta^2 p) \ ,$$

and

$$\begin{aligned} L^*p \ &= \ [-2(2x-1)\alpha + 3(2x-1)^2\beta^2 - \beta^2]v + 2[x(1-x)\alpha - 3x(1-x)(2x-1)\beta^2]\frac{\partial v}{\partial x} \\ &\quad +2x^2(1-x)^2\beta^2\frac{\partial^2 v}{\partial x^2}. \end{aligned}$$

We note that sometimes the boundary terms may not vanish, such boundary terms often lead to singular measures corresponding to densities accumulated at the boundary.

For convenience, we also assume that the coefficients α and β satisfy certain conditions that imply the inequality

$$(-Lu \ , \ u) \geq -\lambda(u \, , \, u) \ , \qquad (5.6)$$

for a given constant λ and any function u which satisfies the boundary conditions required for equations (5.3). In addition, $\alpha(x,y,t)$, $\beta(x,y,t)$ are assumed to be smooth functions and

$$|\alpha(x,y,t)| \leq b_1, \ \forall y \in (0,1), \ t > 0, \ x \in (0,1) \ ,$$

$$|\beta(x,y,t)| \leq b_2, \ \forall y \in (0,1), \ t > 0, \ x \in (0,1)$$

for some positive constants b_1 and b_2.

5.3 Difference approximation in y

Our motivation is that if we first approximate the derivatives in the y direction by differences defined on some given grid, it then might be feasible to apply the one-dimensional Gauss-Galerkin methods at each of the grid points. In this regard, (5.3) is very convenient for variable splitting.

Finite difference approximation is one of the most widely used methods in solving partial differential equations. The difference approximation may be set up easily using standard techniques [13]. For example, the unit interval $[0, 1]$ in the y direction can be partitioned by a uniform grid. Let N_y be the total number of subintervals, $h_y = 1/N_y$ be the length of each subinterval and $\{y_j = jh_y, j = 0, 1, 2, \dots, N_y\}$ be the grid points. Then the second order derivative in y may be approximated by

$$p''(y_j) \approx \frac{p(y_{j+1}) - 2p(y_j) + p(y_{j-1})}{h_y^2}$$

with truncation error of the order $O(h_y^2)$.

There are several choices for the first order derivative in y, for instance, one can either use the center difference

$$\frac{p(y_{j+1}) - p(y_{j-1})}{2h_y}$$

which has truncation error of the order $O(h_y^2)$, or the one-sided difference in y approximated by

$$\frac{p(y_{j+1}) - p(y_j)}{h_y} \text{ or } \frac{p(y_j) - p(y_{j-1})}{h_y},$$

which has truncation error of the order $O(h_y)$.

Even though the center difference is the most accurate approximation among the ones described above, the upwind difference may have superior stability property which may become very important when the diffusion coefficient in the y direction is small or negligible and convection becomes the dominant phenomenon [7]. For this reason, we choose the backward difference approximation for the first order derivative in y.

Using the above difference approximations, we obtain the following semi-discrete approximation of (5.3). (By *semi-discrete*, we mean that the equation is only discretized in one variable while it remains continuous in the other variable.)

$$\frac{\partial p_j}{\partial t} = -2[y_j(1 - y_j)f(y_j) + 2\varepsilon y_j(1 - y_j)(2y_j - 1)]\frac{p_j - p_{j-1}}{h_y}$$

$$+4\varepsilon {y_j}^2(1-y_j)^2\frac{p_{j+1}-2p_j+p_{j-1}}{h_y^2}+L_jp, \qquad (5.7)$$

where

$$L_jp=-2x(1-x)\frac{\partial}{\partial x}(\alpha_jp)-2x(1-x)(2x-1)\frac{\partial}{\partial x}(\beta_j^2p)+2\,x^2\,(1-x)^2\frac{\partial^2}{\partial x^2}(\beta_j^2p), \quad (5.8)$$

for $1 \le j \le N_y-1$, $x \in (0,1)$ and $t>0$, with $p_0=p_{N_y}=0$, $\alpha_j=\alpha(x,y_j,t), \beta_j=\beta(x,y_j,t)$ and each $p_j, 1 \le j \le N_y-1$, must satisfy additional boundary conditions in the x direction that are required for the original solution $p=p(x,y,t)$ of equations (5.3), i.e., p_j satisfies equation (5.5) for any j, and L_j^*v is :

$$L_j^*v=[-2(2x-1)\alpha_j+3(2x-1)^2\beta_j^2-\beta_j^2]v+2[x(1-x)\alpha_j-3x(1-x)(2x-1)\beta_j^2]\frac{\partial v}{\partial x}$$
$$+2x^2\,(1-x)^2\beta_j^2\frac{\partial^2 v}{\partial x^2}. \qquad (5.9)$$

Then for any smooth test function $v \in \mathbf{W}^{2,2}(0,1)\cap \mathbf{C}^2[0,1]$, we have

$$(L_jp_j,v)=(p_j,L_j^*v)\ ,\ \ \forall j=1,2,...,N_y-1\ .$$

The initial condition is given by

$$p_j(x,y,0)=q(x,y_j)\ ,\ \ 1 \le j \le N_y-1\ ,\ \ x \in (0,1)\ . \qquad (5.10)$$

5.4 Gauss-Galerkin methods

The Galerkin methods usually refer to finite dimensional approximations to the weak formulations of partial differential equations. The Gauss-Galerkin methods, in our current context, have special forms. The test functions in the weak form are chosen to be polynomials, while the approximate solutions are taken to be discrete, or *atomic* measures. A unique feature for the Fokker Planck equations is that the inner product of *atomic* measures with polynomials are closely related with the moments of the probability distribution.

The first analysis of the Gauss-Galerkin methods for one dimensional Fokker Planck equations was given by Dawson [5]. A more complete analysis was given in [6], together with a computational algorithm and numerical examples. The model equations considered in [6] is of the form:

$$\frac{\partial p}{\partial t}=-\frac{\partial}{\partial x}(\alpha p)+\frac{1}{2}\frac{\partial^2}{\partial x^2}(\beta^2p)\ ,\ \ x \in (0,1)\,. \qquad (5.11)$$

Thus, equation (5.3) can be viewed as a generalization of the above equation in two dimensional space. On the other hand, one may also view its semi-discretization (5.7) as a system of equations and each of the equations is similar to (5.11).

In (5.11), the spatial interval is finite, thus, the standard L^2 spaces and inner products were used in [5, 6]. When dealing with semi-infinite intervals, these intervals become finite after changing variables.

To solve (5.7) by Gauss-Galerkin methods, let us first formulate its weak form. Taking the inner product of a test function $v \in \mathbf{W}^{2,2}(0,1) \cap \mathbf{C}^2[0,1]$ with the equation (5.7) and integrating by parts, we have for $1 \leq j \leq N_y - 1$,

$$\begin{aligned} \frac{d}{dt}(p_j, v) &= -\frac{c_j}{h_y}(p_{j-1} - p_j, v) + \frac{\varepsilon_j}{h_y^2}(p_{j+1} - 2p_j + p_{j-1}, v) \\ &\quad +([-2(2x-1)\alpha_j + 3(2x-1)^2\beta_j^2 - \beta_j^2]p_j, v) \\ &\quad +(2[x(1-x)\alpha_j - 3x(1-x)(2x-1)\beta_j^2]p_j, \frac{\partial v}{\partial x}) \\ &\quad +2(x^2(1-x)^2\beta_j^2 p_j, \frac{\partial^2 v}{\partial x^2}), \quad t > 0 , \end{aligned} \tag{5.12}$$

where

$$c_j = 2[y_j(1-y_j)f(y_j) + 2\varepsilon y_j(1-y_j)(2y_j - 1)] , \quad \varepsilon_j = 4\varepsilon {y_j}^2(1-y_j)^2 .$$

Let us now approximate the probability measure $p_j dx$ by

$$d\mu_j^n(x,t) = \sum_{k=1}^{n} a_j^k(t)\delta(x_j^k(t))dx , \quad 1 \leq j \leq N_y - 1 , \tag{5.13}$$

i.e., the probability measure is approximated by an n-point discrete Dirac-delta (*atomic*) measure. Here, one may choose different n for different index j, i.e., there could be different number of atoms at different values of y. But, for simplicity, we only consider the case where n is taken to be independent of j. $\{x_j^k = x_j^k(t) , 1 \leq k \leq n\}$ are the n nodes with $a_j^k(t)$ as their corresponding weights. As far as the test functions are concerned, we choose a set of linearly independent functions $\{v_i(x)\}_0^{2n-1}$. In particular, $v_i(x) = x^i$ for $i = 0, 1, 2, ..., 2n-1$ may be the standard choice. Then, (5.12) implies for $0 \leq i \leq 2n-1$, $1 \leq j \leq N_y - 1$ and $t > 0$,

$$\frac{d}{dt}\sum_{k=1}^{n} a_j^k(t)v_i(x_j^k(t)) = \sum_{k=1}^{n}\Bigg[(\frac{c_j}{h_y} + \frac{\varepsilon_j}{h_y^2})a_{j-1}^k(t)v_i(x_{j-1}^k(t))$$

$$-(\frac{c_j}{h_y} + \frac{2\varepsilon_j}{h_y^2})a_j^k(t)v_i(x_j^k(t)) + \frac{\varepsilon_j}{h_y^2}a_{j+1}^k(t)v_i(x_{j+1}^k(t))\Bigg]$$

$$+\sum_{k=1}^{n} a_j^k(t)(L_j^* v_i)(x_j^k(t)) \tag{5.14}$$

where L_j^* is defined as in (5.9) . Note that $v_i(x) = x^i$, $0 \leq i \leq 2n - 1$.

The system (5.14) is a system of $2n \times (N_y - 1)$ ordinary differential equations for the $n(N_y - 1)$ nodes and $n(N_y - 1)$ weights. Initial conditions for the ordinary differential equation system are given by

$$\sum_{k=1}^{n} a_j^k(0)v_i(x_j^k(0)) = (q(x, y_j), v_i(x)), \tag{5.15}$$

for $1 \leq j \leq N_y - 1$, $0 \leq i \leq 2n - 1$.

Using the properties of the Gauss quadrature [14], if $q_j(x) = q(x, y_j)$ is some nonnegative function whose support has positive measure, we may let the weight function in the Gauss quadrature be given by $q_j(x)$ for each j. Then for each j, the nodes $\{x_j^k(0)\}_{k=1}^n$ and weights $\{a_j^k(0)\}_{k=1}^n$ are uniquely determined by (5.15) and the $x_j^k(0)$'s are distinct and lie in $(0, 1)$ (see chapter 9 for details).

We shall make a crucial but also somewhat restrictive assumption that there exists a time interval $[0, T]$ such that the system of ordinary differential equations (5.14) with initial conditions (5.15) has a unique solution in $[0, T]$, with distinct nodes $\{x_j^k(0)\}_{k=1}^n$ in $(0, 1)$. Furthermore, the weights $\{a_j^k(t)\}$ stay positive for $t \in [0, T]$ and for all k and j.

The above assumptions imply that the approximate atomic measure would remain valid in a given time interval and the measure is positive in the sense of distribution. However, these assumptions can be proved rigorously only in some special cases which we will discuss briefly later.

5.5 Some useful results

In this section we present and prove some useful differential inequalities. Let us use the convention that a vector (or matrix) is said to be non-positive ($\preceq 0$) iff all its elements are non-positive, and

$$\vec{X} \preceq \vec{Y} \quad \text{iff} \quad \vec{X} - \vec{Y} \preceq \vec{0} \, .$$

Lemma 3 *Let $A(t)$ be an N by N matrix with smooth and bounded elements and its off-diagonal elements are all non-negative, i.e., $a_{ij}(t) \geq 0, 1 \leq i \neq j \leq N$. Let $\vec{M}(t) \in \mathbf{R}^N$ satisfy that $\vec{M}(0) \preceq \vec{0}$ and for $t > 0$ we have*

$$\frac{d}{dt}\vec{M}(t) \preceq A(t)\vec{M}(t) \, .$$

Then

$$\vec{M}(t) \preceq \vec{0}\ , \quad \forall t > 0\ .$$

Proof. Let $\lambda \geq -\min\{a_{ii}(t)\,,\ 1 \leq i \leq N\}$ be some given constant, then $A_1 = A + \lambda I_N$ is a matrix with all non-negative elements. Define

$$\vec{M}_1(t) = e^{\lambda t}\vec{M}(t)\ .$$

We have $\vec{M}_1(0) \preceq \vec{0}$ and for $t > 0$,

$$\frac{d}{dt}\vec{M}_1(t) \preceq A_1(t)\vec{M}_1(t)\ .$$

So

$$\frac{d}{dt}\vec{M}_1(t) = A_1(t)\vec{M}_1(t) + \vec{G}(t)\ ,$$

with $\vec{G}(t) \preceq \vec{0}$ for $t > 0$. Thus,

$$\begin{aligned} \vec{M}_1(t) \quad &= \quad exp\left\{\int_0^t A_1(u)du\right\}\vec{M}_1(0) \\ &\quad + \int_0^t exp\left\{\int_s^t A_1(u)du\right\}\vec{G}(s)ds\ , \quad \forall t > 0\ . \end{aligned}$$

Since A_1 is non-negative, then

$$B(t,s) = e^{\int_s^t A_1(s)ds}$$

is non-negative for $s \leq t$. Therefore, $\vec{M}_1(0) \preceq \vec{0}$ and $\vec{G}(s) \preceq \vec{0}$ for $s > 0$ imply

$$\vec{M}_1(t) \preceq \vec{0}\ , \ \forall t > 0\ .$$

Thus,

$$\vec{M}(t) \preceq \vec{0}\ , \quad \forall t > 0\ . \quad \square$$

Consequently, we get the following relation between the solution of the system of differential inequalities and the solution of the system of differential equations.

Corollary 1 [Comparison Principle] *Let $A(t)$ be an N by N matrix with smooth and bounded elements and its off-diagonal elements are all non-negative, i.e., $a_{ij}(t) \geq 0, 1 \leq i \neq j \leq N$. Let $\vec{M}(t) \in \mathbf{R}^N$ satisfy*

$$\frac{d}{dt}\vec{M}(t) \preceq A(t)\vec{M}(t) + \vec{G}(t)\ , \quad \forall t > 0\ ,$$

and $\vec{M}_1(t) \in \mathbf{R}^N$ *satisfy*

$$\frac{d}{dt}\vec{M}_1(t) = A(t)\vec{M}_1(t) + \vec{G}(t) \ , \ \ \forall t > 0 \ ,$$

with $\vec{M}_1(0) = \vec{M}(0)$. *Then*

$$\vec{M}(t) \preceq \vec{M}_1(t) \ , \ \ \forall t > 0 \ .$$

In particular, when $N = 1$, we have the following well-known inequality.

Corollary 2 [Gronwall's Inequality] *If for* $t > 0$, *we have*

$$\frac{d}{dt}f(t) \leq a(t)f(t) + g(t) \ ,$$

then,

$$f(t) \leq exp\left\{\int_0^t a(s)ds\right\} f(0) + \int_0^t exp\left\{\int_s^t a(u)du\right\} g(s)ds \ , \ \forall t > 0 \ .$$

Proof: Let

$$h(t) = exp\left\{\int_0^t a(s)ds\right\} f(0) + \int_0^t exp\left\{\int_s^t a(u)du\right\} g(s)ds \ .$$

Then $h(t)$ satisfies that

$$\frac{d}{dt}h(t) = a(t)h(t) + g(t) \ ,$$

with $h(0) = f(0)$. So,

$$f(t) \leq h(t) \ , \ \forall t > 0 \ . \quad \square$$

The following corollary can be viewed as a weak discrete maximum principle.

Corollary 3 *Let* $c_1 > 0$, $c_2 > 0$ *be given constant and* $\{f_j(t)\}_{j=0}^{N+1}$ *satisfies*

$$\frac{d}{dt}f_j(t) = c_1 f_{j-1}(t) - (c_1 + c_2)f_j(t) + c_2 f_{j+1}(t) \ , \ \ 1 \leq j \leq N \ , \ \forall t > 0$$

with $f_0(t) = f_{N+1}(t) = 0$ *and* $f_j(0) = g_j$, $1 \leq j \leq N$. *Then,*

$$f_j(t) \leq \max_{1 \leq j \leq N}\{g_j, 0\} \ , \ \ \forall t > 0 \ .$$

Proof: Define

$$a = \max_{1 \leq j \leq N}\{g_j, 0\}$$

and

$$\vec{M}(t) = (f_1(t) - a, f_2(t) - a, ..., f_N(t) - a)^{\mathrm{T}} \ .$$

Then

$$\frac{d}{dt}\vec{M}(t) = A\vec{M}(t) , \ \forall t > 0 .$$

for some tridiagonal matrix A with elements $a_{j,j-1} = c_1, a_{j-1,j} = c_2$ for $2 \leq j \leq N$ and $a_{j,j} = -(c_1 + c_2)$ for $1 \leq j \leq N$. Since $c_1 > 0, c_2 > 0$, and $\vec{M}(0) \preceq \vec{0}$, by the previous lemma,

$$\vec{M}(t) \preceq \vec{0} , \ \forall t > 0 .$$

This implies

$$f_j(t) \leq \max_{1 \leq j \leq N}\{g_j, 0\} , \qquad \forall 1 \leq j \leq N , \ and \, \forall t > 0 . \ \ \Box$$

Remark : We see from the proof that the differential equations in the corollary can be replaced by inequalities and the conclusion will still be valid. Furthermore, one could prove a stronger version which implies that if $f_k(t) = \max_{1 \leq j \leq N}\{g_j, 0\}$ for some $1 \leq k \leq N$ and $t > 0$, then $f_j(t) = 0$ for all j and all t.

Now, we state some results concerning the classical moment problem [15].

Given a sequence $\{m_n\}_{n=0}^{\infty}$, define the following differences

$$\Delta^0 m_n = m_n , \tag{5.16}$$

$$\Delta^k m_n = \Delta^{k-1} m_n - \Delta^{k-1} m_{n+1} , \ k = 1, 2, \cdots . \tag{5.17}$$

Then we have,

Theorem 5 *A necessary and sufficient condition for the existence of a solution of the Hausdorff moment problem, i.e., the existence of a unique nonnegative measure μ satisfying*

$$m_l = \int_0^1 x^l d\mu , \ l = 0, 1, 2, \cdots ,$$

is that

$$\Delta^k m_l \geq 0 , \ \ k, l = 0, 1, 2, \cdots .$$

5.6 Existence theorems for Gauss-Galerkin approximations

First, let us show that if the coefficients of equation (5.3) are constant, then, the system (5.14) with initial condition (5.15) has a unique global solution with positive weights and distinct nodes.

Theorem 6 *Let the coefficients of equation (5.3) be constant. If q is a probability distribution and the support of $q(x, y_j)$ has positive measure for all $1 \leq j \leq N_y - 1$, and (5.7) has a set of nonnegative solutions $\{p_j,\ j = 1, 2, \ldots, N_y - 1\}$, each having support of positive measure, then the solution of (5.14) with initial condition (5.15) exists for all time and the weights $\{a_j^k(t)\}$ remain positive and the nodes $\{x_j^k(t)\}$ remain distinct.*

Proof: We can use each p_j as the weight function to define a Gauss quadrature, i.e., given n, there exists a unique set of distinct and positive nodes $\{x_j^k(t)\}_{k=1}^n$ and positive weights $\{a_j^k(t)\}_{k=1}^n$ such that

$$\int_0^1 p_j(x,t)v_i(x)dx = \sum_{k=1}^n a_j^k(t)v_i(x_j^k(t)) ,$$

for any polynomial v_i with degree not exceeding $2n - 1$.

Recall that equation (5.3) becomes,

$$\frac{\partial p}{\partial t} = -c\frac{\partial p}{\partial y} + \varepsilon\frac{\partial^2 p}{\partial y^2} - \alpha\frac{\partial p}{\partial x} + \frac{1}{2}\beta^2\frac{\partial^2 p}{\partial x^2} ,$$

Its weak form is,

$$\frac{d}{dt}(p_j, v_i) = \frac{c}{h_y}(p_{j-1} - p_j, v_i) + \frac{\varepsilon}{h_y^2}(p_{j+1} - 2p_j + p_{j-1}, v_i) + \alpha(p_j, \frac{\partial v_i}{\partial x})$$
$$+\frac{1}{2}\beta^2(p_j, \frac{\partial^2 v_i}{\partial x^2}) , \quad t > 0 .$$

Then

$$\frac{d}{dt}\sum_{k=1}^n a_j^k(t)v_i(x_j^k(t)) = \sum_{k=1}^n \Big[(\frac{c}{h_y} + \frac{\varepsilon}{h_y^2})a_{j-1}^k(t)v_i(x_{j-1}^k(t))$$
$$-(\frac{c}{h_y} + \frac{2\varepsilon}{h_y^2})a_j^k(t)v_i(x_j^k(t)) + \frac{\varepsilon}{h_y^2}a_{j+1}^k(t)v_i(x_{j+1}^k(t))\Big]$$
$$+\sum_{k=1}^n a_j^k(t)(L_j^* v_i)(x_j^k(t)) , \quad \forall t > 0 .$$

Thus, (5.14) has a set of solutions for all time. □

6 Convergence of Gauss-Galerkin approximation

6.1 Outline

To prove the convergence of the Gauss-Galerkin approximation, we first show the convergence of the semi-discrete difference approximation and derive an error estimate at the same time under certain regularity assumptions on the exact solution of the Fokker-Planck equation. Then, we show that the Gauss-Galerkin approximation is convergent in some weak sense.

6.2 Convergence of the semi-discrete difference approximation

Here, let us demonstrate that the semi-discrete approximation (5.7) is convergent when the number of grid points in the y direction, N_y, goes to infinity, i.e., when $h_y \to 0$. Using standard techniques in difference approximations, we first examine the truncation error of the approximation by substituting the exact solution $p = p(x,y,t)$ of (5.3) into the equation (5.7). For simplicity, we use $p(y_j)$ to denote $p(x,y_j,t)$. Then,

$$\begin{aligned}\frac{\partial}{\partial t}[p(y_j)] &= -c_j\frac{p(y_j)-p(y_{j-1})}{h_y} + \varepsilon_j\frac{p(y_{j+1})-2p(y_j)+p(y_{j-1})}{h_y^2} \\ &\quad +L_j p(y_j) + \tau_j, \end{aligned} \tag{6.1}$$

for $1 \leq j \leq N_y - 1$, $x > 0$ and $t > 0$, with $p(y_0) = p(y_{N_y}) = 0$, $\alpha_j = \alpha(x,y_j,t)$, $\beta_j = \beta(x,y_j,t)$ and

$$\begin{aligned}\tau_j = \quad & \varepsilon_j[\frac{\partial^2 p}{\partial y^2}(y_j) - \frac{p(y_{j+1})-2p(y_j)+p(y_{j-1})}{h_y^2}] \\ & -c_j[\frac{\partial p}{\partial y}(y_j) - \frac{p(y_j)-p(y_{j-1})}{h_y}].\end{aligned}$$

So,

$$\tau_j = \frac{c_j}{2}\frac{\partial^2 p}{\partial y^2}(y_j)h_y + O(h_y^2) ,$$

and

$$(\tau_j,\tau_j) \leq M h_y^2 , \tag{6.2}$$

for some constant $M > 0$, if we assume that the solution $p = p(x,y,t)$ of (5.3) is sufficiently smooth and its derivatives in y are uniformly bounded.

Now, define $e_j = p(y_j) - p_j$, $j = 1, 2, ..., N_y - 1$, then

$$\frac{\partial e_j}{\partial t} = -c_j\frac{e_j - e_{j-1}}{h_y} + \varepsilon_j\frac{e_{j+1} - 2e_j + e_{j-1}}{h_y^2} + L_j e_j + \tau_j , \tag{6.3}$$

and

$$e_0 = e_{N_y} = 0 . \tag{6.4}$$

Let us now state a convergence theorem.

Theorem 7 *Assume that in a given time interval $[0,T]$, the solution p of (5.3) is smooth and the coefficients satisfy the previously specified conditions. Then, in the given time interval, the solution $\{p_j, 1 \leq j \leq N_y - 1\}$ of the semi-discrete approximation converges to p as $N_y \to \infty$, i.e., $h_y \to 0$.*

Proof: We apply a standard energy type estimate to obtain error bounds on e_j's. Using (6.3) and the inequality (5.6), we get for any j,

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}(e_j, e_j) &= -c_j(\frac{e_j - e_{j-1}}{h_y}, e_j) + \varepsilon_j(\frac{e_{j+1} - 2e_j + e_{j-1}}{h_y^2}, e_j) \\
&\quad +(L_j e_j, e_j) + (\tau_j, e_j) \\
&\leq -\frac{c_j}{2h_y}(e_j - e_{j-1}\,,\, e_j - e_{j-1}) + \frac{c_j}{2h_y}[(e_{j-1}\,,\, e_{j-1}) - (e_j\,,\, e_j)] \\
&\quad +\frac{\varepsilon_j}{h_y^2}(e_{j+1} - e_j\,,\, e_j) - \frac{\varepsilon_j}{h_y^2}(e_j - e_{j-1}\,,\, e_j) \\
&\quad +\lambda(e_j\,,\, e_j) + \frac{1}{2}(e_j\,,\, e_j) + \frac{1}{2}(\tau_j\,,\, \tau_j)\,.
\end{aligned}$$

Here, we have also used the Holder's inequality

$$(\tau_j, e_j) \leq \frac{1}{2}(e_j\,,\, e_j) + \frac{1}{2}(\tau_j\,,\, \tau_j)\,.$$

Now, if we multiply h_y on both sides and sum over all j, we get

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\sum_{j=1}^{N_y-1}(e_j, e_j)h_y &\leq -\frac{c_j}{2}\sum_{j=1}^{N_y-1}(e_j - e_{j-1}\,,\, e_j - e_{j-1}) + \frac{c_j}{2}(e_0\,,\, e_0) \\
&- \frac{c_j}{2}(e_{N_y-1}\,,\, e_{N_y-1}) - \frac{\varepsilon_j}{h_y}\sum_{j=0}^{N_y-1}(e_{j+1} - e_j\,,\, e_{j+1} - e_j) \\
&+ \lambda_1\sum_{j=1}^{N_y-1}(e_j\,,\, e_j)h_y + \frac{1}{2}\sum_{j=1}^{N_y-1}(\tau_j\,,\, \tau_j)h_y \\
&\leq \lambda_1\sum_{j=1}^{N_y-1}(e_j\,,\, e_j)h_y + M_1 h_y^2\,,
\end{aligned}$$

for some positive constants λ_1 and M_1. Using the Gronwall inequality, we get for any $t \in [0,T]$,

$$\sum_{j=1}^{N_y-1}(e_j(t), e_j(t))h_y \leq M_2\sum_{j=1}^{N_y-1}(e_j(0), e_j(0))h_y + M_3 h_y^2 = M_3 h_y^2\,,$$

for some positive constants M_2 and M_3. Convergence then follows. □

6.3 Convergence of the Gauss-Galerkin approximations

We now show that the Gauss-Galerkin approximations, as defined in (5.14) and (5.15), converge to the solution of (5.7), (5.10). Combining with the convergence property of the semi-discrete approximation, this implies the convergence of the Gauss-Galerkin approximation to the solution of the original model equation (5.3).

Given a continuous $f(x)$ function on $[0,1]$, we would like to verify the convergence of the quadrature formula:

$$\Sigma_j^n(f) = \sum_{k=1}^{n} a_{j,n}^k(t) f(x_{j,n}^k(t))$$

to

$$I_j(f) = \int_0^1 f(x) p_j(x,t) dx$$

as $n \to \infty$ for any $t \in [0,T]$ and each $1 \le j \le N_y - 1$. The additional subscript n is used in order to emphasize the dependence of the weights and nodes on n, the number of *atoms*.

We start with the following lemma:

Lemma 4 *Let $\{\mu_j^n(t)\}$ be the Gauss-Galerkin measures defined as in (5.13)-(5.15). Then for any j, $1 \le j \le N_y - 1$, let l be any fixed positive integer and*

$$m_{j,n}^l(t) = \int_0^1 x^l d\mu_j^n(x,t)\,, \quad t \in [0,T]\,. \tag{6.5}$$

We have the set $\{m_{j,n}^l(t)\ : n \ge \frac{1}{2}(l+1)\}$ is uniformly bounded and equicontinuous in $t \in [0,T]$ for each l and j.

Proof: By the definition of $\{\mu_j^n(t)\}$, we have

$$m_{j,n}^l(t) = \sum_{k=1}^{n} a_{j,n}^k(t)(x_{j,n}^k(t))^l\,.$$

Using equation (5.14), we get for $0 \le l \le 2n-1$,

$$\begin{aligned} \frac{d}{dt} m_{j,n}^l(t) &= \left(\frac{c_j}{h_y} + \frac{\varepsilon_j}{h_y^2}\right) m_{j-1,n}^l(t) - \left(\frac{c_j}{h_y} + \frac{2\varepsilon_j}{h_y^2}\right) m_{j,n}^l(t) + \frac{\varepsilon_j}{h_y^2} m_{j+1,n}^l(t) \\ &\quad + \sum_{k=1}^{n} a_{j,n}^k(t) L_j^*(x^l)|_{x=x_{j,n}^k(t)}\,. \end{aligned}$$

Here again L_j^* is as defined in (5.9). Using the boundedness of the coefficients and the assumption that all weights $\{a_{j,n}^k(t)\}$ are positive in $[0,T]$, we get

$$\begin{aligned}\frac{d}{dt}m_{j,n}^l(t) &\leq (\frac{c_j}{h_y}+\frac{\varepsilon_j}{h_y^2})m_{j-1,n}^l(t)-(\frac{c_j}{h_y}+\frac{2\varepsilon_j}{h_y^2})m_{j,n}^l(t)+\frac{\varepsilon_j}{h_y^2}m_{j+1,n}^l(t)\\ &\quad +g^l(m_{j,n}^{l-2}(t),m_{j,n}^{l-1}(t),m_{j,n}^l(t))\ ,\ \ n\geq\frac{1}{2}(l+1)\,,\end{aligned}$$

where g^l is some linear function with nonnegative coefficients. Now, define

$$\begin{aligned}\frac{d}{dt}m_j^l(t) &= (\frac{c_j}{h_y}+\frac{\varepsilon_j}{h_y^2})m_{j-1}^l(t)-(\frac{c_j}{h_y}+\frac{2\varepsilon_j}{h_y^2})m_j^l(t)+\frac{\varepsilon_j}{h_y^2}m_{j+1}^l(t)\\ &\quad +g^l(m_j^{l-2}(t),m_j^{l-1}(t),m_j^l(t)) \qquad (6.6)\end{aligned}$$

and

$$m_j^l(0)=(q(x,y_j),x^l)\ ,\quad 1\leq j\leq N_y-1\ ,\ l\geq 0\ .$$

Using the comparison principle given in Section 6.1, we get

$$m_{j,n}^l(t)\leq m_j^l(t)\ ,\quad \forall\, t\in[0,T],\ 1\leq j\leq N_y-1\ ,\ \ n\geq\frac{1}{2}(l+1)\ .$$

Now, we estimate $\{m_j^l(t)\}$. Define

$$\vec{M}_j(t)=(m_j^0(t),m_j^1(t),\ \ldots,\ m_j^l(t))^{\mathrm{T}}\ .$$

Notice that g^l is some linear function with nonnegative coefficients, we may write (6.6) in matrix form as

$$\frac{d}{dt}\vec{M}_j(t)=(\frac{c_j}{h_y}+\frac{\varepsilon_j}{h_y^2})\vec{M}_{j-1}(t)-(\frac{c_j}{h_y}+\frac{2\varepsilon_j}{h_y^2})\vec{M}_j(t)+\frac{\varepsilon_j}{h_y^2}\vec{M}_{j+1}(t)+A\vec{M}_j(t)\ ,$$

for some nonnegative matrix A. Let $\vec{M}_j^1(t)=e^{-At}\vec{M}_j(t)$,

$$\vec{M}_j^1(t)=(\tilde{m}_j^0(t),\tilde{m}_j^1(t),\ \ldots,\ \tilde{m}_j^l(t))^{\mathrm{T}}\ ,$$

then

$$\frac{d}{dt}\vec{M}_j^1(t)=(\frac{c_j}{h_y}+\frac{\varepsilon_j}{h_y^2})\vec{M}_{j-1}^1(t)-(\frac{c_j}{h_y}+\frac{2\varepsilon_j}{h_y^2})\vec{M}_j^1(t)+\frac{\varepsilon_j}{h_y^2}\vec{M}_{j+1}^1(t)\ ,$$

and

$$\vec{M}_0^1(t)=\vec{M}_{N_y}^1(t)=\vec{0}\ .$$

Now, applying the discrete maximum principle, we get

$$\tilde{m}_j^k(t)\leq\max_{1\leq j\leq N_y-1}\tilde{m}_j^k(0)=\max_{1\leq j\leq N_y-1}(q(x,y_j),x^k)\leq C\ ,\quad \forall\, t\in[0,T],$$

where $1 \leq j \leq N_y - 1$, $n \geq (l+1)/2$ and C is some constant, depending on the initial condition $q = q(x,y)$ only. Meanwhile, since $\vec{M}_j(t) = e^{At}\vec{M}_j^1(t)$ and e^{At} is a matrix with bounded and nonnegative elements for $t \in [0,T]$, we get

$$m_j^k(t) \leq C_l \, , \quad \forall\, t \in [0,T],\ 1 \leq j \leq N_y - 1 \, , \ k \leq l \, ,$$

where C_l are some constant depending on the initial condition $q = q(x,y)$, the time interval $[0,T]$ and the value of l.

Thus, for each given l, we obtain that the set $\{m_{j,n}^l(t) \ : n \geq \frac{1}{2}(l+1)\}$ is uniformly bounded and the bound, in fact, is independent of j and N_y under proper assumptions on the initial condition.

We now consider the equicontinuity of $\{m_{j,n}^l(t) : n \geq \frac{1}{2}(l+1)\}$ in $[0,T]$. By the mean-value theorem

$$|m_{j,n}^l(t_2) - m_{j,n}^l(t_1)| = |\frac{d}{dt}m_{j,n}^l(\zeta)| \cdot |t_2 - t_1| \, , \ \zeta \in (t_1, t_2) \, .$$

From earlier discussion, we have

$$\begin{aligned}
|\frac{d}{dt}m_{j,n}^l(\zeta)| &\leq |(\frac{c_j}{h_y} + \frac{\varepsilon_j}{h_y^2})m_{j-1,n}^l(\zeta) - (\frac{c_j}{h_y} + \frac{2\varepsilon_j}{h_y^2})m_{j,n}^l(\zeta) \\
&\quad + \frac{\varepsilon_j}{h_y^2}m_{j+1,n}^l(\zeta)| + g^l(m_{j,n}^{l-2}(\zeta), m_{j,n}^{l-1}(\zeta), m_{j,n}^l(\zeta)) \\
&\leq (\frac{c_j}{h_y} + \frac{\varepsilon_j}{h_y^2})m_{j-1}^l(\zeta) + (\frac{c_j}{h_y} + \frac{2\varepsilon_j}{h_y^2})m_j^l(\zeta) \\
&\quad + \frac{\varepsilon_j}{h_y^2}m_{j+1}^l(\zeta) + g^l(m_j^{l-2}(\zeta), m_j^{l-1}(\zeta), m_j^l(\zeta)) \\
&\leq B(N_y) \cdot C_l + g^l(C_{l-2}, C_{l-1}, C_l) = M(l, N_y) \, .
\end{aligned}$$

Here, $B(N_y)$ is a constant that may depend on N_y and $M(l,N_y)$ is a constant, possibly depending on l and N_y. So,

$$|m_{j,n}^l(t_2) - m_{j,n}^l(t_1)| \leq M(l,N_y)|t_2 - t_1| \, , \ \forall\, t_1, t_2 \in [0,T] \, .$$

Therefore, for given l and N_y, we have the equicontinuity of $\{m_{j,n}^l(t) \ : n \geq \frac{1}{2}(l+1)\}$ in $[0,T]$. □

From the above lemma, we get

Corollary 4 *Given N_y, for each $1 \leq j \leq N_y - 1$, $\{x^l d\mu_{j,n}(t) \ : n \geq \frac{1}{2}(l+1)\}$ forms a set of uniformly bounded measures.*

Proof: First, for each l, since $\{m^l_{j,n}(t) \ : n \geq \frac{1}{2}(l+1), t \in [0,T]\}$ is uniformly bounded, we get

$$\sum_{i=1}^{n} a^i_{j,n}(t)(x^i_{j,n}(t))^l \leq M \ , \ \forall t \in [0,T] \, .$$

Since by assumption that all weights are positive for $t \in [0,T]$, for $f \in \mathbf{C}[0,1] \cap \mathbf{L}^\infty(0,1)$, we have

$$\begin{aligned} |(f, x^l d\mu^n_j)| &= \left| \int_0^1 x^l f(x) d\mu^n_j(x,t) \right| \\ &= \left| \sum_{i=1}^{n} a^i_{j,n}(t)(x^i_{j,n}(t))^l f(x^i_{j,n}(t)) \right| \\ &\leq \left| \sum_{i=1}^{n} a^i_{j,n}(t)(x^i_{j,n}(t))^l \right| \max_{1 \leq i \leq n} |f(x^i_{j,n}(t))| \\ &\leq M \|f\|_{L^\infty(0,1)} \, . \end{aligned}$$

So, for each j and l, $\{x^l d\mu^n_j\}$ forms a sequence of bounded positive measures. □

With the uniform bound on the sequence $\{m^l_{j,n}(t)\}$ as $n \to \infty$ and the equicontinuity, we now show the following.

Lemma 5 *Given N_y, there exists a sequence $\{k_n\}$, $k_n \to \infty$ and a sequence of functions $\{m^l_{j,*}(t)\}$ such that for every fixed integer l, we have*

$$\lim_{k_n \to \infty} m^l_{j,k_n}(t) = m^l_{j,*}(t) \ , \ \ \forall \, 1 \leq j \leq N_y - 1 \ , \ t \in [0,T] \, .$$

Proof: Using the Ascoli-Arzela theorem [16] and results of the previous lemma, for each l and j, we get a subsequence of $\{m^l_{j,n}(t) \ : n \geq \frac{1}{2}(l+1), t \in [0,T]\}$ that converges uniformly to a limit, denoted by $m^l_{j,*}(t)$. Taking intersections of these subsequences successively and applying a diagonal selection principle, we obtain a sequence $\{k_n\}$, $k_n \to \infty$ such that

$$\lim_{k_n \to \infty} m^l_{j,k_n}(t) = m^l_{j,*}(t) \ , \ \ \forall \, 1 \leq j \leq N_y - 1 \ , \ l \geq 0 \, , \ t \in [0,T] \, . \quad \square$$

Lemma 6 *Given N_y, for each $1 \leq j \leq N_y - 1$ and for any $t \in [0,T]$, the elements of the sequence $\{m^l_{j,*}(t)\}_{l=0}^{\infty}$ are the generalized moments of a nonnegative measure, i.e., there exists a nonnegative measure $P_{j,*}(x,t)$, such that for each $l \geq 0$ and each $1 \leq j \leq N_y - 1$, we have*

$$\int_0^1 x^l dP_{j,*}(x,t) = m^l_{j,*}(t) \ , \quad \forall t \in [0,T] \, .$$

Proof: The proof is based on the necessary and sufficient condition for the moment problem given in an earlier section. Since $\{m^l_{j,n}(t)\ ,l \geq 0\}$ are moments of the measure $d\mu^n_j$, we have for the related difference

$$\Delta^k m^l_{j,n}(t) \geq 0\ , \quad \forall k = 0,1,2,\cdots,\ \text{and}\ l \leq 2n-1\ .$$

Since

$$\lim_{k_n \to \infty} m^l_{j,k_n}(t) = m^l_{j,*}(t)\ ,$$

we get

$$\Delta^k m^l_{j,*}(t) \geq 0\ .$$

Thus, there exists a measure $dP_{j,*}(x,t)$ such that

$$m^l_{j,*}(t) \ = \int_0^1 x^l dP_{j,*}(x,t)\ . \quad \square$$

Corollary 5 *Given N_y, for each $1 \leq j \leq N_y - 1$ and any $f \in \mathbf{C}[0,1]$, we have*

$$\lim_{k_n \to \infty} \int_0^1 f(x) d\mu^{k_n}_j(x,t) = \int_0^1 f(x) dP_{j,*}(x,t)\ , \quad \forall\, t \in [0,T]\ . \tag{6.7}$$

Proof: Notice that for every fixed integer $l \geq 0$, we have

$$\lim_{k_n \to \infty} \int_0^1 x^l d\mu^{k_n}_j(x,t) = \int_0^1 x^l dP_{j,*}(x,t), \ \ \forall\, 1 \leq j \leq N_y - 1\ ,\ t \in [0,T]\ .$$

By the well-known Weierstrass approximation theorem, $\{x^l\}^\infty_{l=0}$ is dense in $\mathbf{C}[0,1]$, so for any continuous function f,

$$\lim_{k_n \to \infty} \int_0^1 f(x) d\mu^{k_n}_j(x,t) = \int_0^1 f(x) dP_{j,*}(x,t)\ , \quad \forall\, t \in [0,T] \ \ \square$$

Let $dP_{j,*}(x,t) = p_{j,*}(x,t)dx$, we now verify that $p_{j,*}(x,t)$ corresponds to the weak solution of the equation (5.7).

Lemma 7 *Given N_y, for each $1 \leq j \leq N_y - 1$ and for each l, we have*

$$\begin{aligned} m^l_{j,*}(t) - m^l_{j,*}(0) \ &= \ \int_0^t (c_j \frac{p_{j-1,*}(s) - p_{j,*}(s)}{h_y}, x^l) ds \\ &\quad + \int_0^t (\varepsilon_j \frac{p_{j+1,*}(s) - 2p_{j,*}(s) + p_{j-1,*}(s)}{h_y^2}, x^l) ds \\ &\quad + \int_0^t (p_{j,*}(s), L^*_j x^l) ds\ , \ \ \forall t \in [0,T]\ . \end{aligned}$$

Here, $p_{0,*} = p_{N_y,*} = 0$.

Proof: It is clear that

$$\begin{aligned} m^l_{j,k_n}(t) - m^l_{j,k_n}(0) &= \int_0^t c_j \frac{m^l_{j-1,k_n}(s) - m^l_{j,k_n}(s)}{h_y} ds \\ &+ \int_0^t \varepsilon_j \frac{m^l_{j+1,k_n}(s) - 2m^l_{j,k_n}(s) + m^l_{j-1,k_n}(s)}{h_y^2} ds \\ &+ \int_0^t \left[\int_0^1 L_j^*(x^l) d\mu_j^{k_n}(x,s)\right] ds , \end{aligned}$$

$$\int_0^1 L_j^*(x^l) d\mu_j^{k_n}(x,s) \le g^l(m^{l-2}_{j,k_n}(s), m^{l-1}_{j,k_n}(s) , m^l_{j,k_n}(s)) ,$$

for some linear function g^l with nonnegative coefficients which is defined before. So, by the Lebesgue dominated convergence theorem, we get from (6.7) and the previous lemmas on the convergence of the moments that

$$\begin{aligned} m^l_{j,*}(t) - m^l_j(0) = \int_0^t c_j \frac{m^l_{j-1,*}(s) - m^l_{j,*}(s)}{h_y} ds \\ + \int_0^t [\varepsilon_j \frac{m^l_{j+1,*}(s) - 2m^l_{j,*}(s) + m^l_{j-1,*}(s)}{h_y^2} + (p_{j,*}(x,s), L_j^*(x^l))] ds \end{aligned}$$

Thus, for any l,

$$\begin{aligned} (p_{j,*}(x,t), x^l) \quad - \quad (q(x,y_j,0), x^l) &= \int_0^t [\frac{c_j}{h_y}(p_{j-1,*}(x,s) - p_{j,*}(x,s), x^l)] ds \\ &+ \int_0^t [\frac{\varepsilon_j}{h_y^2}(p_{j+1,*}(x,s) - 2p_{j,*}(x,s) + p_{j-1,*}(x,s), x^l)] ds \\ &+ \int_0^t (p_{j,*}(x,s), L_j^*(x^l)) ds . \quad \square \end{aligned}$$

Since $\{x^l\}$ is dense in $\mathbf{C}[0,1]$, we see that $\{p_{j,*}\}$ are the weak solutions of (5.7). By the assumption that (5.7) has a unique set of solutions, we see that the limit of $\{\mu_j^{k_n}\}$ is, in fact, independent of the subsequence. It follows that the whole sequence is convergent. Thus, we have proved the following theorem.

Theorem 8 *Under all the previous assumptions, given N_y, for each $1 \le j \le N_y - 1$ and for any continuous function f in $[0,1]$, we have*

$$\lim_{n\to\infty} \int_0^1 f(x) d\mu_j^{k_n}(x,t) = \int_0^1 f(x) p_j(x,t) dx, \ \forall t \in [0,T] , \tag{6.8}$$

where $\{p_j\}$ is the set of weak solutions of (5.7).

Finally, we combine the convergence of semi-discrete approximation and the above result to get the convergence of the Gauss-Galerkin/finite difference approximation to the solution of the model problem (5.3).

Theorem 9 *Let p be the solution of (5.3). Then, under all the previous assumptions, for a given continuous function f in $[0,1]$, we have for any $t \in [0,T]$,*

$$\lim_{N_y \to \infty} \lim_{n \to \infty} \sum_{j=1}^{N_y-1} h_y \left| \int_0^1 f(x) d\mu_j^n(x,t) - \int_0^1 f(x) p(x,y_j,t) dx \right|^2 = 0 ,$$

where p is the solution of (5.3).

Proof: For any $\epsilon > 0$, there exists an N^* such that for any $N_y > N^*$, we have

$$\sum_{j=1}^{N_y-1} h_y \|p_j(x,t) - p(x,y_j,t)\|_{\mathbf{H}}^2 < \frac{\epsilon}{2\|f\|_{\mathbf{H}}^2} , \quad \forall t \in [0,T] ,$$

where $\mathbf{H} = \mathbf{L}^2(0,1)$. So,

$$\sum_{j=1}^{N_y-1} h_y \left| \int_0^1 f(x) p_j(x,t) dx - \int_0^1 f(x) p(x,y_j,t) dx \right|^2 < \frac{\epsilon}{2}.$$

For each N_y, there exists an n^* such that if $n > n^*$, then for each $1 \le j \le N_y - 1$, we have

$$\left| \int_0^1 f(x) d\mu_j^n(x,t) - \int_0^1 f(x) p_j(x,t) dx \right|^2 < \frac{\epsilon}{2} .$$

Hence,

$$\sum_{j=1}^{N_y-1} h_y \left| \int_0^1 f(x) d\mu_j^n(x,t) - \int_0^1 f(x) p_j(x,t) dx \right|^2 < \frac{\epsilon}{2} .$$

Therefore,

$$\sum_{j=1}^{N_y-1} h_y \left| \int_0^1 f(x) d\mu_j^n(x,t) - \int_0^1 f(x) p(x,y_j,t) dx \right|^2 < \epsilon .$$

Letting $\epsilon \to 0$, we get the conclusion in the theorem. □

7 Application to nonlinear oscillation of second order

7.1 Formulation

Consider equation (3.13), which models the problem of general second order nonlinear oscillation with a random force. We rewrite it in the following form

$$\ddot{x} + g(x, \dot{x}) = \dot{w}(t) .$$

The associated Fokker-Planck equation is

$$\frac{\partial p}{\partial t} = -y\frac{\partial p}{\partial x} + \frac{\partial}{\partial y}[g(x,y)p] + D\frac{\partial^2 p}{\partial y^2}\ , \quad x,y \in (-\infty,\infty) \tag{7.1}$$

where $y = \dot{x}$ and with initial data

$$p(x,y,0) = q(x,y)\ .$$

We assume that g and $\frac{\partial g}{\partial y}$ are bounded functions. In order to make the coefficients of equation (7.1) bounded in space (since $\dot{x}$ might not be bounded), we perform the following transformation

$$\tilde{p}(x,y,t) = p(x+yt,y,t)\ ,$$

thus

$$\frac{\partial p}{\partial t} = \frac{\partial \tilde{p}}{\partial t} - y\frac{\partial \tilde{p}}{\partial x}\quad , \quad -y\frac{\partial p}{\partial x} = -y\frac{\partial \tilde{p}}{\partial x},$$

$$\frac{\partial}{\partial y}[g(x,y)p] = \frac{\partial g}{\partial y}p + g\frac{\partial p}{\partial y} = \frac{\partial g}{\partial y}\tilde{p} + g\frac{\partial \tilde{p}}{\partial y} - gt\frac{\partial \tilde{p}}{\partial x},$$

$$D\frac{\partial^2 p}{\partial y^2} = D\frac{\partial}{\partial y}\left(\frac{\partial \tilde{p}}{\partial y} - t\frac{\partial \tilde{p}}{\partial x}\right) = D\frac{\partial^2 \tilde{p}}{\partial y^2} - 2tD\frac{\partial^2 \tilde{p}}{\partial x \partial y} + Dt^2\frac{\partial^2 \tilde{p}}{\partial x^2},$$

and equation (7.1) becomes:

$$\frac{\partial \tilde{p}}{\partial t} = -\tilde{g}t\frac{\partial \tilde{p}}{\partial x} + \frac{\partial}{\partial y}(\tilde{g}\tilde{p}) + D\frac{\partial^2 \tilde{p}}{\partial y^2} - 2tD\frac{\partial^2 \tilde{p}}{\partial x \partial y} + Dt^2\frac{\partial^2 \tilde{p}}{\partial x^2}\ .$$

The coefficients of the above equation are all bounded in a given time interval. Let us change variables similarly as in Section 5.2:

$$\tilde{x} = \frac{1}{2}(1 + cothx)\ \ ,\ \ \tilde{y} = \frac{1}{2}(1 + cothy)\ .$$

Without loss of generality, we write the equation in p, x and y,

$$\frac{\partial p}{\partial t} = a\frac{\partial p}{\partial x} + b\frac{\partial p}{\partial y} + cp + d\frac{\partial^2}{\partial x^2}(dp) + 2d\frac{\partial^2}{\partial x \partial y}(fp) + f\frac{\partial^2}{\partial y^2}(fp)\ . \tag{7.2}$$

Here, $x,y \in (0,1)$, and we assume that $a = a(x,y,t)$, $b = b(x,y,t)$, $c = c(x,y,t)$, $d = d(x,t)$ and $f = f(y)$ are all smooth functions of variables in $[0,1]\times[0,1]\times[0,T]$ and satisfy some further conditions we give later. Moreover, we assume that the initial condition is $p(x,y,0) = q(x,y)$, and $p = 0$ on the boundary.

7.2 Convergence of the semi-discrete difference approximation

The difference approximation may be set up by using standard techniques. For example, the unit interval [0, 1] in the x direction can be partitioned by a uniform grid. Let N_x be the total number of subintervals, $h_x = 1/N_x$ be the length of each subintervals and $x_j = jh_x$, be the grid points. The second order derivative in x may then be approximated by

$$\frac{\partial^2}{\partial x^2}(p(x_j)) \approx \frac{d_{j+1}p_{j+1} - 2d_jp_j + d_{j-1}p_{j-1}}{h_x^2}$$

with truncation error of the order $O(h_x^2)$. For the first order derivative in x we may use the center difference

$$\frac{p(x_{j+1}) - p(x_{j-1})}{2h_x}$$

which has truncation error of the order $O(h_x^2)$. With the above difference approximations, we get the following semi-discrete approximation of (7.2).

$$\begin{aligned} \frac{\partial p_j}{\partial t} &= a_j\frac{p_{j+1} - p_{j-1}}{2h_x} + b_j\frac{\partial p_j}{\partial y} + c_jp_j + d_j\frac{d_{j+1}p_{j+1} - 2d_jp_j + d_{j-1}p_{j-1}}{h_x^2} \\ &\quad + \frac{d_j}{h_x}(\frac{\partial(fp_{j+1})}{\partial y} - \frac{\partial(fp_{j-1})}{\partial y}) + f\frac{\partial^2}{\partial y^2}(fp_j)\,, \end{aligned} \tag{7.3}$$

for $1 \leq j \leq N_x - 1$, $y \in (0,1)$ and $t > 0$, with $p_0 = p_{N_x} = 0$. Here, $a_j = a(x_j, y, t)$, $b_j = b(x_j, y, t)$, $c_j = c(x_j, y, t)$, and $d_j = d(x_j, t)$.

Now let us demonstrate that this semi-discrete approximation is convergent when the number of grid points in the x direction N_x goes to infinity, i.e., when $h_x \to 0$. Using the similar techniques as in Section 6.2, we first examine the truncation error of the approximation by substituting the exact solution $p(x, y, t)$ of (7.2) into equation (7.3). For simplicity we use $p(x_j)$ to denote $p(x_j, y, t)$. Then,

$$\begin{aligned} \frac{\partial}{\partial t}(p(x_j)) &= a_j\frac{p(x_{j+1}) - p(x_{j-1})}{2h_x} + b_j\frac{\partial p(x_j)}{\partial y} + c_jp(x_j) \\ &\quad + d_j\frac{d_{j+1}p(x_{j+1}) - 2d_jp(x_j) + d_{j-1}p(x_{j-1})}{h_x^2} \\ &\quad + \frac{d_j}{h_x}(\frac{\partial[fp(x_{j+1})]}{\partial y} - \frac{\partial[fp(x_{j-1})]}{\partial y}) + f\frac{\partial^2}{\partial y^2}[fp(x_j)] + \tau_j\,, \end{aligned}$$

for $1 \leq j \leq N_x - 1$, $y \in (0,1)$ and $t > 0$, with $p(x_0) = p(x_{N_x}) = 0$, and

$$\tau_j = a_j\,[\frac{\partial}{\partial x}(p(x_j)) - \frac{p(x_{j+1}) - p(x_{j-1})}{2h_x}]$$

$$+d_j\,[\frac{\partial^2}{\partial x^2}(d_j p(x_j)) - \frac{d_{j+1}p(x_{j+1}) - 2d_j p(x_j) + d_{j-1}p(x_{j-1})}{h_x^2}]$$
$$+2d_j\,[\frac{\partial^2}{\partial x\partial y}(fp(x_j)) - \frac{\frac{\partial(fp(x_{j+1}))}{\partial y} - \frac{\partial(fp(x_{j-1}))}{\partial y}}{2h_x}]\,.$$

Thus,

$$\tau_j = a_j\frac{h_x^2}{12}\frac{\partial^3(p(x_j))}{\partial x^3} + d_j\frac{h_x^2}{12}\frac{\partial^4(d\,p(x_j))}{\partial x^4} + d_j\frac{h_x^2}{6}\frac{\partial^4(fp(x_j))}{\partial x^3\partial y} + O(h_x^3)\,,$$

and

$$(\tau_j\,,\,\tau_j) \leq Mh_x^4\,, \tag{7.4}$$

for some constant $M > 0$ if we assume that the solution $p = p(x,y,t)$ of (7.2) is sufficiently smooth and its derivatives in x are uniformly bounded.

Now, define $e_j = p(y_j) - p_j$, $j = 1,2,\ldots,N_x - 1$. Then

$$\begin{aligned}\frac{\partial e_j}{\partial t} &= a_j\frac{e_{j+1} - e_{j-1}}{2h_x} + b_j\frac{\partial e_j}{\partial y} + c_j e_j + d_j\frac{d_{j+1}e_{j+1} - 2d_j e_j + d_{j-1}e_{j-1}}{h_x^2}\\ &+\frac{d_j}{h_x}(\frac{\partial(fe_{j+1})}{\partial y} - \frac{\partial(fe_{j-1})}{\partial y}) + f\frac{\partial^2}{\partial y^2}(fe_j) + \tau_j\,,\end{aligned} \tag{7.5}$$

and

$$e_0 = e_{N_x} = 0\,. \tag{7.6}$$

Theorem 10 *Assume that in a given time interval* $[0,T]$, *the solution* p *of (7.2) is smooth and the coefficients satisfy the previously specified conditions. Then, in the given time interval, the solution* $\{p_j, 1 \leq j \leq N_x - 1\}$ *of the semi-discrete approximation converges to* p *as* $N_x \to \infty$, *i.e.,* $h_x \to 0$.

Proof: Using (7.5), we get for any j,

$$\begin{aligned}\frac{1}{2}\frac{\partial e_j^2}{\partial t} &= a_j e_j\frac{e_{j+1} - e_{j-1}}{2h_x} + b_j e_j\frac{\partial e_j}{\partial y} + c_j e_j^2\\ &+d_j e_j\frac{d_{j+1}e_{j+1} - 2d_j e_j + d_{j-1}e_{j-1}}{h_x^2}\\ &+\frac{d_j e_j}{h_x}(\frac{\partial(fe_{j+1})}{\partial y} - \frac{\partial(fe_{j-1})}{\partial y}) + fe_j\frac{\partial^2}{\partial y^2}(fe_j) + \tau_j e_j\,.\end{aligned}$$

Summing over all j, we will get

$$\frac{1}{2}\sum_{j=1}^{N_x-1}\frac{\partial e_j^2}{\partial t} = \sum_{j=1}^{N_x-1} a_j e_j\frac{e_{j+1} - e_{j-1}}{2h_x} + \sum_{j=1}^{N_x-1} b_j e_j\frac{\partial e_j}{\partial y} + \sum_{j=1}^{N_x-1} c_j e_j^2$$

$$
\begin{aligned}
&+ \sum_{j=1}^{N_x-1} d_j e_j \Big(\frac{d_{j+1}e_{j+1} - d_j e_j}{h_x^2} - \frac{d_j e_j - d_{j-1}e_{j-1}}{h_x^2}\Big) \\
&+ \sum_{j=1}^{N_x-1} \frac{d_j e_j}{h_x}\Big(\frac{\partial(fe_{j+1})}{\partial y} - \frac{\partial(fe_{j-1})}{\partial y}\Big) + \sum_{j=1}^{N_x-1} fe_j \frac{\partial^2}{\partial y^2}(fe_j) \\
&+ \sum_{j=1}^{N_x-1} \tau_j e_j \\
&= \sum_{j=1}^{N_x-1} \frac{a_j e_j e_{j+1}}{2h_x} - \sum_{j=1}^{N_x-1} \frac{a_{j+1} e_j e_{j+1}}{2h_x} + \frac{1}{2}\sum_{j=1}^{N_x-1} b_j \frac{\partial e_j^2}{\partial y} + \sum_{j=1}^{N_x-1} c_j e_j^2 \\
&+ \sum_{j=0}^{N_x-1} \frac{d_j e_j (d_{j+1}e_{j+1} - d_j e_j)}{h_x^2} - \sum_{j=0}^{N_x-1} \frac{d_{j+1}e_{j+1}(d_{j+1}e_{j+1} - d_j e_j)}{h_x^2} \\
&+ \sum_{j=1}^{N_x-1} \frac{d_{j-1}e_{j-1}}{h_x}\frac{\partial(fe_j)}{\partial y} - \sum_{j=1}^{N_x-1} \frac{d_{j+1}e_{j+1}}{h_x}\frac{\partial(fe_j)}{\partial y} \\
&+ \sum_{j=1}^{N_x-1} fe_j \frac{\partial^2(fe_j)}{\partial y^2} + \sum_{j=1}^{N_x-1} \tau_j e_j \, .
\end{aligned}
$$

Integrating over y, we have

$$
\begin{aligned}
\frac{1}{2}\sum_{j=1}^{N_x-1} \frac{d}{dt}\int_0^1 e_j^2 dy \;=\; & \sum_{j=1}^{N_x-1} \int_0^1 e_j e_{j+1} \frac{a_j - a_{j+1}}{2h_x} dy - \frac{1}{2}\sum_{j=1}^{N_x-1} \int_0^1 e_j^2 \frac{\partial b_j}{\partial y} dy \\
& + \sum_{j=1}^{N_x-1} \int_0^1 c_j e_j^2 dy - \sum_{j=0}^{N_x-1} \int_0^1 \frac{\left(d_{j+1}e_{j+1} - d_j e_j\right)^2}{h_x^2} dy \\
& - \sum_{j=1}^{N_x-1} \int_0^1 \frac{d_{j+1}e_{j+1} - d_{j-1}e_{j-1}}{h_x} \frac{\partial(fe_j)}{\partial y} dy \\
& - \sum_{j=1}^{N_x-1} \int_0^1 \left(\frac{\partial(fe_j)}{\partial y}\right)^2 dy + \sum_{j=1}^{N_x-1} \int_0^1 \tau_j e_j dy \\
\le\; & \frac{1}{4}\sum_{j=1}^{N_x-1} \int_0^1 \left|\frac{a_j - a_{j+1}}{h_x}\right| e_{j+1}^2 dy \\
& + \frac{1}{4}\sum_{j=1}^{N_x-1} \int_0^1 \left|\frac{a_j - a_{j+1}}{h_x}\right| e_j^2 dy - \frac{1}{2}\sum_{j=1}^{N_x-1} \int_0^1 e_j^2 \frac{\partial b_j}{\partial y} dy \\
& + \sum_{j=1}^{N_x-1} \int_0^1 c_j e_j^2 dy + \frac{1}{2}\sum_{j=1}^{N_x-1} \int_0^1 e_j^2 dy + \frac{1}{2}\sum_{j=1}^{N_x-1} \int_0^1 \tau_j^2 dy,
\end{aligned}
$$

since

$$
\begin{aligned}
& - \sum_{j=1}^{N_x-1} \int_0^1 \frac{d_{j+1}e_{j+1} - d_{j-1}e_{j-1}}{h_x} \frac{\partial(fe_j)}{\partial y} dy \\
\le\; & \frac{1}{4}\sum_{j=1}^{N_x-1} \int_0^1 \frac{\left(d_{j+1}e_{j+1} - d_{j-1}e_{j-1}\right)^2}{h_x^2} dy + \sum_{j=1}^{N_x-1} \int_0^1 \left(\frac{\partial(fe_j)}{\partial y}\right)^2 dy
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{4}\sum_{j=1}^{N_x-1}\int_0^1 \frac{[(d_{j+1}e_{j+1}-d_je_j)+(d_je_j-d_{j-1}e_{j-1})]^2}{h_x^2}dy \\
&\quad +\sum_{j=1}^{N_x-1}\int_0^1\left(\frac{\partial(fe_j)}{\partial y}\right)^2 dy \\
&\le \frac{1}{2}\sum_{j=1}^{N_x-1}\int_0^1\frac{(d_{j+1}e_{j+1}-d_je_j)^2}{h_x^2}dy+\frac{1}{2}\sum_{j=1}^{N_x-1}\int_0^1\frac{(d_je_j-d_{j-1}e_{j-1})^2}{h_x^2}dy \\
&\quad +\sum_{j=1}^{N_x-1}\int_0^1\left(\frac{\partial(fe_j)}{\partial y}\right)^2 dy \\
&\le \sum_{j=0}^{N_x-1}\int_0^1\frac{(d_{j+1}e_{j+1}-d_je_j)^2}{h_x^2}dy+\sum_{j=1}^{N_x-1}\int_0^1\left(\frac{\partial(fe_j)}{\partial y}\right)^2 dy\,.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\frac{1}{2}\sum_{j=1}^{N_x-1}\frac{d}{dt}\int_0^1 e_j^2dy &\le M_1\sum_{j=1}^{N_x-1}\int_0^1 e_j^2dy+M_2\sum_{j=1}^{N_x-1}\int_0^1\tau_j^2dy \\
&\le M_1\sum_{j=1}^{N_x-1}\int_0^1 e_j^2dy+M_3h_x^3,
\end{aligned}
$$

for some positive constants M_1, M_2 and M_3. Using the Gronwall inequality, we obtain for any $t \in [0,T]$,

$$
\sum_{j=1}^{N_x-1}\int_0^1 e_j^2(t)dy \le M_4\sum_{j=1}^{N_x-1}\int_0^1 e_j^2(0)dy+M_5h_x^3\,,
$$

for some positive constants M_4 and M_5. Convergence now follows. □

7.3 Convergence of the Gauss-Galerkin approximations

We rewrite equation (7.3) as

$$
\begin{aligned}
\frac{\partial p_j}{\partial t} &= a_j\frac{p_{j+1}-p_{j-1}}{2h_x}+c_jp_j+d_j\frac{d_{j+1}p_{j+1}-2d_jp_j+d_{j-1}p_{j-1}}{h_x^2} \\
&\quad +b_j\frac{\partial p_j}{\partial y}+\frac{d_j}{h_x}\left(\frac{\partial(fp_{j+1})}{\partial y}-\frac{\partial(fp_{j-1})}{\partial y}\right)+f\frac{\partial^2}{\partial y^2}(fp_j)\,.
\end{aligned}
$$

To solve (7.3) by Gauss-Galerkin method, let us formulate its weak form. Taking the inner product of a test function $v \in \mathbf{W}^{2,2}(0,1)\cap\mathbf{C}^2[0,1]$ with equation (7.3) and integrating by parts with suitable assumptions on the coefficients, we have, for $1 \le j \le N_x-1$,

$$
\frac{d}{dt}(p_j,v) = \frac{1}{2h_x}(a_j(p_{j+1}-p_{j-1}),v)+(c_jp_j,v)
$$

$$
\begin{aligned}
&+\frac{d_j}{h_x^2}(d_{j+1}p_{j+1} - 2d_jp_j + d_{j-1}p_{j-1}, v) - (p_j, \frac{\partial}{\partial y}(b_jv)) \\
&-\frac{d_j}{h_x}(fp_{j+1}, \frac{\partial v}{\partial y}) + \frac{d_j}{h_x}(fp_{j-1}, \frac{\partial v}{\partial y}) + (fp_j, \frac{\partial^2(fv)}{\partial y^2}) .
\end{aligned} \tag{7.7}
$$

Let us now approximate the probability measure p_jdy by

$$
d\mu_j^n(y,t) = \sum_{k=1}^{n} \alpha_{j,n}^k(t)\delta(y_{j,n}^k(t))dy \ , \ \ 1 \leq j \leq N_x - 1 \, , \tag{7.8}
$$

and choose $v_l(y) = y^l$ for $l = 0, 1, 2, ..., 2n - 1$. Then (7.7) implies for $0 \leq l \leq 2n - 1, 1 \leq j \leq N_x - 1$ and $t > 0$,

$$
\begin{aligned}
\frac{d}{dt}\sum_{k=1}^{n} \alpha_{j,n}^k(t)(y_j^k(t))^l \ = \ & \frac{1}{2h_x}\sum_{k=1}^{n} \alpha_{j+1,n}^k(t)\, a_j(y_{j+1}^k(t))\, (y_{j+1}^k(t))^l \\
& -\frac{1}{2h_x}\sum_{k=1}^{n} \alpha_{j-1,n}^k(t)\, a_j(y_{j-1}^k(t))\, (y_{j-1}^k(t))^l \\
& +\sum_{k=1}^{n} \alpha_{j,n}^k(t)\, c_j(y_j^k(t))\, (y_j^k(t))^l \\
& +\frac{d_jd_{j+1}}{h_x^2}\sum_{k=1}^{n} \alpha_{j+1,n}^k(t)\, (y_{j+1}^k(t))^l \\
& -\frac{2d_j^2}{h_x^2}\sum_{k=1}^{n} \alpha_{j,n}^k(t)\, (y_j^k(t))^l \\
& +\frac{d_{j-1}d_j}{h_x^2}\sum_{k=1}^{n} \alpha_{j-1,n}^k(t)\, (y_{j-1}^k(t))^l \\
& -\sum_{k=1}^{n} \alpha_{j,n}^k(t)\, (y_j^k(t))^l\, \frac{\partial b_j}{\partial y}(y_j^k(t)) \\
& -\sum_{k=1}^{n} l\, \alpha_{j,n}^k(t)\, b_j(y_j^k(t))\, (y_j^k(t))^{l-1} \\
& -\frac{d_j}{h_x}\sum_{k=1}^{n} lf(y_{j+1}^k(t))\ \alpha_{j+1,n}^k(t)\, (y_{j+1}^k(t))^{l-1} \\
& +\frac{d_j}{h_x}\sum_{k=1}^{n} lf(y_{j-1}^k(t))\ \alpha_{j-1,n}^k(t)\, (y_{j-1}^k(t))^{l-1} \\
& +\sum_{k=1}^{n} [f\frac{\partial^2 f}{\partial y^2}](y_j^k(t))\, \alpha_{j,n}^k(t)\, (y_j^k(t))^l \\
& +\sum_{k=1}^{n} 2l\, [f\frac{\partial f}{\partial y}](y_j^k(t))\, \alpha_{j,n}^k(t)\, (y_j^k(t))^{l-1} \\
& +\sum_{k=1}^{n} l(l-1)\, f(y_j^k(t))\, \alpha_{j,n}^k(t)\, (y_j^k(t))^{l-2}.
\end{aligned} \tag{7.9}
$$

Lemma 8 *Let $\{\mu_j^n(t)\}$ be the Gauss-Galerkin measures defined as in (7.8). Then for any j, $1 \le j \le N_x - 1$, let l be any fixed positive integer and*

$$m_{j,n}^l(t) = \int_0^1 y^l d\mu_j^n(y,t)\,, \quad t \in [0,T]\,.$$

We have the set $\{m_{j,n}^l(t)\ : n \ge \frac{1}{2}(l+1)\}$ is uniformly bounded and equicontinuous in $t \in [0,T]$ for each l and j.

Proof: By definition of $\{\mu_j^n(t)\}$, we have

$$m_{j,n}^l(t) = \sum_{k=1}^n a_{j,n}^k(t)(y_{j,n}^k(t))^l\,.$$

Using the boundedness of the coefficients and the assumption that the weights $\{\alpha_{j,n}^k(t)\}$ are positive and the nodes$\{y_j^k(t)\}_{k=1}^n$ are distinct in (0, 1) for $t \in [0,T]$, we get from equation (7.9),

$$\begin{aligned} \frac{d}{dt} m_{j,n}^l(t) &\le \lambda_1\, m_{j-1,n}^l(t) + \lambda_2\, m_{j,n}^l(t) + \lambda_3\, m_{j+1,n}^l(t) \\ &\quad + g^l(m_{j,n}^{l-2}(t), m_{j,n}^{l-1}(t), m_{j,n}^l(t))\,, \quad n \ge \frac{1}{2}(l+1)\,, \end{aligned}$$

where λ_1, λ_2 and λ_3 are some positive constants depending on the coefficients and h_x, and g^l is some linear function with nonnegative coefficients. Now, define

$$\frac{d}{dt} m_j^l(t) = \lambda_1\, m_{j-1}^l(t) + \lambda_2\, m_j^l(t) + \lambda_3\, m_{j+1}^l(t) + g^l(m_j^{l-2}(t), m_j^{l-1}(t), m_j^l(t)) \quad (7.10)$$

and

$$m_j^l(0) = (q(x_j, y), y^l)\,, \quad 1 \le j \le N_x - 1\,,\ l \ge 0\,.$$

Using the comparison principle given in Section 6.1, we get

$$m_{j,n}^l(t) \le m_j^l(t)\,, \quad \forall\, t \in [0,T],\ 1 \le j \le N_x - 1\,,\ n \ge \frac{1}{2}(l+1)\,.$$

We notice that g^l is some linear function with nonnegative coefficients. Equation (7.10) now becomes

$$\frac{d}{dt} m_j^l(t) = \tilde{c}_1\, m_{j-1}^l(t) + \tilde{c}_2\, m_j^l(t) + \tilde{c}_3\, m_{j+1}^l(t) + \tilde{c}_4\, m_j^{l-2}(t) + \tilde{c}_5\, m_j^{l-1}(t)\,, \quad (7.11)$$

where $\tilde{c}_1$, $\tilde{c}_2$, $\tilde{c}_3$, $\tilde{c}_4$ and $\tilde{c}_5$ are some positive constants depending on the coefficients and h_x. Let $M(t) = (M_{j,k})$ be an $(N_x - 1) \times (l+1)$ matrix, with element $M_{j,k}$ being equal to $m_j^k(t)$. We may rewrite (7.11) in matrix form $\frac{d}{dt} M = A\, M$ for some nonnegative matrix A. Then $M(t) = e^{At}\, M(0)$.

Thus

$$m_j^k(t) \le C_l \max_{1\le j\le N_x-1} m_j^k(0) = C_l \max_{1\le j\le N_x-1} (q(x_j, y), y^k) \le M_l, \ \forall\, t \in [0, T],$$

where $1 \le j \le N_x - 1\,, n \ge (l+1)/2$, and M_l is some constant which depends on h_x, the coefficients, the initial condition $q = q(x, y)$, the time interval $[0, T]$ and the value of l.

Therefore, for each given l, we obtain that the set $\{m_{j,n}^l(t) : n \ge \frac{1}{2}(l+1)\}$ is uniformly bounded.

The proof of the equicontinuity is similar to the proof of Lemma 4 in section 6.3. □

Now, we know that $\{m_{j,n}^l(t) : n \ge \frac{1}{2}(l+1)\}$ is uniformly bounded and equicontinuous in $[0, T]$ for each l and j , following the same arguments as in Section 6.3, first we will get the convergence of the Gauss-Galerkin approximation to the semi-discrete approximation. Then if we combine this result with the convergence of the semi-discrete approximation, we have the convergence of the Gauss-Galerkin/finite difference approximation to the solution of equation (7.2).

8 Computational algorithms

We discuss here some methods for the solution of Gauss-Galerkin/finite difference approximation.

8.1 Runge-Kutta and multi-step methods

The Gauss-Galerkin/finite difference approximation yields a system of $2n \times (N_y - 1)$ ordinary differential equations for the $n(N_y - 1)$ nodes and $n(N_y - 1)$ weights (see equation (7.9)). Standard numerical methods such as Runge-Kutta and multi-step methods may be used to solve such equations.

8.2 Predictor-corrector scheme via moments

Due to the special structure of the Gauss-Galerkin approximation and its close relationship with the Gauss quadrature and the moment problems, a alternative way of solving the ordinary differential equation systems has been studied at least for the one-dimensional problem. Here, we present some stability conditions which show why such a procedure is feasible and how it may be applied to two-dimensional cases as well.

First we rewrite the Gauss-Galerkin/finite difference approximation at $t = t_0$ as

$$\frac{d}{dt}m^i_{j,n}(t) = (\frac{c_j}{h_y} + \frac{\varepsilon_j}{h_y^2})m^i_{j-1,n}(t) - (\frac{c_j}{h_y} + \frac{2\varepsilon_j}{h_y^2})m^i_{j,n}(t) + \frac{\varepsilon_j}{h_y^2}m^i_{j+1,n}(t)$$

$$+\sum_{k=1}^{n} a^k_{j,n}(t)(L_j^* v_i)(x^k_{j,n}(t)), \tag{8.1}$$

where L_j^* is defined as in (5.9). Note that $v_i(x) = x^i$, $0 \le i \le 2n-1$. The moments are defined by

$$m^i_{j,n}(t) = \sum_{k=1}^{n} a^k_{j,n}(t)(x^k_{j,n}(t))^i \ , \ \ \forall 0 \le i \le 2n-1 \ .$$

The proposed computational procedure is as follows:

1 Given $\{m^i_{j,n}(t_0)\}_{i=0}^{2n-1}$, compute $\{a^k_{j,n}(t_0)\}$ and $\{x^k_{j,n}(t_0)\}$.

2 Predict $\{m^i_{j,n}(t_0 + \Delta t)\}$ for $i = 0, 1, \ldots, 2n-1$. Evaluate the right hand side of (8.1) at $t = t_0$ and using the forward difference to approximate the time derivative to predict $\{m^i_{j,n}(t_0 + \Delta t)\}$ for $i = 0, 1, \ldots, 2n-1$.

3 With $\{m^i_{j,n}(t_0+\Delta t)\}_{i=0}^{2n-1}$, compute the weights $\{a^k_j(t_0+\Delta t)\}$ and nodes $\{x^k_j(t_0+\Delta t)\}$.

4 Evaluate the right hand side of (8.1) using the new moments $\{m^i_{j,n}(t_0 + \Delta t)\}$, weights $\{a^k_j(t_0 + \Delta t)\}$ and nodes $\{x^k_j(t_0 + \Delta t)\}$. Find the new updates of $\{m^i_{j,n}(t_0 + \Delta t)\}$ for $i = 0, 1, \ldots, 2n-1$.

5 Repeat steps 3 and 4 above or march to the next time-step.

One may consider step 2 as a predictor and steps 3 and 4 as a corrector. Obviously, the core of this procedure is in step 3. The actual implementation of step 3 can be found in section 9.1 (see also [6]). Here we present the next theorem to substantiate the given procedure.

Theorem 11 *Given a set of positive weights $\{a^k_{j,n}(t_0)\}_{k=1}^{n}$, a set of distinct nodes $\{x^k_{j,n}(t_0)\}_{k=1}^{n}$ in $(0, 1)$ and the corresponding moments $\{m^i_{j,n}(t_0)\}_{i=0}^{2n-1}$, there exists a small enough Δt such that if $\{m^i_{j,n}(t_0 + \Delta t)\}_{i=0}^{2n-1}$ are computed from step 2, then* *there exists a set of positive weights* $\{a^k_{j,n}(t_0+\Delta t)\}^n_{k=1}$, *and a set of distinct nodes* $\{x^k_{j,n}(t_0+\Delta t)\}^n_{k=1} \subset (0,1)$ *such that for any* $0 \le i \le 2n-1$,

$$m^i_{j,n}(t_0+\Delta t) = \sum_{k=1}^{n} a^k_{j,n}(t_0+\Delta t)(x^k_{j,n}(t_0+\Delta t))^i \ .$$

Proof : Define $m^i_{j,n}(t) = 0$ for any $i \ge 2n$.

We have

$$\Delta^k m^l_{j,n}(t) \ge 0 \ , \quad k,l = 0,1,2,\ldots,$$

where $\Delta^k m^l_{j,n}(t)$ is defined for each k from $\{m^l_{j,n}(t)\}$ by (5.17). So, if Δt is small enough, by the continuity property we get

$$\Delta^k m^l_{j,n}(t_0+\Delta t) \ge 0 \ , \quad k,l = 0,1,2,\ldots.$$

By the theorem on the Hausdorff moment problem, we obtain the existence of positive weights $\{a^k_{j,n}(t_0+\Delta t)\}^n_{k=1}$ and distinct nodes $\{x^k_{j,n}(t_0+\Delta t)\}^n_{k=1} \subset (0,1)$ such that

$$m^i_{j,n}(t_0+\Delta t) = \sum_{k=1}^{n} a^k_{j,n}(t_0+\Delta t)(x^k_{j,n}(t_0+\Delta t))^i \ , \quad \forall 0 \le i \le 2n-1. \square$$

9 Implementation

In this section, we discuss various issues in the practical implementation of our Gauss-Galerkin/finite difference algorithms. First of all, a general description of the algorithm is given. Numerical results on a number of test problems are then presented. For the model test problem 1, we compare the numerical solution based on the Gauss-Galerkin/finite difference method with that based on the conventional two-dimensional finite difference method. It is demonstrated that the Gauss-Galerkin/finite difference methods are superior to the two-dimensional finite difference methods in achieving high accuracy. For test problem 2, the exact solution is known analytically. This, therefore, allows us to compare the numerical solution with the exact solution. For the third problem, a simple second order random linear oscillation problem is formulated in terms of its Fokker-Planck equation and the partial differential equation is solved by the Gauss-Galerkin/finite difference method. Graphics plots are provided to show various behaviors of the exact solution and the numerical solution.

9.1 Description of the algorithm

The Gauss-Galerkin/finite difference method uses finite difference method in the x variable and Gauss-Galerkin in the y variable. Its implementation consists mainly of the following features (see the appendix for a flow chart).

1 **Initial condition**

Use Simpson's rule to get integrals for the moments of the initial condition.

2 **Difference in the x-variable**

At each time step, center difference is used in x variable.

3 **Finding the associated symmetric tridiagonal matrix**

An important step is to compute the inner product of the associated family of orthogonal polynomials from the corresponding moments. This replies on an expansion of the product of two polynomials in terms of a monomial. We implement the procedure by the algorithm given in Figure 1, which is different from the ones given in theses of Abrouk and HajJafar [1, 6].

4 **Finding the weights and nodes**

We use the following theorem [14] to get the weights and nodes from a tridiagonal matrix

Theorem 12 *Let the polynomials $\{p_j\}$ be defined by the recursions*

$$\begin{aligned} p_0(x) &= 1, \\ p_{i+1}(x) &= (x-\delta_{i+1})p_i(x) - \gamma_{i+1}^2 p_{i+1}(x) \quad \textit{for } i \geq 0, \end{aligned}$$

where $p_{-1}(x) = 0$ and

$$\begin{aligned} \delta_{i+1} &= (xp_i, p_i)/(p_i, p_i) \\ \gamma_{i+1}^2 &= \begin{cases} 0 & \textit{for } i = 0, \\ (p_i, p_i)/(p_{i-1}, p_{i-1}) & \textit{for } i \geq 1. \end{cases} \end{aligned}$$

Algorithm to form the tridiagonal matrix:

```
Input:  ${m_j, 0 \le j \le 2n-1}$ :  the moments
Output:  ${\delta_j}$ :  the diagonal elements,
         ${\gamma_j}$:  the off-diagonal elements.
Begin
    Initialize
        $\alpha_0 = 0,\ \beta_0 = 0,\ c_0 = 1,\ d_0 = -\delta_1,$
        $\alpha_1 = m_0,\ \beta_1 = m_1,\ \delta_1 = \beta_1/\alpha_1,\ d_1 = 1,$
        $\gamma_1 = (m_2/m_0) - (m_1/m_0)^2,\ \alpha_2 = \delta_1^2 m_0 - 2\delta_1 m_1 + m_2,$
        $\beta_2 = \delta_1^2 m_1 - 2\delta_1 m_2 + m_3\ ,\ \delta 2 = \beta_2/\alpha_2.$
    For $j = 2 : n-1$, let
        $e_0 = -\delta_j d_0 - \gamma_{j-1} c_0$
        For $k = 1 : j-2$, let
            $e_k = d_{k-1} - \delta_j d_k - \gamma_{j-1} c_k$
        End for
        $e_{j-1} = d_{j-2} - \delta_j d_{j-1}\ ,\ e_j = d_{j-1},$
        $\alpha_{j+1} = \sum_{i=0}^{j} \sum_{k=0}^{j} e_k e_i m_{k+i},$
        $\beta_{j+1} = \sum_{i=0}^{j} \sum_{k=0}^{j} e_k e_i m_{k+i+1},$
        $\delta_{j+1} = \beta_{j+1}/\alpha_{j+1}\ ,\ \gamma_j = \alpha_{j+1}/\alpha_j.$
        For $i = 0 : j+1$, let
            $c_i = d_i\ ,\ d_i = e_i\ .$
        End for
    End for
    For $j = 1 : n-1$, let
        $\gamma_j = \sqrt{\gamma_j}$
    End for
End
```

Figure 1: Algorithm for the construction of the tridiagonal matrix.

Let the matrix

$$J_j = \begin{pmatrix} \delta_1 & \gamma_2 & & & & & \\ \gamma_2 & \delta_2 & \gamma_3 & & & & \\ & \cdot & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \gamma_j \\ & & & & & \gamma_j & \delta_j \end{pmatrix}.$$

Then

(i). The roots x_i, $i = 1, \ldots, n$, of the n-th orthogonal polynomial p_n are the eigenvalues of the tridiagonal matrix J_n.

(ii). Let $v^{(i)} = (v_1^{(i)}, \ldots, v_n^{(i)})^T$ be an eigenvector of J_n for the eigenvalue x_i. Suppose $v^{(i)}$ is scaled in such a way that

$$(v^{(i)})^T v^{(i)} = (p_0, p_0) = \int_a^b \omega(x)dx.$$

Then the weights are given by $\omega_i = (v_1^{(i)})^2, \quad i = 1, \ldots, n.$

5 **Solving the system of ordinary differential equations**
Predictor-corrector type methods are used to update the moments for the next time step.

9.2 Model test problem 1

Numerical test is performed first for the following problem:

$$u_t = (y^2(1-y)^2 u)_{yy} + u_{xx},$$

with boundary conditions

$$u(x,0,t) = u(x,1,t) = 0\,,$$

$$u_x(0,y,t) = u_x(1,y,t) = 0\,,$$

and initial condition

$$u(x,y,0) = \pi cos^2(\frac{\pi x}{2}) sin(\pi y).$$

The purpose of this test is to show the superior convergence properties of the Gauss-Galerkin approach, as compared with those of the conventional finite difference method.

The Gauss-Galerkin/finite difference method uses finite difference in the x variable and Gauss-Galerkin in the y variable. We use five grid points in the x direction while three nodes in the y direction for each grid point in x. The locations of the nodes ($\diamond$, $\triangle$, and $\star$) corresponding to the grid point $x = 0.5$ are plotted in Figure 2 and the corresponding weights are given in Table 1. We can see that the nodes with positive weights have clearly moved to the boundary of the interval. Similar behavior is observed for all other grid points in x.

time	weights for $\diamond$	weights for $\triangle$	weights for $\star$
0.0	0.298631	0.712243	0.298143
.05	0.289834	0.62341	0.28954
0.1	0.288038	0.561437	0.287862
.45	0.309076	0.395482	0.309062
1.0	0.356011	0.288322	0.35601
2.0	0.411019	0.177964	0.411019
3.0	0.44199	0.11602	0.44199
4.0	0.46051	0.078981	0.46051
5.0	0.472101	0.055799	0.472101
6.0	0.479639	0.040723	0.479639
7.0	0.48471	0.030581	0.48471
8.0	0.488225	0.02355	0.488225
9.0	0.49073	0.018541	0.49073
10.	0.492558	0.014885	0.492558
12.	0.494962	0.010077	0.494962
14.	0.496402	0.007196	0.496402
16.	0.49732	0.00536	0.49732
18.	0.497936	0.004128	0.497936
20.	0.498367	0.003267	0.498367
25.	0.499004	0.001992	0.499004
30.	0.499324	0.001352	0.499324

Table 1: Gauss-Galerkin/finite difference solution: Three nodes in y direction. $\diamond$, $\triangle$, and $\star$ are symbols for nodes. (see Figure 2)

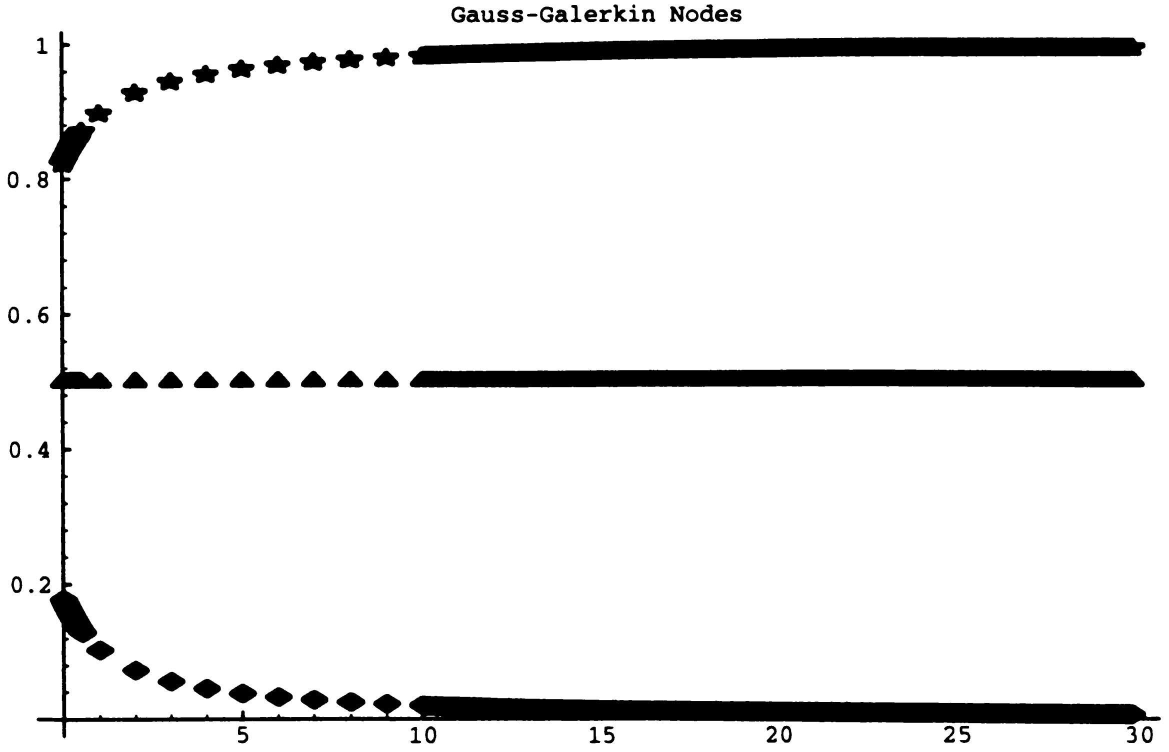


Figure 2: Gauss-Galerkin/finite-difference solution: Movement of nodes to the boundary 0.0 and 1.0 as $t \to \infty$.

At time 30.0, the numerical solution is approaching steady state based on the tolerance we have chosen. The solution becomes uniform in x while it approaches zero in the interior and piles up at the boundary. The nodes and weights at the steady state are given in Table 2 and they are the same for all grid points in x.

y-nodes	0.006261	0.500000	0.993739
weights	0.499324	0.001352	0.499324

Table 2: Gauss-Galerkin/finite difference solution: Three nodes in y direction.

In Table 3, the zero-th through the fifth order moments of the Gauss-Galerkin/finite difference solution at the grid point $x = 0.5$ are given at various times.

Next, conventional two-dimensional finite difference methods in both directions were implemented. In contrast to the high accuracy in estimating moments using Gauss-Galerkin/finite difference methods, the standard finite difference method in two dimensions suffers inaccuracy due to the piling up of solution near the boundary. The related results are presented in Tables 4 and 5. Table 4 consists of the moments computed using 20 grid points, while Table 5 consists of the moments computed using 40 grid points. In addition, the graphs for 20 grid points and 40 grid points at $t = 0.0$ and $t = 30.0$ are shown in Figures 3 and 4.

By comparing the test results, we can see that Gauss-Galerkin/finite difference methods have remarkably high accuracy, and they are much more suitable for solving equations with degenerate diffusion coefficients than conventional two-dimensional finite difference methods. This is mainly because Gauss-Galerkin methods can capture the Dirac-delta measure like singular behavior of the solution very efficiently.

9.3 Model test problem 2

Consider the second order Langevin equation:

$$\ddot{x} + 2\dot{x} = \dot{\omega}(t) \ ,$$

its corresponding Fokker-Planck equation is given as:

$$u_t = -(yu)_x + (2yu)_y + \frac{1}{2}u_{yy} \ ,$$

according to the previous sections.

Time	m0	m1	m2	m3	m4	m5
0.0	1.00000	0.50000	0.29736	0.19604	0.13846	0.10275
1.0	1.00000	0.50000	0.36224	0.29336	0.24855	0.21579
2.0	1.00000	0.50000	0.40020	0.35030	0.31524	0.28760
3.0	1.00000	0.50000	0.42430	0.38645	0.35832	0.33504
4.0	1.00000	0.50000	0.44043	0.41064	0.38752	0.36772
5.0	1.00000	0.50000	0.45168	0.42753	0.40811	0.39106
6.0	1.00000	0.50000	0.45982	0.43973	0.42313	0.40826
7.0	1.00000	0.50000	0.46589	0.44883	0.43440	0.42130
8.0	1.00000	0.50000	0.47053	0.45579	0.44310	0.43142
9.0	1.00000	0.50000	0.47416	0.46124	0.44994	0.43945
10.	1.00000	0.50000	0.47707	0.46560	0.45544	0.44593
11.	1.00000	0.50000	0.47943	0.46915	0.45994	0.45126
12.	1.00000	0.50000	0.48139	0.47208	0.46367	0.45570
13.	1.00000	0.50000	0.48303	0.47454	0.46680	0.45944
14.	1.00000	0.50000	0.48441	0.47662	0.46947	0.46264
15.	1.00000	0.50000	0.48560	0.47841	0.47176	0.46539
16.	1.00000	0.50000	0.48663	0.47995	0.47375	0.46779
17.	1.00000	0.50000	0.48753	0.48130	0.47549	0.46989
18.	1.00000	0.50000	0.48833	0.48249	0.47702	0.47174
19.	1.00000	0.50000	0.48903	0.48354	0.47838	0.47339
20.	1.00000	0.50000	0.48965	0.48448	0.47960	0.47487

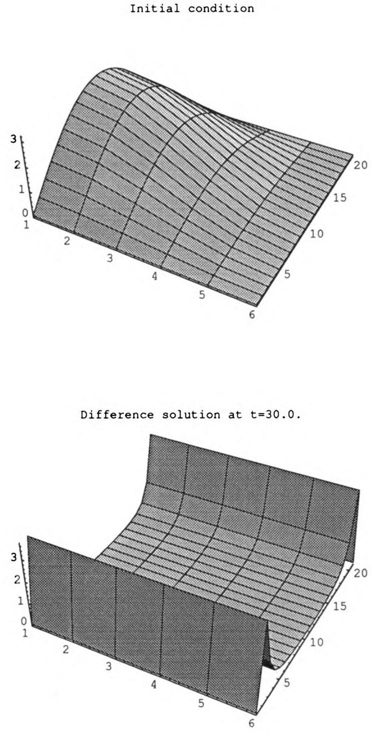
Table 3: Gauss-Galerkin/finite difference solution: Three nodes in y direction.

Time	m0	m1	m2	m3	m4	m5
1.0	0.99796	0.49898	0.35708	0.28613	0.24091	0.20855
2.0	0.99794	0.49897	0.38500	0.32801	0.28930	0.25972
3.0	0.99794	0.49897	0.39766	0.34700	0.31127	0.28300
4.0	0.99794	0.49897	0.40338	0.35559	0.32121	0.29354
5.0	0.99794	0.49897	0.40598	0.35948	0.32571	0.29831
6.0	0.99794	0.49897	0.40715	0.36124	0.32775	0.30047
7.0	0.99794	0.49897	0.40768	0.36203	0.32867	0.30144
8.0	0.99794	0.49897	0.40792	0.36239	0.32908	0.30188
9.0	0.99794	0.49897	0.40803	0.36256	0.32927	0.30208
10.	0.99794	0.49897	0.40808	0.36263	0.32936	0.30217
11.	0.99794	0.49897	0.40810	0.36266	0.32940	0.30221
12.	0.99794	0.49897	0.40811	0.36268	0.32941	0.30223
13.	0.99794	0.49897	0.40811	0.36269	0.32942	0.30224
14.	0.99794	0.49897	0.40812	0.36269	0.32943	0.30225
15.	0.99794	0.49897	0.40812	0.36269	0.32943	0.30225
16.	0.99794	0.49897	0.40812	0.36269	0.32943	0.30225
17.	0.99794	0.49897	0.40812	0.36269	0.32943	0.30225
18.	0.99794	0.49897	0.40812	0.36269	0.32943	0.30225
19.	0.99794	0.49897	0.40812	0.36269	0.32943	0.30225
20.	0.99794	0.49897	0.40812	0.36269	0.32943	0.30225

Table 4: Difference solution: 20 grid points in y direction.

Time	m0	m1	m2	m3	m4	m5
1.0	0.99949	0.49974	0.36190	0.29298	0.24936	0.21838
2.0	0.99949	0.49974	0.39755	0.34645	0.31214	0.28622
3.0	0.99949	0.49974	0.41728	0.37605	0.34705	0.32418
4.0	0.99949	0.49974	0.42820	0.39243	0.36640	0.34524
5.0	0.99949	0.49974	0.43424	0.40149	0.37710	0.35689
6.0	0.99949	0.49974	0.43759	0.40651	0.38303	0.36335
7.0	0.99949	0.49974	0.43944	0.40929	0.38631	0.36692
8.0	0.99949	0.49974	0.44047	0.41083	0.38813	0.36890
9.0	0.99949	0.49974	0.44103	0.41168	0.38913	0.36999
10.	0.99949	0.49974	0.44135	0.41215	0.38969	0.37060
11.	0.99949	0.49974	0.44152	0.41241	0.39000	0.37093
12.	0.99949	0.49974	0.44162	0.41256	0.39017	0.37112
13.	0.99949	0.49974	0.44167	0.41264	0.39026	0.37122
14.	0.99949	0.49974	0.44170	0.41268	0.39031	0.37128
15.	0.99949	0.49974	0.44172	0.41270	0.39034	0.37131
16.	0.99949	0.49974	0.44173	0.41272	0.39036	0.37133
17.	0.99949	0.49974	0.44173	0.41272	0.39037	0.37134
18.	0.99949	0.49974	0.44173	0.41273	0.39037	0.37134
19.	0.99949	0.49974	0.44174	0.41273	0.39038	0.37134
20.	0.99949	0.49974	0.44174	0.41273	0.39038	0.37135

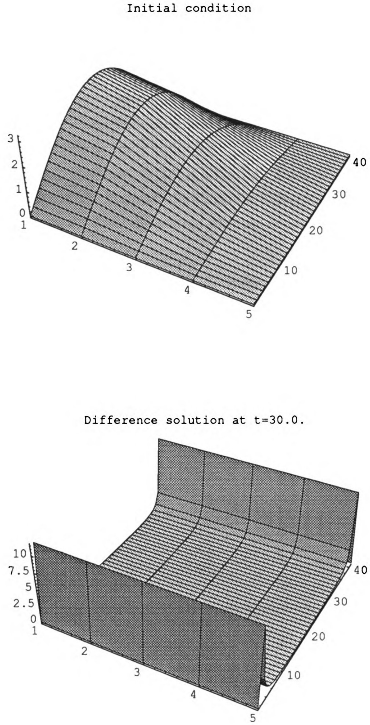
Table 5: Difference solution: 40 grid points in y direction.



Initial condition

Difference solution at t=30.0.

Figure 3: Difference solution at $t = 0.0$ and 30.0: 20 grid points in y.



Initial condition

Difference solution at t=30.0.

Figure 4: Difference solution at $t = 0.0$ and 30.0: 40 grid points in y.

In this numerical test, we solve the following problem

$$u_t = -(yu)_x + (2yu)_y + \frac{1}{2}u_{yy} + f(x,y,t)$$

where f is constructed in such a way that the exact solution to the above equation can be formulated analytically as $u(x,y,t) = X(x)e^{-y^2-t}$. If the interval for x is chosen to be $[-x_1, x_1]$, then $X(x)$ is $(x_1 - x)(x_1 + x)$. With the exact solution, one can compare values of the exact moments with the numerical moments computed by using the Gauss-Galerkin/finite difference method (finite difference method in the x variable and Gauss-Galerkin in the y variable).

Numerical results are presented in Table 6 through Table 11. The zero-th to the seventh order moments of the Gauss-Galerkin/finite difference solution for $x = 0.0$ are given in Tables 6 and 7. Their differences with the corresponding exact moments are shown in Tables 8 and 9.

In Table 10, the maximum of the *time difference*, i.e. $\frac{m_n(0.0,t+\Delta t)-m_n(0.0,t)}{\Delta t}$, of the numerical moments are given. This is used to determine if the solution is approaching steady state. As the time reaches 13.056, the algorithms stops as the solution is near steady state, based on the tolerance we have chosen.

At the steady state, as the weights approach zero, the distribution of the nodes are given in Table 11. The symmetry of the nodal distribution with respect to the origin is also a property of the exact solution. The nodes are almost equally spaced.

9.4 Model test problem 3

As in the previous problem, the Fokker-Planck equation corresponding to the second order equation

$$\ddot{x} + 2\dot{x} = \dot{\omega}(t) ,$$

is known as:

$$u_t = -(yu)_x + (2yu)_y + \frac{1}{2}u_{yy} .$$

We may consider the initial condition

$$u(x,y,0) = \delta(x)\delta(y) ,$$

where δ is the Dirac-delta measure, or more generally, an initial density distribution of the type:

$$u(x,y,0) = u_0(x,y).$$

Time	m0	m1	m2	m3
0.001	24.9875000	0.0000000	12.4937500	0.0000000
0.251	19.4590679	0.0000000	9.7295339	0.0000000
0.500	15.1613706	0.0000000	7.5806848	0.0000000
0.501	15.1537899	0.0000000	7.5768944	0.0000000
0.750	11.8010456	0.0000000	5.9005217	0.0000000
1.000	9.1900891	0.0000000	4.5950431	0.0000000
1.250	7.1568013	0.0000000	3.5783992	0.0000000
1.500	5.5733742	0.0000000	2.7866858	0.0000000
1.750	4.3402771	0.0000000	2.1701375	0.0000000
2.000	3.3800002	0.0000000	1.6899992	0.0000000
2.250	2.6321825	0.0000000	1.3160905	0.0000000
2.500	2.0508434	0.0000000	1.0254210	0.0000000
3.001	1.2431225	0.0000000	0.6215607	0.0000000
4.001	0.4572057	0.0000000	0.2286024	0.0000000
5.000	0.1681552	0.0000000	0.0840773	0.0000000
6.000	0.0618460	0.0000000	0.0309227	0.0000000
7.000	0.0227467	0.0000000	0.0113731	0.0000000
7.500	0.0138021	0.0000000	0.0069008	0.0000000
7.500	0.0137952	0.0000000	0.0068973	0.0000000
10.001	0.0011322	0.0000000	0.0005659	0.0000000
12.501	0.0000933	0.0000000	0.0000465	0.0000000
13.056	0.0000537	0.0000000	0.0000267	0.0000000

Table 6: Gauss-Galerkin/finite difference solution: The values of numerical moments of order zero to three along $x = 0.0$.

Time	m4	m5	m6	m7
0.001	18.7406250	0.0000000	46.8515625	0.0000000
0.251	14.5943008	0.0000000	36.4988288	0.0000000
0.500	11.3710262	0.0000000	28.4400844	0.0000000
0.501	11.3653407	0.0000000	28.4258655	0.0000000
0.750	8.8507808	0.0000000	22.1368620	0.0000000
1.000	6.8925628	0.0000000	17.2391339	0.0000000
1.250	5.3675972	0.0000000	13.4250108	0.0000000
1.500	4.1800274	0.0000000	10.4547550	0.0000000
1.750	3.2552051	0.0000000	8.1416623	0.0000000
2.000	2.5349978	0.0000000	6.3403366	0.0000000
2.250	1.9741349	0.0000000	4.9375505	0.0000000
2.500	1.5381308	0.0000000	3.8470514	0.0000000
3.001	0.9323406	0.0000000	2.3318964	0.0000000
4.001	0.3429033	0.0000000	0.8576424	0.0000000
5.000	0.1261156	0.0000000	0.3154303	0.0000000
6.000	0.0463839	0.0000000	0.1160114	0.0000000
7.000	0.0170595	0.0000000	0.0426676	0.0000000
7.500	0.0103510	0.0000000	0.0258890	0.0000000
7.500	0.0103459	0.0000000	0.0258761	0.0000000
10.001	0.0008488	0.0000000	0.0021228	0.0000000
12.501	0.0000697	0.0000000	0.0001742	0.0000000
13.056	0.0000400	0.0000000	0.0001000	0.0000000

Table 7: Gauss-Galerkin/finite difference solution: The values of numerical moments of order four to seven along $x = 0.0$.

Time	m0	m1	m2	m3
0.001	0.0000000	0.0000000	0.0000000	0.0000000
0.251	0.0000000	0.0000000	0.0000000	0.0000000
0.500	0.0000000	0.0000000	-0.0000005	0.0000000
0.501	0.0000000	0.0000000	-0.0000005	0.0000000
0.750	0.0000000	0.0000000	-0.0000011	0.0000000
1.000	0.0000001	0.0000000	-0.0000014	0.0000000
1.250	0.0000003	0.0000000	-0.0000013	0.0000000
1.500	0.0000005	0.0000000	-0.0000011	0.0000000
1.750	0.0000006	0.0000000	-0.0000008	0.0000000
2.000	0.0000008	0.0000000	-0.0000005	0.0000000
2.250	0.0000009	0.0000000	-0.0000003	0.0000000
2.500	0.0000010	0.0000000	-0.0000002	0.0000000
3.001	0.0000011	0.0000000	0.0000000	0.0000000
4.001	0.0000012	0.0000000	0.0000002	0.0000000
5.000	0.0000011	0.0000000	0.0000002	0.0000000
6.000	0.0000010	0.0000000	0.0000002	0.0000000
7.000	0.0000009	0.0000000	0.0000002	0.0000000
7.500	0.0000009	0.0000000	0.0000002	0.0000000
7.500	0.0000009	0.0000000	0.0000002	0.0000000
10.001	0.0000006	0.0000000	0.0000001	0.0000000
12.501	0.0000004	0.0000000	0.0000001	0.0000000
13.056	0.0000004	0.0000000	0.0000001	0.0000000

Table 8: Gauss-Galerkin/finite difference solution: At $x = 0.0$, difference between numerical moments and exact moments of order zero to three.

Time	m4	m5	m6	m7
0.001	0.0000000	0.0000000	0.0000000	0.0000000
0.251	-0.0000002	0.0000000	0.0130764	0.0000000
0.500	-0.0000017	0.0000000	0.0125145	0.0000000
0.501	-0.0000017	0.0000000	0.0125094	0.0000000
0.750	-0.0000034	0.0000000	0.0099015	0.0000000
1.000	-0.0000039	0.0000000	0.0077171	0.0000000
1.250	-0.0000036	0.0000000	0.0060089	0.0000000
1.500	-0.0000029	0.0000000	0.0046791	0.0000000
1.750	-0.0000023	0.0000000	0.0036439	0.0000000
2.000	-0.0000017	0.0000000	0.0028377	0.0000000
2.250	-0.0000013	0.0000000	0.0022100	0.0000000
2.500	-0.0000009	0.0000000	0.0017219	0.0000000
3.001	-0.0000005	0.0000000	0.0010439	0.0000000
4.001	0.0000000	0.0000000	0.0003841	0.0000000
5.000	0.0000001	0.0000000	0.0001414	0.0000000
6.000	0.0000001	0.0000000	0.0000521	0.0000000
7.000	0.0000001	0.0000000	0.0000193	0.0000000
7.500	0.0000001	0.0000000	0.0000118	0.0000000
7.500	0.0000001	0.0000000	0.0000117	0.0000000
10.001	0.0000001	0.0000000	0.0000011	0.0000000
12.501	0.0000001	0.0000000	0.0000002	0.0000000
13.056	0.0000001	0.0000000	0.0000002	0.0000000

Table 9: Gauss-Galerkin/finite difference solution: At $x = 0.0$, difference between numerical moments and exact moments of order four to seven.

TIME	Maximum of the time difference of moments
0.0005	46.875000000000
0.2505	36.487705858505
0.5005	28.437777611792
0.7505	22.147816506532
1.0005	17.247759848566
1.2505	13.431729474020
1.5005	10.459985598491
1.7505	8.1457350166581
2.0005	6.3435080228516
2.2505	4.9400201311993
3.0005	2.3330626924851
4.0005	0.8580712150043
5.0005	0.3155878452433
6.0005	0.1160692655025
7.0005	4.268883345D-02
7.50045	2.588885695D-02
10.0005	2.123773380D-03
12.5005	1.742318613D-04

Table 10: Gauss-Galerkin/finite difference: Maximum of the time-difference of the numerical moments.

y-nodes	-1.65040	-0.52387	0.52387	1.65040

Table 11: Gauss-Galerkin/finite difference solution: At $x = 0.0$, there are four nodes in the y-direction.

is considered. When the initial condition is the Dirac-delta measure, the exact solution (referred to as $p(x,y,t)$) which is given by

$$p(x,y,t) = \frac{0.5e^{\frac{-2.0\left(-x^2+e^{4.0\,t}x^2-x\,y+2.0\,e^{2.0\,t}x\,y-e^{4.0\,t}x\,y-0.25\,y^2+e^{2.0\,t}y^2-0.75\,e^{4.0\,t}y^2+e^{4.0\,t}t\,y^2\right)}{-1.0+2.0\,e^{2.0\,t}-e^{4.0\,t}-t+e^{4.0\,t}t}}}{\pi^{\frac{3}{2}}\sqrt{-0.0625-0.0625e^{-4.0\,t}+0.125e^{-2.0\,t}+0.0625\,t-0.0625\,te^{-4.0\,t}}},$$

is plotted for various times in Figure 5. The moment functions are given in Figure 6.

The initial condition we choose is of the form

$$u_0(x,y) = \frac{1}{\sqrt{\pi}}e^{-x^2}\delta(y).$$

The exact solution (referred as $g(x,y,t)$, which equals $p \circ u_0$.) is plotted for various times in Figures 7, 8 and 9.

To apply the Gauss-Galerkin/finite difference method, we use the Gauss-Galerkin approximation in the the y direction while using the finite difference approximation in the x direction. As the initial condition is a singular measure, the conditions of our convergence theorems do not apply. Thus, we choose a smooth/regularized condition to be the initial condition in the approximation. One such smoothing is given simply by solving for the exact solution up to time $t = 0.1$, which is

$$g(x,y,0.1) = 0.78394e^{-0.99983x^2+0.09965xy-6.06897y^2},$$

and then applying the Gauss-Galerkin/finite difference method. Given x, the zero-th through the fifth order moments of the exact solution in the y direction have simple analytic forms which are used to prescribe the initial approximations for the Gauss-Galerkin method.

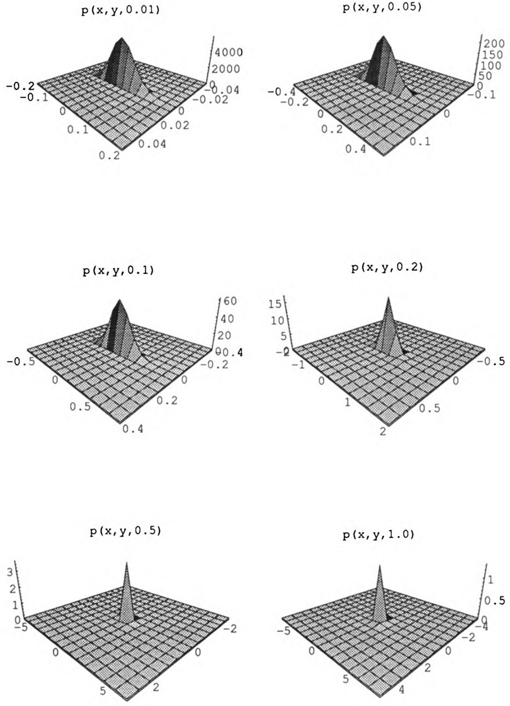
Let

$$\mathcal{M}_n(x,t) = \int_{-\infty}^{\infty} y^n g(x,y,t)dy.$$

It follows from the differential equation that

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{M}_n &= \int_{-\infty}^{\infty} y^n g_t dy \\
&= \int_{-\infty}^{\infty}\left(-y^{n+1}g_x + 2y^{n+1}g_y + 2y^n g + \frac{1}{2}y^n g_{yy}\right)dy \\
&= -\frac{\partial}{\partial x}\int_{-\infty}^{\infty} y^{n+1}gdy - 2n\int_{-\infty}^{\infty} y^n gdy + \frac{n(n-1)}{2}\int_{-\infty}^{\infty} y^{n-2}gdy \\
&= -\frac{\partial}{\partial x}\mathcal{M}_{n+1} - 2n\mathcal{M}_n + \frac{n(n-1)}{2}\mathcal{M}_{n-2}
\end{aligned}$$

This system of equations cannot be solved recursively through analytic approach for general initial conditions. However, some interesting properties of the moments


p(x,y,0.01)

p(x,y,0.05)

p(x,y,0.1)

p(x,y,0.2)

p(x,y,0.5)

p(x,y,1.0)

Figure 5: Exact solution with the Dirac-delta initial condition.

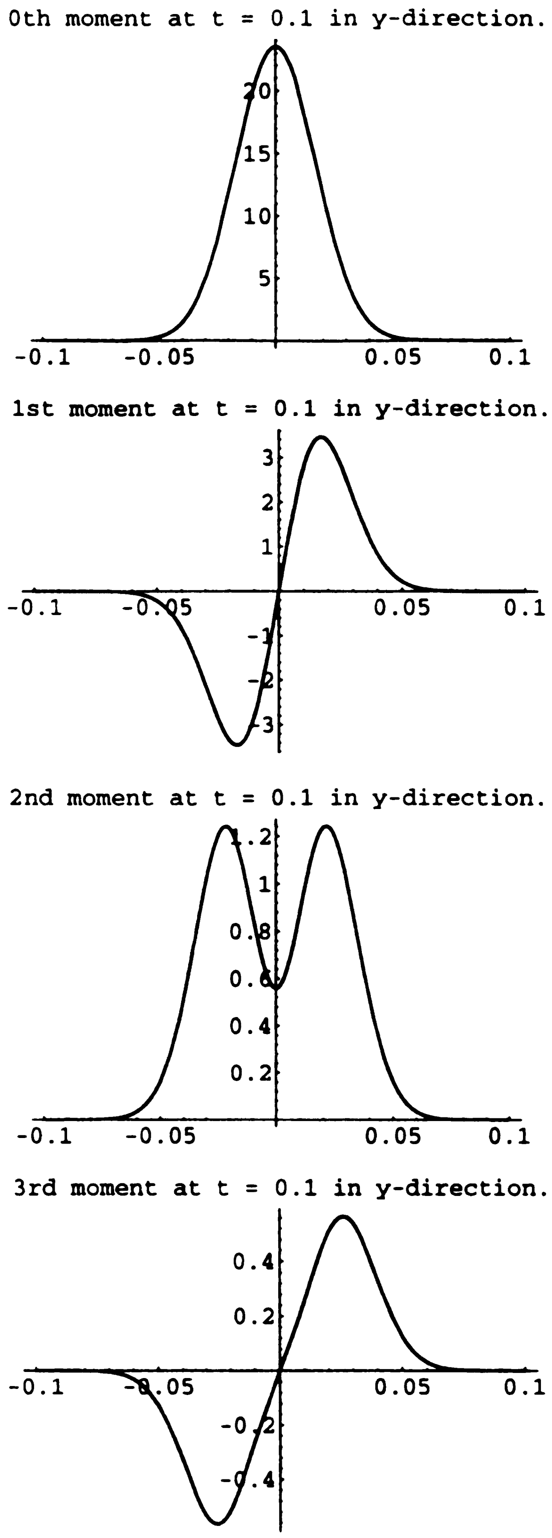


Figure 6: Exact moments with the Dirac-delta initial condition.

may be observed algebraically in the following and graphically in Figure 10 through Figure 13.

If the initial condition is an even function in x, then the even order moments remain even in x while the odd order moments remain odd for any positive t. Thus,

$$\frac{\partial}{\partial x}\mathcal{M}_{2n}(0,t) = 0$$

and

$$\mathcal{M}_{2n+1}(0,t) = 0\ .$$

The equation is given in the whole space. In theory, it can be transformed into an equation defined in a finite region as discussed in the previous sections. However, for practical numerical purposes, we choose to restrict the equations to a finite interval in the y-direction so that finite difference approximation in y can be made. Additional boundary condition has to be implemented as a cut-off of the solution defined in the whole space. Given any finite positive time, the exact solution decays at infinity. Thus, setting the solution to be zero for large values of x would be a reasonable approximation. Naturally, the accuracy of our numerical scheme is largely affected by the cut-off interval that we pick for x. This will be explained later through graphic display.

Given a grid in the x-direction, we typically choose two to three nodes in the y-direction for each x grid point. The length of the cut-off interval in x ranges from $[-2,2]$ to $[-6,6]$ and the number of grid points in x ranges from 20 to 60. The time step-size is often between 10^{-5} to 10^{-2} and the length of time interval is on the order of 1 to 20.

The graphs of the exact and numerical moments as functions of x are given in Figures 10-13. In Figure 10, the moments are computed on the interval $x \in [-2.0, 2.0]$ at $t = 5.21$. For the same value of t, the moments computed on the interval $x \in [-4.0, 4.0]$ are given in Figure 11. By comparing the exact solution with the numerical solutions, we see that the larger the cut-off interval is, the numerical solution will remain a good approximation over a longer time interval. Similar pictures for moments computed on the interval $x \in [-4.0, 4.0]$ and the interval $x \in [-6.0, 6.0]$ at $t = 12.0$ are given in Figures 12 and 13, which substantiate the above claim. This is intuitively clear as demonstrated by the graphs of the exact solution.

Finally, the zero-th order exact moments, numerical moments and the errors at $t = 12$ with 40 grid points in x and two nodes in y are presented in Tables 12 and 13.

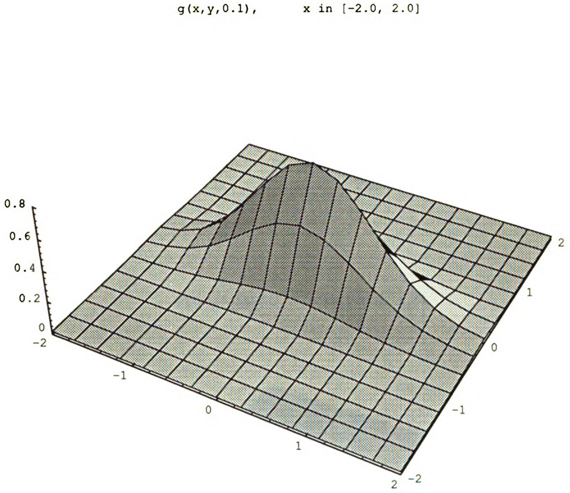
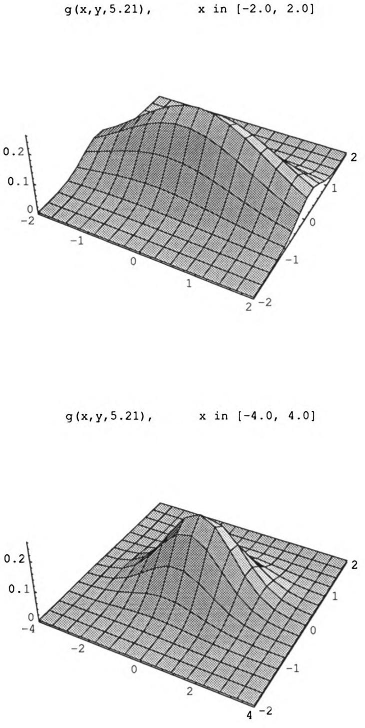


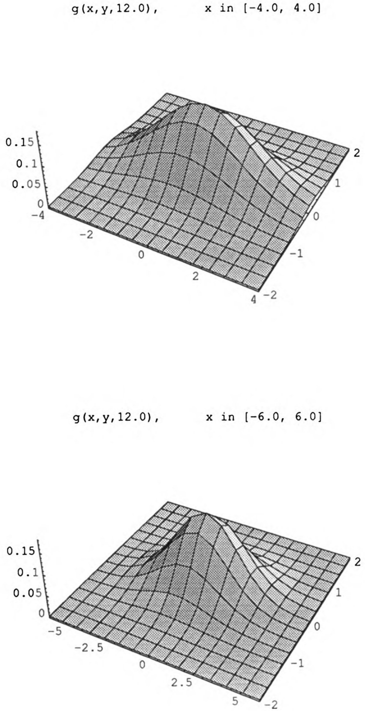
Figure 7: Exact solution for $t = 0.1$.



g(x,y,5.21),   x in [-2.0, 2.0]

g(x,y,5.21),   x in [-4.0, 4.0]

Figure 8: Exact solution for $t = 5.21$.



g(x,y,12.0),   x in [-4.0, 4.0]

g(x,y,12.0),   x in [-6.0, 6.0]

Figure 9: Exact solution for $t = 12.0$.

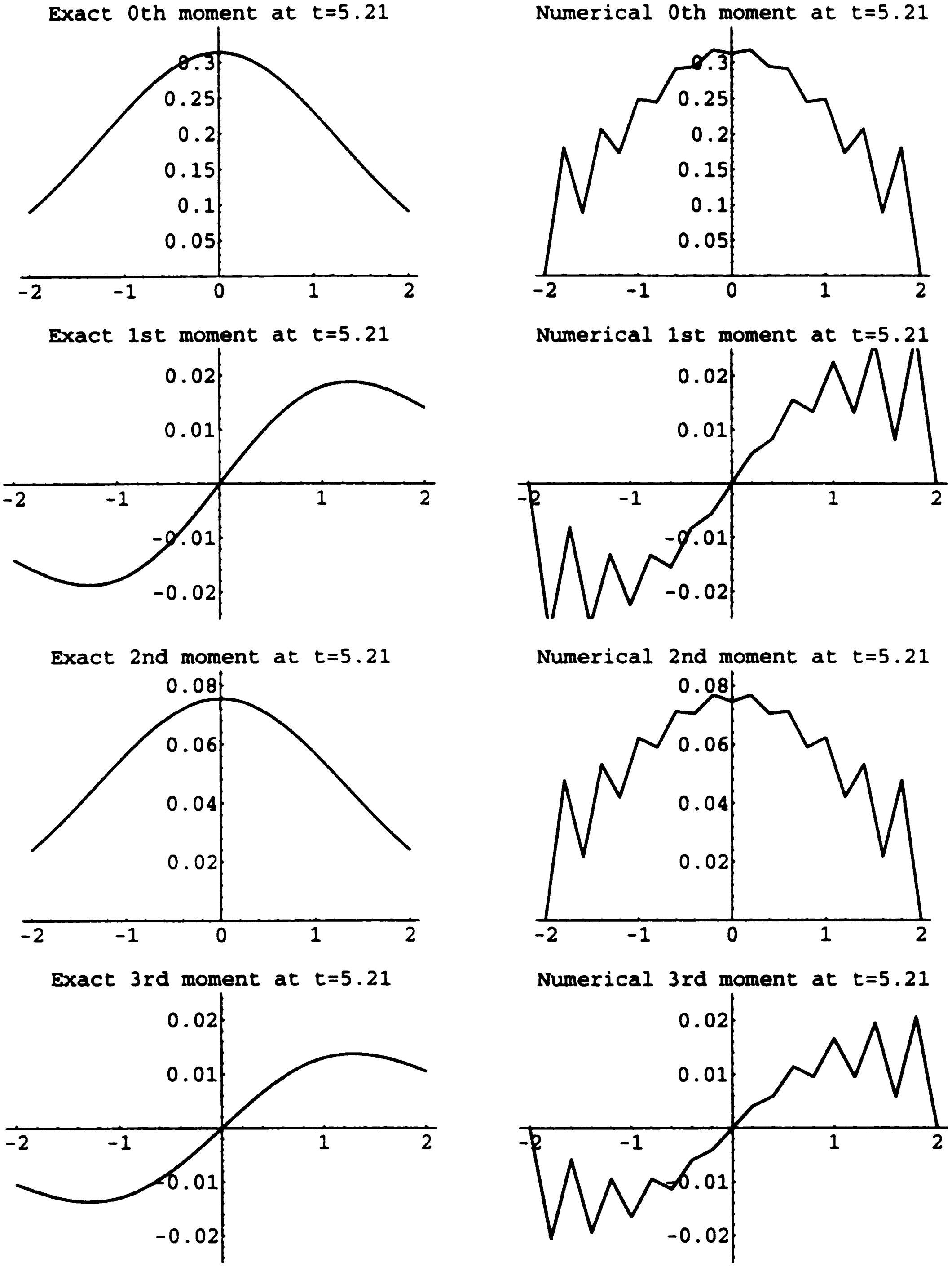


Figure 10: Comparison of the moments when $t = 5.21, x \in [-2.0, 2.0]$.

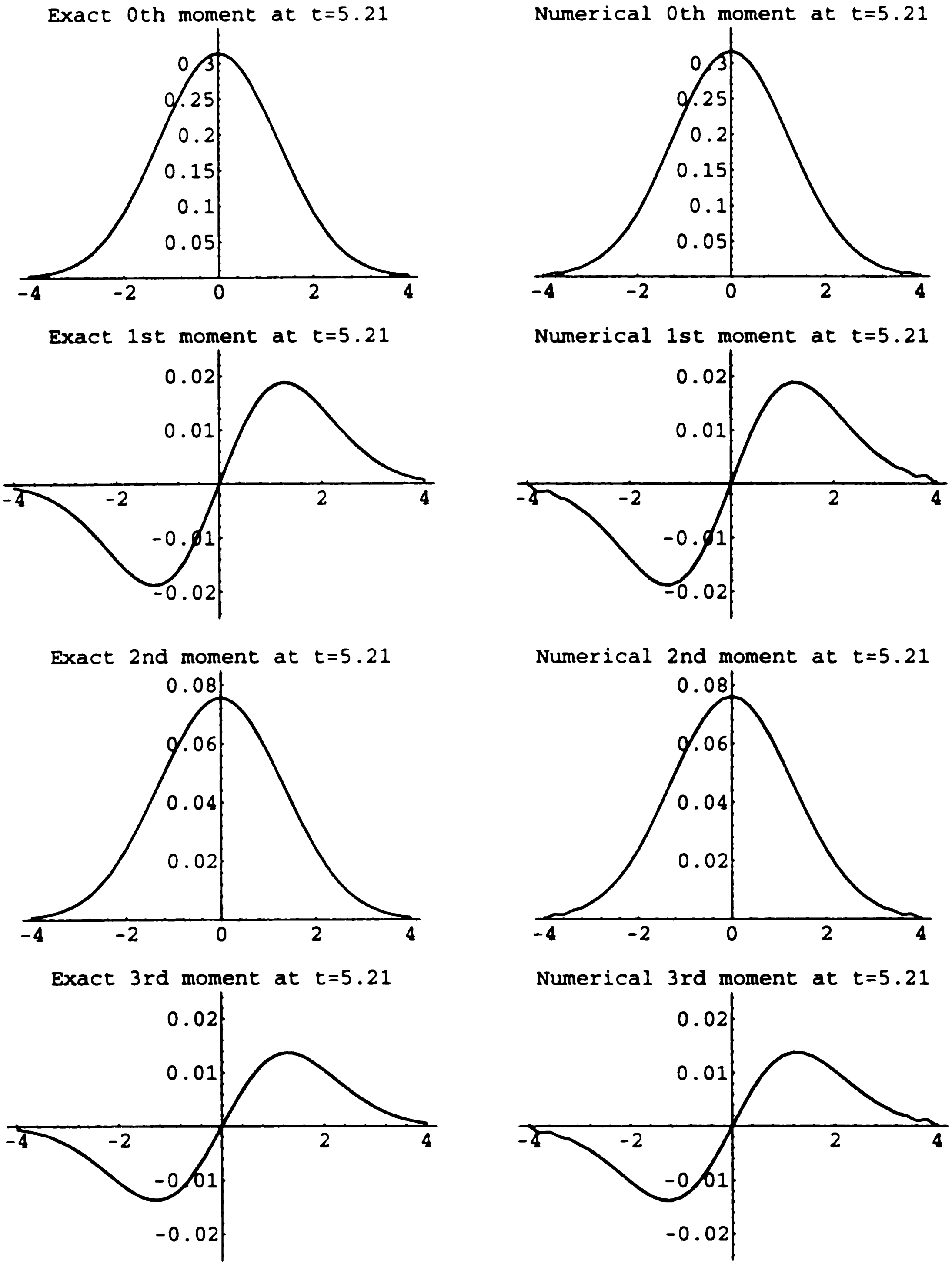
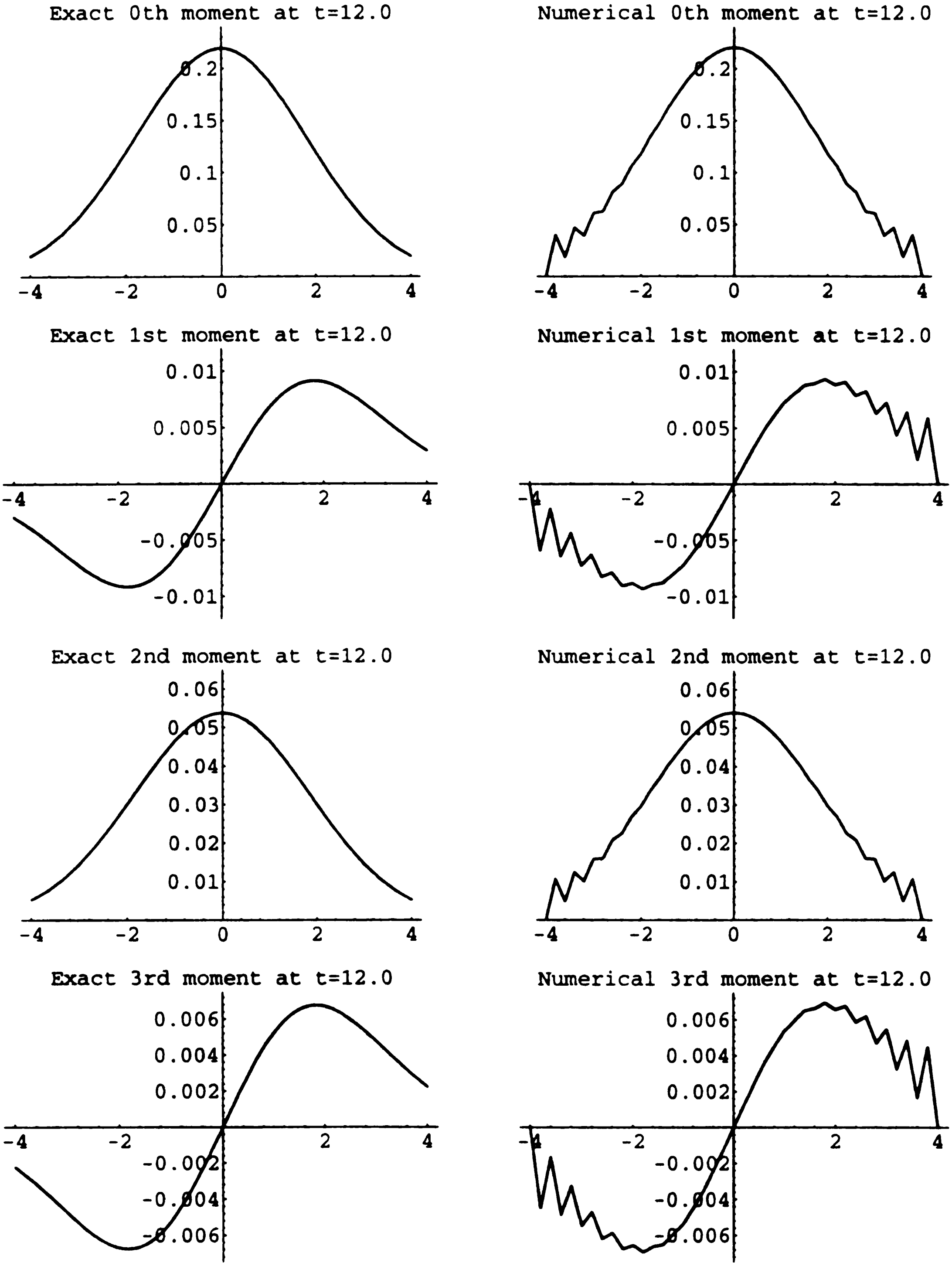


Figure 11: Comparison of the moments when $t = 5.21, x \in [-4.0, 4.0]$.



Exact 0th moment at t=12.0

Numerical 0th moment at t=12.0

Exact 1st moment at t=12.0

Numerical 1st moment at t=12.0

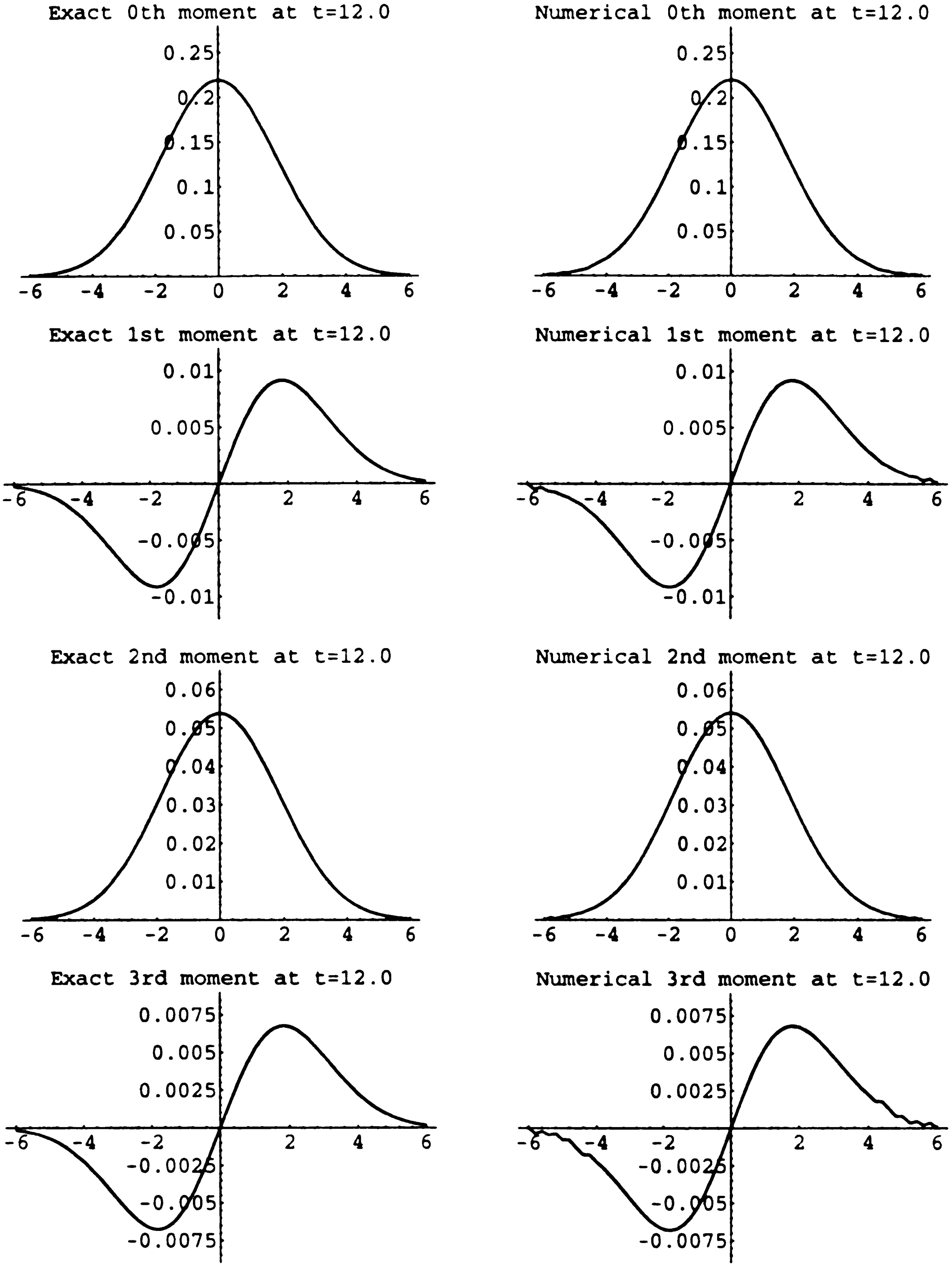
Exact 2nd moment at t=12.0

Numerical 2nd moment at t=12.0

Exact 3rd moment at t=12.0

Numerical 3rd moment at t=12.0

Figure 12: Comparison of the moments when $t = 12.0, x \in [-4.0, 4.0]$.


Exact 0th moment at t=12.0

Numerical 0th moment at t=12.0

Exact 1st moment at t=12.0

Numerical 1st moment at t=12.0

Exact 2nd moment at t=12.0

Numerical 2nd moment at t=12.0

Exact 3rd moment at t=12.0

Numerical 3rd moment at t=12.0

Figure 13: Comparison of the moments when $t = 12.0, x \in [-6.0, 6.0]$.

x	exact moments	numer. moments	error
-4.	0.01958709263951136	0.	0.01958709263951136
-3.8	0.02478757472874219	0.04014208886942286	-0.01535451414068067
-3.6	0.03099229840377004	0.01934606081985256	0.01164623758391748
-3.4	0.03828505035179119	0.04738628312217605	-0.00910123277038486
-3.2	0.04672618773768949	0.04024364299446147	0.006482544743228021
-3.	0.05634393388840611	0.06157565114840212	-0.005231717259996015
-2.8	0.06712583040039726	0.06380420000341576	0.003321630396981495
-2.6	0.07901105596761341	0.0819711045936535	-0.002960048626040101
-2.4	0.0918843928912383	0.0903606037010086	0.001523789190229691
-2.2	0.1055726258102865	0.1072242303532511	-0.001651604542964605
-2.	0.1198440800758245	0.1192245427953118	0.0006195372805126582
-1.8	0.1344118436157715	0.1352757364472455	-0.000863892831474034
-1.6	0.1489409704215598	0.1486639552648338	0.0002770151567260681
-1.4	0.1630596514241039	0.1633949452912657	-0.0003352938671617733
-1.2	0.1763739857375736	0.1761117785044321	0.0002622072331415204
-1.	0.1884856268586013	0.1884167533216087	0.00006887353699266963
-0.8	0.1990112541285562	0.1986229428879876	0.0003883112405686506
-0.6	0.2076025693122749	0.2072437597245973	0.0003588095876776165
-0.4	0.213965375825513	0.2134385613519516	0.0005268144735613712
-0.2	0.2178762877388823	0.2173596789576086	0.0005166087812737419

Table 12: At $t = 12.0$, the zero-th order exact, numerical moments and the errors. ($x = -4.0 \to -0.2$)

x	exact moments	numer. moments	error
0.	0.219195746475355	0.2186121798330846	0.0005835666422704112
0.2	0.2178762877388823	0.2173596789576086	0.0005166087812737419
0.4	0.213965375825513	0.2134385613519516	0.0005268144735613712
0.6	0.2076025693122749	0.2072437597245973	0.0003588095876776165
0.8	0.1990112541285562	0.1986229428879876	0.0003883112405686506
1.	0.1884856268586013	0.1884167533216087	0.00006887353699266963
1.2	0.1763739857375736	0.1761117785044321	0.0002622072331415204
1.4	0.1630596514241039	0.1633949452912657	-0.0003352938671617733
1.6	0.1489409704215598	0.1486639552648338	0.0002770151567260681
1.8	0.1344118436157715	0.1352757364472455	-0.000863892831474034
2.	0.1198440800758245	0.1192245427953118	0.0006195372805126582
2.2	0.1055726258102865	0.1072242303532511	-0.001651604542964605
2.4	0.0918843928912383	0.0903606037010086	0.001523789190229691
2.6	0.07901105596761341	0.0819711045936535	-0.002960048626040101
2.8	0.06712583040039726	0.06380420000341576	0.003321630396981495
3.	0.05634393388840611	0.06157565114840212	-0.005231717259996015
3.2	0.04672618773768949	0.04024364299446147	0.006482544743228021
3.4	0.03828505035179119	0.04738628312217605	-0.00910123277038486
3.6	0.03099229840377004	0.01934606081985256	0.01164623758391748
3.8	0.02478757472874219	0.04014208886942286	-0.01535451414068067
4.	0.01958709263951136	0.	0.01958709263951136

Table 13: At $t = 12.0$, the zero-th order exact, numerical moments and the errors. ($x = 0.0 \rightarrow 4.0$)

10 Discussions and Concluding Remarks

In the previous sections, we have developed, analyzed and implemented the Gauss-Galerkin/finite difference method for two-dimensional Fokker-Planck equations. Theoretical analysis suggests that the method is applicable to a wide range of equations arising in nonlinear oscillations. Numerical tests show that the method is practical and very accurate especially in dealing with solutions that exhibit δ-measure like singularities.

A number of important and interesting issues remain to be studied in the future. Among them, there are various theoretical questions concerning the properties of two-dimensional Fokker-Planck equations, and the Gauss-Galerkin/finite difference approximation such as if the solutions develop singularity and if the Gauss-Galerkin nodes would collide into each other in finite time. On the numerical aspect, we have only presented one possible approach to the two-dimensional Fokker-Planck equations.

In the following, we remark on some other alternatives.

10.1 Possible generalization of Gauss-Galerkin methods

The Gauss-Galerkin/finite difference method we have analyzed and implemented takes a fixed given direction to apply the difference approximation and another fixed direction to apply the Gauss-Galerkin approximation. For some problems, the solution might pile up in both directions, generalization of two-dimensional Gauss-Galerkin methods might be very attractive. It could be performed by introducing the alternating direction approach, i.e., at a given time step, if the Gauss-Galerkin method is applied in the x-direction, then, it is applied in the y-direction at the next time step.

10.2 Simulation of nonlinear oscillation model

The test problems we have studied are limited to the Fokker-Planck equations associated with the linear oscillation models of which the behavior of the exact solutions are more or less clear. The code, however, is written in a general form that allows to have variable coefficients and thus applicable to the Fokker-Planck equations associated with the nonlinear oscillation models. Test problems for *nonlinear* oscillation problems and comparison of this method with other numerical methods will be carried out in the future.

10.3 Other applications

Stochastic equations and their Fokker-Planck equations appear often in other applications such as population genetics and mathematical finance. It remains to be explored in the future that how should Gauss-Galerkin methods be applied in those exciting new areas.

Appendix

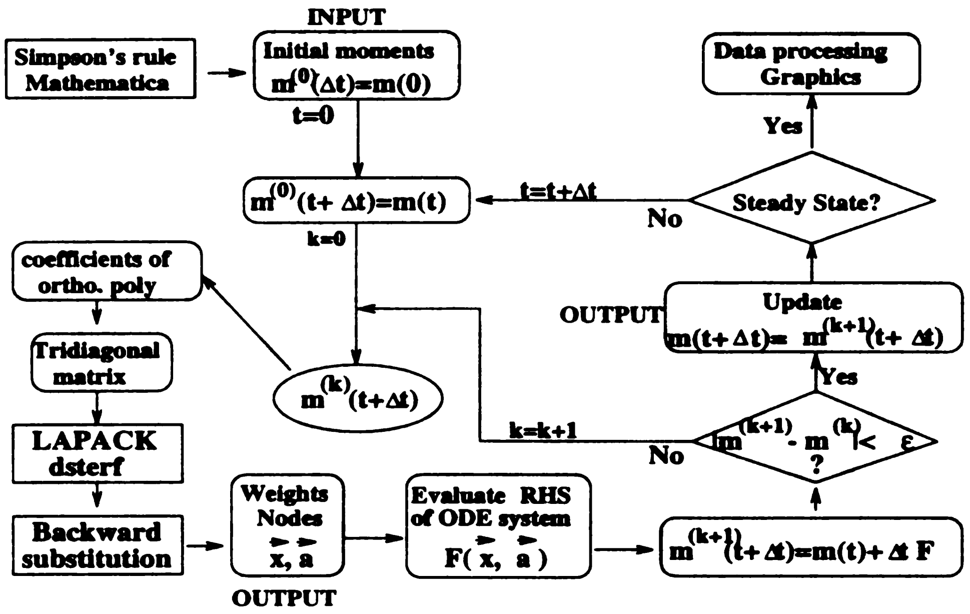


Figure 14: Flow chart for Gauss-Galerkin/finite difference method.

References

[1] Abrouk, N.E. *Some Numerical Methods for Singular Diffusions Arising in Genetics*, Ph.D. thesis, Department of Statistics and Probability, Michigan State University, 1992.

[2] Cauchy, T.K., *Nonlinear Theory of Random Vibrations*, in **Advances in Applied Mechanics**, Vol 11, pp 209-250, Academic Press, 1971.

[3] Chung, K.L., **A Course in Probability Theory**, 2nd ed., Academic Press, 1974.

[4] Crandall, S.H., *Random excitation of nonlinear systems*, **Random Vibrations**, Vol 2, pp 85-101, The M.I.T Press, 1963.

[5] Dawson, D. A., *Galerkin Approximation of Nonlinear Markov Processes*, **Statistics and Related Topics**, pp 317-339, North-Holland, 1981.

[6] Hajjafar, A., H. Salehi, and D.H.Y. Yen, *On a class of Gauss-Galerkin methods for Markov processes*, **Proceedings of the Workshop on Nonstationary Stochastic Processes and Their Applications, August 1-2, 1991**, pp 126-146, World Scientific, 1992.

[7] Harrison, G.W., *Numerical solution of the Fokker-Planck equations using moving finite elements*, **Numerical Methods for Partial Differential Equations**, Vol. 4, pp 219-232, 1988.

[8] Ilin, A.M., and R.Z. Khasminskii, *On equations of Brownian motion*, **Theory of Probability and its Applications**, Vol. 9, pp 421-444, 1964.

[9] Khasminskii, R.Z., *Ergodic properties of recurrent diffusion processes and stabilization of the solution to the Cauchy problem for parabolic equations*, **Theory of Probability and its Applications**, Vol. 5, pp 179-196, 1960.

[10] Kushner, H.J., **Stochastic Stability and Control**, Academic Press, 1967.

[11] Kushner, H.J., *The Cauchy problem for a class of degenerate parabolic equations and asymptotic properties of the related diffusion processes*, **J. of Differential Equations**, Vol. 6, pp 209-231, 1969.

[12] Loeve, M., **Probability Theory**, 3rd ed., the university series in higher mathematics, VNR Company, 1974.

[13] Smith, G.D., **Numerical Solution of Partial Differential Equations**, Oxford University Press, 2nd ed., 1985.

[14] Stoer, J. and R. Bulirsch, **Introduction to Numerical Analysis**, Springer-Verlag, 1983.

[15] Shohat, J.A. and J.D. Tamarkin, **The Problem of Moments**, Mathematical surveys, no.1, AMS, 1963.

[16] Yosida, K., **Functional Analysis**, Springer-verlag, 1985.