mathematics

MDPI

Article

Efficient Solution of Fokker–Planck Equations in Two Dimensions

Donald Michael McFarland [1,2], Fei Ye [1], Chao Zong [1], Rui Zhu [1], Tao Han [2], Hangyu Fu [1], Lawrence A. Bergman [3] and Huancai Lu [2,*]

[1] Sound and Vibration Laboratory, College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310014, China; dmmcf@zjut.edu.cn (D.M.M.); 2112102075@zjut.edu.cn (F.Y.); 2112102199@zjut.edu.cn (C.Z.); 2112102343@zjut.edu.cn (R.Z.); 2112002120@zjut.edu.cn (H.F.)
[2] Ningbo Institute of Digital Twin, Eastern Institute of Technology, Ningbo 315201, China; than@idt.eitech.edu.cn
[3] Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; lbergman@illinois.edu
* Correspondence: huancai.lu@idt.eitech.edu.cn

Abstract: Finite element analysis (FEA) of the Fokker–Planck equation governing the nonstationary joint probability density function of the responses of a dynamical system produces a large set of ordinary differential equations, and computations become impractical for systems with as few as four states. Nonetheless, FEA remains of interest for small systems—for example, for the generation of baseline performance data and reference solutions for the evaluation of machine learning-based methods. We examine the effectiveness of two techniques which, while they are well established, have not to our knowledge been applied to this problem previously: reduction of the equations onto a smaller basis comprising selected eigenvectors of one of the coefficient matrices, and splitting of the other coefficient matrix. The reduction was only moderately effective, requiring a much larger basis than was expected and producing solutions with clear artifacts. Operator splitting, however, performed very well. While the methods can be combined, our results indicate that splitting alone is an effective and generally preferable approach.

Keywords: Fokker–Planck equation; finite element analysis; dimension reduction; operator splitting

MSC: 65M22; 35Q84

check for updates

Academic Editor: Andreas Bück

Received: 9 December 2024
Revised: 10 January 2025
Accepted: 21 January 2025
Published: 31 January 2025

Citation: McFarland, D.M.; Ye, F.; Zong, C.; Zhu, R.; Han, T.; Fu, H.; Bergman, L.A.; Lu, H. Efficient Solution of Fokker–Planck Equations in Two Dimensions. *Mathematics* **2025**, *13*, 491. https://doi.org/10.3390/math13030491

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The Fokker–Planck (FP) equation, also known as the forward Kolmogorov equation, is a deterministic partial equation (PDE) whose solution gives the joint probability density function (pdf) of the responses (states) of a dynamical system subjected to random excitation. While analytical solutions are available for linear systems of any dimension [1,2], nonlinear dynamical systems naturally give rise to more challenging FP equations, for which analysis may be limited to the stationary pdf [3] when it can be carried out at all. When the FP equation is discretized, for example, by using finite elements in a system's phase space, the resulting matrix equation grows much faster than the dimension of the original dynamical system, and computations quickly become intractable. In the following, we try to address this problem for small (but nontrivial) systems.

The first attempts at the application of finite element analysis (FEA) in numerical treatment of the FP equation date back several decades. Bergman and Heinrich [4] first

proposed Petrov–Galerkin FEA to solve the backward Kolmogorov equation (the formal adjoint of the FP equation), and solved the first passage problem for a second-order linear system driven by additive Gaussian white noise. Langley [5] used a finite element formulation to study the stationary random vibration of 2-D nonlinear systems. Spencer and Bergman [6] used a Bubnov–Galerkin FEA method to solve the FP equations of linear, Duffing, and Van der Pol oscillators, and studied the corresponding first passage problems. El-Gebeily and Shabaik [7] reported a straightforward analysis of the transient problem. Masud and Bergman [8] presented a multi-scale FEA formulation, applied it to calculate the solution of the FP equation, and examined its efficacy through selected applications. Floris [9] used spline interpolation in FEA of the FP equation, Kumar et al. [10] considered systems subjected to colored non-Gaussian noise, and Král and Náprstek [11] adopted simplex elements to treat higher-dimensional problems. Further algorithmic refinements and additional applications may be found in Refs. [12–17].

All of these works, however, suffer from the same fundamental limitation. When the FE method is used to solve the FP equation, as the dimension of the state space increases, the number of equations to be solved increases exponentially. The resulting huge amount of computation required brings great challenges. In this paper, we consider two approaches to overcoming this difficulty—namely, reducing the dimension of the FE equation by projecting it onto a smaller basis, eliminating one coefficient matrix in the process; and splitting the remaining matrix to enable efficient time marching of the nonstationary solution. Numerous examples of related work on deterministic problems may be found in the literature, for example, in computational fluid dynamics [18–21] and in nearfield acoustic holography [22]. The novelty of this work lies in the application of these techniques to the FP equation, and in the results obtained for specific choices of basis and splitting scheme.

Dimension reduction has been gradually applied in the solution of FP equations. Lötstedt and Ferm [23] obtained results for a problem involving multiple chemical species by deriving PDEs to approximate the first moments for some of the species and coupling these to the FP equations for the species not so reduced. Leonenko and Phillips [24] used spectral discretization and a high-order reduced-basis approximation in an application to polymer dynamics, and Er [25] proposed a subspace method to solve for the response probability density functions of large-scale nonlinear stochastic dynamical systems. Chen and Lin [26] and Chen and Rui [27] obtained the generalized probability density evolution equation by using decoupled physical equations. Zhu [28] studied the stationary probability density function of a multi-degree-of-freedom nonlinear system subjected to external, independent Poisson white noise, reducing the high-dimensional generalized FP equation to a low-dimensional equation, and Liang et al. [29] improved on this approach. In contrast to most of these studies, the reduction method used herein does not depend on the physics of the problem considered.

Independent of the dimension of the problem, operator splitting provides a numerical method to compute the solution to matrix differential equations that can be more efficient than straightforward integration using the original coefficient matrices. Operator splitting methods have been known for more than a century, starting with the scheme introduced by Lie in the mid-1870s [30]. Usadi and Dawson [31] developed alternating direction implicit (ADI) methods. In the mid-1970s, close relationships between the augmented Lagrangian methods of Hestenes [32] and Powell [33,34] and ADI methods were identified, which inspired the alternating direction method of multipliers (ADMM). Based on a symmetrization principle, Strang [35] introduced a second-order variant of the Lie scheme, which was motivated by a need for the accurate solution of hyperbolic problems. This second-order scheme is unconditionally stable, which makes it popular in computing solutions to PDEs. Speth

et al. [36] further developed balanced splitting and rebalanced splitting methods; compared with earlier splitting methods, these have higher accuracy and can correctly capture steady-state behavior. Goldstein and Osher [37] discussed an algorithm that is equivalent to the ADMM, and this method has been used in many image-processing applications.

Operator splitting has attracted attention for its computational efficiency in the solution of FP equations. Zorzano et al. [38] combined an ADI scheme with the finite difference method to solve FP problems in particle accelerators. For studying the influence of noise in the particle dynamics in a storage ring, a method based on operator splitting was developed by Mais and Zorzano [39], which can also be used in stochastic beam-dynamics problems in accelerators. Zorzano et al. [40] provided a robust finite difference scheme with operator splitting (ADI) for the FP equation in two spatial variables and time.

Because of the rapid growth of the FE equation with the increasing dimension of the underlying dynamical system, it now seems unlikely that this approach to the solution of the FP equation will be extended beyond the three or four states that have represented a practical limit for several years. However, systems of this size do have practical applications, such as design optimization and nonlinear filtering [41], and so there is some value in pursuing their efficient solution. In addition, we are motivated in part by the goal of providing baseline timing and solution results for the FEA of such problems as the field shifts increasingly toward the routine use of alternative formulations, such as physics-informed neural networks (PINNs) [42–44], which are very promising for problems of high dimension but whose numerical properties have not yet been fully tested.

Our interest is thus pragmatic. The finite element analysis of the FP equation is largely a mature subject, with some gaps where known numerical techniques may be applied. We try to provide practical guidance for the use of two of these, model reduction and operator splitting. While we focus on 2-D problems herein, our approach is directly extensible to larger systems, which have structurally similar coefficient matrices; however, both programming complexity and storage requirements grow very quickly, and the efficiency gains we achieve are not enough to offset this. Rather than trying to compete with methods such as the PINN solvers mentioned above (a rough computation based on estimating the numbers of floating-point operations needed suggests that they overtake FEA by the time a system has grown to four states), we seek to improve the efficiency of the solution of a class of problems that may serve as well-understood test cases. For the same reason, we believe these results will be of value to those studying FEA or other techniques [45–47] with an eye primarily to performance and algorithmic properties other than speed.

In the following sections, the FP equation is used to formulate a PDE for the joint pdf of the displacement and velocity of a single-degree-of-freedom (SDOF) oscillator, which is then discretized in the phase plane. The resulting large matrix ordinary differential equation has constant coefficient matrices, and the properties of their eigenvectors for use in a reduced basis are elucidated. One of the coefficient matrices may be usefully split into the sum of two matrices with better stability characteristics, which makes feasible the use of a state transition matrix to advance the solution by a fixed step size. The performance of these techniques is demonstrated through numerical examples where the nonstationary random responses of linear and nonlinear SDOF oscillators are computed. Recommendations based on the findings from these applications are summarized in the results and conclusion.

2. Problem Formulation

Consider an SDOF system (an oscillator) governed by the stochastic differential equations

$$\dot{X}_1 = g_1(\mathbf{X}), \tag{1a}$$
$$\dot{X}_2 = g_2(\mathbf{X}) + w(t), \tag{1b}$$

where the state vector is $\mathbf{X} = \begin{bmatrix} X & \dot{X} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} X_1 & X_2 \end{bmatrix}^{\mathrm{T}}$ and the overdot denotes differentiation with respect to time. The excitation $w(t)$ is Gaussian white noise, which is fully defined by its first two moments,

$$E[w(t)] = 0 \qquad \text{and} \qquad E[w(t)w(t+t')] = 2D\delta(t'), \tag{2}$$

where $\delta(\cdot)$ is the Dirac delta function and D/π is the magnitude of the constant two-sided spectral density function of the excitation. Equation (1) describes a Markov process, and its behavior is completely determined by the transition probability density function $p(\mathbf{x}, t|\mathbf{x}_0)$, which satisfies the (deterministic) Fokker–Planck equation

$$\frac{\partial p}{\partial t} = D\frac{\partial^2 p}{\partial x_2^2} - \frac{\partial}{\partial x_1}[g_1(\mathbf{x})p] - \frac{\partial}{\partial x_2}[g_2(\mathbf{x})p] \tag{3}$$

and the initial condition $\lim_{t\to 0} p(\mathbf{x}, t|\mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0)$. Homogeneous boundary conditions are imposed where $p \to 0$ as $|x_1|, |x_2| \to \infty$. The joint probability density function of the response is related to the transition probability density by

$$f(\mathbf{x}, t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(\mathbf{x}, t|\mathbf{x}_0) f(\mathbf{x}_0)\, \mathrm{d}\mathbf{x}_0, \tag{4}$$

where for convenience the initial condition may be taken in the form of a binormal distribution. Substituting this function into the FP equation and integrating term-wise gives

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x_2^2} - \frac{\partial}{\partial x_1}(g_1 f) - \frac{\partial}{\partial x_2}(g_2 f). \tag{5}$$

The desired joint probability density function $f(\mathbf{x}, t)$ satisfies this PDE subject to the initial conditions on f and the boundary conditions (which are the same as those on p).

A Galerkin finite element method is used where within a single element with domain Ω_e, the probability density function is interpolated according to

$$f(\mathbf{x}, t) = \sum_{r=1}^{n} N_r(\mathbf{x}) f_r^e(t), \tag{6}$$

where the $N_r(\mathbf{x})$, $r = 1, 2, \ldots, n$, are bilinear trial functions of class C^0, n is the number of nodes in a single element, and f_r^e are the nodal values of the density. Substituting Equation (6) into the weak form of Equation (5) leads to

$$\sum_{s=1}^{n} m_{rs}\dot{f}_s^e + \sum_{s=1}^{n} k_{rs} f_s^e = 0, \qquad r = 1, \ldots, n, \tag{7}$$

where the element matrices are given by

$$m_{rs} = \iint_{\Omega_e} N_r N_s \, \mathrm{d}x_1 \, \mathrm{d}x_2 \tag{8a}$$

and

$$k_{rs} = \iint_{\Omega_e} \left[D\frac{\partial N_r}{\partial x_2}\frac{\partial N_s}{\partial x_2} - N_s\left(g_1\frac{\partial N_r}{\partial x_1} + g_2\frac{\partial N_r}{\partial x_2}\right)\right] \mathrm{d}x_1 \, \mathrm{d}x_2. \tag{8b}$$

Assembling the global coefficient matrices by standard procedures produces

$$\mathbf{M}\dot{\mathbf{f}}(t) + \mathbf{K}\mathbf{f}(t) = \mathbf{0}, \tag{9}$$

where $\mathbf{f}(t)$ is the global vector of nodal values of the joint probability density function and $\mathbf{M}$ and $\mathbf{K}$ are constant, square matrices.

We note some properties of the coefficient matrices that are relevant to the discussion to follow. Both $\mathbf{M}$ and $\mathbf{K}$ are real and sparse. The matrix $\mathbf{M}$ is symmetric and positive definite, while $\mathbf{K}$ is nonsymmetric and has 0 as an eigenvalue with multiplicity one; the corresponding eigenvector, scaled so that

$$\int_\Omega f \, \mathrm{d}\Omega = 1, \tag{10}$$

is the solution to

$$\mathbf{K}\mathbf{f} = \mathbf{0}, \tag{11}$$

and gives the stationary joint pdf of the displacement and velocity responses. (Such normalization is not needed when the transient solution is computed starting from initial conditions given in the form of a pdf.) The matrix $\mathbf{M}$ is independent of the oscillator model and the excitation intensity, while $\mathbf{K}$ depends on both.

Experience has shown that for typical linear, Duffing, and Van der Pol oscillators, an effective finite element discretization can be achieved with about 100 elements in each dimension of a uniform rectangular mesh covering Ω [6]. This leads to a state vector $\mathbf{f}$ of dimension 101^2 = 10,201. After imposition of homogeneous essential boundary conditions at the nodes on the boundary $\partial\Omega$, this is reduced to 9801.

3. Model Reduction

Given the matrices $\mathbf{M}$ and $\mathbf{K}$, we can pose three eigenproblems whose solutions may be useful in reducing the dimension of Equation (9) and thus lowering the computational cost of its solution. In this section, we review some familiar results from linear algebra and establish a framework for the systematic selection of the eigenvectors to be included in a reduced basis [48]. We also discuss an ad hoc technique for improving satisfaction of the boundary conditions in the reduced problem.

3.1. Eigenproblems in the Coefficient Matrices

The simplest eigenvalue problem we will consider is

$$\mathbf{M}\boldsymbol{\phi} = \lambda\boldsymbol{\phi}. \tag{12}$$

The $n \times n$ matrix $\mathbf{M}$ is real, symmetric, and positive definite, with eigenvalues λ_i, $i = 1, \ldots, n$, where n is on the order of 10^4 in the problems considered herein. Consequently, the eigenvectors $\boldsymbol{\phi}_i$ are orthogonal, both simply and with respect to $\mathbf{M}$, and can be normalized so that

$$\boldsymbol{\phi}_i^{\mathrm{T}}\mathbf{M}\boldsymbol{\phi}_j = \delta_{ij}. \tag{13}$$

If we define a transformation matrix using a subset of the eigenvectors,

$$\mathbf{S} = \begin{bmatrix} \cdots & \boldsymbol{\phi}_i & \cdots \end{bmatrix}, \tag{14}$$

and let

$$\mathbf{f} = \mathbf{S}\boldsymbol{\eta}, \tag{15}$$

we obtain

$$\dot{\boldsymbol{\eta}} + \hat{\mathbf{K}}\boldsymbol{\eta} = \mathbf{0}, \tag{16}$$

where

$$\hat{\mathbf{K}} = \mathbf{S}^{\mathrm{T}}\mathbf{K}\mathbf{S}. \tag{17}$$

If we have used $m < n$ eigenvectors in defining the transformation, the matrix $\mathbf{S}$ is not square, but $n \times m$; as a result, $\boldsymbol{\eta}$ is of length m, and $\hat{\mathbf{K}}$ is $m \times m$. The dimension of the problem has been reduced from n to m.

The situation is somewhat different if we instead choose to solve

$$\mathbf{K}\boldsymbol{\phi} = \lambda\boldsymbol{\phi}. \tag{18}$$

Because $\mathbf{K}$ is real but not symmetric, its eigenvalues may not all be real. We do again expect an eigenvalue of 0 and its (real) eigenvector giving the stationary pdf. (Computationally, it is usually necessary to accept the eigenvalue that is smallest in modulus as "zero", but we have encountered no cases in which this was problematic.) The other eigenvectors may be complex, so we need to use the Hermitian (conjugate) transpose; otherwise, the foregoing calculations are essentially unchanged. In general, the eigenvectors of $\mathbf{K}$ are not orthogonal. Because neither of the transformed coefficient matrices $\hat{\mathbf{M}}$ and $\hat{\mathbf{K}}$ can be expected to be diagonal while they both will likely be complex, it is not clear that this offers any advantage over the original DE with real coefficient matrices. By defining the transformation of coordinates using a matrix $\mathbf{S}$ containing a subset of the eigenvectors of $\mathbf{K}$, it is again possible to obtain a reduced problem. However, the real-valued reduced problem of the previous subsection, constructed using eigenvectors of $\mathbf{M}$, seems preferable.

The solution of the generalized eigenproblem

$$\mathbf{K}\boldsymbol{\phi} = \lambda\mathbf{M}\boldsymbol{\phi} \tag{19}$$

will produce n eigenvalues and eigenvectors, generally complex-valued. We once again expect an eigenvalue of zero, with its eigenvector giving the stationary solution. The eigenvectors have properties similar to those of $\mathbf{K}$ above, and are not orthogonal with respect to either matrix. The reduced coefficient matrices, while they may be structurally simpler than $\mathbf{M}$ and $\mathbf{K}$, are again both complex.

Subsets of vectors from any of the three problems will work in practice. Because neither of the transformed coefficient matrices is diagonal, and because many standard differential equation solvers require that the resulting complex-valued problem be recast as a real-valued problem of double the dimension, there is little attraction to using eigenvectors of $\mathbf{K}$. The same observations can be made about the use of vectors from the generalized eigenproblem, although there it might be hoped that these drawbacks would be offset by the greater effectiveness of a smaller basis. However, experience has shown that neither set of complex eigenvectors offers any significant advantage over the real eigenvectors obtained from the regular eigenproblem in $\mathbf{M}$, particularly regarding the size of the basis required to achieve a given accuracy. Furthermore, the use of eigenvectors of $\mathbf{M}$ allows the diagonalization of one coefficient matrix, and, with proper normalization, $\hat{\mathbf{M}} = \mathbf{I}$. Therefore, we choose to draw our reduced basis from the eigenvectors of matrix $\mathbf{M}$.

3.2. Selection of Basis Vectors

The questions immediately arise of how to choose m and which m eigenvectors are to be included in $\mathbf{S}$. Some guidance may be had if a vector $\mathbf{u}$ representing a meaningful pdf is

available—for example, results from a past calculation. The projections of this vector onto the eigenvectors are given by

$$\beta_i = \frac{\boldsymbol{\phi}_i^{\mathrm{T}}\mathbf{u}}{\boldsymbol{\phi}_i^{\mathrm{T}}\boldsymbol{\phi}_i}, \qquad i = 1, \ldots, n, \tag{20}$$

when $\boldsymbol{\phi}_i$ is real, and by a similar expression involving the conjugate transpose when the eigenvector is complex. These coefficients can be sorted by absolute value and examined for abrupt changes indicating groups of eigenvectors of similar importance in the new basis. If there are no obvious groupings of the β_i, the results will at least show which $\boldsymbol{\phi}_i$ should be chosen first. Naturally, the value of this ranking depends on the accuracy of $\mathbf{u}$ in representing the solution over time. It may be desirable to construct a composite, artificial "solution" from snapshots taken from a previously computed time series, if one is available. This technique is demonstrated in the examples of a later section.

3.3. Limitations of Solution on Reduced Basis

When computing a nonstationary solution of the FP equation on a reduced basis, it is frequently observed that while the overall character of the pdf obtained at any time step is correct, there are small errors in both the peak values and the tails of the distribution. The latter often manifest as large values near the boundaries, and in combination with the homogeneous boundary conditions lead to large gradients in the pdf and so encourage diffusion of probability mass over the boundaries. This leads to errors that grow over time and ultimately to an underestimated, if not qualitatively wrong, stationary solution. Enlarging the computational domain, refining the mesh, or both have little effect on this phenomenon; a nearly complete basis is often needed if the solution is to satisfy the boundary conditions over time. This significantly limits the dimension reduction that can be employed.

This deterioration of the pdf can be ameliorated by introducing into the original PDE artificial diffusive terms that oppose this loss over the boundaries. Specifically, we consider a modified equation in the form

$$\frac{\partial f}{\partial t} = D\frac{\partial^2 f}{\partial x_2^2} + \alpha_1\frac{\partial^2 f}{\partial x_1^2} + \alpha_2\frac{\partial^2 f}{\partial x_2^2} - \frac{\partial}{\partial x_1}[g_1 f] - \frac{\partial}{\partial x_2}[g_2 f]. \tag{21}$$

For simplicity, the coefficients α_1 and α_2 are taken to be nonzero only in rectangular regions of widths b_1 and b_2 near $\partial\Omega$, starting at zero on the inner edges of these strips and growing quadratically to values of d_1 and d_2 on the boundaries. These terms are easily discretized in the finite element formulation of the problem. The effects of the parameters controlling the subdomains ("boundary layers") where these artificial terms exist and the strength of the resulting inward diffusion are examined in the examples to follow.

4. Operator Splitting

If the convection and diffusion terms in the original equation are separated, perhaps in the expectation that their physics may develop on different time scales, the terms in the element $\mathbf{k}$ matrix can be grouped and that matrix can be divided into

$$\mathbf{k}_1 = \iint_{\Omega_e} D\frac{\partial \mathbf{N}}{\partial x_2}\frac{\partial \mathbf{N}^{\mathrm{T}}}{\partial x_2}\,\mathrm{d}x_1\,\mathrm{d}x_2 \tag{22a}$$

and

$$\mathbf{k}_2 = -\iint_{\Omega_e}\left(g_1\frac{\partial \mathbf{N}}{\partial x_1} + g_2\frac{\partial \mathbf{N}}{\partial x_2}\right)\mathbf{N}^{\mathrm{T}}\,\mathrm{d}x_1\,\mathrm{d}x_2. \tag{22b}$$

Obviously, this splitting satisfies $\mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2$. Then, global matrices $\mathbf{K}_1$ and $\mathbf{K}_2$ can be assembled as the previous $\mathbf{K}$ matrix was and transformed to obtain $\hat{\mathbf{K}}_1 = \mathbf{S}^{\mathrm{T}}\mathbf{K}_1\mathbf{S}$ and $\hat{\mathbf{K}}_2 = \mathbf{S}^{\mathrm{T}}\mathbf{K}_2\mathbf{S}$.

For the system of ordinary differential equations of the form $\dot{\boldsymbol{\eta}} + \hat{\mathbf{K}}\boldsymbol{\eta} = \mathbf{0}$ obtained using $\mathbf{S} = \begin{bmatrix} \boldsymbol{\phi}_1 & \dots & \boldsymbol{\phi}_n \end{bmatrix}$ (so that $\hat{\mathbf{M}} = \mathbf{I}$), we begin with the simple first-order splitting method and

$$\dot{\boldsymbol{\eta}} + (\hat{\mathbf{K}}_1 + \hat{\mathbf{K}}_2)\boldsymbol{\eta} = \mathbf{0}, \tag{23}$$

and separate the original equation into two parts. Over a time step Δ_t, we solve the two equations

$$\dot{\mathbf{q}} + \hat{\mathbf{K}}_1\mathbf{q} = \mathbf{0}, \quad \dot{\mathbf{p}} + \hat{\mathbf{K}}_2\mathbf{p} = \mathbf{0} \tag{24}$$

separately, using the solution of the first as the initial conditions for the second, and then combine the two separate solutions to form a solution to the original Equation (23) [35]. Using the state transition matrices associated with Equation (24), we can advance the complete solution from time step n to step $n+1$ by computing

$$\mathbf{q}_{n+1} = \mathrm{e}^{-\mathbf{K}_1\Delta_t}\mathbf{p}_n, \quad \mathbf{p}_{n+1} = \mathrm{e}^{-\mathbf{K}_2\Delta_t}\mathbf{q}_{n+1}. \tag{25}$$

Combining these equations for a fixed time step Δ_t, the result

$$\boldsymbol{\eta}_{n+1} = \mathrm{e}^{-\hat{\mathbf{K}}_2\Delta_t}\mathrm{e}^{-\hat{\mathbf{K}}_1\Delta_t}\boldsymbol{\eta}_n \tag{26}$$

is obtained, which may be used to efficiently advance the solution over any number of steps by simply multiplying the previous value by a constant matrix. By comparison to the Taylor series for the exact solution of a canonical equation, it can be shown that this splitting method is first-order accurate ($\mathcal{O}(\Delta_t)$) even when the matrices $\mathbf{K}_1$ and $\mathbf{K}_2$ do not commute under multiplication [49].

This scheme is both simple and stable once the matrix exponentials are computed and multiplied to form the single-state transition matrix needed to advance the solution over a fixed step. It is of course possible to choose a time step so large that the exponentiation (carried out in floating-point arithmetic) fails, but our experience has been that using time steps small enough to resolve the dynamics of interest in a system usually avoids any such difficulties.

Symmetric splitting, as introduced by Strang [35], can achieve second-order accuracy. Here, $\mathbf{K}$ is split into

$$\mathbf{K} = \left(\tfrac{1}{2}\mathbf{K}_1 + \mathbf{K}_2\right) + \tfrac{1}{2}\mathbf{K}_1. \tag{27}$$

Examining the Taylor series shows that

$$\mathrm{e}^{\frac{1}{2}\mathbf{K}_1\Delta_t}\mathrm{e}^{\mathbf{K}_2}\mathrm{e}^{\frac{1}{2}\mathbf{K}_1\Delta_t} - \mathrm{e}^{(\mathbf{K}_1+\mathbf{K}_2)\Delta_t} = \mathbf{C}\Delta_t^3 + O(\Delta_t^4), \tag{28}$$

where

$$\mathbf{C} = \big[[\mathbf{K}_1, \mathbf{K}_2], \mathbf{K}_1\big] + 2\big[[\mathbf{K}_1, \mathbf{K}_2], \mathbf{K}_2\big] \tag{29}$$

and where $[\mathbf{A}, \mathbf{B}]$ denotes the matrix commutator $\mathbf{AB} - \mathbf{BA}$. The total number of time steps is T/Δ_t, so the global error is $\mathcal{O}(\Delta_t^2)$.

5. Numerical Examples

The example problems of this section are taken from Spencer and Bergman [6] and comprise two SDOF oscillators, one linear and one nonlinear (a Duffing oscillator with negative linear stiffness but positive cubic stiffness). The notation adopted here is typical of that used to describe vibrating mechanical systems, but these equations of motion have

analogs in many areas of science and engineering. Initial conditions are specified as a binormal distribution. For each example, we give results from reduced problems, from operator splitting, and from these techniques used in combination. The solution of the full FE problem by an adaptive Runge–Kutta (RK) integrator is used as a reference except in those cases where an exact solution is available. (Specifically, we used the function `solve_ivp` from the SciPy Python library (v. 1.8.0), which implements the RK5(4) algorithm of Dormand and Prince [50].) This routine was also used to integrate those reduced problems for which operator splitting was not employed. Time-dependent results are usually reported here in terms of the nondimensional time $\tau = \omega_0 t/2\pi$ (i.e., in units of the (linearized) natural period of the system being considered).

Basis vectors were selected from eigenvectors of $\mathbf{M}$ with the aid of a composite vector $\mathbf{u}$ constructed by superposing twenty snapshots from a full RK solution, equally spaced in time from the relevant initial conditions to apparent stationarity, according to the magnitudes of the projections of this vector onto the eigenvectors. This allows the estimation of the minimum basis size required for the solution of the nonstationary problem, but it is obviously not a technique that will be generally available. Because the required bases were found to be quite large, half or more of the original problem dimension in the Duffing example, we have not investigated alternative means of selecting the best basis vectors; however, we note that the eigenvectors selected on the basis of the projection coefficients obtained using $\mathbf{u}$ and the eigenvectors corresponding to the largest eigenvalues of $\mathbf{M}$ were generally strongly correlated. This suggests that, in the typical case where a composite solution vector is not available, including in the basis, those eigenvectors with the largest eigenvalues will generally give the best results. This choice, and the required size of the basis, will be influenced by the degree to which the nonstationary solution is localized, and thus by the initial conditions. The initial mean values used in our examples are rather large, which makes for interesting but challenging examples.

We summarize here the steps followed in each example when either Runge–Kutta integration or first-order operator splitting was used. The steps for second-order splitting are very similar.

1. Form the finite element global coefficient matrices $\mathbf{M}$ and $\mathbf{K}$. If operator splitting will be used, form $\mathbf{K}_1$ and $\mathbf{K}_2$ in lieu of $\mathbf{K}$. If basis reduction will be used, optionally include in $\mathbf{K}$ or $\mathbf{K}_1$ additional terms representing artificial diffusion.
2. Solve the eigenproblem $\mathbf{M}\boldsymbol{\phi} = \lambda\boldsymbol{\phi}$ and normalize the resulting vectors with respect to $\mathbf{M}$. (Hence, $\hat{\mathbf{M}} = \mathbf{I}$.)
3. Select the eigenvectors to be used to transform the problem (which may be all of them) and form $\mathbf{S}$.
4. Compute $\hat{\mathbf{K}}$ or $\hat{\mathbf{K}}_1$ and $\hat{\mathbf{K}}_2$, as appropriate.
5. Starting from the given initial conditions, march the solution in time (typically to stationarity).
   (a) If using a Runge–Kutta integrator, solve $\dot{\boldsymbol{\eta}} = -\hat{\mathbf{K}}\boldsymbol{\eta}$ directly. If the output is in the form of a sparse structure, evaluate the solution at the desired times (e.g., on uniform steps).
   (b) If using operator splitting, choose a time step Δ_t, evaluate the necessary matrix exponentials, and form the corresponding state transition matrix (cf. Equation (26)). Advance the solution by forming a sequence of matrix–vector products.

Computations were performed on a typical x86 desktop computer using programs written in Python and taking advantage of its compiled numerical libraries, with the exception of the eigensolutions, which were found using Matlab (v. R2021b) running on the same hardware. The reported run times (normalized by the time required for the computation of the RK reference solution) include any transformations associated with problem reduction and recovery of the full solution vector.

5.1. Linear Oscillator

The FP equation for a linear system has an analytical solution [1,2], and it is frequently used as a benchmark problem for testing numerical methods [4,6,8]. We consider the well-known SDOF linear oscillator (LO)

$$g_1(x_1, x_2) = x_2, \tag{30a}$$

$$g_2(x_1, x_2) = -2\zeta\omega_0 x_2 - \omega_0^2 x_1, \tag{30b}$$

with the parameters $\zeta = 0.05$ and $\omega_0 = 1$, the excitation $D = 0.1$, and the initial conditions $(\mu_{10}, \mu_{20}) = (5, 5)$ and $(\sigma_{10}^2, \sigma_{20}^2) = (1/9, 1/9)$. The domain for computation is square, $-10 \leq x_1, x_2 \leq 10$. The FE mesh is uniform, with an element size of 0.2 in both the x_1 and x_2 directions.

We begin with the FEA represented by Equations (8) and (9), applied to the state Equation (30). Using a composite solution vector **u** composed of snapshots taken from a reference Runge–Kutta solution of the full FE equations, the coefficients resulting from the projection of this vector onto all the eigenvectors of **M** are plotted in Figure 1, where they have been sorted from large to small. The norm of the error resulting when **u** is approximated with an increasing number of eigenvectors follows a similar trend.

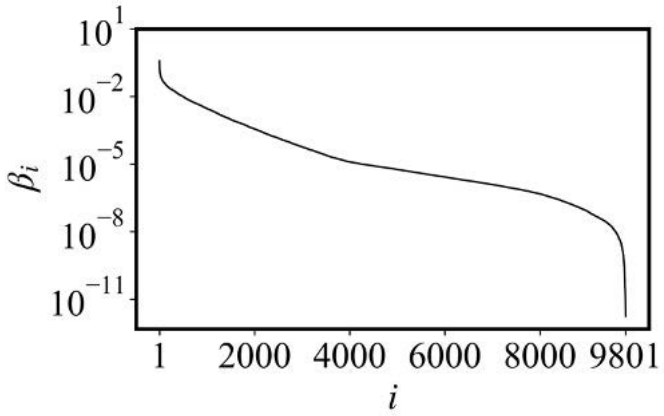
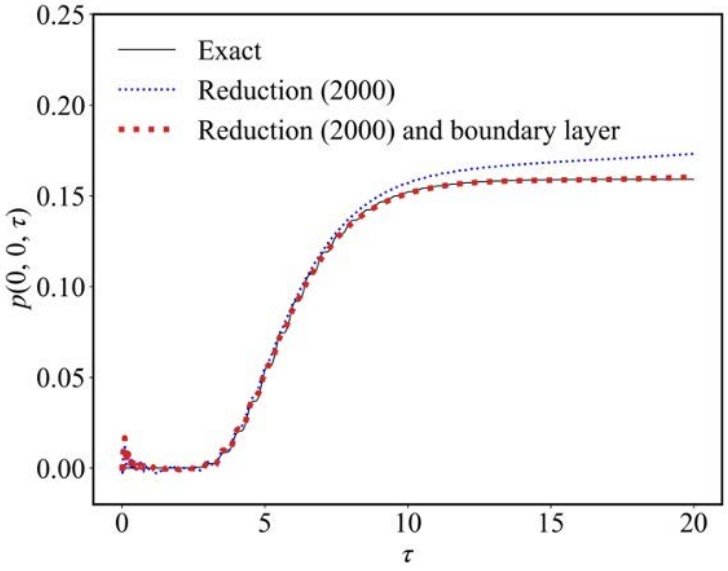


Figure 1. Coefficients resulting from the projection of the composite solution vector **u** onto the eigenvectors of **M** for the linear oscillator.

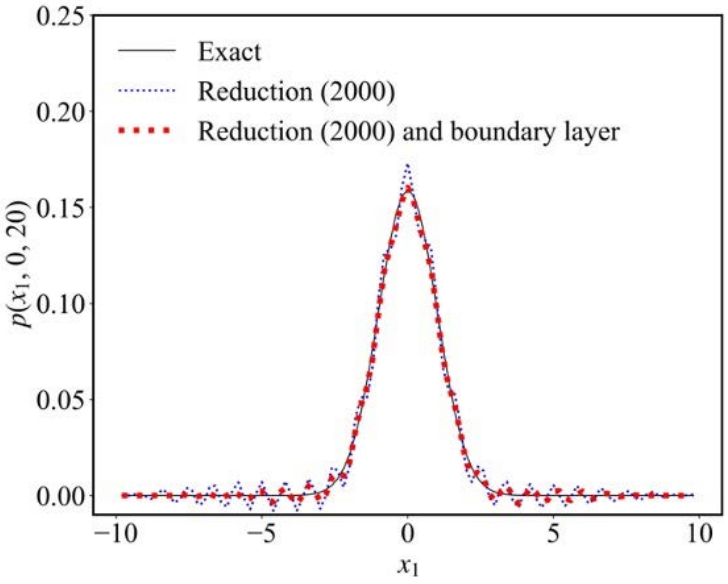
When artificial diffusive terms were used in the solution of this problem on reduced bases, parameters were chosen after a limited amount of experimentation. In the x_1 direction, the parameters $b_1 = 2.4$ and $d_1 = 1.6$ were selected, such that the added terms are nonzero only for $|x_1| > 7.6$ and rise to a maximum value of 1.6 on the boundary of Ω. Likewise, in the x_2 direction, with $b_2 = 2.5$ and $d_2 = 1.3$, the boundary terms enter where $|x_2| > 7.5$, with coefficients that increase to 1.3 on $\partial\Omega$. While an effort was made to minimize the effect of these terms on the original PDE (beyond stabilizing clearly erroneous solutions), the results obtained were not particularly sensitive to the values chosen.

Figure 2 compares solutions computed using a reduced basis of 2000 eigenvectors, with and without artificial diffusive boundary terms, to the exact solution. The basis vectors were selected according to the projection coefficient magnitudes in Figure 1. Figure 2a shows the evolution with time of the nonstationary pdf at the origin of the phase plane. Both reduced solutions exhibit initial large errors because of the distortion of the initial conditions caused by the transformation onto the smaller basis. Both then track the exact solution until approximately $\tau = 6$, when the reduced solution computed without added boundary terms begins to visibly exceed the reference curve. This trend continues, and would eventually result in divergence rather than a stationary solution. In contrast, the numerical solution computed on the same basis but with diffusive terms introduced

near the boundary continues to agree with the exact analytical result, exceeding it only slightly as τ approaches 20. A cross section of the stationary pdf along the displacement axis of the phase plane is plotted in Figure 2b, where it may be seen that the artificial diffusive boundary terms both improve the accuracy of the peak value and reduce spurious oscillations in the tails of the distribution. However, both reduced solutions do oscillate rather than going smoothly to zero in the tails, and consequently, they take on physically meaningless negative values at some points.



(**a**)

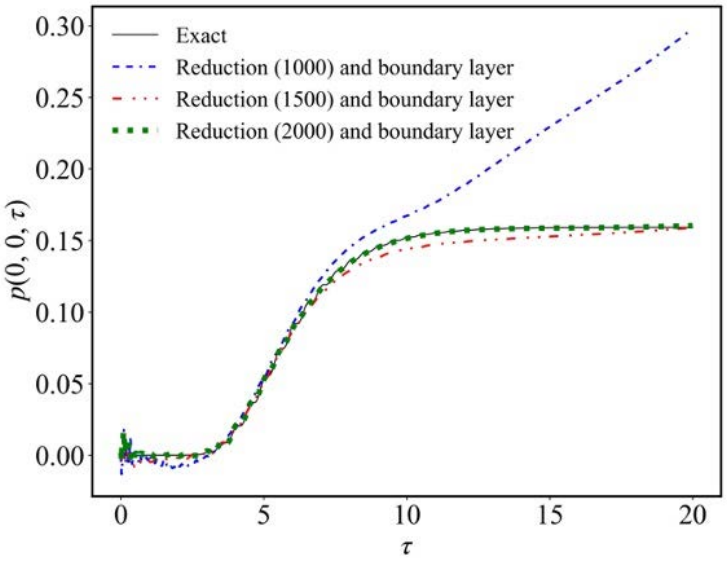


(**b**)

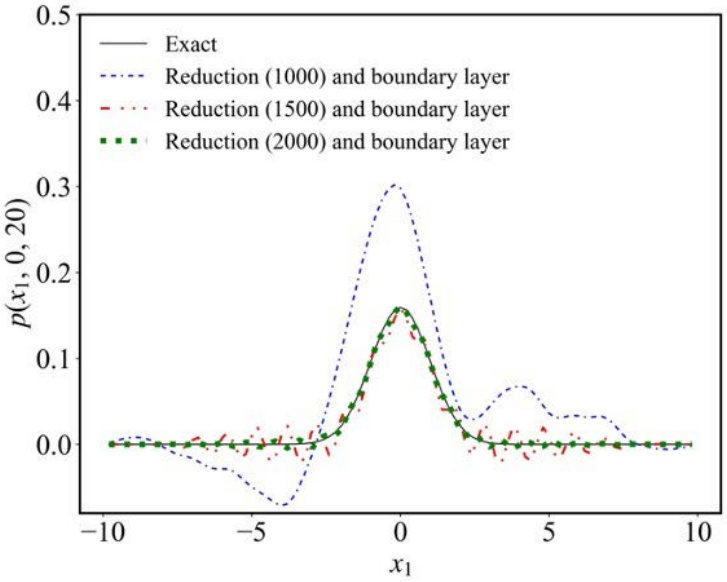
Figure 2. Comparison between solutions computed with and without added boundary terms when the LO problem is reduced onto a basis of 2000 eigenvectors: (**a**) time history of the nonstationary pdf at the origin, $(x_1, x_2) = (0, 0)$; (**b**) cross section of the pdf along $x_2 = 0$ at stationarity ($\tau = 20$).

Figures 3 and 4 reveal the effect of basis size, depicting the same responses discussed above, this time computed using a basis of 1000, 1500, or 2000 eigenvectors. All of the reduced problems were solved using artificial diffusive terms to improve the solutions. Collectively, these figures suggest that the use of fewer than 2000 basis vectors will result in spurious, sometimes negative numerical solutions, if not outright divergence. This is an

important and somewhat surprising result: despite the simplicity and smoothness of the solution, the minimum useful basis is approximately 20% of the size of the original problem.



(**a**)



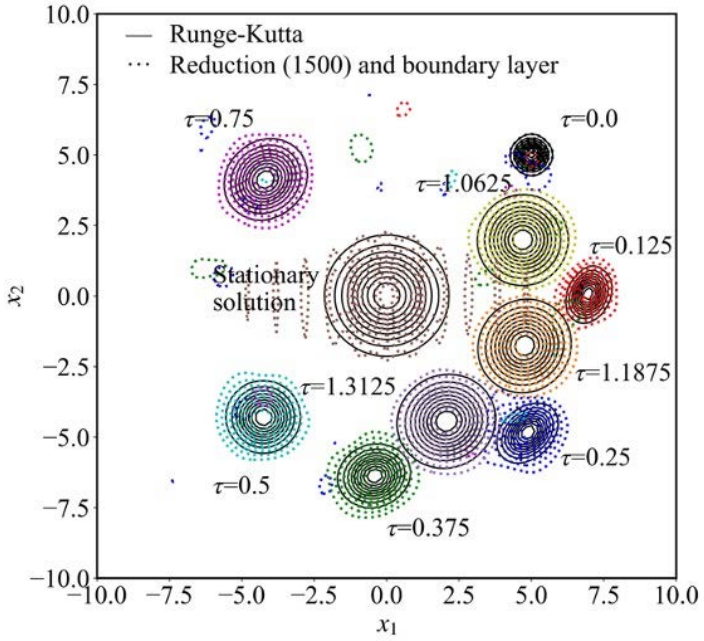
(**b**)

Figure 3. Effect of basis size on LO solutions. All computations used diffusive boundary terms. (**a**) Nonstationary pdf at origin of phase plane. (**b**) Cross section of stationary pdf along displacement axis.

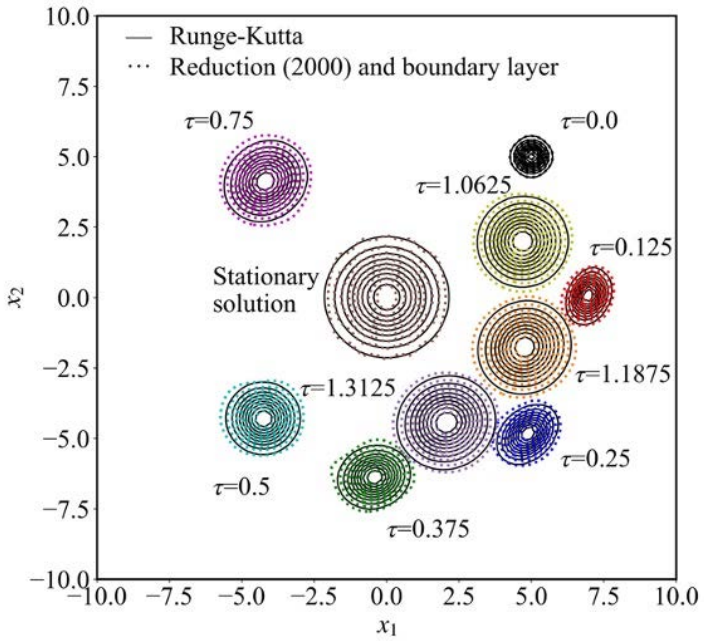
We now return to consideration of the full-dimensional problem and examine the use of operator splitting in its solution, without reduction. When the results of operator splitting are compared to the exact solution, no significant differences are found except that the splitting solution very slightly underestimates the peak value of the stationary pdf. The nonstationary and stationary numerical solutions are practically indistinguishable from the exact ones, and so are not shown here.

In combination with operator splitting, larger and smaller reduced bases, of sizes 1000 and 3000, were used without adding boundary terms. Trends and errors similar to those seen in the earlier reduced formulations for this system were found. We conclude that the use of a basis consisting of 2000 eigenvectors of the **M** matrix along with splitting the **K**

matrix according to convective and diffusive terms produces reasonably accurate results for this problem.



(**a**)



(**b**)

Figure 4. Solutions for the LO computed using (**a**) an inadequate basis of size 1500 and (**b**) an improved basis of size 2000. The computations used added boundary terms. Results of Runge–Kutta integration of the full problem are shown for reference.

Table 1 shows the run times of the different calculation methods applied to this example. Compared with Runge–Kutta integration, operator splitting and solving the problem using a state transition matrix can obviously accelerate the process. The more basis vectors used, naturally, the longer any solver will take. Solving the equation with artificial boundary terms is more time-consuming, but because of the higher accuracy, the run time cost is generally worthwhile. In many cases, the use of only operator splitting, but

not reduction, may result in acceptable run times, obviating any need for artificial diffusive terms.

Table 1. Integration times for various methods of solution for the linear oscillator (τ = 0:0.01:4). The size of the basis used in each method is shown in parentheses. ABT denotes the use of artificial boundary terms. Run times are normalized with respect to the Runge–Kutta solution of the full FE problem.

Method	Time, %
RK (9801)	100.00
RK (3000)	15.90
RK (2000)	11.98
RK (3000) and ABT	17.17
RK (2000) and ABT	12.70
1st-order splitting (9801)	17.36
1st-order splitting (3000)	10.07
1st-order splitting (2000)	9.69
1st-order splitting (3000) and ABT	10.16
1st-order splitting (2000) and ABT	9.81

5.2. Duffing Oscillator

The governing equation of the nonlinear Duffing oscillator is, in state form,

$$g_1(x_1, x_2) = x_2, \tag{31}$$

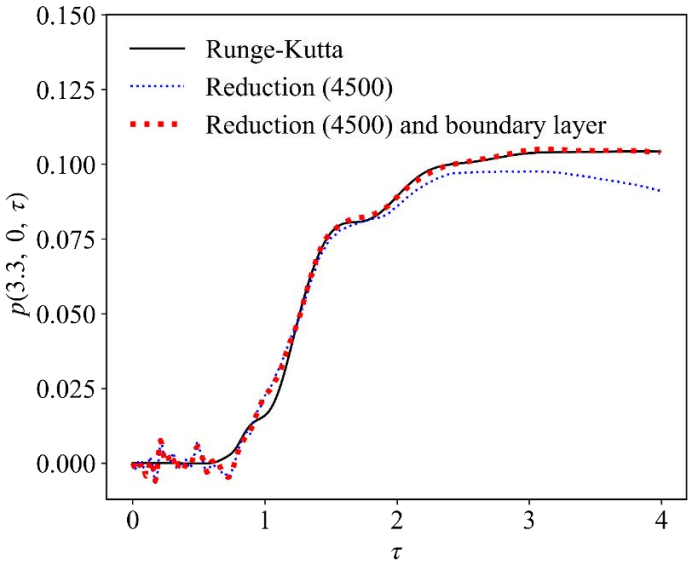
$$g_2(x_1, x_2) = -2\zeta\omega_0 x_2 - \omega_0^2 x_1(\gamma + \epsilon x_1^2), \tag{32}$$

with $\zeta = 0.2$, $\omega_0 = 1$, $\gamma = -1$, $\epsilon = 0.1$, and $D = 0.4$. (Like the previous example, this system is taken from [6].) These parameters are those of a system with negative linear stiffness but positive cubic stiffness, resulting in a bimodal stationary pdf that is doubly symmetric about the origin of the phase plane and approximately Gaussian in x_2. The nonlinearity thus controls the stationary response despite the small numerical value of ϵ. The initial conditions were taken to be $(\mu_{10}, \mu_{20}) = (0, 10)$ and $(\sigma_{10}^2, \sigma_{20}^2) = (1/2, 1/2)$, and the FE domain was $-15 \leq x_1 \leq 15$, $-20 \leq x_2 \leq 20$, with 100 elements used in each dimension.

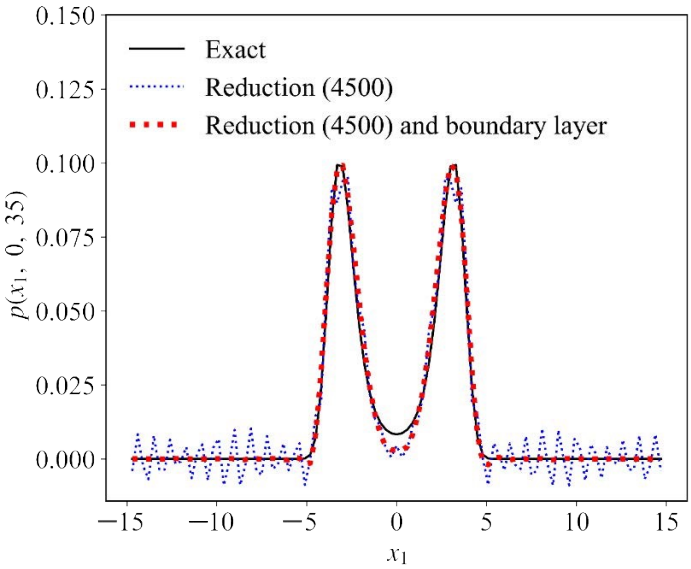
Guided by the projection coefficients of a composite vector **u** and some experimentation, we determined that somewhat more than 4000 basis vectors were needed to capture the dynamics of this problem in reduced form. The diffusive boundary layer coefficients $b_1 = 8$ and $d_1 = 4.6$ and $b_2 = 13$ and $d_2 = 0.1$ were then chosen for the x_1 and x_2 directions, respectively.

The discretized equation for the joint pdf of the Duffing oscillator was reduced onto a basis of 4500 eigenvectors, with and without introducing artificial boundary terms in the PDE. The nonstationary response in Figure 5a is shown for the point $(3.3, 0)$, which is the approximate location in the phase plane of one of the two peaks of the bimodal stationary distribution. Both the full and reduced problems were integrated here by a standard Runge–Kutta routine. When artificial diffusive boundary terms are included in the formulation, the solution of the resulting reduced problem approximates well the solution of the full problem, after a transient error attributed to the inability of the reduced basis to represent the given initial conditions accurately. Without added boundary terms, however, the solution does not approach the correct stationary value, and begins to decrease from it as probability mass is lost over the edges of the FE domain. Figure 5b, a section of the stationary pdf along the x_1 axis, shows spurious oscillations in the tails of the pdf. (The exact stationary solution shown there is from Caughey [3].) Both reduced problems underestimate the pdf in the valley between the peaks of the stationary distribution. While

the inclusion of diffusive boundary terms reduces some of the extraneous ripples, it must be said that these results are not as clean as those obtained for the linear oscillator, despite the use of a much larger reduced basis (here, almost 46% of the size of the full problem).



(**a**)



(**b**)

Figure 5. Comparison between solutions computed with and without artificial boundary terms when the Duffing oscillator problem is reduced onto a basis of 4500 eigenvectors: (**a**) time history of the nonstationary pdf at the location of a stationary peak, $(x_1, x_2) = (3.3, 0)$; (**b**) cross section of the pdf along $x_2 = 0$ at stationarity ($\tau = 35$).

When the results of first- and second-order splitting are compared to those of a Runge–Kutta solution of the unsplit problem, second-order splitting is found to be more accurate, but both solutions from splitting are found to agree quite well with the RK solution during the entire response, including at stationarity.

Responses found using a combination of basis reduction and operator splitting, with and without artificial diffusive boundary terms, followed the trends observed for the linear oscillator. Even when the basis was enlarged to 5000 eigenvectors, the solutions found without boundary terms deviated significantly from the reference RK integration of the

unreduced system. With boundary terms, good results were obtained with the previous basis size of 4500, albeit with the same sorts of errors observed when these equations were earlier solved without splitting.

The times required to solve this example as described above are shown in Table 2. Because of the large basis required, the savings compared to the full-dimensional problem are not as dramatic as for the linear oscillator, but they are still significant. Operator splitting and advancing the solution using a state transition matrix, however, yields great improvement in computational efficiency compared to straightforward time integration using a Runge–Kutta algorithm. Further savings can be achieved by combining dimension reduction with operator splitting, but in view of the errors that result and the practical requirement for adding artificial diffusive terms to the PDE, this approach is relatively unattractive in this example.

Table 2. Integration times for various methods of solution for the Duffing oscillator (τ = 0:0.01:4). The size of the basis used in each method is shown in parentheses. ABT denotes the use of artificial boundary terms. Run times are normalized with respect to the Runge–Kutta solution of the full FE problem.

Method	Time, %
RK (9801)	100.00
RK (4500)	19.84
RK (4500) and ABT	19.76
1st-order splitting (9801)	1.57
2nd-order splitting (9801)	1.61
1st-order splitting (4500)	1.00
1st-order splitting (4500) and ABT	1.03

6. Results

We have investigated two techniques to improve the speed of solution of Fokker–Planck equations as discretized using finite element analysis in the phase plane. In the first of these, reduced bases were constructed from selected eigenvectors of the coefficient matrices **M** and **K** appearing in the resulting first-order differential equation. Basis vectors obtained by solving a regular eigenproblem in **M** were attractive because they were real-valued, could be used to diagonalize **M**, and were independent of the parameters of the underlying dynamical system and the intensity of the excitation. The size of the basis required in order for the reduced problem to capture well the dynamics of the full problem was typically one-fifth to one-half that of the original system. Considering that the number of elements of the reduced coefficient matrices goes roughly as the square of this ratio, this represents an attractive savings in storage and computation; however, it does not always enable the drastic improvement in computational efficiency sometimes associated with a reduced-order model. In addition, even when a large basis was used, the reduced problem was found to be susceptible to spurious oscillations in the tails of the computed pdf, and to loss of probability mass over the edges of the computational domain. These phenomena were effectively addressed by the introduction into the original PDE of artificial diffusive terms near the boundaries of the domain. While greater gains were achieved in the linear example, in the more representative, nonlinear Duffing oscillator example, the typical reduction in run time, compared to integration by a conventional adaptive Runge–Kutta algorithm, was approximately 80%. This is certainly significant; however, we must reiterate that the reduced bases were selected with the benefit of the composite solution vector **u**, and that even a modest reduction in the size of the basis used was found to require the introduction of artificial diffusive terms to stabilize the solution near the boundaries of the

computational domain. In view of these limitations, we cannot enthusiastically recommend the use of this reduction.

The second numerical method applied herein to the same discretized PDE was operator splitting, where the singular coefficient matrix **K** was written as the sum of two nonsingular matrices. This allowed the rapid computation of the joint response pdf with a uniform time step using the constant state transition matrix given by the product of the exponentials of the two component matrices. Both first- and second-order-accurate splitting schemes were considered. It was observed that, although the cost of second-order methods was not much greater, first-order splitting was usually sufficiently accurate while being simpler to program, and that separating the FE coefficient matrix **K** according to the physical source of the terms in the original PDE (i.e., convection or diffusion) was effective in the linear and Duffing oscillator examples studied. The resultant saving in computational effort was comparable to that obtained with basis reduction for the linear oscillator example, a little more than 80%, but savings of over 98% were observed when operator splitting was used to compute the joint pdf of a Duffing oscillator.

When these methods were combined, by using operator splitting to solve equations reduced on a basis of selected eigenvectors, the advantages of both techniques were retained. Run times for the solution phase of nonstationary problems were commonly 1% to 10% of those required for Runge–Kutta integration. This improvement was achieved with little cost in accuracy, although the caveats stated above regarding basis reduction apply here as well.

7. Conclusions

The class of dynamical systems for which the solution of the FP equation by FEA is tractable is small, effectively limited to those with only a few states. While the methods reported here will not, alone, allow the treatment of new, larger systems, they do represent a significant advance in the speed of the computation of the nonstationary pdfs of realistic, complex small systems. It is expected that these results will be useful to anyone needing to accelerate such calculations in a particular application or to test an alternative numerical approach.

Author Contributions: Conceptualization, L.A.B.; methodology, D.M.M.; software, C.Z., R.Z., T.H. and H.F.; validation, F.Y., C.Z., R.Z., T.H., and H.F.; formal analysis, D.M.M. and L.A.B.; data curation, F.Y. and T.H.; writing—original draft preparation, T.H. and H.F.; writing—review and editing, D.M.M. and H.L.; visualization, T.H. and H.F.; supervision, D.M.M. and H.L.; project administration, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China through Grant No. 51975525 and Grant No. 52005443; the Ministry of Science and Technology of China through Grant No. 2017YFC0306202; and the Ningbo Institute of Digital Twin through Grant No. S203.01.22.001 and Grant No. S203.01.22.002.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

The following notation is used in this manuscript:

Variables and functions

b, d	parameters of artificial diffusive terms
D	intensity of random excitation
E	expectation operator
f, $\mathbf{f}$	joint probability density function
g	function defining system dynamics
$\mathbf{I}$	identity matrix
$\mathbf{K}$, $\mathbf{M}$	global coefficient matrices
$\mathbf{k}$, $\mathbf{m}$	element coefficient matrices
$\mathbf{k}_1$, $\mathbf{k}_2$	split element matrices
$\mathbf{K}_1$, $\mathbf{K}_2$	split global matrices
N, $\mathbf{N}$	finite element shape functions
p	transition probability density function
$\mathbf{S}$	transformation matrix
t	time
$\mathbf{u}$	composite solution vector
w	Gaussian white noise input
X, $\mathbf{X}$	state variables (random)
x, $\mathbf{x}$	state variables (deterministic)
α_1, α_2	profiles of artificial diffusive terms
β	projection coefficient
Δ_t	time step
δ	Dirac delta function
ζ	damping ratio
η, $\boldsymbol{\eta}$	transformed coordinates
λ	eigenvalue
τ	normalized time
$\boldsymbol{\phi}$	eigenvector
Ω	computational domain
$\partial\Omega$	boundary of computational domain
ω_0	linearized natural frequency

Subscripts

0	initial value
e	finite element-level quantity
r, s	matrix row and column indices

Diacritical mark

$\hat{(\cdot)}$	dimension-reduced quantity

References

1. Wang, M.C.; Uhlenbeck, G.E. On the Theory of the Brownian Motion II. *Rev. Mod. Phys.* **1945**, *17*, 323–342. [CrossRef]
2. Lin, Y.K. *Probabilistic Theory of Structural Dynamics*; McGraw-Hill: New York, NY, USA, 1967.
3. Caughey, T.K. Nonlinear Theory of Random Vibrations. In *Advances in Applied Mechanics*; Yih, C.S., Ed.; Elsevier: Amsterdam, The Netherlands, 1971; Volume 11, pp. 209–253.
4. Bergman, L.A.; Heinrich, J.C. On the Moments of Time to First Passage of the Linear Oscillator. *Earthq. Eng. Struct. Dyn.* **1981**, *9*, 197–204. [CrossRef]
5. Langley, R.S. A Finite Element Method for the Statistics of Non-Linear Random Vibration. *J. Sound Vib.* **1985**, *101*, 41–54. [CrossRef]
6. Spencer, B.F., Jr.; Bergman, L.A. On the Numerical Solution of the Fokker-Planck Equation for Nonlinear Stochastic Systems. *Nonlin. Dyn.* **1993**, *4*, 357–372. [CrossRef]
7. El-Gebeily, M.A.; Shabaik, H.E.E. Approximate Solution of the Fokker-Planck-Kolmogorov Equation by Finite Elements. *Commun. Numer. Meth. Eng.* **1994**, *10*, 763–771. [CrossRef]
8. Masud, A.; Bergman, L.A. Application of Multi-Scale Finite Element Methods to the Solution of the Fokker-Planck Equation. *Comput. Meth. Appl. Mech. Eng.* **2005**, *194*, 1513–1526. [CrossRef]
9. Floris, C. Numeric Solution of the Fokker-Planck-Kolmogorov Equation. *Engineering* **2013**, *5*, 975–988. [CrossRef]

10. Kumar, P.; Narayanan, S.; Gupta, S. Finite Element Solution of Fokker-Planck Equation of Nonlinear Oscillators Subjected to Colored Non-Gaussian Noise. *Probab. Eng. Mech.* **2014**, *38*, 143–155. [CrossRef]
11. Král, R.; Náprstek, J. Theoretical Background and Implementation of the Finite Element Method for Multi-Dimensional Fokker-Planck Equation Analysis. *Adv. Eng. Softw.* **2017**, *113*, 54–75. [CrossRef]
12. Banks, H.T.; Tran, H.T.; Woodward, D.E. Estimation of Variable Coefficients in the Fokker-Planck Equations using Moving Node Finite Elements. *SIAM J. Numer. Anal.* **1993**, *30*, 1574–1602. [CrossRef]
13. Franca, L.P.; Hauke, G.; Masud, A. Revisiting Stabilized Finite Element Methods for the Advective-Diffusive Equation. *Comput. Meth. Appl. Mech. Eng.* **2006**, *195*, 1560–1572. [CrossRef]
14. Galán, R.F.; Ermentrout, G.B.; Urban, N.N. Stochastic Dynamics of Uncoupled Neural Oscillators: Fokker-Planck Studies with the Finite Element Method. *Phys. Rev. E* **2006**, *76*, 056110. [CrossRef] [PubMed]
15. Kumar, M.; Chakravorty, S.; Singla, P.; Junkins, J.L. The Partition of Unity Finite Element Approach with hp-Refinement for the Stationary Fokker-Planck Equation. *J. Sound Vib.* **2009**, *327*, 144–162. [CrossRef]
16. Sepehrian, B.; Radpoor, M.K. Numerical Solution of Non-Linear Fokker-Planck Equation using Finite Differences Method and the Cubic Spline Functions. *Appl. Math. Comput.* **2015**, *262*, 187–190. [CrossRef]
17. Brunken, J.; Smetana, K. Stable and Efficient Petrov-Galerkin Methods for a Kinetic Fokker-Planck Equation. *SIAM J. Numer. Anal.* **2022**, *60*, 157–179. [CrossRef]
18. Mahajan, A.J.; Dowell, E.H.; Bliss, D.B. Eigenvalue Calculation Procedure for an Euler-Navier-Stokes Solver with Application to Flows Over Airfoils. *J. Comput. Phys.* **1991**, *97*, 398–413. [CrossRef]
19. Mahajan, A.J.; Dowell, E.H.; Bliss, D.B. Role of Artificial Viscosity in Euler and Navier-Stokes Solvers. *AIAA J.* **1991**, *29*, 555–559. [CrossRef]
20. Dowell, E.H.; Hall, K.C.; Romanowski, M.C. Eigenmode Analysis in Unsteady Aerodynamics: Reduced Order Models. *Appl. Mech. Rev.* **1997**, *50*, 371–385. [CrossRef]
21. Hall, K.; Thomas, J.; Dowell, E. Reduced-Order Modelling of Unsteady Small-Disturbance Flows using a Frequency-Domain Proper Orthogonal Decomposition Technique. In Proceedings of the 37th Aerospace Sciences Meeting and Exhibit, Reno, NV, USA, 11–14 January 1999 ; pp. 1–11.
22. Ungnad, S.; Sachau, D. Interior near-field acoustic holography based on finite elements. *J. Acoust. Soc. Am.* **2019**, *146*, 1758–1768. [CrossRef]
23. Lötstedt, P.; Ferm, L. Dimensional Reduction of the Fokker-Planck Equation for Stochastic Chemical Reactions. *Multiscale Model. Sim.* **2006**, *5*, 593–614. [CrossRef]
24. Leonenko, G.M.; Phillips, T.N. On the Solution of the Fokker-Planck Equation using a High-Order Reduced Basis Approximation. *Comput. Meth. Appl. Mech. Eng.* **2009**, *199*, 158–168. [CrossRef]
25. Er, G.K. Methodology for the Solutions of Some Reduced Fokker-Planck Equations in High Dimensions. *Ann. Phys.* **2011**, *523*, 247–258. [CrossRef]
26. Chen, J.; Lin, P. Dimension-Reduction of FPK Equation via Equivalent Drift Coefficient. *Theor. Appl. Mech. Lett.* **2014**, *4*, 013002. [CrossRef]
27. Chen, J.; Rui, Z. Dimension-Reduced FPK Equation for Additive White-Noise Excited Nonlinear Structures. *Probab. Eng. Mech.* **2018**, *53*, 1–13. [CrossRef]
28. Zhu, H.T. Probabilistic Solution of Some Multi-Degree-of-Freedom Nonlinear Systems Under External Independent Poisson White Noises. *J. Acoust. Soc. Am.* **2012**, *131*, 4550–4557. [CrossRef]
29. Liang, C.; Zhang, J.; Lian, J.; Liu, F.; Li, X. Probabilistic Analysis for the Response of Nonlinear Base Isolation System Under the Ground Excitation Induced by High Dam Flood Discharge. *Earthq. Eng. Eng. Vib.* **2017**, *16*, 841–857. [CrossRef]
30. Chorin, A.J.; Hughes, T.J.R.; McCracken, M.F.; Marsden, J.E. Product Formulas and Numerical Algorithms. *Commun. Pure Appl. Math.* **1978**, *31*, 205–256. [CrossRef]
31. Usadi, A.; Dawson, C. 50 Years of ADI Methods: Celebrating the Contributions of Jim Douglas, Don Peaceman, and Henry Rachford. *SIAM News* **2006**, *39*, 3.
32. Hestenes, M.R. Multiplier and Gradient Methods. *J. Optimiz. Theory Appl.* **1969**, *4*, 303–320. [CrossRef]
33. Powell, M.J.D. A Method for Nonlinear Constraints in Minimization Problems. In *Optimization*; Fletcher, R., Ed.; Academic Press: New York, NY, USA, 1969; pp. 283–298.
34. Powell, M.J.D. On Search Directions for Minimization Algorithms. *Math. Program.* **1973**, *4*, 193–201. [CrossRef]
35. Strang, G. On the Construction and Comparison of Difference Schemes. *SIAM J. Numer. Anal.* **1968**, *5*, 506–517. [CrossRef]
36. Speth, R.L.; Green, W.H.; MacNamara, S.; Strang, G. Balanced Splitting and Rebalanced Splitting. *SIAM J. Numer. Anal.* **2013**, *51*, 3084–3105. [CrossRef]
37. Goldstein, T.; Osher, S. The Split Bregman Method for L1-Regularized Problems. *SIAM J. Imaging Sci.* **2009**, *2*, 323–343. [CrossRef]
38. Zorzano, M.P.; Mais, H.; Vazquez, L. Numerical Solution for Fokker-Planck Equations in Accelerators. *Physica D* **1998**, *113*, 379–381. [CrossRef]

39. Mais, H.; Zorzano, M.P. Stochastic Dynamics and Fokker-Planck Equation in Accelerator Physics. *Nuov. Cim. A* **1999**, *112*, 467–474. [CrossRef]
40. Zorzano, M.P.; Mais, H.; Vazquez, L. Numerical Solution of Two Dimensional Fokker-Planck Equations. *Appl. Math. Comput.* **1999**, *98*, 109–117. [CrossRef]
41. Duffy, M.; Chung, S.J.; Bergman, L.A. A general Bayesian Nonlinear Estimation Method using Resampled Smooth Particle Hydrodynamics Solutions of the Underlying Fokker-Planck Equation. *Int. J. Non-Linear Mech.* **2022**, *146*, 104134. [CrossRef]
42. Sirignano, J.; Spiliopoulos, K. DGM: A deep learning algorithm for solving partial differential equations. *J. Comput. Phys.* **2018**, *375*, 1339–1364. [CrossRef]
43. Al-Aradi, A.; Correia, A.; Jardim, G.; Naiff, D.D.F.; Saporito, Y. Extensions of the deep Galerkin method. *Appl. Math. Comput.* **2022**, *430*, 127287. [CrossRef]
44. Sirignano, J.; MacArt, J.; Spiliopoulos, K. PDE-constrained models with neural network terms: Optimization and global convergence. *J. Comput. Phys.* **2023**, *481*, 112016. [CrossRef]
45. Kumar, M.; Pandit, S. An efficient algorithm based on Haar wavelets for numerical simulation of Fokker-Planck equations with constants and variable coefficients. *Int. J. Numer. Meth. Heat Fluid Flow* **2015**, *25*, 41–56. [CrossRef]
46. Ankur.; Jiwari, R. New multiple analytic solitonary solutions and simulation of (2+1)-dimensional generalized Benjamin-Bona-Mahony-Burgers model. *Nonlin. Dyn.* **2023**, *111*, 13297–13325. [CrossRef]
47. Ankur.; Jiwari, R. A new error estimates of finite element method for (2+1)-dimensional nonlinear advection-diffusion model. *Appl. Numer. Math.* **2024**, *198*, 22–42. [CrossRef]
48. Strang, G. *Linear Algebra and Its Applications*, 4th ed.; Brooks: Belmont, CA, USA, 2006.
49. Glowinski, R.; Osher, S.J.; Yin, W. (Eds.) *Splitting Methods in Communication, Imaging, Science, and Engineering*; Springer: Cham, Switzerland, 2016.
50. Dormand, J.R.; Prince, P.J. A family of embedded Runge-Kutta formulae. *J. Comput. Appl. Math.* **1980**, *6*, 19–26. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.