ANALYSIS AND SIMULATION OF RARE EVENTS FOR SPDEs [*]

Charles-Edouard Bréhier[1, 2], Maxime Gazeau[3], Ludovic Goudenège[4] and Mathias Rousset[1,2]

Abstract. In this work, we consider the numerical estimation of the probability for a stochastic process to hit a set B before reaching another set A. This event is assumed to be rare. We consider *reactive trajectories* of the stochastic Allen-Cahn partial differential evolution equation (with double well potential) in dimension 1. Reactive trajectories are defined as the probability distribution of the trajectories of a stochastic process, conditioned by the event of hitting B before A. We investigate the use of the so-called Adaptive Multilevel Splitting algorithm in order to estimate the rare event and simulate reactive trajectories. This algorithm uses a *reaction coordinate* (a real valued function of state space defining level sets), and is based on (i) the selection, among several replicas of the system having hit A before B, of those with maximal reaction coordinate; (ii) iteration of the latter step. We choose for the reaction coordinate the average magnetization, and for B the minimum of the well opposite to the initial condition. We discuss the context, prove that the algorithm has a sense in the usual functional setting, and numerically test the method (estimation of rare event, and transition state sampling).

1. Introduction

1.1. The goal

This paper focuses on the metastable states and associated rare events of reversible stochastic gradient systems in general dimension (finite or infinite). Reversible stochastic gradient systems can be used as a paradigmatic model to study physical metastability, that it is to say systems exhibiting very stable states (in terms of time scale), which are however very unlikely. In $\mathbb{R}^d$ with the usual euclidean structure, typical reversible stochastic gradient systems are given by diffusions solutions of Stochastic Differential Equations (SDEs) of the form:

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon}\,dB_t, \tag{1}$$

where $t \mapsto B_t \in \mathbb{R}^d$ is a standard Wiener process (Brownian Motion). If $V : \mathbb{R}^d \to \mathbb{R}$ is a smooth map such that:

(i) For a small but fixed value of ϵ, the reversible invariant Gibbs measure

$$\mu_\epsilon(dx) := \frac{1}{Z_\epsilon} e^{-\frac{1}{\epsilon} V(x)} \, dx, \tag{2}$$

is a probability measure (the equilibrium distribution) for the appropriate normalisation constant Z_ϵ;
(ii) the local minima of V form a countable set of isolated points;

then it is well-known that when $\epsilon \to 0$, the equilibrium distribution concentrates on the global minimum

$$V(x_0) = \inf_{x \in \mathbb{R}^d} V(x),$$

while all other local minima become metastable states (precise statements are known as the Freidlin-Wentzell large deviation theory, see below). A typical physical context leading to models of the form (1) is given by stochastically perturbed molecular dynamics. In the latter, V is the classical interaction potential energy between the atoms of molecules, and X_t represents the position of the atoms after a mass weighted linear change of coordinates. Moreover, the stochastic perturbation (thermostat) is assumed to be Markovian, with very strong viscosity (the so-called "overdamped" limit), and adimensional temperature proportional to ϵ. Finally, we assume we have quotiented out the continuous symmetries of the system so that V has indeed only isolated local minima.

Let us now define two closed subset $A \subset \mathbb{R}^d$ and $B \subset \mathbb{R}^d$; for instance, one may keep in mind the case where $X_0 = x_0$, $A = \bigcup_{n \geq 2} B(x_n, \delta)$ and $B = B(x_1, \delta)$, the latter being small balls centered at the local minima $x_n, n \geq 1$. We next define the hitting time of A (or equivalently B) as

$$\tau_A \equiv \tau_A^\epsilon = \inf \{t \geq 0; X_t \in A\}.$$

We are interested in computing the rare event probability

$$0 < \mathbb{P}(\{\tau_B < \tau_A\}) \ll 1, \tag{3}$$

as well as sampling according to the *reactive path* ensemble, defined as the conditional probability distribution on trajectories

$$\text{Law}(X_t, \, t \geq 0 | \{\tau_B < \tau_A\}). \tag{4}$$

In the following subsections, we are about to introduce the Allen Cahn equation which will be our model for metastable trajectories in infinite dimension. A brief description in finite dimensions of splitting algorithms (multiple replicas based algorithms) is given as an alternative to crude Monte Carlo simulations. The large deviation theory, introduced next, gives a mathematical justification of the validity of this numerical approach.

1.2. **Infinite dimension**

This work will also focus on gradient systems in *infinite dimension*. As a consequence the stochastic differential equation (1) will be replaced by a Stochastic Partial Differential Equation perturbed with a space-time white noise - see also the abstract formulation (12):

$$\begin{cases} \dfrac{\partial x(t, \lambda)}{\partial t} = \gamma \dfrac{\partial^2 x(t, \lambda)}{\partial \lambda^2} - \nabla V(x(t, \lambda)) + \sqrt{2\epsilon} \dfrac{\partial^2 \omega(t, \lambda)}{\partial t \partial \lambda}, \quad t > 0 \text{ and } \lambda \in (0, 1) \\ \dfrac{\partial x(t, \lambda)}{\partial \lambda}\Big|_{\lambda=1} = \dfrac{\partial x(t, \lambda)}{\partial \lambda}\Big|_{\lambda=0} = 0 \\ x(0, \lambda) = x_0(\lambda), \end{cases} \tag{5}$$

where x_0 is a given initial condition. The second line expresses the homogeneous Neumann boundary conditions at $\lambda = 0$ and $\lambda = 1$. The deterministic forcing term ∇V is a sufficiently regular vector field on a Banach space H, and $\mathcal{V}(x) = \int_0^1 V(x(\lambda)) d\lambda$ is a non-quadratic potential energy. The problem of sampling according to the rare event (3) and the simulation of reactive trajectories (4) remains unchanged.

1.3. **Splitting algorithms**

In finite dimension, more efficient algorithms than the plain Monte-Carlo simulation of independent realizations of (1) were developed introducing a real valued function on the state space, usually called *reaction coordinate*, and denoted by

$$\xi : \mathbb{R}^d \to \mathbb{R}.$$

We will assume that $X_0 = x_0$ with

$$\sup_{x \in A} \xi(x) \leq z_A < \xi(x_0) < z_B \leq \inf_{x \in B} \xi(x).$$

The general key point of splitting methods to compute the rare event $\{\tau_B < \tau_A\}$ consists in simulating n_{rep} replicas of (1), and to duplicate with higher probability the trajectories with highest maximal level ξ. It turns out that this can be done in a consistent way. In Section 4, a general adaptive algorithm inspired by [7, 8] and an associated unbiased estimator of the probability of $\{\tau_B < \tau_A\}$ will be presented, and numerically tested in Section 5.

One objective of this paper is to generalize this approach to an infinite dimensional setting.

1.4. **Large deviations and Transition State Theory**

The behavior of the solutions to either (1) or (5) in the limit $\epsilon \to 0$ is well explained by the theory of large deviations (see [12, th. 4.2, p. 127]). On a finite time window $[0, T]$, it can then be checked that on the space of trajectories (with uniform topology)

$$W(x, y) = \inf_{T \geq 0} \inf_{\gamma_0 = x, \gamma_T = y} I(\gamma) = \inf_{\gamma : x \to y} \left(\sup_{z \in \{\gamma(t); t \in \mathbb{R}^+\}} V(z) - V(x) \right) \tag{6}$$

where γ is taken in the set of all continuous paths such that $\gamma_0 = x$ and $\gamma(t) \xrightarrow[t \to +\infty]{} y$ and where the *good rate function* is given by

$$I(\gamma) := \frac{1}{4} \int_0^T \left| \nabla V(\gamma_t) - \dot{\gamma}_t \right|^2 \, dt,$$

and where extrema are taken over smooth trajectories. The left-hand side of (6) is the so-called *quasi-potential function* defined for any $(x, y) \in \mathbb{R}^d$ as the minimal cost of forcing the system to go from x to y in an indeterminate time. For gradient systems, the latter formula shows that it is the lowest energy barrier that needs to be overcome in order to reach y from x. The quasi-potential yields a rationale to compute rare events related to *exit times*, since for any $\delta > 0$, we have the general formula

$$\lim_{\epsilon \to 0} \mathbb{P}_{X_0 = x}\big(e^{(\inf_{y \in \partial A} W(x, y) - \delta)/\epsilon} < \tau_A^\epsilon < e^{(\inf_{y \in \partial A} W(x, y) + \delta)/\epsilon} \big) = 1.$$

This shows that if $\inf_{y \in \partial A} W(x, y) < \inf_{y \in \partial B} W(x, y)$, then the event $\{\tau_B < \tau_A\}$ is indeed a rare event in the limit $\epsilon \to 0$. Tools of Potential Theory are used to go further allowing to compute the so-called prefactor; see [5] for the general methodology and results in finite dimension. In the double well situation, where x_0 denotes saddle point and $x_{-/+}$ the minima, Kramer's law holds in the sense that

$$\mathbb{E}_{x_-}\big(\tau_{B_\delta(x_+)}\big) = \frac{2\pi}{|\lambda_1(x_0)|} \sqrt{\frac{|\det(\nabla^2 V(x_0))|}{\det(\nabla^2 V(x_-))}} e^{(V(x_0) - V(x_-))/\epsilon} \left[1 + \mathcal{O}\left(\epsilon^{1/2} |\log \epsilon|^{3/2} \right) \right]$$

where $\tau_{B_\delta(x)}$ is a small ball of radius δ centered at x and $\lambda_1(x_0)$ is the single negative eigenvalue of the Hessian $\nabla^2 V(x_0)$.

In the infinite dimensional setting, the choice of functional spaces is essential. In [9], the large deviation result is given in Hölder-like spaces $\mathcal{C}^{2\beta, 0}([0, 1])$ for $\beta \in]0, 1/4[$, with Dirichlet boundary conditions. In [1, 3], the Kramer's law is given respectively with respect to Sobolev and Hölder norms. But this kind of information

is asymptotic, merely theoretical, and the direct computation of the above quantities is not possible in general. This justifies the introduction of efficient numerical methods.

The remainder of this paper is organised as follows. In Section 2, we justify the use of the Allen Cahn equation as a limit of large particules system. Then, we describe a possible numerical scheme for the Allen Cahn equation based on the Baker-Campbell-Hausdorff formula. The numerical estimation of rare events requires to pay attention on the chosen discretisation since its accuracy depends on the time step. A trade-off has to be done between the order of convergence of the method and the computational time which tends to be very long as the time step goes to zero. In Section 3, we show the influence of the parameter γ on the shape of the potential for a system with 2 and 4 particles. This is done to emphasize the role of this bifurcation parameter and the importance of adapting the numerical methods depending on its value. Indeed, local minima may appear and one has to be sure that trajectories are not going to be trapped in one of them. The situation in infinite dimension is similar but more complex to describe. Finally in Section 4, we introduce the Adaptive Multilevel Splitting (AMS) as an alternative method to Monte Carlo to compute rare events (3) and reactive trajectories, i.e. trajectories going from one minimum to another. Stated like this, AMS algorithm is not efficient to deal with situations where bifurcations occur unless theoretical results are known. This question remains open when nothing is known on bifurcations. Note that this algorithm may also be used to compute the time length of reactive trajectories, its probability density function and transition times, namely the time required to go from one metastable state to another one (see [8]). The AMS algorithm is applied in Section 5 where numerical simulations are displayed.

2. Finite and infinite dimensional models

In this Section, we introduce two different mathematical models with metastability, which are linked together through their energies. The first model is a coupled system of N stochastic differential equations called the overdamped Langevin equation. It models the evolution of N interacting particles driven by random forces. The second one is a stochastic partial differential equation known as the Allen-Cahn equation. The main motivation of this study is that the second equation can be viewed, under an appropriate scaling, as the limit of the SDE when N goes to infinity.

We first introduce known results on these two problems and necessary notations especially on the functional spaces. We also present various discretization methods that can be implemented to obtain numerical approximations of these processes.

2.1. **The finite dimensional model**

Throughout this paper, we will use the following terminology: an *atom* denotes a particle and a system of N atoms is defined as a *molecule*, with $N \in \mathbb{N}$ and $N \geq 2$. We model a molecule with N atoms moving on a line and submitted to three kinds of forces:

(1) each atom $i \in [\![1, N]\!]$ is confined by a potential $V : \mathbb{R} \to \mathbb{R}$ with at least two wells. One particular interesting case, considered in this paper, is the symmetric double-well potential $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$;
(2) each atom i interacts with its two nearest neighbors $i - 1$ and $i + 1$ through a quadratic potential;
(3) the movement of each atom is perturbed by a small Gaussian white noise in time. The processes acting on different atoms are mutually independent.

The noise perturbation is represented by a N-dimensional Wiener process $W^N = (W_1^N, \ldots, W_N^N)$ whose covariance is given for any $u, v \in \mathbb{R}^N$ by

$$\mathbb{E}\big[\langle W^N(t), u \rangle \langle W^N(s), v \rangle\big] = \min(t, s) \langle u, v \rangle, \tag{7}$$

for any $s, t \in \mathbb{R}^+$ and where $\langle ., . \rangle$ denotes the canonical scalar product on $\mathbb{R}^N$. The position of the atom i at time t is denoted by $X_i^N(t)$. The configuration of the molecule $(X_i^N)_{1 \leq i \leq N}$ satisfies the following system of

Stochastic Differential Equations for $t > 0$ and any $i \in \{1, \cdots, N\}$

$$dX_i^N(t) = \gamma N(X_{i-1}^N(t) + X_{i+1}^N(t) - 2X_i^N(t))dt \tag{8}$$
$$- \frac{1}{N}V'(X_i^N(t))dt + \sqrt{2\epsilon}dW_i^N(t),$$

with homogeneous discrete Neumann type boundary conditions $X_{N+1}^N := X_N^N$, $X_0^N := X_1^N$ and for an initial configuration $(X_i(0))_{1 \leq i \leq N} = (x_i^N)_{1 \leq i \leq N}$.

The positive parameter $\gamma > 0$ determines the rigidity between the atoms. The value of γ is of importance in the study of metastability. Indeed the nature of some stationary points of the energy defined below is modified, as well as the reactive trajectories between the metastable states. The parameter $\epsilon > 0$ represents the level of the noise acting on the molecule. Physically it can be interpreted as the temperature. One of our concerns is to investigate numerically the evolution of the molecule when ϵ decreases to 0 and we aim to propose efficient algorithms to estimate the transitions rates from one stable state to another. The total energy $\mathcal{E}_\gamma^N$ of the system is given for any $x^N = (x_1^N, \ldots, x_N^N) \in \mathbb{R}^N$

$$\mathcal{E}_\gamma^N(x^N) = \frac{\gamma}{2N}\sum_{i=0}^{N} N^2(x_{i+1}^N - x_i^N)^2 + \frac{1}{N}\sum_{i=1}^{N} V(x_i^N), \tag{9}$$

with the boundary conditions $x_{N+1}^N := x_N^N$, $x_0^N := x_1^N$. The system (8) can be rewritten as the *overdamped Langevin dynamics* in $\mathbb{R}^N$ associated with the energy $\mathcal{E}_\gamma^N$. Indeed, (8) can be rewritten as

$$dX^N(t) = -\nabla\mathcal{E}_\gamma^N(X^N(t))dt + \sqrt{2\epsilon}dW^N(t). \tag{10}$$

The regularity of the potential $V : \mathbb{R} \to \mathbb{R}$ is important for the global well-posedness of this system. We will give sufficient regularity conditions for it.

Remark 2.1. This choice of boundary terms is linked to the homogeneous Neumann boundary conditions imposed on the SPDE below. Moreover the scaling of the different contributions in the energy with respect to the size N in the molecule is related to the convergence when $N \to +\infty$.

2.2. Infinite dimensional model

Let $x : [0, 1] \to \mathbb{R}$ be a function of class $\mathcal{C}^1$ on the space interval $[0, 1]$. For any positive integer N, we denote $\Delta x = \frac{1}{N+1}$ and $x_i^N = x(i\Delta x)$. The generic space variable in the interval $[0, 1]$ is denoted by λ. Then

$$\mathcal{E}_\gamma^N(x_1^N, \ldots, x_N^N) \xrightarrow[N \to +\infty]{} \mathcal{E}_\gamma^\infty(x) = \frac{\gamma}{2}\int_0^1 |x'(\lambda)|^2 d\lambda + \int_0^1 V(x(\lambda))d\lambda$$

as Riemann sums. Inspired by this convergence for smooth functions, we define a new energy on $H^1(0, 1)$, the space of square integrable functions defined on $(0, 1)$ having a derivative (in the distributional sense) belonging to $L^2(0, 1)$:

$$\mathcal{E}_\gamma^\infty(x) = \frac{\gamma}{2}\int_0^1 |x'(\lambda)|^2 d\lambda + \int_0^1 V(x(\lambda))d\lambda. \tag{11}$$

We recall that a Sobolev embedding ensures that $H^1(0, 1)$ is continuously embedded into $\mathcal{C}([0, 1])$; thus the function $V \circ x$ is well-defined and continuous on $[0, 1]$ for any $x \in H^1(0, 1)$.

The stochastic partial differential equation linked to this energy is the Allen-Cahn equation - see (5). In an abstract form, it is written

$$\begin{cases} dX(t) = AX(t)dt - \nabla V(X(t))dt + \sqrt{2\epsilon}dW(t), \\ X(0) = x_0. \end{cases} \tag{12}$$

The interpretation of (12), as a stochastic evolution equation on (an appropriate subset of) the Hilbert space $H = L^2(0,1)$, follows the classical framework of [11]. The operator A denotes the realization on H of the Laplace operator $\frac{\partial^2}{\partial \lambda^2}$ with homogeneous Neumann boundary conditions. Accordingly its domain is

$$\mathcal{D}(A) = \left\{ \phi \in H^2(0,1), \phi'(0) = \phi'(1) = 0 \right\}.$$

The non-linear mapping $F(x)(\lambda) = -\nabla V(x(\lambda))$, for almost all $\lambda \in (0,1)$, is defined on the domain $D(F) = L^p(0,1)$, where p is such that $F(x) \in L^2(0,1)$ for any $x \in L^p(0,1)$. In the case of the Allen-Cahn potential, $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$, and one can take $p = 6$. More general assumptions for V are possible:

(1) V is of class C^3;
(2) V is convex at infinity, i.e. there exists two constants $C > 0$ and $c > 0$ such that for any $|x| > c$ then

$$V''(x) > C > 0;$$

(3) There exists two constants $C > 0$ and $\alpha \geqslant 2$ such that for any $x \in \mathbb{R}$ and $j = 0, 1, 2, 3$

$$\left| V^{(j)}(x) \right| \leqslant C(|x|^{2\alpha - 1} + 1).$$

The noise perturbation in (12) is induced by a centered, Gaussian space-time white noise. Its covariance satisfies for any times s, t and any positions λ, μ

$$\mathbb{E}[\omega(t, \lambda)\omega(s, \mu)] = \min(t, s) \times \min(\lambda, \mu).$$

The noise W is given by a cylindrical Wiener process. Formally, it is defined as the following series

$$W(t) = \sum_{k \in \mathbb{N}} W_k(t)e_k,$$

where $(e_k)_{k \in \mathbb{N}}$ is any complete orthonormal system of the Hilbert space H, and $(W_k)_{k \in \mathbb{N}}$ is a sequence of independent standard scalar Wiener processes. The convergence of this series does not hold in H, but only on larger spaces in which H is embedded thanks to a linear Hilbert-Schmidt operator. The eigenvalues of A are then $\lambda_n = (n\pi)^2$ for $n \in \mathbb{N}$ and the corresponding eigenfunctions are $e_n(\lambda) = \sqrt{2}\cos(n\pi\lambda)$ for $n = 1, 2 \ldots$, and $e_0(\lambda) = 1$; they satisfy

$$\sum_{k=1}^{+\infty} \frac{1}{\lambda_k^{\alpha}} < \infty \Leftrightarrow \alpha > 1/2.$$

If we denote $\widetilde{e}_k = \frac{e_k}{\lambda_k^{1/2+\delta}}$ for $k \geq 1$ and $\widetilde{e}_0 = e_0$, then the series

$$\widetilde{W}(t) = \sum_{k \in \mathbb{N}} W_k(t)\widetilde{e}_k,$$

converges. The white noise is thus well defined in $H^{-1/2-\delta}$ for any $\delta > 0$. The heat kernel will play the role of the Hilbert Schmidt operator and give a rigorous meaning to the stochastic convolution

$$W_A(t) = \int_0^t e^{(t-s)A}dW(s)$$

in $H^{1/2-\delta}$, for any $\delta > 0$.

Proposition 2.2. *Let $\delta > 0$, and fix a suitable parameter p. For any final time $T > 0$ and continuous initial data $x_0 \in L^p(0,1)$ satisfying the Neumann boundary condition, then Equation (12) admits a unique mild solution in $L^2(\Omega, C([0,T] \times [0,1]))$ in the sense that there is a unique solution to the following Duhamel equation*

$$X(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}F(X(s))ds + \sqrt{2\epsilon}\int_0^t e^{(t-s)A}dW(s), \tag{13}$$

such that for every $t \in [0, T]$ we have $X(t) \in L^p(0, 1)$. Moreover it admits a modification which is of class $C^{1/4-\delta}$ in time and $C^{1/2-\delta}$ in space. The solution of the linearized equation is in $H^{1/2-\delta}$ for any $\delta > 0$. Moreover, we have the following stronger estimate:

$$\mathbb{E}[\sup_{t \in [0,T], \lambda \in [0,1]} |x(t, \lambda)|^2] \leq C(T) < +\infty.$$

This proposition is proved using classical truncation and approximation arguments [1]. Moreover, up to a time change, model (8) can be seen as an approximation with a finite differences scheme of Equation (12). For this convergence result, see for instance [1].

2.3. Invariant distribution

In the case of the double-well potential, some critical points of the deterministic dynamics ($\epsilon = 0$) in both finite and infinite dimensions are easily identified. This property will lead to metastability, which can also be identified at the level of invariant distributions of each model.

For the overdamped Langevin dynamics in dimension $N < +\infty$, two global minima are the molecules

$$x_N^\pm = \pm(1, \dots, 1).$$

Another critical point is

$$x_N^0 = (0, \dots, 0);$$

however as explained in Section 3 the precise nature of x_N^0, given by the signature of the Hessian $D^2 \mathcal{E}_\gamma^N(x^0)$, depends on γ and N. Moreover when γ changes new critical points may appear.

In the infinite dimensional case, the corresponding global minima are the constant functions

$$x_\infty^\pm(\lambda) = \pm 1, \quad \text{for all } \lambda \in [0, 1].$$

Remark that these functions satisfy the homogeneous Neumann boundary conditions.

We also have a critical point such that $x_\infty^0(\lambda) = 0$.

A distinction must be done to analyze the invariant laws according to the dimension. There is ergodicity in all cases and a formula of Gibbs type, but with respect to a different type of reference measure which is the Lebesgue measure in $\mathbb{R}^N$ and a suitable (degenerate) Gaussian measure in H in infinite dimension.

Finite dimension: The dynamics is given by the overdamped Langevin equation (10), with the energy $\mathcal{E}_\gamma^N$. Assumptions on V ensure that the probability measure defined with

$$\mu_\gamma^{N,\epsilon}(dx) = \frac{1}{Z_\gamma^{N,\epsilon}} \exp\left(-\frac{\mathcal{E}_\gamma^N(x)}{\epsilon}\right) dx,$$

is the unique invariant law, where dx denotes Lebesgue measure on $\mathbb{R}^N$ and $Z_\gamma^{N,\epsilon}$ is a normalization constant. Moreover, the stationary process is ergodic and reversible with respect to this invariant measure. There is another way to write $\mu_\gamma^{N,\epsilon}$, when $\gamma > 0$, as a Gibbs measure with respect to the measure

$$\nu_\gamma^{N,\epsilon}(dx) = \exp\left(-\frac{\gamma \mathcal{D}^N(x)}{\epsilon}\right) dx,$$

using the following decomposition of the energy into kinetic and potential energy of the atoms in the molecule:

$$\mathcal{E}_\gamma^N = \gamma \mathcal{D}^N + \mathcal{V}^N(x)$$

$$\mathcal{D}^N(x) = \frac{1}{2N} \sum_{i=1}^{N} N^2 (x_{i+1}^N - x_i^N)^2 \tag{14}$$

$$\mathcal{V}^N(x) = \frac{1}{N} \sum_{i=1}^{N} V(x_i^N).$$

The measure $\nu_\gamma^{N,\epsilon}$ cannot be normalized as a probability measure. It is a degenerate Gaussian measure, which can be seen as the unique (up to a multiplicative constant) invariant measure of the linearized process - i.e. with $V = 0$. Using a suitable change of variables, it is expressed as the product of the one-dimensional Lebesgue measure and of a non-degenerate Gaussian probability measure. The degeneracy is explained by the choice of Neumann boundary conditions: from (8) with $V = 0$ (linearization) and $\epsilon = 0$ (deterministic case) we see that the average $\xi(X_1, \ldots, X_N) = \frac{1}{N} \sum i = 1^N X_i$ is preserved. When $\epsilon > 0$, there is no dissipation on this mode, which is solution of $d\xi_t = \sqrt{2\epsilon/N} dB^N(t)$ where $B^N(t) = \frac{1}{N} \sum_{i=1}^{N} W_i^N(t)$.

Infinite dimension: The above discussion needs to be adapted. Recall that if $X(t)$ is solution of the SPDE then almost surely, the quantity $\mathcal{E}_\gamma(X(t))$ is not finite for $t > 0$. However, there is an expression of the invariant law, with a density with respect to a reference measure $\nu_\gamma^{\infty,\epsilon}$. The linear SPDE - with $V = 0$ - is written

$$dY(t) = \gamma A Y(t) + \sqrt{2\epsilon} dW(t).$$

We decompose the Hilbert space H into $H = H_0 + H_0^\perp$ with $H_0 = \mathrm{Span}(e_0)$. Then if we denote $Y(t) = Y_0(t) + Y_0^\perp$, the above linear SPDE is decomposed into two decoupled equations

$$dY_0(t) = \sqrt{2\epsilon} dW_0(t),$$
$$dY_0^\perp(t) = \gamma A Y_0^\perp(t) + \sqrt{2\epsilon} dW_0^\perp(t).$$

$H_0^\perp = \mathrm{Span}\{e_i; i \geq 1\}$ is a stable subspace of A, and A is invertible on $H_0^\perp$, with inverse denoted $A_\perp^1$. The equation on Y_0 admits Lebesgue measure $\nu_0(dx_0) = dx_0$ as unique (up to a multiplicative constant) invariant measure. The equation on $Y_0^\perp$ admits a unique invariant law, which is a Gaussian measure on $H_0^\perp$, denoted by $\nu_0^\perp(dx_0^\perp)$; it is centered and its covariance operator is $\frac{1}{\gamma\epsilon} A_\perp^1$. Therefore the linearized SPDE admits the following invariant measure defined in H:

$$\nu_\gamma^{\infty,\epsilon}(dx) = \nu_0(dx_0) \times \nu_0^{orth}(dx_0^\perp).$$

It is worth noting that for any $p \in [2, +\infty[$ the Banach space is included in the support of $\nu_\gamma^{\infty,\epsilon}$. Assumptions on the potential V now ensure that the SPDE (12) admits a unique invariant probability measure when $\gamma > 0$, denoted by $\mu_\gamma^{\infty,\epsilon}$, with the expression

$$\mu_\gamma^{\infty,\epsilon}(dx) = \frac{1}{Z_\gamma^{\infty,\epsilon}} \exp\left\{-\frac{1}{\epsilon} \int_0^1 V(x(\lambda)) d\lambda\right\} \nu_\gamma^{\infty,\epsilon}(dx),$$

where $Z_\gamma^{\infty,\epsilon}$ is a normalization constant.

Remark 2.3. The case of homogeneous Dirichlet boundary conditions is studied in [13, 15]. There is an interpretation of the Gaussian measure $\nu_\gamma^{\infty,\epsilon}$ as the law of the Brownian Bridge, with a renormalization with respect to the parameters γ and ϵ.

Transitions between the stable equilibrium points appear when the temperature parameter ϵ is positive; the typical time can be expressed thanks to the Kramers law.

2.4. **Discretization of the stochastic processes**

The aim of this part is to introduce somehow classical schemes to solve numerically the Langevin over-damped equation and the stochastic Allen-Cahn equation. For the finite dimensional system we consider the well known Euler scheme

$$X^{n+1} = X^n - \nabla \mathcal{E}_\gamma^N(X^n) \Delta t + \sqrt{2\epsilon\Delta t} G_n$$

where G_n are standard Gaussian random variables. This scheme is proved to be of order $1/2$ in $L^p(\Omega)$ norms for any $p \geq 2$. Regarding the numerical approximation of the stochastic Allen-Cahn equation, various numerical schemes may be proposed that are based on deterministic scheme. We consider here finite difference schemes but other methods may be used efficiently such as finite elements. Spectral methods seem to be a little harder to implement. Indeed one has to use Chebychev polynomials instead of the fast Fourier

transform because Neumann conditions are imposed on the boundary. The nonlinear term is handled using a splitting method. The basic idea of splitting methods (see [14]) is to approach the exact flow $L(t)X_0$ of the nonlinear equation by means of a truncation of the Baker-Campbell-Hausdorff formula.

From now on we consider a possible splitting scheme for the Allen Cahn equation (5), with the potential $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$.

Let us denote by $S(t)X_0$ the solution of the stochastic heat equation

$$\begin{cases} dX(t) &= \gamma\partial_x^2 X(t)dt + \sqrt{2\epsilon}dW(t) \\ X(0) &= X_0, \end{cases}$$

and $T(t)Y_0$ the solution of the Bernoulli differential equation

$$\begin{cases} \partial_t Y(t) &= -\left(Y^3(t) - Y(t)\right) \\ Y(0) &= Y_0. \end{cases}$$

The first equation can be easily solved numerically using a second-order semi-implicit scheme to insure the unconditional stability of the scheme. The second equation is exactly solvable and its solution is given by

$$Y(t) = \frac{Y_0}{\sqrt{Y_0^2 + (1 - Y_0^2)\exp(-2t)}}.$$

The Lie method consists in approximating the exact solution of the stochastic Allen-Cahn equation $L(t)X_0$ by either one of the two methods $S(t)T(t)X_0$ or $T(t)S(t)X_0$ which correspond to a composition of the two previous flows. When $\epsilon = 0$, it is well known that the order of convergence of this scheme is 1 in time. Adding the white noise, the strong order of convergence usually drop to $1/4$. The numerical Lie scheme reads as follows:

$$\begin{cases} X_{n+1/2} - X_n = \dfrac{\gamma\Delta t}{4}\partial_x^2(X_{n+1/2} + X_n) + \dfrac{1}{2}\sqrt{2\epsilon\Delta t}G^n \\[2ex] X_{n+1} = \dfrac{X_{n+1/2}}{\sqrt{X_{n+1/2}^2 + (1 - X_{n+1/2}^2)\exp(-2\Delta t)}} \end{cases} \tag{15}$$

where $G^n = \sum_{j=0}^{J} G_j^n e_j$ and $G_j^n \sim \mathcal{N}(0,1)$ are iid. J is a truncation parameter for the expansion of the noise in an Hilbert basis made of eigenfunctions e_j of the linear operator ∂_x^2, with Neumann boundary conditions.

Higher order schemes may be constructed considering more terms in the Baker-Campbell-Hausdorff formula. For time independent and deterministic operators, a method to construct even order symplectic integrator is proposed in [16]. A natural extension of such schemes is:

$$\begin{cases} X_{n+1/3} - X_n = \dfrac{\gamma\Delta t}{4}\partial_x^2(X_{n+1/3} + X_n) + \dfrac{1}{2}\sqrt{2\epsilon\Delta t}G^n \\[2ex] X_{n+2/3} = \dfrac{X_{n+1/3}}{\sqrt{X_{n+1/3}^2 + (1 - X_{n+1/3}^2)\exp(-2\Delta t)}} \\[2ex] X_{n+1} - X_{n+2/3} = \dfrac{\gamma\Delta t}{4}\partial_x^2(X_{n+1} + X_{n+2/3}) + \dfrac{1}{2}\sqrt{2\epsilon\Delta t}G^n \end{cases} \tag{16}$$

The equation is then discretized with finite differences. We have implemented this method for our simulations; however the analysis of its order of convergence is an open question.

3. BIFURCATIONS

The aim of this Section is to describe the qualitative behavior of the system at the level of the transitions between metastable states, which we will recover with numerical simulations in Section 5.

The total energy $\mathcal{E}_\gamma^N$ depends on a parameter γ that determines the influence of the gradient energy with respect to the potential one. It is refered to as a bifurcation parameter and a small change in its value may suddenly change the behavior of the dynamical system. It can be seen studying the nature and the number of the critical points of the energy. In both finite and infinite dimension, one may thus expect that the reactive trajectories do not experiment the same paths according to the value of γ. To illustrate this, let us consider the energy functional for $N = 2$ as defined in (9) and study its critical points. The approach here is inspired by [4, 10].

Proposition 3.1. *For $N = 2$, the local maximum 0 is unstable in the sense that it degenerates into a local maximum at the bifurcation parameter $\gamma = 1/8$ in the direction given by the vector $(1, -1)$. Then for $\gamma < 1/8$, two saddle points appear at $\pm\sqrt{1 - 8\gamma}$. A new change of regime occurs at $\gamma = 1/12$ and these two saddle points degenerate to local minima; four new saddle points appear when $\gamma < 1/12$.*

For $N = 4$, the saddle point 0 is also unstable. It degenerates at the bifurcation parameters

$$\gamma = \left\{ \frac{1}{16(2 - \sqrt{2})}, \frac{1}{32}, \frac{1}{16(2 + \sqrt{2})} \right\}.$$

For $N = 2$, bifurcations are studied by two approaches: a direct computation and introducing suitable normal forms. Normal forms are simplified functionals exhibiting the same structure of critical points than the full problem and describing the phase transitions in a neighbourhood of a bifurcation value. This is done for $N = 2$ to emphasize the importance of normal forms in cases when explicit computations can not be performed (for example $N = 4$).

Direct computation: The Jacobian matrix is

$$\nabla \mathcal{E}_\gamma^2(x, y) = \frac{1}{2} \begin{pmatrix} x^3 - x + 4\gamma(x - y) \\ y^3 - y - 4\gamma(x - y) \end{pmatrix}.$$

Summing the two equations, we easily obtain that the critical points satisfy the system

$$\begin{cases} (x + y)(x^2 - xy + y^2 - 1) = 0 \\ x^3 - x + 4\gamma(x - y) = 0. \end{cases}$$

Let us now determine the real roots of this system with respect to γ:
 (1) If $x = -y$ then x satisfies

$$x(x^2 + (8\gamma - 1)) = 0. \tag{17}$$

Consequently $x = 0$ or $x = \pm\sqrt{(1 - 8\gamma)}$.
 (2) If $x \neq -y$ then

$$\begin{cases} (x^2 - xy + y^2 - 1) = 0 \\ x^3 - x + 4\gamma(x - y) = 0. \end{cases} \tag{18}$$

From the first equation of system (18), $y = (x \pm \sqrt{4 - 3x^2})/2$. Plugging this expression into the second equation of (18) leads to an equation for x

$$(x - 1)(x + 1)(x^4 - x^2 + 4\gamma x^2 + 16\gamma^2) = 0,$$

whose roots are given by ± 1 and

$$\alpha(\gamma) = \pm\sqrt{\frac{1 - 4\gamma \pm \sqrt{(4\gamma + 1)(1 - 12\gamma)}}{2}}.$$

We now sum up the critical points of the energy according to some values of the bifurcation parameter γ

$$\gamma \geqslant \frac{1}{8} \implies \left\{ \begin{array}{l} x = y = 0 \\ x = y = \pm 1 \end{array} \right.$$

and

$$\frac{1}{12} \leqslant \gamma \leqslant \frac{1}{8} \implies \left\{ \begin{array}{c} x = y = 0 \\ x = y = \pm 1 \\ x = -y = \pm\sqrt{1 - 8\gamma} \end{array} \right.$$

If $\gamma \leqslant \frac{1}{12}$ then $\alpha(\gamma)$ is real valued and the critical points are

$$\left\{ \begin{array}{c} x = y = 0 \\ x = y = \pm 1 \\ x = -y = \pm\sqrt{1 - 8\gamma} \\ x = \alpha(\gamma) \text{ and } y = \frac{x(x^2 - 1 + 4\gamma)}{4\gamma}. \end{array} \right.$$

Let us now determine their nature and see how a small change of γ may influence the variations of the energy. The Hessian of the energy $\mathcal{E}_\gamma^2$ is given by

$$\mathcal{H}\mathcal{E}_\gamma^2(x, y) = \left(\begin{array}{cc} \frac{3}{2}x^2 - \frac{1}{2} + 2\gamma & -2\gamma \\ -2\gamma & \frac{3}{2}y^2 - \frac{1}{2} + 2\gamma \end{array} \right).$$

At the point $(0, 0)$, its spectrum is given by $Sp = \{-\frac{1}{2}, 4\gamma - \frac{1}{2}\}$. When $\gamma > 1/8$, the point $(0, 0)$ is a saddle point while it degenerates into a local maximum for $\gamma \leqslant 1/8$. Another change of regimes occurs at the points $(\pm\sqrt{1 - 8\gamma}, \mp\sqrt{1 - 8\gamma})$. Indeed

$$\mathcal{H}\mathcal{G}(\pm\sqrt{1 - 8\gamma}, \mp\sqrt{1 - 8\gamma}) = \left(\begin{array}{cc} 1 - 10\gamma & -2\gamma \\ -2\gamma & 1 - 10\gamma \end{array} \right) \implies Sp = \{1 - 8\gamma, 1 - 12\gamma\}.$$

Therefore for $1/8 \geqslant \gamma \geqslant 1/12$, these two points are saddles while they degenerate to local minima when $\gamma < 1/12$.

Normal form: Bifurcations at $(0, 0)$ are expected to be described by a normal form in a direction orthogonal to the eigenvector $(1, 1)$. In this simple case, the orthogonal space is easily identified and is $(1, -1)$. We decompose q along this basis

$$q = A \binom{1}{1} + \rho \binom{1}{-1}.$$

Plugging q into the energy, one easily gets $4\mathcal{E}_\gamma^2(\rho, A) = \mathcal{G}_0 + \mathcal{G}_1$, where

$$\left\{ \begin{array}{l} \mathcal{G}_0(A) = A^4 - 2A^2 \\ \mathcal{G}_1(\rho, A) = \rho^4 - 2\rho^2(1 - 8\gamma) + 6\rho^2 A^2. \end{array} \right.$$

We introduce a new quantity

$$G_1(A) := \min_\rho \mathcal{G}_1(\rho, A) = \left\{ \begin{array}{ll} -9A^4 + 6A^2 - 48\gamma A^2 - 1 + 16\gamma - 64\gamma^2 & \text{if } |A| < \sqrt{\frac{1 - 8\gamma}{3}} \\ 0 & \text{otherwise} \end{array} \right. \tag{19}$$

and the normal form is given by

$$G(A) := \mathcal{G}_0(A) + G_1(A) = \left\{ \begin{array}{ll} -8A^4 + 4A^2 - 48\gamma A^2 - 1 + 16\gamma - 64\gamma^2 & \text{if } |A| < \sqrt{\frac{1 - 8\gamma}{3}} \\ A^4 - 2A^2 & \text{otherwise.} \end{array} \right.$$

Then the nature of the normal form changes at $\gamma = 1/8$. For $\gamma > 1/8$, the normal form is locally concave around zero which is a local maximum. For $\gamma < 1/8$, the critical points of this functional are now 0 and $\pm\sqrt{1 - 12\gamma}/2$ possibly complex. For a small perturbation $\gamma_\epsilon = 1/8 - \epsilon$, $0 < \epsilon \ll 1$, the minimum of G is reached at 0 and

$$G(0) = -64\epsilon^2.$$

Then we conclude that $-G(A)$ describes the bifurcation of $\mathcal{E}_\gamma^2$ around $(0,0)$ in the direction $(1, -1)$. Fig 1 plots the isolines of the energy $\mathcal{E}_\gamma^2$ for different values of γ.
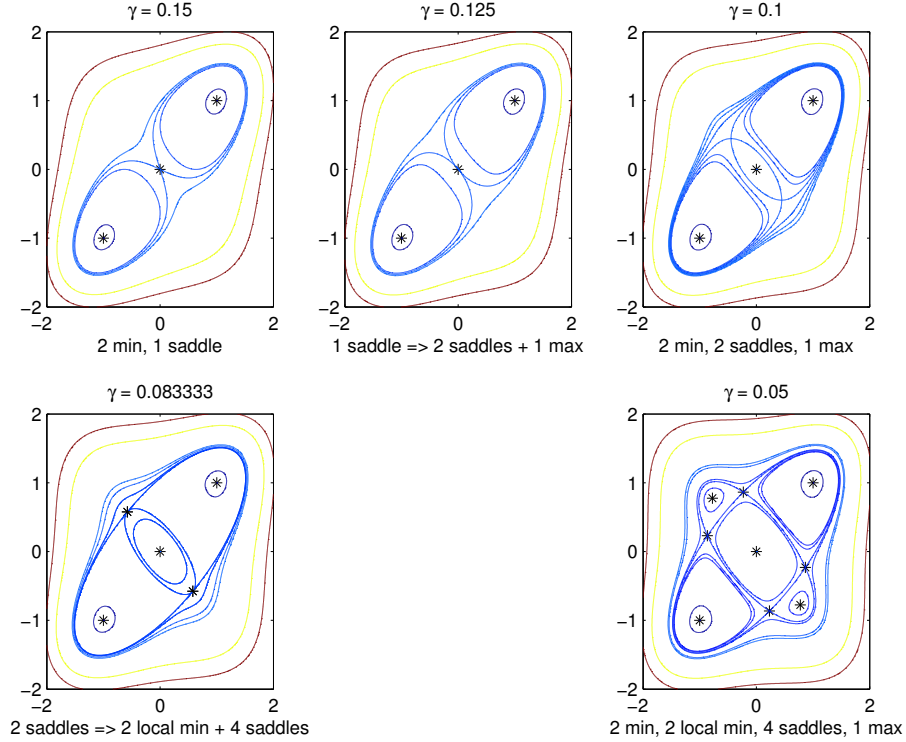

FIGURE 1. Bifurcations for different values of γ. Plot of $\mathcal{E}_\gamma^2$.

For $N = 4$ and $x = (x_1, x_2, x_3, x_4)$, the Jacobian matrix of $\mathcal{E}_\gamma^4(x)$ is

$$\nabla\mathcal{E}_\gamma^4(x) = \begin{pmatrix} \frac{1}{4}\left(x_1^3 - x_1\right) - 4\gamma(x_2 - x_1) \\ \frac{1}{4}\left(x_2^3 - x_2\right) - 4\gamma(x_3 - 2x_2 + x_1) \\ \frac{1}{4}\left(x_3^3 - x_3\right) - 4\gamma(x_4 - 2x_3 + x_2) \\ \frac{1}{4}\left(x_4^3 - x_4\right) - 4\gamma(x_3 - x_4) \end{pmatrix}$$

and its Hessian

$$\mathcal{H}\mathcal{E}_\gamma^4(x) = \begin{pmatrix} \frac{3}{4}x_1^2 - \frac{1}{4} + 4\gamma & -4\gamma & 0 & 0 \\ -4\gamma & \frac{3}{4}x_2^2 - \frac{1}{4} + 8\gamma & -4\gamma & 0 \\ 0 & -4\gamma & \frac{3}{4}x_3^2 - \frac{1}{4} + 8\gamma & -4\gamma \\ 0 & 0 & -4\gamma & \frac{3}{4}x_4^2 - \frac{1}{4} + 4\gamma \end{pmatrix}$$

Obvious critical points are $0, \pm 1$. Direct computations are more complex than in two dimensions and we make use of the normal forms to study bifurcations. Let us first note that the spectrum of $\mathcal{HE}_\gamma^4(0)$ is

$$Sp = \left\{ -\frac{1}{4}, -\frac{1}{4} + 8\gamma - 4\sqrt{2}\gamma, -\frac{1}{4} + 8\gamma, -\frac{1}{4} + 8\gamma + 4\sqrt{2}\gamma \right\}.$$

The associated eigenvectors are respectively

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1+\sqrt{2} \\ 1-\sqrt{2} \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1-\sqrt{2} \\ 1+\sqrt{2} \\ -1 \end{pmatrix}$$

It is obvious that the saddle point 0 degenerates into different types of saddle points at

$$\gamma = \left\{ \frac{1}{16(2-\sqrt{2})}, \frac{1}{32} \right\}$$

and finally to a local maximum at $\gamma = 1/(16(2+\sqrt{2}))$. A full bifurcation diagram for $N = 4$ can be found in [2] but for a different energy. In our case, bifurcations do not appear for the same values of γ but this diagram gives a good insight of what may happen in higher dimension than 2. We study bifurcations in the
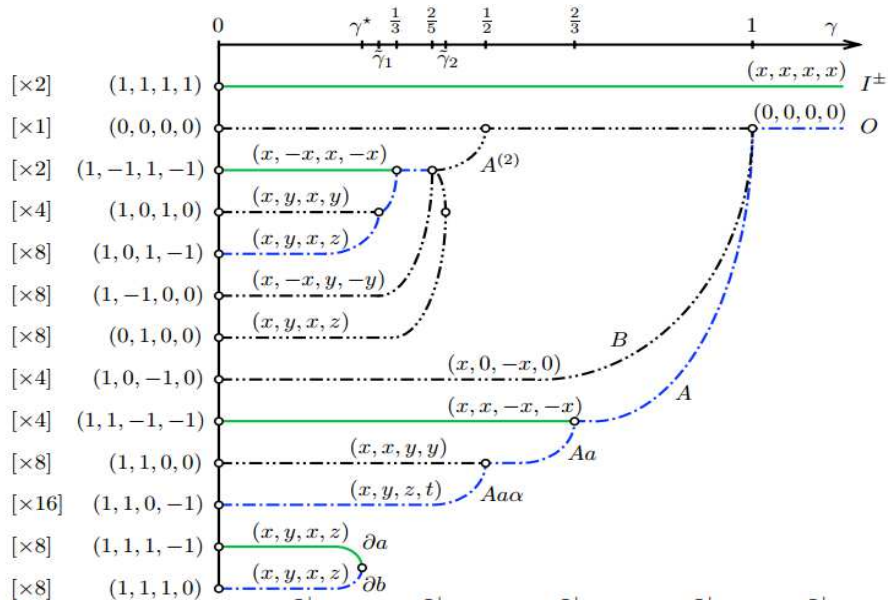

FIGURE 2. From [2]. Bifurcations for different values of γ.

orthogonal of the eigenvector corresponding to the biggest eigenvalues. Accordingly we consider $x = Ae_1 + \rho'$, where $e_1 = (1,1,1,1)$ and ρ' belongs to the orthogonal of e_1. Obviously this situation is more complex than in dimension 2 since the orthogonal space is now of dimension three. Let us first consider the case where $\rho' = \rho(1, -1+\sqrt{2}, 1-\sqrt{2}, -1)^t$. Then,

$$\mathcal{E}_\gamma^4(x) = \frac{1}{4}A^4 - \frac{1}{2}A^2 + \left(\frac{9}{4} - \frac{3}{2}\sqrt{2} \right)\rho^4 + \left(-1 + \frac{\sqrt{2}}{2} + 48\gamma - 32\sqrt{2}\gamma \right)\rho^2 + \left(3 - \frac{3}{2}\sqrt{2} \right)A^2\rho^2.$$

Denoting $\mathcal{G}_1(\rho, A) = \left(\frac{9}{4} - \frac{3}{2}\sqrt{2} \right)\rho^4 + \left(-1 + \frac{\sqrt{2}}{2} + 48\gamma - 32\sqrt{2}\gamma \right)\rho^2 + \left(3 - \frac{3}{2}\sqrt{2} \right)A^2\rho^2$, the critical points are given by 0 and $\pm\frac{1}{3}\sqrt{6 + 3\sqrt{2} - 96\gamma - 18A^2 - 9\sqrt{2}A^2}$. If $\gamma \geq \frac{1}{16(2-\sqrt{2})}$, the expression under the square root

376

is nonpositive. Thus there is only one critical point which is a minimum. Therefore the normal form is given by $G(A) = \frac{1}{4}A^4 - \frac{1}{2}A^2$. It is locally concave around 0 which is a local maximum.

If $\gamma < \frac{1}{16(2-\sqrt{2})}$ the normal form is given by

$$G(A) = \begin{cases} -\frac{5}{4}A^4 + (\frac{1}{2} + 16\sqrt{2}\gamma - 32\gamma)A^2 - \frac{1}{6} + \frac{32-16\sqrt{2}}{3}\gamma - (256 + \frac{512}{3}\sqrt{2})\gamma^2 & \text{if } |A| < \frac{1}{3}\sqrt{3 - 96\gamma + 48\sqrt{2}\gamma}, \\ \frac{1}{4}A^4 - \frac{1}{2}A^2 & \text{otherwise.} \end{cases}$$

Therefore $G'(A) = -5A^3 + (1 + 32\sqrt{2}\gamma - 64\gamma)A$. For a small perturbation $\gamma_\epsilon = \frac{1}{16(2-\sqrt{2})} - \epsilon$, $0 < \epsilon \ll 1$, the point 0 is a local maximum and the only critical point. For smaller γ there is two other critical points near 0 which are local minima. Thanks to the use of normal forms we are able to partly describe bifurcations. Much analysis has to be done in other directions considering linear combinations of orthogonal vectors.

In the case $N = +\infty$, such a discussion is possible, in order to study the nature of the critical point x_∞^0 (see [3]).

This discussion is important from a numerical point of view. Indeed, each bifurcation will introduce new minima that one shoud try to avoid in order to sample faster the required transition. Conversely, computations will be longer because trajectories may be "trapped" around those minima, for positive but small values of the temperature parameter. The numerical simulations of Section 5 will provide observations of these phenomena.

4. ADAPTIVE MULTILEVEL SPLITTING

The goal of splitting methods is to simulate n replicas (copies) of reactive trajectories. Let us recall that reactive trajectories are defined by the conditional distribution (4), the distribution of the SDE or SPDE dynamics (1)-(12) conditioned by the event $\{\tau_B < \tau_A\}$.

Loosely speaking, the small probability of the rare event is enforced using a birth-death mechanism as follows:

- Launch multiple replicas subject to the reference dynamics.
- Kill the replicas having the smallest maximal level, as given by the reaction coordinate (fitness) continuous mapping $\xi : \mathbb{R}^N$ or $H \to \mathbb{R}$.
- Replicate the other replicas.

In the literature, the following cases have been studied:

- $N = 1$, this is the classical AMS algorithm proposed in [7];
- $N > 1$, multiple replica algorithm based on AMS, studied in [8].

The infinite dimensional case $N = \infty$ seems to be treated for the first time in the present work.

We will use the following *reaction coordinate* (fitness mapping) defined as the *average position or magnetization* $\xi : L^2(0,1) \to \mathbb{R}$:

$$\xi(x) := \int_0^1 x(\lambda) \, d\lambda \in \mathbb{R}.$$

For the finite dimensional model, a discrete version of the magnetization is used, based on a quadrature formula for the integral. For instance,

$$\xi^N(x^N) := \frac{1}{N} \sum_{i=1}^N x_i^N \in \mathbb{R}.$$

Note that this choice is not optimal. In Section 3, we described the influence of the parameter γ on the reactive trajectories; in principle, based on this theoretical knowledge (in this special situation), one could build a suitable reaction coordinate for each value of γ. There is no general answer for this problem, and in the sequel we only use the magnetization.

We will present the algorithm in *discrete time*, with time index $k \in \mathbb{N}$, so that it can directly be applied to the numerical discretization.

Consider now n_{rep} replicas of the system dynamics. We thus consider the *(stopping) hitting time* of *open levels* of ξ:

$$t_z(X) := \inf \{ k \geq 0 \, | \, \xi(X_k) > z \} \, .$$

The *rare event* of interest is then rewritten as

$$\{ t_{z_B}(X) \leq \tau_A(X) \} \, ,$$

and we wish to compute its associated *small probability*. Recall that $\tau_A(X) = \inf \{ k \geq 0 \, | \, X_k \in A \}$. If we take $A = \{ x; \xi(x) < z_A \}$, we see that the $\tau_A(X) = \inf \{ k \geq 0 \, | \, \xi(X_k) < z_A \}$.

Typically for the Allen-Cahn equation, $\xi(X_0) \in (z_A, z_B)$, $z_A = -1 + \delta$, $z_B = 1 - \delta$, with some $\delta > 0$. Since we have

$$\xi(x_\infty^-) = -1, \quad \xi(x_\infty^0) = 0, \quad \xi(x_\infty^+) = +1,$$

the magnetization gives some useful piece of information along the transition from one metastable state to the other.

In principle the algorithm can be generalized to *any continuous* $\xi : E \to \mathbb{R}$ on a Polish state space E.

First, let us fix

$$\mathcal{T} \subset \mathbb{N},$$

the set of time indices for which the replica states will be kept in memory. We then denote by

$$k \mapsto X_k^{(n,q)}, \quad n = 1 \dots n_{\text{rep}}$$

indexed by $q \geq 1$ the iteration index of the algorithm, which is different from the *time index* $k \geq 0$.

For $q = 1$ (initial condition), $(k \mapsto X_k^{(n,1)})_{1 \leq n \leq N}$ are iid and stopped at

$$\min \left(\tau_A(X^{(n,1)}), t_{z_B}(X^{(n,1)}) \right),$$

i.e. when the magnetization is either below level z_A or above z_B.

Then iterate on $q \geq 1$ as follows:

(1) q being given, compute:

$$\begin{cases} \text{Max}^{(n,q)} := \max_{k \geq 0} \xi(X_k^{(n,q)}) \\ N_q^{\text{killed}} := \underset{n}{\text{argmin}} \, \text{Max}^{(n,q)} \in [1, N], \end{cases}$$

the replica N_q^{killed} (we assume that it is unique) with minimal maximal ("min-max") level $\text{Max}^{(N_q^{\text{killed}}, q)}$.

(2) Fix $m \in \mathbb{N}_*$ some . Choose N_q^{new} uniformly in $\{1, \dots, N\} - \{N_q^{\text{killed}}\}$ and consider the time (for a small $\delta > 0$):

$$\tau^{(q)} := \inf \left\{ k \in \mathcal{T} | t_{\text{Max}^{(N_q^{\text{killed}}, q)}}(X^{(N_q^{\text{new}}, q)}) \geq k \right\},$$

the first time when the branching replica N_q^{new} has reached the maximum level $\text{Max}^{(N_q^{\text{killed}}, q)}$ of all the killed replicas.

(3) Kill the information of replica N_q^{killed}. Copy the path of replica N_q^{new} until $\tau^{(q)}$ and then re-sample the remaining path with Markov dynamics until either level z_A or z_B is reached.

(4) Stop at $q = Q_{\text{iter}}$ when all replicas have reached level z_B.

Then the general principle of the AMS algorithm can be stated as follows. Assume $n_{\text{rep}} \to +\infty$. Then:

(1) A path of a replica at the end of the algorithm is distributed according to the "reactive trajectory" (4):

$$\text{Law} \left(X_t, \, t \geq 0 \, | \, t_{z_B}(X) < \tau_A(X) \right)$$

(2) The quantity

$$(1 - 1/n_{\text{rep}})^{Q_{\text{iter}}}$$

is a convergent estimator of $\mathbb{P}(t_{z_B}(X) < \tau_A(X))$.

The algorithm presented above corresponds to the way it has been described and implemented in [8]. There also exists a variant where $k_{rep} > 1$ replicas are killed and resampled at each iteration; an appropriate new estimator has to be defined, see [7]. Those estimators are consistent, in the sense that they converge to the transition probability $\mathbb{P}(t_{z_B}(X) < \tau_A(X))$ when the number of replicas n_{rep} tends to $+\infty$.

In fact, in a work in preparation [6], we will generalize the AMS algorithm in order to define *unbiased* estimators of the transition probability. In fact, as soon as discrete-time Markov processes are considered (as for instance when we discretize a SDE/SPDE with an Euler scheme), we can observe that the estimator that we used here is not unbiased. However, the bias tends to 0 when the number of replicas goes to $+\infty$ as well as when the time-step Δt goes to 0.

The required modifications to define the Generalized Adaptive Multilevel Splitting Algorithm are slight but technical. For simplicity, we do not discuss them here and refer to the forthcoming paper [6].

5. NUMERICAL RESULTS

We now want to show the performance of the algorithm introduced in the previous section in two directions: the estimation of the transition probability and the approximation of the reactive trajectories when γ varies.

5.1. **Estimation of the probability**

With the notations of the previous section, we estimate the probability $\mathbb{P}(\tau_B(X) \leq \tau_A(X))$ where the sets A and B correspond to the levels $z_A = -0.99$, $z_B = 0.99$. The Allen-Cahn equation is discretized thanks to the scheme (16) with a time-step $\Delta t = 0.01$, and a finite difference discretization with a regular grid with mesh size $\Delta x = 0.02$. The parameter γ is equal to 1 and we consider the initial condition $X_0(\lambda) = -0.8$.

We perform $N_{MC} = 100$ realizations of the algorithm to compute a Monte-Carlo approximation of the expectation of the estimator. More precisely, we consider different choices of $n_{\text{rep}} \in \{50, 100, 200, 1000\}$ and $\epsilon = 0.05$, so that the probability of the transition is approximately 0.005. In Table 5.1 below, we give an estimated probability, and an empirical standard deviation computed with the Monte-Carlo approximation.

Confidence intervals are obtained when this standard deviation is divided by the square root of the number of realizations: in Table 5.1 below the empirical standard deviation should then be divided by 10.

n_{rep}	estimated probability	empirical standard deviation
50	0.00516	0.00163
100	0.00514	0.00128
200	0.00502	0.000774
1000	0.00501	0.000350

TABLE 1. Estimated probability and empirical standard deviation obtained via the AMS algorithm.

We observed that the precision is improved when we n_{rep} increases. Another useful comparison is given in Table 5.1, where we compare the results for $(n_{\text{rep}} = 100, N_{MC} = 1000)$ and $(n_{\text{rep}} = 1000, N_{MC} = 100)$.

n_{rep}	N_{MC}	estimated probability	empirical standard deviation	time for one realization
100	1000	0.00502	0.00166	67 s
1000	100	0.00501	0.000350	676 s

TABLE 2. Computational time and variance for two values of n_{rep}.

In both cases we obtain the same precision, with approximately the same required computational time. Two arguments are then in favor of choosing the smallest n_{rep} in this situation: first, the estimator is unbiased for every value of n_{rep}, so that taking n_{rep} very large is not necessary; second, it is easy to save computational time with a parallelization of the Monte-Carlo procedure. Parallelization inside the AMS algorithm could also help, and further research is necessary in this direction.

Finally, we compare the performance of the AMS algorithm with a direct Monte-Carlo procedure: we run independent trajectories solving the Allen-Cahn equation, and count 1 if $\tau_B < \tau_A$, 0 if $\tau_A < \tau_B$, and

average over the realizations. The computation of one trajectory only takes about 0.4 s, and if we run 10^6 independent replicas, the Monte-Carlo procedure gives an estimated probability 0.00507, with an empirical standard deviation 7.12×10^{-5}. Compared to the result with $n_{\text{rep}} = 200$ in Table 5.1, we see that for the same precision the AMS algorithm is between 3 and 4 times faster than a direct Monte-Carlo method, for the approximation of a probability of order 5.10^{-3}. Moreover, the smaller the probability becomes, the better the AMS algorithm should be.

Now we study the dependence of the transition probability with respect to the temperature parameter ϵ. For each value of ϵ, we choose the same discretization parameters i.e. $\Delta t = 0.01, \Delta x = 0.02$ and $\gamma = 2$. Moreover, we take $n_{\text{rep}} = 100$ as well as $N_{MC} = 100$ realizations in order to compute an empirical probability with a Monte-Carlo procedure. This number of realizations is small but gives sufficient information for the qualitative behavior of the method. Results are given with the empirical standard deviation of the estimator in Table 5.1:

ϵ	estimated probability	empirical standard deviation
0.30	0.0831	$1.33 \ 10^{-2}$
0.10	0.0276	$5.11 \ 10^{-3}$
0.07	0.0131	$2.81 \ 10^{-3}$
0.05	0.00398	$9.87 \ 10^{-4}$
0.04	0.00143	$4.06 \ 10^{-4}$
0.03	0.000234	$6.17 \ 10^{-5}$

TABLE 3. Dependence of the transition probability with respect to ϵ.

In the following Figure 3 is plotted the logarithm of $\mathbb{P}(\{\tau_B < \tau_A\})$ with respect to $1/\epsilon$ for the values indexed in the above table. We obtain a straight line showing the exponential decrease of $\log(\mathbb{P}(\{\tau_B < \tau_A\}))$ with respect to $1/\epsilon$.
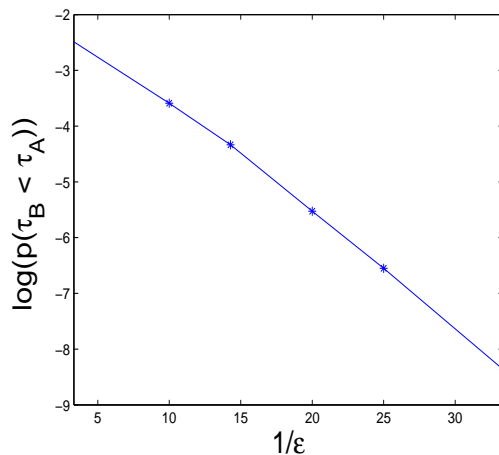

FIGURE 3. Plot of $\log(\mathbb{P}(\{\tau_B < \tau_A\}))$ with respect to $1/\epsilon$.

5.2. Reactive trajectories in dimension 2

In this paragraph, we investigate numerically the qualitative behavior of the reactive trajectories in the AMS algorithm in terms of the parameter γ. As explained in the Section 3, the potential changes with γ and then the reactive trajectories may experiment different paths to go from A to B. At the end of the algorithm we obtain n_{rep} trajectories of the process starting at the same position, and with the property that B is reached before A.

We consider different values for $\gamma \in \{1/4, 1/8, 1/16, 1/32\}$. For each, we have used the algorithm for two different values of the number of replica n_{rep} and of the temperature ϵ: either $(n_{\text{rep}} = 100, \epsilon = 1.10^{-4})$ or $(n_{\text{rep}} = 1000, \epsilon = 5.10^{-4})$. Computations take a few minutes on a personal computer. Figure 4 displays examples of reactive trajectories in dimension $N = 2$ obtained with $(n_{\text{rep}} = 100, \epsilon = 1.10^{-4})$. To represent the qualitative behavior of the approximate reactive trajectories, we also draw histograms of the position in the line $x + y = 0$ when reactive trajectories are crossing it. With $n_{\text{rep}} = 1000$ we obtain almost symmetric histograms of Figure 5.

We recover the expected behavior with respect to γ: first, $(0,0)$ is the unique saddle point; then two better saddle points appear on each side of the line $x + y = 0$; finally two local minima become places where trajectories are trapped during a long time - the importance of this effect on trajectories should decrease when temperature decreases, with the price of more iterations in the algorithm. Another choice of reaction coordinate might improve the performance of the algorithm in this situation; anyway this question is highly problem-dependent.
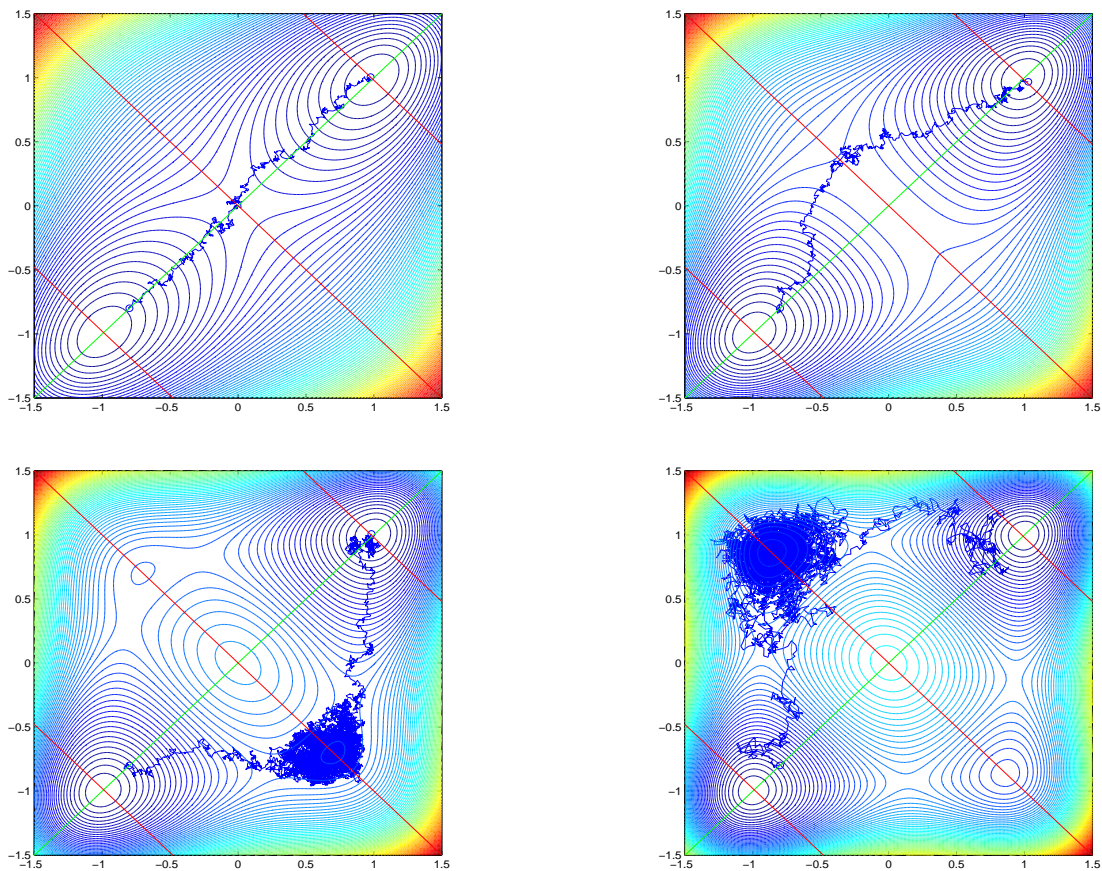

FIGURE 4. Reactive trajectories for the two particles system and for different values of $\gamma = 1/4, 1/8, 1/16, 1/32$.

Notice that since we only kill $k_{\text{rep}} = 1$ replica at each iteration of the algorithm, the time to either reach A or B can be very long even after several steps, due to the presence of additional local minima for instance. This observation is the origin for investigating parallelization strategies, to reduce computational time. Taking $k_{\text{rep}} > 1$ is one natural answer to this problem, but possibly not the only one.
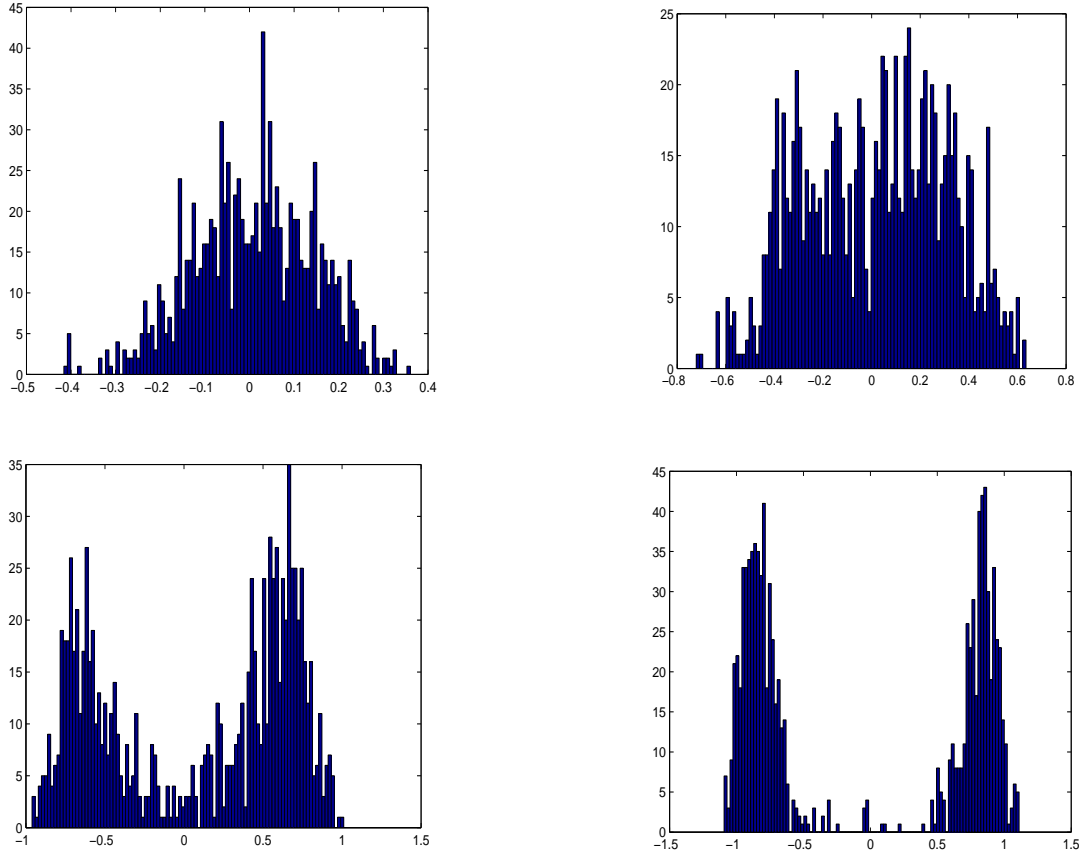
FIGURE 5. Histograms for the two particles system in the $(x, -x)$ line and for different values of $\gamma = 1/4, 1/8, 1/16, 1/32$.

5.3. **Reactive trajectories for the Allen-Cahn equations**

We give a few examples of trajectories obtained in the AMS algorithm for the Allen-Cahn equation. It is discretized as before with $\Delta t = 0.01$ and $\Delta x = 0.02$; we use $n_{\text{rep}} = 100$ and $\epsilon = 0.05$. The initial conditions is given by $x(\lambda) = -0.8$. In Figure 6 are given some examples of reactive trajectories for different values of $\gamma = 1$ and $\gamma = 0.1$.

6. CONCLUSION

In this proceeding, we investigated the numerical estimations of rare events in infinite dimensions. We proposed an efficient algorithm outperforming the standard Monte Carlo simulations and consisting on a generalization of splitting algorithms in finite dimension. Beyond the fact that very small probabilities can be reached, reactive trajectories can also be computed. These paths are known for (8) and even (5), but it can be helpful in more complex situations. Finally, the great advantage of this algorithm is to be highly parallelizable. However, a lot of work remains to be done. In a forthcoming paper, we will study the unbiased property of the estimator for $k_{\text{rep}} \geqslant 1$. Based on this analysis, we wish to develop a parallel version of this algorithm whose efficiency will depend on a ratio between the number of processors and the number of killed replicas k_{rep}.
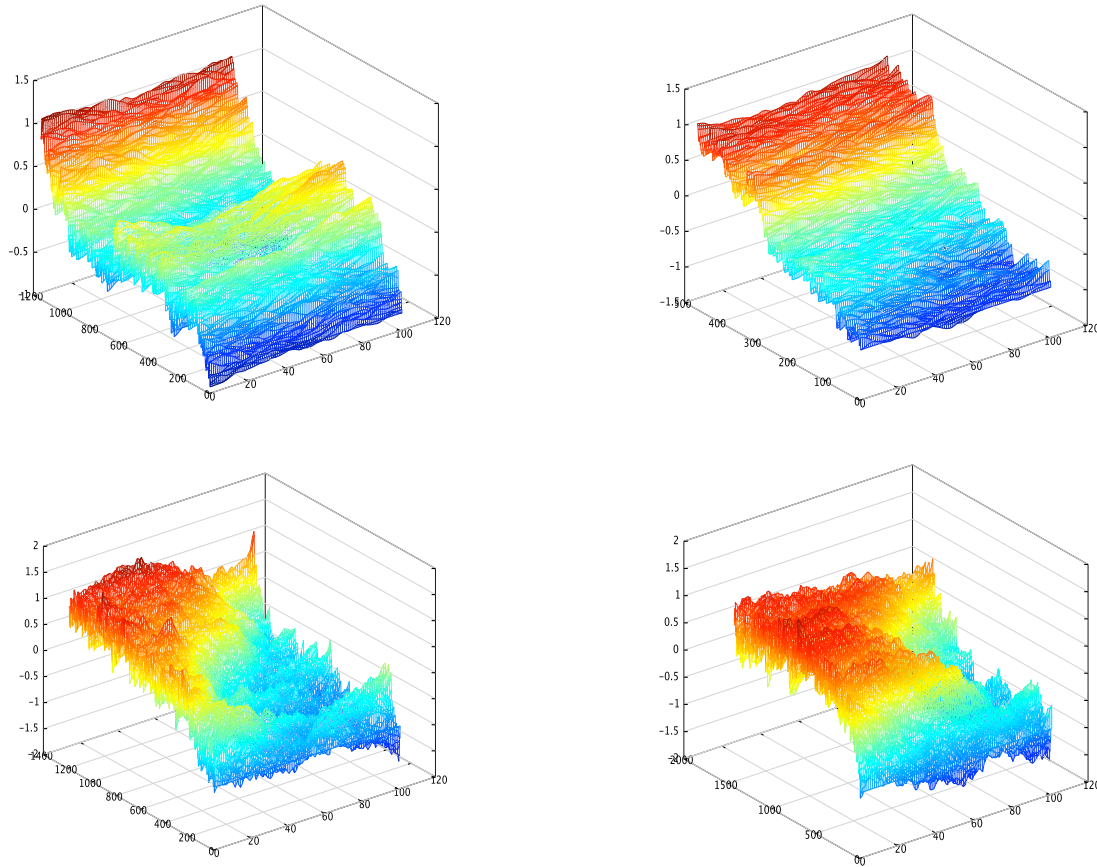
FIGURE 6. Reactive trajectories for the Allen-Cahn equation for $\gamma = 1$ and $\gamma = 0.1$.

discussions with T. Lelièvre and D. Aristoff. Finally, they thank D. Iampietro for its participation to the project.

REFERENCES

[1] F. Barret. Sharp asymptotics of metastable transition times for one dimensional spdes. *arXiv:1201.4440*, 2012.

[2] N. Berglund, B. Fernandez, and B. Gentz. Metastability in Interacting Nonlinear Stochastic Differential Equations I: From Weak Coupling to Synchronisation. *ArXiv Mathematics e-prints*, November 2006.

[3] N. Berglund and B. Gentz. Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers' law and beyond. *Electron. J. Probab.*, 18:no. 24, 58, 2013.

[4] F. Bouchet and E. Simonnet. Random changes of flow topology in two-dimensional and geophysical turbulence. *Phys. Rev. Lett.*, 102:094504, Mar 2009.

[5] A. Bovier. Metastability. In *Methods of contemporary mathematical statistical physics*, volume 1970 of *Lecture Notes in Math.*, pages 177–221. Springer, Berlin, 2009.

[6] C.-E. Bréhier, M. Gazeau, L. Goudenège, T. Lelièvre, and M. Rousset. Unbiasedness of some generalized adaptive multilevel splitting algorithms. in preparation.

[7] F. Cérou and A. Guyader. Adaptive multilevel splitting for rare event analysis. *Stochastic Analysis and Applications*, 25(2):417–443, 2007.

[8] F. Cérou, A. Guyader, T. Lelièvre, and D. Pommier. A multiple replica approach to simulate reactive trajectories. *Journal of Chemical Physics*, 134(5), 2011.

[9] F. Chenal and A. Millet. Uniform large deviations for parabolic SPDEs and applications. *Stochastic Process. Appl.*, 72(2):161–186, 1997.

[10] M. Corvellec. *Transitions de phase en turbulence bidimensionnelle et géophysique*. PhD thesis, Ecole normale supérieure de lyon-ENS LYON, 2012.

[11] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 44. Cambridge University Press, In Encyclopedia of Mathematics and Its Applications, 1992.

[12] M.I. Freidlin and A.D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, third edition, 2012. Translated from the 1979 Russian original by Joseph Szücs.

[13] F. Otto, H. Weber, and M. Westdickenberg. Invariant measure of the stochastic allen-cahn equation: the regime of small noise and large system size. *arXiv preprint arXiv:1301.0408*, 2013.

[14] G. Strang. On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.*, 5:506–517, 1968.

[15] H. Weber. Sharp interface limit for invariant measures of a stochastic allen-cahn equation. *Communications on Pure and Applied Mathematics*, 63(8):1071–1109, 3010.

[16] H. Yoshida. Construction of higher order symplectic integrators. *Physics Letters A*, 150(5):262–268, 1990.