

# GAUSSIAN APPROXIMATIONS FOR TRANSITION PATHS IN BROWNIAN DYNAMICS\*

YULONG LU<sup>†</sup>, ANDREW STUART<sup>‡</sup>, AND HENDRIK WEBER<sup>†</sup>

**Abstract.** This paper is concerned with transition paths within the framework of the overdamped Langevin dynamics model of chemical reactions. We aim to give an efficient description of typical transition paths in the small temperature regime. We adopt a variational point of view and seek the best Gaussian approximation, with respect to Kullback–Leibler divergence, of the non-Gaussian distribution of the diffusion process. We interpret the mean of this Gaussian approximation as the “most likely path,” and the covariance operator as a means to capture the typical fluctuations around this most likely path. We give an explicit expression for the Kullback–Leibler divergence in terms of the mean and the covariance operator for a natural class of Gaussian approximations and show the existence of minimizers for the variational problem. Then the low temperature limit is studied via  $\Gamma$ -convergence of the associated variational problem. The limiting functional consists of two parts: The first part depends only on the mean and coincides with the  $\Gamma$ -limit of the rescaled Freidlin–Wentzell rate functional. The second part depends on both the mean and the covariance operator and is minimized if the dynamics are given by a time-inhomogeneous Ornstein–Uhlenbeck process found by linearization of the Langevin dynamics around the Freidlin–Wentzell minimizer.

**Key words.** transition path, Kullback–Leibler approximation, Onsager–Machlup functional, large deviations, gamma-convergence

**AMS subject classifications.** 28C20, 60G15, 60F10

**DOI.** 10.1137/16M1071845

**1. Introduction.** Determining the behavior of transition paths of complex molecular dynamics is essential for understanding many problems in physics, chemistry, and biology. Direct simulation of these systems can be prohibitively expensive, mainly due to the fact that the dynamical systems can exhibit the phenomenon of *metastability*, which involves disparate time scales: the transition *between* metastable states is logarithmic in the inverse temperature, while fluctuations *within* the metastable states have durations which are exponential in the inverse temperature. In many systems interest focuses on the transition between metastable states and not the local fluctuations within them. This paper addresses the problem of characterizing the most likely transition paths of molecular models of chemical reactions.

We focus on the Brownian dynamics model from molecular dynamics, which takes the form of a gradient flow in a potential, subject to small thermal fluctuations:

$$(1.1) \quad dx(t) = -\nabla V(x(t))dt + \sqrt{2\varepsilon}dW(t);$$

we study the equation subject to the end-point conditions

$$(1.2) \quad x(0) = x_-, \quad x(T) = x_+.$$

\*Received by the editors April 22, 2016; accepted for publication (in revised form) March 3, 2017; published electronically August 10, 2017.

<http://www.siam.org/journals/sima/49-4/M107184.html>

**Funding:** The work of the first author was supported by the EPSRC as part of the MASDOC DTC at the University of Warwick, through grant EP/HO23364/1. The work of the second author was supported by DARPA, the EPSRC, and the ONR. The work of the third author was supported by the EPSRC and by the Royal Society.

<sup>†</sup>Mathematics Institute, University of Warwick, Coventry, CV4 7DL, UK (yulong.lu@warwick.ac.uk, hendrik.weber@warwick.ac.uk).

<sup>‡</sup>Computing & Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125 (astuart@caltech.edu).

Here  $V : \mathbf{R}^d \rightarrow \mathbf{R}$  is the potential function,  $W$  is a standard Brownian motion in  $\mathbf{R}^d$ , and  $\varepsilon > 0$  is a small parameter related to the temperature of the thermal system. The Brownian dynamics model is widely used in the study of molecular dynamics [20]. It is also referred to as the overdamped Langevin equation and can be derived from the second order Langevin dynamics model, which has the form of damped-driven Newtonian dynamics with potential energy  $V$ , by taking a large friction or a small mass limit; see [26, Chapter 7, Exercise 8] and [20] for explicit derivations.

Mathematically we understand the process  $x(t)$ ,  $t \in [0, T]$ , satisfying (1.1), (1.2) to be the initial value problem of (1.1) starting from  $x(0) = x_-$ , subject to the conditioning  $x(T) = x_+$  (see [18]). We propose to study the sample path of this conditioned process as a model for the temporal evolution of molecules making a transition between two atomistic configurations  $x_{\pm}$ . In this paper, we will assume that  $x_{\pm}$  are critical points of  $V$ ; indeed, most interest focuses on the case where both end-points are chosen to be local minima of  $V$ .

When the temperature  $\varepsilon$  is small and when the end-point condition on  $x(T)$  is removed, typical realizations of (1.1) exhibit fluctuations around the local minima of  $V$  for long stretches of time (exponential in  $\varepsilon^{-1}$ ), while the occasional rapid transitions between different minima occur on a much shorter time scale that is only logarithmic in  $\varepsilon^{-1}$ . The difference between these time scales makes it difficult to sample transition paths when  $\varepsilon$  is small. As an alternative to direct sampling, several notions of “most likely transition paths” have been proposed; of particular interest here are the Freidlin–Wentzell and Onsager–Machlup theories.

In the zero temperature limit  $\varepsilon \rightarrow 0$ , the behavior of transition paths can be predicted with overwhelming probability using Freidlin–Wentzell theory [15]. For any fixed  $T$ , the solution process  $\{x(t), t \in [0, T]\}$  to (1.1), (1.2) satisfies a large deviation principle with rate (or action) functional given by

$$(1.3) \quad \bar{S}_T(\varphi) := \frac{1}{4} \int_0^T |\varphi'(t) + \nabla V(\varphi(t))|^2 dt$$

with  $\varphi \in H_{\pm}^1(0, T; \mathbf{R}^d) := \{x \in H^1(0, T; \mathbf{R}^d) : x(0) = x_-, x(T) = x_+\}$ . Loosely speaking, the large deviation principle states that for any small  $\delta > 0$  the probability that the solution  $x$  lies in a tube of width  $\delta$  around a given path  $\varphi$  is approximately given by

$$(1.4) \quad \mathbb{P} \left\{ x : \sup_{t \in [0, T]} |x(t) - \varphi(t)| \leq \delta \right\} \approx \exp(-\varepsilon^{-1} \bar{S}_T(\varphi))$$

for  $\varepsilon$  small enough. Here  $\mathbb{P}$  denotes the law of the process defined in (1.1), (1.2). The large deviation principle thus characterizes the exponential tail of the distribution of the transition paths, but what is of most interest to us is that it leads to a natural variational definition of the most likely path: the minimizer of the rate functional  $\bar{S}_T$  can be interpreted as a most likely path in the sense that the probability of a trajectory in a small neighborhood of this minimizer is exponentially larger in  $\varepsilon^{-1}$  than the probability of hitting neighborhoods of any other paths.

In view of the boundary conditions (1.2), one can rewrite the functional  $\bar{S}_T$  as

$$\begin{aligned}
 \bar{S}_T(\varphi) &:= \frac{1}{4} \int_0^T |\varphi'(t) + \nabla V(\varphi(t))|^2 dt \\
 (1.5) \quad &= \frac{1}{4} \int_0^T |\varphi'(t)|^2 + |\nabla V(\varphi(t))|^2 dt + \frac{1}{2} \int_0^T \varphi'(t) \cdot \nabla V(\varphi(t)) dt \\
 &= \frac{1}{4} \int_0^T |\varphi'(t)|^2 + |\nabla V(\varphi(t))|^2 dt + \frac{1}{2} ((V(x_+) - V(x_-))).
 \end{aligned}$$

The last term in this expression depends only on the boundary conditions and not on the specific choice of  $\varphi$ . Hence minimizing  $\bar{S}_T(\varphi)$  is equivalent to minimizing the following Freidlin–Wentzell functional,

$$(1.6) \quad S_T(\varphi) := \frac{1}{4} \int_0^T |\varphi'(t)|^2 + |\nabla V(\varphi(t))|^2 dt,$$

over  $H_{\pm}^1(0, T; \mathbf{R}^d)$ , and from now on we refer to the minimization of this functional as the Freidlin–Wentzell approach. The Freidlin–Wentzell viewpoint has been enormously influential in the study of chemical reactions. For example, the string method [10, 12] is based on minimization of the action functional (1.3) over paths parameterized by arc-length. See the review article [33] for recent developments in transition path theory.

At finite temperature  $\varepsilon > 0$ , optimal transition paths can be defined as minimizers of the Onsager–Machlup functional [9]. This functional is defined by maximizing small ball probabilities for paths  $x(\cdot)$  solving (1.1), (1.2). To be more precise, we denote by  $\mathbb{P}_0$  the law of the Brownian bridge on  $[0, T]$  connecting  $x_-$  and  $x_+$ , corresponding to vanishing drift ( $V = 0$ ) in (1.1), (1.2), which depends on  $\varepsilon$ . Then, under certain conditions on  $V$  (see (ii) of Remark 2.2), the measure  $\mathbb{P}$  is absolutely continuous with respect to  $\mathbb{P}_0$ , and the Radon–Nikodym density is given by

$$(1.7) \quad \frac{d\mathbb{P}}{d\mathbb{P}_0}(x) = \frac{1}{Z} \exp \left( -\frac{1}{2\varepsilon} \int_0^T \Psi_\varepsilon(x(t)) dt \right),$$

where

$$(1.8) \quad \Psi_\varepsilon(x) := \frac{1}{2} |\nabla V(x)|^2 - \varepsilon \Delta V(x).$$

Equation (1.7) follows from the Girsanov formula and Itô's formula; see [30, section 2]. We define the Onsager–Machlup functional  $I_\varepsilon$  over the space  $H_{\pm}^1(0, T; \mathbf{R}^d)$  by

$$(1.9) \quad I_\varepsilon(x) := \frac{1}{2} \int_0^T \left( \frac{1}{2} |x'(t)|^2 + \Psi_\varepsilon(x(t)) \right) dt = S_T(x) - \frac{\varepsilon}{2} \int_0^T \Delta V(x(t)) dt.$$

In [9] it was shown that for any  $x_1, x_2 \in H_{\pm}^1(0, T; \mathbf{R}^d)$

$$\lim_{\delta \rightarrow 0} \frac{\mathbb{P}(B_\delta(x_1))}{\mathbb{P}(B_\delta(x_2))} = \exp \left( \frac{1}{\varepsilon} (I_\varepsilon(x_2) - I_\varepsilon(x_1)) \right),$$

where  $B_r(x)$  denotes a ball in  $C([0, T]; \mathbf{R}^d)$  with center  $x$  and radius  $r$ . Hence for any fixed  $x_2$ , the above ratio of the small ball probability, as a function of  $x_1$ , is maximized

at minimizers of  $I_\varepsilon$ . In this sense minimizers of  $I_\varepsilon$  are analogous to *maximum a posteriori (MAP) estimators*, which arise for the posterior distribution  $\mathbb{P}$  in Bayesian inverse problems; see [8].

The Onsager–Machlup functional (1.9) differs from the Freidlin–Wentzell functional only by the integral of the Itô correction term  $\varepsilon \Delta V$ . This difference arises because of the order in which the limits  $\varepsilon \rightarrow 0$  and  $\delta \rightarrow 0$  are taken: in Freidlin–Wentzell theory the radius of the ball  $\delta$  is fixed and limit  $\varepsilon \rightarrow 0$  is studied, while in Onsager–Machlup theory  $\varepsilon$  is fixed and limit  $\delta \rightarrow 0$  is studied. For fixed  $T > 0$ , it is clear that  $I_\varepsilon(\varphi) \rightarrow S_T(\varphi)$  as  $\varepsilon \rightarrow 0$ . Hence for fixed time scale  $T$  the Onsager–Machlup theory agrees with the Freidlin–Wentzell theory in the low temperature limit. However, this picture can be different for large  $T$ , more precisely when  $T \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ . In fact, as demonstrated in [29], it is possible that when  $T \gg 1$ , the MAP transition path spends a vast amount of time at a saddle point of  $V$  rather than at minima; moreover, for two paths with the same energy barrier, the one passing through steeper confining walls is always preferred to the other, since a larger value of  $\Delta V$  gives rise to a lower value of  $I_\varepsilon$ . The discussion about the order of limits gives a clue as to why this apparent contradiction occurs: by studying the limit  $\delta \rightarrow 0$  in Onsager–Machlup theory, for fixed temperature  $\varepsilon$ , we remove entropic effects.

Both minimizing the Onsager–Machlup functional (1.9) and finding MAP estimators are attempts to capture key properties of the distribution  $\mathbb{P}$  by identifying a single most likely path. This can be viewed as approximating the measure  $\mathbb{P}$  by a Dirac measure in a well-chosen point. The key idea in this paper is to find better approximations to  $\nu$  by working in a class of measures larger than Diracs. We will study the best Gaussian approximations with respect to Kullback–Leibler divergence. The mean of an optimal Gaussian should capture the concentration of the target measure, while its fluctuation characteristics are described by the covariance of the Gaussian. Furthermore the fluctuations can capture entropic effects. Thus by using the Gaussian approximation, we aim to overcome the shortcomings of the Onsager–Machlup approach. The idea of finding Gaussian approximations for non-Gaussian measures by means of the Kullback–Leibler divergence is not new. For example, in the community of machine learning [31], Gaussian processes have been widely used together with Bayesian inference for regression and prediction. Similar ideas have also been used to study models in ocean-atmosphere science [22] and computational quantum mechanics [2]. Recently, the problem of minimizing the Kullback–Leibler divergence between non-Gaussian measures and certain Gaussian classes was studied from the calculus of variations point of view [28], and numerical algorithms for Kullback–Leibler minimization were discussed in [27].

The present paper builds on the theory developed in [28] and extends it to transition path theory. More specifically, the set of Gaussian measures for approximations is parameterized by a pair of functions  $(m, \mathbf{A})$ , where  $m$  represents the mean and  $\mathbf{A}$  (defined in (3.1)) is used to define the covariance operator for the underlying Gaussian measure. For a fixed temperature  $\varepsilon$ , the Kullback–Leibler divergence is expressed as a functional  $F_\varepsilon$  depending on  $(m, \mathbf{A})$ , and the existence of minimizers is shown in this framework. Then the asymptotic behavior of the best Gaussian approximations in the low temperature limit is studied in terms of the  $\Gamma$ -convergence of the functionals  $\{F_\varepsilon\}$ . The limiting functional (4.14) is identified as the sum of two parts. The first part, depending only on  $m$ , is identical to the  $\Gamma$ -limit of the rescaled Freidlin–Wentzell action functional, implying that for  $\varepsilon \rightarrow 0$  the most likely transition paths defined as the best Gaussian mean  $m$  coincide with large deviation paths. The second part takes entropic effects into account and expresses the penalty for the fluctuations in

terms of  $\mathbf{A}$ ; it vanishes if  $\mathbf{A} = D^2V(m(t))$ , but this choice of  $\mathbf{A}$  is admissible only if the Hessian  $D^2V(m)$  is positive definite. A strictly positive penalty occurs when  $D^2V(m(t))$  has a negative eigenvalue. Therefore minimizing the limiting functional amounts to selecting those optimal paths  $m$  among the large deviation paths that do not spend time in saddles or local maximizers. We stress that although at finite noise intensity  $\varepsilon > 0$  there is no explicit characterization of our most likely transition paths, it is possible to approximately determine them numerically, as demonstrated in [27]; see also section 5.

This paper is organized as follows. In the next section we introduce a time-rescaling of the governing Langevin equation, in terms of  $\varepsilon$ , in which the undesirable effects of the Onsager–Machlup minimization are manifest; we also introduce some notation used throughout the paper. Furthermore, assumptions on the potential  $V$  are discussed. In section 3, we define the subset of Gaussian measures over which Kullback–Leibler minimization is conducted; the existence of minimizers to the variational problem is established at the end of this section. Then in section 4, we study the low temperature limit of the Gaussian approximation using  $\Gamma$ -convergence. The main  $\Gamma$ -convergence result is given in Theorem 4.6. Section 5 discusses some important consequences of the  $\Gamma$ -convergence result, with emphasis on the link to theories of Freidlin and Wentzell and Onsager and Machlup. The proofs of Theorem 4.6 and some related results are presented in section 6.

## 2. Setup and notation.

**2.1. Setup.** As discussed in the previous section, the key issue which motivates our work is the difference in behavior between minimizers of the Freidlin–Wentzell action and the Onsager–Machlup functional. This difference is manifest when  $T \gg 1$  and is most cleanly described by considering the time scale  $T = \varepsilon^{-1}$ . The  $\Gamma$ -limit of the Onsager–Machlup functional (1.9) is studied, as  $\varepsilon \rightarrow 0$ , under this time-rescaling, in [30]; the limit exhibits the undesirable effects described in the preceding section. Our objective is to characterize the  $\Gamma$ -limit for the variational problems arising from best Gaussian approximation with respect to Kullback–Leibler divergences, under the same limiting process.

Applying the time scaling  $t \mapsto \varepsilon^{-1}t$  to (1.1) and noticing the boundary conditions (1.2) yields

$$(2.1) \quad \begin{aligned} dx(t) &= -\varepsilon^{-1} \nabla V(x(t)) dt + \sqrt{2} dW(t), \\ x(0) &= x_-, \quad x(1) = x_+. \end{aligned}$$

The transformed SDE has an order one noise but a strong drift; it will be our object of study throughout the remainder of the paper. For technical reasons, we make the following assumptions on the potential  $V$ .

*Assumptions 2.1.* The potential  $V$  appearing in (2.1) satisfies the following:

(A-1)  $V \in C^5(\mathbf{R}^d)$ .

(A-2) The set of critical points

$$(2.2) \quad \mathcal{E} := \{x \in \mathbf{R}^d, \nabla V(x) = 0\}$$

is finite.

(A-3) The coercivity condition:

$$(2.3) \quad \exists R > 0 \quad \text{such that} \quad \inf_{|x| > R} |\nabla V(x)| > 0.$$

(A-4) The growth condition:

$$(2.4) \quad \exists C_1, C_2 > 0 \text{ and } \alpha \in [0, 2) \text{ such that } \forall x \in \mathbf{R}^d \text{ and } 1 \leq i, j, k \leq d, \\ \limsup_{\varepsilon \rightarrow 0} \max \left( \left| \frac{\partial^3}{\partial x_i \partial x_j \partial x_k} \Psi_\varepsilon(x) \right|, |\Psi_\varepsilon(x)| \right) \leq C_1 e^{C_2 |x|^\alpha}.$$

(A-5)  $V(x) \rightarrow \infty$  when  $|x| \rightarrow \infty$ , and there exists  $R > 0$  such that

$$(2.5) \quad 2\Delta V(x) \leq |\nabla V(x)|^2 \quad \text{for } |x| \geq R.$$

(A-6) The monotonicity condition:

$$(2.6) \quad \exists R > 0 \quad \text{such that } |\nabla V(x_1)| \geq |\nabla V(x_2)| \quad \text{if } |x_1| \geq |x_2| \geq R.$$

*Remark 2.2.*

- (i) Conditions (A-2)–(A-3) are typical assumptions for proving  $\Gamma$ -convergence results for Ginzburg–Landau and related functionals [14, 21]. The smoothness condition (A-1) is needed because our analysis requires the third order Taylor expansion for  $\Psi_\varepsilon$ . Furthermore, we will use conditions (A-4)–(A-6) to analyze the  $\Gamma$ -convergence problem in this paper. These assumptions will be employed to simplify the expectation term in the Kullback–Leibler divergence (see the expression (3.24)).
- (ii) The condition (A-5) is a Lyapunov-type condition which guarantees that at small temperature ( $\varepsilon \leq 1$ ) the solution to the SDE in (2.1) does not explode in finite time. The probability measure determined by this process is absolutely continuous with respect to the reference measure of the Brownian bridge. See [32, Chapter 2] for more discussions about the absence of explosion. Moreover, by the definition of  $\Psi_\varepsilon$ , (A-5) implies that for any  $\delta \in \mathbf{R}$  there exists a constant  $C > 0$  depending only  $R$  and  $\delta$  such that

$$(2.7) \quad |\nabla V(x)|^2 - \varepsilon \delta \Delta V(x) \geq -C\varepsilon \quad \text{for any } x \in \mathbf{R}^d.$$

Such a lower bound will be used to prove the compactness of the functionals of interest (see Proposition 4.5).

- (iii) These conditions are not independent. For instance, the coercivity condition (A-3) can be deduced from the monotonicity condition (A-6) when  $V(x)$  is nonconstant for large  $|x|$ . Hence particularly (A-5) and (A-6) imply (A-3).
- (iv) The set of functions satisfying conditions (A-1)–(A-7) is not empty: they are fulfilled by all polynomials. Therefore many classical potentials, such as the Ginzburg–Landau double-well potential  $V(x) = \frac{1}{4}x^2(1-x)^2$ , are included.

For  $\varepsilon > 0$  we denote by  $\mu_\varepsilon$  the law of the above bridge process  $x$  defined in (2.1), and by  $\mu_0$  the law of the corresponding bridge for vanishing drift ( $V = 0$ ) in (2.1). Then, by arguments identical to those yielding (1.7),  $\mu_\varepsilon$  is absolutely continuous with respect to  $\mu_0$ , and the Radon–Nikodym density is given by

$$(2.8) \quad \frac{d\mu_\varepsilon}{d\mu_0}(x) = \frac{1}{Z_{\mu,\varepsilon}} \exp \left( -\frac{1}{2\varepsilon^2} \int_0^1 \Psi_\varepsilon(x(t)) dt \right),$$

where  $\Psi_\varepsilon$  is given by (1.8) and  $Z_{\mu,\varepsilon}$  is the normalization constant. Note that the extra factor  $\frac{1}{\varepsilon}$  with respect to (1.7) is due to the time rescaling.

**2.2. Notation.** Throughout the paper, we use  $C$  (or occasionally  $C_1$  and  $C_2$ ) to denote a generic positive constant which may change from one expression to the next and is independent of the temperature and any quantity of interest. We write  $A \lesssim B$  if  $A \leq CB$ . Given an interval  $I \subset \mathbf{R}$ , let  $L^p(I)$  and  $W^{m,p}(I)$  with  $m \in \mathbf{N}$ ,  $1 \leq p \leq \infty$  be the standard Lebesgue and Sobolev spaces, respectively, of scalar functions. Let  $H^m(I) = W^{m,2}(I)$ . For  $s \in [0, 1]$ , we set  $H_0^s(I)$  to be the closure of  $C_0^\infty(I)$  in  $H^s(I)$  and equip it with the topology induced by  $H^s(I)$ . Define its dual space  $H^{-s}(I) := (H_0^s(I))'$ . For  $s > 1/2$ , a function of  $H_0^s(I)$  has zero boundary conditions. Thanks to the Poincaré inequality, the  $H^1$ -seminorm is an equivalent norm on  $H_0^1(I)$ . In the case that  $I = (0, 1)$ , we simplify the notations by setting  $H_0^s = H_0^s(0, 1)$  and  $H^{-s} = H^{-s}(0, 1)$ .

We write scalar and vector variables in regular face, whereas matrix-valued variables, function spaces for vectors and matrices are written in boldface. Denote by  $\mathcal{S}(d, \mathbf{R})$  the set of all real symmetric  $d \times d$  matrices, and by  $\mathbf{I}_d$  the identity matrix of size  $d$ . Let  $L^p(0, 1; \mathbf{R}^d)$  and  $L^p(0, 1; \mathcal{S}(d, \mathbf{R}))$  be the spaces of vector-valued and symmetric matrix-valued functions, respectively, with entries in  $L^p(0, 1)$ . Similarly one can define  $H^1(0, 1; \mathbf{R}^d)$ ,  $H_0^s(0, 1; \mathbf{R}^d)$ , and  $H^1(0, 1; \mathcal{S}(d, \mathbf{R}))$ . For simplicity, we use the same notation  $\mathbf{L}^p(0, 1)$  (resp.,  $\mathbf{H}^1(0, 1)$ ) to denote  $L^p(0, 1; \mathcal{S}(d, \mathbf{R}))$  and  $L^p(0, 1; \mathbf{R}^d)$  (resp.,  $H^1(0, 1; \mathcal{S}(d, \mathbf{R}))$  and  $H^1(0, 1; \mathbf{R}^d)$ ). For any  $\mathbf{A} = (A_{ij}) \in L^p(0, 1; \mathcal{S}(d, \mathbf{R}))$  with  $1 \leq p \leq \infty$  we define its norm as

$$\|\mathbf{A}\|_{\mathbf{L}^p(0,1)} := \left( \sum_{i=1}^d \sum_{j=1}^d \|A_{ij}\|_{L^p(0,1)}^2 \right)^{\frac{1}{2}}.$$

For  $\mathbf{A} = (A_{ij}) \in H^1(0, 1; \mathcal{S}(d, \mathbf{R}))$  the norm is defined by

$$\|\mathbf{A}\|_{\mathbf{H}^1(0,1)} := \left( \sum_{i=1}^d \sum_{j=1}^d \|A_{ij}\|_{H^1(0,1)}^2 \right)^{\frac{1}{2}}.$$

We also define  $\mathbf{H}_\pm^1(0, 1) := H_\pm^1(0, 1; \mathbf{R}^d) := \{x \in H^1(0, 1; \mathbf{R}^d) : x(0) = x_-, x(1) = x_+\}$ . Denote by  $\mathbf{BV}(I)$  the set of  $\mathbf{R}^d$ -valued functions of bounded variations on an interval  $I \subset \mathbf{R}$ .

For matrices  $\mathbf{A}, \mathbf{B} \in \mathcal{S}(d, \mathbf{R})$  we write  $\mathbf{A} \geq \mathbf{B}$  when  $\mathbf{A} - \mathbf{B}$  is positive semidefinite. The trace of a matrix  $\mathbf{A}$  is denoted by  $\text{Tr}(\mathbf{A})$ . Denote by  $\mathbf{A}^T$  the transpose of  $\mathbf{A}$ , and by  $|\mathbf{A}|_F$  the Frobenius norm of  $\mathbf{A}$ . Given  $\mathbf{A} \in \mathcal{S}(d, \mathbf{R})$  with the diagonalized form  $\mathbf{A} = \mathbf{P}^T \mathbf{\Lambda} \mathbf{P}$ , we define the matrix  $|\mathbf{A}| := \mathbf{P}^T |\mathbf{\Lambda}| \mathbf{P}$ . For matrices  $\mathbf{A} = (A_{ij})$  and  $\mathbf{B} = (B_{ij})$  we write

$$\mathbf{A} : \mathbf{B} = \text{Tr}(\mathbf{A}\mathbf{B}^T) = \sum_{i=1}^d \sum_{j=1}^d A_{ij} B_{ij}.$$

Define the matrix-valued operator  $\partial_t^2 := \partial_t^2 \cdot \mathbf{I}_d$ . For  $a > 0$  we define

$$\mathbf{L}_a^1(0, 1) := L_a^1(0, 1; \mathcal{S}(d, \mathbf{R})) = \{\mathbf{A} \in L^1(0, 1; \mathcal{S}(d, \mathbf{R})) : \mathbf{A}(t) \geq a \cdot \mathbf{I}_d \text{ a.e. on } (0, 1)\}$$

and

$$\mathbf{H}_a^1(0, 1) := H_a^1(0, 1; \mathcal{S}(d, \mathbf{R})) = \{\mathbf{A} \in H^1(0, 1; \mathcal{S}(d, \mathbf{R})) : \mathbf{A}(t) \geq a \cdot \mathbf{I}_d \text{ a.e. on } (0, 1)\}.$$

We write  $\mathbf{A}_n \rightharpoonup \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$  when  $\mathbf{A}_n$  converges to  $\mathbf{A}$  weakly in  $\mathbf{L}^1(0, 1)$ . Let  $\mathbf{H}_0^1(0, 1) = H_0^1(0, 1; \mathbf{R}^d)$ . Define

$$\mathbf{H}_0^s = \overbrace{H_0^s \times \cdots \times H_0^s}^d,$$

and let  $\mathbf{H}^{-s}$  be the dual. In addition, we define product spaces  $\mathcal{H} := \mathbf{H}_\pm^1(0, 1) \times \mathbf{H}^1(0, 1)$ ,  $\mathcal{H}_a := \mathbf{H}_\pm^1(0, 1) \times \mathbf{H}_a^1(0, 1)$ ,  $\mathcal{X} := \mathbf{L}^1(0, 1) \times \mathbf{L}^1(0, 1)$ , and  $\mathcal{X}_a := \mathbf{L}^1(0, 1) \times \mathbf{L}_a^1(0, 1)$ .

For a vector field  $v = (v_1, v_2, \dots, v_d)$ , let  $\nabla v = (\partial_i v_j)_{i,j=1,2,\dots,d}$  be its gradient, which is a second order tensor (or matrix). Given a potential  $V : \mathbf{R}^d \rightarrow \mathbf{R}$ , denote by  $D^2V$  the Hessian of  $V$ . Given a second order tensor  $\mathbf{T} = (T_{ij})_{i,j=1,2,\dots,d}$ , we denote by  $\nabla \mathbf{T}$  its gradient, which is a rank 3 tensor with  $(\nabla \mathbf{T})_{ijk} = \frac{\partial T_{ij}}{\partial x_k}$ . In particular, we use  $D^3V$  to denote the gradient of the Hessian  $D^2V$ .

Finally, we write  $\nu \ll \mu$  when the measure  $\nu$  is absolutely continuous with respect to  $\mu$ , and write  $\nu \perp \mu$  when both are singular. Throughout the paper, we denote by  $N(m, \Sigma)$  the Gaussian measure on  $\mathbf{L}^2(0, 1)$  with mean  $m$  and covariance operator  $\Sigma$ . Moreover, the Gaussian measures considered in the paper will always have the property that, almost surely, draws from the measure continuous functions on  $[0, 1]$  and thus that pointwise evaluation is well-defined. Given  $h \in \mathbf{L}^2(0, 1)$ , define the translation map  $\mathcal{T}_h$  by setting  $\mathcal{T}_h x = x + h$  for any  $x \in \mathbf{L}^2(0, 1)$ . Denote by  $\mathcal{T}_h^* \mu$  the push-forward measure of a measure  $\mu$  on  $\mathbf{L}^2(0, 1)$  under the map  $\mathcal{T}_h$ .

### 3. Kullback–Leibler minimization.

**3.1. Parametrization of Gaussian measures.** In this subsection, we describe the parametrization of the Gaussian measures that we use in our Kullback–Leibler minimization. To motivate our choice of parameterization we consider the SDE (2.1). This equation has order one noise, but with a strong gradient-form drift which will, most of the time, constrain the sample path to the neighborhood of critical points of  $V$ . The size of the neighborhood will be defined by small fluctuations whose size scales with  $\varepsilon^{\frac{1}{2}}$ . To capture this behavior we seek an approximation to (2.1) of the form  $x = m + z$ , where  $m$  is a path connecting  $x_\pm$  in unit time and where  $z$  describes the small fluctuations. We aim to find  $m$  from an appropriate class of functions, and  $z$  as time-inhomogeneous Ornstein–Uhlenbeck process

$$(3.1) \quad \begin{aligned} dz(t) &= -\varepsilon^{-1} \mathbf{A}(t) z(t) dt + \sqrt{2} dW(t), \\ z(0) &= z(1) = 0. \end{aligned}$$

The time-dependent functions  $(m, \mathbf{A})$  become our unknowns. For subsequent discussions, we require  $m \in \mathbf{H}_\pm^1(0, 1)$ . For  $\mathbf{A}$  we assume that  $\mathbf{A} \in \mathbf{H}^1(0, 1)$ ; i.e.,  $\mathbf{A} \in H^1(0, 1; \mathbf{R}^{d \times d})$  and  $\mathbf{A}(t)$  is symmetric for any  $t \in (0, 1)$ . The symmetry property will simplify the calculation of the change of measures below and will also help when estimating the Greens functions used to show the  $\Gamma$ -convergence in section 4.

Let  $\bar{\nu}_\varepsilon$  be the distribution of the process  $z$  defined by (3.1), and let  $\bar{\mu}_0$  be the corresponding Brownian bridge (with  $\mathbf{A} = 0$ ). The lemma below shows that  $\bar{\nu}_\varepsilon$  is a centered Gaussian with the covariance operator given by the inverse Schrödinger operator  $\Sigma_\varepsilon := 2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1}$  with  $\mathbf{B}_\varepsilon = \varepsilon^{-2} \mathbf{A}^2 - \varepsilon^{-1} \mathbf{A}'$ . Here  $2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1}$  denotes the inverse of the Schrödinger operator  $\frac{1}{2}(-\partial_t^2 + \mathbf{B}_\varepsilon)$  with Dirichlet boundary condition. Let  $\mathbf{M}_\varepsilon(t; s)$  be the fundamental matrix satisfying

$$(3.2) \quad \frac{d}{dt} \mathbf{M}_\varepsilon(t, s) = -\varepsilon^{-1} \mathbf{A}(t) \mathbf{M}_\varepsilon(t, s), \quad \mathbf{M}_\varepsilon(s, s) = \mathbf{I}_d.$$



LEMMA 3.1. Let  $A \in \mathbf{H}^1(0, 1)$ . Then the Radon–Nikodym density of  $\bar{\nu}_\varepsilon$  with respect to  $\bar{\mu}_0$  is given by

$$(3.3) \quad \frac{d\bar{\nu}_\varepsilon}{d\bar{\mu}_0}(z) = \frac{1}{Z_{\bar{\nu}, \varepsilon}} \exp \left( -\frac{1}{4} \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt \right),$$

where  $\mathbf{B}_\varepsilon = \varepsilon^{-2} \mathbf{A}^2 - \varepsilon^{-1} \mathbf{A}'$ , and the normalization constant

$$(3.4) \quad Z_{\bar{\nu}, \varepsilon} = \exp \left( -\frac{1}{2\varepsilon} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt \right) \cdot \left( \int_0^1 \bar{\mathbf{M}}_\varepsilon(t) \bar{\mathbf{M}}_\varepsilon^T(t) dt \right)^{-1/2},$$

where  $\bar{\mathbf{M}}_\varepsilon(t) = \mathbf{M}_\varepsilon(1, t)$ . It follows that  $\bar{\nu}_\varepsilon = N(0, 2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1})$ .

*Proof.* Let  $z$  be the unconditioned Ornstein–Uhlenbeck process that satisfies

$$(3.5) \quad dz(t) = -\varepsilon^{-1} \mathbf{A}(t) z(t) dt + \sqrt{2} dW(t), \quad z(0) = 0.$$

Denote by  $\tilde{\nu}_\varepsilon$  the law of  $z(t)$ ,  $t \in [0, 1]$ , solving (3.5), and by  $\tilde{\mu}_0$  the law of the process  $\sqrt{2}W(t)$ . It follows from Girsanov's theorem that

$$(3.6) \quad \frac{d\tilde{\nu}_\varepsilon}{d\tilde{\mu}_0}(z) = \exp \left( -\frac{1}{2\varepsilon} \int_0^1 \mathbf{A}(t) z(t) \cdot dz(t) - \frac{1}{4\varepsilon^2} \int_0^1 |\mathbf{A}(t) z(t)|^2 dt \right).$$

Simplifying the exponent on the right-hand side of the above by Itô's formula gives

$$(3.7) \quad \frac{d\tilde{\nu}_\varepsilon}{d\tilde{\mu}_0}(z) = \exp \left( -\frac{1}{4} \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt + \frac{1}{2\varepsilon} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt - \frac{1}{4\varepsilon} z(1)^T \mathbf{A}(1) z(1) \right).$$

After conditioning on  $z(1) = 0$  and using [18, Lemma 5.3], (3.3) follows from (3.7). We now calculate the normalization constant  $Z_{\bar{\nu}, \varepsilon}$ . Let  $\rho_1$  be the density of the distribution of  $z(1)$  under the measure  $\bar{\nu}_\varepsilon$ . Let  $\tilde{\mu}_y$  be law of the conditioned process  $(\sqrt{2}W(t) | \sqrt{2}W(1) = y)$ . From (3.7), one can see that for any bounded measurable function  $f : \mathbf{R}^d \rightarrow \mathbf{R}$ ,

$$(3.8) \quad \begin{aligned} \mathbb{E}^{\rho_1}[f(z(1))] &= \mathbb{E}^{\bar{\nu}_\varepsilon}[f(z(1))] \\ &= \mathbb{E}^{\tilde{\mu}_0} \left[ \exp \left( -\frac{1}{4} \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt \right) \right. \\ &\quad \times \exp \left( \frac{1}{2\varepsilon} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt - \frac{1}{4\varepsilon} z(1)^T \mathbf{A}(1) z(1) \right) f(z(1)) \Big] \\ &= \frac{1}{(4\pi)^{d/2}} \int_{\mathbf{R}^d} \exp \left( \frac{1}{2\varepsilon} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt - \frac{1}{4\varepsilon} y^T \mathbf{A}(1) y - \frac{1}{4} |y|^2 \right) f(y) \\ &\quad \times \mathbb{E}^{\tilde{\mu}_y} \left[ \exp \left( -\frac{1}{4} \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt \right) \right] dy, \end{aligned}$$

where we have used the fact that  $z(1) \sim N(0, 2 \cdot \mathbf{I}_d)$  when  $z$  is distributed according to  $\tilde{\mu}_0$ . Then we can read from (3.8) that

$$(3.9) \quad \rho_1(0) = \mathbb{E}^{\bar{\mu}_0} \left[ \exp \left( -\frac{1}{4} \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt \right) \right] \exp \left( \frac{1}{2\varepsilon} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt \right) \frac{1}{(4\pi)^{d/2}}.$$

On the other hand, we know from Appendix B that the solution of (3.5) can be represented as

$$z(t) = \sqrt{2} \int_0^t \mathbf{M}_\varepsilon(t, s) dW(s),$$

where  $\mathbf{M}_\varepsilon$  is the fundamental matrix. In particular, by Itô's isometry the random variable  $z(1)$  is a centered Gaussian with covariance

$$\mathbb{E}[z(1)z(1)^T] = 2 \int_0^1 \overline{\mathbf{M}}_\varepsilon(t) \mathbf{M}_\varepsilon(t)^T dt,$$

where  $\overline{\mathbf{M}}_\varepsilon(t) = \mathbf{M}_\varepsilon(1, t)$ . Therefore we obtain an alternative expression for  $\rho_1$ , namely

$$(3.10) \quad \rho_1(0) = \frac{1}{(4\pi)^{d/2}} \left[ \det \left( \int_0^1 \overline{\mathbf{M}}_\varepsilon(t) \mathbf{M}_\varepsilon(t)^T dt \right) \right]^{-\frac{1}{2}}.$$

Comparing the expressions (3.9) and (3.10) yields (3.4). Finally, by the same arguments used in the proof of [28, Lemma C.1], one can see that  $\bar{\nu}_\varepsilon = N(0, 2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1})$ .  $\square$

We note that the covariance operator  $\Sigma_\varepsilon = 2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1}$  is bounded from  $\mathbf{L}^2(0, 1)$  to  $\mathbf{H}^2(0, 1)$  and is trace-class on  $\mathbf{L}^2(0, 1)$ ; see Lemma C.5 and Remark C.6. The sample paths  $z$  are almost surely continuous, and the covariances are given by

$$(3.11) \quad \mathbb{E}^{\bar{\nu}_\varepsilon}[z(t)^T z(s)] = 2\mathbf{G}_\varepsilon(t, s), \quad t, s \in [0, 1].$$

Here  $\mathbf{G}_\varepsilon(t, s)$  is the Green's tensor (fundamental matrix) of the elliptic operator  $(-\partial_t^2 + \mathbf{B}_\varepsilon)$  under Dirichlet boundary conditions; i.e., for any  $s \in (0, 1)$ ,

$$(3.12) \quad \begin{aligned} (-\partial_t^2 + \varepsilon^{-2} \mathbf{A}^2(\cdot) - \varepsilon^{-1} \mathbf{A}'(\cdot)) \mathbf{G}_\varepsilon(\cdot, s) &= \delta(\cdot - s) \cdot \mathbf{I}_d, \\ \mathbf{G}_\varepsilon(0, s) &= \mathbf{G}_\varepsilon(1, s) = 0. \end{aligned}$$

With a description of the centered fluctuation process  $z$  in hand, we now move on to discuss the noncentered process  $x = m + z$ , whose law is denoted by  $\nu_\varepsilon$ . It is clear that  $\nu_\varepsilon = N(m, \Sigma_\varepsilon)$ . Because of (3.1),  $\nu_\varepsilon$  can also be viewed as the law of the following conditioned Ornstein–Uhlenbeck process:

$$(3.13) \quad \begin{aligned} dx(t) &= (m'(t) - \varepsilon^{-1} \mathbf{A}(t)(x(t) - m(t))) dt + \sqrt{2} dW(t), \\ x(0) &= x_-, \quad x(1) = x_+. \end{aligned}$$

Hence the Gaussian measure  $\nu_\varepsilon$  is parametrized by the pair of functions  $(m, \mathbf{A})$ . To conclude, recalling the space  $\mathcal{H} = \mathbf{H}_\pm^1(0, 1) \times \mathbf{H}_a^1(0, 1)$ , we define the family of Gaussian measures as

$$(3.14) \quad \mathcal{A} = \left\{ N(m, 2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1}) : (m, \mathbf{A}) \in \mathcal{H} \right\},$$

where  $\mathbf{B}_\varepsilon = \varepsilon^{-2} \mathbf{A}^2 - \varepsilon^{-1} \mathbf{A}'$ . For  $a > 0$ , we denote by  $\mathcal{A}_a$  the set of Gaussian measures defined in the same way as (3.14) but with  $\mathcal{H}$  replaced by  $\mathcal{H}_a$ .

**3.2. Calculations of Kullback–Leibler divergence.** To quantify the closeness of probability measures, we use the Kullback–Leibler divergence, or relative entropy. Given two probability measures  $\nu$  and  $\mu$ , with  $\nu$  absolutely continuous with respect to  $\mu$ , the Kullback–Leibler divergence of  $\nu$  and  $\mu$  is

$$D_{\text{KL}}(\nu||\mu) = \mathbb{E}^\nu \log \left( \frac{d\nu}{d\mu} \right),$$

where  $\mathbb{E}^\nu$  denotes the expectation taken with respect to the measure  $\nu$ ; if  $\nu$  is not absolutely continuous with respect to  $\mu$ , then the Kullback–Leibler divergence is defined as  $+\infty$ . Sometimes it is convenient to evaluate the Kullback–Leibler divergence through a reference measure  $\mu_0$ . If the measures  $\mu$ ,  $\nu$ , and  $\mu_0$  are mutually equivalent, then the Kullback–Leibler divergence can be expressed as

$$(3.15) \quad D_{\text{KL}}(\nu||\mu) = \mathbb{E}^\nu \log \left( \frac{d\nu}{d\mu_0} \right) - \mathbb{E}^\nu \log \left( \frac{d\mu}{d\mu_0} \right).$$

In this section, we calculate the Kullback–Leibler divergence between the non-Gaussian measure  $\mu_\varepsilon$  (defined by (2.8)) and the parametrized Gaussian measure  $\nu_\varepsilon = N(m, \Sigma_\varepsilon)$ . Recall that  $\bar{\nu}_\varepsilon$  is the law of the time-inhomogeneous Ornstein–Uhlenbeck process (3.1). Recall also that  $\mu_0$  is the law of the Brownian bridge process corresponding to vanishing drift in the SDE (2.1). It is clear that  $\mu_0 = N(m_0, 2(-\partial_t^2)^{-1})$  with  $m_0(t) = x_-(1-t) + x_+t$ . In order to evaluate the above Kullback–Leibler divergence by using (3.15), we need to calculate the Radon–Nikodym derivative  $d\nu_\varepsilon/d\mu_0$ .

**LEMMA 3.2.** *Let  $m \in \mathbf{H}_\pm^1(0, 1)$  and  $\mathbf{A} \in \mathbf{H}^1(0, 1)$ . Then the Radon–Nikodym density of  $\nu_\varepsilon$  with respect to  $\mu_0$  is given by*

$$(3.16) \quad \frac{d\nu_\varepsilon}{d\mu_0}(x) = \frac{1}{Z_{\nu,\varepsilon}} \exp(-\Phi_{\nu,\varepsilon}(x)),$$

where

$$(3.17) \quad \begin{aligned} \Phi_{\nu,\varepsilon}(x) &= \frac{1}{4} \int_0^1 (x(t) - m(t))^T \mathbf{B}_\varepsilon(t) (x(t) - m(t)) dt \\ &\quad - \frac{1}{2} \int_0^1 m'(t) \cdot dx(t) + \frac{1}{4} \int_0^1 |m'(t)|^2 dt, \end{aligned}$$

and by the normalization constant

$$(3.18) \quad Z_{\nu,\varepsilon} = \exp \left( \frac{|x_1 - x_0|^2}{4} \right) \cdot \exp \left( -\frac{1}{2\varepsilon} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt \right) \cdot \left( \int_0^1 \bar{\mathbf{M}}_\varepsilon(t) \bar{\mathbf{M}}_\varepsilon^T(t) dt \right)^{-1/2},$$

where  $\bar{\mathbf{M}}_\varepsilon(t) = \mathbf{M}_\varepsilon(1, t)$ .

*Proof.* First by definitions of  $\bar{\nu}_\varepsilon$  and  $\bar{\mu}_0$ , we know that  $\nu_\varepsilon = \mathcal{T}_m^* \bar{\nu}_\varepsilon$  and  $\mu_0 = \mathcal{T}_{m_0}^* \bar{\mu}_0$ . Then we have

$$(3.19) \quad \frac{d\nu_\varepsilon}{d\mu_0}(x) = \frac{d\mathcal{T}_m^* \bar{\nu}_\varepsilon}{d\mathcal{T}_{m_0}^* \bar{\mu}_0}(x) = \frac{d\mathcal{T}_m^* \bar{\nu}_\varepsilon}{d\mathcal{T}_m^* \bar{\mu}_0}(x) \cdot \frac{d\mathcal{T}_m^* \bar{\mu}_0}{d\mathcal{T}_{m_0}^* \bar{\mu}_0}(x).$$

Observe that for any Borel set  $A \subset \mathbf{L}^2(0, 1)$ ,

$$\mathcal{T}_m^* \bar{\nu}_\varepsilon(A) = \bar{\nu}_\varepsilon(A - m) = \mathbb{E}^{\bar{\mu}_0} \left[ \frac{d\bar{\nu}_\varepsilon}{d\bar{\mu}_0}(x) \mathbf{1}_{A-m}(x) \right] = \mathbb{E}^{\mathcal{T}_m^* \bar{\mu}_0} \left[ \frac{d\bar{\nu}_\varepsilon}{d\bar{\mu}_0}(x - m) \mathbf{1}_A(x) \right].$$

This together with Lemma 3.1 implies that

$$(3.20) \quad \begin{aligned} \frac{d\mathcal{T}_m^* \bar{\nu}_\varepsilon}{d\mathcal{T}_m^* \bar{\mu}_0}(x) &= \frac{d\bar{\nu}_\varepsilon}{d\bar{\mu}_0}(x - m) \\ &= \frac{1}{Z_{\bar{\nu}, \varepsilon}} \exp \left( -\frac{1}{4} \int_0^1 (x(t) - m(t))^T \mathbf{B}_\varepsilon(t) (x(t) - m(t)) dt \right). \end{aligned}$$

Since  $m \in \mathbf{H}_\pm^1(0, 1)$ ,  $m - m_0 \in \mathbf{H}_0^1(0, 1)$  and hence  $\mathcal{T}_m^* \bar{\mu}_0 \ll \mathcal{T}_{m_0}^* \bar{\mu}_0$ . Furthermore, by the Cameron–Martin formula we have

$$(3.21) \quad \begin{aligned} \frac{d\mathcal{T}_m^* \bar{\mu}_0}{d\mathcal{T}_{m_0}^* \bar{\mu}_0}(x) &= \exp \left( \frac{1}{2} \int_0^1 (m'(t) - m'_0(t)) \cdot d(x(t) - m(t)) \right. \\ &\quad \left. - \frac{1}{4} \int_0^1 |m'(t) - m'_0(t)|^2 dt \right). \end{aligned}$$

Recall that  $m_0(t) = x_-(1-t) + x_+t$ . Using the fact that  $x(0) = x_-$ ,  $x(1) = x_+$  when  $x$  is distributed according to  $\mathcal{T}_m^* \bar{\mu}_0$  (or  $\mathcal{T}_{m_0}^* \bar{\mu}_0$ ), we can simplify the exponent on the right-hand side of (3.21) as follows:

$$(3.22) \quad \begin{aligned} &\frac{1}{2} \int_0^1 (m'(t) - m'_0(t)) \cdot d(x(t) - m_0(t)) - \frac{1}{4} \int_0^1 |m'(t) - m'_0(t)|^2 dt \\ &= \frac{1}{2} \int_0^1 m'(t) \cdot dx(t) - \frac{1}{4} \int_0^1 |m'(t)|^2 dt - \frac{1}{2} \int_0^1 m'_0(t) \cdot d(x(t) - m_0(t)) \\ &\quad - \frac{1}{4} \int_0^1 |m'_0(t)|^2 dt \\ &= \frac{1}{2} \int_0^1 m'(t) \cdot dx(t) - \frac{1}{4} \int_0^1 |m'(t)|^2 dt - \frac{|x_+ - x_-|^2}{4}. \end{aligned}$$

Hence one can obtain (3.16) from (3.19)–(3.22), where the normalization constant

$$Z_{\nu, \varepsilon} = Z_{\bar{\nu}, \varepsilon} \cdot \exp \left( \frac{|x_+ - x_-|^2}{4} \right).$$

This together with (3.4) implies (3.18).  $\square$

According to the definition of  $\mu_\varepsilon$  (given by (2.8)), Lemma 3.2, and the expression (3.15) for the Kullback–Leibler divergence, we obtain that

$$(3.23) \quad D_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon) = \tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon) - \frac{|x_1 - x_0|^2}{4} + \log(Z_{\mu, \varepsilon}),$$

where

$$(3.24) \quad \begin{aligned} \tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon) &= \frac{1}{2\varepsilon^2} \mathbb{E}^{\bar{\nu}_\varepsilon} \int_0^1 \Psi_\varepsilon(z(t) + m(t)) dt + \frac{1}{4} \int_0^1 |m'(t)|^2 dt \\ &\quad - \frac{1}{4} \mathbb{E}^{\bar{\nu}_\varepsilon} \left[ \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt \right] + \frac{1}{2\varepsilon} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt \\ &\quad + \frac{1}{2} \log \left( \det \left( \int_0^1 \bar{\mathbf{M}}_\varepsilon(t) \bar{\mathbf{M}}_\varepsilon(t)^T dt \right) \right). \end{aligned}$$

Here  $\bar{\nu}_\varepsilon = N(0, 2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1})$  and  $\bar{\mathbf{M}}_\varepsilon(t) = \mathbf{M}_\varepsilon(1, t)$  with  $\mathbf{M}_\varepsilon$  defined by (3.2). The form of  $\tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$  is interesting: the first two terms comprise a “fattened” version of the Onsager–Machlup functions (1.9), where the fattening is characterized by the entropic fluctuations of the process  $z$ . The remaining terms penalize those entropic contributions. This characterization will be particularly clear in the small noise limit—see the discussion in section 5.

**3.3. Variational problem.** Recall the set of Gaussian measures

$$\mathcal{A} = \left\{ N(m, 2(-\partial_t^2 + \mathbf{B}_\varepsilon)^{-1}) : (m, \mathbf{A}) \in \mathcal{H} \right\},$$

where  $\mathbf{B}_\varepsilon = \varepsilon^{-2} \mathbf{A}^2 - \varepsilon^{-1} \mathbf{A}'$ , and that  $\mathcal{A}_a$  is defined in the same way with  $\mathcal{H}$  replaced by  $\mathcal{H}_a$  for some  $a > 0$ . Given the measure  $\mu_\varepsilon$  defined by (2.8), i.e., the law of transition paths, we aim to find optimal Gaussian measures  $\nu_\varepsilon$  from  $\mathcal{A}$  or  $\mathcal{A}_a$  minimizing the Kullback–Leibler divergence  $D_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$ . To that end, first in view of (3.23), the constants  $\frac{|x_1 - x_0|^2}{4}$  and  $\log(Z_{\mu, \varepsilon})$  can be neglected in the minimization process since they do not depend on the choice of  $\nu_\varepsilon$ . Hence we are concerned only with minimizing the modified Kullback–Leibler divergence  $\tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$ . Furthermore, instead of minimizing  $\tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$ , we consider the variational problem

$$(3.25) \quad \inf_{\nu \in \mathcal{A}} \left( \varepsilon \tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon) + \varepsilon^\gamma \|\mathbf{A}\|_{\mathbf{H}^1(0,1)}^2 \right),$$

where  $\gamma > 0$  and  $\mathcal{A}$  is given by (3.14). We will also study the minimization problem over the set  $\mathcal{A}_a$ .

The reasons why the problem (3.25) is of interest to us are the following. First, multiplying  $\tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$  by  $\varepsilon$  does not change the minimizers. Yet after this scaling, the  $m$ -dependent terms of  $\tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$  (the first two terms on the right-hand side of (3.24)) and the  $\mathbf{A}$ -dependent terms (middle line of (3.24)) are well-balanced since they are all order one quantities with respect to  $\varepsilon$ . Moreover, the regularization term  $\varepsilon^\gamma \|\mathbf{A}\|_{\mathbf{H}^1(0,1)}^2$  is necessary because the matrix  $\mathbf{B}_\varepsilon$ , along any infimizing sequence for  $\varepsilon \tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$ , will only converge weakly, and the minimizer may not be attained in  $\mathcal{A}$ . This issue is illustrated in [28, Examples 3.8 and 3.9], and a similar regularization is used there.

*Remark 3.3.* The normalization constant  $Z_{\mu, \varepsilon}$  in (3.23) is dropped in our minimization problem. This is one of the advantages of quantifying measure approximations by means of the Kullback–Leibler divergence. However, understanding the asymptotic behavior of  $Z_{\mu, \varepsilon}$  in the limit  $\varepsilon \rightarrow 0$  is quite important, even though this is difficult. In particular, it allows us to study the asymptotic behavior of the scaled Kullback–Leibler divergence  $\varepsilon D_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$ , whereby quantitative information on the quality of the Gaussian approximation in the small temperature limit can be extracted. In the next section we study behavior of the minimizers of  $F_\varepsilon$  in the limit  $\varepsilon \rightarrow 0$ ; we postpone study of  $\varepsilon D_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$ , which requires analysis of  $Z_{\mu, \varepsilon}$  in the limit  $\varepsilon \rightarrow 0$ , to future work.

*Remark 3.4.* We choose the small weight  $\varepsilon^\gamma$  with some  $\gamma > 0$  in front of the regularization term, with the aim of weakening the contribution from the regularization so that it disappears in the limit  $\varepsilon \rightarrow 0$ . For the study of the  $\Gamma$ -limit of  $F_\varepsilon$ , we will consider  $\gamma \in (0, \frac{1}{2})$ ; see Theorem 4.6 in the next section.

*Remark 3.5.* The Kullback–Leibler divergence is not symmetric in its arguments. We do not study  $\tilde{D}_{\text{KL}}(\mu_\varepsilon || \nu_\varepsilon)$ , because minimization of this functional over the class

of Gaussian measures leads simply to moment matching, and this is not appropriate for problems with multiple minimizers; see [3, section 10.7].

The following theorem establishes the existence of minimizers for the problem (3.25).

**THEOREM 3.6.** *Given the measure  $\mu_\varepsilon$  defined by (2.8) with fixed  $\varepsilon > 0$ , there exists at least one measure  $\nu \in \mathcal{A}$  (or  $\mathcal{A}_a$ ) minimizing the functional*

$$(3.26) \quad \nu \mapsto \varepsilon \tilde{D}_{KL}(\nu || \mu_\varepsilon) + \varepsilon^\gamma \|\mathbf{A}\|_{\mathbf{H}^1(0,1)}^2$$

over  $\mathcal{A}$  (or  $\mathcal{A}_a$ ).

*Proof.* We prove the theorem only for the case where the minimizing problem is defined over  $\mathcal{A}_a$ , since the other case can be treated in the same manner. First we show that the infimum of (3.26) over  $\mathcal{A}_a$  is finite for any fixed  $\varepsilon > 0$ . In fact, consider  $\mathbf{A}^* = a \cdot \mathbf{I}_d$  with  $a > 0$  and  $m^*$  being any fixed function in  $\mathbf{H}_+^1(0,1)$ . Then we show that  $F(m^*, \mathbf{A}^*)$  is finite. For this, by (3.24), we need only to show that

$$\mathbb{E}^{\bar{\nu}_\varepsilon} \int_0^1 \Psi_\varepsilon(z(t) + m^*(t)) dt < \infty.$$

Since  $\mathbf{A}^* = a \cdot \mathbf{I}_d$ , from (3.11) one can see that  $z(t) \sim N(0, 2\mathbf{G}_\varepsilon(t, t))$  under the measure  $\bar{\nu}_\varepsilon$ . In addition, it follows from (6.14) that  $|\mathbf{G}_\varepsilon(t, t)|_F \leq C\varepsilon$  a.e. on  $(0, 1)$  for some  $C > 0$ . Then from the growth condition (A-4) on  $\Psi_\varepsilon$  and the fact that  $m^* \in \mathbf{L}^\infty(0, 1)$ ,

$$\begin{aligned} & \mathbb{E}^{\bar{\nu}_\varepsilon} \int_0^1 \Psi_\varepsilon(z(t) + m^*(t)) dt \\ &= \int_0^1 \int_{\mathbf{R}^d} \frac{1}{\sqrt{(4\pi)^d \det(\mathbf{G}_\varepsilon(t, t))}} e^{-\frac{1}{4}x^T \mathbf{G}_\varepsilon(t, t)^{-1}x} \Psi_\varepsilon(x + m^*(t)) dx dt \\ &= \int_0^1 \int_{\mathbf{R}^d} \frac{1}{(4\pi)^{d/2}} e^{-\frac{1}{4}|x|^2} \Psi_\varepsilon\left((\mathbf{G}_\varepsilon(t, t)^{1/2})x + m^*(t)\right) dx dt \\ &\leq C_1 \exp\left(\|m^*\|_{\mathbf{L}^\infty(0,1)}^\alpha\right) \int_{\mathbf{R}^d} e^{-\frac{1}{2}|x|^2 + C_2\varepsilon^\alpha|x|^\alpha} dx < \infty \end{aligned}$$

since  $\alpha \in [0, 2)$ .

Next, we prove that the minimizer exists. By examining the proof of [28, Theorem 3.10], one can see that the theorem is proved if the following statement is valid: if a sequence  $\{\mathbf{A}_n\} \subset \mathbf{H}_a^1(0, 1)$  satisfies  $\sup_n \|\mathbf{A}_n\|_{\mathbf{H}^1(0,1)} < \infty$ , then the sequence  $\{\mathbf{B}_n\}$  with  $\mathbf{B}_n = \varepsilon^{-2}\mathbf{A}_n^2 - \varepsilon^{-1}\mathbf{A}_n'$ , viewed as multiplication operators, contains a subsequence that converges to  $\mathbf{B} = \varepsilon^{-2}\mathbf{A}^2 - \varepsilon^{-1}\mathbf{A}'$  in  $\mathcal{L}(\mathbf{H}^\beta, \mathbf{H}^{-\beta})$  for some  $\mathbf{A} \in \mathbf{H}_a^1(0, 1)$  and some  $\beta \in (0, 1)$ . Hence we only need to show that the latter statement is true. In fact, if  $\sup_n \|\mathbf{A}_n\|_{\mathbf{H}^1(0,1)} < \infty$ , then there exists a subsequence  $\{\mathbf{A}_{n_k}\}$  and some  $\mathbf{A} \in \mathbf{H}^1(0, 1)$  such that  $\mathbf{A}_{n_k} \rightharpoonup \mathbf{A}$  in  $\mathbf{H}^1(0, 1)$ . By Rellich's compact embedding theorem,  $\mathbf{A}_{n_k} \rightarrow \mathbf{A}$  in  $\mathbf{L}^2(0, 1)$ , and passing to a further subsequence, we may assume that  $\mathbf{A}_{n_k} \rightarrow \mathbf{A}$  a.e. on  $[0, 1]$ . This implies that  $\mathbf{A}$  is symmetric and  $\mathbf{A} \geq a \cdot \mathbf{I}_d$  a.e., and hence  $\mathbf{A} \in \mathbf{H}_a^1(0, 1)$ . In addition, it is clear that  $\mathbf{B}_{n_k} \rightharpoonup \mathbf{B}$  in  $\mathbf{L}^2(0, 1)$ . According to Lemma C.3, for any  $\alpha, \beta > 0$  such that  $\beta > \max(\alpha, \alpha/2 + 1/4)$ , a matrix-valued function in  $\mathbf{H}^{-\alpha}(0, 1)$  can be viewed as a multiplication operator in  $\mathcal{L}(\mathbf{H}^\beta, \mathbf{H}^{-\beta})$ . Thanks to the compact embedding from  $\mathbf{L}^2(0, 1)$  to  $\mathbf{H}^{-\alpha}(0, 1)$ , we obtain  $\mathbf{B}_{n_k} \rightarrow \mathbf{B}$  in  $\mathcal{L}(\mathbf{H}^\beta, \mathbf{H}^{-\beta})$ . The proof is complete.  $\square$

*Remark 3.7.* Minimizers of (3.26) are not unique in general. The uniqueness issue is outside the scope of this paper; see more discussions about uniqueness of minimizing the Kullback–Leibler divergence in [28, section 3.4].

**4. Low temperature limit.** In this section, we aim to understand the low temperature limit of the best Gaussian approximations discussed in the previous section. This will be done in the framework of  $\Gamma$ -convergence. First we recall the definition of  $\Gamma$ -convergence (see [4, 24]) and introduce some functionals which are closely related to the Gaussian approximations.

#### 4.1. Notion of $\Gamma$ -convergence and preliminaries.

**DEFINITION 4.1.** Let  $\mathcal{X}$  be a topological space,  $\varepsilon > 0$ , and  $F_\varepsilon : \mathcal{X} \rightarrow \mathbf{R}$  be a family of functionals. We say that  $F_\varepsilon$   $\Gamma$ -converges to  $F : \mathcal{X} \rightarrow \mathbf{R}$  as  $\varepsilon \rightarrow 0$  if the following two conditions hold:

(i) *Liminf inequality:* for every  $u \in \mathcal{X}$  and every sequence  $u_\varepsilon \in \mathcal{X}$  such that  $u_\varepsilon \rightarrow u$ ,

$$F(u) \leq \liminf_{\varepsilon \rightarrow 0} F_\varepsilon(u_\varepsilon).$$

(ii) *Limsup inequality:* for every  $u \in \mathcal{X}$  there exists a sequence  $u_\varepsilon$  such that  $u_\varepsilon \rightarrow u$  and

$$F(u) \geq \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(u_\varepsilon).$$

For studying the low temperature limit of the Gaussian approximations, we consider the following family of functionals:

$$(4.1) \quad F_\varepsilon(m, \mathbf{A}) := \begin{cases} \varepsilon \tilde{D}_{\text{KL}}(\nu_\varepsilon \| \mu_\varepsilon) + \varepsilon^\gamma \|\mathbf{A}\|_{\mathbf{H}^1(0,1)}^2 & \text{if } (m, \mathbf{A}) \in \mathcal{H}, \\ \infty & \text{otherwise in } \mathcal{X}, \end{cases}$$

on the space  $\mathcal{X} = \mathbf{L}^1(0,1) \times \mathbf{L}^1(0,1)$ . Then minimizing (3.26) over  $\mathcal{A}$  is equivalent to the following problem:

$$(4.2) \quad \inf_{(m, \mathbf{A}) \in \mathcal{X}} F_\varepsilon(m, \mathbf{A}).$$

In order to study the  $\Gamma$ -limit of  $F_\varepsilon$ , we equip the space  $\mathcal{X}$  with a product topology such that the convergence  $(m_\varepsilon, \mathbf{A}_\varepsilon) \rightarrow (m, \mathbf{A})$  in  $\mathcal{X}$  means that  $m_\varepsilon \rightarrow m$  in  $\mathbf{L}^1(0,1)$  and that  $\mathbf{A}_\varepsilon \rightharpoonup \mathbf{A}$  in  $\mathbf{L}^1(0,1)$ . The reason for choosing the weak topology for  $\mathbf{A}$  is that the functional  $F_\varepsilon$  is coercive under such topology only; see Proposition 4.5. Now before we proceed to discussing the  $\Gamma$ -convergence of  $F_\varepsilon$ , we first state a useful  $\Gamma$ -convergence result for the classical Ginzburg–Landau functional,

$$(4.3) \quad E_\varepsilon(m) := \begin{cases} \frac{\varepsilon}{4} \int_0^1 |m'(t)|^2 dt + \frac{1}{4\varepsilon} \int_0^1 |\nabla V(m(t))|^2 dt & \text{if } m \in \mathbf{H}_\pm^1(0,1), \\ \infty & \text{otherwise in } \mathbf{L}^1(0,1). \end{cases}$$

Notice that, in the above definition, any  $m$  such that  $E_\varepsilon(m)$  is finite should satisfy the Dirichlet boundary conditions  $m(0) = x_-$  and  $m(1) = x_+$ . We also remark that after performing the scaling transformation  $t \mapsto \varepsilon^{-1}t$ , the functional  $E_\varepsilon$  coincides with the Freidlin–Wentzell functional  $S_T$  (defined in (1.6)) with  $T = \varepsilon^{-1}$ . Indeed, by rewriting  $\tilde{m}(\cdot) = m(\varepsilon^{-1}\cdot)$ , one sees that  $E_\varepsilon(m) = S_T(\tilde{m})$ .

To define the  $\Gamma$ -limit of  $E_\varepsilon$ , we now introduce some additional notation. Recall that  $\mathcal{E}$  defined in (2.2) is the set of critical points of  $V$ . For each pair  $x_-, x_+ \in \mathcal{E}$  we define the set of transition paths

$$\mathbf{X}(x_-, x_+) := \left\{ m \in \mathbf{BV}(\mathbf{R}) \mid \lim_{t \rightarrow \pm\infty} m(t) = x_\pm \text{ and } m' \in \mathbf{L}^2(\mathbf{R}) \right\}$$

and the cost functional

$$(4.4) \quad \mathcal{J}_T(m) = \frac{1}{4} \int_{-T}^T (|m'(t)|^2 + |\nabla V(m(t))|^2) dt,$$

and set  $\mathcal{J}(m) := \mathcal{J}_\infty(m)$ . The minimal transition cost from  $x_-$  to  $x_+$  is then defined as

$$\Phi(x_-, x_+) := \inf \{ \mathcal{J}(m) \mid m \in \mathbf{X}(x_-, x_+) \}.$$

It is worth noting that the function  $\Phi(x_-, x_+)$  is closely related to the so-called quasi potential, which plays an important role in large deviation theory: in fact, suppose that  $x_-, x_+ \in \mathcal{E}$  and that  $V$  satisfies Assumption (2.1). Then according to [14, Lemma 3.2], the function  $\Phi(x_-, x_+)$  has the following equivalent form:

$$(4.5) \quad \Phi(x_-, x_+) = \inf_{T, m} \left\{ \mathcal{J}_T(m) : T > 0, m \in \mathbf{H}^1(-T, T) \text{ and } m(-T) = x_-, m(T) = x_+ \right\}.$$

This definition shows that  $\Phi(x_-, x_+)$  coincides with the quasi potential between  $x_-$  and  $x_+$  (as defined in [15, Chapter 4]) up to the additive constant  $-\frac{1}{2}(V(x_+) - V(x_-))$ .

We also remark that the equivalent formulation (4.5) provides an important ingredient for proving the  $\Gamma$ -convergence of  $E_\varepsilon$ ; see, e.g., [14, 4]. Given  $x_\pm \in \mathcal{E}$ , if either  $x_-$  or  $x_+$  is a local minimum or maximum of potential  $V$  and if  $V$  satisfies (A-1)–(A-3) of Assumption 2.1, it was shown in [30, Lemma 2.1] that the infimum  $\Phi(x_-, x_+)$  is attained by the heteroclinic orbits  $m_*$  of the Hamiltonian system

$$m_*''(t) - D^2V(m_*)\nabla V(m_*) = 0, \quad \lim_{t \rightarrow \pm\infty} m(t) = x_\pm.$$

In this case,

$$(4.6) \quad \Phi(x_-, x_+) = \frac{1}{2} |V(x_+) - V(x_-)|.$$

Denote by  $\mathbf{BV}(0, 1; \mathcal{E})$  the set of functions in  $\mathbf{BV}(0, 1)$  taking values in  $\mathcal{E}$  a.e. on  $[0, 1]$ . For any  $u \in \mathbf{BV}(0, 1; \mathcal{E})$  let  $J(u)$  be the set of jump points of  $u$  on  $(0, 1)$ , and let  $u(t^\pm)$  denote the left- and right-sided limits of  $u$  at time  $t \in [0, 1]$ . The following lemma, concerning the compactness of  $E_\varepsilon$ , will be very useful in identifying its  $\Gamma$ -limit. Its proof can be found in [21, Theorem 1.2].

**LEMMA 4.2.** *Assume that the potential  $V$  satisfies (A-1)–(A-3). Let  $\varepsilon_n \rightarrow 0$ , and let  $\{m_n\} \subset \mathbf{H}_\pm^1(0, 1)$  be such that*

$$\limsup_{n \rightarrow \infty} E_{\varepsilon_n}(m_n) < \infty.$$

*Then there exists a subsequence  $\{m_{n_k}\}$  of  $\{m_n\}$  and an  $m \in \mathbf{BV}(0, 1; \mathcal{E})$  such that  $m_{n_k} \rightarrow m$  in  $\mathbf{L}^1(0, 1)$  as  $k \rightarrow \infty$ .*



We remark that we incorporate the boundary conditions  $m_n(0) = x_-$ ,  $m_n(1) = x_+$  in the statement of the lemma since  $m_n \in \mathbf{H}_\pm^1(0, 1)$ . The following proposition identifies the  $\Gamma$ -limit of  $E_\varepsilon$  with respect to  $\mathbf{L}^1$ -topology; this is based upon Lemma 4.2 and the standard Modica–Mortola-type arguments (see [25, 1, 30]). The proof is given in Appendix D. The same  $\Gamma$ -convergence result was claimed in [30], but the proof there was actually carried out with respect to the topology in the space of bounded variations.

**PROPOSITION 4.3.** *Assuming that  $V$  satisfies the conditions (A-1)–(A-3), the  $\Gamma$ -limit of  $E_\varepsilon$  is*

$$(4.7) \quad E(m) := \begin{cases} \Phi(x_-, m(0^+)) + \sum_{\tau \in J(m)} \Phi(m(\tau^-), m(\tau^+)) \\ \quad + \Phi(m(1^-), x_+) & \text{if } m \in \mathbf{BV}(0, 1; \mathcal{E}), \\ +\infty & \text{otherwise in } \mathbf{L}^1(0, 1). \end{cases}$$

**4.2. Main results.** This subsection presents the main results about the  $\Gamma$ -convergence of the functional  $F_\varepsilon$ ; the proofs will be presented in the next section. Roughly speaking, our arguments indicate that the  $\Gamma$ -limit of  $F_\varepsilon$  on  $\mathcal{X}$  should be

$$(4.8) \quad F(m, \mathbf{A}) := E(m) + \frac{1}{4} \int_0^1 (D^2 V(m(t)) - |\mathbf{A}(t)|)^2 : |\mathbf{A}^{-1}(t)| dt,$$

where  $E(m)$  is defined by (4.7). Recall that  $\mathbf{A} : \mathbf{B} = \text{Tr}(\mathbf{A}\mathbf{B}^T)$ . However, for technical reasons, we are able to prove the claim only under the condition that the matrix  $\mathbf{A}$  is positive definite; see Remark 4.9. To make this clear, let us first redefine  $F_\varepsilon$  to be

$$(4.9) \quad F_\varepsilon(m, \mathbf{A}) := \begin{cases} \varepsilon \tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon) + \varepsilon^\gamma \|\mathbf{A}\|_{\mathbf{H}^1(0,1)}^2 & \text{if } (m, \mathbf{A}) \in \mathcal{H}_a, \\ \infty & \text{otherwise in } \mathcal{X}_a, \end{cases}$$

with some  $a > 0$ . Then we can show that  $F_\varepsilon$  as defined in (4.9)  $\Gamma$ -converges to  $F$  as defined by (4.8) on the space  $\mathcal{X}_a$  for any  $a > 0$ ; see Theorem 4.6. Recall that  $\mathcal{X}_a = \mathbf{L}^1(0, 1) \times \mathbf{L}_a^1(0, 1)$  and that convergence of  $(m_n, \mathbf{A}_n)$  in  $\mathcal{X}_a$  means that the  $m_n$  converge strongly in  $\mathbf{L}^1(0, 1)$  and the  $\mathbf{A}_n$  converge weakly in  $\mathbf{L}_a^1(0, 1)$ .

By the definition of  $F_\varepsilon$  (by (4.9)) and the expression (3.23) for  $\tilde{D}_{\text{KL}}(\nu_\varepsilon || \mu_\varepsilon)$ , we can write

$$(4.10) \quad F_\varepsilon(m, \mathbf{A}) = F_\varepsilon^{(1)}(m, \mathbf{A}) + F_\varepsilon^{(2)}(\mathbf{A}) + \varepsilon^\gamma \|\mathbf{A}\|_{\mathbf{H}^1(0,1)}^2$$

for  $(m, \mathbf{A}) \in \mathcal{H}_a$ , where

$$(4.11) \quad \begin{aligned} F_\varepsilon^{(1)}(m, \mathbf{A}) &:= \frac{\varepsilon}{4} \int_0^1 |m'(t)|^2 dt + \frac{1}{2\varepsilon} \mathbb{E}^{\bar{\nu}_\varepsilon} \left[ \int_0^1 \Psi_\varepsilon(z(t) + m(t)) dt \right], \\ F_\varepsilon^{(2)}(\mathbf{A}) &:= -\frac{\varepsilon}{4} \mathbb{E}^{\bar{\nu}_\varepsilon} \left[ \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt \right] + \frac{1}{2} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt \\ &\quad + \frac{\varepsilon}{2} \log \left( \det \left( \int_0^1 \bar{\mathbf{M}}_\varepsilon(t) \bar{\mathbf{M}}_\varepsilon(t)^T dt \right) \right), \end{aligned}$$

where  $\Psi_\varepsilon$  is given by (1.8) and  $\bar{\mathbf{M}}_\varepsilon$  is defined by (3.2). To identify the  $\Gamma$ -limit of  $F_\varepsilon$ , we need to study the liminf or limsup of the sequence  $\{F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon)\}$  with  $m_\varepsilon \in \mathbf{H}_\pm^1(0, 1)$  and  $\mathbf{A}_\varepsilon \in \mathbf{H}_a^1(0, 1)$ . This is nontrivial in our case, mainly because the functional

$F_\varepsilon$  depends on  $m$  and  $\mathbf{A}$  in an implicit manner through the two expectation terms. Therefore in the first step we shall simplify  $F_\varepsilon$ . The following proposition examines the limiting behavior of the functional  $F_\varepsilon$  from which a simplified and more explicit expression is obtained.

PROPOSITION 4.4. *Let  $(m_\varepsilon, \mathbf{A}_\varepsilon) \in \mathcal{H}_a$ . Assume that for some  $\gamma \in (0, \frac{1}{2})$*

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty \quad \text{and} \quad \limsup_{\varepsilon \rightarrow 0} \|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)} < \infty.$$

*Then for  $\varepsilon > 0$  small enough we have*

(4.12)

$$\begin{aligned} F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon) &= E_\varepsilon(m_\varepsilon) + \frac{1}{4} \int_0^1 (D^2V(m_\varepsilon(t)) - \mathbf{A}_\varepsilon(t))^2 : \mathbf{A}_\varepsilon^{-1}(t) dt \\ &\quad + \int_0^1 (D^3V(m_\varepsilon(t)) \cdot \nabla V(m_\varepsilon(t))) : \mathbf{A}_\varepsilon^{-1}(t) dt + \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 + \mathcal{O}(\varepsilon^{\frac{1}{2}}). \end{aligned}$$

The proof of Proposition 4.4 requires several technical lemmas and is referred to section 6.3. The basic idea for proving Proposition 4.4 is as follows. First one can express the expectation term in  $F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon)$  in terms of the Dirichlet Green's tensor of some Schrödinger operator (see (6.2)). Careful asymptotic analysis of this Green's tensor implies that

$$(4.13) \quad F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon) \approx \frac{1}{4} \int_0^1 \text{Tr}(\mathbf{A}_\varepsilon(t)) dt;$$

see Corollary 6.2 for the precise statement. For the expectation term in  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$ , we approximate  $\Psi_\varepsilon(X)$  by its second order Taylor expansion around the mean  $m_\varepsilon$ . The zero order term of the expansion is  $\Psi_\varepsilon(m_\varepsilon) = \frac{1}{2} |\nabla V(m_\varepsilon)|^2 - \varepsilon \Delta V(m_\varepsilon)$ . Then  $E_\varepsilon(m_\varepsilon)$  is obtained by combining the term  $\frac{\varepsilon}{4} \int_0^1 |m'(t)|^2 dt$  in  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$  with the integral over  $\frac{1}{4\varepsilon} |\nabla V(m_\varepsilon)|^2$ . Additionally, the Itô correction term  $-\varepsilon \Delta V(m_\varepsilon)$ , which is the other zero order term of the Taylor expansion, can be combined with one of the second order terms of the expansion and (4.13) to complete the full quadratic term in (4.12).

As a consequence of Proposition 4.4, we get the following interesting compactness result for the functional  $F_\varepsilon$ .

PROPOSITION 4.5. *Let  $\varepsilon_n \rightarrow 0$ , and let  $\{(m_n, \mathbf{A}_n)\}$  be a sequence in  $\mathcal{H}_a$  such that*

$$\limsup_n F_{\varepsilon_n}(m_n, \mathbf{A}_n) < \infty.$$

*Then there exists a subsequence  $\{(m_{n_k}, \mathbf{A}_{n_k})\}$  of  $\{(m_n, \mathbf{A}_n)\}$  such that  $m_{n_k} \rightarrow m$  in  $\mathbf{L}^1(0,1)$  and  $\mathbf{A}_{n_k} \rightharpoonup \mathbf{A}$  in  $\mathbf{L}^1(0,1)$  with  $m \in \mathbf{BV}(0,1; \mathcal{E})$  and  $\mathbf{A} \in \mathbf{L}_a^1(0,1)$ .*

This compactness result is slightly weaker than the usual compactness property relevant to  $\Gamma$ -convergence (see, e.g., the conclusion in Lemma 4.2), because only weak convergence is obtained for the variable  $\mathbf{A}$ . Building upon the  $\Gamma$ -convergence result of  $E_\varepsilon$ , Proposition 4.4, and Proposition 4.5, the following main theorem establishes the  $\Gamma$ -convergence of  $F_\varepsilon$ .

THEOREM 4.6. *Suppose that  $V$  satisfies assumptions (A-1)–(A-6). Let  $\gamma \in (0, \frac{1}{2})$  in (4.10). Then the  $\Gamma$ -limit of  $F_\varepsilon$  defined by (4.9) on  $\mathcal{X}_a$  is*

$$(4.14) \quad F(m, \mathbf{A}) = E(m) + \frac{1}{4} \int_0^1 (D^2V(m(t)) - \mathbf{A}(t))^2 : \mathbf{A}^{-1}(t) dt,$$

where  $E(m)$  is defined by (4.7).

$\Gamma$ -convergence of  $F_\varepsilon$  implies convergence of the minima.

**COROLLARY 4.7.** *Let  $(m_\varepsilon, \mathbf{A}_\varepsilon) \in \mathcal{H}_a$  be minimizers of  $F_\varepsilon$ . Then up to extracting a subsequence,  $m_\varepsilon \rightarrow m$  in  $\mathbf{L}^1(0, 1)$  and  $\mathbf{A}_\varepsilon \rightharpoonup \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$  for some  $m \in \mathbf{BV}(0, 1; \mathcal{E})$  and  $\mathbf{A} \in \mathbf{L}_a^1(0, 1)$ . Furthermore, the limit  $(m, \mathbf{A})$  is a minimizer of  $F$  on  $\mathcal{X}_a$ .*

**Remark 4.8.** In general, convergence of minima requires both (strong) compactness and  $\Gamma$ -convergence; see, e.g., [4]. In our case we only have weak compactness with respect to  $\mathbf{A}_\varepsilon$  for  $F_\varepsilon$ ; see Proposition 4.5. However, weak convergence of  $\mathbf{A}_\varepsilon$  suffices to pass to the limit because the leading order term of the functional  $F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon)$  is convex with respect to  $\mathbf{A}_\varepsilon$ . See the analysis of the functional (6.41) in the next section.

**Remark 4.9.** Theorem 4.6 shows the  $\Gamma$ -convergence of  $F_\varepsilon$  to  $F$  (given by (4.14)) under the assumption that  $\mathbf{A}$  is bounded away from zero; i.e.,  $\mathbf{A}(\cdot) \geq a \cdot \mathbf{I}_d$  for some  $a > 0$ . However, this assumption is unlikely to be sharp. In fact, under the weaker positivity assumption that  $|\mathbf{A}(\cdot)| \geq a \cdot \mathbf{I}_d$ , one can at least prove the liminf part of the  $\Gamma$ -convergence of  $F_\varepsilon$  to  $F$  defined in (4.8). This is mainly because the leading order of the Green's function  $\mathbf{G}_\varepsilon$  (defined by (3.12)) depends only on  $|\mathbf{A}|$ ; see (A.6) of Lemma A.2. Although the positivity assumption is essential in our arguments for proving Theorem 4.6, we conjecture that the  $\Gamma$ -convergence result is still valid without any positivity assumption. This is to be investigated in future work.

**5. Conclusion.** The Freidlin–Wentzell theory gives a quantitative description of the tail of the distribution of transition paths based on the theory of large deviations. It thereby leads to a natural variational definition of most likely paths in the low temperature limit, namely, the minimizers of the large deviation rate functional. However, this approach exhibits some weaknesses. In particular the large deviation theory of Freidlin and Wentzell makes asymptotic statements in the limit where the noise intensity goes to zero. In practical applications the noise level may not be low enough for these asymptotics to be valid. Furthermore, the large deviation rate functional does not exclude the possibility of the path spending large stretches of time near local maxima or saddles of the potential  $V$ . The Onsager–Machlup theory offers an alternative variational definition of most likely paths at finite temperature in terms of MAP estimators, but as shown in [29], minimizers of the Onsager–Machlup functional may be unphysical at small temperatures because the methodology fails to account for entropic effects; this can lead to transition paths which choose to make transitions through narrow energy barriers rather than (entropically favorable) wider ones with the same height, or to transition paths which (like Freidlin–Wentzell paths, although for different reasons) can spend long times at a saddle point.

We have developed an approach to the problem of identifying the most likely transition path which (like the Onsager–Machlup approach) is well-defined at finite nonvanishing temperature and yet which also recovers the correct limiting behavior in the small temperature limit (minimizers of the Freidlin–Wentzell least action principle). Furthermore, our approach is based on finding the best Gaussian approximation with respect to Kullback–Leibler divergence, and hence captures not only the most likely path, but also the fluctuations around it. In the small temperature limit this gives the appealing interpretation that the fluctuations are defined by an OU process found from linearizing the Brownian dynamics model at the minimizer of the Freidlin–Wentzell action. It is thus important to recognize that our work leads to useful characterizations of transition paths in the Brownian dynamics model, both at

finite  $\varepsilon$  and in the limit  $\varepsilon \rightarrow 0$ . In this paper we have concentrated exclusively on the  $\varepsilon \rightarrow 0$  limit. However, we now make some remarks that have bearing on both of these parameter regimes.

**5.1. Computational methods (fixed  $\varepsilon$ ).** Even though there is no explicit *analytic* characterization for our notion of most likely path at finite temperature, it is possible to calculate one *numerically*. In fact, for a fixed finite temperature  $\varepsilon$ , finding the best Gaussian approximation requires minimizing the functional  $F_\varepsilon$ , which involves the Kullback–Leibler divergence. In [27], a variant of the Robbins–Monro algorithm was introduced to find best Gaussian approximations with respect to the Kullback–Leibler divergence; the numerical results in that paper demonstrate the feasibility of the minimization and also demonstrate that the resulting Gaussian approximation can be used to construct improved Markov chain Monte Carlo algorithms for transition path sampling.

An important aspect of any gradient descent method for minimizing an objective function is the initialization. When the temperature  $\varepsilon$  is small but finite, the analysis in this paper also suggests a good initialization. Proposition 4.4 gives an approximate formula for the functional  $F_\varepsilon$ :

$$F_\varepsilon = \bar{F}_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon) + o(1),$$

where

$$\bar{F}_\varepsilon(m, \mathbf{A}) := E_\varepsilon(m) + \frac{1}{4} \int_0^1 (D^2V(m(t)) - \mathbf{A}(t))^2 : \mathbf{A}^{-1}(t) dt.$$

As a consequence, a minimizer of  $\bar{F}_\varepsilon$  provides a good approximation for the minimizer of  $F_\varepsilon$ . Approximate minimization of the former, to obtain a good initialization, can be carried out by making several alternations of the following two steps: first, freeze  $\mathbf{A}$  and minimize  $\bar{F}_\varepsilon(m, \mathbf{A})$  as a functional of  $m$ —this can be implemented via many minimum action algorithms (see, e.g., [11, 16, 19]); second, with a good approximation for  $m$  being frozen, update  $\mathbf{A}$  to be the minimizer of

$$\int_0^1 (D^2V(m(t)) - \mathbf{A}(t))^2 : \mathbf{A}^{-1}(t) dt,$$

which gives  $\mathbf{A}(t) = |D^2V(m(t))|$  (see Lemma C.9).

**5.2. Interpretation of small  $\varepsilon$  analysis.** The discussion about initializing the numerical minimization for small  $\varepsilon$  also helps to explain our earlier assertions about the desirable structure of our minimizers when  $\varepsilon$  is small. Our main result, Theorem 4.6, shows that in the small temperature limit, the Kullback–Leibler-minimization improves on the predictions obtained by minimizing the large deviation rate functional; the  $\Gamma$ -limit of the functional  $F_\varepsilon$  given in (4.8) consists of two parts whose minimization decouples in the limit  $\varepsilon = 0$ . The first part  $E$  is closely linked with large deviation theory since it is the  $\Gamma$ -limit of the scaled Freidlin–Wentzell functional  $E_\varepsilon$ . (Recall that  $E_\varepsilon$  arises from the Freidlin–Wentzell functional  $\bar{S}_T$ , defined in (1.3), by scaling  $T = \varepsilon^{-1}$  and removing the constant  $\frac{1}{2}(V(x_+) - V(x_-))$ .) The second part keeps track of the Ornstein–Uhlenbeck fluctuations around the optimal path and thereby captures entropic effects. Given a path  $m$ , it can be minimized by choosing  $\mathbf{A}(t) = |D^2V(m(t))|$ . In particular, for this choice  $(D^2V(m(t)) - |\mathbf{A}(t)|)^2 : |\mathbf{A}^{-1}(t)|$  is equal to zero if  $D^2V(m(t))$  is positive definite (corresponding to  $m(t)$  being a local minimizer of  $V$ ) and strictly positive if  $D^2V(m(t))$  has a negative eigenvalue (corresponding to  $m(t)$  being a saddle point or a local maximum). Therefore, the  $\Gamma$ -limit

$F$  can be minimized explicitly as follows: first find a minimizer of  $E$  which amounts to selecting the sequence of critical points connecting  $x_-$  and  $x_+$  that minimizes the transition cost (defined by (4.4)). As shown in section 4, the minimal transition cost equals the Freidlin–Wentzell quasi potential up to a constant, and in simple cases it is given by (4.6). Then in order to minimize also the entropic part, select from among the paths which follow this sequence those that spend no time in saddles or local maximizers. This also shows that the unphysical minimizers of the Onsager–Machlup approach (discussed in the introduction section) are removed in our approach.

## 6. Proofs of main results.

**6.1. Asymptotics of  $F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon)$ .** Let  $\mathbf{G}_\varepsilon(t, s)$  be the Green's tensor (fundamental matrix) of the elliptic operator  $(-\partial_t^2 + \mathbf{B}_\varepsilon)$  under Dirichlet boundary conditions; i.e., for any  $s \in (0, 1)$ ,

$$(6.1) \quad \begin{aligned} (-\partial_t^2 + \varepsilon^{-2} \mathbf{A}_\varepsilon^2(\cdot) - \varepsilon^{-1} \mathbf{A}'_\varepsilon(\cdot)) \mathbf{G}_\varepsilon(\cdot, s) &= \delta(\cdot - s) \cdot \mathbf{I}_d, \\ \mathbf{G}_\varepsilon(0, s) &= \mathbf{G}_\varepsilon(1, s) = 0. \end{aligned}$$

Then by the definition of covariance operator, the expectation term in  $F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon)$  can be calculated in terms of the Green's tensor  $\mathbf{G}_\varepsilon$ . More precisely,

$$(6.2) \quad -\frac{\varepsilon}{4} \mathbb{E}^{\mathbb{P}_\varepsilon} \left[ \int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt \right] = -\frac{\varepsilon}{2} \int_0^1 \mathbf{B}_\varepsilon(t) : \mathbf{G}_\varepsilon(t, t) dt.$$

To simplify  $F_\varepsilon^{(2)}$  we need the asymptotic estimates of  $\mathbf{G}_\varepsilon$  for small  $\varepsilon$ , which we show in the following.

**6.1.1. Asymptotic estimates of the Green's tensor.** For fixed  $s \in (0, 1)$ , the Green's tensor  $\mathbf{G}_\varepsilon(\cdot, s)$  solves the linear elliptic PDE system (6.1) with variable coefficient. We want to approximate  $\mathbf{G}_\varepsilon$  by a simple Green's tensor, for which an explicit asymptotic formula is available. To do this, for any  $s \in (0, 1)$  we define  $\overline{\mathbf{G}}_\varepsilon(\cdot, s)$  such that

$$(6.3) \quad \begin{aligned} (-\partial_t^2 + \varepsilon^{-2} \mathbf{A}_\varepsilon^2(s)) \overline{\mathbf{G}}_\varepsilon(\cdot, s) &= \delta(\cdot - s) \cdot \mathbf{I}_d, \\ \overline{\mathbf{G}}_\varepsilon(0, s) &= \overline{\mathbf{G}}_\varepsilon(1, s) = 0. \end{aligned}$$

According to Lemma A.2, when  $\varepsilon$  is small,

$$(6.4) \quad \overline{\mathbf{G}}_\varepsilon(t, t) = \frac{\varepsilon}{2} (\mathbf{A}_\varepsilon^{-1}(t) + \mathbf{R}_\varepsilon(t))$$

with  $|\mathbf{R}_\varepsilon(t)| \leq C(e^{-\frac{2at}{\varepsilon}} + e^{-\frac{2a(1-t)}{\varepsilon}})$ . Remember that  $a$  is the constant for which we have  $\mathbf{A}_\varepsilon(t) \geq a \cdot \mathbf{I}_d$  a.e., by assumption. Furthermore, the difference  $\tilde{\mathbf{R}}_\varepsilon(t, s) = \mathbf{G}_\varepsilon(t, s) - \overline{\mathbf{G}}_\varepsilon(t, s)$  admits the following bound for small  $\varepsilon$ .

LEMMA 6.1. Let  $\gamma \in (0, \frac{1}{2})$  in (4.9). Let  $\{\mathbf{A}_\varepsilon\} \subset \mathbf{H}_a^1(0, 1)$  such that

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}_a^1(0,1)}^2 < \infty.$$

Then for  $\varepsilon$  sufficiently small we have that

$$(6.5) \quad \sup_{s \in (0,1)} \|\tilde{\mathbf{R}}_\varepsilon(\cdot, s)\|_{\mathbf{L}^\infty(0,1)} \lesssim \varepsilon^{\frac{3}{2}-\gamma}$$

and that

$$(6.6) \quad \sup_{s \in (0,1)} \|\tilde{\mathbf{R}}_\varepsilon(\cdot, s)\|_{\mathbf{L}^2(0,1)} \lesssim \varepsilon^{2-\gamma}.$$

*Proof.* According to (6.1) and (6.3),  $\tilde{\mathbf{R}}_\varepsilon$  satisfies

$$\begin{aligned} (-\partial_t^2 + \varepsilon^{-2} \mathbf{A}_\varepsilon^2(t) - \varepsilon^{-1} \mathbf{A}_\varepsilon'(t)) \tilde{\mathbf{R}}_\varepsilon(t, s) &= \mathbf{F}_\varepsilon(t, s), \\ \tilde{\mathbf{R}}_\varepsilon(0, s) &= \tilde{\mathbf{R}}_\varepsilon(1, s) = 0, \end{aligned}$$

with

$$\mathbf{F}_\varepsilon(t, s) := (\varepsilon^{-2}(\mathbf{A}_\varepsilon^2(s) - \mathbf{A}_\varepsilon^2(t)) + \varepsilon^{-1} \mathbf{A}_\varepsilon'(t)) \overline{\mathbf{G}}_\varepsilon(t, s).$$

Let  $\tilde{R}_\varepsilon^i, \overline{G}_\varepsilon^i, F_\varepsilon^i$  be the  $i$ th column of the matrices  $\tilde{\mathbf{R}}_\varepsilon, \overline{\mathbf{G}}_\varepsilon, \mathbf{F}_\varepsilon$ , respectively:

$$(6.7) \quad \begin{aligned} (-\partial_t^2 + \varepsilon^{-2} \mathbf{A}_\varepsilon^2(t) - \varepsilon^{-1} \mathbf{A}_\varepsilon'(t)) \tilde{R}_\varepsilon^i(t, s) &= F_\varepsilon^i(t, s), \\ \tilde{R}_\varepsilon^i(0, s) &= \tilde{R}_\varepsilon^i(1, s) = 0. \end{aligned}$$

We need only to prove estimates (6.5) and (6.6) for each column  $\tilde{R}_\varepsilon^i$ ,  $i = 1, \dots, d$ . To this end, we first bound the  $\mathbf{L}^1$ -norm of the right-hand side  $F_\varepsilon^i$ . In fact, by Morrey's inequality (see, e.g., [13, Chapter 5]), it holds that

$$|\mathbf{A}_\varepsilon(t) - \mathbf{A}_\varepsilon(s)|_F \lesssim \|\mathbf{A}_\varepsilon'\|_{\mathbf{L}^2(0,1)} \cdot |t - s|^{\frac{1}{2}}$$

for any  $t, s \in [0, 1]$ . This together with (A.5) implies that

$$(6.8) \quad \begin{aligned} &\|F_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^1(0,1)} \\ &\leq 2\varepsilon^{-2} \|\mathbf{A}_\varepsilon\|_{\mathbf{L}^\infty(0,1)} \int_0^1 |(\mathbf{A}_\varepsilon(s) - \mathbf{A}_\varepsilon(t)) \overline{G}_\varepsilon^i(t, s)| dt + \varepsilon^{-1} \int_0^1 |\mathbf{A}_\varepsilon'(t) \overline{G}_\varepsilon^i(t, s)| dt \\ &\lesssim \varepsilon^{-1} \|\mathbf{A}_\varepsilon\|_{\mathbf{L}^\infty(0,1)} \cdot \|\mathbf{A}_\varepsilon'\|_{\mathbf{L}^2(0,1)} \cdot \int_0^1 |t - s|^{\frac{1}{2}} e^{-\frac{\alpha|t-s|}{\varepsilon}} dt \\ &\quad + \|\mathbf{A}_\varepsilon'\|_{\mathbf{L}^2(0,1)} \cdot \|e^{-\frac{\alpha|\cdot-s|}{\varepsilon}}\|_{L^2(0,1)} \\ &\lesssim \varepsilon^{\frac{1}{2}} (\|\mathbf{A}_\varepsilon\|_{\mathbf{L}^\infty(0,1)} + 1) \|\mathbf{A}_\varepsilon'\|_{\mathbf{L}^2(0,1)} \\ &\lesssim \varepsilon^{\frac{1}{2}} (\|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)} + 1) \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)} \lesssim \varepsilon^{\frac{1}{2}-\gamma}, \end{aligned}$$

where we have used the Sobolev embedding  $\mathbf{H}^1(0, 1) \hookrightarrow \mathbf{L}^\infty(0, 1)$  and the assumption that  $\varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty$  in the last two inequalities. Now taking the dot product of (6.7) and  $\overline{G}_\varepsilon^i(\cdot, s)$  and integrating over  $(0, 1)$ , one obtains that

$$(6.9) \quad \begin{aligned} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{H}^1(0,1)}^2 + \frac{a^2}{\varepsilon^2} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^2(0,1)}^2 &\leq \varepsilon^{-1} \int_0^1 |(\tilde{G}_\varepsilon^i(t, s))^T \mathbf{A}_\varepsilon'(t) \tilde{G}_\varepsilon^i(t, s)| dt \\ &\quad + \int_0^1 |\mathbf{F}_\varepsilon(t, s) \cdot \tilde{G}_\varepsilon^i(t, s)| dt. \end{aligned}$$

We claim that the first term on the right-hand side can be neglected when  $\varepsilon$  is small. In fact, using the Sobolev embedding  $\mathbf{H}^{\frac{1}{4}}(0, 1) \hookrightarrow \mathbf{L}^4(0, 1)$  and the interpolation

inequality of Lemma (C.1), we obtain that

$$\begin{aligned}
 \varepsilon^{-1} \int_0^1 |(\tilde{G}_\varepsilon^i(t, s))^T \mathbf{A}'_\varepsilon(t) \tilde{G}_\varepsilon^i(t, s)| dt &\leq \varepsilon^{-1} \|\mathbf{A}'_\varepsilon\|_{\mathbf{L}^2(0,1)} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^4(0,1)}^2 \\
 &\leq C\varepsilon^{-1} \|\mathbf{A}'_\varepsilon\|_{\mathbf{L}^2(0,1)} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{H}^{\frac{1}{4}}(0,1)}^2 \\
 &\leq C\varepsilon^{-1} \|\mathbf{A}'_\varepsilon\|_{\mathbf{L}^2(0,1)} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{H}^1(0,1)}^{\frac{1}{2}} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^2(0,1)}^{\frac{3}{2}} \\
 &\leq C\varepsilon^{-1-\frac{\gamma}{2}} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{H}^1(0,1)}^{\frac{1}{2}} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^2(0,1)}^{\frac{3}{2}} \\
 &\leq \frac{1}{2} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{H}^1(0,1)}^2 + C\varepsilon^{-\frac{4}{3}(1+\frac{\gamma}{2})} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^2(0,1)}^2,
 \end{aligned}$$

where we have used again the assumption that  $\varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty$  in the penultimate inequality, and Young's inequality and the equivalence of the norm on  $\mathbf{H}_0^1(0,1)$  in the last inequality. Hence for  $\gamma \in (0, \frac{1}{2})$  and  $\varepsilon$  sufficiently small, the first term on the right-hand side of (6.9) can be absorbed by the left-hand side. This implies that

$$\begin{aligned}
 (6.10) \quad |\tilde{G}_\varepsilon^i(\cdot, s)|_{\mathbf{H}^1(0,1)}^2 + \frac{a^2}{\varepsilon^2} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^2(0,1)}^2 &\lesssim \int_0^1 |\mathbf{F}_\varepsilon(t, s) \cdot \tilde{G}_\varepsilon^i(t, s)| dt \\
 &\leq \|\mathbf{F}_\varepsilon(\cdot, s)\|_{\mathbf{L}^1(0,1)} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^\infty(0,1)}.
 \end{aligned}$$

In addition, according to Lemma C.2,

$$\begin{aligned}
 \frac{a}{\varepsilon} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^\infty(0,1)}^2 &\leq \frac{2a}{\varepsilon} |\tilde{G}_\varepsilon^i(\cdot, s)|_{\mathbf{H}^1(0,1)} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^2(0,1)} \\
 &\leq |\tilde{G}_\varepsilon^i(\cdot, s)|_{\mathbf{H}^1(0,1)}^2 + \frac{a^2}{\varepsilon^2} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^2(0,1)}^2 \\
 &\lesssim \|\mathbf{F}_\varepsilon(\cdot, s)\|_{\mathbf{L}^1(0,1)} \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^\infty(0,1)}.
 \end{aligned}$$

Therefore we have

$$(6.11) \quad \|\tilde{G}_\varepsilon^i(\cdot, s)\|_{\mathbf{L}^\infty(0,1)} \lesssim \varepsilon \|\mathbf{F}_\varepsilon(\cdot, s)\|_{\mathbf{L}^1(0,1)}.$$

This together with (6.8) yields the estimate (6.5). Finally, the estimate (6.6) follows from (6.10), (6.11), and (6.8).  $\square$

As a consequence of Lemma 6.1,

$$(6.12) \quad \mathbf{G}_\varepsilon(t, t) = \frac{\varepsilon}{2} \mathbf{A}_\varepsilon^{-1}(t) + \varepsilon \mathbf{R}_\varepsilon(t) + \tilde{\mathbf{R}}_\varepsilon(t, t),$$

where

$$(6.13) \quad |\mathbf{R}_\varepsilon(t)|_F \leq C \left( e^{-\frac{2at}{\varepsilon}} + e^{-\frac{2a(1-t)}{\varepsilon}} \right)$$

and  $\tilde{\mathbf{R}}_\varepsilon$  satisfies the estimates in Lemma 6.1. In particular, we have

$$(6.14) \quad |\mathbf{G}_\varepsilon(t, t)|_F \leq C\varepsilon \quad \text{for any } t \in (0, 1).$$

Furthermore, we obtain an asymptotic formula for the expectation term in  $F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon)$ .

**COROLLARY 6.2.** *Let  $\gamma \in (0, \frac{1}{2})$ . Let  $\{\mathbf{A}_\varepsilon\} \subset \mathbf{H}_a^1(0,1)$  such that*

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}_a^1(0,1)}^2 < \infty.$$

Then for  $\varepsilon$  small enough we have

$$(6.15) \quad -\frac{\varepsilon}{4}\mathbb{E}^{\bar{\nu}_\varepsilon}\left[\int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt\right] = -\frac{1}{4}\int_0^1 \text{Tr}(\mathbf{A}_\varepsilon(t)) dt + \mathcal{O}(\varepsilon^{1-2\gamma}).$$

*Proof.* Inserting (6.12) into (6.2) and noting that  $\mathbf{B}_\varepsilon = \varepsilon^{-2}\mathbf{A}_\varepsilon^2 - \varepsilon^{-1}\mathbf{A}'_\varepsilon$ , we get

$$(6.16) \quad -\frac{\varepsilon}{4}\mathbb{E}^{\bar{\nu}_\varepsilon}\left[\int_0^1 z(t)^T \mathbf{B}_\varepsilon(t) z(t) dt\right] = -\frac{1}{4}\int_0^1 \text{Tr}(\mathbf{A}_\varepsilon(t)) dt + \frac{\varepsilon}{4}\int_0^1 \text{Tr}(\mathbf{A}'_\varepsilon(t)\mathbf{A}_\varepsilon^{-1}(t)) dt \\ - \frac{1}{2\varepsilon}\int_0^1 \text{Tr}\left((\mathbf{A}_\varepsilon^2(t) - \varepsilon\mathbf{A}'_\varepsilon(t))(\varepsilon\mathbf{R}_\varepsilon(t) + \tilde{\mathbf{R}}_\varepsilon(t, t))\right) dt.$$

Now we bound the last three terms on the right-hand side. First, using the trace inequality

$$(6.17) \quad \text{Tr}(\mathbf{C}\mathbf{D}) \lesssim |\mathbf{C}|_F |\mathbf{D}|_F,$$

which holds for any matrices  $\mathbf{C}, \mathbf{D}$ , we obtain that

$$(6.18) \quad \left|\frac{\varepsilon}{4}\int_0^1 \text{Tr}(\mathbf{A}'_\varepsilon(t)\mathbf{A}_\varepsilon^{-1}(t)) dt\right| \lesssim \varepsilon \int_0^1 |\mathbf{A}'_\varepsilon(t)|_F |\mathbf{A}_\varepsilon^{-1}(t)|_F dt \lesssim \varepsilon^{1-\frac{\gamma}{2}}.$$

In the second inequality we used the assumption that  $\mathbf{A}_\varepsilon \geq a \cdot \mathbf{I}_d$  so that  $\text{Tr}(\mathbf{A}_\varepsilon^{-1}) \leq d/a$  and hence  $|\mathbf{A}_\varepsilon^{-1}(t)|_F \lesssim 1$ . Next, applying the Cauchy–Schwarz inequality to the last two terms on the right-hand side of (6.16) and using the assumptions on  $\mathbf{A}_\varepsilon$ , the inequality (6.13), and Lemma (6.1), we have

$$(6.19) \quad \left|\frac{1}{2\varepsilon}\int_0^1 \text{Tr}\left((\mathbf{A}_\varepsilon^2(t) - \varepsilon\mathbf{A}'_\varepsilon(t))(\varepsilon\mathbf{R}_\varepsilon(t) + \tilde{\mathbf{R}}_\varepsilon(t, t))\right) dt\right| \\ \lesssim \varepsilon^{-1}\|\mathbf{A}_\varepsilon\|_{\mathbf{L}^\infty(0,1)}^2 \left(\varepsilon\|\mathbf{R}_\varepsilon\|_{\mathbf{L}^1(0,1)} + \int_0^1 |\tilde{\mathbf{R}}_\varepsilon(t, t)| dt\right) \\ + \|\mathbf{A}'_\varepsilon\|_{\mathbf{L}^2(0,1)} \left(\varepsilon\|\mathbf{R}_\varepsilon\|_{\mathbf{L}^2(0,1)} + \left(\int_0^1 |\tilde{\mathbf{R}}_\varepsilon(t, t)|^2 dt\right)^{\frac{1}{2}}\right) \\ \lesssim \varepsilon^{-1-\gamma}(\varepsilon^2 + \varepsilon^{2-\gamma}) + \varepsilon^{-\frac{\gamma}{2}}(\varepsilon^{\frac{3}{2}} + \varepsilon^{2-\gamma}) \lesssim \varepsilon^{1-2\gamma},$$

where we have also used the assumption that  $\gamma \in (0, \frac{1}{2})$ . This finishes the proof.  $\square$

We proceed to proving bounds for the logarithmic term appearing in  $F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon)$ .

LEMMA 6.3. Let  $\gamma \in (0, \frac{1}{2})$ . Let  $\{\mathbf{A}_\varepsilon\} \subset \mathbf{H}_a^1(0, 1)$  such that

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty.$$

Then when  $\varepsilon$  is small enough,

$$(6.20) \quad C\varepsilon \log \varepsilon \leq \frac{\varepsilon}{2} \log \left( \det \left( \int_0^1 \overline{\mathbf{M}}_\varepsilon(t) \overline{\mathbf{M}}_\varepsilon(t)^T dt \right) \right) \leq 0.$$

*Proof.* We first prove the nonpositiveness. Since  $\overline{\mathbf{M}}_\varepsilon(t) = \mathbf{M}_\varepsilon(1, t)$ , where the fundamental matrix  $\mathbf{M}_\varepsilon$  satisfies (3.2) with  $\mathbf{A}$  replaced by  $\mathbf{A}_\varepsilon$ , then the  $i$ th column of  $\overline{\mathbf{M}}_\varepsilon$ , denoted by  $M_\varepsilon^i$ , satisfies

$$\partial_t M_\varepsilon^i(t, s) = -\varepsilon^{-1} \mathbf{A}_\varepsilon(t) M_\varepsilon^i(t, s), \quad M_\varepsilon^i(s, s) = e^i,$$



where  $e^i$  is the unit basis vector of  $\mathbf{R}^d$  in the  $i$ th direction. Taking the dot product of the above equation with  $M_\varepsilon^i(t, s)$  and then integrating from  $s$  to  $t$  implies that

$$|M_\varepsilon^i(t, s)|^2 = -\frac{2}{\varepsilon} \int_s^t M_\varepsilon^i(r, s)^T \mathbf{A}_\varepsilon(r) M_\varepsilon^i(r, s) dr \leq -\frac{2a}{\varepsilon} \int_s^t |M_\varepsilon^i(r, s)|^2 dr.$$

Consequently,  $|M_\varepsilon^i(t, s)| \leq e^{-\frac{a(t-s)}{\varepsilon}}$  for any  $0 \leq s \leq t \leq 1$ . Hence each entry of  $\overline{\mathbf{M}}_\varepsilon(t)$  can be bounded from above by  $e^{-\frac{a(1-t)}{\varepsilon}}$ . As a result, for sufficiently small  $\varepsilon$ , we have

$$\det \left( \int_0^1 \overline{\mathbf{M}}_\varepsilon(t) \overline{\mathbf{M}}_\varepsilon(t)^T dt \right) \leq C\varepsilon < 1.$$

The upper bound of (6.20) thus follows. On the other hand, applying the determinant inequality (B.5) to the matrix function  $\overline{\mathbf{M}}_\varepsilon(t) \overline{\mathbf{M}}_\varepsilon(t)^T$  and the equality (B.3) yields

$$\begin{aligned} \frac{\varepsilon}{2} \log \left( \det \left( \int_0^1 \overline{\mathbf{M}}_\varepsilon(t) \overline{\mathbf{M}}_\varepsilon(t)^T dt \right) \right) &\geq \frac{\varepsilon d}{2} \log \left( \int_0^1 \det(\overline{\mathbf{M}}_\varepsilon(t))^{\frac{2}{d}} dt \right) \\ (6.21) \qquad \qquad \qquad &= \frac{\varepsilon d}{2} \log \left( \int_0^1 \exp \left( -\frac{2}{\varepsilon d} \int_t^1 \text{Tr}(\mathbf{A}_\varepsilon(s)) ds \right) dt \right). \end{aligned}$$

Moreover, from the assumption that  $\varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty$  and the fact that  $\mathbf{H}^1(0,1)$  is embedded into  $\mathbf{L}^\infty(0,1)$ , we obtain that

$$\int_t^1 \text{Tr}(\mathbf{A}_\varepsilon(s)) ds \leq (1-t) \|\mathbf{A}_\varepsilon\|_{\mathbf{L}^\infty(0,1)} \leq C\varepsilon^{-\frac{\gamma}{2}}(1-t).$$

Combining this with (6.21) gives

$$\begin{aligned} \frac{\varepsilon}{2} \log \left( \det \left( \int_0^1 \overline{\mathbf{M}}_\varepsilon(t) \overline{\mathbf{M}}_\varepsilon(t)^T dt \right) \right) &\geq \frac{\varepsilon d}{2} \log \left( \int_0^1 \exp \left( -\frac{2C}{\varepsilon^{1+\frac{\gamma}{2}} d} (1-t) \right) dt \right) \\ (6.22) \qquad \qquad \qquad &= \frac{\varepsilon d}{2} \log \left( \frac{\varepsilon^{1+\frac{\gamma}{2}} d}{2C} \left( 1 - e^{-\frac{2C}{\varepsilon^{1+\frac{\gamma}{2}} d}} \right) \right) \\ &\geq C\varepsilon \log \varepsilon \end{aligned}$$

for sufficiently small  $\varepsilon$ . This completes the proof.  $\square$

Recall the definition of  $F_\varepsilon^{(2)}$  in (4.11). Then the following proposition, containing the asymptotic expression for  $F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon)$ , is a direct consequence of Corollary 6.2 and Lemma 6.3.

**PROPOSITION 6.4.** *Let  $\{\mathbf{A}_\varepsilon\} \subset \mathbf{H}_a^1(0,1)$  such that  $\limsup_{\varepsilon \rightarrow 0} \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty$  with  $\gamma \in (0, \frac{1}{2})$ . Then it holds that*

$$(6.23) \qquad F_\varepsilon^{(2)}(\mathbf{A}_\varepsilon) = \frac{1}{4} \int_0^1 \text{Tr}(\mathbf{A}_\varepsilon(t)) dt + \mathcal{O}(\varepsilon^{1-\gamma})$$

when  $\varepsilon$  is small enough.

**6.2. Asymptotics of  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$ .** In this subsection, we seek an asymptotic expression for  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$  when it is uniformly bounded with respect to  $\varepsilon$ . We start by showing that the boundedness of  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$  implies the boundedness of  $\|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)}$ .

LEMMA 6.5. Assume that  $(m_\varepsilon, \mathbf{A}_\varepsilon) \in \mathcal{H}$  and that  $\limsup_{\varepsilon \rightarrow 0} F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon) < \infty$ . Then we have  $\limsup_{\varepsilon \rightarrow 0} \|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)} < \infty$ .

*Proof.* Recalling that  $\Psi_\varepsilon(x) = \frac{1}{2}|\nabla V(x)|^2 - \varepsilon\Delta V(x)$  and that  $\nu_\varepsilon = N(m_\varepsilon, \Sigma_\varepsilon)$ , we can rewrite  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$  as

$$\begin{aligned} F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon) &= \frac{\varepsilon}{4} \int_0^1 |m'_\varepsilon(t)|^2 dt + \frac{1}{2\varepsilon} \mathbb{E}^{\bar{\nu}_\varepsilon} \left[ \int_0^1 \Psi_\varepsilon(z(t) + m_\varepsilon(t)) dt \right] \\ &:= F_\varepsilon^{(3)}(m_\varepsilon, \mathbf{A}_\varepsilon) + F_\varepsilon^{(4)}(m_\varepsilon, \mathbf{A}_\varepsilon), \end{aligned}$$

where

$$F_\varepsilon^{(3)}(m_\varepsilon, \mathbf{A}_\varepsilon) := \frac{\varepsilon}{4} \int_0^1 |m'_\varepsilon(t)|^2 dt + \frac{1}{8\varepsilon} \int_0^1 \mathbb{E}^{\bar{\nu}_\varepsilon} [|\nabla V(z(t) + m_\varepsilon(t))|^2] dt$$

and

$$\begin{aligned} F_\varepsilon^{(4)}(m_\varepsilon, \mathbf{A}_\varepsilon) &:= \frac{1}{8\varepsilon} \int_0^1 \mathbb{E}^{\bar{\nu}_\varepsilon} [|\nabla V(z(t) + m_\varepsilon(t))|^2 - 4\varepsilon\Delta V(z(t) + m_\varepsilon(t))] dt \\ &= \frac{1}{8\varepsilon} \int_0^1 \mathbb{E}^{\nu_\varepsilon} [|\nabla V(x(t))|^2 - 4\varepsilon\Delta V(x(t))] dt. \end{aligned}$$

First, from (2.7) of Remark 2.2 we can obtain immediately that

$$\liminf_{\varepsilon \rightarrow 0} F_\varepsilon^{(4)}(m_\varepsilon, \mathbf{A}_\varepsilon) > -\infty.$$

This, together with the assumption that  $\limsup_{\varepsilon \rightarrow 0} F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon) < \infty$ , implies

$$\limsup_{\varepsilon \rightarrow 0} F_\varepsilon^{(3)}(m_\varepsilon, \mathbf{A}_\varepsilon) < \infty.$$

We now show that this implies the uniform boundedness of  $\|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)}$ .

We prove a lower bound for  $F_\varepsilon^{(3)}(m_\varepsilon, \mathbf{A}_\varepsilon)$ . Given any  $R > 0$ , define  $T_\varepsilon^R := \{t \in (0, 1) : |m_\varepsilon(t)| > R\}$ , which is an open set on  $(0, 1)$ . By restricting the second integral and expectation over a smaller set, we have

$$(6.24) \quad F_\varepsilon^{(3)}(m_\varepsilon, \mathbf{A}_\varepsilon) \geq \frac{\varepsilon}{4} \int_0^1 |m'_\varepsilon(t)|^2 dt + \frac{1}{8\varepsilon} \int_{T_\varepsilon^R} \mathbb{E}^{\bar{\nu}_\varepsilon} [|\nabla V(m_\varepsilon(t) + z(t))|^2 \mathbf{1}_{\{|z(t)| \leq \varepsilon^{1/4}\}}] dt.$$

Consider  $(t, \omega)$  such that  $|m_\varepsilon(t)| > R$  and  $|z(t, \omega)| \leq \varepsilon^{1/4}$ . If  $\varepsilon > 0$  is small enough to satisfy  $\varepsilon^{1/4} < R/2$ , then

$$|m_\varepsilon(t) + z(t, \omega)| \geq |m_\varepsilon(t)| - |z(t, \omega)| \geq \frac{|m_\varepsilon(t)|}{2}.$$

Combining this with the monotonicity condition (A-6) yields that

$$\begin{aligned} (6.25) \quad & \frac{1}{8\varepsilon} \int_{T_\varepsilon^R} \mathbb{E}^{\bar{\nu}_\varepsilon} [|\nabla V(m_\varepsilon(t) + z(t))|^2 \mathbf{1}_{\{|z(t)| \leq \varepsilon^{1/4}\}}] dt \\ & \geq \frac{1}{8\varepsilon} \int_{T_\varepsilon^R} |\nabla V(m_\varepsilon(t)/2)|^2 \bar{\nu}_\varepsilon \left( \{|z(t)| \leq \varepsilon^{1/4}\} \right) dt \geq \frac{1}{16\varepsilon} \int_{T_\varepsilon^R} |\nabla V(m_\varepsilon(t)/2)|^2 dt \end{aligned}$$

when  $\varepsilon > 0$  is small enough. We have used the fact that  $\bar{\nu}_\varepsilon(\{|z(t)| \leq \varepsilon^{\frac{1}{4}}\}) \geq 1/2$  for any  $t \in (0, 1)$  and small  $\varepsilon$ . This is because  $z(t)$  is a centered Gaussian random variable with covariance  $2\mathbf{G}_\varepsilon(t, t)$  (see (3.11)). In addition, we know from (6.14) that  $|\mathbf{G}_\varepsilon(t, t)|_F \leq C\varepsilon$  for any  $t \in (0, 1)$ , and hence  $\bar{\nu}_\varepsilon(\{|z(t)| \leq \varepsilon^{\frac{1}{4}}\}) \rightarrow 1$  when  $\varepsilon \rightarrow 0$ . Let  $\tilde{m}_\varepsilon = m_\varepsilon/2$ . From (6.24), (6.25), and the uniform boundedness of  $F_\varepsilon^{(3)}(m_\varepsilon, \mathbf{A}_\varepsilon)$  we obtain

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \int_{T_\varepsilon^R} |\tilde{m}'_\varepsilon(t)|^2 dt + \frac{1}{16\varepsilon} \int_{T_\varepsilon^R} |\nabla V(\tilde{m}_\varepsilon(t))|^2 dt < \infty.$$

Then application of the elementary inequality  $2ab \leq a^2 + b^2$  yields

$$\limsup_{\varepsilon \rightarrow 0} \int_{T_\varepsilon^R} |\tilde{m}'_\varepsilon(t)| |\nabla V(\tilde{m}_\varepsilon(t))| dt < \infty.$$

Choosing a sufficiently large  $R$  and by the coercivity condition (A-3), we have

$$(6.26) \quad \limsup_{\varepsilon \rightarrow 0} \int_{T_\varepsilon^R} |\tilde{m}'_\varepsilon(t)| < \infty.$$

Now we conclude the uniform boundedness of  $\|m_\varepsilon\|_{L^\infty(0,1)}$  by applying the same argument used for proving Theorem 1.2 in [21]. Specifically, since  $m_\varepsilon$  is continuous on  $(0, 1)$ ,  $T_\varepsilon^R$  is open on  $(0, 1)$ , and we can write  $T_\varepsilon^R = \cup_{i=1}^\infty (a_\varepsilon^i, b_\varepsilon^i)$ . Suppose that  $T_\varepsilon^R$  is empty; then  $|m_\varepsilon(t)| \leq R$  for all  $t \in (0, 1)$ . Otherwise, consider  $\tilde{m}_\varepsilon(t)$  with  $t \in (a_\varepsilon^i, b_\varepsilon^i)$ . Obviously at least one of the end-points of the subinterval, say  $a_\varepsilon^i$ , is not an end-point of  $(0, 1)$ . Then we should have  $|m_\varepsilon(a_\varepsilon^i)| = R$  and hence  $|\tilde{m}_\varepsilon(a_\varepsilon^i)| = 2R$ . Thus we get from the fundamental theorem of calculus that

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \sup_{t \in (a_\varepsilon^i, b_\varepsilon^i)} |\tilde{m}_\varepsilon(t)| &\leq \limsup_{\varepsilon \rightarrow 0} \left( |\tilde{m}_\varepsilon(a_\varepsilon^i)| + \sup_{t \in (a_\varepsilon^i, b_\varepsilon^i)} \left| \int_{a_\varepsilon^i}^t \tilde{m}'_\varepsilon(s) ds \right| \right) \\ &\leq 2R + \limsup_{\varepsilon \rightarrow 0} \int_{T_\varepsilon^R} |\tilde{m}'_\varepsilon(t)| < \infty, \end{aligned}$$

where the last inequality follows from (6.26). Therefore  $\limsup_{\varepsilon \rightarrow 0} \|m_\varepsilon\|_{L^\infty(0,1)} < \infty$ .  $\square$

Next, the expectation term of  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$  can be simplified under the condition that  $\|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)}$  is uniformly bounded.

**LEMMA 6.6.** *Let  $(m_\varepsilon, \mathbf{A}_\varepsilon) \in \mathcal{H}$ . Assume that  $\limsup_{\varepsilon \rightarrow 0} \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty$  with  $\gamma \in (0, \frac{1}{2})$  and that  $\limsup_{\varepsilon \rightarrow 0} \|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)} < \infty$ . Then for  $\varepsilon > 0$  small enough we have*

$$\begin{aligned} F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon) &= E_\varepsilon(m_\varepsilon) - \frac{1}{2} \int_0^1 \Delta V(m_\varepsilon(t)) dt \\ &\quad + \frac{1}{4} \int_0^1 (D^2 V(m_\varepsilon(t))^2 + D^3 V(m_\varepsilon(t)) \cdot \nabla V(m_\varepsilon(t))) : \mathbf{A}_\varepsilon^{-1}(t) dt + \mathcal{O}(\varepsilon^{\frac{1}{2}}). \end{aligned}$$

*Proof.* Remember that

$$F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon) = \frac{\varepsilon}{4} \int_0^1 |m'_\varepsilon(t)|^2 dt + \frac{1}{2\varepsilon} \mathbb{E}^{\bar{\nu}_\varepsilon} \left[ \int_0^1 \Psi_\varepsilon(m_\varepsilon(t) + z_\varepsilon(t)) dt \right].$$

To evaluate the expectation term of  $F_\varepsilon^{(1)}(m_\varepsilon, \mathbf{A}_\varepsilon)$ , we use the following multivariable Taylor's formula for  $\Psi_\varepsilon$ :

$$\Psi_\varepsilon(x(t)) = \Psi_\varepsilon(m_\varepsilon(t)) + \nabla \Psi_\varepsilon(m_\varepsilon(t)) \cdot z_\varepsilon(t) + \frac{1}{2} z_\varepsilon(t)^T D^2 \Psi_\varepsilon(m_\varepsilon(t)) z_\varepsilon(t) + r_\varepsilon(t),$$

where the reminder term  $r_\varepsilon$  is given in integral form by

$$r_\varepsilon(t) = \sum_{|\alpha|=3} \frac{z_\varepsilon^\alpha(t)}{\alpha!} \int_0^1 \partial^\alpha \Psi_\varepsilon(m_\varepsilon(t) + \xi z_\varepsilon(t)) (1-\xi)^3 d\xi.$$

Here  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d)$  is a multi-index, and we use the notational convention  $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_d^{\alpha_d}$  and  $\partial^\alpha f = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_d^{\alpha_d} f$ . Then, using again the fact that  $z_\varepsilon(t) \sim N(0, 2\mathbf{G}_\varepsilon(t, t))$ , we obtain that

$$\begin{aligned} (6.27) \quad & \frac{1}{2\varepsilon} \mathbb{E}^{\bar{\nu}_\varepsilon} \left[ \int_0^1 \Psi_\varepsilon(m_\varepsilon(t) + z_\varepsilon(t)) dt \right] \\ &= \frac{1}{2\varepsilon} \int_0^1 \Psi_\varepsilon(m_\varepsilon(t)) dt + \frac{1}{2\varepsilon} \int_0^1 D^2 \Psi_\varepsilon(m_\varepsilon(t)) : \mathbf{G}_\varepsilon(t, t) dt + \frac{1}{2\varepsilon} \int_0^1 \mathbb{E}^{\bar{\nu}_\varepsilon} [r_\varepsilon(t)] dt. \end{aligned}$$

Recalling that  $\Psi_\varepsilon(x) = \frac{1}{2} |\nabla V(x)|^2 - \varepsilon \Delta V(x)$ , we have

$$D^2 \Psi_\varepsilon(x) = (D^2 V(x))^2 + D^3 V(x) \cdot \nabla V(x) - \varepsilon D^2(\Delta V(x)).$$

From this equation, the expression (6.12) for  $\mathbf{G}_\varepsilon(t, t)$ , and the uniform boundedness of  $\|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)}$ , the second term on the right-hand side of (6.27) becomes

$$\begin{aligned} (6.28) \quad & \frac{1}{2\varepsilon} \int_0^1 D^2 \Psi_\varepsilon(m_\varepsilon(t)) : \mathbf{G}_\varepsilon(t, t) dt \\ &= \frac{1}{4} \int_0^1 \left( (D^2 V(m_\varepsilon(t)))^2 + D^3 V(m_\varepsilon(t)) \cdot \nabla V(m_\varepsilon(t)) \right) : \mathbf{A}_\varepsilon^{-1}(t) dt + \mathcal{O}(\varepsilon). \end{aligned}$$

Next we claim that the last term on the right-hand side of (6.27) is of order  $\mathcal{O}(\varepsilon^{\frac{3}{2}})$ . Indeed, from the assumption (2.4) and the fact that  $z_\varepsilon(t) = N(0, 2\mathbf{G}_\varepsilon(t, t))$  with  $\mathbf{G}_\varepsilon(t, t)$  satisfying the estimate (6.14), we have

$$\begin{aligned} (6.29) \quad & \mathbb{E}^{\nu_\varepsilon} [r_\varepsilon(t)] \leq \sum_{|\alpha|=3} \frac{1}{\alpha!} \max_{\xi \in [0,1]} \left\{ \mathbb{E}^{\nu_\varepsilon} \left[ |z_\varepsilon(t)|^3 \partial^\alpha \Psi_\varepsilon(m_\varepsilon(t) + \xi z_\varepsilon(t)) \right] \right\} \\ & \leq \frac{C_1}{\sqrt{(4\pi)^d \det(2\mathbf{G}_\varepsilon(t, t))}} \max_{\xi \in [0,1]} \left\{ \int_{\mathbf{R}^d} |x|^3 e^{C_2 |m_\varepsilon(t) + \xi x|^\alpha} \cdot e^{-\frac{1}{4} x^T \mathbf{G}_\varepsilon(t, t)^{-1} x} dx \right\} \\ & \leq \frac{C_1}{\sqrt{(4\pi)^d \det(\mathbf{G}_\varepsilon(t, t))}} e^{2C_2 \|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)}^2} \int_{\mathbf{R}^d} e^{2C_2 |x|^\alpha} |x|^3 e^{-\frac{1}{4} x^T \mathbf{G}_\varepsilon(t, t)^{-1} x} dx \\ & \leq \frac{C_2}{\sqrt{(4\pi)^d}} |\mathbf{G}_\varepsilon(t, t)|_F^{\frac{3}{2}} \cdot e^{2C_2 \|m_\varepsilon\|_{\mathbf{L}^\infty(0,1)}^2} \cdot \int_{\mathbf{R}^d} |x|^3 e^{-\frac{|x|^2}{4}} dx \leq C \varepsilon^{\frac{3}{2}} \end{aligned}$$

when  $\varepsilon$  is small enough. Notice that in the last two inequalities above we used the fact that  $\alpha \in [0, 2)$  and that  $|\mathbf{G}_\varepsilon(t, t)|_F \leq C\varepsilon$ , so that  $e^{2C_2 |x|^\alpha}$  can be absorbed by  $e^{\frac{1}{4} x^T \mathbf{G}_\varepsilon(t, t)^{-1} x}$  for large  $x$ . Then the desired result follows from (6.27), (6.28), and (6.29).  $\square$

### 6.3. Proof of main results.

*Proof of Proposition 4.4.* The proposition follows directly from the definition of  $F_\varepsilon$ , Proposition 6.4, Lemma 6.6, and the following equalities:

$$\begin{aligned}
 (6.30) \quad & -2\Delta V + (D^2V)^2 : \mathbf{A}^{-1} + \text{Tr}(\mathbf{A}) \\
 & = -2\text{Tr}(D^2V) + \text{Tr}((D^2V)^2 \mathbf{A}^{-1}) + \text{Tr}(\mathbf{A}^2 \mathbf{A}^{-1}) \\
 & = -\text{Tr}((\mathbf{A} D^2V \mathbf{A}^{-1}) - \text{Tr}(\mathbf{A}^{-1} D^2V \mathbf{A}) + \text{Tr}((D^2V)^2 \mathbf{A}^{-1}) + \text{Tr}(\mathbf{A}^2 \mathbf{A}^{-1}) \\
 & = (D^2V - \mathbf{A})^2 : \mathbf{A}^{-1},
 \end{aligned}$$

which are valid for any  $V \in C^2(\mathbf{R}^d)$  and any positive definite matrix  $\mathbf{A}$ .  $\square$

The following lemma shows that  $\varepsilon \log(Z_{\mu,\varepsilon})$  is bounded from above.

LEMMA 6.7. *There exists  $C > 0$  depending only on the potential  $V$  such that the following holds:*

$$(6.31) \quad \limsup_{\varepsilon \rightarrow 0} \varepsilon \log(Z_{\mu,\varepsilon}) \leq C.$$

*Proof.* Recall that

$$Z_{\mu,\varepsilon} = \mathbb{E}^{\mu_0} \left[ \exp \left( -\frac{1}{2\varepsilon^2} \int_0^1 |\nabla V(x(t))|^2 - \varepsilon \Delta V(x(t)) dt \right) \right].$$

From (2.7) of Remark 2.2,

$$Z_{\mu,\varepsilon} \leq \exp \left( \frac{C}{\varepsilon} \right)$$

with some  $C > 0$ . This proves (6.31).  $\square$

*Proof of Proposition 4.5.* Assume that  $\limsup_n F_{\varepsilon_n}(m_n, \mathbf{A}_n) < \infty$ . Since the Kullback–Leibler divergence  $D_{\text{KL}}(\nu_{\varepsilon_n} \| \mu_{\varepsilon_n})$  is always nonnegative, it follows from (3.23) and Lemma 6.7 that

$$(6.32) \quad \limsup_{n \rightarrow \infty} \varepsilon_n \tilde{D}_{\text{KL}}(\nu_{\varepsilon_n} \| \mu_{\varepsilon_n}) \geq -C$$

for some  $C > 0$ . This, together with the assumption that  $\limsup_n F_{\varepsilon_n}(m_n, \mathbf{A}_n) < \infty$ , implies that  $\limsup_n \varepsilon_n^\gamma \|\mathbf{A}_n\|_{\mathbf{H}^1(0,1)}^2 < \infty$ . Then from Proposition 6.4 and noting that  $\mathbf{A}(\cdot) - a \cdot \mathbf{I}_d \geq 0$ , we obtain

$$\limsup_n F_{\varepsilon_n}^{(2)}(\mathbf{A}_n) \geq \limsup_n \frac{1}{4} \int_0^1 \text{Tr}(\mathbf{A}_n(t)) dt \geq \frac{da}{4}.$$

Hence we have  $\limsup_n F_{\varepsilon_n}^{(1)}(m_n, \mathbf{A}_n) < \infty$ . Then Lemma 6.5 implies that

$$\limsup_n \|m_n\|_{\mathbf{L}^\infty(0,1)} < \infty.$$

Hence as a consequence of Lemma 6.6,

$$\begin{aligned}
 (6.33) \quad F_{\varepsilon_n}(m_n, \mathbf{A}_n) & = E_{\varepsilon_n}(m_n) + \frac{1}{4} \int_0^1 (D^2V(m_n(t)) - \mathbf{A}_n(t))^2 : \mathbf{A}_n^{-1}(t) dt \\
 & + \int_0^1 (D^3V(m_n(t)) \cdot \nabla V(m_n(t))) : \mathbf{A}_n^{-1}(t) dt + \varepsilon_n^\gamma \|\mathbf{A}_n\|_{\mathbf{H}^1(0,1)}^2 + \mathcal{O}(\varepsilon_n^{\frac{1}{2}}).
 \end{aligned}$$

The second term on the right-hand side above is nonnegative. In addition, owing to the trace inequality (6.17) and the fact that  $\mathbf{A}_n \geq a \cdot \mathbf{I}_d$ ,

$$(6.34) \quad \limsup_n \left| \frac{1}{4} \int_0^1 D^3 V(m_n(t)) \cdot \nabla V(m_n(t)) : \mathbf{A}_n^{-1}(t) dt \right| \\ \leq \limsup_n \int_0^1 \left| (D^3 V(m_n(t)) \cdot \nabla V(m_n(t)))^2 \right|_F |\mathbf{A}_n^{-1}(t)|_F dt < \infty.$$

This implies from (6.33) that  $\limsup_n E_{\varepsilon_n}(m_n) < \infty$ . By the compactness result of Lemma 4.2, there exist  $m \in \mathbf{BV}(0, 1; \mathcal{E})$  and a subsequence  $m_{n_k}$  such that  $m_{n_k} \rightarrow m$  in  $\mathbf{L}^1(0, 1)$ . Moreover, we know from the above reasoning that

$$(6.35) \quad \limsup_n \frac{1}{4} \int_0^1 (D^2 V(m_n(t)) - \mathbf{A}_n(t))^2 : \mathbf{A}_n^{-1}(t) dt < \infty,$$

from which we can conclude that  $\sup_n \|\mathbf{A}_n\|_{\mathbf{L}^1(0,1)} < \infty$ . Indeed,

$$(6.36) \quad \int_0^1 (D^2 V(m_n(t)) - \mathbf{A}_n(t))^2 : \mathbf{A}_n^{-1}(t) dt \\ = \int_0^1 \text{Tr} \left( (D^2 V(m_n(t)))^2 \mathbf{A}_n^{-1}(t) \right) dt - 2 \int_0^1 \text{Tr} (D^2 V(m_n(t))) dt + \int_0^1 \text{Tr}(\mathbf{A}_n(t)) dt.$$

The first term on the right-hand side above is nonnegative. The second term is clearly bounded since  $\|m_n\|_{\mathbf{L}^\infty(0,1)}$  is uniformly bounded. Hence  $\sup_n \|\mathbf{A}_n\|_{\mathbf{L}^1(0,1)} < \infty$  follows from (6.35), (6.36), and the inequality  $|\mathbf{A}|_F \leq \text{Tr}(\mathbf{A})$ , which holds for any positive definite matrix  $\mathbf{A}$ .  $\square$

The proof of  $\Gamma$ -limit of  $F_\varepsilon$  is presented in what follows.

*Proof of Theorem 4.6.* We start by proving the liminf inequality, i.e.,

$$F(m, \mathbf{A}) \leq \liminf_{\varepsilon \rightarrow 0} F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon)$$

for any sequence  $\{(m_\varepsilon, \mathbf{A}_\varepsilon)\}$  such that  $(m_\varepsilon, \mathbf{A}_\varepsilon) \rightarrow (m, \mathbf{A})$  in  $\mathcal{X}$ , or equivalently  $m_\varepsilon \rightarrow m$  and  $\mathbf{A}_\varepsilon \rightarrow \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$ . We may assume that  $\liminf_{\varepsilon \rightarrow 0} F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon) < \infty$ , since otherwise there is nothing to prove. Then by the same argument used in the proof of Proposition 4.5, one can get  $\limsup_{\varepsilon \rightarrow 0} \varepsilon^\gamma \|\mathbf{A}_\varepsilon\|_{\mathbf{H}^1(0,1)}^2 < \infty$ . Let  $\{\varepsilon_k\}$  be a sequence such that  $\varepsilon_k \rightarrow 0$  as  $k \rightarrow \infty$  and  $\lim_{k \rightarrow \infty} F_{\varepsilon_k}(m_{\varepsilon_k}, \mathbf{A}_{\varepsilon_k}) = \liminf_{\varepsilon \rightarrow 0} F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon) < \infty$ . Since  $\mathbf{A}_{\varepsilon_k} \geq a \cdot \mathbf{I}_d$  a.e., it follows from  $\mathbf{A}_{\varepsilon_k} \rightarrow \mathbf{A}$  and Mazur's lemma (Lemma C.10) that the limit  $\mathbf{A} \geq a \cdot \mathbf{I}_d$  a.e. According to Proposition 6.4 and  $\mathbf{A}_{\varepsilon_k} \rightarrow \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$ , it holds that

$$\lim_{k \rightarrow \infty} F_{\varepsilon_k}^{(2)}(\mathbf{A}_{\varepsilon_k}) = \frac{1}{4} \int_0^1 \text{Tr}(\mathbf{A}(t)) dt \geq \frac{da}{4}.$$

Then it follows that  $\lim_{k \rightarrow \infty} F_{\varepsilon_k}^{(1)}(m_{\varepsilon_k}, \mathbf{A}_{\varepsilon_k}) < \infty$ . From Lemma 6.5 we obtain that  $\|m_{\varepsilon_k}\|_{\mathbf{L}^\infty(0,1)}$  is uniformly bounded. Hence as a consequence of Lemma 6.6,

$$F_{\varepsilon_k}^{(1)}(m_{\varepsilon_k}, \mathbf{A}_{\varepsilon_k}) = E_\varepsilon(m_{\varepsilon_k}) - \frac{1}{2} \int_0^1 \Delta V(m_{\varepsilon_k}(t)) dt \\ + \frac{1}{4} \int_0^1 (D^2 V(m_{\varepsilon_k}(t))^2 + D^3 V(m_{\varepsilon_k}(t)) \cdot \nabla V(m_{\varepsilon_k}(t))) : \mathbf{A}_{\varepsilon_k}^{-1}(t) dt + O(\varepsilon_k^{\frac{1}{2}}).$$

In addition, it follows from the uniform boundedness of  $\|m_{\varepsilon_k}\|_{\mathbf{L}^\infty(0,1)}$  and  $\mathbf{A}_{\varepsilon_k} \rightharpoonup \mathbf{A}$  in  $\mathbf{L}^1(0,1)$  that

$$\limsup_{k \rightarrow \infty} \left\{ \left| -\frac{1}{2} \int_0^1 \Delta V(m_{\varepsilon_k}(t)) dt + \frac{1}{4} \int_0^1 (D^2 V(m_{\varepsilon_k}(t))^2 + D^3 V(m_{\varepsilon_k}(t)) \cdot \nabla V(m_{\varepsilon_k}(t))) : \mathbf{A}_{\varepsilon_k}^{-1}(t) dt \right| \right\} < \infty.$$

This in turn implies that  $\limsup_{k \rightarrow \infty} E_{\varepsilon_k}(m_{\varepsilon_k}) < \infty$ . By the compactness result in Lemma 4.2, we have  $m \in \mathbf{BV}((0,1); \mathcal{E})$ . Furthermore, by passing to a subsequence, we may assume further that  $m_{\varepsilon_k} \rightarrow m$  a.e. on  $[0,1]$ . Since  $m$  takes values in  $\mathcal{E}$  a.e. on  $[0,1]$ , we use the definition of  $D^3 V$  and the dominated convergence theorem to conclude that

$$(6.37) \quad \int_0^1 (D^3 V(m_{\varepsilon_k}(t)) \cdot \nabla V(m_{\varepsilon_k}(t))) : \mathbf{A}_{\varepsilon_k}^{-1}(t) dt \rightarrow 0.$$

In fact, similar to (6.34), we have

$$\int_0^1 \left| (D^3 V(m_{\varepsilon_k}(t)) \cdot \nabla V(m_{\varepsilon_k}(t))) : \mathbf{A}_{\varepsilon_k}^{-1}(t) \right| dt < \infty.$$

In addition, since  $\nabla V(m_{\varepsilon_k}(t)) \rightarrow 0$  a.e. on  $[0,1]$  and  $\mathbf{A}_{\varepsilon_k} \geq a \cdot \mathbf{I}_d$ , we have

$$(6.38) \quad \begin{aligned} & \left| (D^3 V(m_{\varepsilon_k}(t)) \cdot \nabla V(m_{\varepsilon_k}(t))) : \mathbf{A}_{\varepsilon_k}^{-1}(t) \right| \\ & \leq \left| D^3 V(m_{\varepsilon_k}(t)) \cdot \nabla V(m_{\varepsilon_k}(t)) \right|_F \left| \mathbf{A}_{\varepsilon_k}^{-1}(t) \right|_F \rightarrow 0 \end{aligned}$$

a.e. on  $[0,1]$ . This proves (6.37). Now we claim that

$$(6.39) \quad \begin{aligned} & \int_0^1 (D^2 V(m(t)) - \mathbf{A}(t))^2 : \mathbf{A}^{-1}(t) dt \\ & \leq \liminf_{k \rightarrow \infty} \int_0^1 (D^2 V(m_{\varepsilon_k}(t)) - \mathbf{A}_{\varepsilon_k}(t))^2 : \mathbf{A}_{\varepsilon_k}^{-1}(t) dt \end{aligned}$$

when  $m_{\varepsilon_k} \rightarrow m$  in  $\mathbf{L}^1(0,1)$  with  $m(\cdot) \in \mathcal{E}$  a.e. and  $\mathbf{A}_{\varepsilon_k} \rightharpoonup \mathbf{A}$  in  $\mathbf{L}^1(0,1)$ . In fact, the weak convergence  $\mathbf{A}_{\varepsilon_k} \rightharpoonup \mathbf{A}$  in  $\mathbf{L}^1(0,1)$  directly implies that

$$\int_0^1 \mathbf{A}_{\varepsilon_k}^2(t) : \mathbf{A}_{\varepsilon_k}^{-1}(t) - \mathbf{A}^2(t) : \mathbf{A}^{-1}(t) dt = \int_0^1 \text{Tr}(\mathbf{A}_{\varepsilon_k}(t) - \mathbf{A}(t)) dt \rightarrow 0.$$

In addition, thanks to the uniform boundedness of  $\|m_{\varepsilon_k}\|_{\mathbf{L}^\infty(0,1)}$  and the strong convergence  $m_{\varepsilon_k} \rightarrow m$  in  $\mathbf{L}^1(0,1)$ , we obtain from the dominated convergence that

$$\begin{aligned} & \int_0^1 (D^2 V(m_{\varepsilon_k}(t)) \mathbf{A}_{\varepsilon_k}(t) + \mathbf{A}_{\varepsilon_k}(t) D^2 V(m_{\varepsilon_k}(t))) : \mathbf{A}_{\varepsilon_k}^{-1}(t) dt \\ & - \int_0^1 (D^2 V(m(t)) \mathbf{A}(t) + \mathbf{A}(t) D^2 V(m(t))) : \mathbf{A}^{-1}(t) dt \\ & = 2 \int_0^1 \text{Tr}(D^2 V(m_{\varepsilon_k}(t)) - D^2 V(m(t))) dt \rightarrow 0. \end{aligned}$$

Therefore to prove (6.39), it suffices to show

$$(6.40) \quad \int_0^1 D^2(m(t))^2 : \mathbf{A}^{-1}(t) dt \leq \liminf_{k \rightarrow \infty} \int_0^1 D^2 V(m_{\varepsilon_k}(t))^2 : \mathbf{A}_{\varepsilon_k}^{-1}(t) dt =: \theta.$$

To that end, let  $\mathbf{B}(\cdot) := D^2 V(m(\cdot))^2$ . Noting that  $m(\cdot) \in \mathcal{E}$  a.e., we know from (A-2) of Assumptions 2.1 that  $\mathbf{B}(\cdot)$  is positive definite a.e. Define the functional

$$(6.41) \quad \mathcal{M}(\mathbf{A}) = \int_0^1 \mathbf{B}(t) : \mathbf{A}^{-1}(t) dt$$

over the set  $\mathbf{L}_a^1(0, 1)$ . Then (6.40) becomes

$$(6.42) \quad \mathcal{M}(\mathbf{A}) \leq \liminf_{k \rightarrow \infty} \mathcal{M}(\mathbf{A}_k).$$

Note that  $\mathbf{L}_a^1(0, 1)$  is a convex subset of the space  $\mathbf{L}^1(0, 1)$ . We first claim that the functional  $\mathcal{M}$  is convex on  $\mathbf{L}_a^1(0, 1)$ . In fact, for any  $\mathbf{A}_1, \mathbf{A}_2 \in \mathbf{L}_a^1(0, 1)$ ,  $\alpha \in (0, 1)$ ,

$$\begin{aligned} \mathcal{M}(\alpha \mathbf{A}_1 + (1 - \alpha) \mathbf{A}_2) &= \int_0^1 \text{Tr}(\mathbf{B}(t)(\alpha \mathbf{A}_1(t) + (1 - \alpha) \mathbf{A}_2(t))^{-1}) dt \\ &= \int_0^1 \text{Tr}(\alpha \mathbf{A}_1(t) \mathbf{B}^{-1}(t) + (1 - \alpha) \mathbf{A}_2(t) \mathbf{B}^{-1}(t)) dt \\ &\leq \alpha \int_0^1 \text{Tr}(\mathbf{A}_1(t) \mathbf{B}^{-1}(t)) dt + (1 - \alpha) \int_0^1 \text{Tr}(\mathbf{A}_2(t) \mathbf{B}^{-1}(t)) dt \\ &= \alpha \mathcal{M}(\mathbf{A}_1) + (1 - \alpha) \mathcal{M}(\mathbf{A}_2), \end{aligned}$$

where we used the trace inequality  $\text{Tr}((\mathbf{C} + \mathbf{D})^{-1}) \leq \text{Tr}(\mathbf{C}^{-1}) + \text{Tr}(\mathbf{D}^{-1})$  for positive definite matrices  $\mathbf{C}, \mathbf{D}$ . Now we prove (6.42) by employing the convexity of  $\mathcal{M}$ . First, by passing a subsequence (without relabeling), we may assume that  $\mathcal{M}(\mathbf{A}_k)$  converges to  $\theta$ . According to Mazur's lemma (Lemma C.10), there exists a convex combination of  $\{\mathbf{A}_k\}$ , defined by

$$\bar{\mathbf{A}}_j = \sum_{k=j}^{N(j)} \alpha_{j,k} \mathbf{A}_k, \quad \alpha_{j,k} \in [0, 1], \quad \sum_{k=j}^{N(j)} \alpha_{j,k} = 1,$$

such that  $\bar{\mathbf{A}}_j \rightarrow A$  strongly in  $\mathbf{L}^1(0, 1)$ . Note that we applied Lemma C.10 to the sequence  $\{\mathbf{A}_k\}_{k \geq j}$  at step  $j$ . Since  $\mathcal{M}$  is convex, we obtain

$$\mathcal{M}(\bar{\mathbf{A}}_j) = \mathcal{M}\left(\sum_{k=j}^{N(j)} \alpha_{j,k} \mathbf{A}_k\right) \leq \sum_{k=j}^{N(j)} \alpha_{j,k} \mathcal{M}(\mathbf{A}_k).$$

Letting  $j \rightarrow \infty$ , since  $k \geq j$  in the sum and  $\mathcal{M}(\mathbf{A}_k) \rightarrow \theta$ , we have

$$(6.43) \quad \liminf_{j \rightarrow \infty} \mathcal{M}(\bar{\mathbf{A}}_j) \leq \theta = \liminf_{k \rightarrow \infty} \mathcal{M}(\mathbf{A}_k).$$

In addition, it holds that

$$(6.44) \quad \mathcal{M}(\bar{\mathbf{A}}_j) \rightarrow \mathcal{M}(\mathbf{A}).$$



Indeed, since  $m \in \mathbf{BV}(0, 1; \mathcal{E})$  and  $\bar{\mathbf{A}}_j \rightarrow \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$ ,

$$\begin{aligned} |\mathcal{M}(\mathbf{A} - \bar{\mathbf{A}}_j)| &= \left| \int_0^1 \text{Tr} \left( D^2 V(m(t)) (\mathbf{A}^{-1}(t) - \bar{\mathbf{A}}_j^{-1}(t)) \right) dt \right| \\ &\leq \int_0^1 \left| \text{Tr} \left( D^2 V(m(t)) \mathbf{A}^{-1}(t) (\bar{\mathbf{A}}_j(t) - \mathbf{A}(t)) \bar{\mathbf{A}}_j^{-1}(t) \right) \right| dt \\ &\lesssim \int_0^1 \left| D^2 V(m(t)) \mathbf{A}^{-1}(t) (\bar{\mathbf{A}}_j(t) - \mathbf{A}(t)) \bar{\mathbf{A}}_j^{-1}(t) \right|_F dt \\ &\lesssim \int_0^1 |D^2 V(m(t))|_F |\mathbf{A}^{-1}(t)|_F |\bar{\mathbf{A}}_j(t) - \mathbf{A}(t)|_F |\bar{\mathbf{A}}_j^{-1}(t)|_F dt \\ &\lesssim \|\bar{\mathbf{A}}_j(t) - \mathbf{A}(t)\|_{\mathbf{L}^1(0,1)} \rightarrow 0. \end{aligned}$$

Therefore (6.42) follows from (6.43) and (6.44) and thereby proves (6.39).

Taking account of the fact that  $E_\varepsilon$   $\Gamma$ -converges to  $E$ , we obtain from Proposition 4.4, (6.37), and (6.39) that

$$\begin{aligned} \liminf_{\varepsilon \rightarrow 0} F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon) &= \lim_{k \rightarrow \infty} F_{\varepsilon_k}(m_{\varepsilon_k}, \mathbf{A}_{\varepsilon_k}) \\ &\geq E(m) + \frac{1}{4} \int_0^1 (D^2 V(m(t)) - \mathbf{A}(t))^2 : \mathbf{A}^{-1}(t) dt = F(m, \mathbf{A}). \end{aligned}$$

Next we prove the limsup inequality; i.e., for a subsequence  $\varepsilon_k \rightarrow 0$ , we want to find a pair of recovering sequences  $(m_k, \mathbf{A}_k)$  converging to  $(m, \mathbf{A})$  such that

$$\limsup_{k \rightarrow \infty} F_{\varepsilon_k}(m_k, \mathbf{A}_k) \leq F(m, \mathbf{A}).$$

It suffices to deal with the case where  $F(m, \mathbf{A}) < \infty$  and hence  $m(t) \in \mathcal{E}$ . Otherwise the limsup inequality is automatically satisfied. First, thanks to the  $\Gamma$ -convergence of  $E_\varepsilon$  to  $E$ , one automatically obtains a recovering sequence  $m_k \in \mathbf{H}_\pm^1(0, 1)$  such that  $m_k \rightarrow m$  in  $\mathbf{L}^1(0, 1)$ ,  $\limsup_k \|m_k\|_{\mathbf{L}^\infty(0,1)} < \infty$ , and  $\limsup_{k \rightarrow \infty} E_\varepsilon(m_k) \leq E(m)$ . We construct a recovering sequence  $\mathbf{A}_k \in \mathbf{H}_a^1(0, 1)$  explicitly by using convolution approximation. Specifically fixing any  $\alpha < \gamma/3$ , we define

$$(6.45) \quad \mathbf{A}_k := \tilde{\mathcal{K}}_{\varepsilon_k^\alpha} \mathbf{A},$$

where  $\tilde{\mathcal{K}}_\varepsilon$  is the convolution operator defined in (C.10). It is proved in (C.8) that  $\mathbf{A}_k \in \mathbf{H}_a^1(0, 1)$  and  $\mathbf{A}_k \rightarrow \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$ . Moreover, by replacing  $\varepsilon$  with  $\varepsilon^\alpha$  in the bound proved in Lemma (C.8), we have

$$(6.46) \quad \varepsilon_k^\gamma \|\mathbf{A}_k\|_{\mathbf{H}^1(0,1)}^2 \lesssim \varepsilon_k^{\gamma-3\alpha} \rightarrow 0.$$

With the above choices for  $m_k$  and  $\mathbf{A}_k$ , we get from Proposition 4.4 that

$$\begin{aligned} &\limsup_{k \rightarrow \infty} F_{\varepsilon_k}(m_k, \mathbf{A}_k) \\ &= \limsup_{k \rightarrow \infty} \left\{ E_{\varepsilon_k}(m_k, \mathbf{A}_k) + \frac{1}{4} \int_0^1 (D^2 V(m_k(t)) - \mathbf{A}_k(t))^2 : \mathbf{A}_k^{-1}(t) dt \right. \\ &\quad \left. + \int_0^1 (D^3 V(m_k(t)) \cdot \nabla V(m_k(t))) : \mathbf{A}_k^{-1}(t) dt + \varepsilon_k^\gamma \|\mathbf{A}_k\|_{\mathbf{H}^1(0,1)}^2 \right\} \\ &\leq E(m) + \frac{1}{4} \int_0^1 (D^2 V(m_k(t)) - \mathbf{A}_k(t))^2 : \mathbf{A}_k^{-1}(t) dt \\ &= F(m, \mathbf{A}). \end{aligned}$$

To pass to the inequality we have used the dominated convergence theorem and (6.37) for the third term on the left-hand side, as well as (6.46) for the fourth term on the left-hand side. The proof is now complete.  $\square$

*Proof of Corollary 4.7.* Let  $(m_\varepsilon, \mathbf{A}_\varepsilon) \in \mathcal{H}$  be a minimizer of  $F_\varepsilon$ . We first argue that  $\limsup_\varepsilon F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon) < \infty$ . In fact, for any fixed  $m \in \mathbf{BV}(0, 1; \mathcal{E})$  and  $\mathbf{A} \in \mathbf{L}_a^1(0, 1)$ , we know from the proof of the limsup inequality of Theorem 4.6 that there exists a recovering sequence  $(\tilde{m}_\varepsilon, \tilde{\mathbf{A}}_\varepsilon)$  such that  $\limsup_\varepsilon F_\varepsilon(\tilde{m}_\varepsilon, \tilde{\mathbf{A}}_\varepsilon) < \infty$ . Since  $(m_\varepsilon, \mathbf{A}_\varepsilon)$  minimizes  $F_\varepsilon$ , we have

$$\limsup_{\varepsilon \rightarrow 0} F_\varepsilon(m_\varepsilon, \mathbf{A}_\varepsilon) \leq \limsup_{\varepsilon \rightarrow 0} F_\varepsilon(\tilde{m}_\varepsilon, \tilde{\mathbf{A}}_\varepsilon) < \infty.$$

Then by Proposition (4.5), there exist a subsequence  $\varepsilon_k$  and the corresponding  $\{(m_k, \mathbf{A}_k)\} \subset \mathcal{H}$  such that  $m_k \rightarrow m$  in  $\mathbf{L}^1(0, 1)$  and  $\mathbf{A}_k \rightarrow \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$  with some  $m \in \mathbf{BV}(0, 1; \mathcal{E})$  and  $\mathbf{A} \in \mathbf{L}_a^1(0, 1)$ . We now show that  $(m, \mathbf{A})$  minimizes  $F$ . In fact, given any  $\tilde{m} \in \mathbf{BV}(0, 1; \mathcal{E})$  and  $\tilde{\mathbf{A}} \in \mathbf{L}_a^1(0, 1)$ , thanks to the  $\Gamma$ -convergence of  $F_\varepsilon$  to  $F$ , one can find a recovering sequence  $(\tilde{m}_k, \tilde{\mathbf{A}}_k) \in \mathcal{H}$  such that

$$\limsup_k F_{\varepsilon_k}(\tilde{m}_k, \tilde{\mathbf{A}}_k) \leq F(\tilde{m}, \tilde{\mathbf{A}}).$$

Since  $(m_k, \mathbf{A}_k)$  minimizes  $F_{\varepsilon_k}$ , we have  $F_{\varepsilon_k}(m_k, \mathbf{A}_k) \leq F_{\varepsilon_k}(\tilde{m}_k, \tilde{\mathbf{A}}_k)$ . Then using the liminf inequality part of the  $\Gamma$ -convergence of  $F_\varepsilon$  to  $F$ , we obtain

$$F(m, \mathbf{A}) \leq \liminf_{k \rightarrow \infty} F_{\varepsilon_k}(m_k, \mathbf{A}_k) \leq \limsup_{k \rightarrow \infty} F_{\varepsilon_k}(\tilde{m}_k, \tilde{\mathbf{A}}_k) \leq F(\tilde{m}, \tilde{\mathbf{A}}).$$

Since  $\tilde{m}, \tilde{\mathbf{A}}$  is arbitrary,  $(m, \mathbf{A})$  is a minimizer of  $F$ .  $\square$

#### Appendix A. Estimates for the constant coefficient Green's functions.

Assume that function  $|\lambda(t)| \geq a$  almost everywhere in  $[0, 1]$  with some  $a > 0$ . For any  $s \in (0, 1)$ , let  $\overline{G}_\varepsilon^\lambda$  be the solution to the equation

$$(A.1) \quad \begin{aligned} (-\partial_t^2 + \varepsilon^{-2} \lambda^2(s)) \overline{G}_\varepsilon^\lambda(t, s) &= \delta(t - s), \quad t \in (0, 1), \\ \overline{G}_\varepsilon^\lambda(0, s) &= \overline{G}_\varepsilon^\lambda(1, s) = 0, \end{aligned}$$

where  $\delta$  is the Dirac function. The solution  $\overline{G}_\varepsilon^\lambda$  is given explicitly as follows:

$$\overline{G}_\varepsilon^\lambda(t, s) = \frac{\varepsilon}{|\lambda(s)| \sinh(|\lambda(s)|/\varepsilon)} \times \begin{cases} \sinh(|\lambda(s)|s/\varepsilon) \sinh(|\lambda(s)|(1-t)/\varepsilon), & s \leq t, \\ \sinh(|\lambda(s)|t/\varepsilon) \sinh(|\lambda(s)|(1-s)/\varepsilon), & s \geq t. \end{cases}$$

Notice that  $\overline{G}_\varepsilon^\lambda(t, s)$  is not a standard Green's function, as it is not symmetric with respect to permutation of its arguments. According to the definition of  $\sinh$ , a few elementary calculations yield the following estimates.

LEMMA A.1. *Let  $|\lambda(t)| \geq a$  a.e. on  $(0, 1)$  for a fixed  $a > 0$ . Then for sufficiently small  $\varepsilon > 0$  the solution  $\overline{G}_\varepsilon^\lambda$  to (A.1) satisfies the following:*

(i) *There exists  $C = C(a) > 0$  such that*

$$(A.2) \quad 0 \leq \overline{G}_\varepsilon^\lambda(t, s) \leq C\varepsilon e^{-\frac{a}{\varepsilon}|s-t|}$$

*for any  $t, s \in [0, 1]$ .*

(ii) There exists  $C = C(a) > 0$  such that

$$\overline{G}_\varepsilon^\lambda(t, t) = \frac{\varepsilon}{2} \left( \frac{1}{|\lambda(t)|} + R(t) \right), \quad t \in [0, 1],$$

with

$$(A.3) \quad |R(t)| \leq C \left( e^{-\frac{2at}{\varepsilon}} + e^{-\frac{2a(1-t)}{\varepsilon}} \right).$$

Considering the Green's tensor  $\overline{\mathbf{G}}_\varepsilon(t, s)$  that solves the matrix equation

$$(A.4) \quad \begin{aligned} (-\partial_t^2 + \varepsilon^{-2} \mathbf{A}(s)) \overline{\mathbf{G}}_\varepsilon(\cdot, s) &= \delta(\cdot - s) \cdot \mathbf{I}_d, \\ \overline{\mathbf{G}}_\varepsilon(0, s) &= \overline{\mathbf{G}}_\varepsilon(1, s) = 0, \end{aligned}$$

with  $\mathbf{A} \in \mathbf{H}^1(0, 1)$  and  $|\mathbf{A}| \geq a$  a.e. on  $(0, 1)$ , we have the following similar estimates.

LEMMA A.2. Let  $\mathbf{A} \in \mathbf{H}^1(0, 1)$  and  $|\mathbf{A}| \geq a$  a.e. on  $(0, 1)$ . For sufficiently small  $\varepsilon > 0$  the solution  $\overline{\mathbf{G}}_\varepsilon$  to (A.4) satisfies the following:

(i) There exists  $C = C(a) > 0$  such that

$$(A.5) \quad |\overline{\mathbf{G}}_\varepsilon(t, s)| \leq C\varepsilon e^{-\frac{a}{\varepsilon}|s-t|}$$

for any  $t, s \in [0, 1]$ .

(ii) There exists  $C = C(a) > 0$  such that

$$(A.6) \quad \overline{\mathbf{G}}_\varepsilon(t, t) = \frac{\varepsilon}{2} (|\mathbf{A}^{-1}(t)| + \mathbf{R}(t))$$

with

$$(A.7) \quad |\mathbf{R}(t)| \leq C \left( e^{-\frac{2at}{\varepsilon}} + e^{-\frac{2a(1-t)}{\varepsilon}} \right).$$

*Proof.* Since  $\mathbf{A}(s)$  is symmetric for any  $s \in (0, 1)$  (by the definition of  $\mathbf{H}^1(0, 1)$ ), there exists an orthogonal matrix  $\mathbf{P}(s)$  such that  $\mathbf{A}(s) = \mathbf{P}^{-1}(s)\mathbf{\Lambda}(s)\mathbf{P}(s)$ , where  $\mathbf{\Lambda}(s) = \text{diag}(\lambda_1(s), \dots, \lambda_d(s))$ . Moreover, by assumption we have  $|\lambda_i(s)| \geq a$  a.e. on  $(0, 1)$  for any  $i = 1, \dots, d$ . Therefore, the problem (A.4) can be diagonalized so that one obtains

$$(A.8) \quad \overline{\mathbf{G}}_\varepsilon(t, s) = \mathbf{P}^{-1}(s) \cdot \text{diag}(\overline{G}_\varepsilon^{\lambda_1}(t, s), \dots, \overline{G}_\varepsilon^{\lambda_d}(t, s)) \cdot \mathbf{P}(s),$$

where  $\overline{G}_\varepsilon^{\lambda_i}(\cdot, s)$  solves (A.1) with  $\lambda$  replaced by  $\lambda_i$ . Then (A.5) follows directly from (A.8) and (A.2), and (A.6) can be deduced from (A.8) and (A.3).  $\square$

**Appendix B. Fundamental matrix of linear systems.** Given  $f : \mathbf{R} \rightarrow \mathbf{R}^d$  and  $\mathbf{A} : \mathbf{R} \rightarrow \mathbf{R}^{d \times d}$ , consider the following linear differential equation:

$$(B.1) \quad dx_\varepsilon(t) = -\varepsilon^{-1} \mathbf{A}(t)x_\varepsilon(t)dt + f(t)dt, \quad x_\varepsilon(t_0) = 0.$$

The solution to (B.1) can be found via the variation of constants method, provided that its fundamental matrix is determined.

DEFINITION B.1 (fundamental matrix). The fundamental matrix  $\mathbf{M}_\varepsilon(t, t_0)$  is the solution matrix that solves the problem

$$(B.2) \quad \frac{d}{dt} \mathbf{M}_\varepsilon(t, t_0) = -\varepsilon^{-1} \mathbf{A}(t) \mathbf{M}_\varepsilon(t, t_0), \quad \mathbf{M}_\varepsilon(t_0, t_0) = \mathbf{I}_d.$$

Suppose that  $\mathbf{A}$  and  $f$  are both continuous; then the solution to the ODE (B.1) can be written in the form

$$x_\varepsilon(t) = \int_{t_0}^t \mathbf{M}_\varepsilon(t, s) f(s) ds.$$

We note that the above formula is still valid when  $f(s)ds$  is replaced by  $dW(s)$ , in which case the integral is understood as Itô's stochastic integration. In the case that  $d = 1$  or if  $\mathbf{A}$  does not depend on  $t$ , we have  $\mathbf{M}_\varepsilon(t, s) = \exp(-\varepsilon^{-1} \int_s^t \mathbf{A}(r) dr)$ . In general, there is no closed-form expression for the fundamental matrix  $\mathbf{M}_\varepsilon$ , and hence the solution to (B.1) has no explicit formula. Nevertheless,  $\mathbf{M}_\varepsilon$  has some nice properties which are useful in studying the asymptotic behavior of the solution to (B.1) when  $\varepsilon \rightarrow 0$ .

LEMMA B.2. *Let  $\mathbf{M}_\varepsilon$  be the fundamental matrix defined by (B.2). Then we have the following:*

- (i) *For all  $t, t_0, t_1 \in \mathbf{R}$ ,  $\mathbf{M}_\varepsilon(t, t_0) = \mathbf{M}_\varepsilon(t, t_1) \mathbf{M}_\varepsilon(t_1, t_0)$ .*
- (ii) *For all  $t, t_0 \in \mathbf{R}$ ,  $\mathbf{M}_\varepsilon(t, t_0)$  is nonsingular and  $\mathbf{M}_\varepsilon^{-1}(t, t_0) = \mathbf{M}_\varepsilon(t_0, t)$ .*
- (iii) *For all  $t, t_0 \in \mathbf{R}$ ,*

$$(B.3) \quad \det(\mathbf{M}_\varepsilon(t, t_0)) = \exp\left(-\varepsilon^{-1} \int_{t_0}^t \text{Tr}(\mathbf{A}(s)) ds\right).$$

*Proof.* The proof can be found in [5, Chapter 6].  $\square$

We finish this appendix with two useful inequalities about the determinants of symmetric positive definite matrices.

LEMMA B.3. *If  $\mathbf{A}, \mathbf{B}$  are real symmetric positive definite matrices of size  $d$ , then*

$$(B.4) \quad (\det(\mathbf{A} + \mathbf{B}))^{\frac{1}{d}} \geq (\det(\mathbf{A}))^{\frac{1}{d}} + (\det(\mathbf{B}))^{\frac{1}{d}}.$$

A proof of this lemma can be found in [23, p. 115]. It shows that the function  $\mathbf{A} \mapsto \det(\mathbf{A})^{\frac{1}{d}}$  is concave. As a consequence, we have the following corollary.

COROLLARY B.4. *Suppose that  $\mathbf{A} \in C([0, 1]; \mathbf{R}^{d \times d})$  is a matrix-valued function and that  $\mathbf{A}(t)$  is symmetric positive definite for any  $t \in [0, 1]$ . Then we have*

$$(B.5) \quad \left[ \det \left( \int_0^1 \mathbf{A}(t) dt \right) \right]^{\frac{1}{d}} \geq \int_0^1 \det(\mathbf{A}(t))^{\frac{1}{d}} dt.$$

*Proof.* Define  $t_i = i/N$  with  $i = 0, 1, \dots, N$ . Since  $\mathbf{A}$  is continuous on  $[0, 1]$ , it holds that

$$\begin{aligned} \left[ \det \left( \int_0^1 \mathbf{A}(t) dt \right) \right]^{\frac{1}{d}} &= \lim_{N \rightarrow \infty} \left[ \det \left( \frac{1}{N} \sum_{i=1}^N \mathbf{A}(t_i) \right) \right]^{\frac{1}{d}} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \left[ \det \left( \sum_{i=1}^N \mathbf{A}(t_i) \right) \right]^{\frac{1}{d}} \\ &\geq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \det(\mathbf{A}(t_i))^{\frac{1}{d}} = \int_0^1 \det(\mathbf{A}(t))^{\frac{1}{d}} dt. \end{aligned}$$

We have used (B.4) in the inequality.  $\square$

**Appendix C. Useful inequalities and lemmas.**

LEMMA C.1 (interpolation inequality). *Let  $u \in H_0^1(0, 1)$ . Then*

$$(C.1) \quad \|u\|_{H^s} \lesssim \|u\|_{H^1}^s \|u\|_{L^2}^{1-s}$$

for any  $s \in [0, 1]$ .

*Proof.* See [17, Corollary 6.11] for the proof.  $\square$

LEMMA C.2. *Let  $u \in H_0^1(0, 1)$ . Then  $\|u\|_{L^\infty}^2 \leq 2\|u\|_{L^2}|u|_{H^1}$ .*

*Proof.* It suffices to prove the inequality when  $u \in C_0^\infty(0, 1)$ . For any  $t \in (0, 1)$ , it follows from the fundamental theorem of calculus and the Cauchy–Schwarz inequality that

$$(C.2) \quad u^2(t) = 2 \int_0^t u(s)u'(s)ds \leq 2\|u\|_{L^2}|u|_{H^1}.$$

The lemma then follows by taking the supremum over  $t$ .  $\square$

LEMMA C.3. *Let  $\alpha, \beta \in (0, 1)$  be such that  $\beta > \max(\alpha, \alpha/2 + 1/4)$ . Then any matrix-valued function  $B \in \mathbf{H}^{-\alpha}(0, 1)$  can be viewed as a bounded multiplication operator from  $\mathbf{H}_0^\beta$  to  $\mathbf{H}^{-\beta}$ . Furthermore we have*

$$(C.3) \quad \|B\|_{\mathcal{L}(\mathbf{H}_0^\beta, \mathbf{H}^{-\beta})} \lesssim \|B\|_{\mathbf{H}^{-\alpha}(0, 1)}.$$

*Proof.* It suffices to consider the proof in the scalar case. Let  $B \in H^{-\alpha}$  and  $\varphi \in H_0^\beta$ . Assume that  $\beta > \alpha$ ; then one can define the multiplication  $B\varphi$  as a distribution in the sense that for any  $\psi \in C_0^\infty(0, 1)$

$$\langle B\varphi, \psi \rangle = \langle B, \varphi\psi \rangle.$$

Moreover, if  $\beta - \alpha/2 > 1/4$ , we have

$$|\langle B\varphi, \psi \rangle| = |\langle B, \varphi\psi \rangle| \leq \|B\|_{H^{-\alpha}} \|\varphi\psi\|_{H^\alpha} \lesssim \|B\|_{H^{-\alpha}} \|\varphi\|_{H^\beta} \|\psi\|_{H^\beta},$$

where the last estimate follows from the following lemma. Therefore the desired estimate (C.3) holds.  $\square$

LEMMA C.4. *Let  $\alpha, \beta$ , and  $\gamma$  be positive exponents such that  $\min(\alpha, \beta) > \gamma$  and  $\alpha + \beta > \gamma + 1/2$ . Then, if  $\varphi \in H^\alpha$  and  $\psi \in H^\beta$ , the product  $\varphi\psi$  belongs to  $H^\gamma$  and  $\|\varphi\psi\|_{H^\gamma} \lesssim \|\varphi\|_{H^\alpha} \|\psi\|_{H^\beta}$ .*

*Proof.* The proof can be found in [17, Theorem 6.18].  $\square$

LEMMA C.5. *Let  $A \in H^1(0, 1)$ , and let  $f \in L^2(0, 1)$ . Set  $B = A^2 - A'$ . Then there exists a unique solution  $u \in H_0^1(0, 1)$  solving the problem*

$$(C.4) \quad \begin{aligned} (-\partial_t^2 + B)u &= f \quad \text{on } (0, 1), \\ u(0) &= u(1) = 0. \end{aligned}$$

Moreover, it holds that

$$(C.5) \quad \|u\|_{H^2(0, 1)} \leq C\|f\|_{L^2(0, 1)},$$

and  $(-\partial_t^2 + B)^{-1}$  is a trace-class operator on  $L^2(0, 1)$ .

*Proof.* Let  $G_0(s, t)$  be the Dirichlet Green's function of  $-\partial_t^2$  on  $(0, 1)$ . In fact,  $G_0(s, t) = s(1-t) \wedge t(1-s)$  for all  $s, t \in [0, 1]$ . From Green's first identity, it is easy to observe that a solution  $u \in H_0^1(0, 1)$  solving (C.4) is a solution  $u \in L^2(0, 1)$  that solves the Lippmann–Schwinger integral equation

$$(C.6) \quad u(t) + \int_0^1 G_0(t, s)B(s)u(s)ds = \int_0^1 G_0(t, s)f(s)ds$$

and vice versa. Now we apply the Fredholm alternative theorem to prove the existence and uniqueness of the solution to (C.6). First the operator

$$(\mathcal{T}u)(t) := (-\partial_t^2)^{-1}(Bu)(t) = \int_0^1 G_0(t, s)B(s)u(s)ds$$

is compact from  $L^2(0, 1)$  to itself. There are several ways to prove this, but the simplest argument is perhaps the observation that  $\mathcal{T}$  is bounded from  $L^2(0, 1)$  to  $W^{1,\infty}(0, 1)$ . Indeed, since  $G_0(t, \cdot) \in W^{1,\infty}(0, 1)$  for any  $t \in [0, 1]$ , we can apply the Cauchy–Schwarz inequality twice to get

$$(C.7) \quad \begin{aligned} \|\mathcal{T}u\|_{W^{1,\infty}(0,1)} &\leq \sup_{t \in [0,1]} \|G_0(t, \cdot)\|_{W^{1,\infty}(0,1)} \|Bu\|_{L^1(0,1)} \\ &\leq \sup_{t \in [0,1]} \|G_0(t, \cdot)\|_{W^{1,\infty}(0,1)} \|B\|_{L^2(0,1)} \|u\|_{L^2(0,1)}. \end{aligned}$$

Then the compactness of  $\mathcal{T}$  follows from the compact embedding  $W^{1,\infty}(0, 1) \hookrightarrow L^2(0, 1)$ . We are left to show the uniqueness of (C.6) or equivalently (C.4). To see this, setting  $f = 0$ , we multiply (C.4) by  $u$ , integrate, and use Green's first identity to get

$$\begin{aligned} 0 &= \int_0^1 u'(t)^2 dt + \int_0^1 (A^2(t) - A'(t))u^2(t) dt \\ &= \int_0^1 u'(t)^2 dt + \int_0^1 A^2(t)u^2(t) dt + 2 \int_0^1 A(t)u(t)u'(t) dt \\ &= \int_0^1 (u'(t) + A(t)u(t))^2 dt. \end{aligned}$$

Therefore we should have  $u'(t) = -A(t)u(t)$ . The only solution to this equation with the Dirichlet boundary conditions is zero. Hence by the Fredholm alternative theorem, the integral equation (C.6) has a unique solution in  $L^2(0, 1)$ . Then the estimate (C.5) follows from (C.7), (C.6), and the estimate that  $\|\mathcal{R}f\|_{H^2(0,1)} \leq C\|f\|_{L^2(0,1)}$ , where

$$(\mathcal{R}f)(t) = \int_0^1 G_0(t, s)f(s)ds.$$

Finally, observe that

$$(-\partial_t^2 + B)^{-1} = (I + (-\partial_t^2)^{-1}B)^{-1}(-\partial_t^2)^{-1} = (I + \mathcal{T})^{-1}(-\partial_t^2)^{-1}.$$

Then it follows from the fact that  $(-\partial_t^2)^{-1}$  is a trace-class operator on  $L^2(0, 1)$  and the boundedness of  $(I + \mathcal{T})^{-1}$  that  $(-\partial_t^2 + B)^{-1}$  is trace-class.  $\square$

*Remark C.6.* Lemma C.5 can be easily extended to the matrix-valued case. More precisely, assume that  $\mathbf{A} \in \mathbf{H}^1(0, 1)$  and that  $\mathbf{A}(t)$  is a symmetric matrix for any  $t \in [0, 1]$ . Let  $\mathbf{B} = \mathbf{A}^2 - \mathbf{A}'$ . Then the inverse of the matrix-valued Schrödinger operator  $(-\partial_t^2 + \mathbf{B})^{-1}$  is bounded from  $\mathbf{L}^2(0, 1)$  to  $\mathbf{H}^2(0, 1)$  and is a trace-class operator on  $\mathbf{L}^2(0, 1)$ .

The next lemma discusses some properties of approximation by convolution.

LEMMA C.7. Let  $K \in C_0^\infty(\mathbf{R})$  such that  $K \geq 0$  and  $\int_{\mathbf{R}} K = 1$ . Define  $K_\varepsilon(\cdot) = \varepsilon^{-1}K(\cdot/\varepsilon)$ . Suppose that  $f \in L^1(\mathbf{R})$ , and define  $\mathcal{K}_\varepsilon f = K_\varepsilon * f$ . Then  $\mathcal{K}_\varepsilon f \in L^1(\mathbf{R}) \cap C^\infty(\mathbf{R})$ . Moreover, we have

$$(C.8) \quad \mathcal{K}_\varepsilon f \rightarrow f \quad \text{in } L^1(\mathbf{R})$$

and

$$(C.9) \quad \|\mathcal{K}_\varepsilon f\|_{H^1(\mathbf{R})} \leq C\varepsilon^{-\frac{3}{2}}\|f\|_{L^1(\mathbf{R})}.$$

*Proof.* The property (C.8), often termed the approximation of identity in  $L^1(\mathbf{R})$ , has been proved in many books; see, e.g., [7]. We now show that  $\mathcal{K}_\varepsilon f \in H^1(\mathbf{R})$  and that (C.9) is valid. This can be seen from the observation that

$$\begin{aligned} \|\mathcal{K}_\varepsilon f\|_{H^1(\mathbf{R})}^2 &= \|\mathcal{K}_\varepsilon f\|_{L^2(\mathbf{R})}^2 + \|\mathcal{K}_\varepsilon' f\|_{L^2(\mathbf{R})}^2 \\ &= \|\varepsilon^{-1}K(\cdot/\varepsilon) * f\|_{L^2(\mathbf{R})}^2 + \varepsilon^{-2}\|\varepsilon^{-1}K'(\cdot/\varepsilon) * f\|_{L^2(\mathbf{R})}^2 \\ &\leq \varepsilon^{-1}\|K\|_{L^2(\mathbf{R})}^2\|f\|_{L^1(\mathbf{R})}^2 + \varepsilon^{-3}\|K'\|_{L^2(\mathbf{R})}^2\|f\|_{L^1(\mathbf{R})}^2 \leq C\varepsilon^{-3}\|f\|_{L^1(\mathbf{R})}^2. \end{aligned}$$

Note that we have used Young's inequalities in the penultimate inequality.  $\square$

We continue to adapt Lemma C.7 to matrix functions defined on a bounded domain. For this purpose, we define two useful operators. Given a function  $f \in L^1(0, 1)$ , we define its extension

$$\mathcal{E}f(x) := \begin{cases} f(x) & \text{if } x \in (0, 1), \\ 0 & \text{otherwise.} \end{cases}$$

Conversely, for a function  $g \in L^1(\mathbf{R})$ , we define the restriction  $\mathcal{R}f := f|_{(0,1)}$ . Likewise, we can define the convolution, extension, or restriction of a matrix function through entrywise operations. The following lemma concerns the convolution approximation of matrix-valued functions.

LEMMA C.8. Let  $\mathbf{A} \in \mathbf{L}_a^1(0, 1)$ . Define

$$(C.10) \quad \tilde{\mathcal{K}}_\varepsilon \mathbf{A} := \mathcal{R}(\mathcal{K}_\varepsilon(\mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d))) + a \cdot \mathbf{I}_d.$$

Then  $\tilde{\mathcal{K}}_\varepsilon \mathbf{A} \in \mathbf{H}_a^1(0, 1)$ . Moreover,  $\tilde{\mathcal{K}}_\varepsilon \mathbf{A} \rightarrow \mathbf{A}$  in  $\mathbf{L}^1(0, 1)$  and  $\|\tilde{\mathcal{K}}_\varepsilon \mathbf{A}\|_{\mathbf{H}^1(0,1)} \leq C\varepsilon^{-\frac{3}{2}}$  with the constant  $C$  depending on  $\mathbf{A}$  and  $a$ .

*Proof.* First it follows from Lemma C.7 that  $\tilde{\mathcal{K}}_\varepsilon \mathbf{A} \in \mathbf{H}^1(0, 1)$ . To show  $\tilde{\mathcal{K}}_\varepsilon \mathbf{A} \in \mathbf{H}_a^1(0, 1)$ , it suffices to show that  $\mathcal{K}_\varepsilon(\mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d))$  is positive semidefinite. Indeed, for any fixed  $x \in \mathbf{R}^d$ ,

$$\begin{aligned} x^T \mathcal{K}_\varepsilon(\mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d))x &= \mathcal{K}_\varepsilon(x^T(\mathbf{A} - a \cdot \mathbf{I}_d)x) \\ &= K_\varepsilon(\cdot) * \mathcal{E}(x^T(\mathbf{A}(\cdot) - a \cdot \mathbf{I}_d)x) \geq 0, \end{aligned}$$

where we have used the assumption that  $\mathbf{A}(\cdot) - a \cdot \mathbf{I}_d$  is positive semidefinite a.e. on  $(0, 1)$ . Next from Lemma C.7 and the fact that  $\mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d) \in \mathbf{L}^1(\mathbf{R})$ , we have

$$\begin{aligned} \|\tilde{\mathcal{K}}_\varepsilon \mathbf{A} - \mathbf{A}\|_{\mathbf{L}^1(0,1)} &= \|\mathcal{R}(\mathcal{K}_\varepsilon(\mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d)) - \mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d))\|_{\mathbf{L}^1(0,1)} \\ &\leq \|\mathcal{K}_\varepsilon(\mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d)) - \mathcal{E}(\mathbf{A} - a \cdot \mathbf{I}_d)\|_{\mathbf{L}^1(\mathbf{R})} \rightarrow 0. \end{aligned}$$

By similar arguments one can show that  $\|\tilde{\mathcal{K}}_\varepsilon \mathbf{A}\|_{\mathbf{H}^1(0,1)} \leq C\varepsilon^{-\frac{3}{2}}$ .  $\square$

The next lemma characterizes explicitly the minimizer of the second component (as a functional of  $\mathbf{A}$ ) of the functional  $F$  defined in (4.8). Recall the notation  $|\mathbf{A}|$  of a matrix  $\mathbf{A}$  defined in section 2.2.

LEMMA C.9. *Let  $\mathbf{B}$  be a fixed symmetric matrix. Let  $\mathbf{A}$  be a minimizer of the functional*

$$\mathcal{G}(\mathbf{A}) := (\mathbf{B} - \mathbf{A})^2 : \mathbf{A}^{-1}$$

*over all positive matrices. Then it holds that  $\mathbf{A} = |\mathbf{B}|$ . With this choice of  $\mathbf{A}$ , we have  $\mathcal{G}(\mathbf{A}) = 2(\text{Tr}(|\mathbf{B}|) - \text{Tr}(\mathbf{B}))$ .*

*Proof.* The lemma follows from the fact that the functional  $\mathcal{G}$  can be rewritten as

$$\begin{aligned} \mathcal{G}(\mathbf{A}) &= \text{Tr}(\mathbf{B}^2 \mathbf{A}^{-1}) + \text{Tr}(\mathbf{A}) - 2\text{Tr}(\mathbf{B}) \\ &= \text{Tr}(\mathbf{B}^2 \mathbf{A}^{-1}) + \text{Tr}(\mathbf{A}) - 2\text{Tr}(|\mathbf{B}|) + 2\text{Tr}(|\mathbf{B}|) - 2\text{Tr}(\mathbf{B}) \\ &= \text{Tr}((|\mathbf{B}| - \mathbf{A})^2 \mathbf{A}^{-1}) + 2\text{Tr}(|\mathbf{B}|) - 2\text{Tr}(\mathbf{B}) \\ &= (|\mathbf{B}| - \mathbf{A})^2 : \mathbf{A}^{-1} + 2\text{Tr}(|\mathbf{B}|) - 2\text{Tr}(\mathbf{B}). \end{aligned} \quad \square$$

Finally, the following lemma from Mazur is useful to obtain a strong convergent subsequence from a weakly convergent sequence. The proof can be found in [6, Corollary 3.8].

LEMMA C.10 (Mazur's lemma). *Let  $X$  be a Banach space, and let  $\{u_n\}_{n \in \mathbb{N}}$  be a sequence in  $X$  that converges weakly to  $u \in X$ . Then there exists a sequence  $\{\bar{u}_j\}_{j \in \mathbb{N}}$  defined by the convex combination of  $\{u_n\}_{n \in \mathbb{N}}$ , namely,*

$$(C.11) \quad \bar{u}_j = \sum_{n=1}^{N(j)} \alpha_{j,n} u_n, \quad \alpha_{j,n} \in [0, 1], \quad \sum_{n=1}^{N(j)} \alpha_{j,n} = 1,$$

*such that  $\bar{u}_j$  converges to  $u$  strongly in  $X$ .*

**Appendix D. Proof of Proposition 4.3.** We first show the liminf inequality. Supposing that  $\{m_\varepsilon\} \subset \mathbf{H}_\pm^1(0, 1)$  and that  $m_\varepsilon \rightarrow m$  in  $\mathbf{L}^1(0, 1)$ , we want to prove that  $E(m) \leq \liminf_{\varepsilon \rightarrow 0} E_\varepsilon(m_\varepsilon)$ . We may assume that  $\liminf_{\varepsilon \rightarrow 0} E_\varepsilon(m_\varepsilon) < \infty$ , since otherwise there is nothing to prove. Let  $\{\varepsilon_n\}$  and  $\{m_n\} \subset \mathbf{H}_\pm^1(0, 1)$  be subsequences such that  $\varepsilon_n \rightarrow 0$  and

$$\lim_{n \rightarrow \infty} E_{\varepsilon_n}(m_n) = \liminf_{\varepsilon \rightarrow 0} E_\varepsilon(m_\varepsilon) < \infty.$$

By Lemma 4.2,  $m \in \mathbf{BV}(0, 1; \mathcal{E})$ , and one can extract a further subsequence (without relabeling) such that  $m_n(t) \rightarrow m(t)$  a.e.  $t \in (0, 1)$ . It is sufficient to deal with the case where  $m$  only has a single jump at  $\tau \in (0, 1)$ , i.e.,

$$(D.1) \quad m(t) = \begin{cases} m(\tau^-) & \text{if } t \in (0, \tau), \\ m(\tau^+) & \text{if } t \in [\tau, 1). \end{cases}$$

Let  $0 < t_1 < \tau < t_2 < 1$  and  $m_n(t_1) \rightarrow m(t_1) = m(\tau^-)$ ,  $m_n(t_2) \rightarrow m(t_2) = m(\tau^+)$ .



Define  $\tilde{m}_n = m_n(\varepsilon_n^{-1}(t - (t_1 + t_2)/2))$ . Then it follows from the equality (4.5) that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{4} \int_{t_1}^{t_2} \varepsilon_n |m'_n(t)|^2 + \frac{1}{\varepsilon_n} |\nabla V(m_n(t))|^2 dt \\ &= \lim_{n \rightarrow \infty} \frac{1}{4} \int_{\frac{t_1 - t_2}{2\varepsilon_n}}^{\frac{t_2 - t_1}{2\varepsilon_n}} |\tilde{m}'_n(t)|^2 + |\nabla V(\tilde{m}_n(t))|^2 dt \\ &\geq \inf_{T, m} \left\{ \frac{1}{4} \int_{-T}^T |m'(t)|^2 + |\nabla V(m(t))|^2 dt : T > 0, m \in \mathbf{H}^1(-T, T) \text{ and} \right. \\ &\quad \left. m(-T) = x_-, m(T) = x_+ \right\} = \Phi(m(\tau^-), m(\tau^+)). \end{aligned}$$

Similarly, taking into account that  $m_n$  satisfies the end-point conditions, one can obtain

$$\lim_{n \rightarrow \infty} \frac{1}{4} \int_0^{t_1} \varepsilon_n |m'_n(t)|^2 + \frac{1}{\varepsilon_n} |\nabla V(m_n(t))|^2 dt \geq \Phi(x_-, m(0^+))$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{4} \int_{t_2}^1 \varepsilon_n |m'_n(t)|^2 + \frac{1}{\varepsilon_n} |\nabla V(m_n(t))|^2 dt \geq \Phi(m(1^-), x_+).$$

Therefore the liminf inequality  $E(m) \leq \liminf_{\varepsilon \rightarrow 0} E_\varepsilon(m_\varepsilon)$  follows.

Now we prove the limsup inequality, and again it suffices to consider  $m$  defined by (D.1). According to (4.5), for any small  $\eta > 0$  there exist  $T > 0$  and  $m_i \in \mathbf{H}^1(-T, T)$ ,  $i = 1, 2, 3$ , such that

$$\begin{aligned} m_1(-T) &= x_-, \quad m_1(T) = m(0^+) \quad \text{and} \quad \mathcal{J}_T(m_1) \leq \Phi(x_-, m(0^+)) + \frac{\eta}{3}, \\ m_2(-T) &= m(\tau^-), \quad m_2(T) = m(\tau^+) \quad \text{and} \quad \mathcal{J}_T(m_2) \leq \Phi(m(\tau^-), m(\tau^+)) + \frac{\eta}{3}, \\ m_3(-T) &= m(1^-), \quad m_3(T) = x_+ \quad \text{and} \quad \mathcal{J}_T(m_3) \leq \Phi(m(1^-), x_+) + \frac{\eta}{3}. \end{aligned}$$

Then for  $\varepsilon > 0$  small enough, we define the recovery sequence

$$m_\varepsilon(t) = \begin{cases} m_1(-T + \varepsilon^{-1}t) & \text{if } t \in (0, 2\varepsilon T), \\ m(0^+) & \text{if } t \in (2\varepsilon T, \tau - \varepsilon T), \\ m_2(\varepsilon^{-1}(t - \tau)) & \text{if } t \in (\tau - \varepsilon T, \tau + \varepsilon T), \\ m(1^-) & \text{if } t \in (\tau + \varepsilon T, 1 - 2\varepsilon T), \\ m_3(\varepsilon^{-1}(t - 1) + T) & \text{if } t \in (1 - 2\varepsilon T, 1). \end{cases}$$

It is clear that  $m_\varepsilon \in \mathbf{H}^1_\pm(0, 1)$  and  $m_\varepsilon \rightarrow m$  in  $\mathbf{L}^1(0, 1)$  as  $\varepsilon \rightarrow 0$ . Furthermore,

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} E_\varepsilon(m_\varepsilon) \\ &= \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{1}{4} \int_0^1 \varepsilon |m'_\varepsilon(t)|^2 + \frac{1}{\varepsilon} |\nabla V(m_\varepsilon(t))|^2 dt \right\} \\ &= \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{1}{4} \int_0^{2\varepsilon T} \varepsilon |m'_\varepsilon(t)|^2 + \frac{1}{\varepsilon} |\nabla V(m_\varepsilon(t))|^2 dt + \frac{1}{4} \int_{\tau - \varepsilon T}^{\tau + \varepsilon T} \varepsilon |m'_\varepsilon(t)|^2 \right. \\ &\quad \left. + \frac{1}{\varepsilon} |\nabla V(m_\varepsilon(t))|^2 dt + \frac{1}{4} \int_{1 - 2\varepsilon T}^1 \varepsilon |m'_\varepsilon(t)|^2 + \frac{1}{\varepsilon} |\nabla V(m_\varepsilon(t))|^2 dt \right\} \\ &\leq \Phi(x_-, m(0^+)) + \Phi(m(\tau^-), m(\tau^+)) + \Phi(m(1^-), x_+) + \eta. \end{aligned}$$

Since  $\eta$  is arbitrary, the limsup inequality follows.  $\square$

**Acknowledgments.** The authors are grateful to Frank Pinski for helpful discussions and insights.

#### REFERENCES

- [1] S. BALDO, *Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids*, Ann. Inst. Henri Poincaré, 7 (1990), pp. 67–90, [https://doi.org/10.1016/S0294-1449\(16\)30304-3](https://doi.org/10.1016/S0294-1449(16)30304-3).
- [2] A. P. BARTÓK, M. C. PAYNE, R. KONDOR, AND G. CSÁNYI, *Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons*, Phys. Rev. Lett., 104 (2010), 136403, <https://doi.org/10.1103/PhysRevLett.104.136403>.
- [3] C. M. BISHOP, *Pattern Recognition and Machine Learning*, Springer, New York, 2006, <https://doi.org/10.1007/978-0-387-45528-0>.
- [4] A. BRAIDES,  *$\Gamma$ -convergence for Beginners*, Oxford Lect. Ser. Math. Appl. 22, Oxford University Press, Oxford, UK, 2002.
- [5] F. BRAUER AND J. A. NOHEL, *Ordinary Differential Equations: A First Course*, University Mathematics, WA Benjamin Advanced Book Program, New York, 1967.
- [6] H. BREZIS, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer Science & Business Media, Springer, New York, 2010.
- [7] P. L. BUTZER AND R. J. NESSEL, *Fourier Analysis and Approximation. Volume 1: One-dimensional Theory*, Birkhäuser-Verlag, Berlin, 1971.
- [8] M. DASHTI, K. LAW, A. STUART, AND J. VOSS, *MAP estimators and posterior consistency in Bayesian nonparametric inverse problems*, Inverse Problems, 29 (2013), 095017, <https://doi.org/10.1088/0266-5611/29/9/095017>.
- [9] D. DÜRR AND A. BACH, *The Onsager-Machlup function as Lagrangian for the most probable path of a diffusion process*, Commun. Math. Phys., 160 (1978), pp. 153–170, <https://doi.org/10.1007/BF01609446>.
- [10] W. E, W. REN, AND E. VANDEN-ELJNDEN, *String method for the study of rare events*, Phys. Rev. B, 66 (2002), 052301, <https://doi.org/10.1103/PhysRevB.66.052301>.
- [11] W. E, W. REN, AND E. VANDEN-ELJNDEN, *Minimum action method for the study of rare events*, Comm. Pure. App. Math., 57 (2004), pp. 637–656, <https://doi.org/10.1002/cpa.20005>.
- [12] W. E, W. REN, AND E. VANDEN-ELJNDEN, *Finite temperature string method for the study of rare events*, J. Phys. Chem. B, 109 (2005), pp. 6688–6693, <https://doi.org/10.1021/jp0455430>.
- [13] L. C. EVANS, *Partial Differential Equations*, 2nd ed., Grad. Stud. Math. 19, AMS, Providence, RI, 2010, <https://doi.org/10.1090/gsm/019>.
- [14] I. FONSECA AND L. TARTAR, *The gradient theory of phase transitions for systems with two potential wells*, Proc. Roy. Soc. Edinburgh Sect. A, 111 (1989), pp. 89–102, <https://doi.org/10.1017/S030821050002504X>.
- [15] M. I. FREIDLIN AND A. D. WENTZELL, *Random Perturbations of Dynamical Systems*, 3rd ed., Grundle Math. Wiss. 260, Springer, New York, 2012, <https://doi.org/10.1007/978-3-642-25847-3>.
- [16] T. GRAFKE, T. SCHÄFER, AND E. VANDEN-ELJNDEN, *Long Term Effects of Small Random Perturbations on Dynamical Systems: Theoretical and Computational Tools*, preprint, <https://arXiv.org/abs/1604.03818>, 2016.
- [17] M. HAIRER, *An Introduction to Stochastic PDEs*, preprint, <https://arXiv.org/abs/0907.4178>, 2009.
- [18] M. HAIRER, A. STUART, AND J. VOSS, *Analysis of SPDEs arising in path sampling II: The nonlinear case*, Ann. Appl. Probab., 17 (2007), pp. 1657–1706, <https://doi.org/10.1214/07-AAP441>.
- [19] M. HEYMANN AND E. VANDEN-ELJNDEN, *The geometric minimum action method: A least action principle on the space of curves*, Comm. Pure. App. Math., 61 (2008), pp. 1052–1117, <https://doi.org/10.1002/cpa.20238>.
- [20] T. LELIÈVRE, G. STOLTZ, AND M. ROUSSET, *Free Energy Computations: A Mathematical Perspective*, World Scientific, Singapore, 2010, <https://doi.org/10.1142/9781848162488>.
- [21] G. LEONI, *A remark on the compactness for the Cahn-Hilliard functional*, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 517–523, <https://doi.org/10.1051/cocv/2013073>.
- [22] A. J. MAJDA AND B. GERSHGORIN, *Improving model fidelity and sensitivity for complex systems through empirical information theory*, Proc. Natl. Acad. Sci. USA, 108 (2011), pp. 10044–10049, <https://doi.org/10.1073/pnas.1105174108>.

- [23] M. MARCUS AND H. MINC, *A Survey of Matrix Theory and Matrix Inequalities*, vol. 14, Dover Publications, New York, 1992.
- [24] G. D. MASO, *An Introduction to  $\Gamma$ -Convergence*, Birkhäuser, Boston, 1993.
- [25] L. MODICA AND S. MORTOLA, *Un esempio di  $\Gamma^-$ -convergenza*, Boll. Un. Mat. Ital. B, 14 (1977), pp. 285–299.
- [26] G. A. PAVLIOTIS AND A. STUART, *Multiscale Methods: Averaging and Homogenization*, Texts in Appl. Math. 53, Springer, New York, 2008.
- [27] F. J. PINSKI, G. SIMPSON, A. M. STUART, AND H. WEBER, *Algorithms for Kullback–Leibler approximation of probability measures in infinite dimensions*, SIAM J. Sci. Comput., 37 (2015), pp. A2733–A2757, <https://doi.org/10.1137/14098171X>.
- [28] F. J. PINSKI, G. SIMPSON, A. M. STUART, AND H. WEBER, *Kullback–Leibler approximation for probability measures on infinite dimensional spaces*, SIAM J. Math. Anal., 47 (2015), pp. 4091–4122, <https://doi.org/10.1137/140962802>.
- [29] F. J. PINSKI AND A. M. STUART, *Transition paths in molecules at finite temperature*, J. Chem. Phys., 132 (2010), 184104.
- [30] F. J. PINSKI, A. M. STUART, AND F. THEIL,  *$\Gamma$ -limit for transition paths of maximal probability*, J. Statist. Phys., 146 (2012), pp. 955–974, <https://doi.org/10.1007/s10955-012-0443-8>.
- [31] C. E. RASMUSSEN AND C. K. I. WILLIAMS, *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, MA, 2006.
- [32] G. ROYER, *An Initiation to Logarithmic Sobolev Inequalities*, AMS, Providence, RI, 2007.
- [33] E. VANDEN-EIJNDEN, *Transition path theory*, in *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, Springer, Amsterdam, 2014, pp. 91–100, <https://doi.org/10.1007/978-94-007-7606-7>.