



The large time profile for Hamilton– Jacobi–Bellman equations

Item Type	Article
Authors	Gomes, Diogo A.;Mitake, Hiroyoshi;Tran, Hung V.
Citation	Gomes, D. A., Mitake, H., & Tran, H. V. (2021). The large time profile for Hamilton–Jacobi–Bellman equations. <i>Mathematische Annalen</i> . doi:10.1007/s00208-021-02320-5
Eprint version	Post-print
DOI	10.1007/s00208-021-02320-5
Publisher	Springer Science and Business Media LLC
Journal	Mathematische Annalen
Rights	Archived with thanks to Mathematische Annalen
Download date	2024-03-02 16:19:57
Link to Item	http://hdl.handle.net/10754/663760

THE LARGE TIME PROFILE FOR HAMILTON–JACOBI–BELLMAN EQUATIONS

DIOGO A. GOMES, HIROYOSHI MITAKE, AND HUNG V. TRAN

ABSTRACT. Here, we study the large-time limit of viscosity solutions of the Cauchy problem for second-order Hamilton–Jacobi–Bellman equations with convex Hamiltonians in the torus. This large-time limit solves the corresponding stationary problem, sometimes called the ergodic problem. This problem, however, has multiple viscosity solutions and, thus, a key question is which of these solutions is selected by the limit. Here, we provide a representation for the viscosity solution to the Cauchy problem in terms of generalized holonomic measures. Then, we use this representation to characterize the large-time limit in terms of the initial data and generalized Mather measures. In addition, we establish various results on generalized Mather measures and duality theorems that are of independent interest.

1. INTRODUCTION

Under a broad range of conditions, as time goes to infinity, viscosity solutions of the Cauchy problem for Hamilton–Jacobi–Bellman equations with convex Hamiltonians converge to stationary solutions. In the last two decades, this matter was investigated in detail. For first-order Hamilton–Jacobi equations, see, for example, [3, 13, 16, 24, 33] and references therein. For second-order Hamilton–Jacobi–Bellman equations, the large-time convergence was studied in the non-degenerate case in [4], and in the general (possibly degenerate) case in [7, 28]. However, the stationary problem may have multiple viscosity solutions and, thus, a natural question is which of these viscosity solutions is selected. Here, in Theorem 1.2, we answer this question by giving a characterization of the limit in terms of generalized Mather measures.

More precisely, we study the large time behavior of the solutions of the following problem.

Problem 1. *Let $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ be the flat n -dimensional torus. Consider a **non-degenerate** diffusion coefficient, $a : \mathbb{T}^n \rightarrow [0, \infty)$. Fix a Hamiltonian $H : \mathbb{T}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, and continuous initial data, $u_0 : \mathbb{T}^n \rightarrow \mathbb{R}$. Find a (viscosity) solution, $u : \mathbb{T}^n \times [0, \infty) \rightarrow \mathbb{R}$.*

Date: June 28, 2021.

2010 Mathematics Subject Classification. 35B40, 37J50, 49L25 .

Key words and phrases. The large time profile; Hamilton–Jacobi–Bellman equations; Ergodic problems; Nonlinear adjoint methods.

The work of DG was partially supported by KAUST baseline funds and KAUST OSR-CRG2017-3452. The work of HM was partially supported by the JSPS grants: KAKENHI #19K03580, #19H00639, #17KK0093, #20H01816. The work of HT was partially supported by NSF grant DMS-1664424 and NSF CAREER grant DMS-1843320.

\mathbb{R} , to the Hamilton–Jacobi–Bellman equation

$$u_t - a(x)\Delta u + H(x, Du) = 0 \quad \text{in } \mathbb{T}^n \times (0, \infty), \quad (1.1)$$

where Du and Δu , respectively, denote the spatial gradient and Laplacian of u , and with the initial data

$$u(x, 0) = u_0(x) \quad \text{on } \mathbb{T}^n.$$

Here, we *always* work with viscosity solutions, and, thus, the term “viscosity” is omitted henceforth.

Our goal is to study the large-time behavior of the solution of Problem 1 and provide a characterization of the asymptotic profile of $u(x, t)$ as $t \rightarrow \infty$. For that, we work under the following assumptions.

Assumption 1. *The functions u_0 , a , and H are of class C^2 .*

Assumption 2. *For every $x \in \mathbb{T}^n$, $D_{pp}^2 H(x, p) > 0$ for all $(x, p) \in \mathbb{T}^n \times \mathbb{R}^n$.*

Assumption 3. *There exist $m > 1$ and $C_0 > 0$ such that, for all $(x, p) \in \mathbb{T}^n \times \mathbb{R}^n$,*

$$\begin{cases} \frac{|p|^m}{C_0} - C_0 \leq H(x, p) \leq C_0(|p|^m + 1), \\ |D_x H(x, p)| \leq C_0(1 + |p|^m), \\ |D_p H(x, p)| \leq C_0(1 + |p|^{m-1}). \end{cases}$$

Because of the strict convexity in Assumption 2 and the growth conditions in Assumption 3, the Lagrangian,

$$L(x, q) = \sup_{p \in \mathbb{R}^n} (p \cdot q - H(x, p)), \quad (1.2)$$

is finite for all $(x, q) \in \mathbb{T}^n \times \mathbb{R}^n$. Moreover, by increasing C_0 , if necessary, we have

$$\frac{|q|^{m'}}{C_0} - C_0 \leq L(x, q) \leq C_0(|q|^{m'} + 1) \quad \text{for all } (x, q) \in \mathbb{T}^n \times \mathbb{R}^n,$$

where

$$\frac{1}{m} + \frac{1}{m'} = 1. \quad (1.3)$$

Under Assumptions 1-3, Problem 1 has a unique Lipschitz solution (see [2, 33] and the references therein). Furthermore, as established in [7, Theorem 1.1], if Assumptions 1-3 hold, there exist $(u_\infty, c) \in \text{Lip}(\mathbb{T}^n) \times \mathbb{R}$ such that

$$u(\cdot, t) + ct \rightarrow u_\infty \quad \text{uniformly on } \mathbb{T}^n \text{ as } t \rightarrow \infty. \quad (1.4)$$

Moreover, (u_∞, c) solves (in the viscosity sense) the following *ergodic problem*.

Problem 2. *Under the setting of Problem 1, find $(v, c) \in \text{Lip}(\mathbb{T}^n) \times \mathbb{R}$ such that*

$$-a(x)\Delta v + H(x, Dv) = c \quad \text{in } \mathbb{T}^n. \quad (1.5)$$

The constant c in Problem 2, the *ergodic constant*, is uniquely determined by H and a under Assumptions 1-3 (see [33, Theorem 4.5, Proposition 4.8] for instance). However, because the operator on the left-hand side of (1.5) is not strictly monotone in v , Problem 2 may have multiple solutions even up to additive constants as the examples in [33, Chapter 6] illustrate. Thus, a key question is which solution is selected by the large-time limit. Here, we answer this question and give a representation formula for the limit function u_∞ that highlights the dependence on the initial data u_0 .

Remark 1. By adding a constant to H , we can assume that the ergodic constant in Problem 2 vanishes, that is, $c = 0$. Therefore, we *always* assume that $c = 0$ henceforth.

The selection problem, that is, to characterize which of the solutions of Problem 2 is selected by the limit $t \rightarrow \infty$, was first addressed for first-order Hamilton–Jacobi equations in [13, Theorem 3.1]. Moreover, the last two authors of the present paper gave an elementary proof of that selection result in [35, Theorem 1.3]. These results characterize u_∞ in the first-order case. If $a > 0$ on \mathbb{T}^n , Problem 2 has a unique solution (up to additive constants), see, for example, [4, 20]. Thus, in this case, the selection problem is simple since only the additive constant needs to be determined. However, prior to this work, there was no characterization of the limit for the general case where a can degenerate. In this case, uniqueness for Problem 2 may not hold and the techniques for first-order equations that use deterministic trajectories do not apply. The present paper closes a key gap in the understanding of the large-time limit of solutions of Problem 1 by establishing a representation result for u_∞ in terms of generalized Mather measures and the initial data.

1.1. Main results on representation formulas and limiting profiles. A holonomic measure on $\mathbb{T}^n \times \mathbb{R}^n$ is a probability measure, $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$, that satisfies

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} q \cdot D\varphi(z) d\mu(z, q) = 0$$

for any $\varphi \in C^1(\mathbb{T}^n)$. In Aubry–Mather theory, the set of holonomic measures is of paramount importance as these measures generalize closed curves. These measures play a critical role in understanding Lagrangian dynamics and viscosity solutions of first-order Hamilton–Jacobi equations.

Generalized Mather measures were introduced in the setting of the stochastic Mather problem in [20] and then further studied for fully nonlinear elliptic equations in [21]. Later, these measures were studied in different contexts and used to study asymptotic problems in [8, 22, 34, 35] and [25].

Here, we define a class of generalized holonomic measures that are suitable for the study of time-dependent problems. Let m be the exponent in Assumption 3 and m' its conjugate exponent as in (1.3). Fix ζ with $1 < \zeta < m'$. Let $t_0, t_1 \in \mathbb{R}$ with $0 \leq t_0 < t_1$. Let $\mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ be the set of all Radon measures on $\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]$, and $\mathcal{R}^+(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ be the set of all non-negative Radon measures on $\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]$. Let $\mathcal{P}(\mathbb{T}^n)$ and $\mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be the set of probability measures on \mathbb{T}^n and $\mathbb{T}^n \times \mathbb{R}^n$, respectively.

For $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$, $\mathcal{H}(\nu_0, \nu_1; t_0, t_1)$ is the set of all $\gamma \in \mathcal{R}^+(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ satisfying

$$\int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} |q|^\zeta d\gamma(z, q, s) < \infty$$

and

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} (\varphi_t(z, s) - a(z)\Delta\varphi(z, s) + q \cdot D\varphi(z, s)) d\gamma(z, q, s) \\ &= \int_{\mathbb{T}^n} \varphi(z, t_1) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, t_0) d\nu_0(z) \end{aligned}$$

for all $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$. Moreover, we set

$$\mathcal{H}(\nu_1; t_0, t_1) := \bigcup_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \mathcal{H}(\nu_0, \nu_1; t_0, t_1).$$

Fix any $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$. It is worth emphasizing that while we do not know that $\mathcal{H}(\nu_0, \nu_1; t_0, t_1) \neq \emptyset$ for each $\nu_0 \in \mathcal{P}(\mathbb{T}^n)$, the set $\mathcal{H}(\nu_1; t_0, t_1)$ is always non-empty as shown in Corollary 2.3. For $\gamma \in \mathcal{H}(\nu_1; t_0, t_1)$, we denote by ν^γ the unique element in $\mathcal{P}(\mathbb{T}^n)$ such that $\gamma \in \mathcal{H}(\nu^\gamma, \nu_1; t_0, t_1)$.

In Section 2, we use the measures in $\mathcal{H}(\nu_1; t_0, t_1)$ to obtain the representation formula for the solutions of Problem 1. This formula generalizes the optimal control representation for first-order Hamilton–Jacobi equations and the stochastic optimal control representation for second-order Hamilton–Jacobi equations (see [14, 31] for example).

Theorem 1.1. *Let u solve Problem 1. Suppose that Assumptions 1–3 hold. Then, for any $\nu \in \mathcal{P}(\mathbb{T}^n)$ and $t > 0$, we have*

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) = \inf_{\gamma \in \mathcal{H}(\nu; 0, t)} \left[\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu^\gamma(z) \right].$$

We believe Theorem 1.1 is new in the literature for possibly degenerate second-order equations. To understand the possible limits as $t \rightarrow \infty$ of the solutions of Problem 1, we rely on a characterization of the solutions of Problem 2 established in [35]. This characterization uses generalized Mather measures. As in the time-dependent case, we begin by defining the set of stationary generalized holonomic measures. Let

$$\mathcal{H} := \left\{ \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) : \int_{\mathbb{T}^n \times \mathbb{R}^n} |q|^\zeta d\mu(z, q) < \infty, \right. \\ \left. \int_{\mathbb{T}^n \times \mathbb{R}^n} (-a(z)\Delta\varphi(z) + q \cdot D\varphi(z)) d\mu(z, q) = 0 \quad \text{for all } \varphi \in C^2(\mathbb{T}^n) \right\}.$$

We consider the variational problem

$$\inf_{\mu \in \mathcal{H}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\mu(x, q). \quad (1.6)$$

A *generalized Mather measure* is a solution of the minimization problem in (1.6). We denote by $\widetilde{\mathcal{M}}$ the set of all generalized Mather measures. Moreover, the set of all generalized projected Mather measures, \mathcal{M} , is the set of all measures in \mathbb{T}^n which are

the projections to \mathbb{T}^n of generalized Mather measures; that is, $\nu \in \mathcal{M}$ if there exists $\mu \in \widetilde{\mathcal{M}}$ such that for all $\varphi \in C(\mathbb{T}^n)$,

$$\int_{\mathbb{T}^n} \varphi(x) d\nu(x) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi(x) d\mu(x, q).$$

Remark 2. Since we normalize that the ergodic constant is 0 as noted in Remark 1, we have

$$\inf_{\mu \in \mathcal{H}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\mu(x, q) = 0, \quad (1.7)$$

as shown in [34, Propositions 2.2 and 2.3]. See also Proposition B.2 in Appendix B for the general diffusion matrix case.

Before stating our main results, we recall a uniqueness result for Problem 2. Let ν_1 and ν_2 be two solutions to Problem 2 with

$$\int_{\mathbb{T}^n} \nu_1(x) d\nu = \int_{\mathbb{T}^n} \nu_2(x) d\nu \quad \text{for all } \nu \in \mathcal{M}.$$

Then, according to [35, Theorem 1.1], we have

$$\nu_1 = \nu_2 \quad \text{on } \mathbb{T}^n.$$

Therefore, to characterize u_∞ uniquely, it is enough to specify the integral of u_∞ with respect to generalized projected Mather measures. This characterization is given in Theorem 1.2 below. The corresponding proof is presented in Section 4.

Theorem 1.2. *Suppose that Assumptions 1-3 hold. Let u solve Problem 1 and u_∞ be the large time limit of u as in (1.4). Then, for any $\nu \in \mathcal{M}$, we have*

$$\int_{\mathbb{T}^n} u_\infty(z) d\nu(z) = \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z) \right], \quad (1.8)$$

where for $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$, $d(\nu_0, \nu_1)$ is the generalized Mañé critical potential from ν_0 to ν_1 given by

$$d(\nu_0, \nu_1) := \inf_{\substack{\gamma \in \mathcal{H}(\nu_0, \nu_1; 0, t) \\ t > 0}} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma(x, q, s). \quad (1.9)$$

We also notice that Theorem 1.2 answers the open question addressed in [33, Section 5.7, (iv)]. We refer to [7, 33, 34, 35] for earlier developments.

Remark 3. The assumption that the initial data, u_0 , is C^2 is not critical as by approximation, we can solve Problem 1 with continuous initial data, and prove the large time convergence in (1.4). The regularity of a and H is more delicate as several estimates require quantitative bounds for their derivatives.

Remark 4. We choose to consider (1.1) with a specific diffusion matrix $A(x) = a(x)I_n$ for $x \in \mathbb{T}^n$, where I_n is the identity matrix of size n , to simplify the presentation. However, the results in Theorems 1.1–1.2 can be generalized for more general diffusion matrices. The precise statements and proofs are more delicate and are discussed in Theorems B.1 and B.3 in Appendix B.

1.2. Duality results. For $t \geq 0$, let $S_t : C(\mathbb{T}^n) \rightarrow C(\mathbb{T}^n)$ be the solution operator for Problem 1; that is, $S_t u_0(\cdot) = u(\cdot, t)$, where u solves Problem 1 with initial data u_0 . We assume the following compactness property for S_t .

Assumption 4. *Let $X_0 := \{u_0 \in C(\mathbb{T}^n) : \min_{\mathbb{T}^n} u_0 = 0\}$. For any $t > 0$, $S_t(X_0)$ is compact in $C(\mathbb{T}^n)$.*

The map S_t often enjoys the compactness in the preceding assumption. Two important cases are examined in Theorem 1.6 below. Under Assumption 4, we have the following remarkable uniform convergence property that is proved at the end of Section 4.

Proposition 1.3. *Consider the setting of Problem 1 and suppose that Assumptions 1-4 hold. Given $\varepsilon > 0$, there exists $T > 0$ such that for any initial data $u_0 \in C^2(\mathbb{T}^n)$, the corresponding solution, u , to Problem 1 satisfies*

$$\|u(\cdot, t) - u_\infty\|_{L^\infty(\mathbb{T}^n)} < \varepsilon$$

for all $t > T$, where

$$u_\infty(x) = \lim_{t \rightarrow \infty} u(x, t) \quad \text{for } x \in \mathbb{T}^n.$$

The study of Hamilton–Jacobi–Bellman equations using duality theory can be traced back to [17, 18, 27]. In the context of Aubry–Mather theory, the elliptic case was considered in [20], the fully nonlinear case in [21], and the parabolic case in [23]. In a sequence of works [29, 30, 32], various duality theorems associated with Hamilton–Jacobi–Bellman equations (1.1) were studied in the non-degenerate case. The connection between Aubry–Mather theory and optimal transport was explored in [6]. We also refer to [14, 31] for recent work concerning stochastic optimal transport.

For $t > 0$, let $h_t : \mathcal{P}(\mathbb{T}^n) \times \mathcal{P}(\mathbb{T}^n) \rightarrow \mathbb{R} \cup \{+\infty\}$ be

$$h_t(\nu_0, \nu_1) = \begin{cases} \inf_{\gamma \in \mathcal{H}(\nu_0, \nu_1; 0, t)} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma & \text{if } \mathcal{H}(\nu_0, \nu_1; 0, t) \neq \emptyset, \\ +\infty & \text{if } \mathcal{H}(\nu_0, \nu_1; 0, t) = \emptyset \end{cases}$$

for $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$. In Section 5, we apply the Legendre–Fenchel–Rockafellar theorem to obtain representation formulas for h_t and d .

Theorem 1.4. *Consider the setting of Problem 1 and suppose that Assumptions 1-3 hold. Suppose that $h_t(\nu_0, \nu_1) < +\infty$ for given $t > 0$ and $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$. Then,*

$$h_t(\nu_0, \nu_1) = \sup_{w \in \mathcal{E}} \left[\int_{\mathbb{T}^n} w(z, t) d\nu_1(z) - \int_{\mathbb{T}^n} w(z, 0) d\nu_0(z) \right], \quad (1.10)$$

where

$$\mathcal{E} := \{w \in \text{Lip}(\mathbb{T}^n \times [0, t]) : w \text{ is a viscosity solution to (1.1)}\}.$$

The preceding result extends the duality formula derived in [36] for the representation of the 2-Wasserstein distance using the Benamou–Brenier dynamic formulation of optimal transport [5]. In the non-degenerate setting, a version of Theorem 1.4 was

proven in [29, 30, 32]. For first-order equations, see [6]. The degenerate setting addressed here is substantially more complex, as the proof of Theorem 1.4 requires a delicate approximation argument.

Under the compactness assumption, Assumption 4, we obtain the following duality result for d restricted to the projected Mather set, \mathcal{M} .

Theorem 1.5. *Consider the setting of Problem 1 and suppose that Assumptions 1-4 hold. Let $\nu_0, \nu_1 \in \mathcal{M}$ and suppose that $d(\nu_0, \nu_1) < \infty$. Then,*

$$d(\nu_0, \nu_1) = \sup_{w \in \mathcal{S}} \left[\int_{\mathbb{T}^n} w(z) d\nu_1(z) - \int_{\mathbb{T}^n} w(z) d\nu_0(z) \right],$$

where

$$\mathcal{S} := \{w \in \text{Lip}(\mathbb{T}^n) : w \text{ is a viscosity solution to (1.5)}\}. \quad (1.11)$$

Finally, we discuss the compactness assumption, Assumption 4.

Theorem 1.6. *Consider the setting of Problem 1 and suppose that Assumptions 1-3 hold. Then Assumption 4 holds if $a \equiv 0$ or $m > 2$.*

It is still an open problem in the literature of whether Assumption 4 holds or not in case $m \in (1, 2]$.

The paper is organized as follows. Section 2 is devoted to the proof of Theorem 1.1. In Section 3, we discuss the properties of generalized holonomic measures. Next, in Section 4, we introduce the generalized Peierls barrier and Aubry set and prove Theorem 1.2. Section 5 is devoted to the proof of Theorems 1.4 and 1.5. In Section 6, we discuss Assumption 4, and give the proof of Theorem 1.6. In Appendix A, we discuss the approximation of viscosity solutions of (1.1) by C^2 -subsolutions of an approximate equation. Finally, in Appendix B, we consider the Hamilton–Jacobi–Bellman equation with a general diffusion matrix, and generalize Theorems 1.1 and 1.2 for this case.

Acknowledgements. We would like to thank Hitoshi Ishii for his suggestions on the approximations of viscosity solutions and subsolutions in Appendix B. We are grateful to Toshio Mikami for the discussions on Theorem 1.1 and for giving us relevant references on the duality result in Theorem 1.4.

2. REPRESENTATION FORMULA FOR SOLUTIONS

In this section, we prove Theorem 1.1, which gives a variational representation formula for solutions to Problem 1 in terms of generalized holonomic measures. We begin by establishing a lower bound for this variational formula.

Lemma 2.1. *Suppose that Assumptions 1-3 hold. Let u solve Problem 1. Then, for all $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$ and $t > 0$,*

$$\int_{\mathbb{T}^n} u(z, t) d\nu_1(z) \leq \inf_{\gamma \in \mathcal{H}(\nu_1; 0, t)} \left[\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u(z, 0) d\nu^\gamma(z) \right].$$

Proof. Let u solve Problem 1. Because Assumptions 1-3 hold, u is globally Lipschitz continuous on $\mathbb{T}^n \times [0, \infty)$ (see [2, Proposition 3.5], [33, Proposition 4.15] for instance).

Fix smooth symmetric standard mollifiers, $\theta \in C_c^\infty(\mathbb{R}^n, [0, \infty))$ and $\rho \in C_c^\infty(\mathbb{R}, [0, \infty))$. More precisely, $\text{supp } \theta \subset \overline{B}(0, 1) \subset \mathbb{R}^n$, $\text{supp } \rho \subset [-1, 1] \subset \mathbb{R}$, $\theta(x) = \theta(-x)$, $\rho(s) = \rho(-s)$, and $\|\theta\|_{L^1(\mathbb{R}^n)} = \|\rho\|_{L^1(\mathbb{R})} = 1$. For $\alpha > 0$, let $\theta^\alpha(x) = \alpha^{-n}\theta(\alpha^{-1}x)$ for $x \in \mathbb{R}^n$, and $\rho^\alpha(t) = \alpha^{-1}\rho(\alpha^{-1}t)$ for $t \in \mathbb{R}$. Let u^α be the function given by (A.1); that is,

$$u^\alpha(x, t) = \int_0^\infty \rho^\alpha(s) \int_{\mathbb{T}^n} \theta^\alpha(y) u(x - y, t - s) dy ds \quad \text{for } (x, t) \in \mathbb{T}^n \times [\alpha, \infty).$$

Set

$$\tilde{u}(x, t) := u^\alpha(x, t + \alpha) \quad \text{for all } (x, t) \in \mathbb{T}^n \times [0, \infty).$$

We notice \tilde{u} is a $C^2(\mathbb{T}^n \times [0, \infty))$ subsolution to (A.2); that is,

$$\tilde{u}_t - a(x)\Delta\tilde{u} + H(x, D\tilde{u}) \leq C\alpha^{1/2} \quad \text{on } \mathbb{T}^n \times [0, \infty)$$

as stated in Proposition A.1 in Appendix A.

Fix any $\gamma \in \mathcal{H}(\nu_1; 0, t)$. By the definition of $\mathcal{H}(\nu_1; 0, t)$,

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\tilde{u}_t(z, s) - a(z)\Delta\tilde{u}(z, s) + q \cdot D\tilde{u}(z, s)) d\gamma(z, q, s) \\ &= \int_{\mathbb{T}^n} \tilde{u}(z, t) d\nu_1(z) - \int_{\mathbb{T}^n} \tilde{u}(z, 0) d\nu^\gamma(z). \end{aligned}$$

Because of the Legendre transform definition in (1.2),

$$q \cdot D\tilde{u}(z, s) \leq L(z, q) + H(z, D\tilde{u}(z, s)).$$

Accordingly, we have

$$\begin{aligned} & \int_{\mathbb{T}^n} \tilde{u}(z, t) d\nu_1 - \int_{\mathbb{T}^n} \tilde{u}(z, 0) d\nu^\gamma(z) \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\tilde{u}_t - a(z)\Delta\tilde{u} + q \cdot D\tilde{u}) d\gamma(z, q, s) \\ &\leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\tilde{u}_t - a(z)\Delta\tilde{u} + H(z, D\tilde{u}) + L(z, q)) d\gamma(z, q, s) \\ &\leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + C\alpha^{1/2}t. \end{aligned}$$

Sending $\alpha \rightarrow 0$, we obtain

$$\int_{\mathbb{T}^n} u(x, t) d\nu_1 \leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u(z, 0) d\nu^\gamma(z).$$

Because $\gamma \in \mathcal{H}(\nu_1; 0, t)$ is arbitrary, the statement follows. \square

Remark 5. If $a > 0$ on \mathbb{T}^n , that is (1.1) is uniformly parabolic, the solution to Problem 1, u , is in $C^2(\mathbb{T}^n \times [0, \infty))$. In this case, in the preceding proof, we do not need to use u^α , the smoothed version of u . However, if a degenerates, we expect only Lipschitz regularity for u . Therefore, we need to build C^2 approximate subsolutions.

This construction is rather technical and is described in Proposition A.1 in Appendix A. The general diffusion case, is examined in Proposition B.12 in Appendix B.

To prove the opposite bound, we consider the following regularized version of Problem 1.

Problem 3. *In the setting of Problem 1 and for $\varepsilon > 0$, find $u^\varepsilon : \mathbb{T}^n \times [0, \infty) \rightarrow \mathbb{R}$ solving*

$$\begin{cases} u_t^\varepsilon - (a(x) + \varepsilon)\Delta u^\varepsilon + H(x, Du^\varepsilon) = 0 & \text{in } \mathbb{T}^n \times (0, \infty), \\ u(x, 0) = u_0(x) & \text{on } \mathbb{T}^n. \end{cases} \quad (2.1)$$

If Assumptions 1-3 hold, Problem 3 has a unique solution, $u^\varepsilon \in C^2(\mathbb{T}^n \times [0, +\infty))$ due to the added viscous term (see [1, 19]). Moreover, u^ε is Lipschitz continuous uniformly in $\varepsilon \in (0, 1)$. Further, by standard viscosity solution theory, $u^\varepsilon \rightarrow u$ locally uniformly on $\mathbb{T}^n \times [0, \infty)$, as $\varepsilon \rightarrow 0$, where u solves Problem 1.

Now, we use the nonlinear adjoint method, see [15, 37], to construct measures that satisfy an approximated holonomy condition.

Lemma 2.2. *For $\varepsilon > 0$, let u^ε solve Problem 3. Then, for any $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$, there exist a Radon measure $\gamma^\varepsilon \in \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times (t_0, t_1))$ and a probability measure $\nu_0^\varepsilon \in \mathcal{P}(\mathbb{T}^n)$ such that*

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} (\varphi_t(z, s) - a(z)\Delta\varphi(z, s) - \varepsilon\Delta\varphi(z, s) + q \cdot D\varphi(z, s)) d\gamma^\varepsilon(z, q, s) \\ &= \int_{\mathbb{T}^n} \varphi(z, t_1) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, t_0) d\nu_0^\varepsilon(z) \end{aligned} \quad (2.2)$$

for all $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$, and γ^ε -almost everywhere, $q = D_p H(x, Du^\varepsilon)$.

Proof. For $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$, the linearization of (2.1) around u^ε is

$$\mathcal{L}^\varepsilon[\varphi] = \varphi_t + D_p H(x, Du^\varepsilon) \cdot D\varphi - (a(x) + \varepsilon)\Delta\varphi.$$

Accordingly, the corresponding adjoint equation is the Fokker-Planck equation

$$\begin{cases} -\sigma_t^\varepsilon - \Delta(a(x)\sigma^\varepsilon) - \operatorname{div}(D_p H(x, Du^\varepsilon)\sigma^\varepsilon) = \varepsilon\Delta\sigma^\varepsilon & \text{in } \mathbb{T}^n \times (t_0, t_1), \\ \sigma^\varepsilon(x, t_1) = \nu_1 & \text{on } \mathbb{T}^n. \end{cases} \quad (2.3)$$

By standard properties of the Fokker-Planck equation (see [19]),

$$\sigma^\varepsilon > 0 \text{ on } \mathbb{T}^n \times [t_0, t_1] \quad \text{and} \quad \int_{\mathbb{T}^n} \sigma^\varepsilon(x, t) dx = 1 \text{ for all } t \in [t_0, t_1].$$

Next, for each $\varepsilon > 0$ and $t \in [t_0, t_1]$, let $\beta_t^\varepsilon \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be the probability measure determined by

$$\int_{\mathbb{T}^n} \psi(x, Du^\varepsilon)\sigma^\varepsilon(x, t) dx = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, p) d\beta_t^\varepsilon(x, p)$$

for all $\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$. For $t \in [t_0, t_1]$, let $\gamma_t^\varepsilon \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be the pullback of β_t^ε by the map $\Phi(x, q) = (x, D_q L(x, q))$, that is,

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, p) d\beta_t^\varepsilon(x, p) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, D_q L(x, q)) d\gamma_t^\varepsilon(x, q)$$

for all $\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$. Note that Φ is invertible as $D_{qq}^2 L > 0$. Hence, γ_t^ε is well defined.

Define the measures $\beta^\varepsilon, \gamma^\varepsilon \in \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ by

$$\int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} f d\beta^\varepsilon = \int_{t_0}^{t_1} \int_{\mathbb{T}^n \times \mathbb{R}^n} f(\cdot, t) d\beta_t^\varepsilon dt,$$

and

$$\int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} f d\gamma^\varepsilon = \int_{t_0}^{t_1} \int_{\mathbb{T}^n \times \mathbb{R}^n} f(\cdot, t) d\gamma_t^\varepsilon dt$$

for any $f \in C_c(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$.

Multiplying the first equation in (2.3) by an arbitrary function, $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$, and integrating on \mathbb{T}^n , by using the integration by parts, we gather

$$\begin{aligned} \varepsilon \int_{\mathbb{T}^n} \sigma^\varepsilon \Delta \varphi dx &= - \int_{\mathbb{T}^n} \varphi \sigma_t^\varepsilon dx + \int_{\mathbb{T}^n} (-a(x) \Delta \varphi + D_p H(x, Du^\varepsilon) \cdot D\varphi) \sigma^\varepsilon dx \\ &= - \int_{\mathbb{T}^n} (\varphi \sigma^\varepsilon)_t dx + \int_{\mathbb{T}^n} (\varphi_t - a(x) \Delta \varphi + D_p H(x, Du^\varepsilon) \cdot D\varphi) \sigma^\varepsilon dx. \end{aligned}$$

Next, integrating on $[t_0, t_1]$, we deduce the identity

$$\begin{aligned} \varepsilon \int_{t_0}^{t_1} \int_{\mathbb{T}^n} \sigma^\varepsilon \Delta \varphi dx dt &= \varepsilon \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} \Delta \varphi d\gamma^\varepsilon(x, q, t) \\ &= - \int_{t_0}^{t_1} \int_{\mathbb{T}^n} (\varphi \sigma^\varepsilon)_t dx dt + \int_{t_0}^{t_1} \int_{\mathbb{T}^n} (\varphi_t - a(x) \Delta \varphi + D_p H(x, Du^\varepsilon) \cdot D\varphi) \sigma^\varepsilon dx dt \\ &= - \left[\int_{\mathbb{T}^n} \varphi(\cdot, t_1) \sigma^\varepsilon(\cdot, t_1) dx - \int_{\mathbb{T}^n} \varphi(\cdot, t_0) \sigma^\varepsilon(\cdot, t_0) dx \right] \\ &\quad + \int_{t_0}^{t_1} \int_{\mathbb{T}^n \times \mathbb{R}^n} (\varphi_t - a(x) \Delta \varphi + q \cdot D\varphi) d\gamma_t^\varepsilon(x, q) \\ &= - \left[\int_{\mathbb{T}^n} \varphi(\cdot, t_1) d\nu_1 - \int_{\mathbb{T}^n} \varphi(\cdot, t_0) d\nu_0^\varepsilon \right] \\ &\quad + \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} (\varphi_t - a(x) \Delta \varphi + q \cdot D\varphi) d\gamma^\varepsilon(x, q, t), \end{aligned}$$

where $d\nu_0^\varepsilon := \sigma^\varepsilon(x, t_0) dx$, which implies the desired result. In the above, we used the fact that $D_p H(x, D_q L(x, q)) = q$. \square

Corollary 2.3. *Under Assumptions 1–3, for all $0 < t_0 < t_1$ and $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$,*

$$\mathcal{H}(\nu_1; t_0, t_1) \neq \emptyset. \quad (2.4)$$

Proof. Let $\nu_0^\varepsilon \in \mathcal{P}(\mathbb{T}^n)$ and $\gamma^\varepsilon \in \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ be the measures given by Lemma 2.2. Because u^ε is Lipschitz in x , there exists $C > 0$, which is independent of ε , such that

$$\|Du^\varepsilon(\cdot, t)\|_{L^\infty(\mathbb{T}^n)} \leq C$$

for all $t \in [t_0, t_1]$. Therefore, there exists a sequence $\{\varepsilon_j\} \rightarrow 0$, $\nu_0 \in \mathcal{P}(\mathbb{T}^n)$, and $\gamma \in \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ such that

$$\nu_0^{\varepsilon_j} \rightharpoonup \nu_0 \quad \text{and} \quad \gamma^{\varepsilon_j} \rightharpoonup \gamma \quad \text{as } j \rightarrow \infty, \quad (2.5)$$

weakly in $\mathcal{P}(\mathbb{T}^n)$ and $\mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$, respectively. Thus, using (2.2), we conclude that $\gamma \in \mathcal{H}(\nu_0, \nu_1; t_0, t_1)$, which implies (2.4). \square

Finally, we use Lemma 2.2 to establish the opposite inequality to the one in Lemma 2.1.

Lemma 2.4. *For any $\nu \in \mathcal{P}(\mathbb{T}^n)$ and $t > 0$, we have*

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) \geq \inf_{\gamma \in \mathcal{H}(\nu; 0, t)} \left\{ \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu^\gamma(z) \right\}.$$

Proof. For $s \in [0, t]$, let γ_s^ε be the measure constructed in the proof of Lemma 2.2 for $t_0 = 0$ and $t_1 = t$. By the definition of Legendre transform in (1.2), we have

$$L(z, q) = D_p H(z, D_q L(z, q)) D_q L(z, q) - H(z, D_q L(z, q)).$$

Therefore,

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n} L(z, q) d\gamma_s^\varepsilon(z, q) \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(z, D_q L(z, q)) \cdot D_q L(z, q) - H(z, D_q L(z, q))) d\gamma_s^\varepsilon(z, q) \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(z, p) \cdot p - H(z, p)) d\beta_s^\varepsilon(z, p) \\ &= \int_{\mathbb{T}^n} (D_p H(x, Du^\varepsilon) \cdot Du^\varepsilon - H(x, Du^\varepsilon)) \sigma^\varepsilon(x, s) dx \end{aligned}$$

for all $s \in [0, t]$. Moreover, integrating by parts and using the adjoint equation, (2.3), and (2.1), we obtain

$$\begin{aligned} & \int_{\mathbb{T}^n} (D_p H(x, Du^\varepsilon) \cdot Du^\varepsilon - H(x, Du^\varepsilon)) \sigma^\varepsilon dx \\ &= \int_{\mathbb{T}^n} -\operatorname{div}(D_p H(x, Du^\varepsilon) \sigma^\varepsilon) u^\varepsilon - H(x, Du^\varepsilon) \sigma^\varepsilon dx \\ &= \int_{\mathbb{T}^n} (\sigma_t^\varepsilon + \varepsilon \Delta \sigma^\varepsilon + \Delta(a \sigma^\varepsilon)) u^\varepsilon + (u_t^\varepsilon - \varepsilon \Delta u^\varepsilon - a \Delta u^\varepsilon) \sigma^\varepsilon dx \\ &= \int_{\mathbb{T}^n} (u^\varepsilon \sigma^\varepsilon)_t dx. \end{aligned}$$

Integrating on $[0, t]$ yields

$$\begin{aligned} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma^\varepsilon(x, q, s) &= \int_0^t \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\gamma_s^\varepsilon(x, q) ds \\ &= \int_{\mathbb{T}^n} u^\varepsilon(x, t) d\nu - \int_{\mathbb{T}^n} u_0(x) d\nu_0^\varepsilon. \end{aligned}$$

Taking subsequences $\{\gamma^{\varepsilon_j}\}$ and $\{\nu_0^{\varepsilon_j}\}$ as in (2.5) yields

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) = \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z).$$

Because $\gamma \in \mathcal{H}(\nu_0, \nu; 0, t)$, we obtain the inequality claimed in the statement. \square

Proof of Theorem 1.1. The statement follows directly by combining Lemma 2.2 with Lemma 2.4. \square

3. PROPERTIES OF GENERALIZED HOLONOMIC MEASURES

Here, we discuss two properties of generalized holonomic measures in Lemmas 3.1 and 3.2. Before proceeding, we recall that the results in [34, Proposition 2.3], [35, Lemma 2.1] (see also the proof of Proposition B.2 in Appendix B) imply that the Mather set and the projected Mather set are not empty; that is,

$$\widetilde{\mathcal{M}} \neq \emptyset, \quad \text{and} \quad \mathcal{M} \neq \emptyset.$$

Lemma 3.1. *Let $t > 0$, $\mu \in \widetilde{\mathcal{M}}$, and set $\nu := \text{proj}_{\mathbb{T}^n} \mu \in \mathcal{M}$. Define $d\gamma(x, q, s) := d\mu(x, q) ds$ for all $s \in [0, t]$; that is,*

$$\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} f(x, q, s) d\gamma(x, q, s) = \int_0^t \int_{\mathbb{T}^n \times \mathbb{R}^n} f(x, q, s) d\mu(x, q) ds \quad (3.1)$$

for all $f \in C_c(\mathbb{T}^n \times \mathbb{R}^n \times [0, t])$. Then,

$$\gamma \in \mathcal{H}(\nu, \nu; 0, t).$$

Proof. For all $\varphi \in C^2(\mathbb{T}^n \times [0, t])$, we have

$$\begin{aligned} &\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\varphi_t + q \cdot D\varphi - a(x)\Delta\varphi) d\gamma \\ &= \int_0^t \int_{\mathbb{T}^n \times \mathbb{R}^n} \varphi_t d\mu ds + \int_0^t \int_{\mathbb{T}^n \times \mathbb{R}^n} (q \cdot D\varphi - a(x)\Delta\varphi) d\mu ds \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n} \left(\int_0^t \varphi_t ds \right) d\mu = \int_{\mathbb{T}^n} \varphi(x, t) d\nu - \int_{\mathbb{T}^n} \varphi(x, 0) d\nu, \end{aligned}$$

which implies the conclusion. \square

Next, we show how to concatenate holonomic measures.

Lemma 3.2. *Let $\nu_1, \nu_2, \nu_3 \in \mathcal{P}(\mathbb{T}^n)$, and $a, b > 0$ be such that $\mathcal{H}(\nu_1, \nu_2; 0, a) \neq \emptyset$, $\mathcal{H}(\nu_2, \nu_3; 0, b) \neq \emptyset$. For any $\gamma_1 \in \mathcal{H}(\nu_1, \nu_2; 0, a)$ and $\gamma_2 \in \mathcal{H}(\nu_2, \nu_3; 0, b)$, set*

$$\gamma(x, q, s) := \begin{cases} \gamma_1(x, q, s) & s \in [0, a], \\ \gamma_2(x, q, s - a) & s \in [a, a + b]. \end{cases}$$

Then, $\gamma \in \mathcal{H}(\nu_1, \nu_3; 0, a + b)$.

Proof. By the definition of $\mathcal{H}(\nu_1, \nu_2; 0, a)$ and $\mathcal{H}(\nu_2, \nu_3; 0, b)$, for any test function $\varphi \in C^2(\mathbb{T}^n \times [0, +\infty))$, we have

$$\begin{aligned} \int_{\mathbb{T}^n} \varphi(\cdot, a) d\nu_2 - \int_{\mathbb{T}^n} \varphi(\cdot, 0) d\nu_1 &= \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, a]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma_1, \\ \int_{\mathbb{T}^n} \varphi(\cdot, b) d\nu_3 - \int_{\mathbb{T}^n} \varphi(\cdot, 0) d\nu_2 &= \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, b]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma_2. \end{aligned}$$

Thus,

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, a+b]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, a]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma + \int_{\mathbb{T}^n \times \mathbb{R}^n \times [a, a+b]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, a]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma_1(x, q, s) \\ & \quad + \int_{\mathbb{T}^n \times \mathbb{R}^n \times [a, a+b]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma_2(x, q, s - a) \\ &= \int_{\mathbb{T}^n} \varphi(\cdot, a) d\nu_2 - \int_{\mathbb{T}^n} \varphi(\cdot, 0) d\nu_1 \\ & \quad + \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, b]} (\varphi_t(x, s + a) - a(x)\Delta\varphi(x, s + a) + q \cdot D\varphi(x, s + a)) d\gamma_2(x, q, s) \\ &= \int_{\mathbb{T}^n} \varphi(\cdot, a) d\nu_2 - \int_{\mathbb{T}^n} \varphi(\cdot, 0) d\nu_1 + \int_{\mathbb{T}^n} \varphi(\cdot, a + b) d\nu_3 - \int_{\mathbb{T}^n} \varphi(\cdot, a) d\nu_2 \\ &= \int_{\mathbb{T}^n} \varphi(\cdot, a + b) d\nu_3 - \int_{\mathbb{T}^n} \varphi(\cdot, 0) d\nu_1, \end{aligned}$$

which implies

$$\gamma \in \mathcal{H}(\nu_1, \nu_3; 0, a + b). \quad \square$$

4. PEIERLS BARRIER AND THE GENERALIZED AUBRY SET

Suppose that Assumptions 1-3 hold. Let u solve Problem 1. As shown in [7], there exists a large-time limit, u_∞ , given by (1.4). This limit function solves Problem 2. As explained in the Introduction, by the results in [35], to characterize a solution of Problem 2, it is enough to determine the value

$$\int_{\mathbb{T}^n} u_\infty(z) d\nu(z) \quad \text{for all } \nu \in \mathcal{M}.$$

In this section, we provide this characterization and prove Theorem 1.2. Moreover, we examine two constructions from Aubry-Mather theory, the Peirls barrier and the projected Aubry set, and extend them in a way suitable for the study of degenerate diffusions.

Recall that, for $t > 0$, we set

$$h_t(\nu_0, \nu_1) = \begin{cases} \inf_{\gamma \in \mathcal{H}(\nu_0, \nu_1; 0, t)} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma & \text{if } \mathcal{H}(\nu_0, \nu_1; 0, t) \neq \emptyset, \\ +\infty & \text{if } \mathcal{H}(\nu_0, \nu_1; 0, t) = \emptyset, \end{cases}$$

for $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$. Also,

$$d(\nu_0, \nu_1) = \inf_{t > 0} h_t(\nu_0, \nu_1). \quad (4.1)$$

Lemma 4.1. *Consider the setting of Problem 1 and suppose that Assumptions 1-3 hold. Fix $t > 0$ and $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$. Then, the map*

$$\nu \mapsto h_t(\nu, \nu_1)$$

is convex.

Proof. Fix $t > 0$ and $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$. Take $\nu_0, \bar{\nu}_0 \in \mathcal{P}(\mathbb{T}^n)$. We must show that for $0 \leq \lambda \leq 1$,

$$h_t(\lambda\nu_0 + (1 - \lambda)\bar{\nu}_0, \nu_1) \leq \lambda h_t(\nu_0, \nu_1) + (1 - \lambda)h_t(\bar{\nu}_0, \nu_1).$$

If any of the terms in the right-hand side is $+\infty$, the result is trivial. Thus, we may assume that $h_t(\nu_0, \nu_1) < +\infty$ and $h_t(\bar{\nu}_0, \nu_1) < +\infty$. Accordingly, $\mathcal{H}(\nu_0, \nu_1; 0, t) \neq \emptyset$ and $\mathcal{H}(\bar{\nu}_0, \nu_1; 0, t) \neq \emptyset$. For $\gamma_1 \in \mathcal{H}(\nu_0, \nu_1; 0, t)$ and $\gamma_2 \in \mathcal{H}(\bar{\nu}_0, \nu_1; 0, t)$, let $\gamma := \lambda\gamma_1 + (1 - \lambda)\gamma_2$ for $\lambda \in [0, 1]$. We claim that

$$\gamma \in \mathcal{H}(\lambda\nu_0 + (1 - \lambda)\bar{\nu}_0, \nu_1; 0, t). \quad (4.2)$$

To establish the claim, we fix $\varphi \in C^2(\mathbb{T}^n \times [0, t])$. Then

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\varphi_t - a\Delta\varphi + q \cdot D\varphi) d(\lambda\gamma_1 + (1 - \lambda)\gamma_2) \\ &= \lambda \left(\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\varphi_t - a\Delta\varphi + q \cdot D\varphi) d\gamma_1 \right) \\ & \quad + (1 - \lambda) \left(\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\varphi_t - a\Delta\varphi + q \cdot D\varphi) d\gamma_2 \right) \\ &= \lambda \left(\int_{\mathbb{T}^n} \varphi(z, t) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, 0) d\nu_0(z) \right) \\ & \quad + (1 - \lambda) \left(\int_{\mathbb{T}^n} \varphi(z, t) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, 0) d\bar{\nu}_0(z) \right) \\ &= \int_{\mathbb{T}^n} \varphi(z, t) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, 0) d(\lambda\nu_0 + (1 - \lambda)\bar{\nu}_0)(z), \end{aligned}$$

thus (4.2) holds. Accordingly, we have

$$h_t(\lambda\nu_0 + (1 - \lambda)\bar{\nu}_0, \nu_1) \leq \lambda \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L d\gamma_1 + (1 - \lambda) \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L d\gamma_2.$$

Taking the infimum on $\gamma_1 \in \mathcal{H}(\nu_0, \nu_1; 0, t)$ and $\gamma_2 \in \mathcal{H}(\bar{\nu}_0, \nu_1; 0, t)$, we obtain

$$h_t(\lambda\nu_0 + (1-\lambda)\bar{\nu}_0, \nu_1) \leq \lambda h_t(\nu_0, \nu_1) + (1-\lambda)h_t(\bar{\nu}_0, \nu_1). \quad \square$$

Next, we define the *generalized Peierls barrier*, $h : \mathcal{P}(\mathbb{T}^n) \times \mathcal{P}(\mathbb{T}^n) \rightarrow \mathbb{R} \cup \{+\infty\}$, by

$$h(\nu_0, \nu_1) = \liminf_{t \rightarrow \infty} h_t(\nu_0, \nu_1) \quad \text{for } \nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n).$$

In weak KAM theory, the Peierls barrier is a function $h : \mathbb{T}^n \times \mathbb{T}^n \rightarrow \mathbb{R}$. However, in our context, points in \mathbb{T}^n are replaced by probability measures in $\mathcal{P}(\mathbb{T}^n)$; thus, h becomes a function on $\mathcal{P}(\mathbb{T}^n) \times \mathcal{P}(\mathbb{T}^n)$. We believe that this is a natural framework for second-order equations.

Now, we introduce an auxiliary function, $m : \mathcal{P}(\mathbb{T}^n) \times \mathcal{P}(\mathbb{T}^n) \rightarrow \mathbb{R}$, given by

$$m(\nu_0, \nu_1) := \sup_{w \in \mathcal{S}} \left[\int_{\mathbb{T}^n} w d\nu_1(x) - \int_{\mathbb{T}^n} w d\nu_0(x) \right],$$

where \mathcal{S} is the set of all Lipschitz viscosity solutions to (1.5) given by (1.11). The classical result for the existence of a solution of the ergodic problem implies that $\mathcal{S} \neq \emptyset$ (see [33], for instance). Besides, by [2, Theorem 3.1], there exists $C > 0$ depending only on H, a such that $\|Dw\|_{L^\infty(\mathbb{T}^n)} \leq C$. Thus, $m(\nu_0, \nu_1)$ is finite for all $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$.

Lemma 4.2. *Consider the setting of Problem 2 and suppose that Assumptions 1-3 hold. Then,*

$$m(\nu_0, \nu_1) \leq d(\nu_0, \nu_1) \quad \text{for all } \nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n). \quad (4.3)$$

Proof. If $d(\nu_0, \nu_1) = \infty$, then (4.3) is trivial. Therefore, we only need to address the case where $d(\nu_0, \nu_1) < \infty$. In this case, there exists $t > 0$ such that $\mathcal{H}(\nu_0, \nu_1; 0, t) \neq \emptyset$. Let w be any Lipschitz viscosity solution to (1.5), and let $w^\alpha \in C^\infty(\mathbb{T}^n)$ be the approximation of w given by (A.1) for $\alpha > 0$; that is,

$$w^\alpha(x) = \int_{\mathbb{T}^n} \theta^\alpha(y) w(x-y) dy \quad \text{for } x \in \mathbb{T}^n,$$

since w is time-independent here. Then, $w^\alpha \in C^\infty(\mathbb{T}^n)$ satisfies

$$H(x, Dw^\alpha) - a(x)\Delta w^\alpha \leq C\alpha^{1/2} \quad \text{in } \mathbb{T}^n.$$

For any $\gamma \in \mathcal{H}(\nu_0, \nu_1; 0, t)$, we have

$$\begin{aligned} & \int_{\mathbb{T}^n} w^\alpha(x) d\nu_1(x) - \int_{\mathbb{T}^n} w^\alpha(x) d\nu_0(x) \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (q \cdot Dw^\alpha(x) - a(x)\Delta w^\alpha(x)) d\gamma(x, q, s) \\ &\leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (L(x, q) + H(x, Dw^\alpha(x)) - a(x)\Delta w^\alpha(x)) d\gamma(x, q, s) \\ &\leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma(x, q, s) + C\alpha^{1/2}t. \end{aligned}$$

Send $\alpha \rightarrow 0$ to yield

$$\int_{\mathbb{T}^n} w(x) d\nu_1(x) - \int_{\mathbb{T}^n} w(x) d\nu_0(x) \leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0,t]} L(x, q) d\gamma(x, q, s).$$

Thus, taking the supremum with respect to $w \in \mathcal{S}$, and then, the infimum with respect to γ gives the claim in the statement. \square

Corollary 4.3. *Consider the setting of Problem 2 and suppose that Assumptions 1-3 hold. Then, for all $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$,*

$$m(\nu_0, \nu_1) \leq d(\nu_0, \nu_1) \leq h(\nu_0, \nu_1). \quad (4.4)$$

Moreover, $h(\nu, \nu) \geq 0$ for all $\nu \in \mathcal{P}(\mathbb{T}^n)$.

Proof. Because of the definition of $d(\nu_0, \nu_1)$ in (1.9) (or (4.1)), it is clear that $d(\nu_0, \nu_1) \leq h(\nu_0, \nu_1)$. Lemma 4.2 gives $m(\nu_0, \nu_1) \leq d(\nu_0, \nu_1)$. Hence, we have (4.4). Finally, because $m(\nu, \nu) = 0$, we conclude that $h(\nu, \nu) \geq 0$. \square

The *generalized projected Aubry set* is the set

$$\mathcal{A} := \{\nu \in \mathcal{P}(\mathbb{T}^n) : h(\nu, \nu) = 0\}.$$

As in standard Aubry-Mather theory, the generalized projected Aubry set contains the generalized projected Mather set, as we establish next.

Proposition 4.4. *Consider the setting of Problem 2 and suppose that Assumptions 1-3 hold. Then,*

$$\mathcal{M} \subset \mathcal{A}.$$

In particular,

$$\mathcal{A} \neq \emptyset.$$

Proof. By Corollary 4.3,

$$0 \leq h(\nu, \nu) \quad \text{for all } \nu \in \mathcal{P}(\mathbb{T}^n).$$

Let μ be any generalized Mather measure, and set $\nu := \text{proj}_{\mathbb{T}^n} \mu$. Let $d\tilde{\gamma}(x, q, s) := d\mu(x, q) ds$ for all $s \in [0, t]$. By Lemma 3.1, $\tilde{\gamma} \in \mathcal{H}(\nu, \nu; 0, t)$ for all $t > 0$. Thus,

$$\begin{aligned} h(\nu, \nu) &= \liminf_{t \rightarrow \infty} h_t(\nu, \nu) = \liminf_{t \rightarrow \infty} \inf_{\gamma \in \mathcal{H}(\nu, \nu; 0, t)} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma(x, q, s) \\ &\leq \liminf_{t \rightarrow \infty} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\tilde{\gamma}(x, q, s) = \liminf_{t \rightarrow \infty} \int_0^t \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\mu ds \\ &= 0, \end{aligned}$$

in light of (1.7), which finishes the proof. Because of the results in [34, 35] (see also the proof of Proposition B.2), we have $\mathcal{M} \neq \emptyset$. Thus, $\mathcal{A} \neq \emptyset$. \square

Now, we prove a stronger version of Theorem 1.2; that is, (1.8) holds for all measures in \mathcal{A} .

Theorem 4.5. *Suppose that Assumptions 1-3 hold. Let u solve Problem 1 and u_∞ be the large time limit of u in (1.4). Then,*

$$\int_{\mathbb{T}^n} u_\infty(z) d\nu(z) = \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z) \right]$$

for all $\nu \in \mathcal{A}$.

Proof. We rewrite the representation formula in Theorem 1.1 as

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) = \inf_{\substack{\gamma \in \mathcal{H}(\nu_0, \nu; 0, t) \\ \nu_0 \in \mathcal{P}(\mathbb{T}^n)}} \left[\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma(x, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z) \right].$$

Accordingly, we have

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) \geq \inf_{\substack{\gamma \in \mathcal{H}(\nu_0, \nu; 0, t) \\ \nu_0 \in \mathcal{P}(\mathbb{T}^n)}} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0 d\nu_0 \right] = \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0 d\nu_0 \right]$$

for all $\nu \in \mathcal{P}(\mathbb{T}^n)$. Thus, by letting $t \rightarrow \infty$, we get

$$\int_{\mathbb{T}^n} u_\infty(z) d\nu(z) \geq \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z) \right].$$

Let $\nu \in \mathcal{A}$. Fix $\varepsilon > 0$, and take $\nu_0 \in \mathcal{P}(\mathbb{T}^n)$, $T > 0$, and $\gamma_1 \in \mathcal{H}(\nu_0, \nu; 0, T)$ such that

$$\begin{aligned} & \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z) \right] + \varepsilon \\ & \geq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, T]} L(z, q) d\gamma_1(z, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z) \geq \int_{\mathbb{T}^n} u(z, T) d\nu(z). \end{aligned}$$

Because $\nu \in \mathcal{A}$, there exists a sequence $\{t_k\}_{k \in \mathbb{N}}$ such that $t_k \rightarrow \infty$ and

$$h_{t_k}(\nu, \nu) \rightarrow \liminf_{t \rightarrow \infty} h_t(\nu, \nu) = h(\nu, \nu) = 0 \quad \text{as } k \rightarrow \infty.$$

Accordingly, there exists $k_0 \in \mathbb{N}$ such that for all $k \geq k_0$,

$$\varepsilon \geq h_{t_k}(\nu, \nu).$$

Moreover, there exists a corresponding measure, $\gamma_k \in \mathcal{H}(\nu, \nu; 0, t_k)$, such that

$$2\varepsilon \geq h_{t_k}(\nu, \nu) + \varepsilon \geq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t_k]} L(x, q) d\gamma_k(x, q, s).$$

We note that $u(\cdot, \cdot + T)$ solves (1.1) with initial data $u(\cdot, T)$. By the representation formula in Theorem 1.1, for $k \geq k_0$,

$$\begin{aligned} & \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0 d\nu_0 \right] + \varepsilon \geq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, T]} L(x, q) d\gamma_1(x, q, s) + \int_{\mathbb{T}^n} u_0 d\nu_0 \\ & \geq \int_{\mathbb{T}^n} u(z, T) d\nu(z) + \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t_k]} L(x, q) d\gamma_k(x, q, s) - 2\varepsilon \\ & \geq \int_{\mathbb{T}^n} u(x, T + t_k) d\nu - 2\varepsilon. \end{aligned}$$

Letting $k \rightarrow \infty$ and then $\varepsilon \rightarrow 0$ gives

$$\int_{\mathbb{T}^n} u_\infty d\nu \leq \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0 d\nu_0 \right],$$

which finishes the proof. \square

Theorem 1.2 follows from the preceding result, as we point out next.

Proof of Theorem 1.2. According to Proposition 4.4, $\mathcal{M} \subset \mathcal{A}$. Thus, the statement follows directly from Theorem 4.5. \square

Next, we record a convexity property for the generalized Mañé critical potential. We notice that we do not know if the maps $\nu \mapsto d(\nu, \bar{\nu})$ and $\nu \mapsto d(\bar{\nu}, \nu)$ are convex for all $\bar{\nu} \in \mathcal{P}(\mathbb{T}^n)$. We can only prove that the maps $\nu \mapsto d(\nu, \bar{\nu})$ and $\nu \mapsto d(\bar{\nu}, \nu)$ are convex for all $\bar{\nu} \in \mathcal{M}$.

Lemma 4.6. *Consider the setting of Problem 1 and suppose that Assumptions 1-3 hold. Then, for any projected Mather measure $\bar{\nu} \in \mathcal{M}$, the maps*

$$\nu \mapsto d(\nu, \bar{\nu}), \quad \nu \mapsto d(\bar{\nu}, \nu)$$

are convex for $\nu \in \mathcal{P}(\mathbb{T}^n)$.

Proof. We only prove that the map $\nu \mapsto d(\bar{\nu}, \nu)$ is convex. Since the proof is similar for the other map, we omit it here.

Take $\nu_1, \nu_2 \in \mathcal{P}(\mathbb{T}^n)$. We claim that for $0 \leq \lambda \leq 1$,

$$d(\bar{\nu}, \lambda\nu_1 + (1-\lambda)\nu_2) \leq \lambda d(\bar{\nu}, \nu_1) + (1-\lambda)d(\bar{\nu}, \nu_2).$$

It is enough to consider the case where both terms in the right-hand side of the preceding expression are finite. Hence, let $t, s > 0$ with $t < s$ so that $\mathcal{H}(\bar{\nu}, \nu_1; 0, t) \neq \emptyset$ and $\mathcal{H}(\bar{\nu}, \nu_2; 0, s) \neq \emptyset$. Let $\gamma_1 \in \mathcal{H}(\bar{\nu}, \nu_1; 0, t)$ and $\gamma_2 \in \mathcal{H}(\bar{\nu}, \nu_2; 0, s)$.

Since $\bar{\nu}$ is a projected Mather measure, we can extend γ_1 to $\bar{\gamma}_1 \in \mathcal{H}(\bar{\nu}, \nu_1; 0, s)$ by a similar argument to the proof of Lemma 3.1. Indeed, let μ be a Mather measure such that $\bar{\nu} = \text{proj}_{\mathbb{T}^n} \mu$. Let

$$d\bar{\gamma}_1(x, q, r) := \begin{cases} d\mu(x, q)dr & \text{for } r \in [0, s-t], \\ d\gamma_1(x, q, r - (s-t)) & \text{for } r \in [s-t, s]. \end{cases}$$

By abuse of notation, we identify $\bar{\gamma}_1$ as γ_1 . Let $\gamma := \lambda\gamma_1 + (1-\lambda)\gamma_2$ for $\lambda \in [0, 1]$. As we show next,

$$\gamma \in \mathcal{H}(\bar{\nu}, \lambda\nu_1 + (1-\lambda)\nu_2; 0, s).$$

To verify the preceding claim, fix $\varphi \in C^2(\mathbb{T}^n \times [0, s])$,

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d(\lambda\gamma_1 + (1-\lambda)\gamma_2) \\ &= \lambda \left(\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma_1 \right) \\ & \quad + (1-\lambda) \left(\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s]} (\varphi_t - a(x)\Delta\varphi + q \cdot D\varphi) d\gamma_2 \right) \end{aligned}$$

$$\begin{aligned}
&= \lambda \left(\int_{\mathbb{T}^n} \varphi(z, s) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, 0) d\bar{\nu}(z) \right) \\
&\quad + (1 - \lambda) \left(\int_{\mathbb{T}^n} \varphi(z, s) d\nu_2(z) - \int_{\mathbb{T}^n} \varphi(z, 0) d\bar{\nu}(z) \right) \\
&= \int_{\mathbb{T}^n} \varphi(z, s) d(\lambda\nu_1 + (1 - \lambda)\nu_2)(z) - \int_{\mathbb{T}^n} \varphi(z, 0) d\bar{\nu}(z).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
d(\bar{\nu}, \lambda\nu_1 + (1 - \lambda)\nu_2) &\leq \lambda \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s]} L d\gamma_1 + (1 - \lambda) \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s]} L d\gamma_2 \\
&= \lambda \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s-t]} L d\mu + \lambda \int_{\mathbb{T}^n \times \mathbb{R}^n \times [s-t, s]} L d\gamma_1 + (1 - \lambda) \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s]} L d\gamma_2 \\
&= \lambda \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L d\gamma_1 + (1 - \lambda) \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, s]} L d\gamma_2
\end{aligned}$$

because of (1.7). Taking the infima on $\gamma_1 \in \mathcal{H}(\bar{\nu}, \nu_1; 0, t)$ for $t > 0$ and $\gamma_2 \in \mathcal{H}(\bar{\nu}, \nu_2; 0, s)$ for $s > 0$, respectively, we obtain

$$d(\bar{\nu}, \lambda\nu_1 + (1 - \lambda)\nu_2) \leq \lambda d(\bar{\nu}, \nu_1) + (1 - \lambda)d(\bar{\nu}, \nu_2). \quad \square$$

We end this section by establishing the improved convergence result in Proposition 1.3. This result, which is of independent interest, is crucial for the proof of the duality formula for the Mañé critical potential.

Proof of Proposition 1.3. We establish the claim by contradiction. Accordingly, we assume that there exists a sequence $\{u_0^k\} \subset C^2(\mathbb{T}^n)$ with $\min_{\mathbb{T}^n} u_0^k = 0$ and a positive number, $\varepsilon > 0$, such that for all $k \in \mathbb{N}$ there exists $t_k > k$ with

$$\|u^k(\cdot, t_k) - u_\infty^k\|_{L^\infty(\mathbb{T}^n)} > \varepsilon.$$

Because of Assumption 4, extracting a subsequence if necessary, we may assume that $u^k(\cdot, 1) \rightarrow u(\cdot, 1)$ uniformly on \mathbb{T}^n as $k \rightarrow \infty$, for some function $u(\cdot, 1) \in C(\mathbb{T}^n)$. Let u_∞ be the large time limit of the solution to (1.1) with initial data $u(\cdot, 1)$.

Because $u^k(\cdot, 1) \rightarrow u(\cdot, 1)$ uniformly on \mathbb{T}^n , we have $u_\infty^k \rightarrow u_\infty$ uniformly on \mathbb{T}^n by the maximum principle (see [33]). By choosing k large enough, we have $\|u^k(\cdot, t) - u(\cdot, t)\|_{L^\infty(\mathbb{T}^n)} \leq \frac{\varepsilon}{3}$ for all $t \geq 1$, and $\|u_\infty^k - u_\infty\|_{L^\infty(\mathbb{T}^n)} \leq \frac{\varepsilon}{3}$. Then,

$$\begin{aligned}
&\|u(\cdot, t_k) - u_\infty\|_{L^\infty(\mathbb{T}^n)} \\
&\geq -\|u(\cdot, t_k) - u^k(\cdot, t_k)\|_{L^\infty(\mathbb{T}^n)} + \|u^k(\cdot, t_k) - u_\infty^k\|_{L^\infty(\mathbb{T}^n)} - \|u_\infty^k - u_\infty\|_{L^\infty(\mathbb{T}^n)} \\
&\geq \frac{\varepsilon}{3},
\end{aligned}$$

which contradicts the uniform convergence of $u(\cdot, t)$ to u_∞ in \mathbb{T}^n as $t \rightarrow \infty$. \square

5. DUALITY FORMULAS

In this section, we use the Legendre-Fenchel-Rockafellar theorem to give dual representations for h_t and d .

5.1. Duality. To use the Legendre-Fenchel-Rockafellar theorem, we need to recall the definition of Legendre transform on a locally convex topological vector space. Let E be a locally convex topological vector space with dual E' and duality pairing (\cdot, \cdot) . Consider a convex function, $f : E \rightarrow (-\infty, +\infty]$. The *Legendre-Fenchel transform*, $f^* : E' \rightarrow [-\infty, +\infty]$, of f is

$$f^*(y) = \sup_{x \in E} ((x, y) - f(x)) \quad \text{for } y \in E'.$$

Let $\Omega = \mathbb{T}^n \times \mathbb{R}^n \times [0, t]$ for $t > 0$ fixed. Let m be as in Assumption 3 and choose ζ with $1 < \zeta < m'$. Let \mathcal{U} be the set

$$\mathcal{U} = \left\{ \mu \in \mathcal{R}(\Omega) : \int_{\Omega} (1 + |q|^{\zeta}) d|\mu| < \infty \right\}.$$

The set \mathcal{U} is the dual of

$$C_{\zeta}(\Omega) := \left\{ \phi \in C(\Omega) : \|\phi\|_{\zeta} := \sup_{\Omega} \left| \frac{\phi(x, q, s)}{1 + |q|^{\zeta}} \right| < \infty, \right. \\ \left. \lim_{|q| \rightarrow \infty} \frac{\phi(x, q, s)}{1 + |q|^{\zeta}} = 0 \quad \text{uniformly for } (x, s) \in \mathbb{T}^n \times [0, t] \right\}.$$

Let

$$\mathcal{U}_1 = \left\{ \mu \in \mathcal{U} : \mu \geq 0, \int_{\Omega} d\mu = t \right\}.$$

Next, for $q \in \mathbb{R}^n$, define the following linear operator on $C^2(\mathbb{T}^n \times [0, t])$

$$A^q \varphi = \varphi_t - a(x) \Delta \varphi + q \cdot D \varphi.$$

Note that A^q can be regarded as a bounded linear mapping from $C^2(\mathbb{T}^n \times [0, t])$ to $C_{\zeta}(\Omega)$. Let $B : C(\mathbb{T}^n \times [0, t]) \rightarrow \mathbb{R}$ be a bounded linear operator. Define

$$\mathcal{U}_2 = \text{cl} \left\{ \mu \in \mathcal{U} : \int_{\Omega} A^q \varphi d\mu(x, q, s) = B\varphi, \forall \varphi \in C^2(\mathbb{T}^n \times [0, t]) \right\},$$

where the closure, cl, is taken with respect to the weak topology in \mathcal{U} . Moreover, we assume that there exists $\bar{\mu} \in \mathcal{U}_2$; that is, \mathcal{U}_2 is non-empty.

For $\phi \in C_{\zeta}(\Omega)$, let

$$f(\phi) = t \sup_{(x, q, s) \in \Omega} (\phi(x, q, s) - L(x, q)). \quad (5.1)$$

Because f is the supremum of convex functions, it is also a convex function. Moreover, because

$$\lim_{|q| \rightarrow \infty} \frac{L(x, q)}{1 + |q|^{\zeta}} = +\infty$$

for all $x \in \mathbb{T}^n$, f is also continuous with respect to the uniform convergence in $C_{\zeta}(\Omega)$.

Let

$$\mathcal{C} = \text{cl} \left\{ \phi \in C_{\zeta}(\Omega) : \phi(x, q, s) = A^q \varphi(x, s) \text{ for some } \varphi \in C^2(\mathbb{T}^n \times [0, t]) \right\},$$

where cl denotes the closure in C_{ζ} . Because A^q is a bounded linear mapping from $C^2(\mathbb{T}^n \times [0, t])$ to $C_{\zeta}(\Omega)$, \mathcal{C} is a closed convex set.

Fix $\bar{\mu} \in \mathcal{U}_2$ and let $g : C(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$ be

$$g(\phi) = \begin{cases} -\int_{\Omega} \phi d\bar{\mu} & \text{if } \phi \in \mathcal{C}, \\ +\infty & \text{otherwise.} \end{cases} \quad (5.2)$$

Because \mathcal{C} is a closed convex set, g is convex and lower semicontinuous. Furthermore, for $\phi = A^q\varphi$, we have $\int_{\Omega} \phi d\bar{\mu} = B\varphi$, according to the definition of \mathcal{U}_2 .

Next, we address the Legendre transform of f . We recall that

$$f^*(\mu) = \sup_{\phi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \phi d\mu - f(\phi) \right).$$

We first give two elementary results on f^* .

Lemma 5.1. *Suppose that Assumptions 1-3 hold. Let f be as in (5.1). Let $\mu \in \mathcal{U}$. If $\mu \not\geq 0$ then $f^*(\mu) = +\infty$.*

Proof. If $\mu \not\geq 0$, there exists a sequence of functions $\phi_n \in C_{\zeta}(\Omega)$ with $\phi_n \leq 0$ such that

$$\int_{\Omega} \phi_n d\mu \rightarrow +\infty.$$

Therefore, because

$$f(\phi_n) = t \sup_{(x,q,s) \in \Omega} [\phi_n(x,q,s) - L(x,q)] \leq -t \min_{(x,q)} L(x,q) \leq C,$$

we have

$$f^*(\mu) \geq \lim_{n \rightarrow \infty} \int_{\Omega} \phi_n d\mu - f(\phi_n) \geq \lim_{n \rightarrow \infty} \int_{\Omega} \phi_n d\mu - C = +\infty. \quad \square$$

Next, we give a lower bound for $f^*(\mu)$ for non-negative μ .

Lemma 5.2. *Suppose that Assumptions 1-3 hold. Let f be as in (5.1). Let $\mu \in \mathcal{U}$. If $\mu \geq 0$ then*

$$f^*(\mu) \geq \int_{\Omega} L d\mu + \sup_{\psi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \psi d\mu - t \sup_{\Omega} \psi \right).$$

Proof. Let L_n be an increasing sequence of functions in $C_{\zeta}(\Omega)$ such that $L_n \rightarrow L$ pointwise. Any ϕ in $C_{\zeta}(\Omega)$ can be written as $\phi = \psi + L_n$, for some ψ in $C_{\zeta}(\Omega)$. Therefore,

$$\sup_{\phi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \phi d\mu - f(\phi) \right) = \sup_{\psi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} L_n d\mu + \int_{\Omega} \psi d\mu - t \sup_{\Omega} (L_n + \psi - L) \right).$$

Because $L_n \leq L$, we have

$$\begin{aligned} \sup_{\phi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \phi d\mu - f(\phi) \right) &\geq \sup_{\psi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} L_n d\mu + \int_{\Omega} \psi d\mu - t \sup_{\Omega} \psi \right) \\ &= \int_{\Omega} L_n d\mu + \sup_{\psi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \psi d\mu - t \sup_{\Omega} \psi \right) \end{aligned}$$

By the monotone convergence theorem,

$$\int_{\Omega} L_n d\mu \rightarrow \int_{\Omega} L d\mu \quad \text{as } n \rightarrow \infty.$$

Thus,

$$f^*(\mu) = \sup_{\phi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \phi d\mu - f(\phi) \right) \geq \int_{\Omega} L d\mu + \sup_{\psi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \psi d\mu - t \sup \psi \right),$$

as required. \square

Using the two preceding lemmas, we now compute the Legendre transform of f .

Proposition 5.3. *Suppose that Assumptions 1-3 hold. Let f be as in (5.1). Let $\mu \in \mathcal{U}$. Then,*

$$f^*(\mu) = \begin{cases} \int_{\Omega} L d\mu & \text{if } \mu \in \mathcal{U}_1, \\ +\infty & \text{otherwise.} \end{cases} \quad (5.3)$$

Proof. By Lemma 5.1, if μ is not non-negative then $f^*(\mu) = \infty$. Moreover, by Lemma 5.2, if $\int_{\Omega} L d\mu = +\infty$ then $f^*(\mu) = +\infty$. Furthermore, if $\int_{\Omega} d\mu \neq t$ then

$$\sup_{\psi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \psi d\mu - t \sup_{\Omega} \psi \right) \geq \sup_{\alpha \in \mathbb{R}} \alpha \left(\int_{\Omega} d\mu - t \right) = +\infty,$$

by choosing $\psi \equiv \alpha$ for $\alpha \in \mathbb{R}$. Therefore, in this case, using also Lemma 5.2, we have $f^*(\mu) = +\infty$.

When $\mu \geq 0$ and $\int_{\Omega} d\mu = t$, Lemma 5.2 implies

$$f^*(\mu) \geq \int_{\Omega} L d\mu,$$

by choosing $\psi = 0$.

Additionally, for any $\phi \in C_{\zeta}(\Omega)$,

$$\int_{\Omega} (\phi - L) d\mu \leq t \sup_{\Omega} (\phi - L) = f(\phi).$$

Therefore,

$$f^*(\mu) = \sup_{\phi \in C_{\zeta}(\Omega)} \left(\int_{\Omega} \phi d\mu - f(\phi) \right) \leq \int_{\Omega} L d\mu.$$

Accordingly, we obtain (5.3). \square

Proposition 5.4. *Suppose that $\mathcal{U}_2 \neq \emptyset$. Let g be as in (5.2). Let $\mu \in \mathcal{U}$. Then*

$$g^*(\mu) = \begin{cases} 0 & \text{if } -\mu \in \mathcal{U}_2 \\ +\infty & \text{otherwise.} \end{cases}$$

Proof. Let $\bar{\mu} \in \mathcal{U}_2$; that is, $\bar{\mu}$ satisfies

$$\int_{\Omega} A^q \varphi d\bar{\mu} = B\varphi,$$

for all $\varphi \in C^2(\mathbb{T}^n \times [0, t])$. Given a measure μ with $-\mu \in \mathcal{U}_2$, decompose $\mu = -\bar{\mu} + \hat{\mu}$ with $\hat{\mu} \in \mathcal{R}(\Omega)$ satisfying

$$\int_{\Omega} A^q \varphi d\hat{\mu} = 0 \quad \text{for all } \varphi \in C^2(\mathbb{T}^n \times [0, t]). \quad (5.4)$$

By continuity, we obtain

$$\int_{\Omega} \phi d\hat{\mu} = 0 \quad \text{for all } \phi \in \mathcal{C};$$

that is,

$$\int_{\Omega} \phi d\mu + \int_{\Omega} \phi d\bar{\mu} = 0$$

for all $\phi \in \mathcal{C}$. Furthermore, if $-\mu \notin \mathcal{U}_2$, then there exists $\varphi \in C^2(\mathbb{T}^n \times [0, t])$ such that $\int_{\Omega} A^q \varphi d(-\mu) \neq B\varphi$, and by (5.4),

$$\int_{\Omega} A^q \varphi d(\mu + \bar{\mu}) \neq 0$$

which implies

$$\int_{\Omega} \hat{\phi} d(\mu + \bar{\mu}) \neq 0 \quad \text{for some } \hat{\phi} \in \mathcal{C}.$$

Thus

$$g^*(\mu) = \sup_{\phi \in \mathcal{C}} \left(\int_{\Omega} \phi d\mu + \int_{\Omega} \phi d\bar{\mu} \right) = \begin{cases} 0 & \text{if } -\mu \in \mathcal{U}_2 \\ +\infty & \text{otherwise.} \end{cases}$$

□

Theorem 5.5. *Suppose that Assumptions 1-3 hold. Suppose that $\mathcal{U}_2 \neq \emptyset$. Then,*

$$\max_{\mu \in \mathcal{U}_1 \cap \mathcal{U}_2} \left(- \int_{\Omega} L d\mu \right) = \inf_{\varphi \in C^2(\mathbb{T}^n \times [0, t])} \left[-B\varphi + t \sup_{(x, s) \in \mathbb{T}^n \times [0, t]} (\varphi_t - a(x)\Delta\varphi + H(x, D\varphi)) \right].$$

Proof. Let f and g be as in (5.1) and (5.2). We have that $f, g : C_{\zeta}(\Omega) \rightarrow (-\infty, +\infty]$ are convex functions. The set $g < +\infty$ is non-empty, and, in this set, f is a continuous function. Therefore, by the Fenchel-Legendre-Rockafellar theorem [38],

$$\max_{\mu \in \mathcal{U}} [-f^*(\mu) - g^*(-\mu)] = \inf_{\phi \in C_{\zeta}(\Omega)} [f(\phi) + g(\phi)].$$

By Propositions 5.3 and 5.4, we have

$$\max_{\mu \in \mathcal{U}} [-f^*(\mu) - g^*(-\mu)] = \max_{\mu \in \mathcal{U}_1 \cap \mathcal{U}_2} \left(- \int_{\Omega} L d\mu \right).$$

Therefore,

$$\begin{aligned} \inf_{\phi \in C_{\zeta}(\Omega)} [f(\phi) + g(\phi)] &= \inf_{\phi \in \mathcal{C}} \left[t \sup_{\Omega} (\phi - L) - \int_{\Omega} \phi d\bar{\mu} \right] \\ &= \inf_{\varphi \in C^2(\mathbb{T}^2 \times [0, t])} \left[-B\varphi + t \sup_{\Omega} (A^q \varphi - L) \right] \\ &= \inf_{\varphi \in C^2(\mathbb{T}^2 \times [0, t])} \left[-B\varphi + t \sup_{\mathbb{T}^n \times [0, t]} \left(\varphi_t - a(x)\Delta\varphi + \sup_{q \in \mathbb{R}^n} (q \cdot D\varphi - L(x, q)) \right) \right] \end{aligned}$$

$$= \inf_{\varphi \in C^2(\mathbb{T}^2 \times [0, t])} \left[-B\varphi + t \sup_{(x, s) \in \mathbb{T}^n \times [0, t]} (\varphi_t - a(x)\Delta\varphi + H(x, D\varphi)) \right],$$

which implies the conclusion. \square

5.2. Convex duality for h_t . We begin by applying Theorem 5.5 to obtain a preliminary dual formula for h_t .

Proposition 5.6. *Consider the setting of Problem 1. Suppose that Assumptions 1-3 hold. Fix $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$, $t > 0$, and assume that $\mathcal{H}(\nu_0, \nu_1; 0, t) \neq \emptyset$. Then,*

$$h_t(\nu_0, \nu_1) = - \inf_{\varphi \in C^2(\mathbb{T}^n \times [0, t])} \sup_{(x, s) \in \mathbb{T}^n \times [0, t]} \left[\int_{\mathbb{T}^n} \varphi(x, 0) d\nu_0 - \int_{\mathbb{T}^n} \varphi(x, t) d\nu_1 + t (\varphi_t(x, s) - a(x)\Delta\varphi(x, s) + H(x, D\varphi(x, s))) \right].$$

Proof. Given ν_0 and ν_1 , we set

$$B\varphi = \int_{\mathbb{T}^n} \varphi(x, t) d\nu_1 - \int_{\mathbb{T}^n} \varphi(x, 0) d\nu_0.$$

Because $\mathcal{H}(\nu_0, \nu_1; 0, t) \neq \emptyset$, we have $\mathcal{U}_2 \neq \emptyset$. Then

$$h_t(\nu_0, \nu_1) = \inf_{\mu \in \mathcal{U}_1 \cap \mathcal{U}_2} \int_{\Omega} L d\mu,$$

where \mathcal{U}_1 and \mathcal{U}_2 are defined in Section 5.1. The identity in the claim follows from Theorem 5.5. \square

As stated in Theorem 1.4, the preceding result can be further refined; it is enough to use as test functions viscosity solutions to the Hamilton–Jacobi–Bellman equation in (1.1). This theorem is proved as follows.

Proof of Theorem 1.4. We denote by $m_t(\nu_0, \nu_1)$ the right-hand side of (1.10).

Let w be a viscosity solution to (1.1) and consider the approximation w^α defined in (A.1). Then,

$$\begin{aligned} & \inf_{\varphi \in C^2(\mathbb{T}^n \times [0, t])} \sup_{(x, s)} \left[\int_{\mathbb{T}^n} \varphi(x, 0) d\nu_0 - \int_{\mathbb{T}^n} \varphi(x, s) d\nu_1 \right. \\ & \quad \left. + t (\varphi_t(x, s) - a(x)\Delta\varphi(x, s) + H(x, D\varphi(x, s))) \right] \\ & \leq \sup_{(x, s)} \left[\int_{\mathbb{T}^n} w^\alpha(x, 0) d\nu_0 - \int_{\mathbb{T}^n} w^\alpha(x, t) d\nu_1 + t (w_t^\alpha - a(x)\Delta w^\alpha + H(x, Dw^\alpha)) \right]. \\ & \leq \left[\int_{\mathbb{T}^n} w^\alpha(x, 0) d\nu_0 - \int_{\mathbb{T}^n} w^\alpha(x, t) d\nu_1 \right] + C\alpha^{1/2}t. \end{aligned}$$

Thus, letting $\alpha \rightarrow 0$, we conclude that

$$\begin{aligned} & \inf_{\varphi \in C^2(\mathbb{T}^n \times [0, t])} \sup_{(x, s)} \left[\int_{\mathbb{T}^n} \varphi(x, 0) d\nu_0 - \int_{\mathbb{T}^n} \varphi(x, s) d\nu_1 \right. \\ & \quad \left. + t (\varphi_t(x, s) - a(x)\Delta\varphi(x, s) + H(x, D\varphi(x, s))) \right] \end{aligned}$$

$$\leq \int_{\mathbb{T}^n} w(x, 0) d\nu_0 - \int_{\mathbb{T}^n} w(x, t) d\nu_1,$$

which implies $h_t(\nu_0, \nu_1) \geq m_t(\nu_0, \nu_1)$.

Now take $\varphi \in C^2(\mathbb{T}^n \times [0, t])$. By subtracting Ct to φ with some $C \in \mathbb{R}$, we can assume that

$$\max_{(x,s) \in \mathbb{T}^n \times [0,t]} (\varphi_t(x, s) - a(x)\Delta\varphi(x, s) + H(x, D\varphi(x, s))) = 0;$$

that is, φ is a classical subsolution of (1.1). Next, let w^φ be the solution of the Hamilton–Jacobi–Bellman equation (1.1) with initial data $w^\varphi(x, 0) = \varphi(x, 0)$. Then, by the comparison, $\varphi(x, t) \leq w^\varphi(x, t)$. Accordingly,

$$\int_{\mathbb{T}^n} w^\varphi(x, 0) d\nu_0 - \int_{\mathbb{T}^n} w^\varphi(x, t) d\nu_1 \leq \int_{\mathbb{T}^n} \varphi(x, 0) d\nu_0 - \int_{\mathbb{T}^n} \varphi(x, t) d\nu_1.$$

Therefore,

$$\begin{aligned} -h_t(\nu_0, \nu_1) &= \inf_{\varphi \in C^2(\mathbb{T}^n \times [0,t])} \sup_{(x,s) \in \mathbb{T}^n \times [0,t]} \left[\int_{\mathbb{T}^n} \varphi(x, 0) d\nu_0 - \int_{\mathbb{T}^n} \varphi(x, t) d\nu_1 + \right. \\ &\quad \left. + t(\varphi_t - a(x)\Delta\varphi + H(x, D\varphi)) \right] \\ &\geq \inf_{\varphi \in C^2(\mathbb{T}^n \times [0,t])} \left(\int_{\mathbb{T}^n} w^\varphi(x, 0) d\nu_0 - \int_{\mathbb{T}^n} w^\varphi(x, t) d\nu_1 \right) \\ &\geq \inf_{w \in \mathcal{E}} \left(\int_{\mathbb{T}^n} w(x, 0) d\nu_0 - \int_{\mathbb{T}^n} w(x, t) d\nu_1 \right) = -m_t(\nu_0, \nu_1), \end{aligned}$$

which implies $h_t(\nu_0, \nu_1) \leq m_t(\nu_0, \nu_1)$. \square

5.3. Convex duality for d on \mathcal{M} . Now, we examine the function d in light of duality methods and prove Theorem 1.5.

Proposition 5.7. *Consider the setting of Problem 1 and suppose that Assumptions 1–3 hold. Let $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$. Then*

$$d(\nu_0, \nu_1) \geq \sup_{w \in \mathcal{S}} \left[\int_{\mathbb{T}^n} w d\nu_1 - \int_{\mathbb{T}^n} w d\nu_0 \right], \quad (5.5)$$

where \mathcal{S} is the set of all Lipschitz viscosity solutions $w \in C(\mathbb{T}^n)$ of (1.5) defined by (1.11).

Proof. Recall that

$$d(\nu_0, \nu_1) = \inf_{t>0} h_t(\nu_0, \nu_1).$$

Let $w \in C(\mathbb{T}^n)$ solve (1.5). Then, $w \in \text{Lip}(\mathbb{T}^n)$. Because of our convention that $c = 0$, w solves (1.1). Therefore

$$h_t(\nu_0, \nu_1) \geq \left[\int_{\mathbb{T}^n} w d\nu_1 - \int_{\mathbb{T}^n} w d\nu_0 \right].$$

Consequently, we obtain (5.5). \square

Lemma 5.8. *Consider the setting of Problem 1 and suppose that Assumptions 1-3 hold. Let $\nu_0 \in \mathcal{M}$. Given $\varphi \in C(\mathbb{T}^n)$, let w^φ be the solution of (1.1) with initial condition $w^\varphi(x, 0) = \varphi(x)$. Then, the map*

$$t \mapsto \int_{\mathbb{T}^n} w^\varphi(x, t) d\nu_0$$

is non-increasing.

Proof. Let μ_0 be a Mather measure such that $\nu_0 = \text{proj}_{\mathbb{T}^n} \mu_0$, and η be a measure such that

$$d\eta(x, q, s) = d\mu_0(x, q) \quad \text{for all } s \geq 0.$$

By the representation formula for w^φ , we have

$$\int_{\mathbb{T}^n} w^\varphi(x, t) d\nu_0 \leq \int_{\mathbb{T}^n} \varphi d\nu_0 + \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\eta(x, q, s) = \int_{\mathbb{T}^n} \varphi d\nu_0.$$

We hence see that for $t > 0$

$$\int_{\mathbb{T}^n} w^\varphi(x, t) d\nu_0 \leq \int_{\mathbb{T}^n} w^\varphi(x, 0) d\nu_0.$$

Thus, by the semigroup property of the solution operator of the Hamilton–Jacobi–Bellman equation (see [33]), we get the claim in the statement. \square

Proposition 5.9. *Consider the setting of Problem 1 and suppose that Assumptions 1-4 hold. Let $\nu_0, \nu_1 \in \mathcal{M}$ and suppose that $d(\nu_0, \nu_1) < +\infty$. Then*

$$d(\nu_0, \nu_1) \leq \sup_{w \in \mathcal{S}} \left[\int_{\mathbb{T}^n} w d\nu_1 - \int_{\mathbb{T}^n} w d\nu_0 \right].$$

Proof. Fix $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$.

Fix $\varepsilon > 0$ and let T be as in the statement of Proposition 1.3. Let $\{t_j\}$ be a sequence with $t_j \rightarrow \infty$ and $h_{t_j}(\nu_0, \nu_1) \rightarrow d(\nu_0, \nu_1)$ as $j \rightarrow \infty$. Select $t_j > T$ such that

$$|d(\nu_0, \nu_1) - h_{t_j}(\nu_0, \nu_1)| \leq \varepsilon.$$

Next, by Theorem 1.4, for every $\varepsilon > 0$ there exists a Lipschitz viscosity solution u to (1.1) such that

$$h_{t_j}(\nu_0, \nu_1) \leq \int_{\mathbb{T}^n} u(\cdot, t_j) d\nu_1 - \int_{\mathbb{T}^n} u(\cdot, 0) d\nu_0 + \varepsilon$$

Let $u_\infty(x) = \lim_{t \rightarrow \infty} u(x, t)$ for $x \in \mathbb{T}^n$. According to Lemma 5.8,

$$\int_{\mathbb{T}^n} u(\cdot, 0) d\nu_0 \geq \int_{\mathbb{T}^n} u_\infty d\nu_0.$$

Moreover, due to the convergence result in Proposition 1.3, we have

$$\int_{\mathbb{T}^n} u(\cdot, t) d\nu_1 \leq \int_{\mathbb{T}^n} u_\infty d\nu_1 + \varepsilon \quad \text{for all } t > T.$$

Therefore,

$$d(\nu_0, \nu_1) \leq \int_{\mathbb{T}^n} u_\infty d\nu_1 - \int_{\mathbb{T}^n} u_\infty d\nu_0 + 3\varepsilon \leq m(\nu_0, \nu_1) + 3\varepsilon.$$

Because ε is arbitrary and u_∞ solves (1.5), we obtain the desired inequality. \square

Proof of Theorem 1.5. Theorem 1.5 is a straightforward result of Propositions 5.7 and 5.9. \square

6. PROOF OF THEOREM 1.6

In this section, we prove Theorem 1.6 in case that $m > 2$. If $a(x) \equiv 0$, the compactness of S_1 follows from the coercivity of the Hamiltonian.

Proof of Theorem 1.6. It is enough to show the compactness of S_1 ; that is, to prove that $\{S_1 v\}_{v \in X_0}$ is uniformly bounded and equicontinuous. The main difficulty here is that $X_0 = \{u_0 \in C(\mathbb{T}^n) : \min_{\mathbb{T}^n} u_0 = 0\}$ is not bounded in $C(\mathbb{T}^n)$.

Take $u_0 \in X_0$. Without loss of generality, assume that

$$u_0(0) = \min_{\mathbb{T}^n} u_0 = 0.$$

It is clear that for any $C \geq \max_{x \in \mathbb{T}^n} |H(x, 0)|$, $\phi(x, t) = -Ct$ is a subsolution to (1.1). Therefore,

$$u(x, t) \geq -Ct \quad \text{in } \mathbb{T}^n \times [0, \infty).$$

We need to construct a supersolution to bound u from above. For this, the superquadratic growth of H is crucial. Here, we borrow some ideas from [11]. For any $C > 0$ and $\lambda > 0$ to be chosen later, let

$$\psi(x, t) = \lambda t^{-\frac{1}{m-1}} (|x|^2 + \lambda t)^{\frac{m'}{2}} + Ct \quad \text{for all } (x, t) \in \mathbb{R}^n \times (0, \infty).$$

Note that $m' = \frac{m}{m-1} \in (1, 2)$ because $m > 2$. We compute that

$$\begin{aligned} \psi_t &= \lambda t^{-m'} (|x|^2 + \lambda t)^{\frac{m'}{2}} \left(\frac{m'}{2} \frac{\lambda t}{|x|^2 + \lambda t} - \frac{1}{m-1} \right) + C, \\ \psi_{x_i} &= \lambda t^{-\frac{1}{m-1}} (|x|^2 + \lambda t)^{\frac{m'}{2}-1} m' x_i, \\ \psi_{x_i x_j} &= \lambda t^{-\frac{1}{m-1}} (|x|^2 + \lambda t)^{\frac{m'}{2}-1} m' \left(\delta_{ij} + \frac{(m'-2)x_i x_j}{|x|^2 + \lambda t} \right). \end{aligned}$$

Therefore, for $(x, t) \in \mathbb{R}^n \times (0, \infty)$, noting that $mm' = m + m'$,

$$|D\psi(x, t)|^m = (m')^m \lambda^m t^{-m'} (|x|^2 + \lambda t)^{\frac{m'}{2}} \left(\frac{|x|^2}{|x|^2 + \lambda t} \right)^{\frac{m}{2}},$$

and

$$a(x)\Delta\psi(x, t) \leq C\lambda t^{-m'} (|x|^2 + \lambda t)^{\frac{m'}{2}} \frac{t}{|x|^2 + \lambda t}.$$

Combining the above computations, for $(x, t) \in \mathbb{R}^n \times (0, \infty)$, we gather

$$\begin{aligned} \psi_t + H(x, D\psi) - a(x)\Delta\psi &\geq \psi_t + \frac{|D\psi|^m}{C} - C - a(x)\Delta\psi \\ &\geq \lambda t^{-m'} (|x|^2 + \lambda t)^{\frac{m'}{2}} \left(\left(\frac{\lambda m'}{2} - C \right) \frac{t}{|x|^2 + \lambda t} - \frac{1}{m-1} + \frac{\lambda^{m-1}}{C} \left(\frac{|x|^2}{|x|^2 + \lambda t} \right)^{\frac{m}{2}} \right). \end{aligned}$$

Because $m > 2$,

$$\frac{m'}{2} = \frac{m-1}{2} > \frac{1}{m-1}.$$

Accordingly, let $\theta = \frac{1}{2} \left(\frac{m'}{2} - \frac{1}{m-1} \right) > 0$. Then, for $\lambda > 0$ sufficiently large,

$$\frac{\lambda m'}{2} - C \geq \lambda \left(\frac{1}{m-1} + \theta \right).$$

Let $z = \frac{|x|^2}{|x|^2 + \lambda t} \in [0, 1]$. Then,

$$\begin{aligned} & \left(\frac{\lambda m'}{2} - C \right) \frac{t}{|x|^2 + \lambda t} - \frac{1}{m-1} + \frac{\lambda^{m-1}}{C} \left(\frac{|x|^2}{|x|^2 + \lambda t} \right)^{\frac{m}{2}} \\ & \geq \left(\frac{1}{m-1} + \theta \right) (1-z) - \frac{1}{m-1} + \frac{\lambda^{m-1}}{C} z^{\frac{m}{2}} \\ & = \frac{\lambda^{m-1}}{C} z^{\frac{m}{2}} + \theta - \left(\frac{1}{m-1} + \theta \right) z > 0 \end{aligned}$$

for $\lambda > 0$ sufficiently large depending only on C, n, m .

Thus, ψ is a classical strict supersolution to (1.1) in $\mathbb{R}^n \times (0, \infty)$, and $\psi(0, 0) = 0$, $\psi(x, 0) = +\infty$ for $x \neq 0$. In particular, $\psi(x, 0) \geq u_0(x)$ for all $x \in \mathbb{R}^n$. The viscosity subsolution test implies that $u - \psi$ cannot have a maximum point at $(x_0, t_0) \in \mathbb{R}^n \times (0, \infty)$. Therefore, $u \leq \psi$ in $\mathbb{R}^n \times (0, \infty)$ by the comparison principle. In particular, for $x \in [0, 1]^n$ and $t = \frac{1}{2}$,

$$-\frac{C}{2} \leq u \left(x, \frac{1}{2} \right) \leq \lambda 2^{\frac{1}{m-1}} \left(|x|^2 + \frac{\lambda}{2} \right)^{\frac{m'}{2}} + \frac{C}{2} \leq C_0,$$

where C_0 depends only on C, n, m . Thus, we see that, for any $u_0 \in C(\mathbb{T}^n)$ such that $\min_{\mathbb{T}^n} u_0 = 0$,

$$\|S_{\frac{1}{2}} u_0\|_{L^\infty(\mathbb{T}^n)} \leq C_0. \quad (6.1)$$

We now use (6.1) together with [11, Theorem 1.2] to conclude that $S_1 u_0 \in C^{0,\alpha}(\mathbb{T}^n)$, where $\alpha \in (0, 1)$ and $\|S_1 u_0\|_{C^{0,\alpha}(\mathbb{T}^n)}$ depend only on C, C_0, m, n . The proof is hence complete. \square

For further Hölder regularity results with respect to superquadratic Hamiltonians, we refer the readers to [2, 9, 10] and the references therein.

A. APPROXIMATIONS OF SOLUTIONS TO PROBLEM 1

Let $\theta \in C_c^\infty(\mathbb{R}^n, [0, \infty))$ and $\rho \in C_c^\infty(\mathbb{R}, [0, \infty))$ be symmetric standard mollifiers; that is, $\text{supp } \theta \subset \overline{B}(0, 1) \subset \mathbb{R}^n$, $\text{supp } \rho \subset \overline{B}(0, 1) \subset \mathbb{R}$, $\theta(x) = \theta(-x)$, $\rho(s) = \rho(-s)$, and $\|\theta\|_{L^1(\mathbb{R}^n)} = \|\rho\|_{L^1(\mathbb{R})} = 1$. For $\alpha > 0$, set $\theta^\alpha(x) := \alpha^{-n} \theta(\alpha^{-1}x)$ for $x \in \mathbb{R}^n$, and $\rho^\alpha(t) := \alpha^{-1} \rho(\alpha^{-1}t)$ for $t \in \mathbb{R}$. For $w \in C(\mathbb{T}^n \times [0, \infty))$, let $w^\alpha \in C^\infty(\mathbb{T}^n \times [\alpha, \infty))$ be

$$w^\alpha(x, t) := \int_0^\infty \rho^\alpha(s) \int_{\mathbb{R}^n} \theta^\alpha(y) w(x-y, t-s) dy ds \quad \text{for } (x, t) \in \mathbb{T}^n \times [\alpha, \infty). \quad (\text{A.1})$$

Proposition A.1. *Let w be a Lipschitz solution to (1.1). For $0 < \alpha < 1$, let w^α be as in (A.1). Then, there exists $C > 0$ depending on H, a , and the Lipschitz constant of w such that*

$$w_t^\alpha - a(x)\Delta w^\alpha + H(x, Dw^\alpha) \leq C\alpha^{1/2} \quad \text{on } \mathbb{T}^n \times [\alpha, \infty). \quad (\text{A.2})$$

We give only a brief outline of the proof. For a detailed proof, see [33, 34].

Proof outline. To obtain the inequality in (A.2), we seek to rewrite its left-hand side in terms of the convolution of the left-hand side in (1.1) with $\rho^\alpha \theta^\alpha$. We will handle each of the terms separately. The first term, w_t^α is trivial. For the last term, $H(x, Dw^\alpha)$, we observe that Jensen's inequality gives, for $(x, t) \in \mathbb{T}^n \times [\alpha, \infty)$,

$$\begin{aligned} H(x, Dw^\alpha(x, t)) &= H\left(x, \int_0^\infty \rho^\alpha(s) \int_{\mathbb{T}^n} \theta^\alpha(y) Dw(x-y, t-s) dy ds\right) \\ &\leq \int_0^\infty \rho^\alpha(s) \int_{\mathbb{T}^n} \theta^\alpha(y) H(x, Dw(x-y, t-s)) dy ds \\ &\leq \int_0^\infty \rho^\alpha(s) \int_{\mathbb{T}^n} \theta^\alpha(y) H(x-y, Dw(x-y, t-s)) dy ds + C\alpha. \end{aligned}$$

Thus, the term $H(x, Dw^\alpha(x, t))$ is controlled by the corresponding term in (1.1) convolved with $\rho^\alpha \theta^\alpha$ and an error term bounded by $C\alpha$.

The second term, $a(x)\Delta w^\alpha$, is where the main difficulty of the estimate lies. Because w is Lipschitz, using equation (1.1), we have

$$-C \leq a(x)\Delta w \leq C \quad \text{on } \mathbb{T}^n \times [0, \infty) \quad (\text{A.3})$$

in viscosity sense.

Because of the simple structure of a , we see further that $\|a\Delta w\|_{L^\infty(\mathbb{T}^n \times [0, \infty))} \leq C$, and w is a subsolution to (1.1) and (A.3) in the distributional sense. We need to control the commutation term,

$$\begin{aligned} &\int_0^\infty \rho^\alpha(s) \int_{\mathbb{T}^n} \theta^\alpha(y) a(x-y) \Delta w(x-y, t-s) dy ds - a(x) \Delta w^\alpha(x, t) \\ &= \int_0^\infty \rho^\alpha(s) \left(\int_{\mathbb{T}^n} \theta^\alpha(y) (a(x-y) - a(x)) \Delta w(x-y, t-s) dy \right) ds \\ &= \int_0^\infty \rho^\alpha(s) R^\alpha(x, t-s) ds, \end{aligned}$$

where

$$R^\alpha(x, t-s) = \int_{\mathbb{T}^n} \theta^\alpha(y) (a(x-y) - a(x)) \Delta w(x-y, t-s) dy.$$

To complete the proof, we show that $|R^\alpha(x, t)| \leq C\alpha^{1/2}$ for all $(x, t) \in \mathbb{T}^n \times [0, \infty)$. We consider two cases

$$(i) \quad \min_{y \in \overline{B(x, \alpha)}} a(y) \leq \alpha, \quad \text{or} \quad (ii) \quad \min_{y \in \overline{B(x, \alpha)}} a(y) > \alpha.$$

In case (i), there exists $\bar{x} \in \overline{B(x, \alpha)}$ such that $a(\bar{x}) \leq \alpha$. Then, there exists a constant $C > 0$ such that,

$$|Da(\bar{x})| \leq Ca(\bar{x})^{1/2} \leq C\alpha^{1/2}.$$

See [7, Lemma 2.6] for example. For any $z \in \overline{B(x, \alpha)}$,

$$|Da(z)| \leq |Da(z) - Da(\bar{x})| + |Da(\bar{x})| \leq C\alpha + C\alpha^{1/2} \leq C\alpha^{1/2}.$$

Moreover, by using Taylor's expansion,

$$\begin{aligned} |a(z) - a(x)| &\leq |a(z) - a(\bar{x})| + |a(x) - a(\bar{x})| \\ &\leq |Da(\bar{x})|(|z - \bar{x}| + |x - \bar{x}|) + C(|z - \bar{x}|^2 + |x - \bar{x}|^2) \leq C\alpha^{3/2} + C\alpha^2 \leq C\alpha^{3/2}. \end{aligned}$$

We use the two above inequalities to control $R^\alpha(x, t)$ as

$$\begin{aligned} |R^\alpha(x, t)| &= \left| \int_{\mathbb{R}^n} (a(x-y) - a(x)) \Delta w(x-y, t) \theta^\alpha(y) dy \right| \\ &= \left| \int_{\mathbb{R}^n} Dw(x-y, t) \cdot Da(x-y) \theta^\alpha(y) dy - \int_{\mathbb{R}^n} Dw(x-y, t) \cdot D\theta^\alpha(y) (a(x-y) - a(x)) dy \right| \\ &\leq C \int_{\mathbb{R}^n} (\alpha^{1/2} \theta^\alpha(y) + \alpha^{3/2} |D\theta^\alpha(y)|) dy \leq C\alpha^{1/2}. \end{aligned}$$

Now, we consider case (ii); that is, $\min_{\overline{B(x, \alpha)}} a > \alpha$. A direct computation shows that

$$\begin{aligned} |R^\alpha(x, t)| &\leq \int_{\mathbb{R}^n} |(a(x-y) - a(x))| \cdot |\Delta w(x-y, t)| \theta^\alpha(y) dy \\ &\leq C \int_{\mathbb{R}^n} \frac{|a(x-y) - a(x)|}{a(x-y)} \theta^\alpha(y) dy \leq C \int_{\mathbb{R}^n} \frac{|Da(x-y)| \cdot |y|}{a(x-y)} \theta^\alpha(y) dy + C\alpha \\ &\leq C \int_{\mathbb{R}^n} \frac{|y|}{a(x-y)^{1/2}} \theta^\alpha(y) dy + C\alpha \leq C \int_{\mathbb{R}^n} \frac{|y|}{\alpha^{1/2}} \theta^\alpha(y) dy + C\alpha \leq C\alpha^{1/2}. \end{aligned}$$

Combining these estimates yields the conclusion. \square

B. THE GENERAL DIFFUSION MATRIX CASE

Now, we consider the case of a general diffusion matrix A . Let \mathbb{S}^n be the space of $n \times n$ real symmetric matrices. Let $A : \mathbb{T}^n \rightarrow \mathbb{S}^n$ be a nonnegative definite diffusion matrix; that is, $\xi^T A(x) \xi \geq 0$ for all $\xi \in \mathbb{R}^n$ and $x \in \mathbb{T}^n$. Assume further that $A \in C^2(\mathbb{T}^n, \mathbb{S}^n)$. We suppose that Assumptions 1–3 always hold in this section and replace (1.1) in Problem 1 by the following general Hamilton–Jacobi–Bellman equation

$$u_t - \text{tr}(A(x)D^2u) + H(x, Du) = 0 \quad \text{in } \mathbb{T}^n \times (0, \infty), \quad (\text{B.1})$$

where Du and D^2u , respectively, denote the spatial gradient and Hessian of u , and with initial data

$$u(x, 0) = u_0(x) \quad \text{on } \mathbb{T}^n.$$

We now extend the results in Theorems 1.1–1.2 for this setting. While the statements are similar, there are several technical points that must be addressed. The main difficulty is that the analog of the approximation result in Proposition A.1 is substantially more involved. This approximation result is examined in Proposition B.12 and requires the approximate equation to be uniformly parabolic. Thus, further approximation arguments are needed in various places, including in the definition of holonomic measures.

B.1. Representation formulas for the general case of diffusion matrices. For $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$, $\eta > 0$, let $\mathcal{H}^\eta(\nu_0, \nu_1; t_0, t_1)$ be the set of all $\gamma \in \mathcal{R}^+(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ satisfying

$$\int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} |q|^\zeta d\gamma(z, q, s) < \infty$$

and

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} (\varphi_t(z, s) - \eta \Delta \varphi(z, s) - \text{tr}(A(z)D^2\varphi(z, s)) + q \cdot D\varphi(z, s)) d\gamma(z, q, s) \\ &= \int_{\mathbb{T}^n} \varphi(z, t_1) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, t_0) d\nu_0(z) \end{aligned}$$

for all $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$. Moreover, we set

$$\mathcal{H}^\eta(\nu_1; t_0, t_1) := \bigcup_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \mathcal{H}^\eta(\nu_0, \nu_1; t_0, t_1).$$

Fix any $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$. It is worth emphasizing that while we do not know that $\mathcal{H}^\eta(\nu_0, \nu_1; t_0, t_1) \neq \emptyset$ for each $\nu_0 \in \mathcal{P}(\mathbb{T}^n)$, the set $\mathcal{H}^\eta(\nu_1; t_0, t_1)$ is always non-empty as shown in the proof of Lemma B.5.

We define $\tilde{\mathcal{H}}(\nu_1; t_0, t_1)$ as follows. We say that $\gamma \in \tilde{\mathcal{H}}(\nu_1; t_0, t_1)$ if $\gamma \in \mathcal{R}^+(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$, and there exist $C > 0$, a sequence $\{\eta_j\}_{j \in \mathbb{N}} \rightarrow 0$, $\{\nu_1^{\eta_j}\}_{j \in \mathbb{N}} \subset \mathcal{P}(\mathbb{T}^n)$, and $\{\gamma^{\eta_j}\}_{j \in \mathbb{N}} \subset \mathcal{R}^+(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ such that

$$\begin{aligned} & \gamma^{\eta_j} \in \mathcal{H}^{\eta_j}(\nu_1^{\eta_j}; t_0, t_1), \text{ and } \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} |q|^\zeta d\gamma^{\eta_j}(z, q, s) < C \quad \text{for each } j \in \mathbb{N}, \\ & \nu_1^{\eta_j} \rightharpoonup \nu_1 \quad \text{weakly in the sense of measures in } \mathcal{R}(\mathbb{T}^n), \\ & \gamma^{\eta_j} \rightharpoonup \gamma \quad \text{weakly in the sense of measures in } \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]) \text{ as } j \rightarrow \infty. \end{aligned}$$

We also note that $\tilde{\mathcal{H}}(\nu_1; t_0, t_1)$ is non-empty for any $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$ as stated in Corollary B.6. For $\gamma \in \tilde{\mathcal{H}}(\nu_1; t_0, t_1)$, let ν^γ be the unique element in $\mathcal{P}(\mathbb{T}^n)$ such that

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} (\varphi_t(z, s) - \eta \Delta \varphi(z, s) - \text{tr}(A(z)D^2\varphi(z, s)) + q \cdot D\varphi(z, s)) d\gamma(z, q, s) \\ &= \int_{\mathbb{T}^n} \varphi(z, t_1) d\nu_1(z) - \int_{\mathbb{T}^n} \varphi(z, t_0) d\nu^\gamma(z) \end{aligned}$$

for all $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$. Accordingly, $\gamma \in \tilde{\mathcal{H}}(\nu^\gamma, \nu_1; t_0, t_1)$.

We now use the measures in $\tilde{\mathcal{H}}(\nu_1; t_0, t_1)$ to obtain a representation formula for solutions of (B.1), which can be viewed as a generalization of Theorem 1.1.

Theorem B.1. *Let u solve (B.1). Suppose that Assumptions 1-3 hold. Then, for any $\nu \in \mathcal{P}(\mathbb{T}^n)$ and $t > 0$, we have*

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) = \inf_{\gamma \in \tilde{\mathcal{H}}(\nu; 0, t)} \left[\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu^\gamma(z) \right].$$

This theorem is proved in the next subsection.

The ergodic problem here is

$$-\operatorname{tr}(A(x)D^2v) + H(x, Dv) = c \quad \text{in } \mathbb{T}^n. \quad (\text{B.2})$$

As previously (cf Remark 1), we add a constant to H , if necessary, so that $c = 0$. As in the time-dependent case, we begin by defining the sets of approximated stationary generalized holonomic measures. For each $\eta > 0$, we denote by

$$\mathcal{H}^\eta := \left\{ \mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) : \int_{\mathbb{T}^n \times \mathbb{R}^n} |q|^\zeta d\mu(z, q) < \infty, \right. \\ \left. \int_{\mathbb{T}^n \times \mathbb{R}^n} (-\eta \Delta \varphi(z) - \operatorname{tr}(A(z)D^2\varphi(z)) + q \cdot D\varphi(z)) d\mu(z, q) = 0 \quad \text{for all } \varphi \in C^2(\mathbb{T}^n) \right\}.$$

In a similar manner to the time-dependent case, $\tilde{\mathcal{H}}$ is the set of all $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ for which there exist $C > 0$, $\{\eta_j\}_{j \in \mathbb{N}} \rightarrow 0$, and $\{\mu^{\eta_j}\}_{j \in \mathbb{N}} \subset \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ such that

$$\mu^{\eta_j} \in \mathcal{H}^{\eta_j}, \text{ and } \int_{\mathbb{T}^n \times \mathbb{R}^n} |q|^\zeta d\mu^{\eta_j}(z, q) < C \quad \text{for all } j \in \mathbb{N}, \\ \mu^{\eta_j} \rightharpoonup \mu \quad \text{weakly in the sense of measures in } \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n) \text{ as } j \rightarrow \infty.$$

We consider the variational problem

$$\inf_{\mu \in \tilde{\mathcal{H}}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\mu(x, q). \quad (\text{B.3})$$

A *generalized Mather measure* is a solution of the minimization problem in (B.3) and $\tilde{\mathcal{M}}$ is the set of all generalized Mather measures. Moreover, \mathcal{M} is the set of all generalized projected Mather measures; that is, the projections to \mathbb{T}^n of generalized Mather measures.

Proposition B.2. *Suppose that Assumptions 1-3 hold. Assume that the ergodic constant c for (B.2) is 0. We have*

$$\inf_{\mu \in \tilde{\mathcal{H}}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\mu(x, q) = 0.$$

This proposition is proved at the end of the paper.

Finally, we have the following representation result, which is a generalized version of Theorem 1.2.

Theorem B.3. *Suppose that Assumptions 1-3 hold. Let u solve (B.1) and u_∞ be as in (1.4). Then, for any $\nu \in \mathcal{M}$, we have*

$$\int_{\mathbb{T}^n} u_\infty(z) d\nu(z) = \inf_{\nu_0 \in \mathcal{P}(\mathbb{T}^n)} \left[d(\nu_0, \nu) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z) \right],$$

where for $\nu_0, \nu_1 \in \mathcal{P}(\mathbb{T}^n)$, d is the generalized Mañé critical potential connecting ν_0 to ν_1 given by

$$d(\nu_0, \nu_1) := \inf_{\substack{\gamma \in \tilde{\mathcal{H}}(\nu_0, \nu_1; 0, t) \\ t > 0}} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma(x, q, s).$$

The proof of the preceding theorem is similar to the one of Theorem 1.2, so we omit it here.

B.2. Proof of Theorem B.1. We begin by proving Theorem B.1.

Lemma B.4. *Suppose that Assumptions 1-3 hold. Let u solve (B.1). Then, for $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$ and $t > 0$,*

$$\int_{\mathbb{T}^n} u(z, t) d\nu_1(z) \leq \inf_{\gamma \in \tilde{\mathcal{H}}(\nu_1; 0, t)} \left[\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u(z, 0) d\nu^\gamma(z) \right].$$

Proof. Note that u is globally Lipschitz continuous on $\mathbb{T}^n \times [0, \infty)$ (see [2, Proposition 3.5], [33, Proposition 4.15] for instance) under Assumptions 1–3.

For $\alpha, \varepsilon, \delta > 0$, let $u^{\alpha, \varepsilon, \delta}$ be the function given by (B.9) in Section B.3 below, and set

$$\tilde{u}(x, t) := u^{\alpha, \varepsilon, \delta}(x, t + \alpha) \quad \text{for all } (x, t) \in \mathbb{T}^n \times [0, \infty).$$

We notice that $\tilde{u} \in C^2(\mathbb{T}^n \times [0, \infty))$ and it is an approximate subsolution to (B.1), as stated in Proposition B.12.

Fix $\gamma \in \tilde{\mathcal{H}}(\nu_1; 0, t)$. By the definition of $\tilde{\mathcal{H}}(\nu_1; 0, t)$, there exist $\{\eta_j\}_{j \in \mathbb{N}} \rightarrow 0$, $\{\nu_1^{\eta_j}\}_{j \in \mathbb{N}} \subset \mathcal{P}(\mathbb{T}^n)$, and $\{\gamma^{\eta_j}\} \subset \mathcal{R}^+(\mathbb{T}^n \times \mathbb{R}^n \times [0, t])$ such that $\gamma^{\eta_j} \in \mathcal{H}^{\eta_j}(\nu_1^{\eta_j}; 0, t)$ for all $j \in \mathbb{N}$, and

$$\begin{aligned} \nu_1^{\eta_j} &\rightharpoonup \nu_1 \quad \text{weakly in the sense of measures in } \mathcal{R}(\mathbb{T}^n), \\ \gamma^{\eta_j} &\rightharpoonup \gamma \quad \text{weakly in the sense of measures in } \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]) \text{ as } j \rightarrow \infty. \end{aligned}$$

Then,

$$\begin{aligned} &\int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\tilde{u}_t(z, s) - \eta_j \Delta \tilde{u}(z, s) - \text{tr}(A(z) D^2 \tilde{u}(z, s)) + q \cdot D \tilde{u}(z, s)) d\gamma^{\eta_j}(z, q, s) \\ &= \int_{\mathbb{T}^n} \tilde{u}(z, t) d\nu_1^{\eta_j}(z) - \int_{\mathbb{T}^n} \tilde{u}(z, 0) d\nu^{\gamma^{\eta_j}}(z). \end{aligned}$$

Because of the definition of the Legendre transform in (1.2),

$$q \cdot D \tilde{u}(z, s) \leq L(z, q) + H(z, D \tilde{u}(z, s)).$$

Accordingly, we have

$$\begin{aligned} &\int_{\mathbb{T}^n} \tilde{u}(z, t) d\nu_1^{\eta_j} - \int_{\mathbb{T}^n} \tilde{u}(z, 0) d\nu^{\gamma^{\eta_j}}(z) \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\tilde{u}_t - \eta_j \Delta \tilde{u} - \text{tr}(A(z) D^2 \tilde{u}) + q \cdot D \tilde{u}) d\gamma^{\eta_j}(z, q, s) \\ &\leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} (\tilde{u}_t - \eta_j \Delta \tilde{u} - \text{tr}(A(z) D^2 \tilde{u}) + H(z, D \tilde{u}) + L(z, q)) d\gamma^{\eta_j}(z, q, s) \\ &\leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma^{\eta_j}(z, q, s) + \kappa(\alpha, \eta_j, \delta, \varepsilon)t, \end{aligned}$$

where $\kappa(\alpha, \eta, \delta, \varepsilon)$ is defined by (B.11) in Section B.3 below. By taking a subsequence if necessary, we have

$$\nu^{\gamma^{\eta_j}} \rightharpoonup \nu^\gamma \quad \text{as } j \rightarrow \infty$$

for some $\nu^\gamma \in \mathcal{P}(\mathbb{T}^n)$ weakly in the sense of measures.

Notice that if we send $\alpha \rightarrow 0$, $j \rightarrow \infty$, and $\varepsilon \rightarrow 0$ in this order, then we have

$$\kappa(\alpha, \eta_j, \delta, \varepsilon) \rightarrow 0.$$

Thus, we obtain

$$\int_{\mathbb{T}^n} u(x, t) d\nu_1 \leq \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u(z, 0) d\nu^\gamma(z).$$

Because $\gamma \in \tilde{\mathcal{H}}(\nu_1; 0, t)$ is arbitrary, the statement follows. \square

To prove the opposite bound, we regularized (B.1) as in Problem 3. That is, for $\varepsilon > 0$, we find $u^\varepsilon : \mathbb{T}^n \times [0, \infty) \rightarrow \mathbb{R}$ solving

$$\begin{cases} u_t^\varepsilon - \operatorname{tr}(A(x)D^2u^\varepsilon) + H(x, Du^\varepsilon) = \varepsilon \Delta u^\varepsilon & \text{in } \mathbb{T}^n \times (0, \infty), \\ u(x, 0) = u_0(x) & \text{on } \mathbb{T}^n. \end{cases} \quad (\text{B.4})$$

If Assumptions 1-3 hold, (B.4) has a unique solution, $u^\varepsilon \in C^2(\mathbb{T}^n \times [0, +\infty))$. Moreover, u^ε is Lipschitz continuous uniformly in $\varepsilon \in (0, 1)$. Further, by the standard viscosity solution theory, $u^\varepsilon \rightarrow u$ locally uniformly, as $\varepsilon \rightarrow 0$, where u solves (B.1).

Now, we use the nonlinear adjoint method, see [15, 37], to construct measures that satisfy an approximated holonomy condition.

Lemma B.5. *For $\varepsilon > 0$, let u^ε solve (B.4). For any $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$, there exist a probability measure $\nu_0^\varepsilon \in \mathcal{P}(\mathbb{T}^n)$ and $\gamma^\varepsilon \in \mathcal{H}(\nu_0^\varepsilon, \nu_1; t_0, t_1)$. Moreover, $q = D_p H(x, Du^\varepsilon(x, t))$ γ^ε -almost everywhere.*

Proof. For $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$, the linearization of (B.4) around the solution, u^ε , is

$$\mathcal{L}^\varepsilon[\varphi] = \varphi_t - a^{ij}\varphi_{x_i x_j} + D_p H(x, Du^\varepsilon) \cdot D\varphi - \varepsilon \Delta \varphi,$$

where we use Einstein's summation convention. Accordingly, the corresponding adjoint equation is the Fokker-Planck equation

$$\begin{cases} -\sigma_t^\varepsilon - (a^{ij}\sigma^\varepsilon)_{x_i x_j} - \operatorname{div}(D_p H(x, Du^\varepsilon)\sigma^\varepsilon) = \varepsilon \Delta \sigma^\varepsilon & \text{in } \mathbb{T}^n \times (t_0, t_1), \\ \sigma^\varepsilon(x, t_1) = \nu_1 & \text{on } \mathbb{T}^n. \end{cases} \quad (\text{B.5})$$

By standard properties of the Fokker-Planck equation,

$$\sigma^\varepsilon > 0 \text{ on } \mathbb{T}^n \times [t_0, t_1), \quad \text{and} \quad \int_{\mathbb{T}^n} \sigma^\varepsilon(x, t) dx = 1 \text{ for all } t \in [t_0, t_1).$$

Next, for each $\varepsilon > 0$ and $t \in [t_0, t_1]$, let $\beta_t^\varepsilon \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be the probability measure determined by

$$\int_{\mathbb{T}^n} \psi(x, Du^\varepsilon)\sigma^\varepsilon(x, t) dx = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, p) d\beta_t^\varepsilon(x, p)$$

for all $\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$. For $t \in [t_0, t_1]$, let $\gamma_t^\varepsilon \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ the pullback of β_t^ε by the map $\Phi(x, q) = (x, D_q L(x, q))$; that is,

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, p) d\beta_t^\varepsilon(x, p) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, D_q L(x, q)) d\gamma_t^\varepsilon(x, q)$$

for all $\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$.

Define the measures $\beta^\varepsilon, \gamma^\varepsilon \in \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$ by

$$\begin{aligned} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} f d\beta^\varepsilon &= \int_{t_0}^{t_1} \int_{\mathbb{T}^n \times \mathbb{R}^n} f(\cdot, t) d\beta_t^\varepsilon dt, \\ \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} f d\gamma^\varepsilon &= \int_{t_0}^{t_1} \int_{\mathbb{T}^n \times \mathbb{R}^n} f(\cdot, t) d\gamma_t^\varepsilon dt \end{aligned}$$

for any $f \in C_c(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1])$.

Multiplying the first equation in (B.5) by an arbitrary function, $\varphi \in C^2(\mathbb{T}^n \times [t_0, t_1])$, and integrating on \mathbb{T}^n , we gather

$$\begin{aligned} \varepsilon \int_{\mathbb{T}^n} \sigma^\varepsilon \Delta \varphi dx &= - \int_{\mathbb{T}^n} \varphi \sigma_t^\varepsilon dx + \int_{\mathbb{T}^n} (-a^{ij}(x) \varphi_{x_i x_j} + D_p H(x, Du^\varepsilon) \cdot D\varphi) \sigma^\varepsilon dx \\ &= - \int_{\mathbb{T}^n} (\varphi \sigma^\varepsilon)_t dx + \int_{\mathbb{T}^n} (\varphi_t - a^{ij}(x) \varphi_{x_i x_j} + D_p H(x, Du^\varepsilon) \cdot D\varphi) \sigma^\varepsilon dx. \end{aligned}$$

Next, integrating on $[t_0, t_1]$, we deduce the identity

$$\begin{aligned} \varepsilon \int_{t_0}^{t_1} \int_{\mathbb{T}^n} \sigma^\varepsilon \Delta \varphi dx dt &= \varepsilon \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} \Delta \varphi d\gamma^\varepsilon(x, q, t) \\ &= - \int_{t_0}^{t_1} \int_{\mathbb{T}^n} (\varphi \sigma^\varepsilon)_t dx dt + \int_{t_0}^{t_1} \int_{\mathbb{T}^n} (\varphi_t - a^{ij}(x) \varphi_{x_i x_j} + D_p H(x, Du^\varepsilon) \cdot D\varphi) \sigma^\varepsilon dx dt \\ &= - \left[\int_{\mathbb{T}^n} \varphi(\cdot, t_1) \sigma^\varepsilon(\cdot, t_1) dx - \int_{\mathbb{T}^n} \varphi(\cdot, t_0) \sigma^\varepsilon(\cdot, t_0) dx \right] \\ &\quad + \int_{t_0}^{t_1} \int_{\mathbb{T}^n \times \mathbb{R}^n} (\varphi_t - a^{ij}(x) \varphi_{x_i x_j} + q \cdot D\varphi) d\gamma_t^\varepsilon(x, q) \\ &= - \left[\int_{\mathbb{T}^n} \varphi(\cdot, t_1) d\nu_1 - \int_{\mathbb{T}^n} \varphi(\cdot, t_0) d\nu_0^\varepsilon \right] \\ &\quad + \int_{\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]} (\varphi_t - a^{ij}(x) \varphi_{x_i x_j} + q \cdot D\varphi) d\gamma^\varepsilon(x, q, t), \end{aligned}$$

where $d\nu_0^\varepsilon := \sigma^\varepsilon(x, t_0) dx$, which implies $\gamma^\varepsilon \in \mathcal{H}^\varepsilon(\nu_0^\varepsilon, \nu_1; t_0, t_1)$. \square

Corollary B.6. *Under Assumptions 1–3, for all $0 < t_0 < t_1$ and $\nu_1 \in \mathcal{P}(\mathbb{T}^n)$,*

$$\tilde{\mathcal{H}}(\nu_1; t_0, t_1) \neq \emptyset.$$

Proof. Let $\nu_0^\varepsilon \in \mathcal{P}(\mathbb{T}^n)$ and $\gamma^\varepsilon \in \mathcal{H}^\varepsilon(\nu_0^\varepsilon, \nu_1; t_0, t_1)$ be measures given by Lemma B.5.

Because

$$\|Du^\varepsilon(\cdot, t)\|_{L^\infty(\mathbb{T}^n)} \leq C \quad \text{for some } C > 0, \text{ for all } t \in [t_0, t_1],$$

there exists a sequence $\{\varepsilon_j\} \rightarrow 0$ such that

$$\nu_0^{\varepsilon_j} \rightharpoonup \nu_0 \in \mathcal{P}(\mathbb{T}^n), \quad \gamma^{\varepsilon_j} \rightharpoonup \gamma \in \mathcal{R}(\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]) \quad \text{as } j \rightarrow \infty, \quad (\text{B.6})$$

weakly in terms of measures on \mathbb{T}^n and $\mathbb{T}^n \times \mathbb{R}^n \times [t_0, t_1]$, respectively. Thus, $\gamma \in \tilde{\mathcal{H}}(\nu_0, \nu_1; t_0, t_1)$, which implies the conclusion. \square

Finally, we use Lemma B.5 to establish the opposite inequality to the one in Lemma B.4.

Lemma B.7. *For any $\nu \in \mathcal{P}(\mathbb{T}^n)$ and $t > 0$, we have*

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) \geq \inf_{\gamma \in \tilde{\mathcal{H}}(\nu; 0, t)} \left\{ \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu^\gamma(z) \right\}.$$

Proof. For $s \in [0, t]$, let γ_s^ε be the measure constructed in the proof of Lemma B.5 for $t_0 = 0$ and $t_1 = t$. By the properties of the Legendre transform

$$L(z, q) = D_p H(z, D_q L(z, q)) \cdot D_q L(z, q) - H(z, D_q L(z, q)).$$

Therefore,

$$\begin{aligned} & \int_{\mathbb{T}^n \times \mathbb{R}^n} L(z, q) d\gamma_s^\varepsilon(z, q) \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(z, D_q L(z, q)) \cdot D_q L(z, q) - H(z, D_q L(z, q))) d\gamma_s^\varepsilon(z, q) \\ &= \int_{\mathbb{T}^n \times \mathbb{R}^n} (D_p H(z, p) \cdot p - H(z, p)) d\beta_s^\varepsilon(z, p) \\ &= \int_{\mathbb{T}^n} (D_p H(x, Du^\varepsilon) \cdot Du^\varepsilon - H(x, Du^\varepsilon)) \sigma^\varepsilon(x, s) dx \end{aligned}$$

for all $s \in [0, t]$. Moreover, integrating by parts and using the adjoint equation and (B.4), we obtain

$$\begin{aligned} & \int_{\mathbb{T}^n} (D_p H(x, Du^\varepsilon) \cdot Du^\varepsilon - H(x, Du^\varepsilon)) \sigma^\varepsilon dx \\ &= \int_{\mathbb{T}^n} -\operatorname{div}(D_p H(x, Du^\varepsilon) \sigma^\varepsilon) u^\varepsilon - H(x, Du^\varepsilon) \sigma^\varepsilon dx \\ &= \int_{\mathbb{T}^n} (\sigma_t^\varepsilon + \varepsilon \Delta \sigma^\varepsilon + (a^{ij} \sigma^\varepsilon)_{x_i x_j}) u^\varepsilon + (u_t^\varepsilon - \varepsilon \Delta u^\varepsilon - a^{ij} u_{x_i x_j}^\varepsilon) \sigma^\varepsilon dx \\ &= \int_{\mathbb{T}^n} (u^\varepsilon \sigma^\varepsilon)_t dx. \end{aligned}$$

Integrating on $[0, t]$ yields

$$\begin{aligned} \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(x, q) d\gamma^\varepsilon(x, q, s) &= \int_0^t \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\gamma_s^\varepsilon(x, q) ds \\ &= \int_{\mathbb{T}^n} u^\varepsilon(x, t) d\nu - \int_{\mathbb{T}^n} u_0(x) d\nu_0^\varepsilon. \end{aligned}$$

Taking subsequences $\{\gamma^{\varepsilon_j}\}$ and $\{\nu_0^{\varepsilon_j}\}$ as in (B.6) yields

$$\int_{\mathbb{T}^n} u(z, t) d\nu(z) = \int_{\mathbb{T}^n \times \mathbb{R}^n \times [0, t]} L(z, q) d\gamma(z, q, s) + \int_{\mathbb{T}^n} u_0(z) d\nu_0(z).$$

Because $\gamma \in \tilde{\mathcal{H}}(\nu_0, \nu; 0, t)$, we obtain the inequality claimed in the statement. \square

Proof of Theorem B.1. The statement follows directly by combining Lemma B.5 with Lemma B.7. \square

B.3. Approximation and proof of Proposition B.2. We first construct an approximation of viscosity solutions of (B.1) by C^2 -subsolutions of an approximate equation. We begin by recalling the definition of the sup and inf-convolutions and some of their basic properties.

Let $w : \mathbb{T}^n \times [0, T] \rightarrow \mathbb{R}$ be a continuous function. The sup-convolution, w^ε , and inf-convolution, w_ε , of w with respect to x for $\varepsilon > 0$ are defined by

$$w^\varepsilon(x, t) := \sup_{y \in \mathbb{T}^n} \left[w(y, t) - \frac{|x - y|^2}{2\varepsilon} \right], \quad w_\varepsilon(x, t) := \inf_{y \in \mathbb{T}^n} \left[w(y, t) + \frac{|x - y|^2}{2\varepsilon} \right].$$

Proposition B.8. *Let $w \in \text{Lip}(\mathbb{T}^n \times [0, \infty))$, and set $L := \|Dw\|_{L^\infty(\mathbb{T}^n \times (0, \infty))}$. We have*

$$w^\varepsilon(x, t) = \max_{|x-y| \leq 2L\varepsilon} \left\{ w(y, t) - \frac{|x - y|^2}{2\varepsilon} \right\}, \quad w_\varepsilon(x, t) = \min_{|x-y| \leq 2L\varepsilon} \left\{ w(y, t) + \frac{|x - y|^2}{2\varepsilon} \right\}.$$

Moreover, $\|Dw^\varepsilon\|_{L^\infty(\mathbb{T}^n \times (0, \infty))} \leq 2L$ and $\|Dw_\varepsilon\|_{L^\infty(\mathbb{T}^n \times (0, \infty))} \leq 2L$.

Proof. We only give a proof for w^ε . Take $x, y \in \mathbb{T}^n$ so that $2L\varepsilon < |x - y|$. Since $w(y, t) - w(x, t) \leq L|x - y| < \frac{|x-y|^2}{2\varepsilon}$, we have $w(y, t) - \frac{|x-y|^2}{2\varepsilon} < w(x, t)$, which implies the first claim.

To get the Lipschitz estimate for w^ε , for a fixed $x \in \mathbb{T}^n$, take $z_x \in \overline{B}(x, 2L\varepsilon)$ so that $w^\varepsilon(x, t) = w(z_x, t) - \frac{|x - z_x|^2}{2\varepsilon}$. Then,

$$\begin{aligned} w^\varepsilon(x, t) - w^\varepsilon(y, t) &\leq \left(w(z_x, t) - \frac{|x - z_x|^2}{2\varepsilon} \right) - \left(w(z_x, t) - \frac{|y - z_x|^2}{2\varepsilon} \right) \\ &= \frac{(y - x) \cdot (x + y - 2z_x)}{2\varepsilon} \leq \frac{|x - y|(|x - z_x| + |y - z_x|)}{2\varepsilon} \\ &\leq \frac{4L + C}{2} |x - y| \end{aligned}$$

if $|x - y| \leq C\varepsilon$ for any $C > 0$, which implies the conclusion. \square

It is well known that the inf-sup convolution $(w^{\varepsilon+\delta})_\delta$ for $\varepsilon, \delta > 0$ gives a $C^{1,1}$ approximation of w in x (see [26] for instance).

Proposition B.9. *We have $(w^{\varepsilon+\delta})_\delta \in \text{Lip}([0, \infty); C^{1,1}(\mathbb{T}^n))$, where we denote by $\text{Lip}([0, \infty); C^{1,1}(\mathbb{T}^n))$ the set of all functions which are Lipschitz on $t \in [0, \infty)$ and $C^{1,1}$ on $x \in \mathbb{T}^n$. Moreover, for each $t > 0$,*

$$-\frac{1}{\varepsilon}I \leq D^2(w^{\varepsilon+\delta})_\delta(x, t) \leq \frac{1}{\delta}I \quad \text{for a.e. } x \in \mathbb{R}^n. \quad (\text{B.7})$$

Proof. It is clear that $(w^{\varepsilon+\delta})_\delta$ is $(1/2\delta)$ -semiconcave. Because the inf and sup convolutions satisfy the semigroup property, that is, $w^{\varepsilon+\delta} = (w^\varepsilon)^\delta$, $w_{\varepsilon+\delta} = (w_\varepsilon)_\delta$ for $\varepsilon, \delta > 0$, we have

$$(w^{\varepsilon+\delta})_\delta = ((w^\varepsilon)^\delta)_\delta,$$

$(w^{\varepsilon+\delta})_\delta$ is $(1/2\varepsilon)$ -semiconvex in light of [12, Proposition 4.5]. Therefore, $(w^{\varepsilon+\delta})_\delta \in \text{Lip}([0, \infty); C^{1,1}(\mathbb{T}^n))$, and (B.7) follows. \square

Lemma B.10. *Let w be a Lipschitz subsolution to (B.1). Then, there exists a modulus of continuity $\omega \in C([0, \infty))$, nondecreasing and with $\omega(0) = 0$, a function $\delta_0(\varepsilon, \eta)$, defined for $\varepsilon, \eta > 0$, such that for $\varepsilon, \eta > 0$, $\delta_0 = \delta_0(\varepsilon, \eta)$ and any $\delta \in (0, \delta_0]$, and any $(x, t) \in \mathbb{T}^n \times (0, \infty)$ where $(w^{\varepsilon+\delta})_\delta$ is twice differentiable in x and differentiable in t we have at (x, t) the following inequality*

$$((w^{\varepsilon+\delta})_\delta)_t - \eta \Delta(w^{\varepsilon+\delta})_\delta - \text{tr}(A(x)D^2(w^{\varepsilon+\delta})_\delta) + H(x, D(w^{\varepsilon+\delta})_\delta) \leq \omega(\varepsilon) + \frac{n\eta}{\varepsilon}.$$

Proof. We first notice that by standard viscosity solution theory, w^ε satisfies

$$w_t^\varepsilon - \text{tr}(A(x)D^2w^\varepsilon) + H(x, Dw^\varepsilon) \leq \omega(\varepsilon)$$

in the sense of viscosity solutions for some nondecreasing function $\omega \in C([0, \infty))$ with $\omega(0) = 0$. Furthermore, because $w^\varepsilon + |x|^2/(2\varepsilon)$ is convex, we see that $-\Delta w^\varepsilon \leq n/\varepsilon$ in $\mathbb{T}^n \times (0, \infty)$ in the sense of viscosity solutions. Thus,

$$w_t^\varepsilon - \eta \Delta w^\varepsilon - \text{tr}(A(x)D^2w^\varepsilon) + H(x, Dw^\varepsilon) \leq \omega(\varepsilon) + \frac{n\eta}{\varepsilon}$$

in the sense of viscosity solutions.

Let $\tilde{w} = (w^{\varepsilon+\delta})_\delta$. Note that $\tilde{w} \geq w^\varepsilon$ on \mathbb{T}^n . Now, let $(\hat{x}, \hat{t}) \in \mathbb{T}^n \times (0, \infty)$ be a point where \tilde{w} is twice differentiable in x and differentiable in t . Select a function, $\varphi \in C^2(\mathbb{T}^n \times (0, \infty))$, such that $\tilde{w} - \varphi$ has a maximum at (\hat{x}, \hat{t}) . At this point either $\tilde{w}(\hat{x}, \hat{t}) = w^\varepsilon(\hat{x}, \hat{t})$ or $\tilde{w}(\hat{x}, \hat{t}) > w^\varepsilon(\hat{x}, \hat{t})$. In the first alternative, that is, when $\tilde{w}(\hat{x}, \hat{t}) = w^\varepsilon(\hat{x}, \hat{t})$, $w^\varepsilon - \varphi$ has a maximum at (\hat{x}, \hat{t}) . Thus,

$$\begin{aligned} & \tilde{w}_t(\hat{x}, \hat{t}) - \eta \Delta \tilde{w}(\hat{x}, \hat{t}) - \text{tr}(A(\hat{x})D^2\tilde{w}(\hat{x}, \hat{t})) + H(x, D\tilde{w}(\hat{x}, \hat{t})) \\ & \leq \varphi_t(\hat{x}, \hat{t}) - \eta \Delta \varphi(\hat{x}, \hat{t}) - \text{tr}(A(\hat{x})D^2\varphi(\hat{x}, \hat{t})) + H(x, D\varphi(\hat{x}, \hat{t})) \leq \omega(\varepsilon) + \frac{n\eta}{\varepsilon}. \end{aligned}$$

In the second alternative, that is, if $\tilde{w}(\hat{x}, \hat{t}) > w^\varepsilon(\hat{x}, \hat{t})$, by [12, Proposition 4.4], $1/\delta$ is one of the eigenvalues of $D^2\tilde{w}(\hat{x}, \hat{t})$. Moreover, by using (B.7), we get

$$\eta \Delta \tilde{w}(\hat{x}, \hat{t}) \geq \eta \left(\frac{1}{\delta} - \frac{n-1}{\varepsilon} \right).$$

Letting $\Lambda := \max_{x \in \mathbb{T}^n} \text{eig}(A(x)) \geq 0$, where $\text{eig}(A(x))$ are the eigenvalues of $A(x)$, we similarly have

$$\text{tr}(A(\hat{x})D^2\tilde{w}(\hat{x}, \hat{t})) \geq -\frac{(n-1)\Lambda}{\varepsilon}.$$

Set $L := \|w_t\|_{L^\infty(\mathbb{T}^n)} + \|Dw\|_{L^\infty(\mathbb{T}^n)}$, and $\bar{C} = \max_{|p| \leq 2L} H(x, p)$. Moreover, we set

$$\delta_0(\varepsilon, \eta) = \frac{\varepsilon\eta}{(n-1)(\Lambda + \eta) + (L + \bar{C})\varepsilon}. \quad (\text{B.8})$$

Combining the preceding estimates, for all $0 < \delta < \delta_0 = \delta_0(\varepsilon, \eta)$, we get

$$\begin{aligned} & \tilde{w}_t(\hat{x}, \hat{t}) - \eta \Delta \tilde{w}(\hat{x}, \hat{t}) - \text{tr}(A(\hat{x})D^2\tilde{w}(\hat{x}, \hat{t})) + H(x, D\tilde{w}(\hat{x}, \hat{t})) \\ & \leq L - \eta \left(\frac{1}{\delta} - \frac{n-1}{\varepsilon} \right) + \frac{(n-1)\Lambda}{\varepsilon} + \bar{C} = -\frac{\eta}{\delta} + \frac{(n-1)(\Lambda + \eta) + (L + \bar{C})\varepsilon}{\varepsilon} \leq 0, \end{aligned}$$

by the choice of δ_0 in (B.8), which finishes the proof. \square

Next, we regularize $(w^{\varepsilon+\delta})_\delta$ further by using standard mollifiers to obtain a C^2 sub-solution. Let $\theta \in C_c^\infty(\mathbb{R}^n, [0, \infty))$ and $\rho \in C_c^\infty(\mathbb{R}, [0, \infty))$ be symmetric standard mollifiers; that is, $\text{supp } \theta \subset \overline{B}(0, 1) \subset \mathbb{R}^n$, $\text{supp } \rho \subset \overline{B}(0, 1) \subset \mathbb{R}$, $\theta(x) = \theta(-x)$, $\rho(s) = \rho(-s)$, and $\|\theta\|_{L^1(\mathbb{R}^n)} = \|\rho\|_{L^1(\mathbb{R})} = 1$. For each $\alpha > 0$, set $\theta^\alpha(x) := \alpha^{-n}\theta(\alpha^{-1}x)$ for $x \in \mathbb{R}^n$, and $\rho^\alpha(t) := \alpha^{-1}\rho(\alpha^{-1}t)$ for $t \in \mathbb{R}$. We define the function $w^{\alpha, \varepsilon, \delta} \in C^\infty(\mathbb{T}^n \times [\alpha, \infty))$ by

$$w^{\alpha, \varepsilon, \delta}(x, t) := \int_0^\infty \rho^\alpha(s) \int_{\mathbb{T}^n} \theta^\alpha(y) (w^{\varepsilon+\delta})_\delta(x - y, t - s) dy ds \quad (\text{B.9})$$

for all $(x, t) \in \mathbb{T}^n \times [\alpha, \infty)$.

Lemma B.11. *Let $w \in \text{Lip}(\mathbb{T}^n \times (0, \infty))$ and $w^{\alpha, \varepsilon, \delta}$ be given by (B.9). Then, there exists a constant $C > 0$, independent of $\alpha, \varepsilon, \delta$, such that for all $(x, t) \in \mathbb{T}^n \times [\alpha, \infty)$, we have*

(i)

$$\begin{aligned} & \left| \text{tr}(A(x)D^2w^{\alpha, \varepsilon, \delta}(x, t)) \right. \\ & \quad \left. - \int_0^\infty \int_{\mathbb{T}^n} \rho^\alpha(s)\theta^\alpha(y) \text{tr}(A(x - y)D^2(w^{\varepsilon+\delta})_\delta(x - y, t - s)) dy ds \right| \\ & \leq C\alpha \max \left\{ \frac{1}{\varepsilon}, \frac{1}{\delta} \right\}, \end{aligned}$$

(ii)

$$\begin{aligned} & H(x, Dw^{\alpha, \varepsilon, \delta}(x, t)) \\ & \quad - \int_0^\infty \int_{\mathbb{T}^n} \rho^\alpha(s)\theta^\alpha(y) H(x - y, D(w^{\varepsilon+\delta})_\delta(x - y, t - s)) dy ds \leq C\alpha. \end{aligned}$$

Proof. We denote, respectively, $w^{\alpha, \varepsilon, \delta}$ and $(w^{\varepsilon+\delta})_\delta$ by w^α and w for simplicity in the proof. We begin by proving the first inequality. We have

$$\begin{aligned} \text{tr}(A(x)D^2w^\alpha(x, t)) &= \text{tr} \left(A(x) \int_0^\infty \int_{\mathbb{T}^n} \rho^\alpha(s)\theta^\alpha(y) D^2w(x - y, t - s) dy ds \right) \\ &= \int_0^\infty \int_{\mathbb{T}^n} \rho^\alpha(s)\theta^\alpha(y) \text{tr}(A(x - y)D^2w(x - y)) dy ds \\ & \quad + \int_0^\infty \int_{|y| \leq \alpha} \rho^\alpha(s)\theta^\alpha(y) \text{tr}((A(x) - A(x - y))D^2w(x - y)) dy ds. \end{aligned}$$

By Proposition B.9, we have, for each $t > 0$,

$$|D^2w(x, t)| \leq C \max \left\{ \frac{1}{\varepsilon}, \frac{1}{\delta} \right\} \quad \text{for a.e. } x \in \mathbb{T}^n,$$

which implies (i).

By the convexity of H , Jensen's inequality implies

$$\begin{aligned} H(x, Dw^\alpha(x, t)) &= H\left(x, \int_0^\infty \int_{\mathbb{T}^n} \rho^\alpha(s) \theta^\alpha(y) Dw(x-y, t-s) dy ds\right) \\ &\leq \int_0^\infty \int_{\mathbb{T}^n} \rho^\alpha(s) \theta^\alpha(y) H(x, Dw(x-y, t-s)) dy ds. \end{aligned}$$

By Proposition B.8 and (A3),

$$|H(x-y, Dw(x-y, t-s)) - H(x, Dw(x-y, t-s))| \leq C\alpha$$

for all $y \in B(x, \alpha)$ and $t-s > 0$. Thus, we finish the proof. \square

Proposition B.12. *Let w be a Lipschitz subsolution to (B.1). For $\alpha, \varepsilon, \delta > 0$, let $w^{\alpha, \varepsilon, \delta} \in C^2(\mathbb{T}^n \times [\alpha, \infty))$ be the function defined by (B.9). For $\eta > 0$, let $\delta_0 = \delta_0(\varepsilon, \eta)$ be the constant given by (B.8). Then, we have*

$$(w^{\alpha, \varepsilon, \delta})_t - \eta \Delta w^{\alpha, \varepsilon, \delta} - \text{tr}(A(x) D^2 w^{\alpha, \varepsilon, \delta}) + H(x, Dw^{\alpha, \varepsilon, \delta}) \leq \kappa(\alpha, \eta, \delta, \varepsilon) \quad (\text{B.10})$$

for all $\delta \in (0, \delta_0]$ and $(x, t) \in \mathbb{T}^n \times [\alpha, \infty)$. Here,

$$\kappa(\alpha, \eta, \delta, \varepsilon) := \omega(\varepsilon) + \frac{n\eta}{\varepsilon} + C \max\left\{\frac{1}{\varepsilon}, \frac{1}{\delta}\right\} \alpha \quad (\text{B.11})$$

where $\omega \in C([0, \infty))$ is a nondecreasing function $\omega \in C([0, \infty))$ with $\omega(0) = 0$, and C is a positive constant.

Proof. Proposition B.12 is a straightforward result of Lemmas B.10, B.11. \square

We finally apply this approximation procedure to give the characterization of the ergodic constant for (B.2) in terms of generalized Mather measures stated in Proposition B.2

Proof of Proposition B.2. Take $\mu \in \tilde{\mathcal{H}}$. By the definition of $\tilde{\mathcal{H}}$, there exist $\{\eta_j\}_{j \in \mathbb{N}} \rightarrow 0$ and $\{\mu^{\eta_j}\}_{j \in \mathbb{N}} \subset \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ such that $\mu^{\eta_j} \in \mathcal{H}^{\eta_j}$, and

$$\mu^{\eta_j} \rightharpoonup \mu \quad \text{weakly in the sense of measures in } \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n).$$

Let v be a Lipschitz viscosity solution to (B.2). For $\alpha, \varepsilon, \delta > 0$, denote by

$$v^{\alpha, \varepsilon, \delta}(x) := \int_{\mathbb{T}^n} \theta^\alpha(y) (v^{\varepsilon+\delta})_\delta(x-y) dy.$$

Then, $v^{\alpha, \varepsilon, \delta} \in C^2(\mathbb{T}^n)$. Let $\tilde{v} := v^{\alpha, \varepsilon, \delta}$, to simplify the notation. Because of Proposition B.12, we have

$$\begin{aligned} \kappa(\alpha, \eta, \delta, \varepsilon) &\geq \int_{\mathbb{T}^n \times \mathbb{R}^n} \left(-\eta_j \Delta \tilde{v} - \text{tr}(A(z) D^2 \tilde{v}) + H(z, D\tilde{v})\right) d\mu^{\eta_j}(z, q) \\ &\geq \int_{\mathbb{T}^n \times \mathbb{R}^n} \left(-\eta_j \Delta \tilde{v} - \text{tr}(A(z) D^2 \tilde{v}) + q \cdot D\tilde{v} - L(z, q)\right) d\mu^{\eta_j}(z, q) \\ &= - \int_{\mathbb{T}^n \times \mathbb{R}^n} L(z, q) d\mu^{\eta_j}(z, q). \end{aligned}$$

We let $\alpha \rightarrow 0$, $j \rightarrow \infty$, and $\varepsilon \rightarrow 0$, in this order, to get

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(z, q) d\mu(z, q) \geq 0 \quad \text{for all } \mu \in \tilde{\mathcal{H}}.$$

For $\varepsilon > 0$, let $(v^\varepsilon, c^\varepsilon) \in C^2(\mathbb{T}^n) \times \mathbb{R}$ solve

$$-\varepsilon \Delta v^\varepsilon - \text{tr}(A(x)D^2v^\varepsilon) + H(x, Dv^\varepsilon) = c^\varepsilon \quad \text{in } \mathbb{T}^n. \quad (\text{B.12})$$

The constant c^ε is unique. Besides, thanks to Assumptions 1-3, we have that v^ε is Lipschitz continuous uniformly in $\varepsilon \in (0, 1)$. Because the ergodic constant c for (1.5) was normalized to be 0, we see that $c^\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$. Let θ^ε solve the associated adjoint equation

$$\begin{cases} -\varepsilon \Delta \theta^\varepsilon - (a^{ij} \theta^\varepsilon)_{x_i x_j} - \text{div}(D_p H(x, Dv^\varepsilon) \theta^\varepsilon) = 0 & \text{in } \mathbb{T}^n, \\ \int_{\mathbb{T}^n} \theta^\varepsilon(x) dx = 1. \end{cases}$$

Next, we define a measure, $\mu^\varepsilon \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$, as follows

$$\int_{\mathbb{T}^n} \psi(x, Dv^\varepsilon) \theta^\varepsilon(x) dx = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, D_q L(x, q)) d\mu^\varepsilon(x, q)$$

for all $\psi \in C_c(\mathbb{T}^n \times \mathbb{R}^n)$.

Multiplying (B.12) by θ^ε , integrating on \mathbb{T}^n , and using the integration by parts, we get

$$\begin{aligned} c^\varepsilon &= \int_{\mathbb{T}^n} \left(-\varepsilon \Delta v^\varepsilon - \text{tr}(A(x)D^2v^\varepsilon) + H(x, Dv^\varepsilon) \right) \theta^\varepsilon dx \\ &= \int_{\mathbb{T}^n} \left(H(x, Dv^\varepsilon) - D_p H(x, Dv^\varepsilon) \cdot Dv^\varepsilon \right) \theta^\varepsilon dx = \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\mu^\varepsilon. \end{aligned}$$

By a similar argument to the one in Lemma 2.2, we see that $\mu^\varepsilon \in \mathcal{H}^\varepsilon$. By extracting a subsequence if necessary, there exists a sequence $\{\varepsilon_j\} \rightarrow 0$ such that

$$\mu^{\varepsilon_j} \rightharpoonup \mu \quad \text{weakly in the sense of measures in } \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$$

as $j \rightarrow \infty$ for some $\mu \in \tilde{\mathcal{H}}$, and

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, q) d\mu = 0.$$

This also shows that $\mu \in \tilde{\mathcal{M}}$, and finishes the proof. \square

DATA STATEMENT

We have no data associate for the submission.

REFERENCES

- [1] H. Amann and M. G. Crandall. On some existence theorems for semi-linear elliptic equations. *Indiana Univ. Math. J.*, 27(5):779–790, 1978.
- [2] S. N. Armstrong and H. V. Tran. Viscosity solutions of general viscous Hamilton-Jacobi equations. *Math. Ann.*, 361(3-4):647–687, 2015.
- [3] G. Barles and P. E. Souganidis. On the large time behavior of solutions of Hamilton-Jacobi equations. *SIAM Journal on Mathematical Analysis*, 31(4):925–939, 2000.
- [4] G. Barles and P. E. Souganidis. Space-time periodic solutions and long-time behavior of solutions to quasi-linear parabolic equations. *SIAM J. Math. Anal.*, 32(6):1311–1323, 2001.
- [5] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.
- [6] P. Bernard and B. Buffoni. Optimal mass transportation and Mather theory. *J. Eur. Math. Soc. (JEMS)*, 9(1):85–121, 2007.
- [7] F. Cagnetti, D. Gomes, H. Mitake, and H. Tran. A new method for large time behavior of degenerate viscous Hamilton-Jacobi equations with convex Hamiltonians. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 32(1):183–200, 2015.
- [8] F. Camilli, I. Capuzzo-Dolcetta, and D. Gomes. Error estimates for the approximation of the effective Hamiltonian. *Appl. Math. Optim.*, 57(1):30–57, 2008.
- [9] P. Cannarsa and P. Cardaliaguet. Hölder estimates in space-time for viscosity solutions of Hamilton-Jacobi equations. *Comm. Pure Appl. Math.*, 63(5):590–629, 2010.
- [10] I. Capuzzo-Dolcetta, F. Leoni, and A. Porretta. Hölder estimates for degenerate elliptic equations with coercive Hamiltonians. *Transactions of the American Mathematical Society*, 362(9):4511–4536, 2010.
- [11] P. Cardaliaguet and L. Silvestre. Hölder continuity to Hamilton-Jacobi equations with superquadratic growth in the gradient and unbounded right-hand side. *Comm. Partial Differential Equations*, 37(9):1668–1688, 2012.
- [12] M. G. Crandall, M. Kocan, P. Soravia, and A. Swiech. On the equivalence of various weak notions of solutions of elliptic PDEs with measurable ingredients. In *Progress in elliptic and parabolic partial differential equations (Capri, 1994)*, volume 350 of *Pitman Res. Notes Math. Ser.*, pages 136–162. Longman, Harlow, 1996.
- [13] A. Davini and A. Siconolfi. A generalized dynamical approach to the large time behavior of solutions of Hamilton-Jacobi equations. *SIAM J. Math. Anal.*, 38(2):478–502 (electronic), 2006.
- [14] S. Dweik, N. Ghoussoub, Y.-H. Kim, and A. Z. Palmer. Stochastic optimal transport with free end time. *Preprint in arXiv:1909.04814*, 2019.
- [15] L. C. Evans. Adjoint and compensated compactness methods for Hamilton-Jacobi PDE. *Arch. Ration. Mech. Anal.*, 197(3):1053–1088, 2010.
- [16] A. Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math.*, 327:267–270, 1998.
- [17] W. Fleming and D. Vermes. Generalized solutions in the optimal control of diffusions. In *Stochastic differential systems, stochastic control theory and applications (Minneapolis, Minn., 1986)*, volume 10 of *IMA Vol. Math. Appl.*, pages 119–127. Springer, New York, 1988.
- [18] W. Fleming and D. Vermes. Convex duality approach to the optimal control of diffusions. *SIAM J. Control Optim.*, 27(5):1136–1155, 1989.
- [19] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.
- [20] D. Gomes. A stochastic analogue of Aubry-Mather theory. *Nonlinearity*, 15(3):581–603, 2002.
- [21] D. Gomes. Duality principles for fully nonlinear elliptic equations. In *Trends in partial differential equations of mathematical physics*, volume 61 of *Progr. Nonlinear Differential Equations Appl.*, pages 125–136. Birkhäuser, Basel, 2005.

- [22] D. Gomes. Generalized Mather problem and selection principles for viscosity solutions and Mather measures. *Adv. Calc. Var.*, 1(3):291–307, 2008.
- [23] D. Gomes and E. Valdinoci. Duality theory, representation formulas and uniqueness results for viscosity solutions of Hamilton-Jacobi equations. In *Dynamics, games and science. II*, volume 2 of *Springer Proc. Math.*, pages 361–386. Springer, Heidelberg, 2011.
- [24] N. Ichihara and H. Ishii. Long-time behavior of solutions of Hamilton-Jacobi equations with convex and coercive Hamiltonians. *Arch. Ration. Mech. Anal.*, 194(2):383–419, 2009.
- [25] H. Ishii, H. Mitake, and H. V. Tran. The vanishing discount problem and viscosity Mather measures. Part 1: The problem on a torus. *J. Math. Pures Appl. (9)*, 108(2):125–149, 2017.
- [26] J.-M. Lasry and P.-L. Lions. A remark on regularization in Hilbert spaces. *Israel J. Math.*, 55(3):257–266, 1986.
- [27] R. M. Lewis and R. B. Vinter. Relaxation of optimal control problems to equivalent convex programs. *J. Math. Anal. Appl.*, 74(2):475–493, 1980.
- [28] O. Ley and V. D. Nguyen. Large time behavior for some nonlinear degenerate parabolic equations. *J. Math. Pures Appl. (9)*, 102(2):293–314, 2014.
- [29] T. Mikami. Marginal problem for semimartingales via duality. *Gakuto International Series. Mathematical Sciences and Applications*, 30:133–152, 2008.
- [30] T. Mikami. Two end points marginal problem by stochastic optimal transportation. *SIAM J. Control Optim.*, 53(4):2449–2461, 2015.
- [31] T. Mikami. Stochastic optimal transport revisited. *Preprint in arXiv:2003.11811v2*, 2020.
- [32] T. Mikami and M. Thieullen. Duality theorem for the stochastic optimal control problem. *Stochastic Process. Appl.*, 116(12):1815–1835, 2006.
- [33] H. Mitake and H. Tran. Dynamical properties of Hamilton-Jacobi equations via the nonlinear adjoint method: large time behavior and discounted approximation. In *Dynamical and geometric aspects of Hamilton-Jacobi and linearized Monge-Ampère equations—VIASM 2016*, volume 2183 of *Lecture Notes in Math.*, pages 125–228. Springer, Cham, 2017.
- [34] H. Mitake and H. Tran. Selection problems for a discount degenerate viscous Hamilton-Jacobi equation. *Adv. Math.*, 306:684–703, 2017.
- [35] H. Mitake and H. Tran. On uniqueness sets of additive eigenvalue problems and applications. *Proc. Amer. Math. Soc.*, 146(11):4813–4822, 2018.
- [36] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.
- [37] H. Tran. Adjoint methods for static Hamilton-Jacobi equations. *Calc. Var. Partial Differential Equations*, 41(3-4):301–319, 2011.
- [38] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

(D. A. Gomes) KING ABDULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY (KAUST), CEMSE DIVISION, THUWAL 23955-6900, SAUDI ARABIA.

Email address: diogo.gomes@kaust.edu.sa

(H. Mitake) GRADUATE SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF TOKYO 3-8-1 KOMABA, MEGURO-KU, TOKYO, 153-8914, JAPAN

Email address: mitake@ms.u-tokyo.ac.jp

(Hung V. Tran) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN MADISON, VAN VLECK HALL, 480 LINCOLN DRIVE, MADISON, WI 53706, USA

Email address: hung@math.wisc.edu