



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Physics Letters A 344 (2005) 149–155

PHYSICS LETTERS A

[www.elsevier.com/locate/pla](http://www.elsevier.com/locate/pla)

# A numerical method for some stochastic differential equations with multiplicative noise

Charles R. Doering\*, Khachik V. Sargsyan, Peter Smereka

*Department of Mathematics and Michigan Center for Theoretical Physics, University of Michigan, Ann Arbor, MI 48109-1109, USA*

Received 10 February 2005; received in revised form 15 June 2005; accepted 15 June 2005

Available online 24 June 2005

Communicated by A.P. Fordy

## Abstract

Diffusion processes intended to model the continuous state space limit of birth–death processes, chemical reactions, and other discrete particle systems often involve multiplicative noise where the diffusion vanishes near one (or more) of the state space boundaries. Standard direct numerical simulation schemes for the associated stochastic differential equations run the risk of “overshooting”, i.e., of varying outside the meaningful state space domain where simple analytic expressions for the diffusion coefficient may take on unphysical (negative or complex) values. We propose a simple scheme to overcome this problem and apply it to an exactly soluble stochastic ordinary differential equation (SODE), and to a related parabolic stochastic partial differential equation (SPDE) that admits exact analytic solution for the stationary correlation function. Armed with these analytic benchmark solutions, we demonstrate that the scheme produces approximate solutions for the SODE with distributions that display first-order convergence in the Wasserstein metric. For the SPDE, the scheme produces first order convergence for the stationary correlation function in  $L_2$ .

© 2005 Elsevier B.V. All rights reserved.

## 1. Introduction

To motivate this study, consider a birth–death process  $N(t) \in \{0, 1, 2, \dots\}$  with birth rate  $\lambda_n = \mathcal{N}\bar{\lambda}(n/\mathcal{N}) \geq 0$  and death rate  $\mu_n = \mathcal{N}\bar{\mu}(n/\mathcal{N}) \geq 0$  where  $\bar{\lambda}(u)$  and  $\bar{\mu}(u)$  are smooth functions of their arguments and  $\mathcal{N} \gg 1$  is a population scale. That is, the

“continuum limit” of the population level is taken to be a real-valued process  $U(t) = N(t)/\mathcal{N} \in [0, \infty)$  as  $\mathcal{N} \rightarrow \infty$ . In the full continuum limit where fluctuations are (presumably) negligible,  $U(t)$  should obey the ordinary differential equation

$$\frac{dU}{dt} = f(U), \quad (1)$$

where  $f(u) = \bar{\lambda}(u) - \bar{\mu}(u)$ .

To keep some discrete fluctuation effects in a continuum description, start from the master (forward

\* Corresponding author.

E-mail address: [doering@umich.edu](mailto:doering@umich.edu) (C.R. Doering).

Kolmogorov) equation for the probability mass function,  $p_n(t)$ , of  $N(t)$ ,

$$\frac{dp_n}{dt} = \lambda_{n-1}p_{n-1} - (\lambda_n + \mu_n)p_n + \mu_{n+1}p_{n+1}. \quad (2)$$

Then the “near-continuum” limit, simply obtained via Taylor expansion keeping only the leading order effect of the fluctuations, produces a Fokker–Planck (forward Kolmogorov) equation for the density  $\rho(u, t) \approx \mathcal{N}p_{\mathcal{N}u}(t)$  of  $U(t)$  of the form

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial u} \left\{ -f(u) + \frac{1}{2} \frac{\partial}{\partial u} g(u)^2 \right\} \rho \quad (3)$$

with drift  $f(u)$  as in (1) and non-negative diffusion  $g(u)^2 = (\bar{\lambda}(u) + \bar{\mu}(u))/\mathcal{N}$ . From this we conclude that the near-continuum process  $U(t)$  is a Markov diffusion process satisfying the Itô stochastic differential equation

$$\frac{dU}{dt} = f(U) + g(U)\xi(t), \quad (4)$$

where  $\xi(t)$  is delta-correlated Gaussian white noise.<sup>1</sup>

Typically at low population levels  $\lambda_n \sim n$  and  $\mu_n \sim n$  so that  $\bar{\lambda}(u) \sim u$ ,  $\bar{\mu}(u) \sim u$  and the multiplicative noise function  $g(u) \sim \sqrt{u}$  as  $u \downarrow 0$ . This vanishing of the diffusion near the zero-population boundary ( $u = 0$  or  $U = 0$ ) of the state space is what we mean by *singular* diffusion for such problems. If  $f(u) \sim u$  near the zero-population boundary, then for small  $U$  the stochastic differential equation for this kind of system takes the form

$$\frac{dU}{dt} = \alpha U + \beta \sqrt{U} \xi(t), \quad (5)$$

where  $\alpha$  and  $\beta$  are just coefficients. Fluctuations are extremely important at these small values of  $U$ : even with a small coefficient  $\beta \sim \mathcal{N}^{-1/2}$ , the  $\sqrt{U}$  factor multiplying the noise can greatly overwhelm the drift  $\sim U$  when  $U$  is sufficiently small. Leaving questions of the non-Lipshitz nature of the generic square-root singularity aside, in this Letter we focus on issues related to the direct numerical simulation of such singular diffusions.

<sup>1</sup> We will use the standard physics notation. In the mathematical literature this stochastic differential equation would be written  $dU = f(U)dt + g(U)dW$ .

A straightforward Euler method for (6) leads to the iteration

$$U_{k+1} = U_k + \alpha U_k \Delta t + \beta \sqrt{U_k} \Delta W_k, \quad (6)$$

where  $U_k \approx U(k\Delta t)$  and the  $\Delta W_k$  are independent Gaussians with mean zero and variance  $\Delta t$ . The problem at low  $U$ -levels is that even when  $U_k > 0$  it is easy for fluctuations to drive  $U_{k+1} < 0$ , out of the sensible domain. Then because of the square root, the iteration cannot meaningfully proceed. What we propose in this Letter is a computational scheme that can overcome this difficulty. This problem has been previously recognized and addressed by some schemes (see, for example, [1,2]) to varying degrees of success [3]. The method proposed here is not so sophisticated, but is easy to implement and, as will be shown, quite effective.

We also bring our approach to bear on generalizations of such problems including spatial dependence, diffusion and noise. For example, generalizations of local birth–death, reaction, or other dynamics including spatial diffusion lead to a stochastic partial differential equations with multiplicative noise of the form

$$\frac{\partial U(x, t)}{\partial t} = D \frac{\partial^2 U}{\partial x^2} + f(U) + g(U)\eta(x, t), \quad (7)$$

where the Gaussian white noise  $\eta(x, t)$  is delta-correlated in both  $t$  and  $x$ . If  $U$  only makes sense for non-negative values and  $g(U) \sim \sqrt{U}$  for small  $U$  while  $f$  is Lipshitz, then the same kind of “overshoot” problems with simulations have to be addressed.

The rest of this Letter is organized as follows. In Section 2 we develop our approach for another, soluble, stochastic ordinary differential equation (SODE) with a square root singularity for the multiplicative noise at the state space boundaries. Given the exact solution for the stationary distribution we also study the scheme’s quantitative accuracy in terms of convergence in the Wasserstein metric. In Section 3 we apply the numerical technique to a stochastic partial differential equation (SPDE) with the same sort of multiplicative space–time white noise, and examine numerical convergence for the stationary correlation function. The final Section 4 is a brief summary of the results and discussion of applications for this methods to a problem of current interest.

## 2. A stochastic ordinary differential equation

To introduce and evaluate the scheme we consider the SODE

$$\frac{d\Phi(t)}{dt} = -\gamma\Phi(t) + \sigma\sqrt{1 - \Phi(t)^2} \xi(t) \tag{8}$$

interpreted in the Itô sense;  $\xi(t)$  is Gaussian white noise with  $\langle \xi(t) \rangle = 0$  and  $\langle \xi(t)\xi(s) \rangle = \delta(t - s)$ . The diffusion coefficient vanishes at  $\Phi = \pm 1$  so we restrict the initial condition to satisfy

$$-1 \leq \Phi(0) \leq 1. \tag{9}$$

We focus on the case  $\gamma > 0$  where the drift subsequently maintains  $\Phi \in [-1, 1]$ . (The  $\gamma = 0$  case needs to be treated separately; the process stays real only for a finite time if  $\gamma < 0$  and we will not consider this situation.) Positive  $\gamma$  can be scaled away by redefining time, so that the single remaining parameter is  $\gamma/\sigma^2$ , but we retain the parameters for consistency with the notation in [4].

Then the stationary probability density of the process  $\Phi(t)$  is

$$\rho_s(\phi) = \frac{\Gamma(\gamma/\sigma^2 + 1/2)}{\Gamma(\gamma/\sigma^2)\Gamma(1/2)} (1 - \phi^2)^{-1+\gamma/\sigma^2} \tag{10}$$

for  $\phi \in (-1, 1)$  and, of course,  $\rho_s(\phi) = 0$  for  $|\phi| > 1$ . The primary difficulty in the straightforward numerical solution of (8) is the multiplicative noise term: standard discretizations will not always respect the constraint  $|\Phi| \leq 1$ . When the numerical approximation strays outside  $[-1, 1]$  the scheme will be destroyed.

We propose an extremely simple numerical method for solving (1) that avoids this problem. Define

$$G(\phi) := \begin{cases} \sqrt{1 - \phi^2}, & |\phi| \leq 1, \\ 0, & \text{otherwise,} \end{cases} \tag{11}$$

and discretize the SODE via the elementary Euler scheme:

$$\phi_{n+1} = \phi_n - \gamma\phi_n\Delta t + \sigma G(\phi_n)\Delta W_n, \tag{12}$$

where  $\Delta t$  is the time step and  $\Delta W_n$  are independent Gaussian random variables with mean 0 and variance  $\Delta t$ . The initial value for the iterations is  $\phi_0 = \Phi(0)$ . The “truncated”  $G(\phi)$  in place of  $\sqrt{1 - \phi^2}$  avoids the overshoot problem in the sense that the iterates  $\phi_n$  always remain real even should they wander outside the

interval  $[-1, 1]$ . The deterministic dynamics then directs the variable back into the interval (this property is certainly desirable and in some cases probably necessary for this scheme). Now we investigate convergence properties for this numerical scheme.

Fig. 1 shows a comparison of the stationary distribution for the iterates of the numerical scheme (12) to the exact analytic distribution  $\rho_0(\phi)$  in (10). When  $\gamma \geq \sigma^2$  case we actually observe  $L_1$ -convergence of the probability distribution functions. The  $\gamma < \sigma^2$  case requires special attention because for small values of  $\gamma/\sigma^2$  the invariant measure approaches a combination of  $\delta$ -functions at  $\pm 1$ , so arbitrarily small errors these peaks’ positions prohibit  $L_1$ -convergence. This is the main motivation to study convergence of the distribution of the solution of the finite difference approximation (8) using the Wasserstein metric as the measure of the distance between probability distributions.

The Wasserstein metric has recently been applied to a number of problems involving flows, diffusion, and mass transport [5–7]. The Wasserstein distance  $d_p(F_1, F_2)$  of order  $1 \leq p < \infty$  between two probability distributions on  $\mathbb{R}^n$  is defined by the formula

$$d_p(F_1, F_2)^p = \inf_{H \in \mathcal{H}(F_1, F_2)} \int |x - y|^p dH(x, y), \tag{13}$$

where  $|\cdot|$  is Euclidian norm on  $\mathbb{R}^n$  and  $\mathcal{H}(F_1, F_2)$  is the set of all joint probability distributions on  $\mathbb{R}^n \times \mathbb{R}^n$  with marginals  $F_1$  and  $F_2$ . The infimum in (13) is attained by the joint distribution function  $H(x, y) = \min\{F_1, F_2\}$ . See [8,9].

Equivalently, the Wasserstein distance between two probability distributions with cumulative distribution functions  $F_1$  and  $F_2$  can be defined by:

$$d_p(F_1, F_2)^p = \int_0^1 |F_1^{-1}(t) - F_2^{-1}(t)|^p dt, \tag{14}$$

where the power  $-1$  denotes the pseudo-inverse function

$$F^{-1}(t) = \inf \{x: F(x) > t\}. \tag{15}$$

The equivalence of these definitions is proven in [8, 10]. The Wasserstein distance is a well-defined metric on the set of all probability distributions with finite  $p$ -moments [5,11]. Moreover, convergence in the Wasserstein metric is equivalent to the weak convergence plus the convergence of  $p$ th moments; see

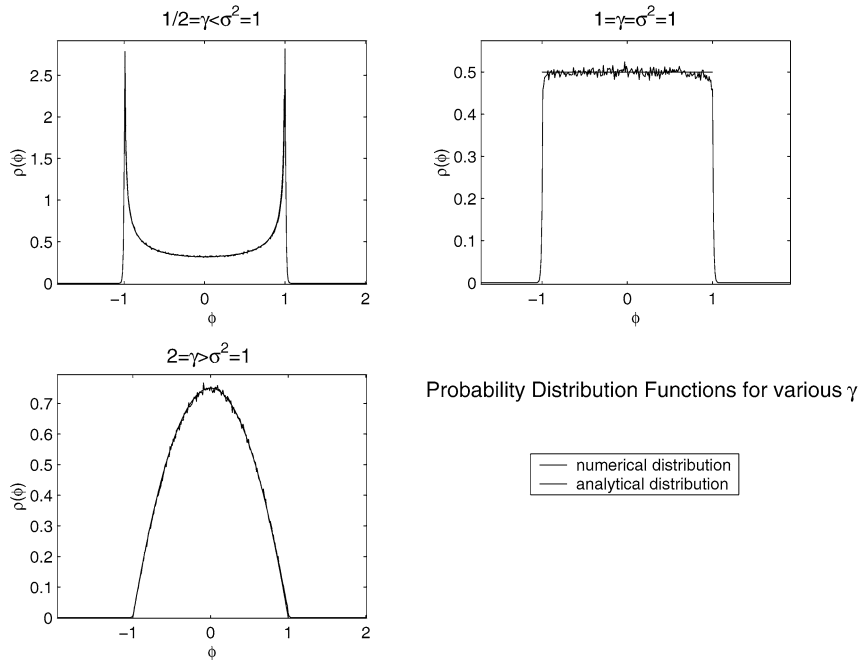


Fig. 1. Probability distribution functions, from simulation data and the exact formula (10), for several different values of  $\gamma$  with  $\sigma^2 = 1$ . The simulation statistics here correspond to  $10^6$  independent realizations.  $\Delta t = 0.01$ .

[5,12]. By Hölder’s inequality, the distances  $d_p(F_1, F_2)$  are an increasing sequence for  $p < \infty$ , so we can define Wasserstein distance for  $p = \infty$  as the limit of that sequence. (Note: in some literature these definitions are associated with the name of Kantorovich as well and the popular  $p = 2$  case is called the Kantorovich–Wasserstein metric.)

Fig. 2 shows that the stationary distribution of the iterates of the numerical scheme (12) converges to the exact stationary distribution  $\rho_0(\phi)$  of the solution of the SODE (8) as  $\Delta t \rightarrow 0$  in the  $\infty$ -Wasserstein metric (and therefore also in every  $p$ -Wasserstein metric for  $p \geq 1$ ). These data suggest that the convergence rate of the method is first order  $\Delta t$ .

As previously mentioned, Wasserstein convergence implies the convergence of moments. Hence the moments of the iterates of the discrete scheme (12) converge to those of (8), namely,

$$\langle \Phi^{2n} \rangle_{\text{stat}} := \int_{-1}^1 \phi^{2n} \rho_0(\phi) d\phi = \prod_{i=1}^n \frac{1}{1 + \frac{2\gamma}{\sigma^2(2i-1)}}, \quad (16)$$

$$\langle \Phi^{2n-1} \rangle_{\text{stat}} = 0. \quad (17)$$

Finally, we note that in the high noise (small  $\gamma/\sigma^2$ ) limit,

$$\rho_0(\phi) \rightarrow \frac{1}{2} (\delta(\phi - 1) + \delta(\phi + 1)) \quad (18)$$

as  $\sigma \rightarrow \infty$ . For the  $\gamma = 0$  case there is no unique stationary distribution for (8) but there are invariant measures of the form  $\frac{1+m}{2} \delta(\phi - 1) + \frac{1-m}{2} \delta(\phi + 1)$  where  $m$  is the mean of the initial distribution. The mean and all odd moments of this kind of distribution are  $m$  while the even moments are 1.

### 3. A stochastic partial differential equation

Consider the SPDE

$$\frac{\partial}{\partial t} \Phi(x, t) = -\gamma \Phi + \beta \frac{\partial^2 \Phi}{\partial x^2} + \sigma \sqrt{1 - \Phi^2} \eta(x, t), \quad (19)$$

where  $\eta(x, t)$  is a space–time Gaussian white noise with  $\langle \eta(x, t) \rangle = 0$  and

$$\langle \eta(x, t) \eta(y, s) \rangle = \delta(x - y) \delta(t - s).$$

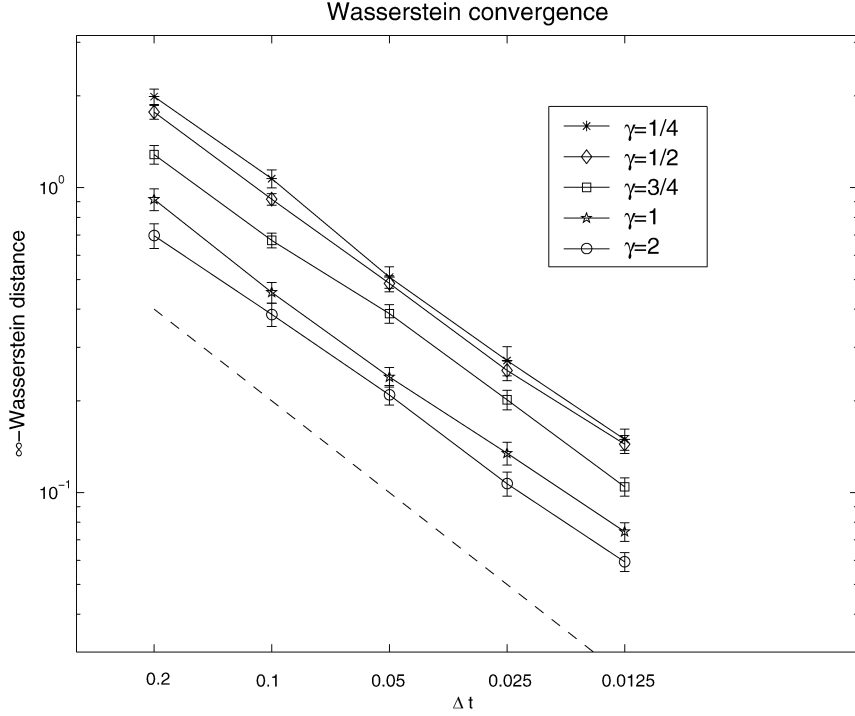


Fig. 2. The  $\infty$ -Wasserstein distance from the numerical distribution to the exact  $\rho_0(\phi)$  at time  $T = 10$ . The relaxation time toward the stationary distribution, i.e., the lowest positive eigenvalue of the associated Fokker–Planck (forward Kolmogorov) operator, is  $1/\gamma$  so  $T = 10$  is a sufficiently long time for these parameter values. Other parameters:  $\sigma = 1$  and  $10^6$  independent realizations to minimize the statistical noise with initial condition  $\Phi(0)$  is uniformly distributed on  $[0, 1/2]$ . The error bars are based on ten independent simulations. The dashed line  $\sim \Delta t$  is provided for comparison.

Again  $\gamma > 0$  and  $\beta > 0$ . We consider this equation to be the continuum limit of a spatial discretization consisting of a set of coupled Itô equations.

The full discretization on a lattice of the length  $L$  with  $N$  lattice points is:

$$\begin{aligned} \phi_i^{n+1} = & \phi_i^n - \gamma \phi_i^n \Delta t + \frac{\beta \Delta t}{h^2} (\phi_{i-1}^n - 2\phi_i^n + \phi_{i+1}^n) \\ & + h^{-1/2} \sigma G(\phi_i^n) \sqrt{\Delta t} \Delta W_i^n, \end{aligned} \quad (20)$$

where the lattice space is  $h = L/N$ ,  $G$  is the truncated function in (11), and the  $\Delta W_i^n$  are independent Gaussians with mean zero and variance  $\Delta t$ . Note that noise amplitude is rescaled by the factor  $h^{-1/2}$  to achieve the proper normalization in the spatial continuum limit. Spatial boundary conditions (periodic, Dirichlet, etc.) may be implemented as usual for such diffusion equations.

This SPDE permits the exact evaluation of the stationary correlation function [4], making it a convenient

benchmark for numerical methods [13]. In the continuous time limit, the formula for the stationary spatial covariance is

$$\begin{aligned} C_2(i, j) & := \lim_{t \rightarrow \infty} \langle \Phi_i(t) \Phi_j(t) \rangle \\ & = \langle \Phi_i \Phi_j \rangle_{\text{stat}} = \left[ (M^{-1})_{ii} - \frac{2h}{\sigma^2} \right] (M^{-1})_{ij}, \end{aligned} \quad (21)$$

where the matrix  $M = -\gamma I + \beta \Delta_h$  with  $\Delta_h$  the discrete Laplacian matrix. For periodic boundary conditions  $M$  can be inverted using discrete Fourier transform:

$$(M^{-1})_{lm} = \frac{1}{N} \sum_{k=1}^N d_k e^{\frac{2\pi i}{N}(k-1)(m-l)} \quad (22)$$

with

$$d_k = \left( -\gamma - \frac{4\beta}{h^2} \sin^2 \frac{\pi}{N}(k-1) \right)^{-1}.$$

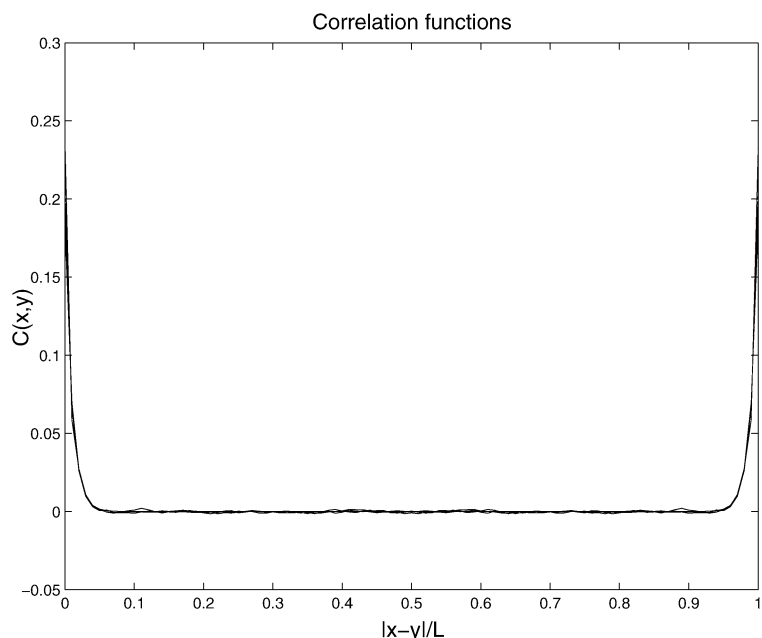


Fig. 3. Numerically computed stationary correlation functions for the parameter values listed in Table 1.

Table 1

$\Delta t$	0.2	0.1	0.05	0.025	0.0125
$N = \frac{100}{\Delta t}$	500	1000	2000	4000	8000
$L_2$ -norm	0.0423	0.0200	0.0092	0.0045	0.0021

We expect that the fully discrete covariance function generated by (20),

$$C_{\text{Discr}}(i, j) = \lim_{n \rightarrow \infty} \langle \phi_i^n \phi_j^n \rangle \quad (23)$$

converges to  $C_2(i, j)$  as the time step decreases. Table 1 shows the numerical results for the  $L_2$  norm of the difference between  $C_{\text{Discr}}(i, j)$  and  $C_2(i, j)$  for  $h = 1$  using  $10^3$  independent ensembles, for some decreasing values of  $\Delta t$ .

These data indicate first order convergence in  $L_2$  for the stationary correlation function. Fig. 3 is a plot of the correlation functions.

#### 4. Discussion

We have presented a simple and effective method for numerically approximating some singular diffusion problems arising in applications. Comparison of

the simulation results with exact solutions for the benchmark problem considered here indicate that one can expect reasonable convergence properties for appropriate statistical quantities like measures and moments.

As we have pointed out in the introduction, the square-root singularity appears naturally in applications. Precisely this form of multiplicative noise appears at one boundary in the SPDE describing directed percolation, and other numerical methods have been tailored to deal with the “overshoot” problem [14]. The method presented here has already been applied to simulate wavefront propagation in the stochastic Fisher–Kolmogorov–Petrovsky–Piscounov (SFKPP) equation,

$$\frac{\partial U(x, t)}{\partial t} = D \frac{\partial^2 U}{\partial x^2} + \gamma U(1 - U) + \sigma \sqrt{U(1 - U)} \eta(x, t). \quad (24)$$

This equation arises as the fluctuating hydrodynamic description of a long-ranged contact process [15]. It is also dual to a simple reaction–diffusion particle process [16] that has become a paradigm for the effect of discreteness and fluctuations on “pulled” fronts in mesoscopic systems [17,18]. Note that as for the

benchmark problem studied in the previous section, the square root singularity at both boundaries (in this case  $U = 0$  and 1) of the process' state space.

The deterministic ( $\sigma = 0$ ) version of (24) is the classical Fisher–Kolmogorov–Petrovsky–Piscounov equation that supports wavefront solutions travelling at a variety of speeds  $c \in [2\sqrt{D\gamma}, \infty)$ . Discreteness effects, modelled by the noise in the sense that the population scale  $\mathcal{N} \sim \sigma^{-2}$ , were conjectured to lead to a unique front speed

$$c \approx 2\sqrt{D\gamma} \left[ 1 - \mathcal{O} \left( \left[ \log \frac{\sqrt{D\gamma}}{\sigma^2} \right]^{-2} \right) \right]$$

for large  $\mathcal{N}$  (i.e., small noise amplitude  $\sigma$ ). Simulations of the SFKPP equation in [16] using the method described in this Letter, as well as other numerical techniques [19], have supported this conjecture. Moreover, in [16] it was also conjectured the wavefront speed should obey  $c \sim 2D\gamma/\sigma^2$  as  $\sqrt{D\gamma}/\sigma^2 \rightarrow 0$  (strong noise). The numerical scheme described in this Letter, as well as another recently developed method [3] have been utilized to confirm this strong noise scaling.

### Acknowledgements

We thank Esteban Moro, and our colleagues J.G. Conlon, J. Devita, L.M. Sander and R.M. Ziff for helpful discussions. This research was supported in part by NSF Awards DMS-0244419 and DMS-0207402. One of us (C.R.D.) is also grateful for funding from the Alexander von Humboldt Foundation and the hospi-

ality of International University Bremen, where part of this work was performed.

### References

- [1] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, New York, 1992.
- [2] G.N. Milstein, E. Platen, H. Schurz, SIAM J. Numer. Anal. 35 (1998) 1010.
- [3] E. Moro, Phys. Rev. E 70 (2004) 045102(R).
- [4] C.R. Doering, Phys. Lett. A 122 (3–4) (1987) 133.
- [5] R. Jordan, D. Kinderlehrer, F. Otto, J. Math. Anal. 29 (1) (1998) 1.
- [6] S.T. Rachev, Probability Metrics and the Stability of Stochastic Models, Wiley, New York, 1991.
- [7] S. Rachev, L. Ruschendorf, Mass Transportation Problems, vol. 2, Springer, New York, 1998.
- [8] J.A. Carrillo, G. Toscani, Wasserstein Metric and Large-Time Asymptotics of Nonlinear Diffusion Equations, World Scientific, Singapore, 2003.
- [9] M. Frechet, Ann. Univ. Lyon Sect. A 14 (1951) 53.
- [10] C. Villani, Topics in Optimal Transportation, American Mathematical Society, Providence, 2002.
- [11] C.R. Givens, R.M. Shortt, Michigan Math. J. 31 (1984) 231.
- [12] J.A. Cuesta-Albertos, C. Matran, A. Tuero-Diaz, J. Multivariate Anal. 60 (1997) 72.
- [13] M. Seesselberg, F. Petruccione, Comput. Phys. Commun. 74 (1993) 303.
- [14] R. Dickman, Phys. Rev. E 50 (1994) 4404.
- [15] C. Mueller, R. Tribe, Probab. Theory Related Fields 102 (1995) 519.
- [16] C.R. Doering, C. Mueller, P. Smereka, Physica D 325 (2003) 243.
- [17] E. Brunet, B. Derrida, Phys. Rev. E 56 (1997) 2597.
- [18] D.A. Kessler, Z. Ner, L.M. Sander, Phys. Rev. E 58 (1998) 107.
- [19] L. Pechenik, H. Levine, Phys. Rev. E 59 (1999) 3893.