

Looking at Fokker–Planck Dynamics with a Noisy Instrument

Gabor Schmera,¹ Adi Bulsara,¹ David Pierson,² Frank Moss,² and Enrico Di Cera³

Received July 23, 1992; final December 15, 1992

We consider the class of experiments which can be characterized by a Fokker–Planck dynamics corresponding to the overdamped motion of a state point in a suitable stochastic potential. We assume that the general form of the potential is known (or can be guessed with reasonable accuracy), but that its parameters are to be determined experimentally by measurements made with a noisy instrument. This possible method for determining the potential parameters, which exploits the system's own internal stochastic motion in order to explore rapidly its available parameter space, is substantially more efficient than traditional methods involving time averages of single point measurements, and yet does not appear to have been previously considered. The method could be important when, for example, the experiment must be completed in a limited time owing either to the expense of the experimental materials or to the temporary stationarity of the preparation, situations which are commonly encountered in experimental biochemistry and biology.

KEY WORDS: Fokker–Planck equation; instrumental noise; reaction potentials; fluctuating chemical reactions; enzyme reaction.

1. INTRODUCTION

A large class of nonlinear dynamical systems which are subject to either inherent or external fluctuations, or noise, are describable by Fokker–Planck equations.^(1,2) Physical realizations run from continuously stirred chemical reactions, to lasers, liquid crystals, or semiconductors, to neurons

¹ NCCOSC-RDT&E Division, Materials Research Branch, San Diego, California 92129.

² Department of Physics, University of Missouri at St. Louis, St. Louis, Missouri 63121.

³ Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110.

and excitable membranes all driven by noise.^(1, 2) In many cases it may be of importance to measure experimentally certain parameters which determine the state point motions of the system, for example, parameters in a chemical reaction, such as ratios of rate constants, which determine the temporal evolution of the reaction coordinate. We assume such parameters to be measurable quantities. If the probability densities experimentally obtained from measurements of the time series of the state point $x_1(t)$ are describable as approximate solutions of a Fokker-Planck equation, then the parameters associated with the potential function $U(x_1)$ are in principle obtainable and of interest. It is our purpose in this paper to explore, by means of purely numerical experiments, how such parameters could be obtained from experimental data on $x_1(t)$ and the accuracy one might expect from such measurements. Moreover, we assume that the measurement of $x_1(t)$ is carried out with a noisy instrument whose noise is well characterized, that is, its probability density and intensity are well known, for example, by means of separate measurements.

A schematic of such an experiment might be rendered as shown in Fig. 1. The experimental system, driven by its own internal noise of intensity D_1 , generates the time series $x_1(t)$, which is detected and measured by the instrument adding noise $x_2(t)$ in the process. The output of the instrument is $x(t) = x_1 + x_2$, which can be digitized and input to a computer. The computer obtains an ensemble average of the experimental probability density $P_e(x)$, which we assume to be stationary over the course of the experiment. For simplicity, both noises are assumed to be white (that is, of infinite bandwidth) and Gaussian, which is often a good approximation in the case of the instrument noise, but is rarely so in the case of the internal noise of the system. It remains a future project to generalize this work to include the case of colored system and/or instrument noise.

Our task is to guess the *form* of the potential $U(x_1)$. This leads to an ideal probability density $P_1(x_1)$ containing the unknown parameters of U . Convoluting P_1 with the known instrument noise, we obtain an ideal density $P(x)$, which is then to be matched, using nonlinear least squares

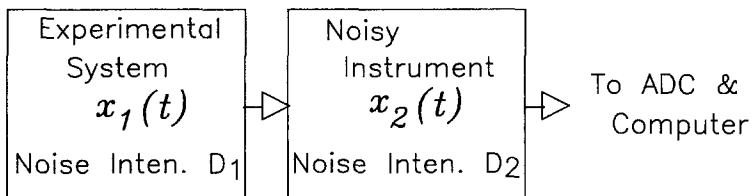


Fig. 1. A schematic representation of the method.

fit, to the experimentally obtained version $P_e(x)$. The parameters are obtained from the fit.

This technique stands in contrast to traditional methods of measuring the potential parameters by making painstaking time averages of $x(t)$ for each member of a set of changed experimental conditions.^(3,4) Such measurements are, of necessity, made in nonequilibrium conditions which may also change the potential parameters. Moreover, many such experiments, especially on enzyme reactions of importance in biochemistry, involve expensive materials. In other applications, particularly in biology, the system may not remain for long in a stationary state. In both cases the data need to be obtained as rapidly and efficiently as possible. The present technique, wherein all parameters are obtained from a single time series, measured in the steady state (if not in equilibrium) could therefore be of substantial importance in these cases.

In this paper, as a preliminary exercise, we investigate the efficacy of this method using an idealized potential, the standard quartic, for which an exact solution of the Fokker–Planck equation is available. Numerically generated estimates of the potential parameters can therefore be compared to the exact values. Subsequently, we test the method using a potential function of specific importance in enzyme biochemistry.⁽⁵⁾ Our strategy is to associate a Langevin equation to the Fokker–Planck equation containing the potential, digitally simulate the Langevin equation to produce a noisy dynamics which mimics the expected experimental result, and add the instrumental noise, all of which results in a noisy time series whose probability density represents the experimental data. The nonlinear least squares fit is then performed between this “experimental” probability density and the “ideal” or guessed one, in this test case obtained as an exact solution of the Fokker–Planck equation. Our question is, how accurately can the potential parameters be obtained, given a certain statistical accuracy for the experimental probability density and given a certain (known) instrumental noise intensity?

This paper is organized as follows. In Section 2 we describe the method and lay out the equations employed for testing the method. In Section 3 we outline the numerical and Monte Carlo procedures used in generating the substitute experimental data, and we discuss the estimation of the potential parameters and questions of accuracy. In Section 4 we describe the numerical production of test data for an example enzyme reaction. We consider a widely used potential function appropriate to this reaction, and we show the method is used to extract the parameters of the potential. In both the enzyme and standard quartic cases the Fokker–Planck equation can be solved exactly, but the convolution integral of the solution with the instrument noise must be performed numerically. We

outline, however, an approximate procedure for estimating solutions. In Section 5 we summarize our work and make some suggestions for the possible future development and exploitation of the method.

2. DESCRIPTION OF THE METHOD

For the purpose of testing this method, and to gain some insight into its accuracy, we have assumed a particularly simple, one-dimensional, bistable potential—the standard quartic—which has been well studied using Fokker–Planck methods (see, e.g., ref. 6):

$$U(x_1) = ax_1^4 - bx_1^2; \quad a, b > 0 \quad (1)$$

The Langevin equation, which describes the temporal evolution of the state point $x_1(t)$, is given by

$$\dot{x}_1 = -4ax_1^3 + 2bx_1 + \zeta(t) \quad (2)$$

where the internal system noise $\zeta(t)$ is Gaussian and white with zero mean,

$$\langle \zeta(t) \zeta(s) \rangle = 2D_1 \delta(t-s) \quad (3)$$

With these conditions, the analogous Fokker–Planck equation

$$\frac{\partial P_1(x_1, t)}{\partial t} = -\frac{\partial}{\partial x_1} [(-4ax_1^3 + 2bx_1) P_1] + D_1 \frac{\partial^2}{\partial x_1^2} P_1(x_1, t) \quad (4)$$

can be solved exactly for the stationary density,

$$P_1(x_1) = N^{-1} \exp(-Ax_1^4 + Bx_1^2) \quad (5)$$

where $A = a/D_1$ and $B = b/D_1$ are the potential parameters to be determined. The normalization is given by

$$\begin{aligned} N &= 2 \int_0^1 \frac{1}{x_1} \exp[-A(\ln x_1)^4 + B(\ln x_1)^2] dx_1 \\ &= \frac{\pi}{(2A)^{1/4}} \exp\left(\frac{B^2}{8A}\right) Q_{-1/2}\left(-\frac{B}{(2A)^{1/2}}\right) \end{aligned} \quad (6)$$

where $Q_{-1/2}$ is a parabolic cylinder function.

Now we add the instrumental noise, so that the experimentally

accessible time series is $x(t) = x_1(t) + x_2(t)$, where x_2 is Gaussian and white with intensity D_2 defined by

$$P_2(x_2) = \frac{1}{2(\pi D_2)^{1/2}} \exp\left(-\frac{x_2^2}{4D_2}\right) \tag{7}$$

and where the variance is $\sigma_2^2 = 2D_2$.

Now the probability density available at the output of the instrument is the convolution,

$$\begin{aligned} P(x) &= P_1 \text{ conv } P_2 \\ &= \frac{1}{2N(\pi D_2)^{1/2}} \int_{-\infty}^{\infty} \exp\left[-At^4 + Bt^2 - \frac{(x-t)^2}{4D_2}\right] dt \\ &= \frac{2}{N(4\pi D_2)^{1/2}} \exp\left(-\frac{x^2}{4D_2}\right) \\ &\quad \times \int_0^1 \frac{1}{t} \exp\left[-A(\ln t)^4 - \left(\frac{1}{4D_2} - B\right)(\ln t)\right] \cosh\left(\frac{x}{2D_2} \ln t\right) dt \end{aligned} \tag{8}$$

It is this density which will be fit to the experimental density $P_e(x)$ in order to extract the parameters A and B . Of course in the real situation, that is, when dealing with a real experiment, we would have to guess the form of the potential $U(x)$, but here, for the purpose of testing the feasibility, we can “guess” the exact result. Obviously, the overall accuracy of the method obtained in any practical situation will depend on the goodness of this guess.

3. NUMERICAL GENERATION OF TEST DATA AND EXAMPLE RESULTS FOR THE STANDARD QUARTIC POTENTIAL

In order to generate a time series which would play the role of experimental data, we have numerically integrated Eq. (2) for $a = 0.25$ and $b = 0.5$. This was done with completely standard techniques using the Heun algorithm for the integration and the Box-Mueller algorithm to generate the Gaussian white noise.^(7,8) The step size was 0.001, and the number of steps used in the integration was 3×10^6 . To the time series $x_1(t)$ the second Gaussian white noise $x_2(t)$ was added, and the probability density of the sum $P(x)$ was generated. The number of bins in the density was 1000. This result is shown in Fig. 2 for several values of D_2 , but with $D_1 = 0.1$, which was a constant throughout this work. This value of D_1 results in $A = 2.5$ and $B = 5.0$, which play the roles of the unknown potential parameters. The ideal, or “guessed,” function for $P(x)$ was then fitted to the numerically

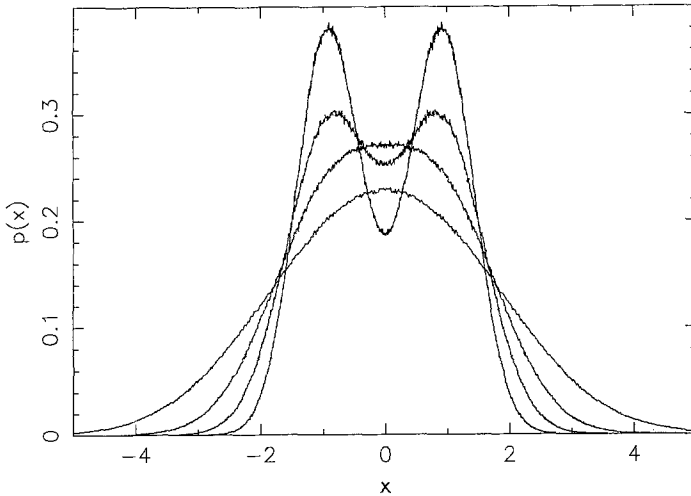


Fig. 2. Numerically generated probability densities from Eq. (2) with internal noise $D_1 = 0.1$ and with varying amounts of added instrumental noise (reading top to bottom): $D_2 = 0.1, 0.2, 0.4,$ and 1.0 . Note that the residual bistability is evident only in the cases of $D_2 = 0.1$ and 0.2 .

generated, or “experimental,” result using the Levenberg–Marquardt nonlinear least squares method (see, e.g., ref. 9). Thus, if successful, our fitting procedure should result in approximations to the aforementioned exact values of A and B . We have generated $P(x)$ for four different values of D_2 . Note that the largest value of D_2 is ten times larger than D_1 , and completely obscures the fact that the system density $P_1(x)$ is bistable.

In order to obtain an estimate of the accuracy of the method, a Monte Carlo simulation was carried out. For all the same parameters, Eq. (2) was numerically solved 100 times, to generate 100 sample densities $P(x)$. The nonlinear least squares fit to the ideal density was carried out for each case, and from these results we were able to generate means, medians, and standard deviations (about the mean) for the two fitting parameters A and B . The results are shown in Table I. The distributions of A and B obtained

Table I. Results of Monte Carlo Simulation for $A \equiv 2.50$ and $B \equiv 5.0$ for 100 Trials with $D_1 = 0.1$

| D_2/D_1 | A_{med} | σ_A | B_{med} | σ_B |
|-----------|------------------|------------|------------------|------------|
| 1.0 | 2.5 | ± 0.1 | 5.0 | ± 0.1 |
| 2.0 | 2.5 | ± 0.1 | 5.1 | ± 0.2 |
| 4.0 | 2.6 | ± 0.2 | 5.1 | ± 0.4 |
| 10.0 | 2.6 | ± 0.8 | 5.0 | ± 1.5 |

from the simulation are not Gaussian, nor even symmetrical about the mean. Using the means to represent the values of the parameters thus results in a bias (in this case, positive) such that the mean value systematically increased with increasing D_2 . We found, however, that there was no systematic bias in the median values, so those are quoted in Table I. The standard deviations quoted were nevertheless taken about the mean values. The results of Table I indicate that, at least in this numerical experiment, the potential parameters can be obtained with satisfactory accuracy by nonlinear least squares fit to a noisy density *even when the instrumental noise is a factor of ten larger than the internal noise*, as indicated by the last entry in Table I.

Even though this result seems quite encouraging, it is necessary to bear in mind that this numerical test was carried out under ideal, and perhaps unrealistic, conditions. First, and probably most important, the *exact* form of the system density function, Eq. (5), was known, thus eliminating a significant uncertainty which might arise in a real experiment. Second, in the course of the Monte Carlo test, this density was generated numerically to rather high accuracy, as indicated by the small fluctuations about the trend curves shown in Fig. 2. In an actual experiment, this density would be measured and averaged for some limited amount of time, and could be expected to show much larger fluctuations.

In order to gain some insight into how the method might perform with a potential function which might actually be encountered in biochemistry experiments, in the next section we apply the method to a potential representation of an enzyme reaction obeying Michaelis-Menten kinetics.⁽⁵⁾

4. APPLICATION OF THE METHOD TO AN ENZYME KINETICS

As a more realistic application, we now consider a nonlinear Langevin equation, which has been used previously⁽⁵⁾ to represent an enzyme kinetics,

$$\dot{x}_1 = -\frac{ax_1}{b+x_1} + (2D_1)^{1/2} \xi(t) \quad (9)$$

where x_1 is the substrate concentration ($0 \leq x_1 \leq \infty$), a is the maximum catalytic rate of the enzyme, and b is the substrate concentration at which the catalytic rate is half of its maximum value ($a, b \geq 0$). There is a reflecting boundary at $x_1=0$. The potential function, Fokker-Planck equation, and its solution, the stationary probability density function $P_1(x)$, corresponding to this Langevin equation are

$$U(x_1) = ax_1 - ab[\ln(b + x_1)] \quad (10)$$

$$\frac{\partial P_1(x_1, t)}{\partial t} = -\frac{\partial}{\partial x_1} \left[-\frac{ax_1}{b + x_1} P_1(x_1, t) \right] + D_1 \frac{\partial^2}{\partial x_1^2} P_1(x_1, t) \quad (11)$$

and

$$P_1(x_1) = N^{-1} \exp[-\beta x_1(b + x_1)^\alpha] \quad (12)$$

where $\alpha = ab/D_1$, $\beta = a/D_1$, and the normalization constant is $N = \beta^{-(\alpha+1)} e^{a\beta} \Gamma(1 + \alpha, \beta a)$, with Γ representing the incomplete gamma function.

Convolving P_1 with the Gaussian instrumental noise P_2 results in the density,

$$P(y) = N^{-1} (\gamma/\pi)^{1/2} \int_0^\infty (b+x)^\alpha \exp[-\beta x - \gamma(y-x)^2] dx \quad (13)$$

where α and β defined above are the parameters to be found from the fit for this case, and $\gamma = (4D_2)^{-1}$. This density can only be computed in an approximate form, which can be found by rewriting it as

$$P(y) = N^{-1} (\gamma/\pi)^{1/2} \exp[-\beta(y - \beta/4\gamma)] I(y) \quad (14)$$

and defining

$$I(y) = \int_0^\infty (b+x)^\alpha \exp[-\gamma(x-z)^2] dx \quad (15)$$

where $z = y - \beta/2\gamma$. The Gaussian term in the integral above is peaked around z , which can be positive or negative depending on the magnitude of y relative to $\beta/2\gamma$. When $z \geq 0$ the prefactor $(b+x)^\alpha$ can be expanded around z up to second order to yield

$$(b+x)^\alpha \simeq (b+z)^\alpha + \alpha(b+z)^{\alpha-1} (x-z) + \alpha(\alpha-1)(b+z)^{\alpha-2} (x-z)^2/2$$

When $z < 0$ the prefactor is expanded around $x=0$. Such a separation is made possible by the fact that $I(y)$ is computed for $x \geq 0$. For large enough instrumental noise D_2 the value of z becomes negative, and the resulting value of $I(y)$ can be a very small number, since most of the Gaussian function is contained in the negative half-plane. The case for small D_2 gives to large and positive z , so that one must be concerned with the extremities of $P(y)$, and any discrepancy arising from the approximation is likely to be negligible. It should be pointed out that for large D_2 the auxiliary condition $\alpha \leq 1$ must be met for the approximation to hold well, while no such

condition need be considered for small D_2 . For $\alpha > 1$ the prefactor grows rapidly, but the net contribution to the integral remains bounded. Using the Taylor expansion as above, $I(y)$ becomes

$$I(y) \exp(\gamma z^2) = (2\gamma)^{-1/2} (b+z)^\alpha Q_{-1}(\omega) + (2\gamma)^{-2} \alpha (b+z)^{\alpha-1} Q_{-2}(\omega) + \frac{1}{2} (2\gamma)^{-3/2} \alpha(\alpha-1) (b+z)^{\alpha-2} Q_{-3}(\omega) \quad \text{for } z \geq 0 \quad (16)$$

and

$$I(y) \exp(\gamma z^2) = (2\gamma)^{-1/2} b^\alpha D_{-1}(\omega) + (2\gamma)^{-2} \alpha b^{\alpha-1} Q_{-2}(\omega) + \frac{1}{2} (2\gamma)^{-3/2} \alpha(\alpha-1) b^{\alpha-2} Q_{-3}(\omega) \quad \text{for } z < 0 \quad (17)$$

where $\omega = -(2\gamma)^{1/2} z$, and the Q 's are parabolic cylinder functions. We do not here use this approximation, but only include it for general interest and utility. The analysis here proceeds by numerical integration of Eqs. (9) and (13).

In order to play the rôle of an experiment, the densities $P(y)$ were constructed by Monte Carlo simulation. As in the previous section, a stretch of 10^6 values of x was simulated by numerical integration of Eq. (9). Values of $a = 1.0$ and $b = 1.0$ ($\alpha = 10.0$ and $\beta = 10.0$) were chosen for the parameters, and also as before $D_1 = 0.1$. The density $P(y)$ so obtained is shown in Fig. 3 for three different values of the instrumental noise D_2 ranging to a maximum value of 10.

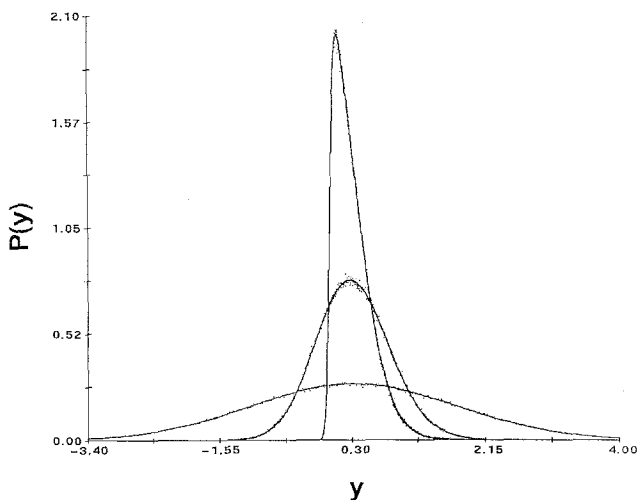


Fig. 3. Probability density function $P(y)$ for the nonlinear Langevin equation which describes the enzyme kinetics. The values of α , β , and D_1 are 10.0, 10.0, and 0.1, respectively. The curves represent the values of D_2/D_1 as follows: 0.01 (top), 0.1 (middle), and 1.0 (bottom).

The nonlinear least squares fit was applied in both cases in order to obtain best fit estimates of D_2 , b , and β in the case of the nonlinear Langevin equation and D_2 and α in the linear case. The nonlinear fit was done by minimizing $\chi^2 = \sum_i [P_e(y_i) - P(y_i)]^2 \sigma_i^{-2}$ for each set of parameter values using a version of the Levenberg–Marquardt algorithm.⁽⁹⁾ Here P_e is the value obtained from numerical integration of the Langevin equation (or from experiment), P is the theoretical or “guess” function, σ_i is the standard deviation associated with $P_e(y_i)$, and y_i is the location (midpoint) of the i th bin. For each value of D_2 , 100 density functions were constructed and analyzed. The results of these fits were collected in order to obtain estimates of the means, medians, and standard deviations which might be expected from such a procedure.

The results of the Monte Carlo simulations are shown in Table II. In the linear case the simulated values of α are recovered with reasonable accuracy and insignificant bias up to an instrumental noise intensity which exceeds the internal noise by 100 times. Even in the extreme case where $D_2/D_1 = 80$, a bias of only 8% is observed, although the standard deviation of α becomes more than 30%. Consistent with the basic principles of nonlinear least squares fitting techniques,⁽⁹⁾ parameter resolution is more difficult in the nonlinear case, as indicated by the rather large values of the standard deviations for $D_2/D_1 > 10$. This table indicates, however, that b and β can still be resolved with reasonable accuracy even when the instrument noise far exceeds the internal noise of the system.

Table II. Results of Monte Carlo Simulation of Nonlinear and Linear Dynamics with $D_1 = 0.1$ ^a

| D_2/D_1 | Enzyme (nonlinear) dynamics | | Linear dynamics |
|-----------|---|-------------------------------|---------------------------------------|
| | $\alpha_{\text{med}} \pm \sigma_\alpha$ | $b_{\text{med}} \pm \sigma_b$ | $\beta_{\text{med}} \pm \sigma_\beta$ |
| 0.01 | 10.0 ± 0.1 | 1.0 ± 0.2 | 10.0 ± 1.2 |
| 0.1 | 10.0 ± 0.1 | 1.0 ± 0.2 | 10.1 ± 1.2 |
| 1.0 | 10.0 ± 0.1 | 1.0 ± 0.2 | 10.3 ± 1.3 |
| 10.0 | 10.0 ± 0.3 | 1.1 ± 0.4 | 10.3 ± 2.6 |
| 50.0 | 9.8 ± 1.7 | 0.9 ± 0.6 | 9.4 ± 3.8 |
| 80.0 | 10.0 ± 3.5 | 1.1 ± 0.7 | 9.1 ± 4.3 |

^a The median and standard deviations are given as determined by 100 realizations for exact parameter values, $\alpha = 10.0$, $b = 1.0$, and $\beta = 10.0$. In order to provide a comparison to an extremely simple system, we have applied the procedure also to the linear Langevin equation, $\dot{x} = -ax + (2D_1)^{1/2} \xi(t)$, again with $\alpha = a/D_1$, for the values $a = 1.0$ and $D_1 = 0.1$ and for the same range of values of D_2 .

5. CONCLUSIONS AND DISCUSSION

We have shown that two unknown parameters of a one-dimensional potential in which there exists a Fokker–Planck dynamics in the large damping limit can be determined even in the presence of a large external noise. The method used is that of a nonlinear least squares fit of an experimentally obtained probability density to an estimated density resulting from a guess of the form of the potential. The method was applied to the standard quartic potential, for which an exact solution exists as a benchmark. It was further illustrated by application to a potential characteristic of catalytic enzymatic reactions. The numerical simulations presented here indicate that the potential parameters can be estimated with considerable accuracy. It therefore seems possible to exploit the considerable amount of information encapsulated by a probability density function obtained as the convolution of properties inherent to a stochastic dynamics with experimental error arising from instrumental noise.

The approach developed here for the resolution of model parameters in two prototypic examples of general interest may prove of value in the analysis of other experimental systems where the effects of internal noise need to be decoupled from the instrumental noise. The ideal cases treated here in a strictly numerical experiment seem to result in considerably more accurate estimates of the parameters than would probably be possible in a real experiment. For example, we have made no studies on the effects of the inevitably nonideal nature of real experimental data. In order to further develop this method and to test its potential usefulness, the effects of asymmetries, shifts, and deviations of functional form of the actual data from the “guess” of $P(x)$ need to be studied, as well as the simple statistical accuracy of limited samples.

ACKNOWLEDGMENTS

This work was supported by grants from the Office of Naval Research N00014-90-J-1327 (F.M.) and N00014-90-AF-00001 (A.B.) and by the National Science Foundation DMB9104963 (E.D.C.). We are grateful to Andre Longtin for a valuable discussion on noise reduction.

REFERENCES

1. W. Horsthemke and R. Lefever, *Noise-Induced Transitions: Theory and Applications in Physics, Chemistry and Biology* (Springer-Verlag, Berlin, 1984).
2. H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications* (Springer-Verlag, Berlin, 1989).

3. G. E. P. Box and G. M. Jenkins, *Time Series Analysis* (Holden Day, Oakland, California, 1976).
4. L. Brand and M. L. Johnson, *Numerical Computer Methods* (Academic Press, San Diego, California, 1992).
5. E. Di Cera and P. E. Phillipson, *J. Chem. Phys.* **93**:6006 (1990).
6. F. Moss and P. V. E. McClintock, *Noise in Nonlinear Dynamical Systems*, Vols. 1-3 (Cambridge University Press, Cambridge, 1989).
7. A. Greiner, W. Strittmatter, and J. Honerkamp, *J. Stat. Phys.* **51**:95 (1988).
8. Thomas C. Gard, *Introduction to Stochastic Differential Equations* (Marcel Dekker, New York, 1988).
9. J. Bard, *Nonlinear Parameter Estimation* (Academic Press, New York, 1974).
10. E. Di Cera and P. E. Phillipson, *J. Chem. Phys.* **93**:6006 (1990).