



University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): F. J. Pinski, A. M. Stuart and F. Theil

Article Title:  $\Gamma$ -Limit for Transition Paths of Maximal Probability

Year of publication: 2011

Link to published article:

<http://link.aps.org/doi/10.1007/s10955-012-0443-8>

Publisher statement: The original publication is available at [www.springerlink.com](http://www.springerlink.com)

# $\Gamma$ –Limit for Transition Paths of Maximal Probability

F.J. Pinski \*

Physics Department  
University of Cincinnati  
PO Box 210011  
Cincinnati OH 45221, USA

and

A.M. Stuart and F. Theil †  
Mathematics Institute  
Warwick University  
Coventry CV4 7AL, UK

March 11, 2011

## Abstract

Chemical reactions can be modeled via diffusion processes conditioned to make a transition between specified molecular configurations representing the state of the system before and after the chemical reaction. In particular the model of Brownian dynamics – gradient flow subject to additive noise – is frequently used. If the chemical reaction is specified to take place on a given time interval, then the most likely path taken by the system is a minimizer of the Onsager-Machlup functional. The  $\Gamma$ –limit of this functional is determined explicitly in the case where the temperature is small and the transition time scales as the inverse temperature.

## 1 Introduction

In this paper we study the problem of determining the most likely paths taken by a system undergoing a chemical reaction. We employ the model known as *Brownian dynamics* [1]: the atomic positions are assumed to be governed by a gradient flow in an interaction potential, subject to additive thermal noise. The resulting

---

\*E-mail address: frank.pinski@uc.edu

†E-mail address: {a.m.stuart,f.theil}@warwick.ac.uk.

stochastic differential equation is conditioned to make a transition between two different atomic configurations representing the state of the system before and after the chemical reaction [2]. If the chemical reaction is specified to take place on a given time interval, then the most likely path taken by the system is a minimizer of the Onsager-Machlup functional [3, 4]. Minimizers of the Onsager-Machlup functional have received considerable attention in the chemistry literature (see [7] and [5] and the references therein) and the work in this paper provides a mathematically rigorous explanation for what is observed in these computational studies. In [5] numerical computations are presented which study minimizers of the Onsager-Machlup functional for a variety of model problems in low dimension, together with some higher dimensional problems from physics and chemistry such as vacancy diffusion and the Lennard-Jones cluster of 38 atoms. The minimizers exhibit a number of interesting effects, including multiple pathways between configurations, together with transition paths which concentrate at saddle points of the potential, and not at minima. The computations in [5] were performed at low temperature over fixed long intervals. To gain mathematical insight into these computations, the work in this paper is concerned with determining the  $\Gamma$ -limit (see [6, 17]) of the Onsager-Machlup functional in the case where the temperature is small and the transition time scales as the inverse temperature. The explicit  $\Gamma$ -limit which we exhibit may also be of independent interest. It can be optimized via knowledge of the critical points of the potential, the trace of the Hessian of the potential at the critical points, and certain heteroclinic orbits connecting them. Notably these heteroclinic orbits are in a Hamiltonian system; heteroclinic orbits in the forward or backward gradient flow found from our model at zero temperature are solutions of this Hamiltonian problem but not necessarily vice versa.

Regarding the role of the Hessian in the  $\Gamma$ -limit, it is pertinent to mention the paper [8] in which second derivative of the potential plays a role in a large deviations principle for SDEs.

In section 2 we provide a precise mathematical description of the conditioned SDE which forms our mathematical model. We provide an informal derivation of the Onsager-Machlup functional whose minimizers determine most likely transition paths and we explain the sense in which this informal argument can be made rigorous. In section 3 we compute the  $\Gamma$ -limit of the Onsager-Machlup functional building on related analysis in [9, 10]. The paper concludes, in section 4, with an informal characterization of the  $\Gamma$ -limit, together with numerical experiments which illustrate this characterization.

## 2 Set-Up

Consider the following conditioned SDE for  $x \in C([0, T]; \mathbb{R}^N)$  making a transition between two states  $x^-$  and  $x^+$  in time  $T$ :

$$\begin{aligned} dx &= -DV(x)dt + \sqrt{2\varepsilon} dW, \\ x(0) &= x^- \quad \text{and} \quad x(T) = x^+. \end{aligned} \tag{2.1}$$

Here  $V : \mathbb{R}^N \rightarrow \mathbb{R}$  is the potential,  $W$  is a standard Brownian motion in  $\mathbb{R}^N$  and  $\varepsilon \ll 1$ . In many applications in physics/chemistry  $N = nd$ ,  $d \in \{1, 2, 3\}$  being the physical dimension and  $n$  the number of particles. Then  $\varepsilon$  is the (non-dimensional) temperature and  $x \in C([0, T]; \mathbb{R}^N)$  denotes the configurational path

of the  $n$  atoms making a transition between  $x^\pm$ . Throughout we assume that  $x^\pm$  are chosen as critical points of  $V$ . We will be particularly interested in choosing these critical points to be minima, so that the problem (2.1) describes a chemical reaction; but, in the numerical computations, we will also choose saddle points in order to illustrate certain mathematical phenomena.

Escape from local minima occurs on timescales which are exponentially long in  $\varepsilon^{-1}$ , whilst transitions themselves occur on scales which are logarithmic in  $\varepsilon^{-1}$ . We focus interest on an intermediate timescale between these two regimes, which is  $\varepsilon^{-1}$ . This timescale is hence long enough to capture a single transition, but not long enough to capture the typical waiting time in a potential minimum. In particular, transition paths calculated in this scaling which pass through an intermediate minimum  $x^0$  of  $V$  will not necessarily exhibit the “typical” behavior that would be exhibited if the time to make the transition from  $x^-$  to  $x^+$  via  $x^0$  were left free. The behavior of the most likely single transitions in this regime is captured by minimizers of the Freidlin-Wentzell action [11, 12]. Numerical methods to capture multiple such transitions are the subject of active study, and the reader should consult [13], and the references therein, for details. From an applied perspective, our work serves to highlight the potential pitfalls of using the Onsager-Machlup approach to compute transition paths, and this point is discussed in detail in [5]. As mentioned above, the work in this paper gives a mathematical explanation for the numerical results observed in [5].

The results may also be of independent interest from the point of view of the calculus of variations. In particular the results show how heteroclinic orbits in a certain Hamiltonian flow form the building block for the construction of the  $\Gamma$ -limit. This fact may also be of interest in applications where attention has focused on heteroclinic orbits in the forward or backward gradient flow found from (2.1) with  $\varepsilon = 0$ ; these form particular instances of heteroclinic orbits in the Hamiltonian flow, but non-gradient connections are also possible, as we will demonstrate.

To enforce the scaling of interest we choose  $T = \varepsilon^{-1}$  and rescale time as  $t = \varepsilon^{-1}s$ , to obtain

$$\begin{aligned} dx &= -\frac{1}{\varepsilon}DV(x)ds + \sqrt{2}dW, \\ x(0) &= x^- \quad \text{and} \quad x(1) = x^+. \end{aligned} \tag{2.2}$$

In this scaling we are studying transitions on a unit time-interval, in which the systematic motion of the molecules is large and the thermal noise is of order 1.

Conditioning the diffusion process to hit a point at time 1 is of course a measure zero event with respect to the Brownian motion. However the regular conditional probability measure  $\pi$  governing the stochastic boundary value problem (2.2) exists and may be characterized via its density with respect to the Brownian bridge measure  $\pi_0$  arising in the case  $V \equiv 0$ . The density relating the two measures is found from the Girsanov formula, together with an integration by parts (use of Itô formula) and use of the boundary conditions on the path [14, 15], and is given by

$$\frac{d\pi}{d\pi_0}(x) \propto \exp\left(-\frac{1}{2\varepsilon^2} \int_0^1 G(x; \varepsilon) dt\right) \tag{2.3}$$

where the *path potential*  $G$  is (with  $|\cdot|$  denoting the Euclidean norm)

$$G(x; \varepsilon) = \frac{1}{2}|DV(x)|^2 - \varepsilon\Delta V(x). \tag{2.4}$$

Functions are infinite dimensional and there is no Lebesgue measure in infinite dimensions. Nonetheless, it is instructive to think heuristically of the Brownian bridge obtained from (2.2) with  $V \equiv 0$  as having a probability density (with respect to Lebesgue measure) of the form [16]

$$\exp\left(-\frac{1}{4}\int_0^1\left|\frac{dx}{ds}\right|^2 ds\right). \quad (2.5)$$

(This is the formal limit obtained from the probability density function for a discretized Brownian bridge). Let  $H_{\pm}^1((0, 1))$  denote the subset of  $H^1((0, 1); \mathbb{R}^N)$  comprised of functions satisfying  $x(0) = x^-$  and  $x(1) = x^+$ . Then, combining (2.3) and (2.5), we may think of the probability density for  $\pi$  as being proportional to  $\exp(-\frac{1}{2\varepsilon}I_{\varepsilon}(x))$  where the Onsager-Machlup functional  $I_{\varepsilon} : H_{\pm}^1((0, 1)) \rightarrow \mathbb{R}$  is defined by

$$I_{\varepsilon}(x) := \int_0^1\left(\frac{\varepsilon}{2}\left|\frac{dx}{ds}\right|^2 + \frac{1}{\varepsilon}G(x; \varepsilon)\right)ds. \quad (2.6)$$

This intuitive definition of  $I_{\varepsilon}(x)$ , via the logarithm of the pathspace probability density function, suggests that minimizers of  $I_{\varepsilon}$  are related to paths of maximal probability. This can be formulated precisely as follows [3, 4]. Let  $B^{\delta}(z)$  denote a ball of radius  $\delta$  in  $C([0, 1]; \mathbb{R}^N)$  centered on  $z \in H_{\pm}^1((0, 1))$ . Then, for any  $z_1, z_2 \in H_{\pm}^1((0, 1))$ ,

$$\lim_{\delta \rightarrow 0} \frac{\pi(B^{\delta}(z_1))}{\pi(B^{\delta}(z_2))} = \exp\left(\frac{1}{2\varepsilon}(I_{\varepsilon}(z_2) - I_{\varepsilon}(z_1))\right).$$

Thus, for small ball radius  $\delta$ , the logarithm of the ratio of the probabilities of the two balls is equal to the difference in  $I_{\varepsilon}$  evaluated at the ball centers. For this reason we are interested in minimizers of  $I_{\varepsilon}$ .

The following notation will be useful. We define

$$\begin{aligned} J_{\varepsilon}(x) &= \int_0^1\left(\frac{\varepsilon}{2}\left|\frac{dx}{ds}\right|^2 + \frac{1}{2\varepsilon}|DV(x)|^2\right)ds, \\ J(x) &= \int_{-\infty}^{\infty}\left(\frac{1}{2}\left|\frac{dx}{ds}\right|^2 + \frac{1}{2}|DV(x)|^2\right)ds. \end{aligned} \quad (2.7)$$

Then

$$I_{\varepsilon}(x) = J_{\varepsilon}(x) - \int_0^1 \Delta V(x(s)) ds. \quad (2.8)$$

Our goal in what follows is to demonstrate that the  $\Gamma$ -limit of  $I_{\varepsilon}$  is finite only when evaluated on the set of BV functions supported on the critical points of  $V$ . And, furthermore, that on this set the value of the  $\Gamma$ -limit is determined by (suitably rescaled) minima of  $J$ , subject to end-point conditions, together with the integral of  $\Delta V(x)$ . Thus we conclude this section with some observations concerning minima of  $J$ .

The Euler-Lagrange equations for the functional  $J$  are

$$\frac{d^2x}{ds^2} - D^2V(x)DV(x) = 0, \quad (2.9)$$

where  $D^2V$  denotes the Hessian of  $V$ . These equations are Hamiltonian and conserve the energy

$$E = \frac{1}{2} \left| \frac{dx}{ds} \right|^2 - \frac{1}{2} |DV(x)|^2.$$

Thus heteroclinic orbits connecting critical points of  $V$  via the equation (2.9), for which  $E$  is necessarily zero, satisfy

$$\left| \frac{dx}{ds} \right| = |DV(x)|. \quad (2.10)$$

Hence, if there is a heteroclinic orbit connecting critical points of  $V$  under either of the forward or backward gradient flows

$$\frac{dx}{ds} = \pm DV(x) \quad (2.11)$$

then this will also determine a heteroclinic orbit in the Hamiltonian system. However, the converse is not necessarily the case: there are heteroclinic orbits in the Hamiltonian flow which are not heteroclinic orbits in the gradient dynamics.

We say that a potential  $V : \mathbb{R}^N \rightarrow \mathbb{R}$  is *admissible* if  $V \in C^3(\mathbb{R}^N, \mathbb{R})$  and

1. the set of critical points

$$\mathcal{E} = \{x \in \mathbb{R}^N \mid DV(x) = 0\}$$

is finite;

2. the Hessian  $D^2V(x)$  has no zero eigenvalues for every  $x \in \mathcal{E}$ ;
3. the weak coercivity condition

$$\exists R > 0 \text{ such that } \inf_{|x| > R} |DV(x)| > 0 \quad (2.12)$$

is satisfied.

Admissibility implies, in particular, that all critical points for the gradient flows (2.11) are hyperbolic.

For each pair  $x^-, x^+ \in \mathcal{E}$  the set of *transition paths* is defined as

$$X(x^-, x^+) = \left\{ y \in BV(\mathbb{R}) \mid \lim_{t \rightarrow \pm\infty} y(t) = x^\pm \text{ and } \dot{y} \in L^2(\mathbb{R}) \right\}.$$

**Lemma 2.1.** For any  $x \in X(x^-, x^+)$

$$J(x) \geq |V(x^+) - V(x^-)|. \quad (2.13)$$

Furthermore, if the infimum of  $J$  over  $X(x^-, x^+)$  is attained at  $x^* \in X(x^-, x^+)$ , then  $x^* \in C^2(\mathbb{R})$  satisfies the Euler-Lagrange equation (2.9) and has the property

$$J(x^*) = \int_{-\infty}^{\infty} |DV(x^*(s))|^2 ds. \quad (2.14)$$

Assume that the potential is admissible. If either  $x^-$  or  $x^+$  is a local minimum or maximum and there exists a heteroclinic orbit  $x^*$  connecting  $x^-$  and  $x^+$  under the Hamiltonian dynamics (2.9), then  $x^*$  solves (2.11) and

$$J(x^*) = |V(x^+) - V(x^-)|. \quad (2.15)$$

In particular  $x^*$  is a minimizer of  $J$  which attains the lower bound (2.13).

*Proof.* Clearly

$$\begin{aligned} J(x) &\geq \int_{-\infty}^{\infty} \left| \frac{dx}{ds} \right| |DV(x)| ds \\ &\geq \int_{-\infty}^{\infty} \left| \frac{d}{ds} (V(x)) \right| ds \\ &\geq \left| \int_{-\infty}^{\infty} \frac{d}{ds} (V(x)) ds \right|. \end{aligned}$$

Integrating and using the end point conditions gives the first result. Standard regularity results imply that minimizers  $x$  of  $J$  are  $C^2(\mathbb{R})$  and satisfy the Euler-Lagrange equation (2.9). Equation (2.14) then follows from (2.10).

To prove the final result we first note that the linearization of equation (2.9) at a critical point  $x_c \in \mathcal{E}$  is given by

$$\frac{d^2 y}{ds^2} - D^2 V(x_c) y = 0, \quad (2.16)$$

whilst the forward gradient flow, given by (2.11) with a minus sign, has linearization

$$\frac{dy}{ds} + D^2 V(x_c) y = 0. \quad (2.17)$$

If  $V$  is admissible, and  $x_c$  is a critical point of  $V$ , then  $(x_c, 0)$  is hyperbolic for the Hamiltonian flow and, from (2.16), has stable and unstable manifolds both of dimension  $N$ .

We assume next that  $x^+$  is a local minimum of  $V$ . Let  $\mathcal{M}$  be the  $N$ -dimensional manifold defined by

$$\mathcal{M} = \{(x, -DV(x)) \mid x \in \mathbb{R}^N\} \subset \mathbb{R}^{2N}.$$

Differentiating equation (2.11) with respect to  $t$  shows that  $\mathcal{M}$  is invariant under the flow of equation (2.9). Since  $D^2V(x^+)$  is positive definite there exists  $\varepsilon > 0$  such that all solutions  $y(t)$  of equation (2.9) with the properties  $|y(0) - x^+| < \varepsilon$  and  $\dot{y}(0) = -DV(y(0))$  converge to  $x^+$  as  $t$  tends to infinity. Define next the stable manifold

$$\mathcal{M}(x^+) = \left\{ (x, v) \in \mathbb{R}^{2N} \mid \lim_{t \rightarrow \infty} y(t) = x^+ \right\},$$

where  $y$  solves (2.9) and satisfies the initial condition  $y(0) = x, \dot{y}(0) = v$ . Since the dimension of  $\mathcal{M}(x^+)$  is  $N$  the sets  $\mathcal{M}(x^+) \cap \{|x - x^+| < \varepsilon\}$  and  $\mathcal{M} \cap \{|x - x^+| < \varepsilon\}$  coincide. Thanks to the invariance of  $\mathcal{M}$  it follows that  $\mathcal{M}(x^+) \subset \mathcal{M}$ . As a result one obtains that

$$J(x^*) = - \int_{-\infty}^{\infty} DV(x^*(t)) \dot{x}^* = V(x^-) - V(x^+) \geq 0,$$

where the last inequality holds because (2.11) implies that  $V$  is a decreasing Lyapunov function along the trajectory  $x^*$ . Thus, formula (2.15) has been proven under the assumption that  $x^+$  is a local minimum. The remaining three cases ( $x^+$  is a local maximum,  $x^-$  is a local minimum/maximum) can be dealt with in an analogous fashion.  $\square$

The result shows that only saddle-saddle connections can give rise to heteroclinic orbits in the Hamiltonian system (2.9) which cannot be found by in one of the gradient flows (2.11). As a closing remark in this section we observe that the existence of heteroclinic orbit in (2.9) is a generic property of non-degenerate systems. For the Hamiltonian flow the effective dimension of the space is  $2N - 1$ , because it is constrained to a level set of the Hamiltonian. Since the stable and unstable manifolds for the Hamiltonian system are both of dimension  $N$  (cf eq. (2.16)) we expect that, generically, there will be a 1 dimensional manifold connecting any pair of equilibria under the Hamiltonian flow.

### 3 The $\Gamma$ -Limit

We now determine the  $\Gamma$ -limit for the Onsager-Machlup function (2.6). A thorough introduction to  $\Gamma$ -convergence can be found in [17] and [6]. Problems closely related to ours are studied using  $\Gamma$ -convergence in [9, 10].

We start with some basic definitions and propositions which serve to explain the form of the  $\Gamma$ -limit. The key result is contained in Proposition 3.2 which demonstrates that every transition is achieved via a *finite* collection of intermediate transitions. The limit Theorem 3.4 is then stated and proved, building on a number of lemmas which follow it.

Recall the definition of admissible potential  $V$ , as well as that of a transition path.

**Proposition 3.1.** *Let  $V$  be admissible and define for each pair  $x^-, x^+ \in \mathcal{E}$  the function  $\Phi(x^-, x^+)$  by*

$$\Phi(x^-, x^+) = \inf \{ J(y) \mid y \in X(x^-, x^+) \}.$$



Then there exists  $c > 0$  such that  $\Phi(x^-, x^+) > c$  for all pairs  $x^\pm \in \mathcal{E}$  with  $x^- \neq x^+$ .

We can establish a direct representation of the transition energy  $\Phi$  which involves only minima and avoids the usage of infima.

**Proposition 3.2.** *Let  $V$  be admissible and  $x^\pm \in \mathcal{E}$  be two critical points of  $V$ . Then there exists a finite sequence  $\{x_i\}_{i=0}^k \in \mathcal{E}$  such that  $x_0 = x^-$ ,  $x_k = x^+$  and*

$$\Phi(x^-, x^+) = \sum_{i=1}^k \min \{J(y) \mid y \in X(x_{i-1}, x_i)\},$$

A straight-forward refinement of the analysis shows that the sequence  $x_i$  is injective. Now we state the definition of  $\Gamma$ -convergence

**Definition 3.3.** *Let  $X$  be a Banach-space,  $\varepsilon > 0$  a parameter and  $I_\varepsilon : X \rightarrow \mathbb{R}$  a family of functionals. The functional  $I_0 : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is the  $\Gamma$ -limit of  $I_\varepsilon$  as  $\varepsilon \rightarrow 0$  if for all  $x \in X$  and all sequences  $x_\varepsilon \in X$  which converge weak-\* to  $x$  as  $\varepsilon$  tends to 0 the liminf-inequality*

$$\liminf_{\varepsilon \rightarrow 0} I_\varepsilon(x_\varepsilon) \geq I_0(x)$$

holds, and for all  $z \in X$ , there exists a recovery sequence  $z_\varepsilon \in X$  which converges weak-\* to  $z$  and satisfies

$$\limsup_{\varepsilon \rightarrow 0} I_\varepsilon(z_\varepsilon) \leq I_0(z).$$

**Theorem 3.4.** *Let  $V$  be an admissible potential. Then the  $\Gamma$ -limit of the functional  $I_\varepsilon$  as  $\varepsilon$  tends to 0 is*

$$I_0(x) = \begin{cases} \sum_{\tau \in \mathcal{D}} \Phi(x^-(\tau), x^+(\tau)) - \int_0^1 \Delta V(x(s)) \, ds & \text{if } x \in BV([0, 1]) \\ & \text{and } x(s) \in \mathcal{E} \text{ a.e. } s \in [0, 1], \\ +\infty & \text{else,} \end{cases}$$

where  $\mathcal{D}(x)$  is the set of discontinuity points of  $x$  and  $x^\pm(\tau)$  are the left and right-sided limits of  $x$  at  $\tau$ .

Proposition 3.2 shows that the infimum  $\Phi$  can be written as a finite sum of minima. These minima are obtained from evaluation of (2.14), where  $x^*$  solves (2.9) subject to  $x^*(t) \rightarrow x_{i-1}/x_i$  as  $t \rightarrow \pm\infty$ . Furthermore, if either  $x_{i-1}$  or  $x_i$  is a local extremum of  $V$ , then  $x^*$  is also a heteroclinic orbit in one of the gradient flows (2.11), then the minimum is given by (2.15). Thus Theorem 3.4 shows that the  $\Gamma$ -limit of  $I_\varepsilon$  can be computed through knowledge of the critical points of  $V$  and the set of (Hamiltonian or gradient) heteroclinic orbits connecting them.

We start the proof of Theorem 3.4 by proving a lemma which delivers a lower bound for the amount of energy needed to reach one of the stationary points  $x \in \mathcal{E}$  when starting nearby. We will often write  $x^\pm$  as a shorthand for  $\{x^-, x^+\}$ .

**Lemma 3.5.** *Let  $V$  be an admissible potential. There exists numbers  $C, \varepsilon_0 > 0$  such that for and all  $\varepsilon \in [0, \varepsilon_0]$ ,  $x^\pm \in \mathcal{E}$  and all paths  $y \in X(x^-, x^+)$  with  $\text{dist}(y(0), x^\pm) \geq \varepsilon$  the estimate*

$$J(y) \geq \frac{1}{2C}\varepsilon^2$$

holds.

*Proof.* We first show that there exist numbers  $\delta, C > 0$  such that, for all  $x, y \in \mathbb{R}^N$  with  $x \in \mathcal{E}$  and with the property  $|y - x| < \delta$ , the estimates

$$\frac{1}{C}|y - x|^2 \leq |DV(y)|^2 \leq C|y - x|^2 \quad (3.1)$$

Indeed, thanks to the smoothness of  $V$  and nondegeneracy condition  $\det(D^2V(x)) \neq 0$  for  $x \in \mathcal{E}$ , for sufficiently small  $\varepsilon$  and large  $\tilde{C}$

$$\begin{aligned} |DV(y) - D^2V(x) \cdot (y - x)| &\leq \tilde{C}|y - x|^2, \\ |z| &\leq \tilde{C}|D^2V(x) \cdot z| \quad \forall z \in \mathbb{R}^N \end{aligned}$$

hold. Thus

$$\begin{aligned} |DV(y)| &\geq |D^2V(x) \cdot (y - x)| - \tilde{C}|y - x|^2 \\ &\geq \frac{1}{C}|y - x| - \tilde{C}|y - x|^2 \geq \frac{1}{2C}|y - x| \end{aligned}$$

if  $|y - x| \leq \min\{\delta, \frac{1}{2\tilde{C}^2}\}$ . This shows that the first inequality in (3.1) holds. The second inequality is obtained in a similar way.

Let  $y \in X(x^-, x^+)$  and recall that by Sobolev's imbedding theorem  $y$  is continuous. For each  $\varepsilon \in (0, \min\{\delta, \frac{1}{2\tilde{C}^2}\})$  we define

$$t^+(\varepsilon) = \min \{s \in \mathbb{R} \mid |y(t') - x^+| < \varepsilon \text{ for all } t' > s\}, \quad (3.2)$$

and  $t^-$  analogous. Then we obtain

$$\begin{aligned} J(y) &\geq \int_{t^+(\varepsilon)}^\infty |DV(y(s))| |\dot{y}(s)| ds \geq \frac{1}{2C} \int_{t^+(\varepsilon)}^\infty (x^+ - y) \cdot \dot{y}(s) ds \\ &= \frac{1}{4C} |y(t^+(\varepsilon)) - x^+|^2 = \frac{1}{4C} \varepsilon^2. \end{aligned}$$

The claim follows since we can derive the same estimate with  $t^-$  instead of  $t^+$ .  $\square$

*Proof of Proposition 3.1.* Let  $r > 0$  be the separation of the stationary points of  $V$ , i.e

$$r = \min\{|x - x'| \mid x, x' \in \mathcal{E}, x \neq x'\},$$

and let  $y_l \in X(x^-, x^+)$  be such that  $\lim_{l \rightarrow \infty} J(y_l) = \Phi(x^-, x^+)$ . There exists a sequence  $t_l \in \mathbb{R}$  such that  $\text{dist}(y_l(t_l), \mathcal{E}) \geq r/2$  and thus Lemma 3.5 applied to the translated sequence  $y_l(\cdot - t_l)$  implies that  $\inf_l J(y_l) \geq c$ , with  $c = \frac{1}{2C} \min\{r/2, \varepsilon_0\}^2$ .  $\square$

The proof of Proposition 3.2, showing the existence of minimizing connecting orbits, is established with the direct method of the calculus of variations, with the aid of the following lemma.

**Lemma 3.6.** *Let  $x^\pm \in \mathcal{E}$ . If there exists a minimizing sequence  $\{y_l\}_{l \in \mathbb{N}} \in X(x^-, x^+)$  such that the density  $\rho_l(t) = \frac{1}{2} (|\dot{y}_l(t)|^2 + |DV(y_l(t))|^2)$  is tight in  $L^1(\mathbb{R})$ , then there exists a minimizer  $y^\pm \in X(x^-, x^+)$  such that  $J(y^\pm) = \Phi(x^-, x^+)$ .*

*Proof.* First we show that the boundary conditions are not lost during the passage to the limit. The tightness of  $\rho_l$  implies that for each  $\varepsilon > 0$  there exists  $\tau^\pm(\varepsilon) \in \mathbb{R}$  with  $\tau^-(\varepsilon) < \tau^+(\varepsilon)$  and the property that, for all  $l$ ,

$$\limsup_{l \rightarrow \infty} \frac{1}{2} \int_{\mathbb{R} \setminus (\tau^-, \tau^+)} (|\dot{y}_l|^2 + |DV(y_l(s))|^2) ds \leq \frac{1}{4C} \varepsilon^2, \quad (3.3)$$

where  $C$  is the constant in Lemma 3.5. Let

$$t_l^+(\varepsilon) = \min \{s \in \mathbb{R} \mid |y_l(t') - x^+| < \varepsilon \text{ for all } t' > s\},$$

and  $t_l^-$  analogously. Without loss of generality we assume that  $t^-(\varepsilon) \leq 0 \leq t_l^+(\varepsilon)$  and  $|y_l(0) - x^\pm| \geq \varepsilon$ . Then Lemma 3.5 implies that

$$\limsup_{l \rightarrow \infty} \frac{1}{2} \int_{t_l^+}^{\infty} (|\dot{y}_l|^2 + |DV(y_l(s))|^2) ds \geq \frac{1}{4C} \varepsilon^2.$$

Together with inequality (3.3) we obtain that  $\limsup_{l \rightarrow \infty} t_l^+(\varepsilon) \leq \tau^+(\varepsilon)$ . Another application of this argument with  $t^+$  replaced by  $t^-$  delivers the desired result concerning the boundary conditions: for each  $\varepsilon > 0$  there exists  $-\infty < \tau^-(\varepsilon) < \tau^+(\varepsilon) < \infty$  such that

$$\begin{cases} \limsup_{l \rightarrow \infty} \sup_{s \geq \tau^+} |y_l(s) - x^+| \leq \varepsilon, \\ \limsup_{l \rightarrow \infty} \sup_{s \leq \tau^-} |y_l(s) - x^-| \leq \varepsilon. \end{cases} \quad (3.4)$$

From the bound on  $\dot{y}_l$  in  $L^2(\tau^-(\varepsilon), \tau^+(\varepsilon))$  we deduce that the length of the vector  $y_l(0) \in \mathbb{R}^N$  is bounded. Since  $J(y_l)$  is bounded the sequence  $\dot{y}_l$  is bounded in  $L^2(\mathbb{R})$ . We extract a subsequence  $y_l$  (not relabeled) such that  $\dot{y}_l$  converges weakly to  $\eta \in L^2(\mathbb{R})$  and  $y_l(0)$  converges to  $z \in \mathbb{R}^d$ .

We define the limiting path  $y^\pm$  by

$$y^\pm(t) = z + \int_0^t \eta(s) ds$$

Equation (3.4) implies that  $\lim_{s \rightarrow \pm\infty} y^\pm(s) = x^\pm$ , i.e. the boundary conditions are satisfied.

Next we demonstrate that  $y^\pm$  is a minimizer, i.e.

$$\lim_{l \rightarrow \infty} J(y_l) = J(y^\pm).$$

Note that  $J$  is weakly lower semicontinuous: the first term is weakly lower semicontinuous because it is convex whilst the second term is weakly continuous. To see the weak continuity note that

$$\begin{aligned} \int_{\mathbb{R}} |DV(y^\pm(s))| \, ds &= \lim_{R \rightarrow \infty} \int_{-R}^R |DV(y^\pm(s))| \, ds \\ &= \lim_{R \rightarrow \infty} \lim_{l \rightarrow \infty} \int_{-R}^R |DV(y_l(s))| \, ds = \lim_{l \rightarrow \infty} \int_{-\infty}^{\infty} |DV(y_l(s))| \, ds. \end{aligned}$$

The last equation is again due to the tightness of  $J(y_l)$ , the second but last equation holds because of Sobolev's embedding theorem. The weak lower semicontinuity of  $J$ , coupled with the fact that  $\{y_l\}$  is a minimizing sequence, shows that  $J(y^\pm) = \lim_{l \rightarrow \infty} J(y_l)$ .

Finally we show that  $\dot{y}^\pm \in L^1(\mathbb{R})$ , which entails the claim  $y^\pm \in X(x^-, x^+)$ . Define the functional

$$G_t(y) = \frac{1}{2} \int_t^\infty (|\dot{y}(s)|^2 + |DV(y(s))|^2) \, ds,$$

and the function  $g(t) = G_t(y^\pm)$ . Note that  $g$  is monotonically decreasing with  $t$  and  $(y^\pm(s))_{s \geq t}$  minimizes  $G_t(y)$  for every  $t$ , subject to the boundary condition  $y(t) = y^\pm(t)$ . Testing  $G_t(\cdot)$  with a function which is affine on  $[t, t+1]$  and assumes the value  $x^+$  for all  $s \geq t+1$  together with estimate (3.1) delivers the bound

$$g(t) \leq \frac{1}{2} |y^\pm(t) - x^+|^2 \int_0^1 (1 + Cs^2) \, ds = \frac{1}{2} \left(1 + \frac{C}{3}\right) |y^\pm(t) - x^+|^2. \quad (3.5)$$

Recall that  $\lim_{t \rightarrow \infty} y^\pm(t) = x^+$  and choose  $t^+ > 0$  such that  $|y(t) - x^+| < \varepsilon_0$  for all  $t > t^+$ , where  $\varepsilon_0$  is defined in Lemma 3.5. The function  $g$  satisfies for all  $t > t^+$  the differential inequality

$$-\frac{dg}{dt}(t) \geq \frac{1}{2} |DV(y^\pm(t))|^2 \geq \frac{1}{2C} |y^\pm(t) - x^+|^2 \geq \mu g(t),$$

with  $\mu = \frac{3}{(3+C)C}$ . The last inequality is (3.5), the penultimate inequality is due to (3.1). Gronwall's inequality implies the exponential decay bound  $g(t) \leq \varepsilon_0 e^{-\mu(t-t^+)}$ , and Lemma 3.5 delivers that  $y^\pm(t)$  converges exponentially fast to  $x^+$  as  $t$  tends to infinity. Since  $y^\pm$  satisfies the Euler-Lagrange equation (2.9) we obtain that  $\dot{y}^\pm(t)$  converges exponentially fast to 0 and in particular  $\dot{y}^\pm \in L^1([t^+, \infty))$ . The same argument also shows that  $\dot{y}^\pm \in L^1((-\infty, t^-])$  for some  $t^-$ , and thus  $y^\pm \in L^1(\mathbb{R})$ .  $\square$

The proof of Proposition 3.2 also relies on Lions' concentration compactness lemma which we state here for completeness.

**Lemma 3.7.** [18] *Let  $(\rho_l)_{l \geq 1}$  be a sequence in  $L^1(\mathbb{R})$  satisfying*

$$\rho_l \geq 0 \text{ in } \mathbb{R} \text{ and } \lim_{l \rightarrow \infty} \int_{\mathbb{R}} \rho_l(t) dt = \lambda,$$

where  $\lambda > 0$  is fixed. Then, there exists a subsequence (not relabeled) satisfying one of the following three possibilities:

1. (compactness) there exists  $t_l \in \mathbb{R}$  such that  $\rho_l(\cdot - t_l)$  is tight, i.e.

$$\forall \varepsilon > 0, \exists R < \infty \text{ such that } \int_{t_l - R}^{t_l + R} \rho_l(s) ds \geq \lambda - \varepsilon;$$

2. (vanishing)

$$\lim_{l \rightarrow \infty} \sup_{t \in \mathbb{R}} \int_{t-R}^{t+R} \rho_l(s) ds = 0 \text{ for all } R > 0;$$

3. (splitting) there exists  $0 < \alpha < \lambda$  such that for all  $\varepsilon > 0$  there exists  $l_0 \geq 1$  and  $\rho_l^1, \rho_l^2 \in L^1_{\geq 0}(\mathbb{R})$  such that for all  $l \geq l_0$

$$\left\{ \begin{array}{l} \|\rho_l^1 + \rho_l^2 - \rho_l\|_{L^1} + \|\rho_l^1\|_{L^1} - \alpha + \|\rho_l^2\|_{L^1} + \alpha - \lambda \leq \varepsilon, \\ \lim_{l \rightarrow \infty} \text{dist}(\text{supp}(\rho_l^1), \text{supp}(\rho_l^2)) = \infty. \end{array} \right.$$

*Proof of Proposition 3.2.* Let  $y_l \in X(x^-, x^+)$  be a minimizing sequence of  $J$  and define  $\rho_l(t)$  as in Lemma 3.6. We now use Lemma 3.7 in the following way: we show that vanishing cannot occur, and that splitting can only occur a finite number of times.

To rule out vanishing we define, for real number  $\varepsilon$  satisfying

$$\varepsilon \in \left(0, \frac{1}{2} \min\{|x - z| \mid x, z \in \mathcal{E}, x \neq z\}\right),$$

the starting and ending time of the final transition:

$$\begin{aligned} t_l^+ &= \sup \left\{ t \in \mathbb{R} \mid \inf_{s \leq t} |y_l(s) - x^+| \geq \varepsilon \right\}, \\ t_l^- &= \sup \left\{ t \leq t_l^+ \mid \text{dist}(y_l(t), \mathcal{E} \setminus \{x^+\}) \leq \varepsilon \right\}. \end{aligned}$$

By construction  $t_l^- \leq t_l^+$  and Jensen's inequality implies that

$$\begin{aligned} \int_{t_l^-}^{t_l^+} \rho_l(s) ds &\geq \frac{1}{2} \int_{t_l^-}^{t_l^+} \left( \frac{\text{dist}(x^+, \mathcal{E} \setminus \{x^+\}) - 2\varepsilon}{t_l^+ - t_l^-} \right)^2 ds \\ &= \frac{(\text{dist}(x^+, \mathcal{E} \setminus \{x^+\}) - 2\varepsilon)^2}{2(t_l^+ - t_l^-)}. \end{aligned}$$

Since  $\limsup_{l \rightarrow \infty} \int_{\mathbb{R}} \rho_l(s) ds < \infty$  this implies that  $\liminf_{l \rightarrow \infty} (t_l^+ - t_l^-) > 0$ . Furthermore, since  $V$  admissible there exists a number  $c > 0$  such that  $\inf_{s \in [t_l^-, t_l^+]} \rho_l(s) \geq c$  and thus

$$\limsup_{l \rightarrow \infty} (t_l^+ - t_l^-) \leq \frac{1}{c} \limsup_{l \rightarrow \infty} \int_{t_l^-}^{t_l^+} \rho_l(s) ds < \infty.$$

This pair of inequalities shows that  $\rho_l$  does not vanish.

If the sequence  $\rho_l$  splits, then there exist sequences  $a_l \leq b_l$  such that  $[a_l, b_l] \cap \text{supp}(\rho_l^1 + \rho_l^2) = \emptyset$  and

$$\lim_{l \rightarrow \infty} (a_l - b_l) = \infty, \quad (3.6)$$

$$\liminf_{l \rightarrow \infty} \int_{-\infty}^{a_l} \rho_l(t) dt > 0 \text{ and } \liminf_{l \rightarrow \infty} \int_{b_l}^{\infty} \rho_l(t) dt > 0. \quad (3.7)$$

We claim now that there exists a third sequence  $t_l \in [a_l, b_l]$  and a stationary point  $x \in \mathcal{E}$  such that  $\lim_{l \rightarrow \infty} y_l(t_l) = x$ . This is because

$$\limsup_{l \rightarrow \infty} \int_{a_l}^{b_l} \rho_l(t) dt < \infty,$$

which together with (3.6) and the weak coercivity assumption (2.12) delivers that

$$\liminf_{l \rightarrow \infty} \{\text{dist}(y_l(t), \mathcal{E}) \mid t \in [a_l, b_l]\} = 0.$$

We define next the sequences  $y_l^1 \in X(x^-, x)$  and  $y_l^2 \in X(x, x^+)$  as follows:

$$y_l^1(t) = \begin{cases} y_l(t) & \text{if } t \leq t_l, \\ (t_l - t + 1)y_l(t_l) + (t - t_l)x & \text{if } t_l < t < t_l + 1, \\ x & \text{if } t \geq t_l + 1, \end{cases}$$

$$y_l^2(t) = \begin{cases} x & \text{if } t \leq t_l - 1, \\ (t - t_l + 1)y_l(t_l) + (t_l - t)x & \text{if } t_l - 1 < t < t_l, \\ y_l(t) & \text{if } t \geq t_l. \end{cases}$$

Clearly  $y_l^1 + y_l^2 - x \in X(x^-, x^+)$  and the convergence  $\lim_{l \rightarrow \infty} y_l(t_l) = x$  implies that

$$\lim_{l \rightarrow \infty} J(y_l^1 + y_l^2 - x) = \lim_{l \rightarrow \infty} (J(y_l^1) + J(y_l^2)) = \lim_{l \rightarrow \infty} J(y_l). \quad (3.8)$$

Moreover, equation (3.7) implies

$$\liminf_{l \rightarrow \infty} J(y_l^1) > 0 \text{ and } \liminf_{l \rightarrow \infty} J(y_l^2) > 0. \quad (3.9)$$

We will show now that  $x \notin \{x^\pm\}$ . Indeed, if  $x = x^-$ , then  $y_l^2 \in X(x^-, x^+)$  and by (3.8)

$$\liminf_{l \rightarrow \infty} J(y_l^2) \leq \lim_{l \rightarrow \infty} J(y_l) - \liminf_{l \rightarrow \infty} J(y_l^1).$$

Equation (3.9) delivers a contradiction to the assumption that  $y_l$  is a minimizing sequence. An analogous argument can be constructed for the case  $x = x^+$  and thus  $x \in \mathcal{E} \setminus \{x^\pm\}$ .

The construction yields the following equation

$$\begin{aligned} \Phi(x^-, x^+) &\geq \lim_{l \rightarrow \infty} J(y_l^1 + y_l^2 - x) \geq \liminf_{l \rightarrow \infty} J(y_l^1) + \liminf_{l \rightarrow \infty} J(y_l^2) \\ &\geq \Phi(x^-, x) + \Phi(x, x^+). \end{aligned}$$

Progressing inductively we obtain a sequence  $x_0, x_1, \dots, x_k \in \mathcal{E}$  such that  $x_0 = x^-, x_k = x^+, x_{i-1} \neq x_i$  for all  $i = 1 \dots k$  and

$$\Phi(x^-, x^+) \geq \sum_{i=1}^k \Phi(x_{i-1}, x_i).$$

Proposition 3.1 implies that  $k \leq \Phi(x^-, x^+)/c$ , placing a finite bound on the number of splittings possible. The minimizing sequences associated with  $X(x_{i-1}, x_i)$  are eventually tight and so achieve their contribution to the infimum at a minimizer of  $J$ .  $\square$

*Proof of Theorem 3.4.* Now the representation formula for the  $\Gamma$ -limit will be verified. Let  $x \in BV([0, 1], \mathbb{R}^d)$  be a limit path and  $x_\varepsilon \in H^1(0, 1)$  such that  $x_\varepsilon$  converges to  $x$  weak-\* in  $BV([0, 1])$ . First we show that the Laplacian can be treated separately. Since weak-\* convergence entails boundedness in  $BV$  there is a constant  $C > 0$  such that  $\limsup_{\varepsilon \rightarrow 0} \|x_\varepsilon\|_{L^\infty} \leq C$ . Furthermore, thanks to Helly's theorem we can select a subsequence which converges pointwise for almost every  $s \in [0, 1]$  and therefore by dominated convergence

$$\lim_{\varepsilon \rightarrow 0} \int_0^1 \Delta V(x_\varepsilon(s)) \, ds = \int_0^1 \Delta V(x(s)) \, ds.$$

Note that only continuity of  $\Delta V$  is required for this step. This implies that

$$\liminf_{\varepsilon \rightarrow 0} I_\varepsilon(x_\varepsilon) = \liminf_{\varepsilon \rightarrow 0} J_\varepsilon(x_\varepsilon) - \int_0^1 \Delta V(x(s)) \, ds \quad (3.10)$$

with  $J_\varepsilon$  as defined in (2.7).

Next we consider the case where the set of times  $s$  where the limit function  $x(s) \in \mathcal{E}$  does not have full measure, so that  $I_0(x) = \infty$ . Define

$$M = \text{meas}(\{t \in [0, 1] \mid x(t) \notin \mathcal{E}\}).$$

There exists  $\delta > 0$  such that

$$\liminf_{\varepsilon > 0} \text{meas}(\{t \in [0, 1] \mid \text{dist}(x(t), \mathcal{E}) > \delta\}) > M/2.$$

Since  $V$  satisfies the weak coercivity condition the infimum of  $|DV(x)|$  exceeds some  $\lambda > 0$ , on this set with measure exceeding  $M/2$ . Thus

$$\liminf_{\varepsilon \rightarrow 0} J_\varepsilon(x_\varepsilon) \geq \liminf_{\varepsilon \rightarrow 0} \frac{1}{2\varepsilon} \int_0^1 |DV(x_\varepsilon(s))|^2 ds \geq \liminf_{\varepsilon \rightarrow 0} \frac{M}{4\varepsilon} \lambda^2 = +\infty.$$

Finally we consider the case where  $I_0(x)$  is finite. The proof of the lim inf-inequality follows directly from the definition, as we now show. Since  $x \in BV([0, 1])$  and  $I_0(x)$  is finite, the set of discontinuity points  $\mathcal{D} \subset [0, 1]$  is finite. We assume next that  $\mathcal{D} \subset (0, 1)$ , i.e.  $x$  is continuous at the end-points. If  $x$  jumps at one or both end points the same argument can be repeated with obvious modifications of the cut-offs.

Thanks to this assumption for each  $\tau \in \mathcal{D}$  the left-sided and right-sided limits  $x^\pm(\tau)$  exist. Since  $x_\varepsilon$  converges weak-\* in  $BV([0, 1])$  to the piecewise constant function  $x$  Helly's theorem implies that  $\lim_{\varepsilon \rightarrow 0} x_\varepsilon(t) = x(t)$  for almost every  $t \in [0, 1] \setminus \mathcal{D}$ . Choose

$$\delta = \frac{1}{2} \min \{ \tau - \tau' \mid \tau, \tau' \in \mathcal{D} \cup \{0, 1\}, \tau < \tau' \}.$$

and for each  $\tau \in \mathcal{D}$  a pair of such points  $t^\pm \in (0, 1)$  such that  $\tau - \delta \leq t^- < \tau < t^+ \leq \tau + \delta$  and  $\lim_{\varepsilon \rightarrow 0} x_\varepsilon(t^\pm) = x^\pm(\tau)$ . Setting  $y_\varepsilon(t) = x_\varepsilon(\varepsilon t)$  one obtains

$$\begin{aligned} J_\varepsilon(x_\varepsilon) &= \frac{1}{2} \int_0^1 \left( \varepsilon |\dot{x}(s)|^2 + \frac{1}{\varepsilon} |DV|^2 \right) ds = \frac{1}{2} \int_0^{1/\varepsilon} \left( |\dot{y}_\varepsilon(s)|^2 + |DV(y_\varepsilon(s))|^2 \right) ds \\ &\geq \sum_{\tau \in \mathcal{D}} \frac{1}{2} \int_{(t^-(\tau)-\delta)/\varepsilon}^{(t^+(\tau)+\delta)/\varepsilon} \left( |\dot{y}_\varepsilon(s)|^2 + |DV(y_\varepsilon(s))|^2 \right) ds. \end{aligned}$$

To justify the boundary condition we modify the path segments  $y_\varepsilon(t - \tau/\varepsilon)$  in a way such that the energy changes only slightly. Define now the function

$$\bar{y}_\varepsilon^\tau(t) = \begin{cases} x^-(\tau) & \text{if } s \leq \frac{t^-}{\varepsilon} - 1 \\ \left(\frac{t^-}{\varepsilon} - s\right)x^-(\tau) + \left(s - \frac{t^-}{\varepsilon} + 1\right)x_\varepsilon(t^-) & \text{if } \frac{t^-}{\varepsilon} - 1 < s < \frac{t^-}{\varepsilon} \\ y_\varepsilon(s) & \text{if } \frac{t^-}{\varepsilon} \leq s \leq \frac{t^+}{\varepsilon}, \\ \left(s - \frac{t^+}{\varepsilon}\right)x^+(\tau) + \left(\frac{t^+}{\varepsilon} + 1 - s\right)x_\varepsilon(t^+) & \text{if } \frac{t^+}{\varepsilon} < s < \frac{t^+}{\varepsilon} + 1, \\ x^+(\tau) & \text{if } s \geq \frac{t^+}{\varepsilon} + 1. \end{cases}$$

The construction above implies that  $\bar{y}_\varepsilon^\tau \in X(x^-(\tau), x^+(\tau))$  and thus

$$J(\bar{y}_\varepsilon^\tau) \geq \Phi(x^-(\tau), x^-(\tau)).$$



Furthermore, the modification of the function  $y_\varepsilon$  affects the energy only in a negligible way, i.e.

$$\sum_{\tau \in \mathcal{D}} \frac{1}{2} \int_{(\tau-\delta)/\varepsilon}^{(\tau+\delta)/\varepsilon} \left( |\dot{y}_\varepsilon(s)|^2 + |DV(y_\varepsilon(s))|^2 \right) ds \geq \sum_{\tau \in \mathcal{D}} J(\bar{y}_\varepsilon^\tau) + o(1).$$

Hence, the liminf-inequality

$$\lim_{\varepsilon \rightarrow 0} J_\varepsilon(y_\varepsilon) \geq \sum_{\tau \in \mathcal{D}} \Phi(x^-(\tau), x^+(\tau))$$

holds. The same argument also shows that

$$\liminf_{\varepsilon \rightarrow 0} J_\varepsilon(x_\varepsilon) = +\infty,$$

if the set  $\mathcal{D}$  is infinite.

To construct a recovery sequence we fix  $\delta > 0$ . By definition for each  $\tau \in \mathcal{D}$  there exists

$$y_\delta^\tau \in X(x^-(\tau), x^+(\tau))$$

such that the support of  $\dot{y}_\delta^\tau$  is compact for all  $\tau \in \mathcal{D}$  and  $J(y_\delta^\tau) \leq \Phi(x^-(\tau), x^+(\tau)) + \delta$ . It can be checked that for fixed  $\delta > 0$  the sequence

$$y_{\varepsilon, \delta}(t) = x(0) + \frac{1}{\varepsilon} \sum_{\tau \in \mathcal{D}} \int_0^t \dot{y}_\delta^\tau((s - \tau)/\varepsilon) ds$$

converges weak-\* in  $BV([0, 1])$  to  $x$  as  $\varepsilon \rightarrow 0$ .

Moreover, there exists a function  $\delta(\varepsilon)$  such that  $\lim_{\varepsilon \rightarrow 0} \delta(\varepsilon) = o(1)$  and the supports of the functions  $\dot{y}_\delta^\tau((\cdot - \tau)/\varepsilon)$ ,  $\tau \in \mathcal{D}$  are disjoint as  $\varepsilon$  tends to 0. Hence,

$$\begin{aligned} J_\varepsilon(y_{\varepsilon, \delta(\varepsilon)}) &= \sum_{\tau \in \mathcal{D}} J_\varepsilon(y_\delta^\tau((\cdot - \tau)/\varepsilon)) \leq \sum_{\tau \in \mathcal{D}} \Phi(x^-(\tau), x^+(\tau)) + \#\mathcal{D} \delta(\varepsilon) \\ &= \sum_{\tau \in \mathcal{D}} \Phi(x^-(\tau), x^+(\tau)) + o(1) \end{aligned}$$

as  $\varepsilon$  tends to 0. □

## 4 Numerical Experiments

The aim of these numerical experiments is to illustrate that the  $\Gamma$ -limit derived in section 3 accurately captures the behaviour of the problem of minimizing  $I_\varepsilon$  given by (2.6) when  $\varepsilon$  is small. The intuitive picture of the  $\Gamma$ -limit is that it is comprised of minimizers with the following properties:

1. The minimizers are BV functions supported on the set of critical points of  $V$  (see Theorem 3.4).
2. The contribution to the limit functional from the jumps in these BV functions (ie from  $J_\varepsilon$ ) can be expressed in terms of sums of integrals  $\int_\infty^{-\infty} |DV(x^*(s))|^2 ds$  where  $x^*$  is a heteroclinic orbit for the Hamiltonian equations (2.9) (see Lemma 2.1, equation (2.14)).
3. Unless both equilibria  $x^\pm$  are saddle points, then these heteroclinic orbits will also be heteroclinic for the gradient system (2.11) and then the contribution to the limit functional from  $J_\varepsilon$  can be expressed in terms of sums  $|V(x^*(\infty)) - V(x^*(-\infty))|$  (see Lemma 2.1, equation (2.15)).
4. Minimizers  $x^*$  will seek to organize the support of the limiting BV function so as to minimize the value of  $\int_0^1 \Delta V(x^*)(s) ds$ ; this is the second contribution to the infimum defined in Theorem 3.4.

A variety of numerical computations, all of which exhibit these phenomena on a range of problems, including high dimensional systems arising in vacancy diffusion and the Lennard-Jones 38 cluster, may be found in the paper [5]. The purpose of this section is to illustrate the four points above on a single low dimensional example and relate the results in an explicit way to the theory developed in earlier sections. We employ the potential  $V : \mathbb{R}^2 \rightarrow \mathbb{R}$  given by

$$V(x_1, x_2) = (x_1^2 + x_2^2) ((x_1 - 1)^2 + x_2^2) (x_1^2 + (x_2 - 1)^2)$$

which is shown in Figure 1. The potential has three wells of equal depth, situated at  $M_0 = (0, 0)$ ,  $M_1 = (1, 0)$ , and  $M_2 = (0, 1)$ . Saddle points exist at  $S_1 = ((2 + \sqrt{2})/6, (2 - \sqrt{2})/6)$  and (by symmetry) at  $S_2 = ((2 - \sqrt{2})/6, (2 + \sqrt{2})/6)$ . The potential is zero at the minima and attains a value of  $2/27$  at the saddles. The Laplacian of  $V$  has value zero at the saddle points, 4 at the minimum  $M_0$  and 8 at the minima  $M_1$  and  $M_2$ .

In the following numerical experiments we use gradient descent to minimize  $I_\varepsilon$  or  $J_\varepsilon$  given by (2.8) and (2.7). In all the experiments we employ a value of  $\varepsilon^{-1} = 10^{-3}$  which proved to be small enough to exhibit the behaviour of the  $\Gamma$ -limit. We solve the parabolic PDE arising from the  $L^2$  gradient flow for  $I_\varepsilon$  (resp.  $J_\varepsilon$ ) by means of a linearly implicit method with stepsize chosen to ensure decrease of  $I_\varepsilon$  (resp.  $J_\varepsilon$ ) at each time-step.

Figure 2 shows heteroclinic orbits for the gradient flow (2.11) (in green) and for the Hamiltonian flow (2.9) (in blue). Figure 3 shows minimizers of  $J_\varepsilon$ , both connecting  $M_1$  and  $M_2$ . It is instructive to compare this figure with the preceding Figure 2. The green curve in Figure 3 connects  $M_1$  to  $M_2$  via  $M_0$  and is comprised of 4 segments each of which has been verified to be approximately given by the gradient heteroclinic orbits (2.11). The blue curve connects  $M_1$  to  $M_2$  via  $S_1$  and  $S_2$  and has been verified to be approximately given by Hamiltonian heteroclinic orbits satisfying (2.9). Furthermore, in the gradient case we have verified that the minimizers obey the sum rule (2.15) This illustrates the connection between minimizers of  $J_\varepsilon$  and solutions of the Euler Lagrange equations for minimizers of  $J$ , and points 2. and 3. in particular.

Figures 4 and 5 show minimizers of  $J_\varepsilon$  and  $I_\varepsilon$  which connect the two saddle points  $S_1$  and  $S_2$ . Several approximate minimizers are shown in each case, found

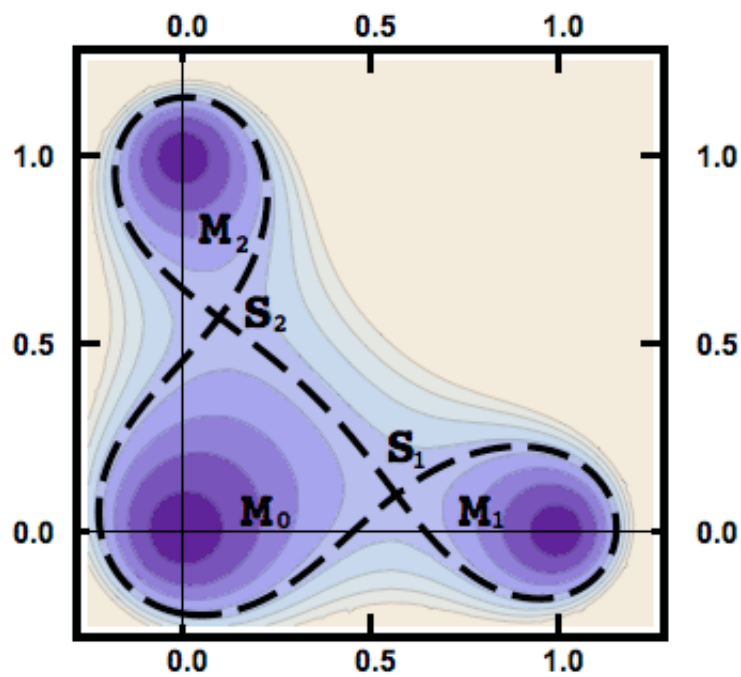


Figure 1: Contour plot of the potential. The dashed black line is the equipotential contour that connects the saddle points,  $S_1$  and  $S_2$ . The three equal-depth minima are labeled:  $M_0$ ,  $M_1$ , and  $M_2$ .

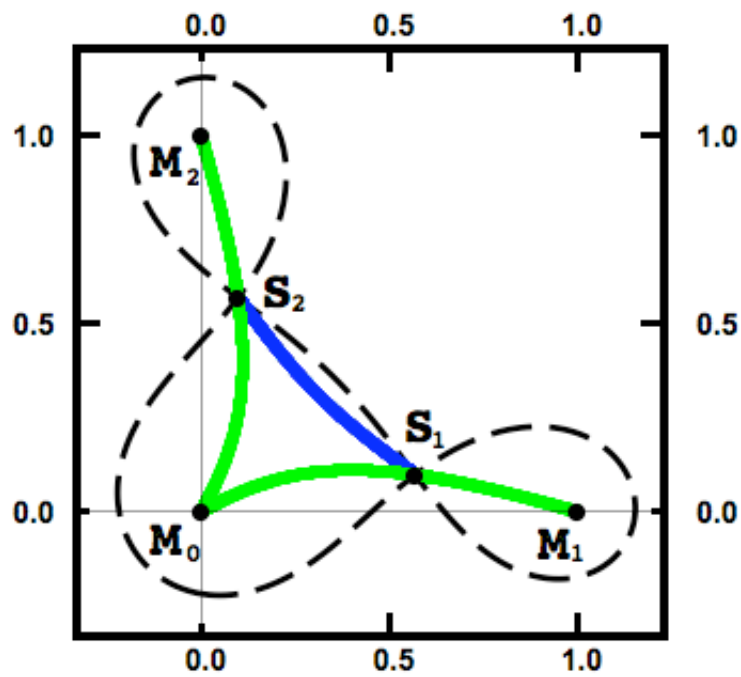


Figure 2: The dashed black line is the equipotential contour that connects the saddle points,  $S_1$  and  $S_2$ . The green line is the path that follows the gradient of the potential from  $M_1$  to  $M_0$  via  $S_1$  and then to  $M_2$  via  $S_2$ . The blue line is a Hamiltonian minimizer (not a gradient flow) that connects  $S_1$  to  $S_2$ .

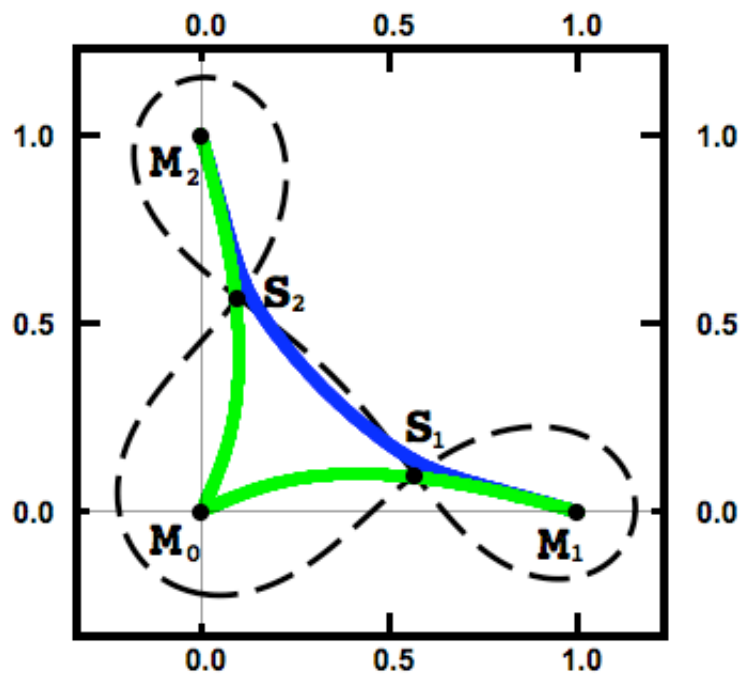


Figure 3: Solutions that minimize  $J_\varepsilon$  at  $\varepsilon = 10^{-3}$ . The dashed black line is the equipotential contour that connects the saddle points,  $S_1$  and  $S_2$ . The green line is the path that starts at  $M_1$  and proceeds to  $M_2$  via  $M_0$ . The blue line is the path path that starts at  $M_1$  and proceeds to  $M_2$  avoiding  $M_0$ .

from different starting points for the gradient flow. The experiments illustrate point 1. as they show that the solutions concentrate on critical points of  $V$ : in this case simply the two saddles. They also illustrate point 4. as they show that, in this case, the minimizers of  $J_\varepsilon$  and  $I_\varepsilon$  are indistinguishable; this is because the Laplacian of  $V$  is zero at the saddle points.

Figures 6 and 7 also show minimizers of  $J_\varepsilon$  and  $I_\varepsilon$  which connect the two saddle points  $S_1$  and  $S_2$ . However the starting points for the gradient flow differ from those used to generate Figures 4 and 5; in particular they are based on a function which passes through the minimum  $M_0$ . As a consequence the minimizers also pass through  $M_0$ . For  $J_\varepsilon$  there are then multiple approximate minimizers, all supported on  $S_1, S_2$  and  $M_0$ . However the support can be organized more or less arbitrarily (provided only two transitions occur) to obtain approximately the same value of  $J_\varepsilon$ ; we show a solution where the support is organized symmetrically. The situation for  $I_\varepsilon$  is quite different: the effect of the Laplacian of  $V$ , which is 4 at  $M_0$  and 0 at  $S_1$  and  $S_2$ , means that minimizers place most of their support at  $M_0$ . The experiments thus again illustrate point 1. as they show that the solutions concentrate on critical points of  $V$ . They also illustrate point 4.

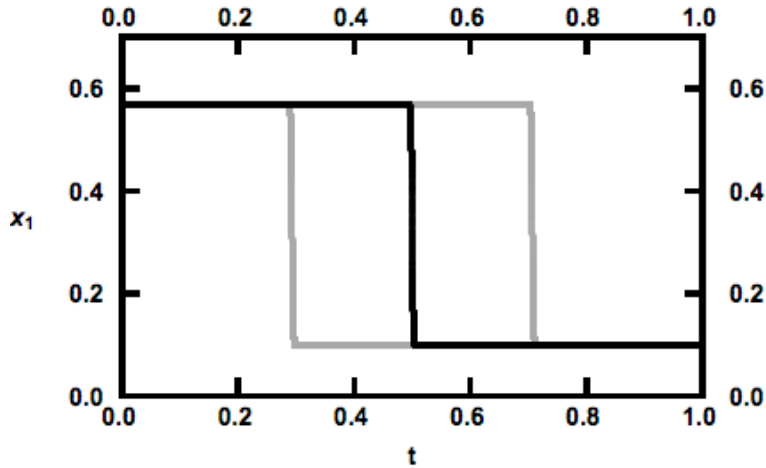


Figure 4: The  $x_1$  component of the paths that minimize  $J_\varepsilon$  at  $\varepsilon = 10^{-3}$ . The paths start at  $S_1$  and are conditioned to end at  $S_2$ . The ending path depends on the nature of the starting path. The results of using three different starting paths are displayed. None of the initial paths approach the origin and thus avoid passing through  $M_0$ . The black curve corresponds to a symmetric minimizer.

Figures 8 and 9 again show minimizers of  $J_\varepsilon$  and  $I_\varepsilon$ , now connecting the two minima  $M_1$  and  $M_2$ , and constructed to pass through the other minimum  $M_0$  and the two saddle points  $S_1$  and  $S_2$ . For  $J_\varepsilon$  there are then multiple approximate minimizers, all supported on the five critical points, one of which is shown in Figure 8, a solution where the support is organized symmetrically. The situation for  $I_\varepsilon$  is again very different: the effect of the Laplacian of  $V$ , which is 8 at  $M_1$  and  $M_2$ , means that minimizers place most of their support at these two points, as shown in Figure 9. The single interface in fact contains several transitions, and hence several contributions to the  $\Gamma$ -limit. Furthermore this single interface can be placed ar-

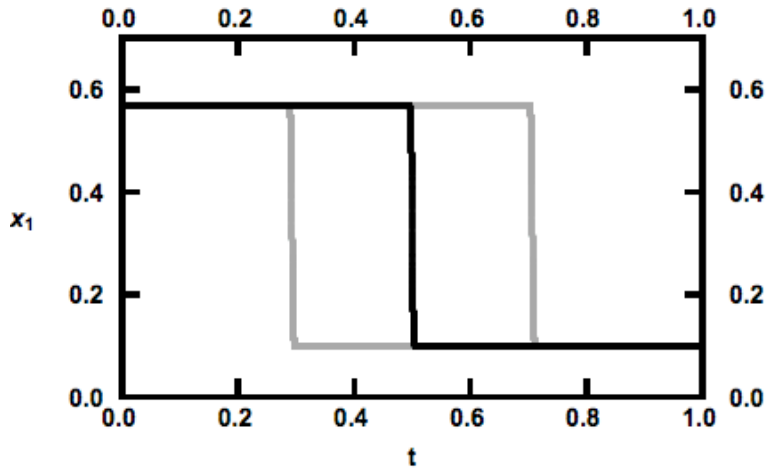


Figure 5: The  $x_1$  component of the paths that minimize  $I_\varepsilon$  at  $\varepsilon = 10^{-3}$ . The paths start at  $S_1$  and are conditioned to end at  $S_2$ . The ending path depends on the nature of the starting path. The results of using three different starting paths are displayed. None of the initial paths approach the origin and thus avoid passing through  $M_0$ . The black curve corresponds to a symmetric minimizer. Note that this plot is indistinguishable from the previous one.

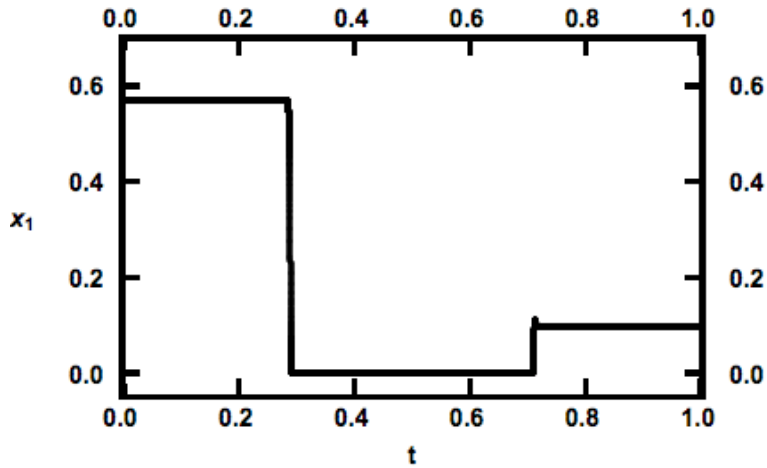


Figure 6: The  $x_1$  component of the path that minimize  $J_\varepsilon$  at  $\varepsilon = 10^{-3}$ . The path starts at  $S_1$  and is conditioned to end at  $S_2$ . The initial path contains the origin  $M_0$  and thus the ending path spends an arbitrary fraction of the time at  $M_0$ . The ending path depends on the nature of the starting path, in particular where the path crosses the origin. Only the symmetric minimizer is displayed here.

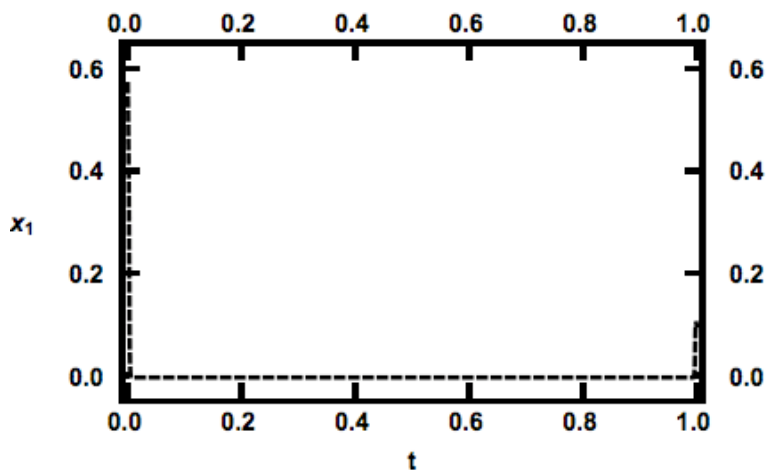


Figure 7: The  $x_1$  component of the path that minimize  $I_\varepsilon$  at  $\varepsilon = 10^{-3}$ . The path starts at  $S_1$  and is conditioned to end at  $S_2$ . The initial path contains the origin  $M_0$  and thus the ending path is dominated by  $M_0$ .

bitrarily; we have shown a symmetric case. The experiments once again illustrate points 1. and 4.

**Acknowledgements.** The authors are grateful to Eric Vanden Eijnden for helpful discussions. AMS is grateful to EPSRC and ERC for financial support.

## References

- [1] M.P.Allen and D.J.Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 1987.
- [2] P. Bolhuis, C. Dellago, P.L. Geissler, and D. Chandler. Transition path sampling: throwing ropes over rough mountain passes, in the dark. *Ann. Rev. Phys. Chem.*, 53:291–318, 2002.
- [3] D. Dürr and A. Bach. The Onsager-Machlup function as lagrangian for the most probable path of a diffusion process. *Communications in Mathematical Physics*, 160:153–170, 1978.
- [4] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland Publishing Co., Amsterdam, second edition, 1989.
- [5] F. Pinski and A.M. Stuart. Transition paths in molecules: gradient descent in pathspace. *J. Chem. Phys.*, 132:184104, 2010.
- [6] A. Braides.  $\Gamma$ -convergence for beginners, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.



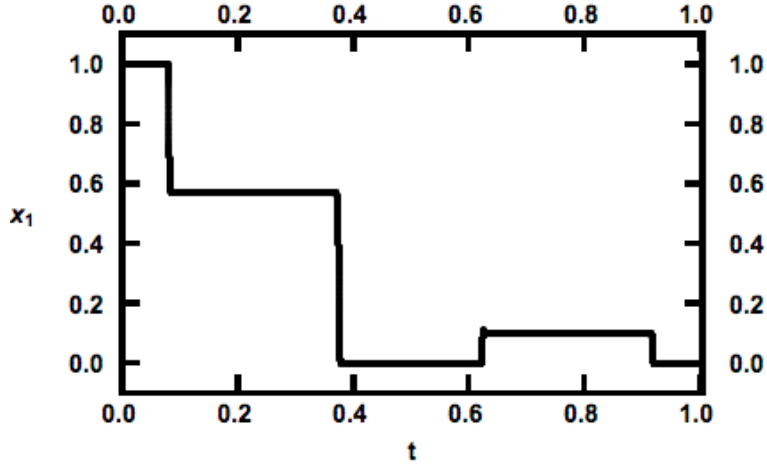


Figure 8: The  $x_1$  component of the path that minimize  $J_\varepsilon$  at  $\varepsilon = 10^{-3}$ . The path starts at  $M_1$  and is conditioned to end at  $M_2$ . The path breaks into several segments. The initial path contains the origin and thus the "ending" path passes through  $M_0$ . The ending path depends on the nature of the starting path, in particular where the path crosses the origin. Only the symmetric minimizer is displayed here.

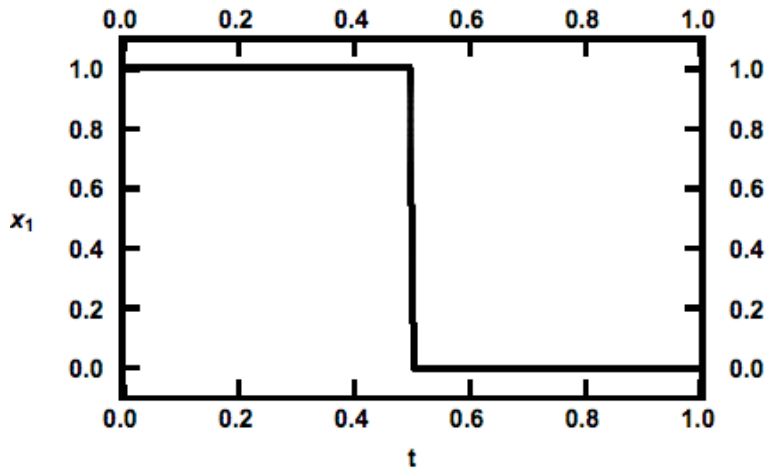


Figure 9: The  $x_1$  component of the path that minimize  $I_\varepsilon$  at  $\beta = 1000$ . The path starts at  $M_1$  and is conditioned to end at  $M_2$ . The initial path does not approach the origin  $M_0$  and thus the ending path also avoids  $M_0$ . Only the symmetric minimizer is displayed here. Note that all the "activity" is scrunched into a small time interval.

- [7] R. Olender and R. Elber. Yet another look at the steepest descent path. *J. Mol. Struct: THEOCHEM*, 63:398–399, 1997.
- [8] J. Voss. Large deviations for one dimensional diffusions with a strong drift. *Electron. J. Probab.*, 13:1479–1526, 2008.
- [9] R.V. Kohn and P. Sternberg. Local minimizers and singular perturbations. *Proc. Roy. Soc. Edinburgh*, 111:69–84, 1989.
- [10] S. Baldo. Minimal interface criterion for phase transitions in mixtures of cahn-hilliard fluids. *Ann. Inst. H. Poincaré*, 7:67–90, 1990.
- [11] M.I. Freidlin and A.D. Wentzell. *Random Perturbations of dynamical systems*. Springer-Verlag, New York, 1984.
- [12] W.E, W. Ren, and E. Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66:052301, 2002.
- [13] W.E, W. Ren, and E. Vanden-Eijnden. Transition pathways in complex systems: reaction coordinates, isocommitor surfaces and transition tubes. *Chem. Phys. Lett.*, 413:242–247, 2005.
- [14] E. Reznikoff and E. Vanden Eijnden. Invariant measures of stochastic PDEs. *C.R. Acad. Sci, Paris*, 340:305–308, 2005.
- [15] M. Hairer, A.M. Stuart, and A.M. Voss. Analysis of SPDEs arising in path sampling. part 2: The nonlinear case. *Ann. Appl. Prob.*, 340:305–308, 2007.
- [16] A.J. Chorin and O.H. Hald. *Stochastic tools in mathematics and science*, volume 1 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2006.
- [17] G. Dal Maso. *An Introduction to  $\Gamma$ -convergence*. Birkhauser, Boston, 1993.
- [18] P.L. Lions. The concentration-compactness principle in the calculus of variations. the locally compact case, part 1. *Ann. Inst. Henri Poincaré: Anal. Nonl.*, 1:109–145, 1984.