A NEW METHOD FOR LARGE TIME BEHAVIOR OF DEGENERATE VISCOUS HAMILTON–JACOBI EQUATIONS WITH CONVEX HAMILTONIANS

FILIPPO CAGNETTI, DIOGO GOMES, HIROYOSHI MITAKE AND HUNG V. TRAN

ABSTRACT. We investigate large-time asymptotics for viscous Hamilton–Jacobi equations with possibly degenerate diffusion terms. We establish new results on the convergence, which are the first general ones concerning equations which are neither uniformly parabolic nor first order. Our method is based on the nonlinear adjoint method and the derivation of new estimates on long time averaging effects. It also extends to the case of weakly coupled systems.

CONTENTS

1. INTRODUCTION

In this paper we obtain new results on the study of the large time behavior of Hamilton–Jacobi equations with possibly degenerate diffusion terms

$$u_t + H(x, Du) = \operatorname{tr}\left(A(x)D^2u\right) \qquad \text{in } \mathbb{T}^n \times (0, \infty), \qquad (1.1)$$

where $\mathbb{T}^n$ is the n-dimensional torus $\mathbb{R}^n/\mathbb{Z}^n$. Here Du, D^2u are the (spatial) gradient and Hessian of the real-valued unknown function u defined on $\mathbb{T}^n \times [0, \infty)$. The functions $H : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ and $A : \mathbb{T}^n \to \mathbb{M}^{n \times n}_{\text{sym}}$ are the Hamiltonian and the diffusion matrix, respectively, where $\mathbb{M}^{n \times n}_{\text{sym}}$ is the set of $n \times n$ real symmetric matrices. The basic hypotheses that we require are that H is uniformly convex in the second variable, and A is nonnegative definite.

Our goal in this paper is to study the large time behavior of viscosity solutions of (1.1). Namely, we prove that

$$\|u(\cdot, t) - (v - ct)\|_{L^\infty(\mathbb{T}^n)} \to 0 \quad \text{as } t \to \infty, \tag{1.2}$$

where (v, c) is a solution of the *ergodic problem*

$$H(x, Dv) = \text{tr}\left(A(x)D^2v\right) + c \quad \text{in } \mathbb{T}^n. \tag{1.3}$$

In view of the quadratic or superquadratic growth of the Hamiltonian, there exists a unique constant $c \in \mathbb{R}$ such that (1.3) holds true for some $v \in C(\mathbb{T}^n)$ in the viscosity sense. We notice that in the uniformly parabolic case (A is positive definite), v is unique up to additive constants. It is however typically the case that v is not unique even up to additive constants when A is degenerate, which makes the convergence (1.2) delicate and hard to be achieved. We will state clearly the existence result of (1.3), which itself is important, in Section 2.

It is worth emphasizing here that the study of the large-time asymptotics for this type of equations was only available in the literature for the uniformly parabolic case and for the first order case. There was no results on the large-time asymptotics for (1.1) with possibly degenerate diffusion terms up to now as far as the authors know.

In the last decade, a number of authors have studied extensively the large time behavior of solutions of (first order) Hamilton–Jacobi equations (i.e., (1.1) with $A \equiv 0$), where H is coercive. Several convergence results have been established. The first general theorem in this direction was proven by Namah and Roquejoffre in [20], under the assumptions: $p \mapsto H(x, p)$ is convex, $H(x, p) \geq H(x, 0)$ for all $(x, p) \in \mathbb{T}^n \times \mathbb{R}^n$, and $\max_{x \in \mathbb{T}^n} H(x, 0) = 0$. Fathi then gave a breakthrough in this area in [10] by using a dynamical systems approach from the weak KAM theory. Contrary to [20], the results of [10] use uniform convexity and smoothness assumptions on the Hamiltonian but do not require any condition on the structure above. These rely on a deep understanding of the dynamical structure of the solutions and of the corresponding ergodic problem. See also the paper of Fathi and Siconolfi [11] for a beautiful characterization of the Aubry set. Afterwards, Davini and Siconolfi in [7] and Ishii in [12] refined and generalized the approach of Fathi, and studied the asymptotic problem for Hamilton–Jacobi equations on $\mathbb{T}^n$ and on the whole n-dimensional Euclidean space, respectively. Besides, Barles and Souganidis [2] obtained additional results, for possibly non-convex Hamiltonians, by using a PDE method in the context of viscosity solutions. Barles, Ishii and Mitake [1] simplified the ideas in [2] and presented the most general assumptions (up to now). In general, these methods are based crucially on delicate stability results of extremal

curves in the context of the dynamical approach in light of the finite speed of prop-
agation, and of solutions for time large in the context of the PDE approach. It is
also important to point out that the PDE approach in [2, 1] does not work with the
presence of any second order terms.

In the uniformly parabolic setting (i.e., A uniformly positive definite), Barles and
Souganidis [3] proved the long-time convergence of solutions. Their proof relies on
a completely distinct set of ideas from the ones used in the first order case as the
associated ergodic problem has a simpler structure. Indeed, the strong maximum
principle holds, the ergodic problem has a unique solution up to constants. The proof
for the large-time convergence in [3] strongly depends on this fact.

It is clear that all the methods aforementioned (for both the cases $A \equiv 0$ and A
uniformly positive definite) are not applicable for the general degenerate viscous cases
because of the presence of the second order terms and the lack of both the finite speed
of propagation as well as the strong comparison principle. We briefly describe the key
ideas on establishing (1.2) in subsection 1.1. Here the nonlinear adjoint method, which
was introduced by Evans in [8], plays the essential role in our analysis. Our main
results are stated in subsection 1.2.

1.1. **Key Ideas.** Let us now briefly describe the key ideas on establishing (1.2). With-
out loss of generality, we may assume the ergodic constant is 0 henceforth. In order
to understand the limit as $t \to \infty$, we introduce a rescaled problem. For $\varepsilon > 0$, set
$u^\varepsilon(x,t) = u(x,t/\varepsilon)$. Then $(u^\varepsilon)_t(x,t) = \varepsilon^{-1}u_t(x,t/\varepsilon)$, $Du^\varepsilon(x,t) = Du(x,t/\varepsilon)$, and u^ε
solves

$$\begin{cases} \varepsilon u_t^\varepsilon + H(x, Du^\varepsilon) = \mathrm{tr}\left(A(x)D^2u^\varepsilon\right) & \text{in } \mathbb{T}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x), & \text{on } \mathbb{T}^n. \end{cases}$$

By this rescaling, $u^\varepsilon(x,1) = u(x,1/\varepsilon)$ and we can easily see that to prove (1.2) is
equivalent to prove that

$$\|u^\varepsilon(\cdot,1) - v\|_{L^\infty(\mathbb{T}^n)} \to 0 \quad \text{as } \varepsilon \to 0.$$

To show the above, we first introduce the following approximation:

$$\begin{cases} \varepsilon w_t^\varepsilon + H(x, Dw^\varepsilon) = \mathrm{tr}\left(A(x)D^2w^\varepsilon\right) + \varepsilon^4\Delta w^\varepsilon & \text{in } \mathbb{T}^n \times (0,\infty), \\ u^\varepsilon(x,0) = u_0(x), & \text{on } \mathbb{T}^n. \end{cases}$$

Then, we observe that w^ε is smooth, and

$$\|w^\varepsilon(\cdot,1) - u^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} \to 0 \quad \text{as } \varepsilon \to 0.$$

It is thus enough to derive the convergence of $w^\varepsilon(\cdot,1)$ as $\varepsilon \to 0$. To prove this, we show
that

$$\varepsilon\|w_t^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} \to 0 \quad \text{as } \varepsilon \to 0, \tag{1.4}$$

which is a way to prove the convergence (1.2). Indeed, (1.2) is a straightforward
consequence of (1.4) by using the stability of viscosity solutions. We notice that this
principle appears in the papers of Fathi [10], Barles and Souganidis [2] in the case

$A \equiv 0$ in a completely different way. More precisely, Barles and Souganidis [2] first realized the importance of (1.4), and they gave a beautiful proof of the fact that $\max\{u_t, 0\}, \min\{u_t, 0\} \to 0$ as $t \to \infty$ in the viscosity sense. In view of this fact, they succeeded to deal with some cases of non-convex Hamilton–Jacobi equations. On the other hand, we emphasize that the proofs in [10, 2] do not work at all for the second order cases, and therefore, one cannot apply it to (1.1). One of our key contributions in this paper is the establishment of (1.4) in the general setting.

In order to prove (1.4), we use the nonlinear adjoint method introduced by Evans [8] and give new ingredients on the averaging action as $t \to \infty$ (or equivalently as $\varepsilon \to 0$ by rescaling), as clarified below. Let L_{w^ε} be the *formal linearized operator* of the regularized equation around w^ε, i.e.,

$$L_{w^\varepsilon} f := \frac{\partial}{\partial \eta} \Big[\varepsilon(w^\varepsilon + \eta f)_t + H(x, Dw^\varepsilon + \eta Df) - \mathrm{tr}\big(A(x)(D^2 w^\varepsilon + \eta D^2 f)\big) - \varepsilon^4 \Delta(w^\varepsilon + \eta f) \Big]\Big|_{\eta=0},$$

for any $f \in C^2(\mathbb{T}^n \times (0, \infty))$. Then we consider the following adjoint equation:

$$\begin{cases} L_{w^\varepsilon}^* \sigma^\varepsilon = 0 & \text{in } \mathbb{T}^n \times (0, 1), \\ \sigma^\varepsilon(x, 1) = \delta_{x_0} & \text{on } \mathbb{T}^n, \end{cases}$$

where $L_{w^\varepsilon}^*$ is the formal adjoint operator of L_{w^ε}, and δ_{x_0} is the Dirac delta measure at some point $x_0 \in \mathbb{T}^n$. We then see that $\sigma^\varepsilon(\cdot, t)$ is a probability measure for all $t \in (0, 1)$ and conservation of energy holds, namely,

$$\frac{d}{dt} \int_{\mathbb{T}^n} \big[H(x, Dw^\varepsilon) - \mathrm{tr}\big(A(x)D^2 w^\varepsilon\big) - \varepsilon^4 \Delta w^\varepsilon \big] \sigma^\varepsilon(x, t)\, dx = 0.$$

The conservation of energy in particular gives us a different and completely new way to interpret $\varepsilon w_t^\varepsilon(\cdot, 1)$ as

$$\varepsilon w_t^\varepsilon(x_0, 1) = \int_0^1 \int_{\mathbb{T}^n} \big[H(x, Dw^\varepsilon) - \mathrm{tr}\big(A(x)D^2 w^\varepsilon\big) - \varepsilon^4 \Delta w^\varepsilon \big] \sigma^\varepsilon(x, t)\, dx\, dt. \qquad (1.5)$$

The most important part of the paper is then about showing that the right hand side of (1.5) vanishes as $\varepsilon \to 0$, which requires new ideas and estimates (see Lemmas 2.8 and 3.7). We also notice that the averaging action appears implicitly in (1.5) and plays the key role here. More precisely, if we rescale the above integral back to its actual scale, it turns out to be

$$\frac{1}{T} \int_0^T \int_{\mathbb{T}^n} \big[H(x, Dw^\varepsilon) - \mathrm{tr}\big(A(x)D^2 w^\varepsilon\big) - \varepsilon^4 \Delta w^\varepsilon \big] \sigma^\varepsilon(x, t)\, dx\, dt, \qquad (1.6)$$

where $T = 1/\varepsilon \to \infty$.

The nonlinear adjoint method for Hamilton-Jacobi equations was introduced by Evans [8] to study the vanishing viscosity process, and gradient shock structures of viscosity solutions of non convex Hamilton–Jacobi equations. Afterwards, Tran [22] used it to establish a rate of convergence for static Hamilton–Jacobi equations and was able to relax the convexity assumption of the Hamiltonians in some cases. Cagnetti, Gomes and Tran [4] then used it to study the Aubry–Mather theory in the non convex

settings and established the existence of Mather measures. See also [5, 9] for further developments of this new method in the context of Hamilton–Jacobi equations. We notice further that σ^ε here is strongly related to the Mather measures in the context of the weak KAM theory. See [4] for more details.

1.2. **Main Results.** We state the assumptions we use *throughout* the paper as well as our main theorems.

For some given $\theta, C > 0$, we denote by $\mathcal{C}(\theta, C)$ the class of all pairs of (H, A) satisfying

(H1) $H \in C^2(\mathbb{T}^n \times \mathbb{R}^n)$, and $D^2_{pp}H \geq 2\theta I_n$, where I_n is the identity matrix of size n,

(H2) $|D_x H(x, p)| \leq C(1 + |p|^2)$,

(H3) $A(x) = (a^{ij}(x)) \in \mathbb{M}^{n \times n}_{\mathrm{sym}}$ with $A(x) \geq 0$, and $A \in C^2(\mathbb{T}^n)$.

Theorem 1.1 (Main Theorem 1). *Assume that $(H, A) \in \mathcal{C}(\theta, C)$. Let u be the solution of (1.1) with initial data $u(\cdot, 0) = u_0 \in C(\mathbb{T}^n)$. Then there exists $(v, c) \in C(\mathbb{T}^n) \times \mathbb{R}$ such that (1.2) holds, where the pair (v, c) is a solution of the ergodic problem (1.3).*

We also consider the weakly coupled system

$$(u_i)_t + H_i(x, Du_i) + \sum_{j=1}^m c_{ij}u_j = \mathrm{tr}\left(A_i(x)D^2u_i\right) \quad \text{in } \mathbb{T}^n \times (0, \infty), \text{ for } i = 1, \ldots, m. \quad (1.7)$$

where $(H_i, A_i) \in \mathcal{C}(\theta, C)$ is, for each i, the Hamiltonian and diffusion matrix $H_i : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R}$ and $A_i : \mathbb{T}^n \to \mathbb{M}^{n \times n}_{\mathrm{sym}}$ and u_i are the real-valued unknown functions on $\mathbb{T}^n \times [0, \infty)$ for $i = 1, \ldots, m$. The coefficients c_{ij} are given constants for $1 \leq i, j \leq m$ which are assumed to satisfy

(H4) $c_{ii} > 0$, $c_{ij} \leq 0$ for $i \neq j$, $\sum_{j=1}^m c_{ij} = 0$ for any $i = 1, \ldots, m$.

We remark that (H4) ensures that (1.7) is a monotone system.

Under these conditions, we prove

Theorem 1.2 (Main Theorem 2). *Assume that $(H_i, A_i) \in \mathcal{C}(\theta, C)$, and c_{ij} satisfies (H4) for $1 \leq i, j \leq m$. Let $(u_1, \ldots, u_m)$ be the solution of (1.7) with initial data $(u_{01}, \ldots, u_{0m}) \in C(\mathbb{T}^n)^m$. There exists $(v_1, \ldots, v_m, c) \in C(\mathbb{T}^n)^m \times \mathbb{R}$ such that*

$$\|u_i(\cdot, t) - (v_i - ct)\|_{L^\infty(\mathbb{T}^n)} \to 0 \quad \text{as } t \to +\infty, \quad \text{for } i = 1, \ldots m, \quad (1.8)$$

where $(v_1, \ldots, v_m, c)$ is a solution of the ergodic problem for systems:

$$H_i(x, Dv_i) + \sum_{j=1}^m c_{ij}v_j = \mathrm{tr}\left(A_i(x)D^2v_i\right) + c \quad \text{in } \mathbb{T}^n, \text{ for } i = 1, \ldots, m.$$

The study of the large-time behavior of solutions to the weakly coupled system (1.7) of first order cases (i.e. $A_i \equiv 0$ for $1 \leq i \leq m$) was started by [15] and [6] independently under rather restrictive assumptions for Hamiltonians. Recently, Mitake and Tran [17] were able to establish convergent results under rather general assumptions on Hamiltonians. Their proof is based on the dynamical approach, inspired by the papers

[7, 12], together with a new representation formula for the solutions. See [19] for a PDE approach inspired by [2]. We also refer to [16] for a related work on homogenization of weakly coupled systems of Hamilton–Jacobi equations with fast switching rates.

We prove Theorem 1.2 by following the ideas described above. We notice that the coupling terms cause some additional difficulties and are needed to be handled carefully. We in fact establish a new estimate (see Lemma 3.7 (ii) and Subsection 3.4) to control these coupling terms in order to achieve (1.4). This is completely different from the single case.

After this paper was completed, we learnt that Ley and Nguyen [13] obtained recently related convergence results for some specific degenerate parabolic equations. They however assume rather restrictive and technical conditions on the degenerate diffusions so that they could combine the PDE approaches in [2] and [3] to achieve the results. On the other hand, they can deal with a type of fully nonlinear case, which is not included in ours.

This paper is organized as follows: Sections 2 and 3 are devoted to prove Theorems 1.1 and 1.2 respectively. We provide details and explanations to the method described above in subsections 2.1, 2.3, 3.1, and 3.3 and give key estimates in subsections 2.2 and 3.2.

2. Degenerate Viscous Hamilton–Jacobi equations

In this section we study degenerate viscous Hamilton–Jacobi equations. To keep the formulation as simple as possible, we first consider the equation

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

where $u_0 \in C(\mathbb{T}^n)$. Throughout this section we *always* assume that $(H, aI_n) \in \mathcal{C}(\theta, C)$, i.e., a is supposed to be in $C^2(\mathbb{T}^n)$ with $a \geq 0$. We remark that all the results proved for this particular case hold with trivial modifications for the general elliptic operator $\text{tr}\,(A(x)D^2 u)$, except estimate (2.5) and Lemma 2.8 (ii). The corresponding results will be considered in the end of this section.

The next three propositions concern basic existence results, both for (C) and for the associated stationary problem. The proofs are standard, hence omitted. We refer the readers to the companion paper [18] by Mitake and Tran for the detailed proofs of Propositions 2.2 and 2.3.

Proposition 2.1. *Let $u_0 \in C(\mathbb{T}^n)$. There exists a unique solution u of* (C) *which is uniformly continuous on $\mathbb{T}^n \times [0, \infty)$. Furthermore, if $u_0 \in \text{Lip}\,(\mathbb{T}^n)$, then $u \in \text{Lip}\,(\mathbb{T}^n \times [0, \infty))$.*

Proposition 2.2. *There exists a unique constant $c \in \mathbb{R}$ such that the ergodic problem*

$$\text{(E)} \qquad H(x, Dv(x)) = a(x)\Delta v + c \qquad in\ \mathbb{T}^n,$$

admits a solution $v \in \text{Lip}\,(\mathbb{T}^n)$. We call c the ergodic constant of (E).

In order to get the existence of the solutions of the ergodic problem (E), we use the vanishing viscosity method as described by the following Proposition.

Proposition 2.3. *For every $\varepsilon \in (0,1)$ there exists a unique constant $\overline{H}_\varepsilon$ such that the following ergodic problem:*

$$(\mathrm{E})_\varepsilon \qquad H(x, Dv^\varepsilon) = (\varepsilon^4 + a(x))\Delta v^\varepsilon + \overline{H}_\varepsilon \quad \text{in } \mathbb{T}^n$$

has a unique solution $v^\varepsilon \in \mathrm{Lip}\,(\mathbb{T}^n)$ up to some additive constants. In addition,

$$|\overline{H}_\varepsilon - c| \leq C\varepsilon^2, \qquad \|Dv^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \leq C, \tag{2.1}$$

for some positive constant C independent of ε. Here c is the ergodic constant of (E).

In view of the quadratic or superquadratic growth of the Hamiltonian H, we can get (2.1) by the Bernstein method. See the proof of [18, Proposition 1.1] for details. By passing to some subsequences if necessary, $v^\varepsilon - v^\varepsilon(x_0)$ for a fixed $x_0 \in \mathbb{T}^n$ converges uniformly to a Lipschitz function $v : \mathbb{T}^n \to \mathbb{R}$ which is a solution of (E) as $\varepsilon \to 0$.

We prove Theorem 1.1 in a sequence of subsections by using the method described in Introduction.

2.1. Regularizing Process and Proof of Theorem 1.1.

We only need to study the case where $c = 0$, where c is the ergodic constant, by replacing, if necessary, H and $u(x,t)$ by $H - c$ and $u(x,t) + ct$, respectively. Therefore, from now on, we always assume that $c = 0$ in this section. Also, without loss of generality we can assume that $u_0 \in \mathrm{Lip}\,(\mathbb{T}^n)$, since the general case $u_0 \in C(\mathbb{T}^n)$ can be obtained by a standard approximation argument. In particular, thanks to Proposition 2.1, we see that u is Lipschitz continuous on $\mathbb{T}^n \times [0, \infty)$.

As stated in Introduction, we consider a rescaled problem. Setting $u^\varepsilon(x,t) = u(x, t/\varepsilon)$ for $\varepsilon > 0$, where u is the solution of (C), one can easily check that u^ε satisfies

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

Notice however that in this way we do not have a priori uniform Lipschitz estimates on ε, since the Lipschitz bounds on u give us that

$$\|u_t^\varepsilon\|_{L^\infty(\mathbb{T}^n \times [0,1])} \leq C/\varepsilon, \qquad \|Du^\varepsilon\|_{L^\infty(\mathbb{T}^n \times [0,1])} \leq C. \tag{2.2}$$

In general, the function u^ε is only Lipschitz continuous. For this reason, we add a viscosity term to $(\mathrm{C})_\varepsilon$, and we consider the regularized equation

$$(\mathrm{A})_\varepsilon^\eta \qquad \begin{cases} \varepsilon w_t^{\varepsilon,\eta} + H(x, Dw^{\varepsilon,\eta}) = (a(x) + \eta)\Delta w^{\varepsilon,\eta} & \text{in } \mathbb{T}^n \times (0, \infty), \\ w^{\varepsilon,\eta}(x,0) = u_0(x) & \text{on } \mathbb{T}^n, \end{cases}$$

for $\eta > 0$. The advantage of considering $(\mathrm{A})_\varepsilon^\eta$ lies in the fact that the solution is smooth, and this will allow us to use the nonlinear adjoint method. The adjoint

equation corresponding to $(A)_\varepsilon^\eta$ is

$$(AJ)_\varepsilon^\eta \qquad \begin{cases} -\varepsilon\sigma_t^{\varepsilon,\eta} - \operatorname{div}(D_pH(x, Dw^{\varepsilon,\eta})\sigma^{\varepsilon,\eta}) = \Delta\big((a(x) + \eta)\sigma^{\varepsilon,\eta}\big) & \text{in } \mathbb{T}^n \times (0,1), \\ \sigma^{\varepsilon,\eta}(x,1) = \delta_{x_0} & \text{on } \mathbb{T}^n, \end{cases}$$

where δ_{x_0} is the Dirac delta measure at some point $x_0 \in \mathbb{T}^n$.

Lemma 2.4 (Elementary Properties of $\sigma^{\varepsilon,\eta}$). *We have $\sigma^{\varepsilon,\eta} \geq 0$ and*

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

This is a straightforward result of adjoint operator and easy to check. Heuristically, the vanishing viscosity method gives that the rate of convergence of $w^{\varepsilon,\eta}$ to w^ε as $\eta \to 0$ is

$$\sqrt{\text{viscosity coefficient}}/(\text{the coefficient of } w_t^{\varepsilon,\eta})$$

and therefore, we naturally expect that we need to choose $\eta = \varepsilon^\alpha$ with $\alpha > 2$. We mostly choose $\eta = \varepsilon^4$ hereinafter, and therefore we specially denote $(A)_\varepsilon^{\varepsilon^4}$, $(AJ)_\varepsilon^{\varepsilon^4}$ by $(A)_\varepsilon$, $(AJ)_\varepsilon$ and $w^{\varepsilon,\varepsilon^4}$, $\sigma^{\varepsilon,\varepsilon^4}$ by w^ε, σ^ε.

The proof of Theorem 1.1 is based on the two following results, Proposition 2.5 and Theorem 2.7.

Proposition 2.5. *Let w^ε be the solution of $(A)_\varepsilon$. There exists $C > 0$ independent of ε such that*

$$\|w^\varepsilon(\cdot,1)\|_{C^1(\mathbb{T}^n)} \leq C, \quad \|u^\varepsilon(\cdot,1) - w^\varepsilon(\cdot,1)\|_{L^\infty(\mathbb{T}^n)} \leq C\varepsilon.$$

The proof of Proposition 2.5 can be derived using standard arguments. Nevertheless, we give the proof below, since some of the estimates involved will be used later.

Before starting the proof, we state a basic property of the function $a \in C^2(\mathbb{T}^n)$.

Lemma 2.6. *There exists a constant $C > 0$ such that*

$$|Da(x)|^2 \leq Ca(x) \quad \text{for all } x \in \mathbb{T}^n. \tag{2.3}$$

Proof. In view of [21, Theorem 5.2.3], $a^{1/2} \in \operatorname{Lip}(\mathbb{T}^n)$. It is then immediate to get (2.3) by noticing that, for $a(x) > 0$,

$$D\left(a^{1/2}\right)(x) = \frac{Da(x)}{2a^{1/2}(x)}. \qquad \square$$

Proof of Proposition 2.5. Since we are assuming that the ergodic constant for (E) is 0 now and H is quadratic or superquadratic on p, we can easily get the first estimate. We only prove the second one.

Let $w^{\varepsilon,\eta}$ be the solution of $(A)_\varepsilon^\eta$ and set $\varphi(x,t) = |Dw^{\varepsilon,\eta}|^2/2$. Then, φ satisfies

$$\varepsilon\varphi_t + D_pH \cdot D\varphi + D_xH \cdot Dw^{\varepsilon,\eta} = (\eta + a)(\Delta\varphi - |D^2w^{\varepsilon,\eta}|^2) + (Da \cdot Dw^{\varepsilon,\eta})\Delta w^{\varepsilon,\eta}.$$

We next notice that for $\delta > 0$ small enough, in light of Lemma 2.6,

$$a_{x_k}w_{x_k}^{\varepsilon,\eta}\Delta w^{\varepsilon,\eta} \leq C|Da| \cdot |\Delta w^{\varepsilon,\eta}| \leq \frac{C}{\delta} + \delta|Da|^2|D^2w^{\varepsilon,\eta}|^2 \leq C + \frac{1}{2}a|D^2w^{\varepsilon,\eta}|^2. \tag{2.4}$$

Hence

$$\varepsilon\varphi_t + D_pH \cdot D\varphi + \frac{1}{2}(\eta + a)|D^2w^{\varepsilon,\eta}|^2 \leq (\eta + a)\Delta\varphi + C.$$

Multiply the above by σ^ε and integrate over $[0,1] \times \mathbb{T}^n$ to yield

$$\int_0^1 \int_{\mathbb{T}^n} (a(x) + \eta)|D^2w^{\varepsilon,\eta}|^2\sigma^{\varepsilon,\eta}\,dx\,dt \leq C \tag{2.5}$$

for some $C > 0$.

Note that $w^{\varepsilon,\eta}$ is differentiable with respect to η by standard regularity results for elliptic equations. Differentiating the equation in $(A)_\varepsilon^\eta$ with respect to η, we get

$$\varepsilon\left(w_\eta^{\varepsilon,\eta}\right)_t + D_pH(x, Dw^{\varepsilon,\eta}) \cdot Dw_\eta^{\varepsilon,\eta} = \Delta w^{\varepsilon,\eta} + (a(x) + \eta)\Delta w_\eta^{\varepsilon,\eta}, \quad \text{in } \mathbb{T}^n,$$

where f_η denotes the derivative of a function f with respect to the parameter η. Multiplying the above by $\sigma^{\varepsilon,\eta}$ and integrating by parts on $[0,1] \times \mathbb{T}^n$ yield

$$\varepsilon w_\eta^{\varepsilon,\eta}(x_0, 1) = \int_0^1 \int_{\mathbb{T}^n} \Delta w^{\varepsilon,\eta}\sigma^{\varepsilon,\eta}\,dx\,dt,$$

where we used the fact that $w_\eta^{\varepsilon,\eta}(x, 0) \equiv 0$. Thanks to (2.5), by the Hölder inequality

$$\varepsilon|w_\eta^{\varepsilon,\eta}(x_0, 1)| \leq C\left(\int_0^1 \int_{\mathbb{T}^n} |D^2w^{\varepsilon,\eta}|^2\sigma^{\varepsilon,\eta}\,dx\,dt\right)^{1/2} \leq \frac{C}{\sqrt{\eta}}.$$

By choosing properly the point x_0 we have thus

$$\|w_\eta^{\varepsilon,\eta}(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} \leq \frac{C}{\varepsilon\sqrt{\eta}},$$

which gives

$$\|w^{\varepsilon,\eta}(\cdot, 1) - u^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} = \|w^{\varepsilon,\eta}(\cdot, 1) - w^{\varepsilon,0}(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} \leq \frac{C\sqrt{\eta}}{\varepsilon}.$$

Finally, observing that $w^\varepsilon = w^{\varepsilon,\varepsilon^4}$, choosing $\eta = \varepsilon^4$ we get the result. $\qquad\square$

Next theorem gives (1.4) in the special case of problem $(A)_\varepsilon$.

Theorem 2.7. *We have*

$$\lim_{\varepsilon \to 0} \varepsilon\|w_t^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} = 0.$$

More precisely, there exists a positive constant C, independent of ε, such that

$$\varepsilon\|w_t^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} = \|H(\cdot, Dw^\varepsilon(\cdot, 1)) - (\varepsilon^4 + a(\cdot))\Delta w^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} \leq C\varepsilon^{1/4}.$$

The proof of Theorem 2.7 is postponed to the end of this section. We can now give the proof of Theorem 1.1.

Proof of Theorem 1.1. By Proposition 2.5, we can choose a sequence $\{\varepsilon_m\} \to 0$ such that $\{w^{\varepsilon_m}(\cdot, 1)\}$ converges uniformly to a continuous function v. In view of Theorem 2.7, v is a solution of (E), and thus a (time independent) solution of the equation in $(C)_\varepsilon$. We let $t_m = 1/\varepsilon_m$ and use Proposition 2.5 to deduce that

$$\|u(\cdot, t_m) - v\|_{L^\infty(\mathbb{T}^n)} \to 0 \text{ as } m \to \infty.$$

Let us show that the limit does not depend on the sequence $\{t_m\}_{m \in \mathbb{N}}$. Now, for any $x \in \mathbb{T}^n$, $t > 0$ such that $t_m \le t < t_{m+1}$, we use the comparison principle to yield that

$$|u(x, t) - v(x)| \le \|u(\cdot, t_m + (t - t_m)) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)} \le \|u(\cdot, t_m) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)}.$$

Thus,

$$\lim_{t \to \infty} |u(x, t) - v(x)| \le \lim_{m \to \infty} \|u(\cdot, t_m) - v(\cdot)\|_{L^\infty(\mathbb{T}^n)} = 0,$$

which gives the conclusion. □

2.2. **Convergence Mechanisms: Degenerate Equations.** We show now the following key lemma, which provides integral bounds on first and second order derivatives of the difference $w^\varepsilon - v^\varepsilon$ on the support of σ^ε and a, where v^ε is a solution of $(E)_\varepsilon$, and w^ε and σ^ε are solutions of

$$(A)_\varepsilon \qquad \begin{cases} \varepsilon w_t^\varepsilon + H(x, Dw^\varepsilon) = (a(x) + \varepsilon^4)\Delta w^\varepsilon & \text{in } \mathbb{T}^n \times (0, \infty), \\ w^\varepsilon(x, 0) = u_0(x) & \text{on } \mathbb{T}^n, \end{cases}$$

$$(AJ)_\varepsilon \qquad \begin{cases} -\varepsilon \sigma_t^\varepsilon - \operatorname{div}(D_p H(x, Dw^\varepsilon)\sigma^\varepsilon) = \Delta\big((a(x) + \varepsilon^4)\sigma^\varepsilon\big) & \text{in } \mathbb{T}^n \times (0, 1), \\ \sigma^\varepsilon(x, 1) = \delta_{x_0} & \text{on } \mathbb{T}^n, \end{cases}$$

respectively.

Lemma 2.8 (Key Estimates). *There exists a positive constant C, independent of ε, such that the following hold:*

(i) $\displaystyle \int_0^1 \int_{\mathbb{T}^n} \left(\frac{1}{\varepsilon} |D(w^\varepsilon - v^\varepsilon)|^2 + \varepsilon^7 |D^2(w^\varepsilon - v^\varepsilon)|^2 \right) \sigma^\varepsilon \, dx \, dt \le C,$

(ii) $\displaystyle \int_0^1 \int_{\mathbb{T}^n} a^2(x) |D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \le C\sqrt{\varepsilon}.$

Proof. Subtracting equation $(A)_\varepsilon$ from $(E)_\varepsilon$, thanks to the uniform convexity of H, we get

$$\begin{aligned} 0 &= \varepsilon(v^\varepsilon - w^\varepsilon)_t + H(x, Dv^\varepsilon) - H(x, Dw^\varepsilon) - (\varepsilon^4 + a(x))\Delta(v^\varepsilon - w^\varepsilon) - \overline{H}_\varepsilon \\ &\ge \varepsilon(v^\varepsilon - w^\varepsilon)_t + D_p H(x, Dw^\varepsilon) \cdot D(v^\varepsilon - w^\varepsilon) + \theta |D(v^\varepsilon - w^\varepsilon)|^2 \\ &\quad - (\varepsilon^4 + a(x))\Delta(v^\varepsilon - w^\varepsilon) - \overline{H}_\varepsilon. \end{aligned}$$

Multiply the above inequality by σ^ε and integrate by parts on $[0,1] \times \mathbb{T}^n$ to deduce that

$$
\theta \int_0^1 \int_{\mathbb{T}^n} |D(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \leq \overline{H}_\varepsilon - \int_0^1 \int_{\mathbb{T}^n} \varepsilon((v^\varepsilon - w^\varepsilon)\sigma^\varepsilon)_t \, dx dt
$$

$$
+ \int_0^1 \int_{\mathbb{T}^n} \left[\varepsilon \sigma_t^\varepsilon + \operatorname{div}(D_p H(x, Dw^\varepsilon)\sigma^\varepsilon) + \Delta((\varepsilon^4 + a)\sigma^\varepsilon) \right] (v - w) \, dx dt
$$

$$
= \overline{H}_\varepsilon + \varepsilon \left[\int_{\mathbb{T}^n} (w^\varepsilon - v^\varepsilon)\sigma^\varepsilon \, dx \right]_{t=0}^{t=1}
$$

$$
= \overline{H}_\varepsilon + \varepsilon(w^\varepsilon(x_0, 1) - v^\varepsilon(x_0)) - \varepsilon \int_{\mathbb{T}^n} (u_0(x) - v^\varepsilon(x))\sigma^\varepsilon(x, 0) \, dx
$$

$$
= \overline{H}_\varepsilon + \varepsilon w^\varepsilon(x_0, 1) - \varepsilon \int_{\mathbb{T}^n} (v^\varepsilon(x_0) - v^\varepsilon(x))\sigma^\varepsilon(x, 0) \, dx - \varepsilon \int_{\mathbb{T}^n} u_0(x)\sigma^\varepsilon(x, 0) \, dx
$$

$$
\leq \overline{H}_\varepsilon + C\varepsilon + C\varepsilon \|Dv^\varepsilon\|_{L^\infty(\mathbb{T}^n)} - \varepsilon \int_{\mathbb{T}^n} u_0(x)\sigma^\varepsilon(x, 0) \, dx \leq C\varepsilon,
$$

where we used Propositions 2.5, 2.3 (recall that we set $c = 0$) in the last two inequalities. We hence get

$$
\int_0^1 \int_{\mathbb{T}^n} |D(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \leq C\varepsilon, \tag{2.6}
$$

which is the first part of (i).

Next, subtract $(A)_\varepsilon$ from $(E)_\varepsilon$ and differentiate with respect to x_i to get

$$
\varepsilon(v^\varepsilon - w^\varepsilon)_{x_i t} + D_p H(x, Dv^\varepsilon) \cdot Dv_{x_i}^\varepsilon - D_p H(x, Dw^\varepsilon) \cdot Dw_{x_i}^\varepsilon
$$
$$
+ H_{x_i}(x, Dv^\varepsilon) - H_{x_i}(x, Dw^\varepsilon) - (\varepsilon^4 + a)\Delta(v^\varepsilon - w^\varepsilon)_{x_i} - a_{x_i}\Delta(v^\varepsilon - w^\varepsilon) = 0.
$$

Let $\varphi(x, t) = |D(v^\varepsilon - w^\varepsilon)|^2/2$. Multiplying the last identity by $(v^\varepsilon - w^\varepsilon)_{x_i}$ and summing up with respect to i, we achieve that

$$
\varepsilon \varphi_t + D_p H(x, Dw^\varepsilon) \cdot D\varphi + \left[\left(D_p H(x, Dv^\varepsilon) - D_p H(x, Dw^\varepsilon) \right) \cdot Dv_{x_i}^\varepsilon \right] (v_{x_i}^\varepsilon - w_{x_i}^\varepsilon)
$$
$$
+ \left(D_x H(x, Dv^\varepsilon) - D_x H(x, Dw^\varepsilon) \right) \cdot D(v^\varepsilon - w^\varepsilon) - (\varepsilon^4 + a(x)) \left(\Delta\varphi - |D^2(v^\varepsilon - w^\varepsilon)|^2 \right)
$$
$$
- \left[Da \cdot D(v^\varepsilon - w^\varepsilon) \right] \Delta(v^\varepsilon - w^\varepsilon) = 0.
$$

By using various bounds on the above as in the proof of Proposition 2.5, we derive that

$$
\varepsilon \varphi_t + D_p H(x, Dw^\varepsilon) \cdot D\varphi - (\varepsilon^4 + a(x))\Delta\varphi + (\varepsilon^4 + a(x)/2)|D^2(v^\varepsilon - w^\varepsilon)|^2
$$
$$
\leq C + C(|D^2 v^\varepsilon| + 1)|D(v^\varepsilon - w^\varepsilon)|^2. \tag{2.7}
$$

The last term in the right hand side of (2.7) is a dangerous term. We now take advantage of (2.5) and (2.6) to handle it. Using the fact that $\|Dv^\varepsilon\|_{L^\infty}$ and $\|Dw^\varepsilon\|_{L^\infty}$

are bounded, we have

$$C|D^2v^\varepsilon| \, |D(v^\varepsilon - w^\varepsilon)|^2 \leq C|D^2(v^\varepsilon - w^\varepsilon)| \, |D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon| \, |D(v^\varepsilon - w^\varepsilon)|^2$$

$$\leq \frac{\varepsilon^4}{2}|D^2(v^\varepsilon - w^\varepsilon)|^2 + \frac{C}{\varepsilon^4}|D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon|. \qquad (2.8)$$

Combine (2.7) and (2.8) to deduce

$$\varepsilon\varphi_t + D_pH(x, Dw^\varepsilon) \cdot D\varphi - (\varepsilon^4 + a(x))\Delta\varphi + \frac{\varepsilon^4}{2}|D^2(v^\varepsilon - w^\varepsilon)|^2$$

$$\leq C|D(v^\varepsilon - w^\varepsilon)|^2 + \frac{C}{\varepsilon^4}|D(v^\varepsilon - w^\varepsilon)|^2 + C|D^2w^\varepsilon|. \quad (2.9)$$

We multiply (2.9) by σ^ε, integrate over $[0,1] \times \mathbb{T}^n$, to yield that, in light of (2.5) and (2.6),

$$\varepsilon^4 \int_0^1 \int_{\mathbb{T}^n} |D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \leq C\varepsilon + \frac{C}{\varepsilon^4}\varepsilon + C\int_0^1 \int_{\mathbb{T}^n} |D^2w^\varepsilon|\sigma^\varepsilon \, dx \, dt$$

$$\leq \frac{C}{\varepsilon^3} + C\Big(\int_0^1 \int_{\mathbb{T}^n} |D^2w^\varepsilon|^2\sigma^\varepsilon \, dx \, dt\Big)^{1/2}\Big(\int_0^1 \int_{\mathbb{T}^n} \sigma^\varepsilon \, dx \, dt\Big)^{1/2} \leq \frac{C}{\varepsilon^3} + \frac{C}{\varepsilon^2} \leq \frac{C}{\varepsilon^3}.$$

Finally, we prove (ii). Setting $\psi(x,t) = a(x)|D(v^\varepsilon - w^\varepsilon)(x,t)|^2/2 = a(x)\varphi(x,t)$ and multiplying (2.7) by $a(x)$ we get

$$\varepsilon\psi_t + D_pH(x, Dw^\varepsilon) \cdot D\psi - (D_pH(x, Dw^\varepsilon) \cdot Da)\varphi - (\varepsilon^4 + a(x))\Delta\psi$$

$$+ (\varepsilon^4 + a(x))(\Delta a\varphi + 2Da \cdot D\varphi) + a(x)(\varepsilon^4 + a(x)/2)|D^2(v^\varepsilon - w^\varepsilon)|^2$$

$$\leq Ca(x)(|D^2v^\varepsilon| + 1)|D(v^\varepsilon - w^\varepsilon)|^2.$$

We use the facts that Da, Δa are bounded on $\mathbb{T}^n$ to simplify the above as follows

$$\varepsilon\psi_t + D_pH(x, Dw^\varepsilon) \cdot D\psi - (\varepsilon^4 + a(x))\Delta\psi + a(x)(\varepsilon^4 + a(x)/2)|D^2(v^\varepsilon - w^\varepsilon)|^2$$

$$\leq C|D(v^\varepsilon - w^\varepsilon)|^2 - 2(\varepsilon^4 + a(x))Da \cdot D\varphi + Ca(x)|D^2v^\varepsilon| \, |D(v^\varepsilon - w^\varepsilon)|^2. \quad (2.10)$$

Next, we have to control the last two terms on the right hand side of (2.10). Observe first that for $\delta > 0$ small enough

$$2|(\varepsilon^4 + a(x))Da \cdot D\varphi| \leq C(\varepsilon^4 + a(x))|Da| \, |D^2(v^\varepsilon - w^\varepsilon)| \, |D(v^\varepsilon - w^\varepsilon)|$$

$$\leq \delta(\varepsilon^4 + a(x))|Da|^2|D^2(v^\varepsilon - w^\varepsilon)|^2 + \frac{C}{\delta}|D(v^\varepsilon - w^\varepsilon)|^2$$

$$\leq \frac{1}{8}(\varepsilon^4 + a(x))a(x)|D^2(v^\varepsilon - w^\varepsilon)|^2 + C|D(v^\varepsilon - w^\varepsilon)|^2, \qquad (2.11)$$

where we used Lemma 2.6 in the last inequality. On the other hand,

$$
\begin{aligned}
&a(x)|D^2 v^\varepsilon| \ |D(v^\varepsilon - w^\varepsilon)|^2 \\
&\leq a(x)|D^2 w^\varepsilon| \ |D(v^\varepsilon - w^\varepsilon)|^2 + a(x)|D^2(v^\varepsilon - w^\varepsilon)| \ |D(v^\varepsilon - w^\varepsilon)| \\
&\leq \sqrt{\varepsilon} a(x)|D^2 w^\varepsilon|^2 + \frac{C}{\sqrt{\varepsilon}}|D(v^\varepsilon - w^\varepsilon)|^2 + \frac{a(x)^2}{8}|D^2(v^\varepsilon - w^\varepsilon)|^2 + C|D(v^\varepsilon - w^\varepsilon)|^2.
\end{aligned}
\tag{2.12}
$$

We combine (2.10), (2.11), and (2.12) to deduce that

$$
\begin{aligned}
\varepsilon\psi_t + D_p H(x, Dw^\varepsilon) \cdot D\psi &- (\varepsilon^4 + a(x))\Delta\psi + \frac{a(x)^2}{4}|D^2(v^\varepsilon - w^\varepsilon)|^2 \\
&\leq (C + C\varepsilon^{-1/2})|D(v^\varepsilon - w^\varepsilon)|^2 + \varepsilon^{1/2} a(x)|D^2 w^\varepsilon|^2.
\end{aligned}
$$

We multiply the above inequality by σ^ε, integrate over $\mathbb{T}^n \times [0,1]$ and use (2.5) to yield the result. $\qquad\square$

Remark 1. Let us give some comments on the estimates in Lemma 2.8.
1. In case $a \equiv 0$, estimate (i) gives us much better control of $D(w^\varepsilon - v^\varepsilon)$ and $D^2(w^\varepsilon - v^\varepsilon)$ on the support of σ^ε. More precisely, a priori estimates only imply that $D(w^\varepsilon - v^\varepsilon)$ and $\varepsilon^4 \Delta(w^\varepsilon - v^\varepsilon)$ are bounded. By using the adjoint equation, we can get further formally that $\varepsilon^{-1/2} D(w^\varepsilon - v^\varepsilon)$ and $\varepsilon^{7/2} D^2(w^\varepsilon - v^\varepsilon)$ are bounded on the support of σ^ε. We notice that while we need to require the uniform convexity of H to obtain the first term in (i), the second term is achieved without any convexity assumption on H. A version of the second term in (i) was first derived by Evans [8].
2. If the equation in (C) is uniformly parabolic, i.e., $a(x) > 0$ for all $x \in \mathbb{T}^n$, then the second term of (i) is not needed anymore as estimate (ii) is much stronger. On the other hand, if a is degenerate, then (ii) only provides estimation of $|D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon$ on the support of a and it is hence essential to use the second term in (i) to control the part where $a = 0$.

2.3. **Averaging Action and Proof of Theorem** 2.7.

Lemma 2.9 (Conservation of Energy). *The following hold:*
(i) $\displaystyle \frac{d}{dt}\int_{\mathbb{T}^n} (H(x, Dw^\varepsilon) - (\varepsilon^4 + a(x))\Delta w^\varepsilon)\sigma^\varepsilon \, dx = 0,$

(ii) $\displaystyle \varepsilon w_t^\varepsilon(x_0, 1) = \int_0^1 \int_{\mathbb{T}^n} (H(x, Dw^\varepsilon) - (\varepsilon^4 + a(x))\Delta w^\varepsilon)\sigma^\varepsilon \, dx \, dt.$

We stress the fact that identity Lemma 2.9 (ii) is extremely important. As stated in Introduction, if we scale back the time, the integral in the right hand side becomes (1.6), that is the averaging action as $t \to \infty$. Relation (ii) together with Lemma 2.8 allow us to conclude the proof of Theorem 2.7.

Proof. We only need to prove (i) as (ii) follows directly from (i). This is a straightforward result of adjoint operators and comes from a direct calculation:

$$
\frac{d}{dt} \int_{\mathbb{T}^n} (H(x, Dw^\varepsilon) - (\varepsilon^4 + a(x))\Delta w^\varepsilon)\sigma^\varepsilon \, dx
$$

$$
= \int_{\mathbb{T}^n} (D_p H(x, Dw^\varepsilon) \cdot Dw_t^\varepsilon - (\varepsilon^4 + a(x))\Delta w_t^\varepsilon)\sigma^\varepsilon \, dx
$$

$$
+ \int_{\mathbb{T}^n} (H(x, Dw^\varepsilon) - (\varepsilon^4 + a(x))\Delta w^\varepsilon)\sigma_t^\varepsilon \, dx
$$

$$
= - \int_{\mathbb{T}^n} \Big(\operatorname{div} \big(D_p H(x, Dw^\varepsilon)\sigma^\varepsilon \big) + \Delta(\varepsilon^4 + a(x))\sigma^\varepsilon) \Big) w_t^\varepsilon \, dx - \int_{\mathbb{T}^n} \varepsilon w_t^\varepsilon \sigma_t^\varepsilon \, dx = 0. \quad \square
$$

We now can give the proof of Theorem 2.7, which is the main principle to achieve large time asymptotics, by using the averaging action above and the key estimates in Lemma 2.8.

Proof of Theorem 2.7. Let us first choose x_0 such that

$$
|\varepsilon w_t^\varepsilon(x_0, 1)| = |H(x_0, Dw^\varepsilon(x_0, 1)) - (\varepsilon^4 + a(x_0))\Delta w^\varepsilon(x_0, 1)|
$$

$$
= \|H(\cdot, Dw^\varepsilon(\cdot, 1)) - (\varepsilon^4 + a(\cdot))\Delta w^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)}.
$$

Thanks to Lemma 2.9 and Proposition 2.3,

$$
\varepsilon \|w_t^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} = \|H(\cdot, Dw^\varepsilon(\cdot, 1)) - (\varepsilon^4 + a(\cdot))\Delta w^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)}
$$

$$
= \left| \int_0^1 \int_{\mathbb{T}^n} (H(x, Dw^\varepsilon) - (\varepsilon^4 + a)\Delta w^\varepsilon)\sigma^\varepsilon \, dx \, dt \right|
$$

$$
\leq \int_0^1 \int_{\mathbb{T}^n} |(H(x, Dw^\varepsilon) - (\varepsilon^4 + a)\Delta w^\varepsilon) - (H(x, Dv^\varepsilon) - (\varepsilon^4 + a)\Delta v^\varepsilon)|\sigma^\varepsilon \, dx \, dt + |\overline{H}_\varepsilon|
$$

$$
\leq \int_0^1 \int_{\mathbb{T}^n} \left[C|D(w^\varepsilon - v^\varepsilon)| + (\varepsilon^4 + a)|\Delta(w^\varepsilon - v^\varepsilon)| \right] \sigma^\varepsilon \, dx \, dt + C\varepsilon^2.
$$

We finally use the Hölder inequality and Lemma 2.8 to get

$$
\varepsilon \|w_t^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)}
$$

$$
\leq C \left(\int_0^1 \int_{\mathbb{T}^n} |D(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2} + C\varepsilon^4 \left(\int_0^1 \int_{\mathbb{T}^n} |D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2}
$$

$$
+ C \left(\int_0^1 \int_{\mathbb{T}^n} a^2(x)|D^2(w^\varepsilon - v^\varepsilon)|^2 \sigma^\varepsilon \, dx \, dt \right)^{1/2} + C\varepsilon^2 \leq C\varepsilon^{1/4}. \quad \square
$$

2.4. **General Case.** In this subsection we consider the general case (1.1). As pointed out before we only need to address the analogs to estimate (2.5) and Lemma 2.8 (ii).

These are

$$\int_0^1 \int_{\mathbb{T}^n} \left[a^{ij}(x) w^\varepsilon_{x_i x_k} w^\varepsilon_{x_j x_k} + \varepsilon^4 |D^2 w^\varepsilon|^2 \right] \sigma^\varepsilon \, dx dt \leq C, \tag{2.13}$$

$$\int_0^1 \int_{\mathbb{T}^n} a^{ij}(x) a^{ll}(x) (v^\varepsilon - w^\varepsilon)_{x_i x_k} (v^\varepsilon - w^\varepsilon)_{x_j x_k} \sigma^\varepsilon \, dx dt \leq C\sqrt{\varepsilon}. \tag{2.14}$$

In the previous formulas, and throughout this section we will use Einstein's convention except in a few places where the summation signs are explicitly written to avoid ambiguities. The proofs of (2.13) and (2.14) follow the same lines as before. Recall that w^ε and σ^ε satisfy

$$\begin{cases} \varepsilon w^\varepsilon_t + H(x, Dw^\varepsilon) = a^{ij}(x) w^\varepsilon_{x_i x_j} + \varepsilon^4 \Delta w^\varepsilon & \text{in } \mathbb{T}^n \times (0, \infty), \\ w^\varepsilon(x, 0) = u_0(x) & \text{on } \mathbb{T}^n, \end{cases} \tag{2.15}$$

and

$$\begin{cases} -\varepsilon \sigma^\varepsilon_t - \operatorname{div}(D_p H(x, Dw^\varepsilon) \sigma^\varepsilon) = \partial_{x_i x_j} \left(a^{ij}(x) \sigma^\varepsilon \right) + \varepsilon^4 \Delta \sigma^\varepsilon & \text{in } \mathbb{T}^n \times (0, 1), \\ \sigma^\varepsilon(x, 1) = \delta_{x_0} & \text{on } \mathbb{T}^n, \end{cases}$$

and v^ε is a solution to the approximate cell problem

$$H(x, Dv^\varepsilon) = a^{ij}(x) v^\varepsilon_{x_i x_j} + \varepsilon^4 \Delta v^\varepsilon \quad \text{in } \mathbb{T}^n. \tag{2.16}$$

We need the following estimates, which are from [21, Lemma 3.2.3],

$$|Da^{ij}| \leq C \left((a^{ii})^{1/2} + (a^{jj})^{1/2} \right) \qquad \text{for } 1 \leq i, j \leq n, \tag{2.17}$$

$$(\operatorname{tr}(A_{x_k} S))^2 \leq C \operatorname{tr}(SAS) \qquad \text{for } S \in \mathbb{M}^{n \times n}_{\text{sym}}, \ 1 \leq k \leq n, \tag{2.18}$$

for some constant C depending only on n and $\|D^2 A\|_{L^\infty(\mathbb{T}^n)}$.

We address first (2.13). To do so, as before, setting $\varphi := |Dw^\varepsilon|^2/2$, we obtain

$$\varepsilon \varphi_t + D_p H(x, Dw^\varepsilon) \cdot D\varphi \leq a^{ij}(\varphi_{x_i x_j} - w^\varepsilon_{x_i x_k} w^\varepsilon_{x_j x_k}) + \varepsilon^4(\Delta \varphi - |D^2 w^\varepsilon|^2) + a^{ij}_{x_k} w^\varepsilon_{x_i x_j} w^\varepsilon_{x_k} + C.$$

The key term to estimate is $a^{ij}_{x_k} w^\varepsilon_{x_i x_j} w^\varepsilon_{x_k}$, as all the others do not pose any further problem. This is done as follows with help of (2.18):

$$a^{ij}_{x_k} w^\varepsilon_{x_i x_j} w^\varepsilon_{x_k} = \operatorname{tr}(A_{x_k} D^2 w^\varepsilon) w^\varepsilon_{x_k} \leq \frac{1}{2} \operatorname{tr}(D^2 w^\varepsilon A D^2 w^\varepsilon) + C = \frac{1}{2} a^{ij} w^\varepsilon_{x_i x_k} w^\varepsilon_{x_j x_k} + C.$$

Then the proof follows exactly as before.

Concerning estimate (2.14), as before we subtract (2.15) from (2.16) and differentiate with respect to x_k to get

$$\varepsilon (v^\varepsilon - w^\varepsilon)_{x_k t} + D_p H(x, Dv^\varepsilon) \cdot Dv^\varepsilon_{x_k} - D_p H(x, Dw^\varepsilon) \cdot Dw^\varepsilon_{x_k}$$
$$+ H_{x_k}(x, Dv^\varepsilon) - H_{x_k}(x, Dw^\varepsilon) - (\varepsilon^4 \delta^{ij} + a^{ij})((v^\varepsilon - w^\varepsilon)_{x_k})_{x_i x_j} - a^{ij}_{x_k}(v^\varepsilon - w^\varepsilon)_{x_i x_j} = 0,$$

where δ^{ij} is the Kronecker delta. We multiply the previous equation by $a^{ll}(v^\varepsilon - w^\varepsilon)_{x_k}$ and set $\psi = a^{ll}|D(v^\varepsilon - w^\varepsilon)|^2/2$. After some tedious computations we conclude that the

main order term from which (2.14) follows is

$$a^{ij}a^{ll}(v^\varepsilon - w^\varepsilon)_{x_i x_k}(v^\varepsilon - w^\varepsilon)_{x_j x_k} = \text{tr}\,(A)\text{tr}\,(D^2(v^\varepsilon - w^\varepsilon)AD^2(v^\varepsilon - w^\varepsilon))$$

$$= \left(\sum_l d^l\right)\sum_{k,m}\left(\sum_i \sqrt{d^m}p^{mi}(v^\varepsilon - w^\varepsilon)_{x_i x_k}\right)^2, \quad (2.19)$$

where A is diagonalized as $A = P^T DP$ with $D = \text{diag}\,\{d^1,\ldots,d^n\}$ with $d^i \geq 0$, and $P^T P = I_n$. As before, a number of error terms need to be controlled. The procedure is completely analogous, except for two error terms which need to be addressed in a slightly different way. These are

$$a^{ij}a^{ll}_{x_i}(v^\varepsilon - w^\varepsilon)_{x_j x_k}(v^\varepsilon - w^\varepsilon)_{x_k} \quad \text{and} \quad a^{ij}_{x_k}a^{ll}(v^\varepsilon - w^\varepsilon)_{x_i x_j}(v^\varepsilon - w^\varepsilon)_{x_k}.$$

The first term is handled in the following way: using (2.17) we have

$$a^{ij}a^{ll}_{x_i}(v^\varepsilon - w^\varepsilon)_{x_j x_k}(v^\varepsilon - w^\varepsilon)_{x_k} = a^{ll}_{x_i}p^{mi}p^{mj}d^m(v^\varepsilon - w^\varepsilon)_{x_j x_k}(v^\varepsilon - w^\varepsilon)_{x_k}$$

$$\leq C\left(\sum_l \sqrt{d^l}\right)\sum_{k,m}d^m\left|p^{mj}(v^\varepsilon - w^\varepsilon)_{x_j x_k}\right|\left|D(v^\varepsilon - w^\varepsilon)\right|$$

$$\leq \frac{1}{4}\left(\sum_l d^l\right)\sum_{k,m}\left(\sum_j \sqrt{d^m}p^{mj}(v^\varepsilon - w^\varepsilon)_{x_j x_k}\right)^2 + C|D(v^\varepsilon - w^\varepsilon)|^2.$$

Concerning the second term, using (2.18) we obtain

$$a^{ij}_{x_k}a^{ll}(v^\varepsilon - w^\varepsilon)_{x_i x_j}(v^\varepsilon - w^\varepsilon)_{x_k} = \text{tr}\,(A)\text{tr}\,(A_{x_k}D^2(v^\varepsilon - w^\varepsilon))(v^\varepsilon - w^\varepsilon)_{x_k}$$

$$\leq \frac{1}{4}\text{tr}\,(A)\text{tr}\,(D^2(v^\varepsilon - w^\varepsilon)AD^2(v^\varepsilon - w^\varepsilon)) + C|D(v^\varepsilon - w^\varepsilon)|^2.$$

This shows therefore that the two error terms are well controlled by (2.19).

3. Weakly Coupled Systems of Hamilton–Jacobi Equations

In this section, we prove Theorem 1.2. Our proof follows along the lines of the scalar case, together with additional estimates for the coupling terms. To simplify the presentation we start first with the following weakly coupled system of first-order Hamilton–Jacobi equations:

$$\text{(SC)} \quad \begin{cases} (u_1)_t + H_1(x, Du_1) + u_1 - u_2 = 0 & \text{in } \mathbb{T}^n \times (0,\infty), \\ (u_2)_t + H_2(x, Du_2) + u_2 - u_1 = 0 & \text{in } \mathbb{T}^n \times (0,\infty), \\ u_i(x,0) = u_{0i}(x) & \text{on } \mathbb{T}^n \text{ for } i = 1,2. \end{cases}$$

Throughout this section we *always* assume that $u_{0i} \in C(\mathbb{T}^n)$ and that the pairs $(H_i, 0) \in \mathcal{C}(\theta, C)$ for $i = 1, 2$.

We observe that almost all results we prove for (SC) are valid with trivial modifications for general weakly coupled systems with possibly degenerate diffusion terms. Indeed, if we combine the arguments in Section 2 and Section 3 below, then we can immediately get the result on the large-time behavior for weakly coupled systems of

degenerate viscous Hamilton–Jacobi equations (1.7). We present the simplest case here since we want to concentrate on the difficulty coming from the coupling terms of the system. We derive new estimates for the coupling terms (see part (ii) of Lemma 3.7 and Subsection 3.4), which help us to control the large time average on the coupling terms and achieve the desired results.

We first state the basic existence results for (SC) and for the associated stationary problem. The proofs of the next three propositions are standard, hence omitted.

Proposition 3.1. *Let $(u_{01}, u_{02}) \in C(\mathbb{T}^n)^2$. There exists a unique solution (u_1, u_2) of (SC) which is uniformly continuous on $\mathbb{T}^n \times [0, \infty)$. Furthermore, if $u_{0i} \in \mathrm{Lip}\,(\mathbb{T}^n)$, then $u_i \in \mathrm{Lip}\,(\mathbb{T}^n \times [0, \infty))$ for $i = 1, 2$.*

We refer to [5, Section 4] for the following results.

Proposition 3.2. *There exists a unique constant $c \in \mathbb{R}$ such that the ergodic problem:*

$$\text{(SE)} \qquad \begin{cases} H_1(x, Dv_1) + v_1 - v_2 = c & \text{in } \mathbb{T}^n, \\ H_2(x, Dv_2) + v_2 - v_1 = c & \text{in } \mathbb{T}^n, \end{cases}$$

has a solution $(v_1, v_2) \in \mathrm{Lip}\,(\mathbb{T}^n)^2$. We call c the ergodic constant of (SE).

Proposition 3.3. *For every $\varepsilon > 0$ sufficiently small there exists a unique $\overline{H}_\varepsilon \in \mathbb{R}$ such that the following ergodic problem:*

$$\text{(SE)}_\varepsilon \qquad \begin{cases} H_1(x, Dv_1^\varepsilon) + v_1^\varepsilon - v_2^\varepsilon = \varepsilon^4 \Delta v_1^\varepsilon + \overline{H}_\varepsilon & \text{in } \mathbb{T}^n, \\ H_2(x, Dv_2^\varepsilon) + v_2^\varepsilon - v_1^\varepsilon = \varepsilon^4 \Delta v_2^\varepsilon + \overline{H}_\varepsilon & \text{in } \mathbb{T}^n, \end{cases}$$

has a unique solution $(v_1^\varepsilon, v_2^\varepsilon) \in \mathrm{Lip}\,(\mathbb{T}^n)^2$ up to additional constants. In addition,

$$|\overline{H}_\varepsilon - c| \leq C\varepsilon^2, \quad \|Dv_i^\varepsilon\|_{L^\infty(\mathbb{T}^n)} \leq C, \ \text{for } i = 1, 2,$$

for some positive constant C independent of ε. Here c is the ergodic constant of (SE).

Without loss of generality, we may assume $c = 0$ as in Section 2 henceforth.

3.1. **Regularizing Process and Proof of Theorem** 1.2. In the following we will assume that (u_1, u_2) is Lipschitz on $\mathbb{T}^n \times [0, \infty)$, as it was done in Section 2. Once again, we will follow the method stated in Introduction.

We perform a change of time scale. For $\varepsilon > 0$, let us set $u_i^\varepsilon(x, t) = u_i(x, t/\varepsilon)$, which is the solution of

$$\text{(SC)}_\varepsilon \qquad \begin{cases} \varepsilon(u_1^\varepsilon)_t + H_1(x, Du_1^\varepsilon) + u_1^\varepsilon - u_2^\varepsilon = 0 & \text{in } \mathbb{T}^n \times (0, \infty), \\ \varepsilon(u_2^\varepsilon)_t + H_2(x, Du_2^\varepsilon) + u_2^\varepsilon - u_1^\varepsilon = 0 & \text{in } \mathbb{T}^n \times (0, \infty), \\ u_i^\varepsilon(x, 0) = u_{0i}(x) & \text{on } \mathbb{T}^n \text{ for } i = 1, 2, \end{cases}$$

and we approximate (SC) by adding viscosity terms to the equations:

$$\text{(SA)}_\varepsilon \qquad \begin{cases} \varepsilon(w_1^\varepsilon)_t + H_1(x, Dw_1^\varepsilon) + w_1^\varepsilon - w_2^\varepsilon = \varepsilon^4 \Delta w_1^\varepsilon & \text{in } \mathbb{T}^n \times (0, \infty), \\ \varepsilon(w_2^\varepsilon)_t + H_2(x, Dw_2^\varepsilon) + w_2^\varepsilon - w_1^\varepsilon = \varepsilon^4 \Delta w_2^\varepsilon & \text{in } \mathbb{T}^n \times (0, \infty), \\ w_i^\varepsilon(x, 0) = u_{0i}(x) & \text{on } \mathbb{T}^n \text{ for } i = 1, 2. \end{cases}$$

We can conclude the proof of Theorem 1.2 with the following two results.

Lemma 3.4. *Let $(w_1^\varepsilon, w_2^\varepsilon)$ be the solution of $(SA)_\varepsilon$. There exists $C > 0$ independent of ε such that $\|w_i^\varepsilon(\cdot, 1)\|_{C^1(\mathbb{T}^n)} \leq C$, $\|u_i^\varepsilon(\cdot, 1) - w_i^\varepsilon(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} \leq C\varepsilon$ for $i = 1, 2$.*

Theorem 3.5. *We have*

$$\lim_{\varepsilon \to 0} \max_{i=1,2} \varepsilon \|(w_i^\varepsilon)_t(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} = 0.$$

3.2. **Convergence Mechanisms: Weakly Coupled Systems.** The adjoint system corresponding to $(SA)_\varepsilon$ is

$$(SAJ)_\varepsilon \quad \begin{cases} -\varepsilon(\sigma_1^\varepsilon)_t - \operatorname{div}(D_p H_1(x, Dw_1^\varepsilon)\sigma_1^\varepsilon) + \sigma_1^\varepsilon - \sigma_2^\varepsilon = \varepsilon^4 \Delta \sigma_1^\varepsilon & \text{in } \mathbb{T}^n \times (0, 1), \\ -\varepsilon(\sigma_2^\varepsilon)_t - \operatorname{div}(D_p H_2(x, Dw_2^\varepsilon)\sigma_2^\varepsilon) + \sigma_2^\varepsilon - \sigma_1^\varepsilon = \varepsilon^4 \Delta \sigma_2^\varepsilon & \text{in } \mathbb{T}^n \times (0, 1), \\ \sigma_i^\varepsilon(x, 1) = \delta_{ik}\delta_{x_0} & \text{on } \mathbb{T}^n \text{ for } i = 1, 2. \end{cases}$$

where $\delta_{ik} = 1$ if $i = k$ and $\delta_{ik} = 0$ if $i \neq k$, and $x_0 \in \mathbb{T}^n$ and $k \in \{1, 2\}$ are to be chosen later. Notice that for any choice of k, either $\sigma_1^\varepsilon(\cdot, 1) = 0$ or $\sigma_2^\varepsilon(\cdot, 1) = 0$. Let us record some elementary properties of $(\sigma_1^\varepsilon, \sigma_2^\varepsilon)$ first.

Lemma 3.6 (Elementary properties of $(\sigma_1^\varepsilon, \sigma_2^\varepsilon)$). *We have $\sigma_i^\varepsilon \geq 0$ for $i = 1, 2$ and*

$$\sum_{i=1}^2 \int_{\mathbb{T}^n} \sigma_i^\varepsilon(x, t) \, dx = 1 \quad \text{for all } t \in [0, 1].$$

We next derive key integral bounds for $(w_1^\varepsilon, w_2^\varepsilon)$, $(v_1^\varepsilon, v_2^\varepsilon)$ and their derivatives on the supports of $(\sigma_1^\varepsilon, \sigma_2^\varepsilon)$.

Lemma 3.7 (Key estimates for weakly coupled systems). *The followings hold true:*

(i) $\displaystyle \int_0^1 \int_{\mathbb{T}^n} \sum_{i=1}^2 \left(\frac{1}{\varepsilon} |D(w_i^\varepsilon - v_i^\varepsilon)|^2 + \varepsilon^7 |D^2(w_i^\varepsilon - v_i^\varepsilon)|^2 \right) \sigma_i^\varepsilon \, dx \, dt \leq C,$

(ii) $\displaystyle \int_0^1 \int_{\mathbb{T}^n} [(w_1^\varepsilon - v_1^\varepsilon) - (w_2^\varepsilon - v_2^\varepsilon)]^2 (\sigma_1^\varepsilon + \sigma_2^\varepsilon) \, dx \, dt \leq C\varepsilon.$

Lemma 3.7 (ii) is a new observation on the study of weakly coupled systems, which gives us the large time average control on the coupling terms. This is actually the key point in the derivation of the main result for systems (Theorem 1.2 and Theorem 3.5) as one can see in the proof of Lemma 3.8.

Proof. We will only prove (ii), since part (i) can be derived by repeating the proof of Lemma 2.8.

Thanks to Lemma 3.4, we can always add to the pair $(v_1^\varepsilon, v_2^\varepsilon)$ an arbitrarily large constant C (independent of ε) such that

$$2C \geq v_i^\varepsilon \geq w_i^\varepsilon \text{ in } \mathbb{T}^n, \text{ for } i = 1, 2. \tag{3.1}$$

Let $\varphi_i = (v_i^\varepsilon - w_i^\varepsilon)^2/2$ for $i = 1, 2$. Subtract the first equation of $(\text{SA})_\varepsilon$ from the first equation of $(\text{SE})_\varepsilon$, and multiply the result by $v_1^\varepsilon - w_1^\varepsilon$ to get

$$\varepsilon(v_1^\varepsilon - w_1^\varepsilon)(v_1^\varepsilon - w_1^\varepsilon)_t + (v_1^\varepsilon - w_1^\varepsilon)(H_1(x, Dv_1^\varepsilon) - H_1(x, Dw_1^\varepsilon))$$
$$+ (v_1^\varepsilon - w_1^\varepsilon)^2 - (v_1^\varepsilon - w_1^\varepsilon)(v_2^\varepsilon - w_2^\varepsilon) = \varepsilon^4(v_1^\varepsilon - w_1^\varepsilon)\Delta(v_1^\varepsilon - w_1^\varepsilon) + \overline{H}_\varepsilon(v_1^\varepsilon - w_1^\varepsilon).$$

We employ the convexity of H_1 and (3.1) to deduce that

$$\varepsilon(\varphi_1)_t + D_p H_1(x, Dw_1^\varepsilon) \cdot D\varphi_1 + \varphi_1 - \varphi_2$$
$$+ \frac{1}{2}[(v_1^\varepsilon - w_1^\varepsilon) - (v_2^\varepsilon - w_2^\varepsilon)]^2 \leq \varepsilon^4 \Delta\varphi_1 - \varepsilon^4|D(v_1^\varepsilon - w_1^\varepsilon)|^2 + C|\overline{H}_\varepsilon|. \quad (3.2)$$

Similarly,

$$\varepsilon(\varphi_2)_t + D_p H_2(x, Dw_2^\varepsilon) \cdot D\varphi_2 + \varphi_2 - \varphi_1$$
$$+ \frac{1}{2}[(v_1^\varepsilon - w_1^\varepsilon) - (v_2^\varepsilon - w_2^\varepsilon)]^2 \leq \varepsilon^4 \Delta\varphi_2 - \varepsilon^4|D(v_2^\varepsilon - w_2^\varepsilon)|^2 + C|\overline{H}_\varepsilon|. \quad (3.3)$$

Multiplying (3.2), (3.3) by σ_1^ε, σ_2^ε respectively, and integrating by parts

$$\frac{1}{2}\int_0^1 \int_{\mathbb{T}^n} [(v_1^\varepsilon - w_1^\varepsilon) - (v_2^\varepsilon - w_2^\varepsilon)]^2(\sigma_1^\varepsilon + \sigma_2^\varepsilon)\, dx\, dt$$
$$\leq -\sum_{i=1}^{2} \varepsilon \left[\int_{\mathbb{T}^n} \varphi_i \sigma_i^\varepsilon\, dx\right]_{t=0}^{t=1} + C|\overline{H}_\varepsilon| - \varepsilon^4 \sum_{i=1,2} \int_{\mathbb{T}^n} |D(v_i^\varepsilon - w_i^\varepsilon)|^2 \sigma_i^\varepsilon\, dx \leq C\varepsilon,$$

which implies (ii). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3.3. **Averaging Action and Proof of Theorem 3.5.** For each $i \in \{1, 2\}$, setting $j = 3 - i$ we have $\{i, j\} = \{1, 2\}$. The following result concerning conservation of energy and averaging action is analogous to Lemma 2.9 and therefore we omit the proof.

Lemma 3.8 (Conservation of Energy for weakly coupled systems). *The following hold:*

(i) $\dfrac{d}{dt}\displaystyle\int_{\mathbb{T}^n} \sum_{i=1}^{2}(H_i(x, Dw_i^\varepsilon) + w_i^\varepsilon - w_j^\varepsilon - \varepsilon^4 \Delta w_i^\varepsilon)\sigma_i^\varepsilon\, dx = 0.$

(ii) $-(k-1)\varepsilon(w_1^\varepsilon)_t(x_0, 1) - (2-k)\varepsilon(w_2^\varepsilon)_t(x_0, 1)$

$$= \int_0^1 \int_{\mathbb{T}^n} \sum_{i=1}^{2}(H_i(x, Dw_i^\varepsilon) + w_i^\varepsilon - w_j^\varepsilon - \varepsilon^4 \Delta w_i^\varepsilon)\sigma_i^\varepsilon\, dx\, dt,$$

where $k = 1, 2$.

Proof of Theorem 3.5. Without loss of generality, we assume that there exists $x_0 \in \mathbb{T}^n$ such that

$$\varepsilon|(w_1^\varepsilon)_t(x_0, 1)| = \varepsilon \max_{i=1,2} \|(w_i^\varepsilon)_t(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)}.$$

We then choose $k = 1$ in $(\text{SAJ})_\varepsilon$ and use Lemma 3.8 to get

$$
\varepsilon \max_{i=1,2} \|(w_i^\varepsilon)_t(\cdot, 1)\|_{L^\infty(\mathbb{T}^n)} = \left| \int_0^1 \int_{\mathbb{T}^n} \sum_{i=1}^2 (H_i(x, Dw_i^\varepsilon) + w_i^\varepsilon - w_j^\varepsilon - \varepsilon^4 \Delta w_i^\varepsilon) \sigma_i^\varepsilon \, dx \, dt \right|
$$

$$
\leq \left| \int_0^1 \int_{\mathbb{T}^n} \sum_{i=1}^2 \left\{ H_i(x, Dw_i^\varepsilon) + w_i^\varepsilon - w_j^\varepsilon - \varepsilon^4 \Delta w_i^\varepsilon) \right. \right.
$$

$$
\left. \left. - (H_i(x, Dv_i^\varepsilon) + v_i^\varepsilon - v_j^\varepsilon - \varepsilon^4 \Delta v_i^\varepsilon) \right\} \sigma_i^\varepsilon \, dx \, dt \right| + |\overline{H}_\varepsilon|
$$

$$
\leq \int_0^1 \int_{\mathbb{T}^n} \sum_{i=1}^2 \left[C|D(w_i^\varepsilon - v_i^\varepsilon)| + \varepsilon^4 |\Delta(w_i^\varepsilon - v_i^\varepsilon)| \right] \sigma_i^\varepsilon \, dx \, dt
$$

$$
+ \int_0^1 \int_{\mathbb{T}^n} \left| [(w_1^\varepsilon - v_1^\varepsilon) - (w_2^\varepsilon - v_2^\varepsilon)](\sigma_1^\varepsilon - \sigma_2^\varepsilon) \right| \, dx \, dt + |\overline{H}_\varepsilon|.
$$

Thus,

$$
\left| \int_0^1 \int_{\mathbb{T}^n} \sum_{i=1}^2 (H_i(x, Dw_i^\varepsilon) + w_i^\varepsilon - w_j^\varepsilon - \varepsilon^4 \Delta w_i^\varepsilon) \sigma_i^\varepsilon \, dx \, dt \right|
$$

$$
\leq C \sum_{i=1}^2 \left\{ \left[\int_0^1 \int_{\mathbb{T}^n} |D(w_i^\varepsilon - v_i^\varepsilon)|^2 \sigma_i^\varepsilon \, dx \, dt \right]^{\frac{1}{2}} + \varepsilon^4 \left[\int_0^1 \int_{\mathbb{T}^n} |D^2(w_i^\varepsilon - v_i^\varepsilon)|^2 \sigma_i^\varepsilon \, dx \, dt \right]^{\frac{1}{2}} \right\}
$$

$$
+ \sum_{i=1}^2 \left[\int_0^1 \int_{\mathbb{T}^n} |(w_1^\varepsilon - v_1^\varepsilon) - (w_2^\varepsilon - v_2^\varepsilon)|^2 \sigma_i^\varepsilon \, dx \, dt \right]^{1/2} + |\overline{H}_\varepsilon| \leq C\sqrt{\varepsilon},
$$

where the last inequality follows by using Lemma 3.7.   □

3.4. **General Case.** We address now the general case of systems of m-equations of the form

$$
(u_i)_t + H_i(x, Du_i) + \sum_{j=1}^m c_{ij} u_j = 0 \quad \text{in } \mathbb{T}^n \times (0, \infty),
$$

with $(H_i, 0) \in \mathcal{C}(\theta, C)$ and c_{ij} satisfying (H4) for any $1 \leq i, j \leq m$. As stated before, the key point is to generalize the coupling terms as in part (ii) of Lemma 3.7. More precisely, we show that

$$
\lim_{\varepsilon \to 0} \int_0^1 \int_{\mathbb{T}^n} \sum_{j=1}^m |c_{ij}| \left[(w_j^\varepsilon - v_j^\varepsilon) - (w_i^\varepsilon - v_i^\varepsilon) \right]^2 \sigma_i^\varepsilon \, dx \, dt = 0. \tag{3.4}
$$

Set $\varphi_i = (v_i^\varepsilon - w_i^\varepsilon)^2 / 2$ for $i = 1, \ldots, m$. Then we can compute that

$$
\varepsilon(v_i^\varepsilon - w_i^\varepsilon)(v_i^\varepsilon - w_i^\varepsilon)_t + (v_i^\varepsilon - w_i^\varepsilon)(H_i(x, Dv_i^\varepsilon) - H_i(x, Dw_i^\varepsilon))
$$

$$
+ \sum_{j=1}^m c_{ij}(v_i^\varepsilon - w_i^\varepsilon)(v_j^\varepsilon - w_j^\varepsilon) = \varepsilon^4 (v_i^\varepsilon - w_i^\varepsilon) \Delta(v_i^\varepsilon - w_i^\varepsilon) + \overline{H}_\varepsilon(v_i^\varepsilon - w_i^\varepsilon).
$$

The last term in the right hand side of the above identity can be written as

$$\sum_{j=1}^{m} c_{ij}(v_i^\varepsilon - w_i^\varepsilon)(v_j^\varepsilon - w_j^\varepsilon) = \sum_{j \neq i} |c_{ij}| \left\{ (v_i^\varepsilon - w_i^\varepsilon)^2 - (v_i^\varepsilon - w_i^\varepsilon)(v_j^\varepsilon - w_j^\varepsilon) \right\}$$

$$= \sum_{j \neq i} |c_{ij}| \left\{ \frac{1}{2}(v_i^\varepsilon - w_i^\varepsilon)^2 - \frac{1}{2}(v_j^\varepsilon - w_j^\varepsilon)^2 + \frac{1}{2}[(v_i^\varepsilon - w_i^\varepsilon) - (v_j^\varepsilon - w_j^\varepsilon)]^2 \right\}$$

$$= \sum_{j=1}^{m} c_{ij}\varphi_j + \frac{1}{2}\sum_{j=1}^{m} |c_{ij}| [(v_j^\varepsilon - w_j^\varepsilon) - (v_i^\varepsilon - w_i^\varepsilon)]^2.$$

Hence

$$\varepsilon(\varphi_i)_t + D_p H_i(x, Dw_i^\varepsilon) \cdot D\varphi_i + \sum_{j=1}^{m} c_{ij}\varphi_j$$

$$+ \frac{1}{2}\sum_{j=1}^{m} |c_{ij}| [(v_j^\varepsilon - w_j^\varepsilon) - (v_i^\varepsilon - w_i^\varepsilon)]^2 \leq \varepsilon^4 \Delta\varphi_i - \varepsilon^4 |D(v_i^\varepsilon - w_i^\varepsilon)|^2 + C|\overline{H}_\varepsilon|.$$

Then (3.4) follows immediately.

References

[1] G. Barles, H. Ishii, H. Mitake, *A new PDE approach to the large time asymptotics of solutions of Hamilton-Jacobi equation*, to appear in Bull. Math. Sci.

[2] G. Barles, P. E. Souganidis, *On the large time behavior of solutions of Hamilton–Jacobi equations*, SIAM J. Math. Anal. **31** (2000), no. 4, 925–939.

[3] G. Barles, P. E. Souganidis, *Space-time periodic solutions and long-time behavior of solutions to quasi-linear parabolic equations*, SIAM J. Math. Anal. 32 (2001), no. 6, 1311–1323.

[4] F. Cagnetti, D. Gomes, H. V. Tran. *Aubry-Mather measures in the non convex setting*, SIAM J. Math. Anal. **43** (2011), no. 6, 2601–2629.

[5] F. Cagnetti, D. Gomes, H. V. Tran, *Adjoint methods for obstacle problems and weakly coupled systems of PDE*, ESAIM Control Optim. Calc. Var. **19** (2013), no. 3, 754–779.

[6] F. Camilli, O. Ley, P. Loreti and V. Nguyen, *Large time behavior of weakly coupled systems of first-order Hamilton–Jacobi equations*, NoDEA Nonlinear Differential Equations Appl. **19** (2012), 719–749.

[7] A. Davini, A. Siconolfi, *A generalized dynamical approach to the large-time behavior of solutions of Hamilton–Jacobi equations*, SIAM J. Math. Anal. **38** (2006), no. 2, 478–502.

[8] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton–Jacobi PDE*, Archive for Rational Mechanics and Analysis **197** (2010), 1053–1088.

[9] L. C. Evans, *Envelopes and nonconvex Hamilton–Jacobi equations*, to appear in Calculus of Variations and PDE.

[10] A. Fathi, *Sur la convergence du semi-groupe de Lax-Oleinik*, C. R. Acad. Sci. Paris Sér. I Math. **327** (1998), no. 3, 267–270.

[11] A. Fathi, A. Siconolfi, *Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, Invent. Math. **155** (2004), no. 2, 363–388.

[12] H. Ishii, *Asymptotic solutions for large-time of Hamilton–Jacobi equations in Euclidean n space*, Ann. Inst. H. Poincaré Anal. Non Linéaire, **25** (2008), no 2, 231–266.

[13] O. Ley, V. D. Nguyen, Large time behavior for some nonlinear degenerate parabolic equations, preprint. (arXiv:1306.0748)

[14] P.-L., Lions, G. Papanicolaou, S. R. S. Varadhan, Homogenization of Hamilton–Jacobi equations, unpublished work (1987).

[15] H. Mitake, H. V. Tran, *Remarks on the large-time behavior of viscosity solutions of quasi-monotone weakly coupled systems of Hamilton–Jacobi equations*, Asymptot. Anal., **77** (2012), 43–70.

[16] H. Mitake, H. V. Tran, *Homogenization of weakly coupled systems of Hamilton–Jacobi equations with fast switching rates*, to appear in Archive for Rational Mechanics and Analysis.

[17] H. Mitake, H. V. Tran, *A dynamical approach to the large-time behavior of solutions to weakly coupled systems of Hamilton–Jacobi equations*, to appear in Journal de Mathématiques Pures et Appliquées.

[18] H. Mitake, H. V. Tran, *Large-time behavior for obstacle problems for degenerate viscous Hamilton–Jacobi equations*, submitted. (arXiv:1309.4831)

[19] V. D. Nguyen, *Some results on the large time behavior of weakly coupled systems of first-order Hamilton-Jacobi equations*, preprint. (arXiv:1209.5929)

[20] G. Namah, J.-M. Roquejoffre, *Remarks on the long time behaviour of the solutions of Hamilton–Jacobi equations*, Comm. Partial Differential Equations **24** (1999), no. 5-6, 883–893.

[21] D. W. Stroock and S. R. S. Varadhan, Multidimensional diffusion processes. Reprint of the 1997 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2006. xii+338 pp.

[22] H. V. Tran, *Adjoint methods for static Hamilton-Jacobi equations*, Calculus of Variations and PDE **41** (2011), 301–319.

(F. Cagnetti) CENTER FOR MATHEMATICAL ANALYSIS, GEOMETRY, AND DYNAMICAL SYSTEMS, DEPARTAMENTO DE MATEMÁTICA, INSTITUTO SUPERIOR TÉCNICO, 1049-001 LISBOA, PORTUGAL
    E-mail address: `cagnetti@math.ist.utl.pt`

(D. Gomes) CENTER FOR MATHEMATICAL ANALYSIS, GEOMETRY, AND DYNAMICAL SYSTEMS, DEPARTAMENTO DE MATEMÁTICA, INSTITUTO SUPERIOR TÉCNICO, 1049-001 LISBOA, PORTUGAL AND KING ABDULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY (KAUST), CSMSE DIVISION, THUWAL 23955-6900, SAUDI ARABIA
    E-mail address: `dgomes@math.ist.utl.pt`

(H. Mitake) DEPARTMENT OF APPLIED MATHEMATICS, FACULTY OF SCIENCE, FUKUOKA UNIVERSITY, FUKUOKA 814-0180, JAPAN
    E-mail address: `mitake@math.sci.fukuoka-u.ac.jp`

(H. V. Tran) DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF CHICAGO, 5734 S. UNIVERSITY AVENUE, CHICAGO, ILLINOIS 60637, USA
    E-mail address: `hung@math.uchicago.edu`