

Equation-free Model Reduction in Agent-based Computations: Coarse-grained Bifurcation and Variable-free Rare Event Analysis

Ping Liu¹, C.I. Siettos², C.W. Gear³, I.G. Kevrekidis^{3,4} *

¹ Department of Molecular, Cellular and Developmental Biology,
Yale University, West Haven, CT 06516, USA

² School of Applied Mathematics and Physical Sciences,
National Technical University of Athens, Athens, Greece

³ Department of Chemical and Biological Engineering,
Princeton University, Princeton, New Jersey 08544, USA

⁴ Program in Applied and Computational Mathematics (PACM),
and Department of Chemical and Biological Engineering,
Princeton University, Princeton, New Jersey 08544, USA

Abstract. We study the coarse-grained, reduced dynamics of an agent-based market model due to Omurtag and Sirovich [18]. We first describe the large agent number, deterministic limit of the system dynamics by performing numerical bifurcation calculations on a continuum approximation of their model. By exploring a broad parameter space, we observe several interesting phenomena including turning points leading to unstable stationary agent density distributions as well as a type of “termination point.” Close to these deterministic turning points we expect the stochastic underlying model to exhibit rare event transitions. We then proceed to discuss a coarse-grained approach to the quantitative study of these rare events. The basic assumption is that the dynamics of the system can be decomposed into fast (noise) and slow (single reaction coordinate) dynamics, so that the system can be described by an effective, coarse-grained Fokker-Planck(FP) equation. An explicit form of this effective FP equation is not available; in our computations we bypass the lack of a closed form equation by numerically estimating its components - the drift and diffusion coefficients - from ensembles of short bursts of microscopic simulations with judiciously chosen initial conditions. The reaction coordinate is first constructed based on our understanding of the continuum model close to the turning points, and it gives results reasonably close to those from brute-force direct simulations. When no guidelines for the selection of a good reaction coordinate are available, data-mining tools, in particular Diffusion Maps, can be used to determine a suitable reaction coordinate. In the third part of this work we demonstrate this “variable-free” approach by constructing a reaction coordinate simply based on the data from the simulation itself. This Diffusion Map based, empirical coordinate gives results consistent with the direct simulation.

Keywords and phrases: agent based modeling, coarse-graining, equation-free computation, rare events

Mathematics Subject Classification: 35Q91, 37G10, 37H20, 65P30

1. Introduction

In recent years agent-based modeling has become a powerful mathematical/computational modeling technique with a broad spectrum of applications, varying from ecology [1,2,10] to economics and financial systems [23,24,29], and from traffic and supply chain networks [11,19,28] to biological [3,27,30] and social [9,12,17] systems.

In agent-based modeling, agents are treated as unique and discrete entities. They interact with each other and make their own “decisions” through detailed interaction rules. Modelers expect to study the macroscopic, population level dynamical behavior (what is called *emergent behavior*) through extensive direct simulations in time. Such a bottom-up approach is flexible because fine details of the interactions can be easily specified at the agent level. On the other hand, the modeler is interested in macroscopic, population level information, and this information can be much more efficiently extracted from continuum, population level evolution equations (e.g. partial differential equations), for which a wealth of computational tools and algorithms exists. The derivation of such population-level equations in closed form is, however, often practically impossible; and even in the cases where accurate macroscopic equations can be derived, a simple change in a microscopic rule would invalidate them. To overcome this obstacle, over the last years Kevrekidis et al. [13,14] developed a systematic computer-assisted methodology aimed at extracting and understanding coarse-grained, emergent dynamics by bridging agent-based simulations with macroscopic, systems-level, continuum numerical analysis. This methodology is termed the “Equation-Free approach” as it bypasses the derivation of *macroscopic* evolution equations when these equations conceptually exist but are not available in closed form. This is achieved by using appropriately initialized short bursts of the agent-based simulator to estimate *on the fly* the coarse-grained numerical quantities required for continuum algorithms (e.g. residuals, actions of Jacobians, etc.).

In our previous work [25] we have demonstrated how the Equation-Free concept [13–15] can be applied to perform coarse bifurcation analysis and controller design using agent-based models. Our illustrative example is a financially motivated agent-based model proposed by Omurtag and Sirovich [18]. It simulates the behavior of a large population of investors that buy or sell on the basis of information from the external environment and the action of other agents. In that work, by using what we refer to as the “coarse time stepper” [22] we constructed the coarse bifurcation diagram of the agent density distribution with respect to a bifurcation parameter quantifying the strength of interactions between agents. We obtained a stable branch of coarse stationary solutions and also an unstable stationary branch which cannot be found by long-term temporal simulations. It was shown that the instability observed in the system dynamics is the result of a coarse saddle-node bifurcation. A wash-out controller was also designed based on the information (Jacobian and control matrix) extracted on demand from the coarse time stepper, and was used to stabilize the open-loop unstable macroscopic system behavior.

Close to the coarse-grained turning points, however, the noise inherent in the stochastic agent-based model will manifest itself in fluctuations that can eventually “overshoot” the stable branch and cause a form on intense (explosive) instability; this phenomenon is consistent with the occurrence of rare events in stochastic differential equations. In this article we extend the scope of our coarse-graining computational technology, and show how to obtain reduced models capable of quantifying these rare event transitions.

The paper is structured as follows: After a brief description of the financial market model (Omurtag and Sirovich, [18]) we perform a two-parameter bifurcation study of its dynamics (to be precise, of the dynamics of its continuum approximation) to set the stage for the occurrence of rare events in the neighborhood of continuum saddle-node (turning point) bifurcations. At selected parameter regimes we construct representative one-parameter bifurcation diagrams. We also demonstrate that coarse bifurcation diagrams of the agent-based model (constructed using the Equation-Free approach) match qualitatively the ones of the approximate continuum model. By exploring a wide parameter space, we observe several interesting phenomena: turning points of the unstable solution branch (beyond the ones presented in previous work),

*Corresponding author. E-mail: yannis@princeton.edu

a special triangle-like probability density profile, as well as so-called “termination points” (beyond which no steady state solution exists).

We subsequently demonstrate how the Equation-Free approach can be extended to perform rare event analysis using the agent-based model. Far from the continuum turning points the problem appears to be well-approximated by a deterministic model, and the coarse-grained dynamics are governed by a Fokker-Planck-type partial differential equation (PDE). However, at the region near the turning points, due to the inherent model stochasticity, the very slow one-dimensional dynamics, and also due to the closeness (in phase space) between the stable and unstable stationary states, the problem is more suitably modeled by a (one-dimensional) Langevin-type stochastic differential equation (SDE). In this context, the deterministic stable stationary states are actually metastable ones: it is seen that after some (possibly long) time the state of the population may escape from the region of attraction of the “stable” stationary state. This escape is the rare event we study. The distribution of escape times for this rare-event is analyzed in Section 3.

The implementation of equation-free algorithms requires knowledge of the appropriate coarse-grained variables (the collective statistics of the agent based model, in terms of which the coarse-grained equations, whether deterministic or stochastic Markovian, would be formulated). In many cases such coarse variables are known from physical experience/understanding, or from an approximate continuum model; in our case, the Fokker-Planck approximation does suggest a good coarse variable for the one-dimensional effective Langevin SDE we will study in Section 3. In recent years there have been significant breakthroughs in using data-mining techniques that pass low-dimensional nonlinear manifolds through high-dimensional simulation data and simultaneously provide information about the dimensionality of such manifolds. One can think of techniques such as Isomap [26], or Diffusion Maps [4, 5, 16] that we use here, as nonlinear extensions of Principal Component Analysis (Proper Orthogonal Decomposition in the model reduction literature [20]). In the third part of this article we demonstrate how the application of Diffusion Maps to simulation data detects appropriate coarse-grained variables in terms of which our macroscopic effective model can be written. In section 4, we conduct the rare event analysis with these coarse variables and observe that the diffusion map approach gives comparable results to the ones using knowledge-based coarse variables.

2. The model and its bifurcation analysis

2.1. The agent-based model

In our previous published work [25] we studied an agent-based financial market model initially described in Omurtag and Sirovich [18]; this model simulates the actions of buying and selling by a large population of interacting individuals in the presence of mimesis. In this model, each agent’s propensity to buy or sell is indicated by its scalar preference state. The preference state $x_{i=1,2\dots N}(t)$ of each individual $i = 1, 2\dots N$ evolves with time in the one-dimensional “artificial world” $x_i(t) \in (-1, 1)$ according to two coupled processes. The first one is the exponential decay of $x_i(t)$ towards the neutral state 0 at the constant decay rate γ . This decay implements the assumption that each agent tends to gradually forget its current preference state and tends to be neutral in its preference for buying or selling in the absence of information.

The second one is a stochastic process denoted by $I_i(t)$ which represents the effect of the incoming information to each preference state $x_i(t)$. $I_i(t)$ basically tells us three things: (1) when the incoming information arrives; (2) what kind of information (i.e. “good” or “bad”) it is; (3) how this information affects the state $x_i(t)$. The arrival of incoming information is assumed to follow a Poisson distribution with average arrival frequency $(\nu^+ + \nu^-)$, where ν^+ denotes the average arrival frequency for “good news” and ν^- denotes the one for “bad news”. The average arrival frequencies are defined as follows:

$$\nu^\pm = \nu_{ex}^\pm + gR^\pm \quad (2.1)$$

where the parameter ν_{ex}^+ (resp. ν_{ex}^-) represents the contribution of the *external* information each individual draws from its environment (e.g. mass media news or opinions of stock market consultants). R^+ and R^- are the buy and sell rate respectively, defined as the number of buys or sells that occur in the market per unit time, normalized by the total number of agents in the market (we assume there is a very large external pool, so that R^+ and R^- do not need to be equal). The parameter g is a feedback constant which tells us to what extent the average arrival frequency is determined by the buy (or sell) rate. Since the buy or sell rate is a collective property of the population (i.e. all the agents in the market), the second term of Eq.(2.1) gR^\pm tells us how each individual agent's preference state is affected by the collective behavior of the entire population.

Each ‘‘quantum’’ of information that arrives has probability $\nu^+ / (\nu^+ + \nu^-)$ to be good news and probability $\nu^- / (\nu^+ + \nu^-)$ to be bad news. When good news arrives, the value of $x_i(t)$ is increased instantaneously by a fixed amount ϵ^+ ($\epsilon^+ > 0$); similarly, when bad news arrives, $x_i(t)$ is decreased instantaneously by a fixed amount $-\epsilon^- > 0$ ($\epsilon^- < 0$). If after a positive jump the value of $x_i(t)$ exceeds the positive boundary (i.e., $x_i(t) > 1$), then a ‘‘buy’’ is considered to occur, and the number of buys for that time interval is increased by one. Similarly after $x_i(t)$ crosses the negative boundary (i.e., $x_i(t) < -1$), the number of ‘‘sells’’ is increased by one. In either case $x_i(t)$ is set back to the neutral state (i.e., $x_i(t) = 0$) after the number of buys or sells is updated. In this way, each individual agent's decision can also affect the population's collective behavior.

This discrete jump process and the previously mentioned exponential decay process together form the evolutionary rule for each agent's preference state x_i , which can be summarized as the following (stochastic differential) equation

$$\frac{dx_i}{dt} = -\gamma x_i(t) + I_i(t). \quad (2.2)$$

2.2. Population-level dynamics and the Fokker-Planck approximation

It is possible to derive a concise description of the dynamics of a *large assembly* of agents by keeping track of the probability of finding an agent at some preference state, rather than tracking the preference states of *every single* agent in the population. This is accomplished by deriving an equation for $\rho(x, t)$, the probability density of agents at preference state x at time t . We can also imagine $\rho(x, t)$ as the density of agents averaged over a large number of replicas of the population. From continuity considerations, Omurtag and Sirovich [18] derived the following population-level equation which describes the evolution of the probability density $\rho(x, t)$.

$$\frac{\partial \rho}{\partial t} = -\gamma \frac{\partial(x\rho)}{\partial x} + \sum_{k=+,-} \nu^k (\rho(x - \epsilon^k, t) - \rho(x, t)) + (R^+ + R^-) \delta(x) \quad (2.3)$$

For small ϵ^k , $k = +, -$ by expanding $\rho(x - \epsilon^k, t)$ about $\rho(x, t)$ in terms of ϵ^k and truncating terms higher than second order in ϵ^k , they obtained the following Fokker-Planck approximation:

$$\frac{\partial \rho}{\partial t} = \gamma \frac{\partial((x - p)\rho)}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 \rho}{\partial x^2} + (R^+ + R^-) \delta(x) \quad (2.4)$$

where $p = (\nu^+ \epsilon^+ + \nu^- \epsilon^-) / \gamma$ and $\sigma^2(t) = \nu^+ (\epsilon^+)^2 + \nu^- (\epsilon^-)^2$. The fact that agents moving to/beyond $x = \pm 1$ are restored to the origin $x = 0$ dictates that $\rho(x = \pm 1, t) = 0$. The buy and sell rates are defined as fluxes across the boundaries $R^\pm = \mp \frac{1}{2} \sigma^2 \partial \rho / \partial x |_{x=\pm 1}$. By setting the time derivative to zero, we can arrive at an *analytical* solution for the stationary states of Equation (2.4) in the following form:

$$\rho_L(x) = f_1(x; \epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-) \quad (2.5)$$

$$\rho_R(x) = f_2(x; \epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-) \quad (2.6)$$

$$R^+ = h_1(\epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-) \quad (2.7)$$

$$R^- = h_2(\epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-) \quad (2.8)$$

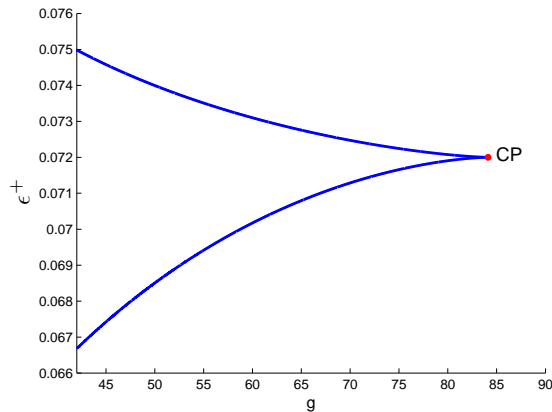


FIGURE 1. Two-parameter partial bifurcation diagram generated by MATCONT based on the analytical FP solution. The two parameters are the positive jump size ϵ^+ and the feedback factor g . The negative jump size ϵ^- is fixed at -0.072 . (CP: cusp point)

where $\rho_L(x)$ and $\rho_R(x)$ denote the probability density profiles at the left and right half of the space domain, respectively. The detailed expressions for f_1, f_2, h_1, h_2 are given in the Appendix. By inserting the above set of nonlinear equations into software like AUTO/MATCONT [6, 7], we can perform the bifurcation analysis for the Fokker-Planck approximation (Eq.(2.4)) of the model. Fig. 1 shows the two-parameter bifurcation diagram generated by MATCONT based on the analytical expressions (2.5) - (2.8). The two free parameters in Fig.1 are the positive jump size ϵ^+ and the feedback factor g . We fix the negative jump size ϵ^- at -0.072 . A cusp point appears when the positive jump size equals the negative one. Next, we select two representative regimes in the two-parameter bifurcation diagram to perform illustrative one-parameter cuts. In addition to ϵ^- , we also fix ϵ^+ leaving the feedback factor g as the only free parameter for this one-parameter bifurcation analysis. We study two representative cases – the *symmetric* and the *asymmetric* case respectively. In the asymmetric case, ϵ^+ and ϵ^- have different magnitude ($\epsilon^+ = 0.075$, $\epsilon^- = -0.072$), while in the symmetric case their magnitudes are equal (in particular, we set $\epsilon^+ = -\epsilon^- = 0.072$). Fig. 2(a) and Fig. 2(b) show the one-parameter bifurcation diagrams for the asymmetric and the symmetric case respectively. The diagrams are plotted in three dimensions because we are plotting two observed variables versus one free parameter. The first observed variable is the mean preference state \bar{x} , while the second is the quadratic mean of the buy and sell rates $[\left((R^+)^2 + (R^-)^2\right)/2]^{\frac{1}{2}}$. Fig. 3(a) justifies the stabilities of various branches through plots of the dominant eigenvalues of the Jacobian matrix associated with the (sufficiently finely) discretized Partial Differential Equation (PDE). Fig. 3(b) shows the probability density profiles of the preference states at representative parameter values. Fig. 4 demonstrates that the discretized PDE solution converges to the analytical solution.

Fig. 4(a) and Fig. 4(b) indicate that in the region of relatively small g values, there exists a pitchfork bifurcation for the symmetric case; and, as a result of the perturbation of the pitchfork, a saddle-node type bifurcation for the asymmetric case. The pitchfork bifurcation (branch point, BP) occurs when g is around 84, while the turning point of the saddle-node bifurcation occurs when g is around 42. For the symmetric (resp. asymmetric) case, by extending the middle (resp. lower) unstable branch to much greater g values, we observe an additional turning point which occurs when g is around 276 (resp. 248). The turning point can be seen more clearly in the side-view. As Fig. 3(a) shows, after passing this latter turning point the number of eigenvalues greater than zero increases from one to two, meaning the solution acquires even more unstable directions. The bifurcation diagrams also indicate that, for parameter values of g greater than the value of this latter turning point, there exist no steady state

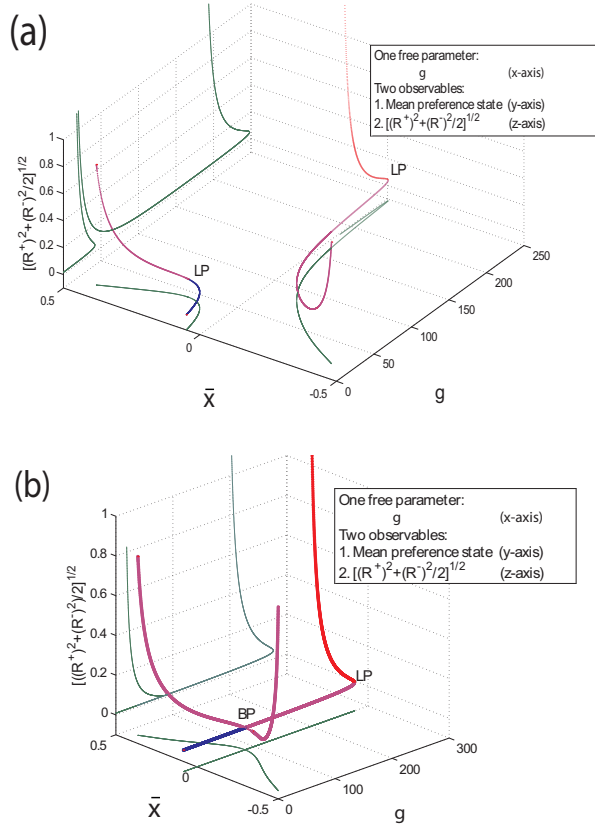


FIGURE 2. Three dimensional views of one-parameter bifurcation diagrams generated from the analytical solutions for (a) the asymmetric case ($\epsilon^+ = 0.075$, $\epsilon^- = -0.072$) and (b) the symmetric case ($\epsilon^+ = -\epsilon^- = 0.072$). The bifurcation parameter is the feedback intensity g . The two observable quantities plotted are the mean preference state \bar{x} and the quadratic mean of the buy and sell rate $[(R^+)^2 + (R^-)^2 / 2]^{1/2}$ respectively. Blue color indicates stability; purple indicates unstable branches with a single positive eigenvalue; red indicates the unstable branch with two positive eigenvalues; the green curves are the various projections of the 3D plot to the corresponding 2D planes. (LP: limit point, BP: branch point)

solutions, nor even unstable ones. More interestingly (see Fig.3), after this latter turning point, the probability density profile of the agents' preference states changes from a more normal-like distribution to a more triangular-shaped one. This is because in the region close to, and beyond, the additional turning point, the arrival frequencies ν^+ and ν^- both have a very large magnitude which leads to a large σ^2 (see Eq. (2.4)) value and, simultaneously, a small p value. These two observations together give rise to the triangular probability density shape.

By extending the unstable branches towards lower values of g we observed so called “termination points” – parameter values at which the respective unstable branches terminate. We found that these termination point for both the symmetric and the asymmetric cases is located between $g = 13$ and $g = 14$. For parameter values less than the critical termination value, no unstable steady state solution exists any more.

Fig. 5 shows the coarse one-parameter bifurcation diagrams for the agent-based code computed based on the coarse timestepper $x(t + \Delta T) = \phi_{\Delta T}(x(t))$ constructed in our previous work [25]. At the macro-

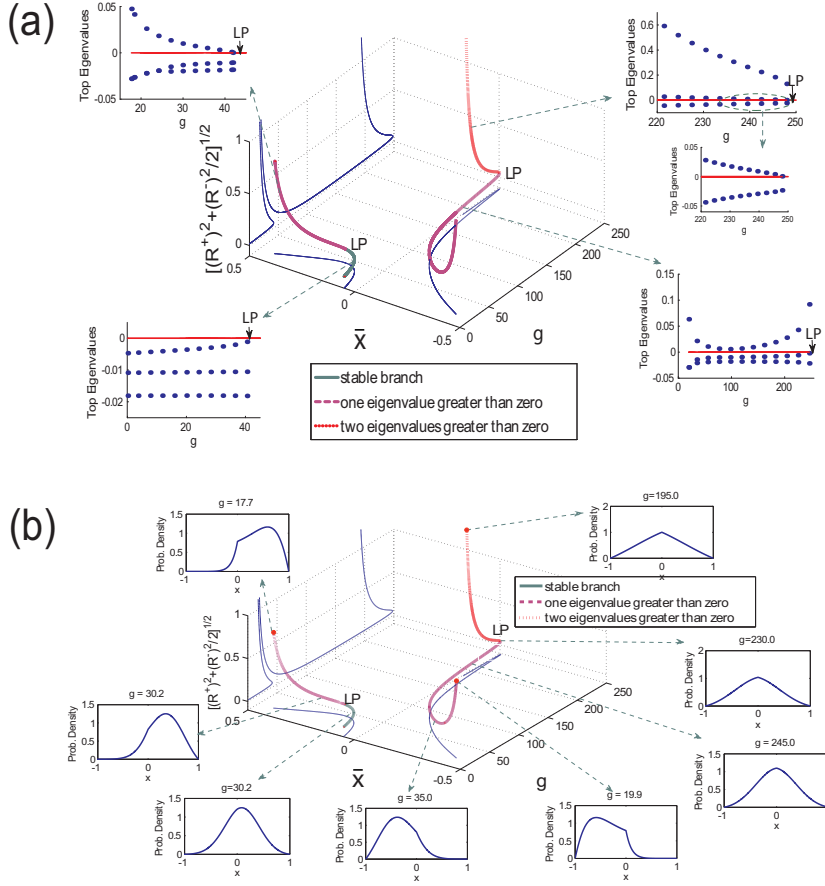


FIGURE 3. Bifurcation results for the asymmetric case ($\epsilon^+ = 0.075, \epsilon^- = -0.072$): (a) Stability of various fixed point branches. For each branch the leading eigenvalues of the discretized PDE system at various fixed points are plotted. (b) Probability density profiles of the preference states at different representative parameter values. (LP: limit point)

scopic level, we approximate the distribution of the agents' states by 37 distinct percentile points of the cumulative distribution function. The precise number of percentile points we use does not matter so much (as long as they sufficiently resolve the cumulative distribution), as we rely on the microscopic simulator to "heal" the discrepancies caused by the approximation. The coarse variables are these 37 percentile points plus the buy and sell rates R^+ and R^- , i.e. 39 coarse variables in total. The stationary states on the coarse bifurcation diagram have been obtained as fixed points of the agent-based timestepper averaged over 1000 realizations. The shape and stability type of various branches are seen to agree qualitatively well with the bifurcation diagrams shown before for the PDE approximation of the continuous model. The leftmost inset in Fig. 5b shows that the stationary solution, calculated using our equation-free methodology (through the coarse timestepper) also agrees well with the results from the long term agent-based simulation.

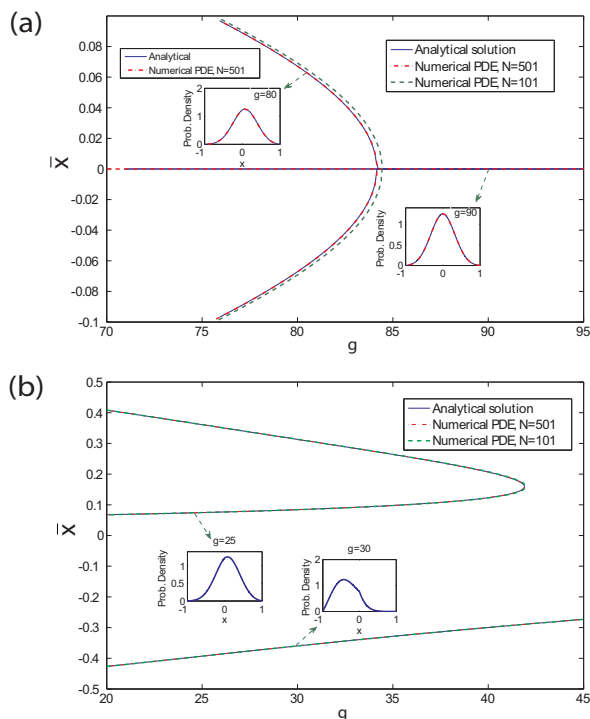


FIGURE 4. Comparison between the analytical solution and the numerically discretized PDE solutions for (a) the symmetric case, and (b) the asymmetric case. The numerical PDE solution, obtained by discretizing in different numbers of bins ($N=501$ in red dash dot and $N=101$ in green dash), approaches the analytical solution (in blue solid line). Detailed probability density profiles are also compared at specific locations of the various different bifurcation branches to show this convergence.

3. Rare event analysis

For the asymmetric case of the agent-based problem, at regions close to the coarse saddle-node bifurcation point the stable stationary state lies quite close (in phase space) to the “approaching” unstable, saddle-type stationary state. Due to this fact and also due to the stochasticity of the process itself, a system starting within the region of attraction of the what is a stable steady state in the deterministic approximation, will eventually escape from this region. An interesting question naturally arises - how long (on average) does it take for this escape event to take place for a system at a particular value of the parameter g ? We can run the agent-based code repeatedly to obtain the statistics for the escape times and then calculate the average value, but that would be too costly. In this section we will apply the equation-free approach to perform an “economical” coarse-grained rare event analysis to estimate these escape times and their distribution. We observe that this approach saves a very large fraction of the “brute force” computational time required by direct simulations while giving quantitatively comparable results.

This section starts with a brief discussion of the relevant theoretical background; then a “knowledge-based” 1D coarse reaction coordinate is proposed, and the hypothesis of an effective one-dimensional Fokker-Planck equation along this reaction coordinate is tested; finally the rare event analysis results are presented.

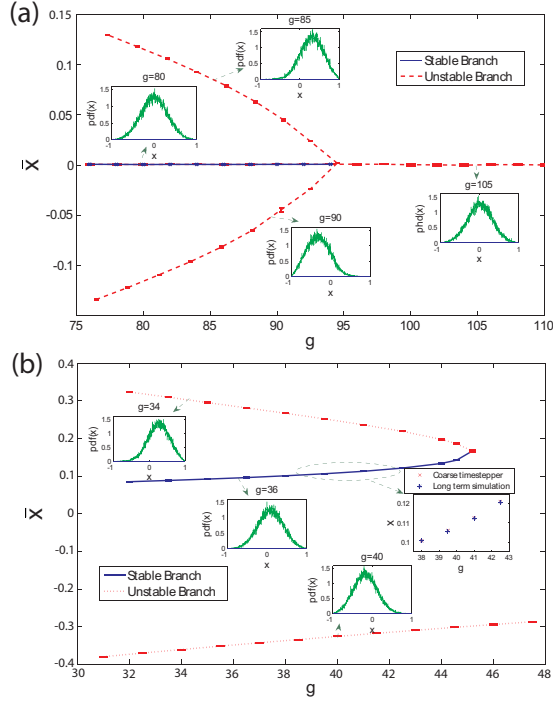


FIGURE 5. Coarse bifurcation diagram computed by wrapping the Newton-GMRES scheme around the coarse timestepper of the agent-based code with $N_{agents} = 50000$, $N_{copies} = 1000$ and $Time\ horizon = 0.25 \times 4$. Fig. 5a shows the symmetric case ($\epsilon^+ = -\epsilon^- = 0.072$) and Fig. 5b shows the asymmetric case ($\epsilon^+ = 0.075, \epsilon^- = -0.072$). The solid line corresponds to the branch with stable fixed points, while the dotted one corresponds to the branch with unstable fixed points. The standard error for each fixed point is estimated based on 20 separate runs. The stability of each branch is determined based on the algebraically largest eigenvalue calculated by wrapping a matrix-free Arnoldi algorithm around the coarse timestepper. The rightmost inset compares the coarse fixed points obtained using the equation free methodology with the ones from long term microscopic simulations.

3.1. Theoretical background

For a general one-dimensional stochastic process $\psi(t)$, the evolution of the probability density $P(\psi, t)$ of ψ obeys the following integral equation [21]:

$$P(\psi, t + \tau) = \int W(\psi, t + \tau | \psi', t) P(\psi', t) d\psi' \quad (3.1)$$

where $W(\psi, t + \tau | \psi', t)$ is the transition probability from point ψ' at time t to point ψ at time $(t + \tau)$. The differential form of this equation, known as the Kramers-Moyal expansion [21], is as follows:

$$\frac{\partial P(\psi, t)}{\partial t} = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial \psi}\right)^n D^{(n)}(\psi, t) P(\psi, t), \quad (3.2)$$

where,

$$D^{(n)}(\psi, t) = \frac{1}{n!} \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle [\xi(t + \tau) - \xi(t)]^n \rangle_{\xi(t) = \psi} \quad (3.3)$$

are the moments of the transition probability W . The angular brackets here denote ensemble averaging and ξ denotes a realization of the stochastic process. At the starting point t this has a δ -function distribution, $\xi(t) = \psi$.

If the process is Gaussian and Markovian, then all the higher order terms other than the first two in Eq. (3.3) will vanish, and the Kramers-Moyal expansion Eq.(3.2) reduces to the forward Fokker-Planck partial differential equation:

$$\frac{\partial P(\psi, t)}{\partial t} = \left[-\frac{\partial}{\partial \psi} v(\psi, t) + \frac{\partial^2}{\partial \psi^2} D(\psi, t) \right] P(\psi, t). \quad (3.4)$$

Furthermore, if the stochastic process is temporally homogeneous (i.e. if it is invariant with respect to shifts in the time origin), which is true for processes without external time-dependent forcing, then the forward Fokker-Planck equation can be written as

$$\frac{\partial P(\psi, t)}{\partial t} = \left[-\frac{\partial}{\partial \psi} v(\psi) + \frac{\partial^2}{\partial \psi^2} D(\psi) \right] P(\psi, t). \quad (3.5)$$

where, $v(\psi) \equiv D^{(1)}(\psi)$ is the *drift coefficient*, and $D(\psi) \equiv D^{(2)}(\psi)$ is the *diffusion coefficient*. The drift and diffusion coefficients completely determine the Markov process $\psi(t)$ in the stochastic sense [8]. The forward Fokker-Planck equation is useful in the sense that, with appropriate initial and boundary conditions, it uniquely determines the evolution of the Markov state density function $P(\psi, t | \psi_0, t_0)$ of the process. However, it does not describe the evolution of $\psi(t)$ itself, which we may also want to study. Fortunately the drift and diffusion coefficients also provide us a way to study the evolution of $\psi(t)$ itself through the Langevin-type stochastic differential equation:

$$\psi(t + dt) = \psi(t) + D^{1/2}(\psi(t))dW(dt) + A(\psi(t))dt, \quad (3.6)$$

where, $A(\psi(t))$ and $D(\psi(t))$ are the drift and diffusion coefficients, $dW(dt)$ is a normal random variable with mean 0 and variance dt . As long as we obtain the drift and diffusion coefficients, we can essentially reconstruct both the forward Fokker-Planck equation and the corresponding effective Langevin equation. After the reconstruction of these two very useful analytical tools, several global characteristics of the system will be available, such as the potential function $G(\psi)$ (for a scalar problem) as well as the mean escape time τ (for escapes between potential wells). The potential function $G(\psi)$ can be obtained from the equilibrium probability distribution which is a solution of the steady-state Fokker-Planck equation,

$$0 = -\frac{d}{d\psi} [A(\psi)P_s(\psi)] + \frac{d^2}{d\psi^2} [D(\psi)P_s(\psi)]. \quad (3.7)$$

Substituting the ansatz $P_{eq}(\psi) \propto \exp[-G(\psi)]$ into Eq.(3.7), we obtain the potential function

$$G(\psi) = -\int \frac{v(\psi')}{D(\psi')} d\psi' + \ln D(\psi) + const. \quad (3.8)$$

Another function which is similar to, but different from the potential function is defined as follows,

$$G_0(\psi) = -\int \frac{v(\psi')}{D(\psi')} d\psi' + const. \quad (3.9)$$

$G_0(\psi)$ is called the *pseudo-potential function* (to differentiate it from the actual potential function $G(\psi)$). $G_0(\psi)$ will be identical to $G(\psi)$ when the diffusion coefficient $D(\psi)$ is constant. Although it is not the actual potential function, $G_0(\psi)$ is useful in mean escape time calculations. In fact, the mean escape time from a potential well can be written as [8]:

$$\tau(\psi_0 \rightarrow \psi_e) = \int_{\psi_0}^{\psi_e} dy e^{G_0(y)} \int_{\infty}^y dz \frac{e^{-G_0(z)}}{D(z)}, \quad (3.10)$$

where ψ_0 is a starting point inside the well and ψ_e is a point beyond the potential well, where the probability density is nearly zero. When the y-integrand and z-integrand contribute significantly only for small regions around ψ_e and ψ_0 respectively, an analytical approximation of Eq.(3.10) is available as follows [8]:

$$\tau(\psi_0 \rightarrow \psi_b) \approx 2\pi \left| \frac{D(\psi_b)}{D(\psi_0)v'(\psi_b)(v'(\psi_0) - D''(\psi_0))} \right|^{\frac{1}{2}} e^{(G_0(\psi_b) - G_0(\psi_0))}. \quad (3.11)$$

where ψ_b is the peak of the barrier adjacent to the potential well. For a constant diffusion coefficient, Eq. (3.11) further simplifies to the well-known Kramers' escape time:

$$\tau(\psi_0 \rightarrow \psi_b) \approx 2\pi \frac{e^{[G_0(\psi_b) - G_0(\psi_0)]}}{\sqrt{G_0''(\psi_0)|G_0''(\psi_b)|}}. \quad (3.12)$$

3.2. A knowledge-based reaction coordinate for the agent-based model

In this section we will construct a knowledge-based one-dimensional reaction coordinate for the parameter regime where the rare event analysis will be performed. After the construction of this reaction coordinate, a Fokker-Planck equation will be estimated and the mean escape time will be calculated in the following sections.

In our rare event analysis there are three different levels of description, namely, the microscopic level, the intermediate (agent density PDE) coarse level and the effective Langevin long-time coarse level. The microscopic level description refers to the level at which the agent-based simulator operates. At this level, the microscopic variable $X(t)$ is defined as the input/output variable of the agent-based simulator. This microscopic variable here was of dimension 25002; 25000 of these refer to the preference states of the 25000 agents we specified for the system, and the remaining two refer to the current positive and negative information arrival frequencies (i.e. ν^+ and ν^- respectively). Next, the intermediate coarse level refers to the level the coarse, agent density PDE-based timestepper operates. The intermediate coarse variable(s) $x(t)$ are the input/output variables of the coarse timestepper. This level is 39-dimensional; 37 of these coarse variables refer to our discretization of the single agent density function: they are the 37 percentile points of the cumulative probability distribution of the agents' preference state; the remaining two are the current positive and negative information arrival frequencies. The intermediate coarse level is a population-level description of the agents.

The long-time coarse level is a one-dimensional description of the escape of the collective system state from metastable wells. The scalar variable associated with this level is denoted as $\psi(t)$ - a single *reaction coordinate*.

The rare events are studied in the parameter regime close to (just before) the deterministic saddle-node turning point. In this parameter region, the single unstable eigenvector of the Jacobian matrix of the coarse timestepper *at the saddle* is close to the dominant stable eigenvector (i.e. the eigenvector corresponding to the eigenvalue with smallest real part) of the Jacobian matrix of the coarse timestepper *at the stable node*. These two eigenvectors approach each other closer and closer as the parameter g comes closer to the turning point. Eventually as g reaches the turning point these two eigenvectors coincide with each other and form the *center eigendirection* of the turning point. Based on this knowledge, for parameter values slightly before the turning point, the collective behavior of the agent population can be approximately described as moving along the vector pointing from the saddle x_{saddle} to the node x_{node} . This assumption implies that all other variables quickly approach a slow, attracting, one-dimensional manifold *parametrized by the projection on this direction*; the statistics of the simulation quickly become functions of a single long-time coarse variable. Our candidate one-dimensional reaction coordinate ψ can thus be defined as follows:

$$\psi = (x_{node} - x)^T \frac{(x_{saddle} - x_{node})}{\|(x_{saddle} - x_{node})\|}. \quad (3.13)$$

3.3. Validity of the Fokker-Planck equation assumptions.

Based on the coarse reaction coordinate proposed in the last section, we will briefly validate two important assumptions behind the effective Fokker-Planck equation description: that the stochastic process is (approximately) Gaussian and Markovian.

The process $\psi(t)$ is demonstrated qualitatively to be Gaussian in the following way:

1. by initializing many (in this case, 20000 copies of) separate microscopic configurations corresponding to the same coarse coordinate value ψ_0 , we construct a “pseudo-delta function” at some specific coarse coordinate value ψ_0 ;
2. we then propagate this pseudo-delta function by running the microscopic simulator for each of those 20000 microscopic initial configurations for some time t , and throughout this process we record the microscopic configurations at different transient times t_i with $i = 1, 2, 3, \dots, n$;
3. at each t_i by restricting the microscopic configurations to their corresponding coarse coordinate values $\psi_j(t_i)$, $j = 1, 2, 3, \dots, 20000$, we construct the evolutionary probability density distributions of the coarse variable $\psi(t)$ at different times t_i .

As Fig. 6 shows, the propagated distributions of the coarse variable are well fitted (in the eye norm) by Gaussian distributions, which validates the assumption that the stochastic process $\psi_i(t)$ is (approximately) Gaussian.

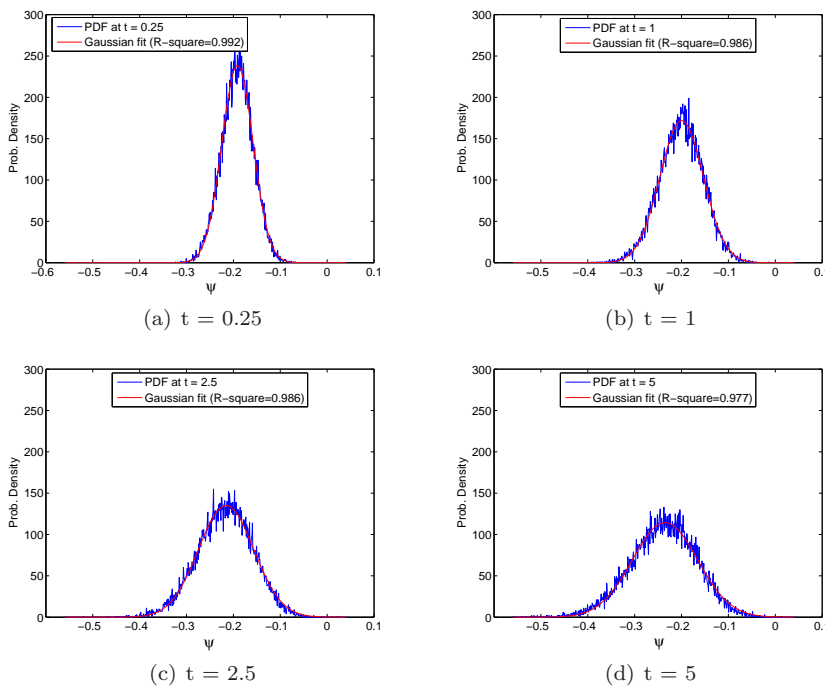


FIGURE 6. Probability density distributions of the coarse variable $\psi(t)$ after several path realizations initialized at $\psi_0 = -0.25$.

Showing that the process is (approximately) Markovian is tantamount to showing that the noise term (the second term) in the Langevin Eq. (3.6) has zero correlation time. Manipulation of the Langevin equation shows that the correlation time of the noise is proportional to the correlation time of $d\hat{\psi}(t)/dt$, where

$$\hat{\psi}(t) = \psi(t) - \langle \psi(t) \rangle \quad (3.14)$$

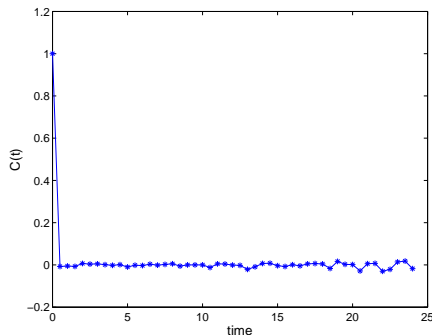


FIGURE 7. Autocorrelation function $C(t)$ of the noise term in the effective Langevin equation normalized so that $C(0) = 1$. The function shown is computed for an initial coarse variable value of $\psi_0 = -0.25$ and is typical of all coarse coordinate values in the interval between ψ_{node} and ψ_{saddle} , the region we are interested in for our rare event analysis.

is the fluctuation of $\psi(t)$. Therefore, to show that the noise term has zero correlation time we just need to show that $d\hat{\psi}(t)/dt$ has zero correlation time. The term $d\hat{\psi}(t)/dt$ is estimated using forward differences,

$$\left. \frac{d\hat{\psi}(t)}{dt} \right|_{t=t_i} \approx \frac{\hat{\psi}(t_{i+1}) - \hat{\psi}(t_i)}{\Delta t}, \quad (3.15)$$

where Δt is the step size of the coarse time stepper. A normalized autocorrelation function of $d\hat{\psi}(t)/dt$ (and therefore of the noise) for the initial coarse variable value $\psi_0 = -0.25$ is shown in Fig. 7. This clearly supports the hypothesis that the coarse stochastic process $\psi(t)$ is Markovian.

3.4. Mean escape time results – based on the empirical coordinate.

In this section, we will first demonstrate how the drift and diffusion profiles (see Eq. (3.13)) can be estimated as functions of the reaction coordinate through short time simulations of the agent-based code. Then, based on these approximate drift and diffusion coefficients, the mean escape time will be calculated.

The coarse reaction coordinate was discretized into 30 bins of fixed size 0.15. The grid points range from $\psi = -3.3$ to $\psi = -1.2$ in order to span the coordinate space between the saddle and node (and slightly beyond them). Since the drift and diffusion coefficients are actually the first two moments of Eq.(3.3), at each grid point ψ_i , they can be approximated from data obtained through multiple parallel runs via the following equations:

$$\nu(\psi_i) \approx \frac{1}{\Delta t} \langle \xi(t_0 + \Delta t) - \xi(t_0) \rangle |_{\xi(t_0)=\psi_i}, \quad (3.16)$$

$$D(\psi_i) \approx \frac{1}{2\Delta t} \langle [\xi(t_0 + \Delta t) - \xi(t_0)]^2 \rangle |_{\xi(t_0)=\psi_i}, \quad (3.17)$$

where the angular brackets denote ensemble averaging, and ξ denotes a realization of the stochastic process with a δ -function distribution at the starting point t_0 , $\xi(t_0) = \psi_i$.

In practice, the initial delta function is approximated in the following way: for each coarse grid point ψ_i , first we find its corresponding intermediate coarse variable x_i along the direction from the saddle x_{saddle} to the node x_{node} using Eq.(3.13); then we initialize 4000 copies of microscopic configurations $X_j, j = 1, 2, \dots, 4000$ consistent with the intermediate coarse variable x_i by randomly allocating agents' preference states between each two adjacent percentile points stored in x_i .

After the initialization, we run the agent-based simulator for each copy of the 4000 microscopic configurations for a short time Δt (one coarse time interval); we then restrict the evolved microscopic variables back to the corresponding *intermediate* coarse variables $x_j(t_0 + \Delta t)$, and finally further restrict each intermediate coarse variable to the corresponding *long-time* coarse coordinate value $\xi_j(t_0 + \Delta t)$. We can now use Eq.(3.16) and Eq.(3.17) to estimate the drift and diffusion coefficients at grid point ψ_i . The above procedure is repeated for each grid point. Next, the profile of the effective potential function $G(\psi)$ and the pseudo-potential function $G_0(\psi)$ can be approximated by numerical integration using Eq.(3.8) and Eq.(3.9). The results are shown in Fig.8.

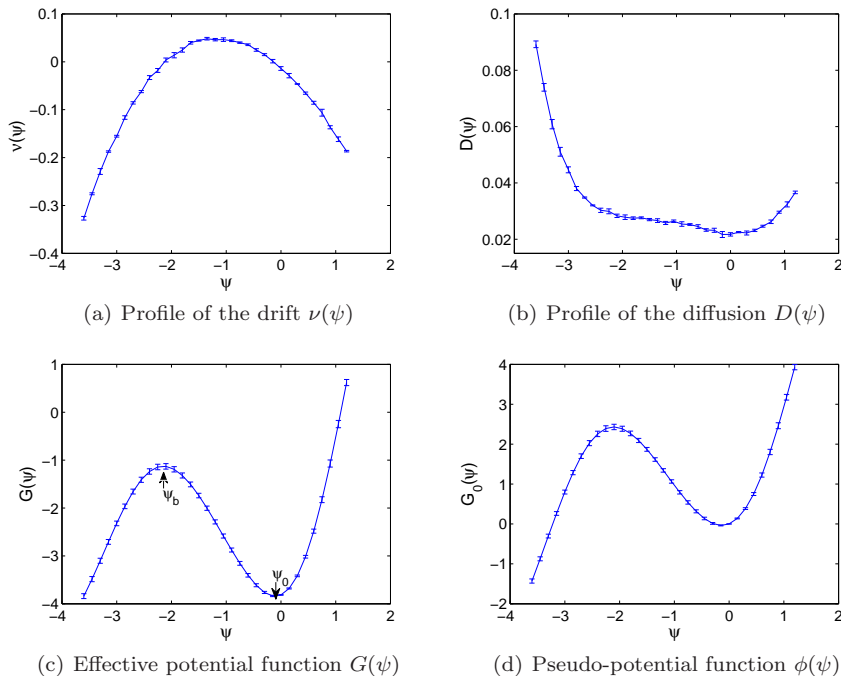


FIGURE 8. Profiles of estimated coarse drift and coarse diffusion coefficient as a function of the knowledge-based reaction coordinate (as well as of the effective potential and pseudo-potential) obtained through ensembles of short time agent-based simulations. The error bars are estimated based on five simulation runs.

Based on these estimated characteristic functions, we can apply the equations mentioned at the end of section 3 to estimate the mean escape time. Because Eq. (3.10) is computationally more expensive to apply, and requires data beyond the potential barrier, which are difficult to properly estimate, we decided to apply Eq.(3.11) for our coarse grained calculation. We specify the starting state ψ_0 for the rare event as state at the bottom of the effective potential well (as Fig.8(c) shows); the mean escape time obtained using Eq.(3.11) is $\tau = 788$, which compares reasonably well with the mean escape time $\tau = 599$ observed from the full-scale agent-based simulations (see Fig.9). Fig. 8(b) shows that the variation of the diffusion coefficient is relatively small between ψ_b and ψ_0 . If we treat the diffusion coefficient as constant with $D \approx (D(\psi_0) + D(\psi_b))/2$, and then apply Kramers' Eq.(3.12), we obtain a mean escape time of $\tau = 725$, in even better agreement with the brute force simulation result.

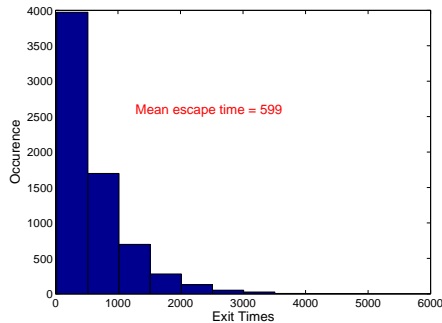


FIGURE 9. Histogram of escape times based on 6860 computational experiments with the full-scale agent-based simulator. The mean escape time is 599, with sample standard deviation 6.3 estimated based on the mean escape times of four samples, each consisting of 1715 computational experiments.

In terms of computational cost, the full-scale simulation requires 1.64×10^7 steps of agent-based simulation, while the coarse grained approach only takes 5.28×10^5 steps, as little as 3.2% of the amount required for the full-scale simulation.

4. Data-based (Diffusion Map) construction of the coarse reaction coordinate.

In this section, we will construct a candidate coarse reaction coordinate using a data mining tool – diffusion maps [5]. This technique is especially useful when physical intuition/prior knowledge of the system is not available. The diffusion map technique is based on the construction of a Markov matrix describing a random walk over a data set, where the probability for one data point to “hop” to another is described by a *pairwise similarity metric*. The *similarity kernel* is defined as follows:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\varepsilon}\right) \quad (4.1)$$

where $\varepsilon > 0$ is a measure of the local neighborhood of the points in the dataset over which the Euclidean distance is meaningful [5]. Next one defines $p_\varepsilon(x) = \sum_j K(x, x_j)$ and constructs the following generic kernel:

$$\tilde{K}(x_i, x_j) = \frac{K(x_i, x_j)}{p_\varepsilon^\alpha(x_i)p_\varepsilon^\alpha(x_j)} \quad (4.2)$$

where the parameter α depends on the application. In our case we use $\alpha = 0.5$ which corresponds to Fokker-Planck diffusion [5]. After normalization by the sum of each row $D_i = \sum_j \tilde{K}(x_i, x_j)$, the Markov matrix M is constructed with elements of the form

$$M_{i,j} = \frac{\tilde{K}_{i,j}}{D_i}. \quad (4.3)$$

Some of the first dominant eigenvectors of this Markov matrix will help parametrize the low-dimensional manifold underlying (intrinsic to the geometry of) the original data set. In our example, where we already know that the relevant space is one-dimensional, it is the component of a data point in the first nontrivial eigenvector of M that coarsely describes the data point itself. The Euclidean distance *in diffusion map space* corresponds to the diffusion distance [5] in the original space.

In our study of the agent-based problem, the first step is to generate data to build the Markov matrix Eq.(4.3). The data is sampled over the potential well -and slightly beyond- by running simulations

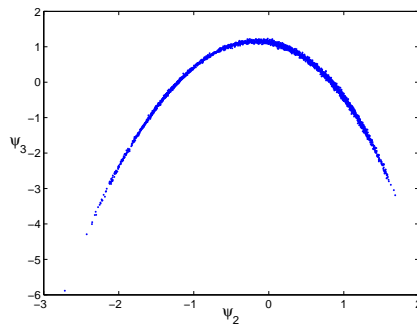


FIGURE 10. The first two nontrivial eigenvectors of the diffusion map matrix M ; the strong correlation is indicative of one-dimensional coarse data.

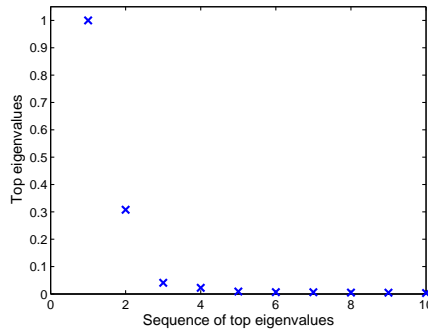


FIGURE 11. Top 10 dominant eigenvalues.

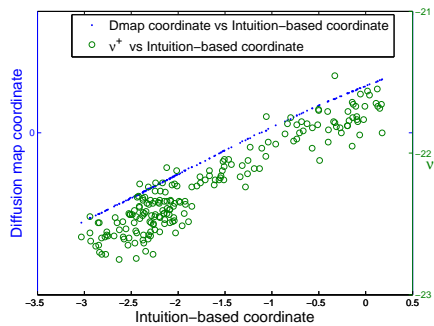


FIGURE 12. Comparison between the knowledge-based coarse coordinate and the diffusion map-generated one. The correlation with an arbitrarily chosen coarse observable, the positive information arrival frequency ν^+ , is also shown.

starting from the lifted microscopic configuration X_{saddle} up to one of the two “stop” points: one at the microscopic configuration corresponding to $x_{saddle} + \delta_1(x_{saddle} - x_{node})$; and the other at a microscopic configuration corresponding to $x_{node} - \delta_2(x_{saddle} - x_{node})$ (both δ_1 and δ_2 are small positive numbers). In our case, this sampling process takes 893 steps of the microscopic agent-based simulator, a relatively inexpensive computational effort.

After the construction of the Markov matrix outlined above, the first several dominant eigenvectors of that matrix are computed. The plot for the first and second dominant eigenvectors is shown in Fig.10. It reveals that the top two dominant eigenvectors actually both lie along the same one-dimensional space (they are very strongly correlated, indeed the second is a function of the first). Fig.11 shows the plot of the top 50 dominant eigenvalues. We observed a spectral gap between the first (nontrivial) dominant eigenvalue and the remaining ones (eigenvalue 1 is the trivial eigenvalue). The existence of a single dominant eigenvalue confirms the existence of an underlying one-dimensional slow manifold for the dynamics of our coarse grained system. In fact, the dominant slow time scale in this case corresponds to the time scale of the “rare event” while the remaining fast time scales correspond to the relaxation times within the well. Fig.12 compares three possible coarse reaction coordinates. It shows a good correlation between the diffusion map-generated coordinate and the knowledge-based coordinate we have discussed. It also shows a poor correlation between these two and a somewhat arbitrarily chosen coarse observable - the positive information arrival frequency ν^+ .

We applied this diffusion map-generated coordinate to the same data set (the short time simulation data) we used in the previous section to estimate the effective drift and diffusion coefficient profiles - $\nu(\psi_{dmap})$, $D(\psi_{dmap})$, $G(\psi_{dmap})$, and $G_0(\psi_{dmap})$. We then calculated the mean escape times based on these profiles. Eq.(3.11) then gives $\tau = 862$; using Kramers' Eq.(3.12) and assuming constant diffusion coefficient gives $\tau = 596$, both reasonably close to the results arrived at with the knowledge-based coordinate as well as the brute force simulation.

5. Summary

In this article we first presented a detailed two-parameter bifurcation study of the continuous version of a financial market model introduced by Omurtag and Sirovich. Several nontrivial dynamical phenomena were observed: turning point on the unstable branch, a “special” triangular probability density profile, and so-called “branch termination” points. We then discussed a coarse-grained approach to the computation of rare events for the detailed agent-based model of the system. The underlying one-dimensional effective Fokker-Planck equation was estimated from ensembles of short bursts of microscopic simulations with judiciously chosen initial conditions. The resulting mean escape time leading to a “bubble burst” (a rare event for the problem) compares reasonably well to the actual mean escape time from costly temporal agent-based simulations. A data-mining based approach - Diffusion Maps - was successfully applied to detect a good coarse coordinate, which then was used to obtain escape-time results that were consistent both with direct simulation and with knowledge-based coarse graining.

Appendix: Steady State Solution of the Continuum Model

We use subscripts ‘L’ and ‘R’ to denote the properties at the two sides of the spatial domain - $[-1, 0]$ and $[0, 1]$ - respectively.

The boundary conditions of the continuum model Eq.(2.4) are:

$$\rho_L(t, -1) = 0 \quad (5.1)$$

$$\rho_R(t, 1) = 0 \quad (5.2)$$

$$\rho_L(t, 0) = \rho_R(t, 0) \quad (5.3)$$

$$R^- = \frac{1}{2}\sigma^2 \frac{\partial \rho}{\partial x}(t, -1) \quad (5.4)$$

$$R^+ = -\frac{1}{2}\sigma^2 \frac{\partial \rho}{\partial x}(t, 1). \quad (5.5)$$

The integral condition reads,

$$\int_{-1}^1 \rho(x) dx = 1. \quad (5.6)$$

At steady state, the continuum model is written as follows,

$$0 = \gamma \frac{\partial((x-p)\rho)}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 \rho}{\partial x^2}. \quad (5.7)$$

Integrating Eq. (5.7) in space,

$$(x-p)\rho + \frac{1}{2} \left(\frac{\sigma^2}{\gamma} \right) \frac{\partial \rho}{\partial x} = c. \quad (5.8)$$

Let

$$\beta = \frac{\sigma^2}{\gamma}; \quad (5.9)$$

Eq.(5.8) then becomes

$$(x-p)\rho + \frac{1}{2}\beta \frac{\partial \rho}{\partial x} = c. \quad (5.10)$$

The solution of the homogeneous part of Eq.(5.10) (i.e. with $c = 0$) is

$$u(x) = \alpha e^{-(x-p)^2/\beta}. \quad (5.11)$$

Looking for a polynomial expression $w(x)$, Eq.(5.10) has the following general solution:

$$\rho(x) = u(x) + c \cdot w(x) \quad (5.12)$$

where $w(x)$ is given by

$$w(x) = (x-p) \left(1 - \frac{z}{3} + \frac{z^2}{3 \times 5} - \frac{z^3}{3 \times 5 \times 7} + \dots \right) \quad (5.13)$$

with $z = 2(x-p)^2/\beta$. Substituting Eq.(5.12) into Eq.(5.1) and (5.2),

$$c_L = -\frac{\alpha_L}{w(-1)} e^{-(-1-p)^2/\beta} \quad (5.14)$$

$$c_R = -\frac{\alpha_R}{w(1)} e^{-(1-p)^2/\beta}. \quad (5.15)$$

The probability density profiles at the two sides of the domain can be expressed as,

$$\rho_L = \alpha_L \left(e^{-(x-p)^2/\beta} + c_L w(x) \right) \quad (5.16)$$

$$\rho_R = \alpha_R \left(e^{-(x-p)^2/\beta} + c_R w(x) \right). \quad (5.17)$$

Substituting Eq.(5.16) and Eq.(5.17) into Eq.(5.3) we obtain

$$\alpha_L = F \cdot \alpha_R \quad (5.18)$$

with

$$F = \frac{e^{-\frac{p^2}{\beta}} - c_R w(0)}{e^{-\frac{p^2}{\beta}} - c_L w(0)}. \quad (5.19)$$

Substituting Eq.(5.18) into Eq.(5.16) and Eq.(5.17), and then applying the integral condition Eq.(5.6), we have:

$$\alpha_L = A + B + F(D + E) \quad (5.20)$$

$$\alpha_R = \frac{A + B}{F} + D + E. \quad (5.21)$$

By defining a function $G(x) = \frac{1}{2} \sum_{i=1}^{\infty} \frac{x^i}{i(2i-1)!!}$, A, B, D, E are expressed in terms of p and β as follows,

$$\begin{aligned} A &= -\sqrt{\frac{\beta\pi}{4}} \operatorname{erf}\left(p \cdot \sqrt{\frac{1}{\beta}}\right) \\ &\quad + \sqrt{\frac{\beta\pi}{4}} \operatorname{erf}\left((p+1) \cdot \sqrt{\frac{1}{\beta}}\right) \end{aligned} \quad (5.22)$$

$$B = \frac{c_R \beta}{2} \left(G\left(\frac{2(1-p)^2}{\beta}\right) - G\left(\frac{2p^2}{\beta}\right) \right) \quad (5.23)$$

$$\begin{aligned} D &= -\sqrt{\frac{\beta\pi}{4}} \cdot \operatorname{erf}\left((p-1) \sqrt{\frac{1}{\beta}}\right) \\ &\quad + \sqrt{\frac{\beta\pi}{4}} \cdot \operatorname{erf}\left(p \sqrt{\frac{1}{\beta}}\right) \end{aligned} \quad (5.24)$$

$$E = \frac{c_L \beta}{2} \left(G\left(\frac{2p^2}{\beta}\right) - G\left(\frac{2(-1-p)^2}{\beta}\right) \right). \quad (5.25)$$

So far, we have expressed $c_L, c_R, \alpha_L, \alpha_R$ in terms of p and β . Recall that

$$\nu^+ = \nu_{ex}^+ g R^+ \quad (5.26)$$

$$\nu^- = \nu_{ex}^- g R^- \quad (5.27)$$

$$\sigma^2 = \nu^+ (\epsilon^+)^2 + \nu^- (\epsilon^-)^2 \quad (5.28)$$

$$p = \frac{\nu^+ \epsilon^+ + \nu^- \epsilon^-}{\gamma}. \quad (5.29)$$

From Eq.(5.26-5.29) and Eq.(5.9), we can see that p and β are functions of the model parameters $\epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g$, and variables R^+, R^- . Therefore, $c_L, c_R, \alpha_L, \alpha_R$ are functions of these model parameters and variables as well. According to Eq.(5.16-5.17), the density profiles can be written in the following form,

$$\rho_L(x) = f_1(x; \epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-) \quad (5.30)$$

$$\rho_R(x) = f_2(x; \epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-). \quad (5.31)$$

Finally, by substituting Eq.(5.30-5.31) into boundary conditions Eq.(5.4-5.5), we obtain a set of nonlinear equations as follows:

$$R^+(x) = h_1(\epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-) \quad (5.32)$$

$$R^-(x) = h_2(\epsilon^+, \epsilon^-, \nu_{ex}^+, \nu_{ex}^-, \gamma, g, R^+, R^-). \quad (5.33)$$

Eq.(5.32-5.33) can then be plugged into AUTO/MATCONT for the fixed point computation and bifurcation analysis. For each fixed point pair (R^+, R^-) , the associated probability density profile can be obtained using Eq.(5.16-5.17).

Acknowledgements. This work was partially supported by the US Department of Energy, the National Science Foundation and the US Air Force Office of Scientific Research.

References

- [1] E. Blanchart, N. Marilleau, J.L. Chotte, A. Drogoul, E. Perrier, C. Cambier. *Sworm: an agent-based model to simulate the effect of earthworms on soil structure*. European Journal of Soil Science, 60 (1) (2009), 13–21.
- [2] J. Buhl, D.J.T. Sumpter, I.D. Couzin, J.J. Hale, E.Despland, E.R. Miller, S.J. Simpson. *From disorder to order in marching locusts*. Science, 312 (5778) (2006), 1402–1406.

- [3] F. Castiglione, F. Pappalardo, M. Bernaschi, S. Motta. *Optimization of HAART with genetic algorithms and agent-based models of HIV infection*. *Bioinformatics*, 23 (24) 2007, 3350–3355.
- [4] R.R. Coifman, S. Lafon, A. Lee, M. Maggioni, B. Nadler, F. Warner, S. Zucker. *Geometric diffusions as a tool for harmonic analysis and structure definition of data: Multiscale methods*. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (21) (2005), 7432–7437.
- [5] R.R. Coifman, S. Lafon, A.B. Lee, M. Maggioni, B. Nadler, F. Warner, S.W. Zucker. *Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps*. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (21) (2005), 7426–7431.
- [6] A. Dhooge, W. Govaerts, Y.A. Kuznetsov. *MATCONT: a MATLAB package for numerical bifurcation analysis of ODEs*. *ACM Trans. Math. Softw.*, 29 (2) (2003), 141–164.
- [7] E.J. Doedel, R.C. Paffenroth, A.R. Champneys, T.F. Fairgrieve, Y.A. Kuznetsov, B.E. Oldeman, B. Sandstede, X. Wang. *AUTO 2000: Continuation and bifurcation software for ordinary differential equations (with HomCont)*. Technical Report. California Institute of Technology, Pasadena, CA, 2001.
- [8] D.T. Gillespie. *Markov Process*. Academic Press, San Diego, 1992.
- [9] L. Hamill, N. Gilbert. *Social circles: A simple structure for agent-based social network models*. *The Journal of Artificial Societies and Social Simulation*, 12 (2) (2009), 3.
- [10] F.L. Hellweger, V. Bucci. *A bunch of tiny individuals-individual-based modeling for microbes*. *Ecological Modelling*, 220(1) (2009), 8–22.
- [11] E. Ilie-Zudor, L. Monostori. *Agent-based framework for pre-contractual evaluation of participants in project-delivery supply-chains*. *Assembly Automation*, 29 (2) (2009), 137–153.
- [12] M.A. Janssen, L.N. Alessa, M. Barton, S. Bergin, A. Lee. *Towards a community framework for agent-based modelling*. *The Journal of Artificial Societies and Social Simulation*, 11 (2) (2008), 6.
- [13] I.G. Kevrekidis, C.W. Gear, G. Hummer. *Equation-free: The computer-aided analysis of complex multiscale systems*. *AIChE Journal*, 50 (7) (2004), 1346–1355.
- [14] I.G. Kevrekidis, C.W. Gear, J.M. Hyman, P.G. Kevrekidis, O. Runborg, C. Theodoropoulos. *Equation-free coarse-grained multiscale computation: enabling microscopic simulators to perform system-level tasks*. *Comm. Math. Sciences*, 1 (4) (2003), 715–762.
- [15] I.G. Kevrekidis, G. Samaey. *Equation-free multiscale computation: Algorithms and applications*. *Annual Review of Physical Chemistry*, 60 (2009), 321–344.
- [16] B. Nadler, S. Lafon, R.R. Coifman, I.G. Kevrekidis. *Diffusion Maps, Spectral Clustering and Reaction Coordinates of Dynamical Systems*. *Applied and Computational Harmonic Analysis*, 21 (2006), 113–127.
- [17] I. Nishizaki, H. Katagiri, T. Oyama. *Simulation analysis using multi-agent systems for social norms*. *Computational Economics*, 34 (1) (2009), 37–65.
- [18] A. Omurtag, L. Sirovich. *Modeling a large population of traders: Mimesis and stability*. *Journal of Economic Behavior & Organization*, 61 (4) (2006), 562–576.
- [19] A. Pan, S.Y.S. Leung, K.L. Moon, K.W. Yeung. *Optimal reorder decision-making in the agent-based apparel supply chain*. *Expert Systems with Applications*, 36 (4) (2009), 8571–8581.
- [20] R. Pinnau. *Model Reduction via Proper Orthogonal Decomposition*. In *Model Order Reduction: Theory, Research Aspects and Applications*. W. A. Schilders, H. van der Vorst, J. Rommes, Eds. vol. 13. Springer, Berlin Heidelberg, (2008), 95–109.
- [21] H. Risken. *The Fokker-Planck Equation. Methods of Solution and Applications*. Second edition. Springer, Berlin, 1989.
- [22] O. Runborg, C. Theodoropoulos, I.G. Kevrekidis. *Effective bifurcation analysis: a time-stepper-based approach*. *Nonlinearity*, 15 (2) (2002), 491–511.
- [23] E. Samanidou, E. Zschischang, D. Stauffer, T. Lux. *Agent-based models of financial markets*. *Reports on Progress in Physics*, 70 (3) (2007), 409–450.
- [24] T. Shimokawa, K. Suzuki, T. Misawa. *An agent-based approach to financial stylized facts*. *Physica A*, 379 (1) (2007), 207–225.
- [25] C.I. Siettos, C.W. Gear, I.G. Kevrekidis. *An equation-free approach to agent-based computation: bifurcation analysis and control of stationary states*. *Europhys. Lett.*, 99 (4) (2012), 48007.
- [26] J.B. Tenenbaum, V. de Silva, and J.C. Langford. *A global geometric framework for nonlinear dimensionality reduction*. *Science*, 290 (5500) (2000), 2319.
- [27] B.C. Thorne, A.M. Bailey, S.M. Peirce. *Combining experiments with multi-cell agent-based modeling to study biological tissue patterning*. *Briefings in Bioinformatics*, 8 (4) (2007), 245–257.
- [28] D. Tykhonov, C. Jonker, S. Meijer, T. Verwaart. *Agent-based simulation of the trust and tracing game for supply chains and networks*. *The Journal of Artificial Societies and Social Simulation*, 11 (3) (2008), 1.
- [29] F.H. Westerhoff. *The use of agent-based financial market models to test the effectiveness of regulatory policies*. *Jahrbucher Fur Nationalokonomie Und Statistik*, 228 (2-3) (2008), 195–227.
- [30] L. Zhang, Z.H. Wang, J.A. Sagotsky, T.S. Deisboeck. *Multiscale agent-based cancer modeling*. *Journal of Mathematical Biology*, 58 (4-5) (2009), 545–559.