Exact results in the large system size limit for the dynamics of the chemical master equation, a one dimensional chain of equations

A. Martirosyan[1] and David B. Saakian[2,3,4,*]

[1]*Yerevan State University, Alex Manoogian 1, Yerevan 375025, Armenia*

[2]*Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan*

[3]*Yerevan Physics Institute, Alikhanian Brothers St. 2, Yerevan 375036, Armenia*

[4]*National Center for Theoretical Sciences: Physics Division, National Taiwan University, Taipei 10617, Taiwan*

(Received 27 March 2011; revised manuscript received 14 June 2011; published 12 August 2011)

We apply the Hamilton-Jacobi equation (HJE) formalism to solve the dynamics of the chemical master equation (CME). We found exact analytical expressions (in large system-size limit) for the probability distribution, including explicit expression for the dynamics of variance of distribution. We also give the solution for some simple cases of the model with time-dependent rates. We derived the results of the Van Kampen method from the HJE approach using a special ansatz. Using the Van Kampen method, we give a system of ordinary differential equations (ODEs) to define the variance in a two-dimensional case. We performed numerics for the CME with stationary noise. We give analytical criteria for the disappearance of bistability in the case of stationary noise in one-dimensional CMEs.

DOI: 10.1103/PhysRevE.84.021122 PACS number(s): 02.50.Ey, 87.10.−e

I. INTRODUCTION

The statistical physics of a living cell requires a theory for chemical reactions with few molecules [1,2]. One of the mathematical tools used here is the chemical master equation (CME) [3,4], describing the dynamics of probability $P(X,t)$ of having different (integer) numbers X of molecules. The same CME could be applied to other areas of science as well, even in the financial market theory [5]. Recently, Ge and Qian [6] considered a kinetic model [7] for the phosphorylation-dephosphorylation cycle in the cell, and the corresponding CME was investigated. We considered the model with the bistability phenomenon, and derived some results for the existence of bistability. Here we solve exactly the dynamics of the CME, a very important topic according to [6]. The accurate (exact) solution of the CME dynamics is important for financial market modeling [5]. In this paper we will derive the exact dynamics for the CME.

The master equation is formulated as a system of linear differential equations for $P(X,t)$ of having X molecules, $0 \leqslant X \leqslant N$, where N is a large integer:

$$\frac{dP(X,t)}{Ndt} = R_+\left(\frac{X-1}{N}\right)P(X-1,t) + R_-\left(\frac{X+1}{N}\right)P(X+1,t) - \left[R_+\left(\frac{X}{N}\right) + R_-\left(\frac{X}{N}\right)\right]P(X,t). \quad (1)$$

Here R_+ is the growth rate and R_- is the degradation rate.

Actually, we should modify the equation at the border:

$$\frac{dP(0,t)}{Ndt} = R_-\left(\frac{1}{N}\right)P(1,t) - R_+(0)P(0,t) \quad (2)$$

and

$$\frac{dP(N,t)}{Ndt} = R_+\left(1-\frac{1}{N}\right)P\left(1-\frac{1}{N},t\right) - R_-(1)P(N,t). \quad (3)$$

The large parameter N describes the system volume.

Similar master equations have been considered in [8] and evolution theory in [9,13]. Let us introduce the coordinate x and function $p(x,t)$:

$$x = X/N, \quad p(x,t;N) \equiv NP(X,t). \quad (4)$$

Assuming that the probability distribution is a smooth function of x,

$$P(X+1,t) - P(X\pm 1,t) \ll 1, \quad (5)$$

one obtains the Fokker-Planck equation for the model by Eq. (1).

The investigation of CMEs via Fokker-Planck equations meets some problems [14,15]. An alternative approach is to assume that $p(x,t)$ is not a smooth function of x, but the function $u(x,t)$ is, where

$$p(x,t;N) = \exp[Nu(x,t)]. \quad (6)$$

Thus the $p(x,t;N)$ might be unsmooth in the limit of $N \to \infty$, but still have smooth $u(x,t)$. Such an ansatz has been assumed for the statics in [8], and for the dynamics in [11,12], while considering the evolution models. This ansatz gives the solution of the dynamics and steady state with accuracy $O(1/N)$, while the approximation of the master equation via the Fokker-Planck equation, assuming the smoothness of $p(x,t)$, is certainly wrong. We already discussed this topic in [11], and then again in [13].

In Sec. II we calculate the dynamics of variance for population distribution, using HJE. In Sec. III we solve the CME dynamics for the simple case of time-dependent rates. In the Appendix we rederive the variance of distribution using the Van Kampen method. We also used that method to calculate the variance for a general one-dimensional (1D) multistep CME,

*saakian@yerphi.am

as well as give ODEs to derive the variance in the case of two-dimensional (2D) CMEs.

II. THE MASTER EQUATIONS WITH CONSTANT RATES

A. Hamilton-Jacobi equation for the chemical master equation

Using ansatz by Eq. (6), the model equation (1) can be written as Hamilton-Jacobi equations for $u \equiv \ln\, p(x,t)/N$,

$$\frac{\partial u}{\partial t} + H(u',x) = 0, \tag{7}$$

where $u' = \partial u/\partial x$,

$$H(u',x) = R_+(x) + R_-(x) - R_+(x)e^{-u'} - R_-(x)e^{u'}. \tag{8}$$

Equations (7) and (8) were derived in [16,17], where the corresponding Hamilton equations have been investigated mainly to calculate the time of extinction from metastabile state. We are interested in the investigation of the whole distributions, using the traditional mathematical method of characteristics. To solve the HJE, we consider the Hamilton equation for x and the corresponding momentum, getting the system for the characteristics [18,19]

$$\begin{aligned} \dot{x} &= H_v(x,v) = R_+(x)e^{-v} - R_-(x)e^{v}, \\ \dot{v} &= -H_x(x,v), \\ \dot{u} &= v\, H_v(x,v) - H(x,v) = v\dot{x} + q, \end{aligned} \tag{9}$$

subject to initial conditions $x(0) = x_0$, $v(0) = v_0(x_0)$, and $u(0) = u_0(x_0)$. Here $v := \partial u/\partial x$, $v_0(x) := u_0'(x)$, and $q := \partial u/\partial t$. The respective solution to Eq. (9) in (x,t) space is called the *characteristic* of Eq. (7).

Our Hamiltonian is time independent. Then Eqs. (7) and (9) result in

$$\dot{q} = 0. \tag{10}$$

Along the characteristic $x = x(t)$ the variable q is constant, so q is selected to parametrize these curves.

Consider the equation

$$q = -[R_+(x) + R_-(x)] + R_+(x)e^{-v} + R_-(x)e^{v}. \tag{11}$$

It has a solution for

$$q \geqslant 0, \tag{12}$$

if at some point

$$R_+(x) = R_-(x), \tag{13}$$

and we take $q \geqslant 0$.

Using Eq. (11), we transform the first equation in Eqs. (9) into

$$\dot{x} = \pm\sqrt{(q + R_+ + R_-)^2 - 4R_+R_-}. \tag{14}$$

Consider the following initial distribution:

$$u_0(x) = -a(x - x_0)^2, \tag{15}$$

with large a. The maximum of the distribution corresponds to the point $u' = 0$, therefore $q = 0$. Thus for the maximum of distribution we should consider a characteristic with $q = 0$.

Integrating Eq. (14), we derive

$$T = \int_{x_0}^{x} \frac{dy}{R_+(y) - R_-(y)}. \tag{16}$$

Such an equation was derived in [6].

We can now define the dynamics of the full distribution. Let us define the function

$$T(x,q) = \int_{x_0}^{x} \frac{dy}{\sqrt{[q + R_+(y) + R_-(y)]^2 - 4R_+(y)R_-(y)}}. \tag{17}$$

To calculate $u(x,t)$ we first calculate q for the given x from the equation

$$T(x,q) = t. \tag{18}$$

Equation (18) defines an implicit function

$$q = Q(x,t). \tag{19}$$

B. The elasticity

It is important to calculate $u''(x,t)$ at the point of the maximum of distribution, the "elasticity" [6].

We calculate $q_x' \equiv \partial^2 u/\partial x \partial t$ from Eq. (19):

$$\frac{\partial q}{\partial x} = \frac{\partial Q(x,t)}{\partial x}. \tag{20}$$

We can calculate the last derivative at fixed t from the expression

$$T(x,q) = \text{const}, \quad \frac{\partial T(x,q)}{\partial x} + \frac{\partial T(x,q)}{\partial q} q_x' = 0. \tag{21}$$

It is equivalent to calculate q_x' from Eq. (17) at fixed t. Thus we get the following equation for q' at the point of maximum $[u'(x,t) = 0]$:

$$\frac{1}{R_+(x) - R_-(x)} - q' \int_{x_0}^{x} \frac{dy\,[R_+(y) + R_-(y)]}{[R_+(y) - R_-(y)]^3} = 0. \tag{22}$$

From Eq. (11) we can obtain

$$q_x' = -[R_+(x) - R_-(x)]v'. \tag{23}$$

Thus eventually we get

$$\frac{-1}{v_x'} = [R_+(x) - R_-(x)]^2 \int_{x_0}^{x} \frac{dy\,[R_+(y) + R_-(y)]}{[R_+(y) - R_-(y)]^3}. \tag{24}$$

Equation (24) is the main result of our work.

A comparison of the analytical results with the numerics is given in Fig. 1.

C. Probability distribution

Equation (17) defines $q(x,t)$ for given x,t; we can define $v(x,t)$ also using Eq. (7). To calculate $u(x,t)$ at the given (x,t), let us consider the trajectory of points $(x(\tau),\tau)$ connecting that point with the starting point $(x_0,0)$.

We have chosen the trajectory to have $q(x(\tau),\tau) = q$. We take $x(\tau)$ just as a solution of the equation

$$\tau = T(x(\tau),q). \tag{25}$$

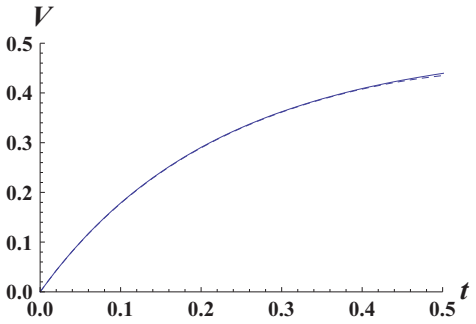

FIG. 1. (Color online) The graphics for the elasticity $V(t) \equiv \frac{-1}{v'(t)}$ for the model with $N = 100$, $R_+(x) = \exp(-x)$, $R_-(x) = \exp(x)$, $x(0) = 0.5$. The smooth line is the analytical result by Eq. (24) and the dashed line is the numerical result. The difference is less than 0.5%.

At any point of our trajectory $(x(\tau),\tau)$, we can calculate $v(x(\tau),\tau)$, while $q(x(\tau),\tau) = q$ is constant. Equation (11) gives

$$R_- e^{2v} - (q + R_+ + R_-)e^v + R_+ = 0,$$
$$v(y) = -\ln 2R_- + \ln(q + R_+(y) + R_-(y) \pm\sqrt{[q + R_+(y) + R_-(y)]^2 - 4R_+(y)R_-(y)}), \tag{26}$$

where we denoted $x(\tau) = y$.

We derive the solution of the original Eq. (2), integrating the equation $\dot{u} = v\dot{x} + q$ along our trajectory [the characteristics connecting the points (x,t) and $(x_0,0)$]:

$$u(x,t) = u(x,0) + \int_{x_0}^{x} dy\, v(y) + qt = \int_{x_0}^{x} dy[-\ln 2R_-(y) + \ln(q + R_+(y) + R_-(y) \pm\sqrt{[q + R_+(y)+R_-(y)]2 - 4R_+(y)R_-(y)}] + qt. \tag{27}$$

Having the expression $u(x,t)$ we can calculate $p(x,t)$.

D. The restricted meaning of probability distributions in master equations

In the case of evolution models [9,13], we have master equations similar to Eqs. (1)–(3); only the negative term $\sim P(X,t)$ has a coefficient other than $-(R_+ + R_-)$, and therefore there is no balance condition.

Contrary to the case of master equations in evolution models [9–13] where all the initial distributions have a meaning, now there are some restrictions. We should clarify the meaning of the probabilities $P(X,t)$. At every moment of time the system has only *one* value of X, and $P(X,t)$ just gives such probabilities. We should solve the system (1)–(3) for the given initial value $P(X_0,t) = 1$ and $P(X,t) = 0$ for other X; other initial distributions have no meaning.

Another difference is connected with the stable point solutions. Equation (13) gives the steady-state solution. If that

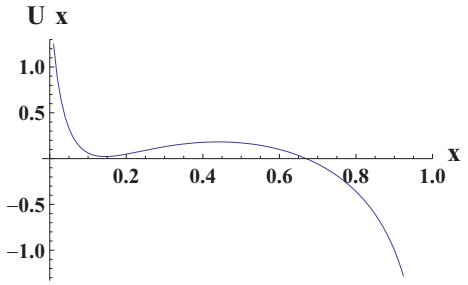

FIG. 2. (Color online) The graphics for the $U(x) = \ln[R_+(x)/R_-(x)]$ at $K_1 = 44$, $K_2 = 10$, $a = 0$. There are three solutions for the equation $U(x) = 0$.

equation has two stable solutions x_1, x_2 and the probability of these two positions is the same, i.e.,

$$\int_{x_1}^{x_2} dx\, \log \frac{R_+(x)}{R_-(x)} = 0, \tag{28}$$

we again should accurately interpret the HJE results [6]. For the rates given by Eq. (27) at $a = 0$, $k_2 = 10$, the transition is at $k_1 = 43.1274$. One can use the HJE method to calculate the mean period of time that solution, initially located at one stationary point, will move to the other stationary point [20]. Then it should again return back, as every moment the system could exist only with one value of X. In the case of the evolution model, the system goes to the equilibrium state instead of oscillatiing between two stable solutions.

E. The dynamics for the stationary but random rates

Consider now the case when the rates are smooth functions of x plus some random noises. The noise in the rates is well confirmed experimentally [21]. We took the case of rates from [6],

$$R_+(k_1,x) = (1 - x)(0.5 + k_1 x^2), \quad R_-(k_2,x) = x(k_2 + 0.01x^2), \tag{29}$$

where now k_1 and k_2 are random variables,

$$k_1 = K_1 \exp[a\xi_1(x) - a^2/2)x^2], \quad k_2 = K_2 \exp[a\xi_2(x) - a^2/2], \tag{30}$$

and $\xi_1(x),\xi_2(x)$ have normal distributions. We consider the model with $N = 100$, and performed numerics for different values of parameters K_1, K_2, a. For $K_1 = 50$, $K_2 = 10$, at $a \approx 0.9$ there is a phase transition: instead of two steady-state solutions we get one steady state $x \approx 0.088$.

We can analytically estimate the transition point (from one stabile point to bistability), considering the behavior of

$$U(x) = \left\langle \ln \frac{R_+(k_1,x)}{R_-(k_2,x}\right\rangle\Big|_{k_1,k_2}. \tag{31}$$

We found that the function $U(x)$ changes its behavior with the level of the noise. At $a = 0$ it has only three roots $U(x) = 0$, while for $a > 0.75$ there is one root.

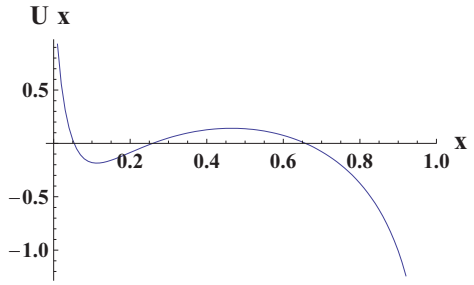

FIG. 3. (Color online) The graphics for $U(x) = \langle \ln[R_+(x)/R_-(x)] \rangle$ at $K_1 = 44$, $K_2 = 10$, and $a = 0.8$. There is only one solution for $U(x) = 0$.

III. MASTER EQUATION WHEN THE RATES VARY WITH TIME

A. The simplest solvable case

Consider the case when birth and death rate coefficients in CMEs change with the time as $g(t)$ and $f(t)$, while they are the same for all x. We have the following Hamiltonian:

$$-H = g(t)e^{-v} + f(t)e^{v} - g(t) - f(t), \quad v = \frac{\partial u}{\partial x}. \tag{32}$$

We get the following system of equations:

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} = 0. \tag{33}$$

Thus the p is constant:

$$v = v_0. \tag{34}$$

For the coordinate we have a simple equation:

$$\frac{dx}{dt} = \frac{\partial H}{\partial v} = g(t)e^{-v_0} - f(t)e^{v_0}. \tag{35}$$

Thus we have a simple solution

$$x - x_0 = e^{-v_0}\int_0^t g(\tau) - e^{v_0}\int_0^t f(\tau). \tag{36}$$

We find v_0 as a function of x,t from the solution of the last equation.

$$u(x,t) = u(x_0,0) + v_0(x - x_0) + \int_0^t [g(\tau)e^{-v_0} + f(\tau)e^{v_0} - g(\tau) - f(\tau)]d\tau. \tag{37}$$

For the maximum of distribution we have $v_0 = 0$, therefore we get

$$x - x_0 = \int_0^t [g(\tau) - f(\tau)]d\tau. \tag{38}$$

It is interesting to find the variance of distribution. Differentiating Eq. (36), we get, putting $v_0 = 0$ at the maximum point,

$$-v'(x,t) = \frac{1}{\int_0^t [g(\tau) + f(\tau)]}d\tau. \tag{39}$$

Thus the variance of distribution decreases with the time.

B. Rates as linear functions of the number of molecules

Consider now the Hamiltonian

$$-H = [ax + g(t)]e^{-v} + [bx + f(t)]e^{v} - (a + b)x - f(t) - g(t). \tag{40}$$

We have

$$\frac{dv}{dt} = ae^{-v} + be^{v} - (a + b). \tag{41}$$

We can solve this case:

$$\int_{v_0}^{v} \frac{dv}{ae^{-v} + be^{v} - (a + b)} = t. \tag{42}$$

We can define the function $v = V(v_0,t)$ from the latter equation,

$$V(v_0,t) = \log\left[\frac{ae^{at} - ae^{bt} - be^{at+v_0} + ae^{bt+v_0}}{ae^{at} - be^{bt} - be^{at+v_0} + be^{bt+v_0}}\right]. \tag{43}$$

We also have

$$\frac{dV}{dv_0}(t) = \frac{-be^{at+v_0} + ae^{bt+v_0}}{ae^{at} - ae^{bt} - be^{at+v_0} + ae^{bt+v_0}} - \frac{-be^{at+v_0} + be^{bt+v_0}}{ae^{at} - be^{bt} - be^{at+v_0} + be^{bt+v_0}}. \tag{44}$$

Now we solve the equation for x:

$$\frac{dx}{dt} = (b - a)x + f(t)e^{V} - e^{-V}g(t). \tag{45}$$

Its solution gives

$$x - x_0 = e^{(b-a)t}\int_0^t [f(\tau)e^{V(v_0,\tau)} - g(\tau)e^{-V(v_0,\tau)}]e^{-(b-a)\tau}\,d\tau. \tag{46}$$

Equations (42) and (46) together define the trajectory of characteristics.

To get the solution for the maximum, we put $v = 0$ on the left-hand side of Eq. (42). In this way we get a simple explicit equation for v_0 as a function of time t:

$$\frac{ae^{at} - ae^{bt} - be^{at+v_0} + ae^{bt+v_0}}{ae^{at} - be^{bt} - be^{at+v_0} + be^{bt+v_0}} = 1. \tag{47}$$

Putting that solution into Eq. (46), we find the trajectory of the maximum of distribution.

To calculate v' we should differentiate the expression in Eq. (46) via x. Using the relation $V'(v_0,t) = 1$, we derive

$$\frac{1}{v'} = e^{(b-a)t}\int_0^t [f(\tau)e^{V(v_0,\tau)} + g(\tau)e^{-V(v_0,\tau)}]\frac{dV}{dv_0}(\tau)e^{-(b-a)\tau}d\tau, \tag{48}$$

where v_0 is defined by Eq. (46) as a function of t.

IV. CME IN MULTIDIMENSIONAL SPACE

Consider now the CME in multidimensional space, when we have d-dimensional $\vec{X}$.

$$\frac{dP(\vec{X},t)}{N\,dt} = \sum_{n_l=-K}^{K} R_{\vec{n}}\left(\frac{\vec{X}-\vec{n}}{N}\right)P(\vec{X}-\vec{n},t) - \sum_{\vec{n}} R_{\vec{n}}\left(\frac{\vec{X}}{N}\right)P(\vec{X},t). \tag{49}$$

In Eqs. (49), (50), (56), (57) we dropped in the sums the terms with $n=0$. We get a HJE for the function $u(\vec{x})$ with the following Hamiltonian:

$$\frac{\partial H}{\partial t} + H(\vec{x},\vec{p}) = 0, \qquad -H(\vec{x},\vec{p}) = \sum_{n_l=-K,1\leqslant l\leqslant d}^{K} R_{\vec{n}}(\vec{x})[\exp[-\vec{n}\,\vec{p}]-1]. \tag{50}$$

Let us investigate the motion of the maximum of distribution. We denote $\vec{y}(t)=\vec{x}$ the maximum of distribution at moment of time t, and assume the following ansatz for $u(x,t)$ near the maximum of distribution:

$$u(\vec{x},t) = -1/2\langle\vec{x}-\vec{y}(t)|V|\vec{x}-\vec{y}(t)\rangle \tag{51}$$

putting this ansatz into

$$u''_{x_l t} + H'_{x_l}(\vec{x},\vec{p}) + \sum_m H'_{\vec{p}_m}(\vec{x},\vec{p})\frac{dp_m}{dx_l} = 0. \tag{52}$$

Consider the point $\vec{x}=\vec{y}(t)$, $\vec{p}=0$, and use the identity

$$\frac{dp_m}{dx_l} = -V_{ml}. \tag{53}$$

We get

$$\sum_m V_{lm}\frac{dy_m}{dt} - \left(\sum_{n_h=-K}^{K} R_{n_1,\ldots,n_d}\sum_m n_m\right)V_{lm} = 0. \tag{54}$$

We get the following ODE for the dynamics of the maximum of distribution:

$$\frac{dy_l(t)}{dt} = \sum_m Q_m, \quad Q_m = \sum_{1\leqslant h\leqslant d}\ \sum_{-K\leqslant n_h\leqslant K} R_{n_1,\ldots,n_d} n_m. \tag{55}$$

V. DERIVATION OF ELASTICITY IN 1D CASE

Let us get a system of ODEs for V. We consider the multistep version of CME in one dimension.

$$\frac{dP(X,t)}{N\,dt} = \sum_{n=-K}^{K} R_n\left(\frac{X-n}{N}\right)P(X-n,t) - \sum_n R_n\left(\frac{X}{N}\right)P(X,t). \tag{56}$$

Now we have the following Hamilton-Jacobi equation:

$$-H(u',x) = \sum_{n=-K}^{K} R_n(x)[e^{-nu'}-1]. \tag{57}$$

Now we have to consider the higher terms in the expansion of $u(x,t)$ near the maximum:

$$u = -V[x-y(t)]^2/2 - T[x-y(t)]^3/6 \tag{58}$$

Equation (56) gives

$$\frac{dy(t)}{dt} = b, \quad b = \sum_{-K\leqslant n\leqslant K} R_n(x)n. \tag{59}$$

We also denote

$$a(x) = \sum_{-K\leqslant n\leqslant K} R_n(x)n^2, \tag{60}$$

With the ansatz by Eq. (58) we have

$$u'''_{xxt} = -\dot{V} + Tb, \tag{61}$$

On the other hand, differentiating the right-hand side of HJE, we have

$$-u'''_{xxt} = H''_{pp}V^2 - 2H''_{xp}V - TH'_p = aV^2 - 2b'V - bT, \tag{62}$$

We derived Eqs. (61) and (62) putting $x=y(t)$.

Then

$$2b\frac{d}{dt}V = -4bb'V + 2baV^2 \tag{63}$$

or using $dy/dt = b$

$$b\frac{d}{dx}\frac{1}{V} = 2b'\frac{1}{V} - a, \tag{64}$$

which gives Eq. (24) for the one-step CME case.

For the multistep CME we have the following expression for the elasticity:

$$\frac{-1}{v'_x} = [b(x)]^2\int_{x_0}^{x}\frac{dy\,a(y)}{[b(y)]^3}. \tag{65}$$

VI. CONCLUSION

In conclusion, we calculated exact probability dynamics Eq. (27) for the chemical master distribution using the Hamilton-Jacobi equation method. Using the method of characteristics for solution of the HJE, we gave explicit expressions for the variance of distribution equation (24). The latter is important both for chemical [2] and financial applications. We also derived the variance directly from the HJE, using a special ansatz. The latter is equivalent to the Van Kampen method [4]. Using the Van Kampen method, in the Appendix we give an exact expression for the variance for a general 1D model, as well as a system of ODEs to define the variance in the 2D case. Both the HJE and the Van Kampen methods gave identical results for the variance and the maximum point dynamics of the CME: The HJE gives

a directly differential equation for the variance, while the Van Kampen method originally gives a differential equation for the probability distributions, Eq. (A5), which later proved the differential equation for the variance. The HJE also gives the exact steady-state distribution, and is more adequate for investigation of metastabile points [16,17]. Using the HJE, we derived exact dynamics equations (44)–(46) for the case of 1D CMEs when the rates are linear functions of the coordinate (number of molecules) plus some time-dependent functions.

We performed some numerics in the case of static noise, and also give analytical criteria for the level of noise when the bistability disappears. Our choice of potential for the averaging equation (31) is rather arbitrary (contrary to all other results in this paper, which are derived rigorously and are exact). A further investigation of the problem is necessary. Perhaps it is possible to investigate the more realistic case of nonstationary noise [21].

ACKNOWLEDGMENTS

D.B.S. thanks H. Qian and T. Lux for useful discussions, DARPA Prophency programm and Academia Sinica for financial support.

APPENDIX: THE CALCULATION OF VARIANCE DYNAMICS USING THE VAN KAMPEN METHOD

A. 1D multistep models

Consider the CME defined through the system of equations:

$$\frac{dP(X,t)}{N\,dt} = \sum_{l=-K}^{K} R_l\left(\frac{X+l}{N}\right)P(X+l,t) - \sum_{l} R_l\left(\frac{X}{N}\right)P(X,t). \tag{A1}$$

In Eq. (A1) we drop in the sum the term $l=0$. Following Van Kampen, we assume that near the maximum

$$X = N\phi(t) + \sqrt{N}\xi(t), \qquad P(X,t) = \Pi(\xi,t). \tag{A2}$$

For the maximum point $\phi(t) = \sum_i i\,P(i,t)$ we will derive Eq. (A4) later.

We consider terms with different scaling by N in the equation

$$\frac{d\Pi(\xi,t)}{dt} - \sqrt{N}\phi'(t)\frac{\partial\Pi}{\partial\xi} = N\sum_{l=-K}^{K} R_l\left[\phi(t)+\frac{\xi}{\sqrt{N}}-\frac{1}{N}\right]P(X-l,t) - N\sum_{l} R_l\left[\phi(t)+\frac{\xi}{\sqrt{N}}\right]P(X,t). \tag{A3}$$

Collecting together the $\sim\sqrt{N}$ terms, we get

$$\frac{d\phi(t)}{dt} = \sum_{l=-K}^{K} l R_l[\phi(t)]. \tag{A4}$$

The N^0 terms give an equation

$$\frac{d\Pi(t)}{dt} = -b'\frac{\partial}{\partial\xi}(\xi)\Pi + \frac{a}{2}\frac{\partial^2}{\partial\xi^2}\Pi, \qquad b = \sum_l l R_l, \quad a = \sum_l l^2 R_l. \tag{A5}$$

We derive the following equations for $\langle\xi_i\rangle$:

$$\frac{d\langle\xi\rangle}{dt} = b'(t)\langle\xi_i\rangle. \tag{A6}$$

The initial condition $\langle\xi(0)\rangle = 0$ gives $\langle\xi(t)\rangle = 0$.

For the variance we get

$$\frac{d}{dt}\langle\xi(t)^2\rangle = a[\psi(t)] + 2b'[\psi(t)]\langle\xi_i(t)\rangle^2. \tag{A7}$$

We can consider ODEs via ψ instead of t. Then Eq. (A4) gives $d\psi/dt = b$ and

$$b\frac{d}{d\psi}\langle\xi^2\rangle = a(\psi) + 2b'(\psi)\langle\xi\rangle^2. \tag{A8}$$

Eventually we get Eq. (64), with $R_+ - R_- \to b \equiv \sum_l R_l l$, $R_+ + R_- \to a \equiv \sum_l l^2 R_l$.

B. 2D case

We will apply the Van Kampen method, giving ODEs to calculate the variance of distribution in the 2D case.

Consider the following system of equations for $P(X,Y,t)$, $0 \leqslant X \leqslant N$, $0 \leqslant Y \leqslant N$:

$$\begin{aligned}\frac{dP(X,Y,t)}{dt} = &-\left[R_{1+}\left(\frac{X}{N},\frac{Y}{N}\right)+R_{1-}\left(\frac{X}{N}\right),\frac{X}{N}\right]+\left[R_{2+}\left(\frac{X}{N},\frac{Y}{N}\right)+R_{2-}\left(\frac{X}{N}\right),\frac{X}{N})\right]P(X,Y,t)\\ &+R_{1+}\left(\frac{X-1}{N},\frac{Y}{N}\right)P(X-1,Y,t)+R_{1-}\left(\frac{X+1}{N},\frac{Y-1}{N}\right)P(X+1,Y,t)+R_{2+}\left(\frac{X}{N},\frac{Y-1}{N}\right)P(X,Y-1,t)\\ &+R_{2-}\left(\frac{X}{N},\frac{Y+1}{N}\right)P(X,Y+1,t).\end{aligned} \tag{A9}$$

Following [4], we introduce the fluctuating variables ξ_1,ξ_2 and replace $P(X,Y,t)$ by $\Pi(\xi_1,\xi_2,t)$ (see [4]):

$$X = N\phi_1(t) + \sqrt{N}\xi_1(t), \quad Y = N\psi_2(t) + \sqrt{N}\xi_2(t), \quad P(X,Y,t) = \Pi(\xi_1,\xi_2,t), \tag{A10}$$

where $\psi_1(t),\psi_2(t)$ give the solution in the case of infinite N:

$$\frac{d\psi_1(t)}{dt} = b_1(\psi_1(t),\psi_2(t)), \quad \frac{d\psi_2(t)}{dt} = b_2(\psi_1(t),\psi_2(t)), \quad b_1(\psi_1,\psi_2) = R_{1+}(\psi_1,\psi_2) - R_{1-}(\psi_1,\psi_2), \quad b_2(\psi_1,\psi_2) = R_{2+}(\psi_1,\psi_2) - R_{2-}(\psi_1,\psi_2). \tag{A11}$$

We can solve Eq. (A11) and calculate $\psi_1(t), \psi_2(t)$.

Following the methods of [4], we derive

$$\begin{aligned}\frac{d\Pi(t)}{dt} &= -\sum_\alpha \left(R'_{\alpha+} - R'_{\alpha-}\right)\frac{\partial}{\partial \xi_\alpha}(\xi_\alpha \Pi(\xi_1,\xi_2,t)) \\ &\quad + \sum_\alpha \frac{R_{\alpha+} + R_{\alpha-}}{2}\frac{\partial^2}{\partial^2 \xi_\alpha}\Pi(\xi_1,\xi_2,t) \\ &\equiv -\sum_\alpha \frac{\partial b_\alpha}{\partial \xi_\alpha}\frac{\partial}{\partial \xi_\alpha}\Pi + \frac{1}{2}\sum_\alpha a_\alpha \frac{\partial^2}{\partial \xi_\alpha^2}\Pi, \qquad \text{(A12)}\end{aligned}$$

where we denoted $R'_{\alpha\pm} \equiv \frac{\partial R_{\alpha\pm}(\psi_1(t),\psi_2(t))}{\partial \xi_\alpha}$, $R_{\alpha\pm} \equiv R_{\alpha\pm}(\psi_1(t),\psi_2(t))$.

We derive the following equations for $\langle \xi_i \rangle$:

$$\frac{d\langle \xi_i \rangle}{dt} = b'_i(t)\langle \xi_i \rangle. \qquad \text{(A13)}$$

The initial condition $\langle \xi(0) \rangle = 0$ gives $\langle \xi(t) \rangle = 0$.

For the variance we get

$$\begin{aligned}\frac{d}{dt}\langle \xi_i(t)^2 \rangle &= a_i(\psi_1(t),\psi_2(t)) \\ &\quad + 2b'_i(\psi_1(t),\psi_2(t))\langle \xi_i(t) \rangle^2 \qquad \text{(A14)}\end{aligned}$$

We can calculate the variance numerically as a function of t, using the solution of Eq. (A11).

[1] R. Phillips, J. Kondev, and J. Theriot, *Physical Biology of the Cell* (Garland Science, New York, 2008).

[2] D. A. Beard and H. Qian, in *Chemical Biophysics: Quantitative Analysis of Cellular Systems*, Cambridge Texts in Biomedical Engineering (Cambridge University Press, New York, 2008).

[3] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1985).

[4] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier Science, North-Holland, 2007).

[5] T. Lux, J. Econ. Behav. Organ. **72**, 638 (2009).

[6] H. Ge and H. Qian, Phys. Rev. Lett. **103**, 148103 (2009).

[7] J. E. Ferrell and W. Xiong, Chaos **11**, 227 (2001).

[8] G. Hu, Physica A **136**, 607 (1986).

[9] H. Woodcock and P. G. Higgs, J. Theor. Biol. **179**, 61 (1996).

[10] E. Baake and H. Wagner, Genet. Res. **78**, 93 (2001).

[11] D. B. Saakian, J. Stat. Phys. **128**, 781 (2007).

[12] K. Sato and K. Kaneko, Phys. Rev. E **75**, 061909 (2007).

[13] D. B. Saakian, O. Rozanova, and A. Akmetzhanov, Phys. Rev. E **78**, 041908 (2008).

[14] J. Keizer, Proc. Natl. Acad. Sci. USA **75**, 3023 (1978).

[15] R. and S. J. Chapman, Eur. J. Appl. Math. **16**, 427 (2005).

[16] M. Assaf and B. Meerson, Phys. Rev. E **81**, 021116 (2010).

[17] C. Escudero and A. Kamenev, Phys. Rev. E **79**, 041149 (2009).

[18] L. C. Evans, *Partial Differential Equations* (Providence, RI, 1999).

[19] A. Melikyan, *Generalized Characteristics of First Order PDEs* (Birkhäuser, Boston, 1998).

[20] M. Vellela and H. Qian, J. R. Soc. Interface **6**, 925 (2009).

[21] X. S. Xie, in *Single Molecule Spectroscopy in Chemistry, Physics and Biology Nobel Symposium*, edited by Astrid Graslund *et al.*, Springer Series in Chemical Physics (Springer-Verlag, Berlin, 2010), pp. 435–448.