

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/176381>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

A NEW MONOTONICITY CONDITION FOR ERGODIC BSDEs AND ERGODIC CONTROL WITH SUPER-QUADRATIC HAMILTONIANS

JOE JACKSON AND GECHUN LIANG

ABSTRACT. We establish the existence (and in an appropriate sense uniqueness) of Markovian solutions for ergodic BSDEs under a novel monotonicity condition. Our monotonicity condition allows us to prove existence even when the driver f has arbitrary (in particular super-quadratic) growth in z , which reveals an interesting trade-off between monotonicity and growth for ergodic BSDEs. The technique of proof is to establish a probabilistic representation of the derivative of the Markovian solution, and then use this representation to obtain a-priori estimates. Our study is motivated by applications to ergodic control, and we use our existence result to prove the existence of optimal controls for a class of ergodic control problems with potentially super-quadratic Hamiltonians. We also treat a class of drivers coming from the construction of forward performance processes, and interpret our monotonicity condition in this setting.

1. INTRODUCTION

1.1. **Ergodic BSDEs.** In this paper, we study the ergodic backward stochastic differential equation (EBSDE)

$$\begin{cases} dX_t = b(X_t)dt + dW_t, \\ dY_t = -f(X_t, Z_t)dt + \lambda dt + Z_t \cdot dW_t. \end{cases} \quad (1)$$

Here W is a d -dimensional Brownian motion, and the data consists of a vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and a function $f : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ called the driver. We are interested in Markovian solutions, i.e. triples (u, v, λ) where $u : \mathbb{R}^d \rightarrow \mathbb{R}$, $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$, and $\lambda \in \mathbb{R}$, such that for each $x \in \mathbb{R}^d$ and $0 \leq t \leq T < \infty$, we have

$$Y_t^x = Y_T^x + \int_t^T f(X_s^x, Z_s^x)ds - \lambda(T-t) - \int_t^T Z_s^x \cdot dW_s,$$

where X^x solves $X_t^x = x + \int_0^t b(X_s^x)ds + W_t$ and $(Y^x, Z^x) = (u(X^x), v(X^x))$. We recall that there is a natural connection between the EBSDE (1) and the partial differential

During the preparation of this work the first author has been supported by the National Science Foundation under Grant No. DGE1610403 (2020-2023). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

equation (PDE)

$$\frac{1}{2}\Delta u + b \cdot \nabla u + f(x, \nabla u) = \lambda \text{ in } \mathbb{R}^d, \quad (2)$$

in the sense that if there is a pair (u, λ) with $u : \mathbb{R}^d \rightarrow \mathbb{R}$ and $\lambda \in \mathbb{R}$ satisfying (2), then Itô's formula reveals that $(u, \nabla u, \lambda)$ is a Markovian solution to the EBSDE (1).

1.2. Related literature.

1.2.1. *Probabilistic approaches.* Ergodic BSDEs were first introduced in [11] in an infinite dimensional setting. Applied to the somewhat simpler equation (1), the results of [11] show existence and uniqueness of a Markovian solution when the drift b is dissipative, i.e.

$$(b(x) - b(\bar{x})) \cdot (x - \bar{x}) \leq -\delta|x - \bar{x}|^2 \quad (3)$$

for some $\delta > 0$, and the driver f is Lipschitz. In [7], well-posedness is established under weaker conditions on b , but still with the Lipschitz condition on f . In [14], ergodic BSDEs are used to study the long-time behavior of parabolic PDEs, still in an infinite-dimensional framework. In the finite-dimensional setting, [18] and [6] consider generalizations of the ergodic BSDE (1) involving reflection and jumps, respectively. In [17], [5], and [13] a connection is developed between EBSDEs and forward performance processes. In [17], the state space is finite-dimensional, and the driver f is of quadratic growth in z and satisfies the regularity condition

$$|f(x, z) - f(\bar{x}, z)| \leq C_x(1 + |z|)|x - \bar{x}|. \quad (4)$$

In [17], existence of a Markovian solution is established for (1) under the condition $\delta > C_x$. Similar “sufficiently dissipative” conditions appear in [13] and [5]. Finally, we mention that super-quadratic BSDEs have been studied on a finite horizon, most notable in [8] where a number of results, both positive and negative were obtained.

1.2.2. *Analytical approaches.* The works listed above study the EBSDE (1) through probabilistic methods. There is also a large literature dealing directly with the associated PDE (2) through both probabilistic and analytical techniques. Here the focus is not only on well-posedness of the equation (2), but also on its connection to the behavior of the parabolic problem

$$\begin{cases} \partial_t u + \frac{1}{2}\Delta u + b \cdot \nabla u + f(x, \nabla u) = 0 \text{ in } (0, T) \times \mathbb{R}^d, \\ u(T, x) = g(x), \quad x \in \mathbb{R}^d \end{cases} \quad (5)$$

as $T \rightarrow \infty$. The works [1], [15], [16], and [19] all contain well-posedness results for the PDE (2) when the driver f has quadratic growth in z . But typically the PDE approach requires strict convexity¹ of f in z , see e.g. [15]. The recent work [4] contains some remarkable

¹More precisely, it is typical in the PDE literature to consider the equation $-\frac{1}{2}\Delta v - b \cdot \nabla v + H(x, Dv) = \gamma$ with $H = H(x, z)$ strictly convex in z . Make the transformation $u = -v$, $\lambda = -\gamma$, $f(x, z) = -H(x, -z)$ to see that this is equivalent to (2) with f strictly convex in z .

results on the qualitative behavior of the discounted analogue of (2) with super-quadratic Hamiltonian $\frac{1}{m}|\nabla u|^m$ for $m > 2$, but the focus there is on qualitative behavior rather than general well-posedness results.

1.3. Our results.

1.3.1. *The main existence result.* The purpose of the present paper is to study the well-posedness of the ergodic BSDE (1) when f is just locally Lipschitz, and in particular can have arbitrary growth in z . Our main result is Theorem 2.3, which gives the existence (and in an appropriate sense uniqueness) of a Markovian solution to (1) when f splits as $f = g + h$, and the data (b, g, h) is such that

- b is Lipschitz and dissipative,
- $g = g(x, z)$ satisfies
 - $\sup_{x \in \mathbb{R}^d} |g(x, 0)| < \infty$,
 - g is Lipschitz (in both arguments) on $\mathbb{R}^d \times B_R$ for each $R > 0$
 - the map $z \mapsto \partial_x g(x, z)$ is monotone, in the sense that $(\partial_x g(x, z) - \partial_x g(x, \bar{z})) \cdot (z - \bar{z}) \leq 0$.
- $h = h(x, z)$ satisfies
 - $\sup_{x \in \mathbb{R}^d} |h(x, 0)| < \infty$
 - $x \mapsto h(x, z)$ is Lipschitz, uniformly in z
 - h is Lipschitz (in both arguments) on $\mathbb{R}^d \times B_R$, for each $R > 0$

To the best of our knowledge, Theorem 2.3 is the first existence result for ergodic BSDEs which allows the driver to grow arbitrarily fast in z . Moreover, we emphasize that there is no requirement on the dissipativity constant δ verifying $(b(x) - b(\bar{x})) \cdot (x - \bar{x}) \leq -\delta|x - \bar{x}|^2$ beyond positivity. Thus, while we do not handle the “weak dissipative” case treated in [7], neither do we require that δ be “large enough” as in [17]. We also note that we have established in the Appendix a rigorous connection between the EBSDE (1) and the PDE (2), which allows us to assert that if (u, v, λ) is the unique Markovian solution to (1), then in fact u is a classical solution of the PDE (2) and $v = \nabla u$ a.e.

We remark that Theorem 2.3 can be extended to the case where $dX_t = b(X_t)dt + \kappa dW_t$ for an appropriate $\kappa \in \mathbb{R}^{d \times d}$ in a relatively straightforward manner. In this case, the relevant monotonicity condition becomes $(\partial_x g(x, \kappa z) - \partial_x g(x, \kappa \bar{z})) \cdot (z - \bar{z}) \leq 0$, i.e. we need monotonicity of the map $z \mapsto \partial_x g(x, \kappa z)$ rather than $z \mapsto \partial_x g(x, z)$. We do not include this slight generalization here for the sake of notational simplicity.

1.3.2. *Discussion of the monotonicity condition.* The monotonicity condition

$$(\partial_x g(x, z) - \partial_x g(x, \bar{z})) \cdot (z - \bar{z}) \leq 0 \tag{6}$$

is the main novelty of the paper, and it is worth discussing it in detail. First, recall that when studying the corresponding PDE (2), it is common to assume strict convexity of f

in z , which amounts to the requirement that

$$(\partial_z f(x, z) - \partial_z f(x, \bar{z})) \cdot (z - \bar{z}) > C|z - \bar{z}|^2 \quad (7)$$

for some $C > 0$. Thus the existing PDE approaches require a strict monotonicity condition on $z \mapsto \partial_z f(x, z)$, while our main condition is instead the monotonicity of $z \mapsto \partial_x g(x, z)$. As for the probabilistic perspective, we note that monotonicity conditions are fairly common in the BSDE literature. For example, for infinite horizon BSDEs of the form

$$dY_t = -f(Y_t, Z_t)dt + Z_t \cdot dW_t,$$

the requirement that $(f(y, z) - f(\bar{y}, z)) \cdot (y - \bar{y}) \leq -\delta|y - \bar{y}|^2$ is standard, see [2]. For FBSDEs of the type

$$\begin{cases} dX_t = b(X_t, Y_t, Z_t)dt + \sigma(X_t, Y_t, Z_t)dW_t, \\ dY_t = -f(X_t, Y_t, Z_t)dt + Z_t \cdot dW_t, \quad Y_T = g(X_T), \end{cases}$$

a joint monotonicity property of the maps b, σ, f is often imposed, see Section 8.4 of [20] and the references therein. But the monotonicity condition considered here is unique from probabilistic perspective in that it really involves monotonicity of a certain *derivative* of f , rather than of f itself.

Here are some examples of drivers f which satisfy the hypotheses of our main existence result:

- If f is globally Lipschitz and $f(x, 0)$ is bounded, as in [11] and [7], then it satisfies our condition (take $g = 0$).
- $f(x, z) = g(z) + h(x)$, where g is locally Lipschitz and h is Lipschitz and bounded.
- $f(x, z) = g(x, z) = g_1(x)g_2(z)$, where $d = 1$ and $g_1 : \mathbb{R} \rightarrow \mathbb{R}$ is bounded, Lipschitz, and non-increasing (non-decreasing), $g_2 : \mathbb{R} \rightarrow \mathbb{R}$ is locally Lipschitz and non-decreasing (non-increasing)

1.3.3. Strategy of the proof. The main step in proving Theorem 2.3 is to establish an a-priori estimate of the Lipschitz constant of u , where (u, v, λ) is the Markovian solution of (1). This is done with a mix of probabilistic and analytical techniques. More specifically, we first differentiate the PDE (2) which we expect u to solve in order to find a PDE system for ∇u . Then, we represent ∇u in terms of a certain system of infinite-horizon BSDEs, and finally we prove L^∞ estimates for this BSDE system to obtain the desired estimate on the Lipschitz constant of u . While studying the equations for the derivatives of u is a natural approach to obtain Lipschitz estimates, we point out the following interesting features of our approach:

- From the perspective of PDEs, it might seem more natural to view the equation for each partial derivative $\partial_{x_i} u$ in isolation, rather than viewing the equations together as a system of PDEs for ∇u . In our case, it is essential that we view the equations for ∇u as a system - if not, we cannot take advantage of the dissipativity condition (3) for b .

- Differentiating the PDE for u first and then representing ∇u in terms of a system of infinite-horizon BSDEs (as we do in the present paper) does not lead to the same system of infinite-horizon BSDEs that we (formally) get from directly differentiating the BSDE in the initial condition of X , in the sense of Proposition 4.1 in [9]. In fact, it is not clear that this latter, more probabilistic, approach can be used to obtain the desired estimates. Thus it seems essential that we use a mix of probabilistic and analytical approaches to obtain our main estimate.
- The L^∞ estimate we obtain for the system of BSDEs representing ∇u does not follow directly from existing results for infinite horizon BSDEs as in e.g. [2], except for in the special case $d = 1$. Instead, we must combine existing techniques with a judicious application of Girsanov's Theorem, which in turn is only possible because of the special structure of the PDE system for ∇u .

A formal derivation of the estimate is presented in sub-section 3.1, and a rigorous proof (which requires verifying that u is sufficiently smooth via elliptic regularity theory) is given in sub-section 3.2.

1.3.4. *Application to ergodic control.* In Section 4, we apply our results to a control problem in which a controller chooses a process α taking values in \mathbb{R}^d in order to minimize the cost

$$J(\alpha) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T r(X_s, \alpha_s) ds \right],$$

subject to $dX_t = (b(X_t) + \alpha_t)dt + dW_t$. The inputs to the model are the vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and the running cost $r : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$. We focus on the case that the drift is linear in α and the data is sufficiently smooth because this makes it easier to interpret our monotonicity condition. As shown in [11], an optimal control (in the weak formulation) can be characterized by a BSDE with driver $f(x, z) = H(x, z)$, where H is the Hamiltonian given by

$$\begin{aligned} H(x, z) &:= \inf_a h(x, z, a), \text{ where} \\ h(x, z, a) &:= a \cdot z + r(x, a). \end{aligned} \tag{8}$$

In this setting, we obtain new results on the existence of optimal controls when the Hamiltonian H has potentially superquadratic growth in z .

In Proposition 4.5, we show that the driver H satisfies our conditions when $\partial_a^2 r > 0$ and $\partial_a \partial_x r (\partial_a^2 r(x, z))^{-1} \geq 0$, in the sense that $z^T \partial_a \partial_x r(x, a) (\partial_a^2 r(x, a))^{-1} z \geq 0$ for all $x, a, z \in \mathbb{R}^d$. This is an interesting condition which to the best of the authors' knowledge is new. We first note that this condition is automatically satisfied when $\partial_a \partial_x r = 0$, i.e. when the running cost $r(x, a) = r_0(x) + r_1(a)$ is separable. In this case, we have

$$H(x, z) = \tilde{H}(z) + r_0(x), \text{ where } \tilde{H}(z) = \inf_{a \in \mathbb{R}^d} (a \cdot z + r_1(a)).$$

Thus Theorem 2.3 applies when \tilde{H} is just locally Lipschitz (and in particular can have arbitrary growth) and r_0 is bounded and Lipschitz. Thanks to Proposition 4.2, this implies the existence of optimal controls for this class of ergodic control problems with super-quadratic Hamiltonians. Our result also covers running costs of the form $r(x, a) = \frac{1}{m}|a|^m + r_0(x)$ for r_0 Lipschitz and bounded and $1 < m \leq 2$, which is similar to the case considered in [4] - indeed, this class of examples does not necessarily satisfy $\partial_a^2 r > 0$, but the Hamiltonian is given explicitly by $H(x, z) = \frac{1}{q}|z|^q + r_0(x)$ with $q > 2$ the conjugate of m , which clearly satisfies the hypotheses of our main result.

To interpret the conditions $\partial_a^2 r > 0$ and $\partial_a \partial_x r(x, z) \partial_a^2 g(x, z)^{-1} \geq 0$ when r is not separable, we recall that it is common in studying standard control problems through the stochastic maximum principle to require that the running cost is convex in a and that the Hamiltonian is jointly convex in x and a (see Theorem 4.14 of [3]). In this setting, the requirement that $(x, a) \mapsto h(x, p, a)$ is convex is equivalent to the requirement that the $2d \times 2d$ matrix written in blocks as $\begin{pmatrix} \partial_x^2 r & \partial_x \partial_a r \\ \partial_a \partial_x r & \partial_a^2 r \end{pmatrix}$ is positive definite, and of course the convexity of r in a means that $\partial_a^2 r > 0$. To summarize,

- when approaching a standard control problem with linear drift through the maximum principle, it is typical to assume $\partial_a^2 r > 0$, $\begin{pmatrix} \partial_x^2 r & \partial_x \partial_a r \\ \partial_a \partial_x r & \partial_a^2 r \end{pmatrix} \geq 0$,
- for ergodic control problems, our main monotonicity condition reduces to the requirement that $\partial_a^2 r > 0$, $\partial_a \partial_x r (\partial_a^2 r)^{-1} \geq 0$.

Thus while the conditions $\partial_a^2 r > 0$, $\partial_a \partial_x r (\partial_a^2 r)^{-1} \geq 0$ are somewhat difficult to interpret, they are reminiscent of similar conditions which appear naturally in the control literature. For a concrete class of control problems to which this structural condition applies outside the separated case $r(x, a) = r_0(x) + r_1(a)$, consider the running cost

$$r(x, a) = \frac{1}{2}|a|^2 + c_1(x)a + c_0(x).$$

Then

$$\partial_a^2 r(x, a) = I_{d \times d}, \quad \partial_a \partial_x r = Dc_1(x),$$

with Dc_1 the Jacobian matrix of $c_1 : \mathbb{R}^d \rightarrow \mathbb{R}^d$, so our results apply when c_0 is Lipschitz, c_1 is locally Lipschitz, and $Dc_1(x) \geq 0$.

We show in sub-section 4.2 that Theorem 2.3 yields existence of optimal controls for a risk-sensitive version of the ergodic control problem as well. Indeed, in this case the cost functional is of the form

$$J^\delta(\alpha) = \limsup_{T \rightarrow \infty} \frac{1}{T} \left(\frac{1}{\delta} \ln \mathbb{E}[\exp(\delta \int_0^T r(X_t^x, \alpha_t)] dt) \right) \quad (9)$$

for some $\delta > 0$, and the relevant driver is

$$f(x, z) = H(x, z) + \frac{\delta}{2}|z|^2, \quad (10)$$

with H as above. It is easy to see that the driver in (10) satisfies the monotonicity condition (6) if and only if the driver $H(x, z)$ does, so we can obtain very similar results in this setting as in the non-risk-sensitive case. Finally, we show in sub-section (4.3) that because of the Markovian structure of the ergodic control problem, the weak formulation we consider is actually equivalent to an (arguably) more natural strong formulation.

1.3.5. *Application to forward performance processes.* In [17], EBSDEs are used to construct forward performance processes in factor form. In their setting, the diffusion X is the stochastic factor process, and there is a vector S of stock prices satisfying

$$dS_t^i = S_t^i(\eta^i(X_t)dt + \sum_{j=1}^d \sigma^{ij}(X_t)dW_t^j).$$

Thus the dynamics of the stock prices S^i are impacted by the position of the diffusion X . We refer to [17] for the definition of forward performance processes and their applications. For our purposes, it is enough to note that Proposition 13 of [17] reduces the construction of a forward performance process in factor form to solving the EBSDE (1) with driver

$$f(x, z) = \frac{1}{2}\delta^2 \text{dist}^2(\Pi, \frac{\theta(x) + z}{\delta}) - z^T \theta(x) + \frac{1}{2}|\theta(x)|^2, \quad (11)$$

where $\delta > 0$, $\Pi \subset \mathbb{R}^d$ is closed and convex and θ is the market price of risk vector given by $\theta(x) = \sigma(x)^T[\sigma(x)\sigma(x)^T]^{-1}\eta(x)$. In the present work, we simply take the market price of risk vector $\theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$ as a given, and ask - what conditions on θ and Π guarantee that the driver f appearing in (11) satisfies our conditions? By explicitly computing $\partial_x f$ for f given by (51), we are able to state in Proposition 5.1 a general existence criterion which can be checked on a case-by-case basis.

1.4. **Organization of the paper.** In the remainder of the introduction, we fix notations and conventions. In Section 2 we discuss some preliminaries. In Section 3 we state our main existence result, Theorem 2.3. In sub-section 3.1, we provide a formal derivation of the main a-priori estimate, and sub-section 3.2 is devoted to a rigorous proof of Theorem 2.3. In Sections 4 and 5 we present applications to ergodic control and forward performance processes, respectively. Finally, the Appendix collects some auxiliary results concerning the relevant PDEs and their connection to BSDEs.

1.5. Notation and conventions.

Probabilistic setup. We work on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which hosts a d -dimensional Brownian motion W . We denote by $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ the augmented filtration of the Brownian motion W . We denote by $\mathbb{E}[\cdot]$ expectation with respect to the probability measure \mathbb{P} .

Spaces of processes. We denote by \mathcal{P} the space of progressively measurable processes taking values in some Euclidean space. When necessary, we specify the target in a natural way, e.g. $\mathcal{P}(\mathbb{R}^n)$ denotes the space of progressively measurable processes taking values in \mathbb{R}^n . \mathcal{P}^2 is the space of $Z \in \mathcal{P}$ such that

$$\|Z\|_{\mathcal{P}^2}^2 := \mathbb{E}[\int_0^\infty |Z_t|^2 dt] < \infty$$

For $1 \leq p \leq \infty$, we denote by L^p the usual space of p -integrable random variables, and by \mathcal{S}^p the space of continuous, adapted processes X such that

$$\|X\|_{\mathcal{S}^p} := \left\| \sup_{0 \leq t \leq \infty} |X_t| \right\|_{L^p} < \infty.$$

We also define and notate localized versions of these spaces in the usual way, e.g. \mathcal{P}_{loc}^2 is the space of locally square integrable processes. Finally, we note that we use the standard notation $\mathcal{E}(\cdot)$ for the stochastic exponential, i.e. if M is a martingale then $\mathcal{E}(M)$ denotes the local martingale $\mathcal{E}(M)_t = \exp(M_t - \frac{1}{2}\langle M \rangle_t)$, $\langle M \rangle$ being the quadratic variation of M . Spaces of functions. We say that a function u defined on some open set in \mathbb{R}^d is C^k if it is k -times continuously differentiable. Given an open set $\Omega \subset \mathbb{R}^d$, we define

$$[u]_{C^\alpha(\Omega)} = \sup_{x, y \in \Omega, x \neq y} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

We denote by $C^\alpha(\Omega)$ the space of all functions u defined on Ω and taking values in some Euclidean space such that

$$\|u\|_{C^\alpha(\Omega)} := \|u\|_{L^\infty(\Omega)} + [u]_{C^\alpha(\Omega)} < \infty$$

We denote by $C^{k,\alpha}(\Omega)$ the space of all functions u taking values in some Euclidean space such that

$$\|u\|_{C^{k,\alpha}(\Omega)} := \sup_{|\beta| < k} \|D^\beta u\|_{L^\infty(\Omega)} + \sum_{|\beta|=k} \|D^\beta u\|_{C^\alpha(\Omega)} < \infty.$$

Here $|\beta|$ denotes the order of a multi-index β . We use the shorthand $C^{k,\alpha} = C^{k,\alpha}(\mathbb{R}^d)$. Finally, we define and notate localized versions of all these spaces, e.g. $u \in C_{loc}^\alpha(\Omega)$ means that for all open sets U compactly contained in Ω , $u \in C^\alpha(U)$. On one occasion, we will have need of the Sobolev spaces $H_0^1(\Omega)$, $W^{2,p}(\Omega)$, etc., and we refer to [12] for the relevant definitions. Finally, we note that throughout the paper we will use B_R to denote the unit ball in \mathbb{R}^d .

Notation for stochastic integrals. When $Z \in \mathcal{P}_{loc}^2(\mathbb{R}^d)$, we will write $\int Z \cdot dW$ for the local martingale $\sum_{j=1}^d \int Z^j dW^j$, where Z^j is the j^{th} component of Z . At times, we will also deal with integrands $Z \in \mathcal{P}_{loc}^2(\mathbb{R}^{d \times d})$. In this case, we will write Z^i for the i^{th} row of Z , and then $\int Z dW$ will denote the \mathbb{R}^d -valued local martingale whose i^{th} component is given by $(\int Z dW)^i = \int Z^i dW = \sum_{ij} \int Z^{ij} dW^j$, where now Z^{ij} denotes the element of Z in row i and column j (this is just the usual way of interpreting $Z dW$ though

matrix multiplication). The use of these two different conventions (rather than viewing all integrands as taking values in $\mathbb{R}^{n \times d}$ for some n and always interpreting ZdW by matrix multiplication, for example) is motivated by the fact that we will usually have $Z = \nabla u(X)$ for some $u : \mathbb{R}^d \rightarrow \mathbb{R}$ (in which case the first convention is natural) or $z = Dv(X)$ for some $v : \mathbb{R}^d \rightarrow \mathbb{R}^d$ (in which case the second convention is more natural).

Notation for derivatives. We use the variables $x = (x_1, \dots, x_d)$ and $z = (z_1, \dots, z_d)$ for elements of \mathbb{R}^d . The notation ∂_{x_i} means differentiation in the variable x_i , and similarly for ∂_{z_j} . For $b = b(x) : \mathbb{R}^d \rightarrow \mathbb{R}^d$, we use Db to denote the matrix $(Db)_{ij} = \partial_j b^i$. For a map $f = f(x, z) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, we write $\partial_x^2 f$ for the Hessian of f in the variable x , i.e. $(\partial_x^2 f(x, z))_{ij} = \partial_{x_i} \partial_{x_j} f(x, z)$. Likewise, $\partial_z^2 f(x, z)$ is the Hessian in the variable z , while $\partial_z \partial_x f$ denotes the matrix $(\partial_z \partial_x f(x, z))_{ij} = \partial_{z_j} \partial_{x_i} f(x, z)$. The notation $\partial_x \partial_z f$ is similar. Thus $\partial_z \partial_x f$ is the derivative of the map $z \mapsto \partial_x f(x, z)$, and of course $\partial_z \partial_x f = (\partial_x \partial_z f)^T$.

2. THE MAIN RESULT

We recall the ergodic BSDE

$$\begin{cases} dX_t = b(X_t)dt + dW_t, \\ dY_t = -f(X_t, Z_t)dt + \lambda dt + Z_t \cdot dW_t. \end{cases} \quad (12)$$

The data of the problem consists of a vector field $b = b(x) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and a driver $f = f(x, z) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$. The diffusion X takes values in \mathbb{R}^d , and the unknowns Y and Z take values in \mathbb{R} and \mathbb{R}^d , respectively. We are looking for a **Markovian solution** to (12), i.e. a triple (u, v, λ) with $\lambda \in \mathbb{R}$ and $(u, v) : \mathbb{R}^d \rightarrow \mathbb{R} \times \mathbb{R}^d$ such that for any $x \in \mathbb{R}^d$, the processes $Y^x = u(\cdot, X^x)$, $Z^x = v(\cdot, X^x)$ satisfy

$$Y_t^x = Y_T^x + \int_t^T f(X_s^x, Z_s^x) ds - \lambda(T - t) - \int_t^T Z_s^x \cdot dW_s$$

for all $0 \leq t \leq T < \infty$. Here X^x denotes the solution of the SDE

$$X_t^x = x + \int_0^t b(X_s^x) ds + W_t.$$

We recall also the corresponding PDE

$$\frac{1}{2} \Delta u + b \cdot \nabla u + f(x, \nabla u) = \lambda \text{ in } \mathbb{R}^d. \quad (13)$$

We will make the following assumptions regarding the data b and f .

Assumption 2.1. The vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and the driver $f = g + h : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfy

- (1) (dissipativity and Lipschitz property for b) The vector field b is Lipschitz, and there is a constant $\delta > 0$ such that

$$(b(x) - b(\bar{x})) \cdot (x - \bar{x}) \leq -\delta |x - \bar{x}|^2 \quad (14)$$

holds for all $x, \bar{x} \in \mathbb{R}^d$.

- (2) (growth and regularity for g) The function $g = g(x, z)$ is Lipschitz on $\mathbb{R}^d \times B_R$ for each $R > 0$. Moreover $\sup_{x \in \mathbb{R}^d} |g(x, 0)| < \infty$.
- (3) (monotonicity of $\partial_x g$) The estimate

$$(\partial_x g(x, z) - \partial_x g(x, \bar{z})) \cdot (z - \bar{z}) \leq 0 \quad (15)$$

holds for almost every $(x, z, \bar{z}) \in (\mathbb{R}^d)^3$, where $\partial_x g$ is the weak derivative of g in the variable x .

- (4) (growth and regularity for h) The function $h = h(x, z) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is globally Lipschitz in x , i.e. $|h(x, z) - h(\bar{x}, z)| \leq L|x - \bar{x}|$ for all $x, \bar{x}, z \in \mathbb{R}^d$ and some $L > 0$. Moreover, h is Lipschitz (in both arguments) on $\mathbb{R}^d \times B_R$ for each $R > 0$, and $\sup_{x \in \mathbb{R}^d} |h(x, 0)| < \infty$.

Remark 2.2. The first condition (14) is standard in the ergodic BSDE literature, and guarantees that the diffusion X is exponentially ergodic, in the sense that

$$|X_t^x - X_t^{\bar{x}}| \leq \exp(-2\delta t)|x - \bar{x}|^2$$

a.s. for each $t \geq 0$.

The growth and regularity conditions on g and h ensure that if π is an appropriate truncation function, then $\tilde{f}(x, z) = f(x, \pi(z))$ satisfies standard Lipschitz conditions for ergodic BSDEs, which will be important for the truncation technique we use to prove Theorem 2.3. Setting $g = 0$, we see that we easily cover the case that f is globally Lipschitz and $f(\cdot, 0)$ is bounded, which is the condition used in [11] and [7], albeit in a more general setting.

The third condition concerns the monotonicity of the map $z \mapsto \partial_x g(x, z)$, and is our main qualitative assumption. We note that when $d = 1$ and g is C^2 , the condition (15) is equivalent to the simpler condition $\partial_{zx} g \leq 0$.

We are now ready to state the main result. We define the quantity M by

$$M := \|\partial_x g\|_{L^\infty(\mathbb{R}^d \times B_1)} + \|\partial_x h\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \quad (16)$$

The main result is the following:

Theorem 2.3. *Suppose that Assumption 2.1 holds. Then, there exists a Markovian solution (u, v, λ) to (12). The function u has linear growth, and v satisfies*

$$\|v\|_{L^\infty} \leq \frac{M}{\delta}.$$

Moreover, the solution is unique in the class of Markovian solutions with v bounded and u of linear growth, in the sense that if $(\bar{u}, \bar{v}, \bar{\lambda})$ is another solution with the same properties then $v = \bar{v}$, $\lambda = \bar{\lambda}$, while u and \bar{u} differ by a constant. Finally, u is in fact a classical solution of the PDE (13), and $v = \nabla u$ a.e.

Remark 2.4. The uniqueness part of Theorem 2.3 follows easily from uniqueness for standard Lipschitz ergodic BSDEs, which can be obtained by adapting the proof of Theorem 3.11 in [7]. More precisely, if (u, v, λ) and $(\bar{u}, \bar{v}, \bar{\lambda})$ are Markovian solutions to (12), then they are also both solutions to the ergodic BSDE

$$dY_t = \lambda dt - \tilde{f}(X_t, Z_t)dt + Z_t \cdot dW_t, \quad (17)$$

where $\tilde{f}(X_t, Z_t) = \tilde{f}(X_t, \pi(Z_t))$ and $\pi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is the orthogonal projection onto the ball of radius $\max\{\|v\|_{L^\infty}, \|\bar{v}\|_{L^\infty}\}$. Furthermore \tilde{f} is globally Lipschitz thanks to Assumption 2.1. We can thus conclude from uniqueness for the equation (17) that $v = \bar{v}$, $\lambda = \bar{\lambda}$, and u and \bar{u} differ by a constant, as required. Thus, in what follows we focus on the question of existence.

3. EXISTENCE FOR EBSDES UNDER MONOTONICITY

The goal of this section is to present a proof of Theorem 2.3.

For later use, we recall the following elementary fact, which gives a differential characterization of the dissipativity condition (14).

Proposition 3.1. *Suppose that $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is C^1 . Then ϕ satisfies the dissipativity condition*

$$(\phi(x) - \phi(\bar{x})) \cdot (x - \bar{x}) \leq -\epsilon|x - \bar{x}|^2 \quad (18)$$

holds for $\epsilon \geq 0$ if and only if the condition

$$z^T D\phi(x)z \leq -\epsilon|z|^2$$

holds for all $x, z \in \mathbb{R}^d$.

Our approach to proving Theorem 2.3 will be by truncation. More precisely, first choose a smooth increasing function $\rho : [0, \infty) \rightarrow [0, \frac{M}{\delta} + 1]$ such that

- $\rho(x) = x$ for $x \leq \frac{M}{\delta}$,
- $\rho(x) = \frac{M}{\delta} + 1$ for all $x > \frac{M}{\delta} + 2$.

Then we set

$$\pi(z) = \frac{\rho(|z|)}{|z|}z, \quad (19)$$

and note that π is smooth, radial, bounded, and $\pi(z) = z$ when $|z| \leq \frac{M}{\delta}$. Next, we set

$$\tilde{f}(x, z) = f(x, \pi(z)), \quad (20)$$

and note that \tilde{f} is Lipschitz, so the ergodic BSDE

$$dY_t = \lambda dt - \tilde{f}(X_t, Z_t)dt + Z_t \cdot dW_t \quad (21)$$

has a Markovian solution (u, v, λ) with v bounded. If we can show that $\|v\|_{L^\infty} \leq \frac{M}{\delta}$, then (u, v, λ) is in fact a Markovian solution (12). Actually, the proof only works as outlined here when b and f are sufficiently smooth. The general case is more subtle, but the basic strategy is the same.

3.1. Formal derivation of the a-priori estimate. As explained above, proving Theorem 2.3 boils down to establishing the estimate

$$\|v\|_{L^\infty} \leq \frac{M}{\delta} \quad (22)$$

whenever (u, v, λ) is a Markovian solution to (12) with f globally Lipschitz and satisfying Assumption 2.1. For the reader's convenience, we give in this sub-section a formal derivation of (22) which provides intuition while avoiding the technical details.

We suppose in our formal argument that $d = 1$ (this simplifies a key step in the derivation), and we also suppose that u is in fact a smooth solution of the PDE

$$\frac{1}{2} \partial_x^2 u + b \partial_x u + f(x, \partial_x u) = \lambda, \quad (23)$$

and that $v = \partial_x u$. Then we differentiate (23) to find that

$$\frac{1}{2} \partial_x^2 v + b' v + \partial_x f(x, v) + (\partial_z f(x, v) + b) \partial_x v = 0.$$

We see that in fact $(\tilde{Y}, \tilde{Z}) = (v(X), \partial_x v(X))$ solves the infinite horizon BSDE

$$d\tilde{Y}_t = -\psi(X_t, \tilde{Y}_t, \tilde{Z}_t) dt + \tilde{Z}_t \cdot dW_t,$$

with

$$\psi(x, y, z) = b'(x)y + \partial_x g(x, y) + \partial_x h(x, v(x)) + \partial_z f(x, v(x))z$$

The dissipativity condition (14) and the monotonicity condition (15) combine to show that ψ is monotone in y , and $(\psi(x, y, z) - \psi(x, \bar{y}, z))(y - \bar{y}) \leq -\delta|y - \bar{y}|^2$, so that estimates for infinite horizon BSDEs (see the proof of Lemma 3.1 in [2]) show that

$$\|v\|_{L^\infty} = \|\tilde{Y}\|_{\mathcal{S}^\infty} \leq \frac{1}{\delta} \sup_{x \in \mathbb{R}^d} |\psi(x, 0, 0)| \leq \frac{M}{\delta},$$

as desired.

3.2. Proof of Theorem 2.3. This sub-section is devoted to a proof of Theorem 2.3. We use heavily the connection between the ergodic BSDE (12) and the PDE (13) established in the appendix.

Lemma 3.2 (a-priori estimate). *Suppose Assumption 2.1 holds, and in addition $f = g + h$ is globally Lipschitz and b, g, h are $C_{loc}^{1,\alpha}$. Let (u, v) be the unique Markovian solution to the ergodic BSDE (12) with u of linear growth and v bounded. Then we have*

$$\|v\|_{L^\infty} \leq \frac{M'}{\delta},$$

where $M' = \sup_{x \in \mathbb{R}^d} |\partial_x g(x, 0)| + \|\partial_x h\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)}$.

Proof. By Lemma A.4, we find that u is a C^3 solution to (13), and that we may assume $v = \nabla u$ is C^2 . Thus, we can differentiate (13) to find an equation for v . Namely, writing $v^i = \partial_i u$, we obtain the PDE system

$$\frac{1}{2} \Delta v^i + \partial_i b \cdot v + b \cdot \nabla v^i + \partial_{x_i} f(x, v) + \partial_z f(x, v) \nabla v^i = 0. \quad (24)$$

We notice that if we set X^x to be the unique solution of the SDE

$$X_t^x = x + \int_0^t b(X_s^x) ds + W_t,$$

then since v is C^2 , we may apply Itô's formula and (24) to verify that the $\mathbb{R}^d \times \mathbb{R}^{d \times d}$ -valued pair of processes (Y^x, Z^x) given by $(Y^x, Z^x) = (v(X^x), Dv(X^x))$ solves the discounted BSDE

$$\begin{aligned} Y_t^{x,i} &= Y_T^{x,i} + \int_t^T (\partial_i b(X_s^x) \cdot Y_s^x + \partial_{x_i} f(X_s^x, Y_s^x)) ds \\ &\quad + \int_t^T \partial_z f(X_s^x, Y_s^x) \cdot Z_s^{x,i} ds - \int_t^T Z_s^{x,i} \cdot dW_s. \end{aligned}$$

We rewrite the above as

$$\begin{aligned} Y_t^{x,i} &= Y_T^{x,i} + \int_t^T (\partial_i b(X_s^x) \cdot Y_s^x + \partial_{x_i} f(X_s^x, Y_s^x)) ds - \int_t^T Z_s^{x,i} \cdot d\tilde{W}_s \\ &= Y_T^{x,i} + \int_t^T (\partial_i b(X_s^x) \cdot Y_s^x + \partial_{x_i} g(X_s^x, Y_s^x) + \partial_{x_i} h(X_s^x, Y_s^x)) ds - \int_t^T Z_s^{x,i} \cdot d\tilde{W}_s \end{aligned} \quad (25)$$

where

$$\tilde{W} := W - \int \partial_z f(X^x, Y^x) dt$$

is by Girsanov's Theorem a Brownian motion on $[0, T]$ under the equivalent probability measure \mathbb{Q}_T given by

$$d\mathbb{Q}_T = \mathcal{E}\left(\int \partial_z f(X^x, Y^x) \cdot dW\right)_T d\mathbb{P}.$$

Now we apply Itô's formula together with (25) to compute

$$\begin{aligned}
d|Y_t^x|^2 &= \left(-2(Y_t^x)^T Db(X_t^x)Y_t - 2(Y_t^x)^T \partial_x g(X_t^x, Y_t^x) \right. \\
&\quad \left. - 2(Y_t^x)^T \partial_x h(X_t^x, Y_t^x) + |Z_t^x|^2 \right) dt + (Y_t^x)^T Z_t^x d\tilde{W}_t \\
&= \left(-2(Y_t^x)^T Db(X_t^x)Y_t - 2(Y_t^x)^T (\partial_x g(X_t^x, Y_t^x) - \partial_x g(X_t^x, 0)) \right. \\
&\quad \left. - 2(Y_t^x)^T \partial_x g(X_t^x, 0) - 2(Y_t^x)^T \partial_x h(X_t^x, Y_t^x) + |Z_t^x|^2 \right) dt + (Y_t^x)^T Z_t^x d\tilde{W}_t. \tag{26}
\end{aligned}$$

The monotonicity condition gives

$$\beta_t := -2(Y_t^x)^T Db(X_t^x)Y_t - 2(Y_t^x)^T (\partial_x g(X_t^x, Y_t^x) - \partial_x g(X_t^x, 0)) + |Z_t^x|^2 \geq 2\delta|Y_t^x|^2, \tag{27}$$

and Young's inequality together with the growth conditions on $\partial_x g$ and $\partial_x h$ gives

$$\begin{aligned}
|\gamma_t| &:= |2(Y_t^x)^T \partial_x g(X_t^x, 0) + 2(Y_t^x)^T \partial_x h(X_t^x, Y_t^x)| \\
&\leq 2|Y_t^x| (|\partial_x g(X_t^x, 0)| + |\partial_x h(X_t^x, Y_t^x)|) \\
&\leq \frac{1}{\delta} (|\partial_x g(X_t^x, 0)| + |\partial_x h(X_t^x, Y_t^x)|)^2 + \delta|Y_t^x|^2 \leq \frac{(M')^2}{\delta} + \delta|Y_t^x|^2. \tag{28}
\end{aligned}$$

Combining (26), (27), and (28), we see that we have

$$d|Y_t^x|^2 = (\delta|Y_t^x|^2 + \alpha_t)dt + (Y_t^x)^T Z_t^x d\tilde{W}_t, \tag{29}$$

where

$$\alpha_t = \beta_t + \gamma_t - \delta|Y_t^x|^2 \geq -\frac{(M')^2}{\delta}.$$

Now from (29) it is easy to check that

$$|Y_0^x|^2 = \exp(-\delta T)|Y_T^x|^2 - \int_0^T \exp(-\delta s)\alpha_s ds - \int_0^T \exp(-\delta s)(Y_s^x)^T Z_s^x d\tilde{W}_s. \tag{30}$$

We already know that Y^x is bounded, so from equation (25) we see that the \mathbb{Q}_T -local martingale $\int Z^x \tilde{W}$ is bounded. It follows that the stochastic integral in (30) is a true martingale on $[0, T]$ under \mathbb{Q}_T , and so

$$\begin{aligned}
|Y_0^x|^2 &= \exp(-\delta T)\mathbb{E}^{\mathbb{Q}_T}[|Y_T^x|^2] + \mathbb{E}^{\mathbb{Q}_T}\left[\int_0^T -\exp(-\delta s)\alpha_s ds\right] \\
&\leq \exp(-\delta T)\|Y\|_{\mathcal{S}^\infty}^2 + \int_0^T \exp(-\delta s)\frac{(M')^2}{\delta} dt \\
&= \exp(-\delta T)\|Y\|_{\mathcal{S}^\infty}^2 + \frac{(M')^2}{\delta^2}(1 - \exp(-\delta T)),
\end{aligned}$$

where $\mathbb{E}^{\mathbb{Q}_T}$ denotes expectation under \mathbb{Q}_T . Finally, sending $T \rightarrow \infty$ gives the estimate

$$|v(x)|^2 = |Y_0^x|^2 \leq \frac{(M')^2}{\delta^2},$$

i.e. $\|v\|_{L^\infty} \leq \frac{M'}{\delta}$. \square

Remark 3.3. A careful reading of the proof of Lemma 3.2 shows that we do not in fact need the full monotonicity condition (15) for the conclusion of the lemma to hold, but only the weaker condition

$$(\partial_x g(x, z) - \partial_x g(x, 0)) \cdot z \leq 0 \quad (31)$$

for all $x, z \in \mathbb{R}^d$.

Finally, we can complete the proof of Theorem 2.3.

Proof of Theorem 2.3. We split the argument into several steps.

Step 1 (checking that mollification preserves structural properties): Let $(\rho_\epsilon)_{0 < \epsilon < 1}$ be a standard approximation to the identity on \mathbb{R}^d (in particular satisfying $\text{supp}(\rho_\epsilon) \subset B_1$ for all ϵ), and define

- $b_\epsilon(x) = b * \rho_\epsilon(x) = \int_{\mathbb{R}^d} b(x-y)\rho_\epsilon(y)dy$
- $g_\epsilon(x, z) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(x-y_1, z-y_2)\rho_\epsilon(y_1)\rho_\epsilon(y_2)dy_1dy_2$,
- $h_\epsilon(x, z) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(x-y_1, z-y_2)\rho_\epsilon(y_1)\rho_\epsilon(y_2)dy_1dy_2$.
- $f_\epsilon = g_\epsilon + h_\epsilon$.

Then b_ϵ and f_ϵ are smooth and, moreover we have

$$\begin{aligned} (b_\epsilon(x) - b_\epsilon(\bar{x})) \cdot (x - \bar{x}) &= \left(\int_{\mathbb{R}^d} [b(x-y) - b(\bar{x}-y)]\rho_\epsilon(y)dy \right) \cdot (x - \bar{x}) \\ &= \int_{\mathbb{R}^d} ([b(x-y) - b(\bar{x}-y)]\rho_\epsilon(y)) \cdot (x - y - \bar{x} + y)dy \leq -\delta|x - \bar{x}|^2, \end{aligned} \quad (32)$$

i.e. b_ϵ satisfies the dissipativity condition (14) uniformly in ϵ . Similar computations show that

$$(\partial_x g_\epsilon(x, z) - \partial_x g_\epsilon(x, \bar{z})) \cdot (z - \bar{z}) \leq 0, \quad (33)$$

and also

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} |\partial_x g_\epsilon(x, 0)| &\leq \|\partial_x g\|_{L^\infty(\mathbb{R}^d \times B_1)}, \\ \|\partial_x h_\epsilon\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} &\leq \|\partial_x h\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)}, \end{aligned}$$

so that in particular

$$\sup_{x \in \mathbb{R}^d} |\partial_x g_\epsilon(x, 0)| + \|\partial_x h_\epsilon\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \leq M. \quad (34)$$

Step 2 (estimates for mollified and truncated equation): We set $\tilde{g}_\epsilon(x, z) = g_\epsilon(x, \pi(z))$, $\tilde{h}_\epsilon(x, z) = h_\epsilon(x, \pi(z))$, and $\tilde{f}_\epsilon = \tilde{g}_\epsilon + \tilde{h}_\epsilon$, where π is as introduced in (19) and notice that thanks to (34)

$$(\partial_x \tilde{g}_\epsilon(x, z) - \partial_x \tilde{g}_\epsilon(x, 0)) \cdot z = \rho(|z|)(\partial_x \rho(|z|)g_\epsilon(x, \pi(z)) - \partial_x \tilde{g}_\epsilon(x, 0)) \cdot \pi(z) \leq 0. \quad (35)$$

That is, \tilde{g}_ϵ satisfies the weak monotonicity condition (31). Since g_ϵ and h_ϵ are smooth, so are \tilde{g}_ϵ and \tilde{h}_ϵ , and from (34) we see that

$$\sup_{x \in \mathbb{R}^d} |\partial_x \tilde{g}_\epsilon(x, 0)| + \|\partial_x \tilde{h}_\epsilon\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \leq M. \quad (36)$$

Using (32), (35), and (36), we may apply Lemma 3.2 (actually a slight extension, see Remark 3.3) to conclude that

$$\|v^\epsilon\|_{L^\infty} \leq \frac{M}{\delta}, \quad (37)$$

where $(u^\epsilon, v^\epsilon, \lambda^\epsilon)$ is the unique Markovian solution to

$$dY_t^\epsilon = \lambda^\epsilon dt - \tilde{f}_\epsilon(X_t, Z_t^\epsilon) dt + Z_t^\epsilon \cdot dW_t. \quad (38)$$

Step 3 (Passing the a-priori estimate through the limit): Recalling that u^ϵ is unique only up to an additive constant, let us normalize u^ϵ so that $u^\epsilon(0) = 0$. By Proposition A.4, u^ϵ actually is a C^2 solution of the the corresponding PDE and $v^\epsilon = \nabla u^\epsilon$ a.e., thus elliptic regularity arguments (very similar to those presented in the proof of Proposition A.4 in the appendix) show that for some $0 < \gamma < 1$ and for each $R > 0$, u^ϵ is uniformly bounded in $C^{2,\gamma}(B_R)$, and that moreover for some $\epsilon_n \downarrow 0$, we have

$$\|u^{\epsilon_n} - u\|_{C^{2,\gamma}(B_R)} \rightarrow 0 \text{ as } n \rightarrow \infty,$$

for each $R > 0$, where (u, v, λ) is the unique Markovian solution of the truncated EBSDE (21) (also normalized so that $u(0) = 0$). In particular, $v^\epsilon = \nabla u^{\epsilon_n} \rightarrow \nabla u = v$ almost everywhere, and so we get $\|v\|_{L^\infty} \leq \frac{M}{\delta}$. Thus, the unique Markovian solution (u, v, λ) of (21) is also a Markovian solution of (12), as desired. As explained in Remark 2.4, uniqueness follows from uniqueness for Lipschitz drivers, and the fact that u is a regular solution of (13) is established in Proposition A.4. \square

4. APPLICATIONS TO CONTROL

In this section, we apply Theorem 2.3 to an ergodic control problem. We do not pursue the greatest possible generality, because we are primarily interested in answering the question - what does the monotonicity condition (15) mean in terms of ergodic control? In particular, we focus in both cases on the case of linear drift, convex costs, and sufficiently smooth data, since this is the setting where the monotonicity condition (15) can most easily be understood.

4.1. Ergodic control. In this sub-section, we recall the basic connection between ergodic control and the EBSDE (12), and then we check which conditions on the data of the ergodic control problem guarantee that the corresponding EBSDE satisfies Assumption 2.1.

4.1.1. *Set-up and connection to ergodic BSDE.* The data will be a cost function $r = r(x, a) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ and a vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$.

Assumption 4.1 (basic assumptions for control problem). The functions $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $r : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfy

- (1) (dissipativity and regularity for the drift) b satisfies the first condition in Assumption 2.1, i.e. b is Lipschitz and satisfies the dissipativity condition (14).
- (2) (local boundedness of the cost) $r : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfies the estimate

$$|r(x, a)| \leq \kappa(|a|) \quad (39)$$

for some increasing function $\kappa : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and all $x, a \in \mathbb{R}^d$.

- (3) (convexity of the cost) r is uniformly convex in a .

For each $x \in \mathbb{R}^d$, we denote by X^x the solution of the SDE

$$X_t^x = x + \int_0^t b(X_s^x) ds + W_t.$$

We define \mathcal{A} to be the set of bounded and progressively measurable processes taking values in \mathbb{R}^d . For any $\alpha \in \mathcal{A}$, we can write

$$X_t^x = (b(X_t^x) + \alpha_t) dt + dW_t^\alpha,$$

where $W^\alpha = W - \int \alpha dt$ is a Brownian motion on the interval $[0, T]$ under the probability measure $\mathbb{P}^{\alpha, T}$, given by

$$d\mathbb{P}^{\alpha, T} = \mathcal{E}\left(\int \alpha \cdot dW\right)_T d\mathbb{P}.$$

Now we define for each $x \in \mathbb{R}^d$ the cost functional $J_x : \mathcal{A} \rightarrow \mathbb{R}$, given by

$$J_x(\alpha) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}^{\mathbb{P}^{\alpha, T}} \left[\int_0^T r(X_t^x, \alpha_t) dt \right], \quad (40)$$

which is well-defined under Assumption 4.1. We recall the relevant Hamiltonian

$$\begin{aligned} H(x, z) &= \inf_{a \in \mathbb{R}^d} h(x, z, a), \\ h(x, z, a) &:= a \cdot z + r(x, a). \end{aligned} \quad (41)$$

Since r is uniformly convex, there is an optimizer \hat{a} of h , i.e. a measurable map $\hat{a}(x, z) : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$h(x, z, \hat{a}(x, z)) = H(x, z).$$

We pose the ergodic BSDE

$$dY_t = \lambda dt - H(X_t, Z_t) dt + Z_t \cdot dW_t. \quad (42)$$

We now state without proof the connection between the control problem and an ergodic BSDE, which is just a slight adaptation of Theorem 7.1 in [11] to our setting.

Proposition 4.2. *Suppose that the ergodic BSDE (42) has a Markovian solution (u, v, λ) such that u is of linear growth and v is bounded, and that $x \mapsto \hat{a}(x, v(x))$ is bounded. Then, for each $x \in \mathbb{R}^d$, we have*

$$\inf_{\alpha \in \mathcal{A}} J_x(\alpha) = \lambda = J_x(\alpha^*),$$

where $\alpha^* = \hat{a}(X^x, v(X^x)) \in \mathcal{A}$.

4.1.2. *Conditions for existence of an optimal control.* Now we establish conditions on the running cost r under which the ergodic BSDE (42) is covered by Theorem 2.3. The main difficulty is to check what conditions on r will force the driver $H(x, z) = h(x, z, \hat{a}(x, z))$ to satisfy the monotonicity condition

$$(\partial_x H(x, z) - \partial_x H(x, \bar{z})) \cdot (z - \bar{z}) \leq 0. \quad (43)$$

We start with a Lemma, which is proved by differentiating the optimality condition

$$z^i + \partial_a r^i(x, \hat{a}(x, z)) = 0. \quad (44)$$

implicitly in z .

Lemma 4.3. *If r is C^2 , then*

$$\partial_z \partial_x H(x, z) = -\partial_a \partial_x r(x, \hat{a}(x, z)) (\partial_a^2 r(x, \hat{a}(x, z)))^{-1}.$$

In particular, by Proposition 3.1, H satisfies the monotonicity condition (15) whenever $\partial_a \partial_x r(x, a) (\partial_a^2 r(x, a))^{-1}$ is non-negative, i.e.

$$z^T \partial_a \partial_x r(x, a) (\partial_a^2 r(x, z))^{-1} z \geq 0 \quad (45)$$

holds for all $x, z, a \in \mathbb{R}^d$.

Remark 4.4. Notice that in the case $d = 1$, the condition that $\partial_a \partial_x r (\partial_a^2 r)^{-1}$ is non-negative is equivalent to the condition that $\partial_a \partial_x r \geq 0$, since we have already assumed that $\partial_x^2 r > 0$. Notice also that in any dimension, the condition is satisfied as soon as $\partial_a \partial_x r = 0$, i.e. if r has a separated structure

$$r(x, a) = r_1(x) + r_2(a).$$

Proposition 4.5. *Suppose that in addition to Assumption 4.1, r is C^2 . Suppose further $\partial_a \partial_x r (\partial_a^2 r)^{-1}$ is non-negative. Finally, assume that H is bounded and Lipschitz on $\mathbb{R}^d \times B_R$ for each $R > 0$. Then the drift b and the driver H satisfy Assumption 2.1, and so by Theorem 2.3 the ergodic BSDE (42) has a unique Markovian solution (u, v, λ) with u of linear growth and v bounded.*

4.2. Risk-sensitive ergodic control. Now we apply the reasoning from the previous sub-section to a risk-sensitive ergodic control problem, as introduced in [10]. As in the previous sub-section, the data of the control problem consists of a vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and a cost function $r : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$. We will again enforce Assumption 4.1, and define \mathcal{A} , $\mathbb{P}^{\alpha, T}$, H , \hat{a} , etc. in the same way. But this time we fix $\delta > 0$, and define the cost functionals $J_x^\delta : \mathcal{A} \rightarrow \mathbb{R}$ by

$$J_x^\delta(\alpha) = \limsup_{T \rightarrow \infty} \frac{1}{T} \left(\frac{1}{\delta} \ln \mathbb{E}^{\mathbb{P}^{\alpha, T}} \left[\exp \left(\delta \int_0^T r(X_t^x, \alpha_t) dt \right) \right] \right). \quad (46)$$

It turns out that this control problem is related to the ergodic BSDE

$$\begin{aligned} dY_t &= \lambda dt - f(X_t, Z_t) dt + Z_t \cdot dW_t, \\ \text{with } f(x, z) &= H(x, z) + \frac{\delta}{2} |z|^2, \end{aligned} \quad (47)$$

where the Hamiltonian H is as defined in the previous sub-section. We now state the connection between the risk-sensitive control problem and the ergodic BSDE (47).

Proposition 4.6. *Suppose that the ergodic BSDE (47) has a Markovian solution (u, v, λ) such that u is of linear growth and v is bounded, and that $x \mapsto \hat{a}(x, v(x))$ is bounded. Then, for each $x \in \mathbb{R}^d$, we have*

$$\inf_{\alpha \in \mathcal{A}} J_x^\delta(\alpha) = \lambda = J_x^\delta(\alpha^*),$$

where $\alpha^* = \hat{a}(X^x, v(X^x)) \in \mathcal{A}$.

This result is a fairly standard generalization of Proposition 4.2, and the proof of Proposition 4 in [17] applies directly to this slightly more general setting. Thus we omit the proof. Because the driver f in (47) differs from the driver H appearing in the previous sub-section only by the term $\frac{\delta}{2}|z|^2$, it is clear that f satisfies Assumption 2.1 if and only if H does. Thus we get the following analogue of Proposition 4.5.

Proposition 4.7. *Suppose that in addition to Assumption 4.1, r is C^2 . Suppose further that $\partial_a \partial_x r (\partial_a^2 r)^{-1}$ is non-negative. Finally, assume that H is bounded and Lipschitz on $\mathbb{R}^d \times B_R$ for each $R > 0$. Then the drift b and the driver f satisfy Assumption 2.1, and so by Theorem 2.3 the EBSDE (47) has a unique Markovian solution (u, v, λ) with u of linear growth and v bounded.*

4.3. Weak vs. strong formulation. The control problems discussed in the previous two subsections are in weak formulation, in the sense that the underlying diffusion X^x is fixed and the controller affects the dynamics of X^x by changing the probability measure. This formulation is often easier to work with than the strong formulation, but it is interesting to note that because we are in the Markovian setting and the controls produced are bounded, we can treat the strong formulation by the same techniques.

As in the weak formulation, we fix a vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and a cost function $r : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfying Assumption 4.1. Then we define \mathcal{A}^S to be the set of all bounded and measurable maps $\alpha = \alpha(t, x) : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, and note that by a classical result which can be traced to Veretennikov (see the introduction of [21] for a modern account of existence of strong solutions for SDEs with irregular drift), for each $x \in \mathbb{R}^d$ and $\alpha \in \mathcal{A}^S$, there is a unique solution $X^{x,\alpha}$ to the SDE

$$\begin{cases} dX_t^{x,\alpha} = (b(X_t^{x,\alpha}) + \alpha(t, X_t^{x,\alpha}))dt + dW_t, \\ X^{x,\alpha} = x. \end{cases}$$

For each $x \in \mathbb{R}^d$, we define $J_x^S : \mathcal{A}^S \rightarrow \mathbb{R}$ by

$$J_x^S(\alpha) = \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T r(X_t^{x,\alpha}, \alpha(t, X_t^{x,\alpha})) dt \right].$$

This formulation will be equivalent to the weak formulation in the following sense.

Proposition 4.8. *Suppose that the ergodic BSDE (42) has a Markovian solution (u, v, λ) such that u has linear growth, v is bounded and $\alpha^*(t, x) = \alpha^*(x) := \hat{a}(x, v(x))$ is bounded. Then*

$$\inf_{\alpha \in \mathcal{A}^S} J_x^S(\alpha) = J_x^S(\alpha^*) = \lambda.$$

In particular, $\inf_{\alpha \in \mathcal{A}^S} J_x^S(\alpha) = \inf_{\alpha \in \mathcal{A}} J_x(\alpha)$.

Proof. Let (u, v) be the Markovian solution to (42) which is supposed to exist, and for fixed $x \in \mathbb{R}^d$ set $(Y^x, Z^x) = (u(X^x), v(X^x))$, where $X^x = X^{x,\alpha^*}$. We claim that

$$Y_t^x = Y_T^x - \lambda(T-t) + \int_t^T r(X_s^x, \hat{a}(X_s^x, Z_s^x)) ds - \int_t^T Z_s^x \cdot dW_s. \quad (48)$$

Indeed, the fact that (u, v) is a Markovian solution for (42) implies that if $X = X^0$, and $(Y, Z) = (u(X), v(X))$ then we have

$$\begin{aligned} Y_t &= Y_T - \lambda(T-t) + \int_t^T h(X_s, Z_s, \hat{a}(X_s, Z_s)) ds - \int_t^T Z_s \cdot dW_s \\ &= Y_T - \lambda(T-t) + \int_t^T \left(h(X_s, Z_s, \hat{a}(X_s, Z_s)) + b(X_s) \cdot v(X_s) \right) ds - \int_t^T Z_s \cdot dX_s, \end{aligned}$$

i.e.

$$\begin{aligned} u(X_t) &= u(X_T) - \lambda(T-t) + \int_t^T \left(h(X_s, v(X_s), \hat{a}(X_s, v(X_s))) + b(X_s) \cdot v(X_s) \right) ds \\ &\quad - \int_t^T v(X_s) \cdot dX_s. \end{aligned}$$

But for each $T > 0$ there is an equivalent probability measure \mathbb{Q}^T under which the law of X^x on $[0, T]$ is equal to the law of X under \mathbb{P} , and so the equation

$$\begin{aligned} u(X_t^x) = u(X_T^x) - \lambda(T-t) + \int_t^T \left(h(X_s^x, v(X_s^x), \hat{a}(X_s^x, v(X_s^x))) + b(X_s^x) \cdot v(X_s^x) \right) ds \\ - \int_t^T v(X_s^x) \cdot dX_s^x \end{aligned} \quad (49)$$

holds under \mathbb{Q}^T , hence also under \mathbb{P} . Combining (49) with the identities

$$\begin{aligned} dX_s^x &= (b(X_s^x) + \hat{a}(X_s^x, v(X_s^x))) + dW_s, \\ h(X_s^x, v(X_s^x), \hat{a}(X_s^x, v(X_s^x))) &= \hat{a}(X_s^x, v(X_s^x)) \cdot v(X_s^x) + r(X_s^x, \hat{a}(X_s^x, v(X_s^x))), \end{aligned}$$

we arrive at

$$u(X_t^x) = u(X_T^x) - \lambda(T-t) + \int_t^T r(X_s^x, \hat{a}(X_s^x, v(X_s^x))) ds - \int_t^T v(X_s^x) \cdot dW_s.$$

Recalling that $(Y^x, Z^x) = (u(X^x), v(X^x))$, we get (48). From (48), we in turn get

$$\lambda = \frac{1}{T} \left(Y_T^x - Y_0^x + \int_0^T r(X_s^x, \hat{a}(X_s^x, Z_s^x)) dt - \int_0^T Z_s^x \cdot dW_s \right)$$

For any fixed $\alpha \in \mathcal{A}^S$, we can further manipulate the expression for λ to read

$$\begin{aligned} \lambda = \frac{1}{T} \left(Y_T^x - Y_0^x + \int_0^T (h(X_s^x, Z_s^x, \hat{a}(X_s^x, Z_s^x)) - h(X_s^x, Z_s^x, \alpha(s, X_s^x))) ds \right. \\ \left. + \int_0^T r(X_s^x, \alpha(s, X_s^x)) ds - \int_0^T Z_s^x \cdot dW_s^\alpha \right), \end{aligned} \quad (50)$$

where $W^\alpha = W - \int \alpha(\cdot, X^x) dt + \int \alpha^*(X^x) dt$ is a Brownian motion on $[0, T]$ under $\mathbb{P}^{\alpha, T}$, where $d\mathbb{P}^{\alpha, T} = \mathcal{E}(\int [\alpha(\cdot, X^x) - \alpha^*(X^x)] \cdot dW)_T$. We note that we have

$$dX_t^x = (b(X_t) + \alpha(t, X_t)) dt + dW_t^\alpha$$

under $\mathbb{P}^{\alpha, T}$, i.e. X^x is a weak solution of the same SDE which defines $X^{x, \alpha}$. By uniqueness in law for this equation, we deduce that the law of X^x under $\mathbb{P}^{\alpha, T}$ is the same as that of $X^{x, \alpha}$ under \mathbb{P} , and in particular from (50) we get

$$\begin{aligned} \lambda = \frac{1}{T} \mathbb{E}^{\mathbb{P}^{\alpha, T}} [Y_T^x - Y_0^x + \int_0^T (h(X_s^x, Z_s^x, \hat{a}(X_s^x, Z_s^x)) - h(X_s^x, Z_s^x, \alpha(s, X_s^x))) ds] \\ + \frac{1}{T} \mathbb{E} \left[\int_0^T r(X_s^{x, \alpha}, \alpha(s, X_s^{x, \alpha})) ds \right]. \end{aligned}$$

Since $(h(X_s^x, Z_s^x, \hat{a}(X_s^x, Z_s^x)) - h(X_s^x, Z_s^x, \alpha(s, X_s^x))) \leq 0$, we can take $T \rightarrow \infty$ to get $\lambda \leq J_x^S(\alpha)$. Similar arguments show that $\lambda = J_x^S(\alpha^*)$, and this completes the proof. \square

5. EQUATION FROM FORWARD PERFORMANCE PROCESSES

In this section, we consider the driver

$$f(x, z) = \frac{1}{2}\delta^2 \text{dist}^2(\Pi, \frac{\theta(x) + z}{\delta}) - z^T \theta(x) + \frac{1}{2}|\theta(x)|^2, \quad (51)$$

where Π is a convex and closed subset of \mathbb{R}^d , and $\theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$. This driver is used in Section 4 of [17] to construct forward performance processes in factor form. In their context, the diffusion X is a stochastic factor process, and $\theta = \theta(x)$ is the market price of risk corresponding to the vector of stock prices S , which solves

$$dS_t^i = S_t^i(\eta^i(X_t)dt + \sum_{j=1}^d \sigma^{ij}(X_t)dW_t^j)$$

for η and σ satisfying appropriate conditions. That is, $\theta = \sigma^T[\sigma\sigma^T]^{-1}\eta$.

Because we are primarily interested in interpreting our monotonicity condition (15) in this setting, we refer to [17] for more details on forward performance processes and their connection to the driver (51). Here, we simply take the market price of risk $\theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and the convex closed set $\Pi \subset \mathbb{R}^d$ as given, and we focus on checking what conditions on θ and Π will guarantee that f satisfies the hypotheses of Theorem 2.3. We will see that in a few simple cases the monotonicity condition (15) boils down to a monotonicity condition on θ . As usual, we fix a vector field $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ which satisfies the first condition in Assumption 2.1.

Let us suppose that $h = 0$ in the decomposition $f = g + h$. Then we need to check under what conditions on Π and θ we have the monotonicity condition

$$(\partial_x f(x, z) - \partial_x f(x, \bar{z})) \cdot (z - \bar{z}) \leq 0. \quad (52)$$

First, we recall that since Π is closed and convex, there is a well-defined projection $P_\Pi : \mathbb{R}^d \rightarrow \Pi$, which maps $x \in \Pi^c$ to the unique element of $\text{argmin}_{p \in \Pi} |x - p|$. Moreover, $x \mapsto \text{dist}(\Pi, x)$ is Lipschitz, and its weak gradient $\nabla \text{dist}(\Pi, \cdot)$ is given by

$$\nabla \text{dist}(\Pi, \cdot)(x) = \begin{cases} \frac{x - P_\Pi(x)}{|x - P_\Pi(x)|} & x \in \Pi^c \\ 0 & x \in \Pi. \end{cases}$$

Using this, we find that

$$\nabla \text{dist}^2(\Pi, \cdot)(x) = \begin{cases} 2(x - P_\Pi(x)) & x \in \Pi^c, \\ 0 & x \in \Pi. \end{cases}$$

Now we can compute $\partial_x f$ where f is the driver (51). We find that

$$\begin{aligned} [\partial_x f(x, z)]^T &= \delta \left(\frac{\theta(x) + z}{\delta} - P_\Pi \left(\frac{\theta(x) + z}{\delta} \right) \right) D\theta(x) - z^T D\theta(x) + \theta(x)^T D\theta(x) \\ &= -\delta P_\Pi \left(\frac{\theta(x) + z}{\delta} \right) D\theta(x) + 2\theta(x)^T D\theta(x). \end{aligned}$$

Using this we check that the condition (52) is equivalent to the condition

$$\left(P_\Pi \left(\frac{\theta(x) + z}{\delta} \right) - P_\Pi \left(\frac{\theta(x) + \bar{z}}{\delta} \right) \right) D\theta(x)(z - \bar{z}) \geq 0. \quad (53)$$

From this, we easily deduce the following Proposition:

Proposition 5.1. *Suppose that θ is Lipschitz and bounded, and that the condition (53) holds. Then b and f satisfy Assumption 2.1, and so the ergodic BSDE (12) admits a unique Markovian solution by Theorem 2.3.*

Remark 5.2. Some computations reveal that the condition (53) holds (and hence Proposition 5.1 applies) if

- $d = 1$ and $\theta' \geq 0$, or
- $\Pi = \mathbb{R}^d$ and $D\theta \geq 0$, in the sense that $z^T D\theta(x)z \geq 0$ for all $x, z \in \mathbb{R}^d$.

Thus in these two simple cases, the monotonicity condition (15) on the driver f reduces to a certain monotonicity condition on the market price of risk vector θ . Unfortunately, this does not hold in general, i.e. it is possible to choose Π and θ such that $D\theta$ is non-negative, but (53) is not satisfied.

APPENDIX A. SOME RESULTS FROM ELLIPTIC REGULARITY THEORY

In this appendix, we establish rigorously the connection between the discounted BSDE

$$\begin{cases} dX_t = b(X_t)dt + dW_t, \\ dY_t = \rho Y_t dt - f(X_t, Z_t)dt + Z_t \cdot dW_t \end{cases} \quad (54)$$

and the PDE (55)

$$\frac{1}{2} \Delta u^\rho + b \cdot \nabla u^\rho + f(x, \nabla u^\rho) = \rho u^\rho, \quad (55)$$

and also the connection between the ergodic BSDE (12) and the PDE (13). We start with a preliminary Lemma which is standard in the ergodic BSDE literature.

Lemma A.1. *Suppose that in addition to Assumption 2.1, f is Lipschitz, and let (u, v, λ) be the unique solution to the ergodic BSDE (12). Then we have*

- $v = \nabla u$ a.s.
- there exists $x_0 \in \mathbb{R}^d$ and a sequence $\rho_n \downarrow 0$ such that u is locally uniform limit of the sequence $u^{\rho_n} - u^{\rho_n}(x_0)$, and λ is limit of $\rho_n u^{\rho_n}(x_0)$, where (u^{ρ_k}, v^{ρ_k}) is the Markovian solution of the discounted BSDE (54) with $\rho = \rho_k$.

The first assertion of Lemma A.1 is a special case of Theorem 5.3 in [11], while the second assertion is clear from e.g. the appendix of [17].

The next lemma establishes existence and regularity for a certain semi-linear elliptic equation.

Lemma A.2 (Semi-linear Dirichlet problem). *Suppose that in addition to Assumption 2.1, f is Lipschitz and bounded. For each $R > 0$ and $\phi : \partial B_R \rightarrow \mathbb{R}$ Lipschitz, there exists a unique Lipschitz classical solution u to the Dirichlet problem*

$$\begin{cases} \frac{1}{2}\Delta u + b \cdot \nabla u + f(x, \nabla u) = \rho u \text{ in } B_R \\ u = \phi \text{ on } \partial B_R. \end{cases} \quad (56)$$

Moreover, if b and f are $C_{loc}^{k,\alpha}$, then u is in C^{k+2} in B_R .

Proof. We proceed in three steps.

Step 1 (reduction to zero boundary condition): We begin by posing the linear problem

$$\begin{cases} \frac{1}{2}\Delta \tilde{u} + b \cdot \nabla \tilde{u} = \rho \tilde{u} \text{ in } B_R \\ \tilde{u} = \phi \text{ on } \partial B_R. \end{cases}$$

Since ϕ and b are Lipschitz, existence and regularity results for linear elliptic equations give a unique classical solution \tilde{u} which is Lipschitz on B_R . Moreover, if $b \in C_{loc}^{k,\alpha}$, then \tilde{u} lies in $C_{loc}^{k,\alpha}(B_R)$, thanks to the interior Schauder estimates. Subtracting the equations for u and \tilde{u} gives an equation for $w := u - \tilde{u}$, namely

$$\begin{cases} \frac{1}{2}\Delta w + b \cdot \nabla w + g(x, \nabla w) = \rho w \text{ in } B_R \\ w = 0 \text{ on } \partial B_R, \end{cases} \quad (57)$$

where $g(x, z) = f(x, z + \nabla \tilde{u}(x)) : B_R \times \mathbb{R}^d \rightarrow \mathbb{R}$. We note that g is uniformly Lipschitz in z , and locally Lipschitz in x .

Step 2 (Existence of a weak solution): Now we focus on producing a solution in H_0^1 to the equation (57). We define a map $\Phi : H_0^1(B_R) \rightarrow H_0^1(B_R)$ by $\Phi(\tilde{w}) = w$, where w is the unique solution (which exists by Theorem 8.3 of [12]) to the linear problem

$$\begin{cases} \frac{1}{2}\Delta w + b \cdot \nabla w + g(x, \nabla \tilde{w}) = \rho w \text{ in } B_R \\ w = \phi \text{ on } \partial B_R. \end{cases} \quad (58)$$

Notice g is uniformly Lipschitz in z , so that for $\tilde{w}, \tilde{w}' \in H_0^1(B_R)$, we have

$$\|g(\cdot, \nabla \tilde{w}) - g(\cdot, \nabla \tilde{w}')\|_{L^2(B_R)} \leq C \|\nabla \tilde{w} - \nabla \tilde{w}'\|_{L^2(B_R)} \leq C \|\tilde{w} - \tilde{w}'\|_{H_0^1(B_R)}$$

for some constant C , and thus by Corollary 8.7 of [12]

$$\|\Phi(\tilde{w}) - \Phi(\tilde{w}')\|_{H_0^1(B_R)} \leq C \|\tilde{w} - \tilde{w}'\|_{H_0^1(B_R)}.$$

That is, $\Phi : H_0^1(B_R) \rightarrow H_0^1(B_R)$ is continuous. Furthermore, by Calderon-Zygmund estimates (see e.g. Theorem 8.10 of [12]), the image of Φ is contained in a bounded subset of $H^2(\Omega)$, and in particular is compact in H_0^1 . Thus by the Schauder fixed point theorem, there exists a fixed point of Φ , i.e. a solution of (57).

Step 3 (bootstrapping to higher regularity): Now that we have a solution w to (57), we note that (again by Calderon-Zygmund estimates) we have that $w \in W^{2,p}(B_1)$ for all $p > 1$, hence in $C^{1,\beta}$ for some β by Sobolev embedding. Since g is locally Lipschitz, it follows that $x \mapsto g(x, \nabla w(x))$ is actually C_{loc}^β , so in fact by local Schauder estimates for linear elliptic equations (see Theorem 6.6 and Corollary 6.9 of [12]) we have $w \in C^{2,\beta}(B_R)$. If $b, f \in C_{loc}^{1,\alpha}$, then we see that $g \in C_{loc}^{1,\alpha}(B_R)$. Since $w \in C^{2,\beta}(B_R)$, we deduce that $x \mapsto g(x, \nabla w(x))$ is in $C_{loc}^{1,\gamma}(B_R)$ for some γ . Thus applying again the interior Schauder estimates (see Theorem 6.17 of [12]) we have $w \in C_{loc}^{3,\gamma}(B_R)$. If b, f are $C_{loc}^{k,\alpha}$, we just continue this bootstrapping argument to find that in fact $u \in C_{loc}^{k+2,\gamma}(B_R)$ for some γ , which completes the proof of existence.

Step 4 (uniqueness): For uniqueness, one checks that if u is a Lipschitz classical solution to (56), then for each $x \in \mathbb{R}^d$, $(Y^x, Z^x) = (u(X^x), \nabla u(X^x))$ solves the BSDE

$$\begin{cases} dY_t = \rho Y_t dt - f(X_t, Z_t) dt + Z_t \cdot dW_t, \\ Y_\tau = g(X_\tau) \end{cases}$$

on $[0, \tau]$, where τ is the stopping time $\tau = \inf\{t \geq 0 : |X_t^x| = R\}$. Thus uniqueness for the PDE (56) follows from uniqueness for such BSDEs, see e.g. Theorem 3.3 of [2]. \square

Lemma A.3 (Regularity of Markovian solutions to discounted BSDE). *Suppose that in addition to Assumption 2.1, f is globally Lipschitz, and let (u^ρ, v^ρ) be the unique Markovian solution to the discounted BSDE (54). Then in fact u^ρ is a classical solution to the PDE (55), and $v^\rho = \nabla u^\rho$ almost surely. Moreover, if b and f are $C_{loc}^{k,\alpha}$, then u^ρ is C^{k+2} .*

Proof. For $R > 0$, consider the Dirichlet problem

$$\begin{cases} \frac{1}{2} \Delta \tilde{u} + b \cdot \nabla \tilde{u}^\delta + f(x, \nabla \tilde{u}^\delta) = \rho \tilde{u}^\delta \text{ in } B_R, \\ \tilde{u}^\rho = u^\rho \text{ on } \partial B_R. \end{cases} \quad (59)$$

First, assume also that f is bounded. By Lemma A.2, we have a unique Lipschitz classical solution \tilde{u} . Now we fix $x \in B_R$, and we define

$$\begin{aligned} (Y^x, Z^x) &= (u^\delta(X^{x,\tau}), v^\delta(X^{x,\tau})), \\ (\tilde{Y}^x, \tilde{Z}^x) &= (\tilde{u}^\delta(X^{x,\tau}), \nabla \tilde{u}^\delta(X^{x,\tau})) \end{aligned}$$

where $X^{x,\tau}$ denotes the process X^x stopped at $\tau := \inf\{t \geq 0 : |X_t^x| = R\}$. One can check that $(\tilde{Y}^x, \tilde{Z}^x)$ solves the BSDE

$$\begin{cases} dY_t = [\rho Y_t - f(X_t, Z_t)dt + Z_t \cdot dW_t] 1_{t \leq \tau}, \\ Y_\tau = u^\delta(X_\tau^x) \end{cases} \quad (60)$$

on the stochastic interval $[0, \tau]$. Of course, (Y, Z) also satisfies (60), so by uniqueness for (see [2] Theorem 3.3) we must have

$$u^\delta(x) = Y_0^x = \tilde{Y}_0^x = \tilde{u}^\delta(x),$$

and $v^\delta = \nabla \tilde{u}$ almost surely. Lemma A.2 also shows that u is C^{k+1} as soon as b and f are $C_{loc}^{k,\alpha}$. Finally, to remove the hypothesis that f is bounded, we consider the drivers $f^j(x, z) = f(x, \pi^j(z))$, where π^j is a sequence of maps $\mathbb{R}^d \rightarrow \mathbb{R}^d$ which each smooth, have Lipschitz constant 1, are bounded, and satisfy $\pi^j(z) = z$ for $|z| \leq j$. Then let (u^j, v^j) be the Markovian solution to the discounted BSDE

$$dY_t = \rho Y_t - f^j(X_t, Z_t)dt + Z_t \cdot dW_t,$$

and note that we have already shown that u^j is a C^{k+2} solution of

$$\frac{1}{2}\Delta u^j + b \cdot \nabla u^j + f^j(x, \nabla u^j) = \rho u^j$$

and that $v^j = \nabla u^j$. Finally, we apply a-priori estimates for Lipschitz discounted BSDEs to get that $\sup_j \|v^j\|_{L^\infty} < \infty$, and so in fact for large enough j $(u^j, v^j) = (u, v)$, where (u, v) is the unique Markovian solution to (54). This completes the proof. \square

Proposition A.4 (Regularity for Markovian solutions of ergodic BSDE). *Suppose that in addition to Assumption 2.1, f is globally Lipschitz, and let (u, v) be the unique Markovian solution to the ergodic BSDE (12). Then u is in fact a classical solution to the PDE (13), and $v = \nabla u$ a.e. Moreover, if b and f are $C_{loc}^{k,\alpha}$, then u is C^{k+2} .*

Proof. We know that u is the locally uniform limit of $\tilde{u}^{\rho_n} := u^{\rho_n} - u^{\rho_n}(0)$, where for $\rho > 0$, (u^ρ, v^ρ) is the unique Markovian solution of the discounted BSDE (54). By Lemma A.3, \tilde{u}^{ρ_n} is in fact a smooth solution of the PDE

$$\frac{1}{2}\Delta \tilde{u}^{\rho_n} + b \cdot \nabla \tilde{u}^{\rho_n} + f(x, \nabla \tilde{u}^{\rho_n}) = \rho_n \tilde{u}^{\rho_n} + \rho_n u^{\rho_n}(0) \text{ in } \mathbb{R}^d.$$

Let us rewrite this equation as

$$\mathcal{L}\tilde{u}^{\rho_n} = F^n(x) \text{ in } \mathbb{R}^d, \quad (61)$$

where

$$F^n(x) = \rho_n \tilde{u}^{\rho_n} + \rho_n u^{\rho_n}(0) - f(x, \nabla \tilde{u}^{\rho_n}), \quad \mathcal{L}\tilde{u}^{\rho_n} = \frac{1}{2}\Delta \tilde{u}^{\rho_n} + b \cdot \nabla \tilde{u}^{\rho_n}.$$

By Lemma 3.1 of [2] (see also the Appendix of [17]), we have

$$\|u^{\rho_n}\|_{L^\infty} \leq \frac{\sup_{x \in \mathbb{R}^d} f(x, 0)}{\rho_n}, \quad (62)$$

and in particular the sequence of real number $\rho_n u^{\rho_n}(0)$ is bounded. Moreover, since we know that u^{ρ_n} is Lipschitz uniformly in n (again see the Appendix of [17]), we observe that for each $R > 0$,

- $\|f(\cdot, \nabla \tilde{u}^{\rho_n})\|_{L^\infty(\mathbb{R}^d)}$ is bounded uniformly in n ,
- $\|\rho_n \tilde{u}^{\rho_n}\|_{L^\infty(B_{2R})}$ converges to zero,
- $\rho_n u^{\rho_n}(0)$ is bounded uniformly in n .

Putting these observations together, we see that for each $R > 0$, $\|F^n\|_{L^\infty(B_{2R})}$ is bounded uniformly in n . By the interior Calderon-Zygmund estimates, we have for each $R > 0$ and any $1 < p < \infty$ the estimate

$$\|\tilde{u}^{\rho_n}\|_{W^{2,p}(B_R)} \leq C(\|\tilde{u}^{\rho_n}\|_{L^p(B_{2R})} + \|F^n\|_{L^p(B_{2R})}),$$

for some C independent of n , and in particular we see that for each $1 < p < \infty$ and each $R > 0$, $\|\tilde{u}^{\rho_n}\|_{W^{2,p}(B_R)}$ is bounded uniformly in n . By choosing p large enough and applying Sobolev embedding, we find that in fact $\|\tilde{u}^{\rho_n}\|_{C^{1,\alpha}(B_R)}$ is bounded uniformly in n for some $0 < \beta < 1$. But now similar reasoning as above shows that $\|F^n\|_{C^{1,\beta}(B_{2R})}$ is bounded uniformly in n . Now from the interior Schauder estimates we get

$$\|\tilde{u}^{\rho_n}\|_{C^{2,\beta}(B_R)} \leq C(\|\tilde{u}^{\rho_n}\|_{C^\beta(B_{2R})} + \|F^n\|_{C^\beta(B_{2R})})$$

for some C independent of n , and in particular we find that $\|\tilde{u}^{\rho_n}\|_{C^{2,\beta}(B_R)}$ is bounded uniformly in n . Since $C^{2,\beta}(B_R)$ compactly embeds in $C^{2,\gamma}$ when $0 < \gamma < \beta < 1$, we find that (passing to a further subsequence if necessary), we in fact have

$$\|u - \tilde{u}^{\rho_n}\|_{C^{2,\gamma}(B_R)} \rightarrow 0 \text{ as } n \rightarrow \infty$$

for some $0 < \gamma < \beta$ and all $R > 0$. From Lemma A.1, it is now easy to see that u is a classical solution of (13). To show that u is C^{k+2} when b and f are $C_{loc}^{k,\alpha}$, we continue the above bootstrapping argument to show that if b and f are $C_{loc}^{k,\alpha}$, then in fact $\|\tilde{u}^{\rho_n}\|_{C^{(k+2),\gamma}}$ is bounded uniformly in n for some γ , and then by the same reasoning as above

$$\|u - \tilde{u}^{\rho_n}\|_{C^{k+2,\gamma}(B_R)} \rightarrow 0 \text{ as } n \rightarrow \infty,$$

for each $R > 0$, and so in particular $u \in C_{loc}^{k+2,\gamma}(\mathbb{R}^d)$. Finally, by Lemma A.1, we have $v = \nabla u$ a.e. \square

REFERENCES

1. A. Bensoussan and J. Frehse, *On bellman equations of ergodic control in $Rn.$* , J. Reine Angew. Math. **429** (1992), 125–160.
2. Philippe Briand and Ying Hu, *Stability of BSDEs with random terminal time and homogenization of semilinear elliptic pdes*, Journal of Functional Analysis **155** (1998), 455–494.

3. Rena A. Carmona, *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications*, SIAM, 2016.
4. Emmanuel Chasseigne and Naoyuki Ichihara, *Sharp estimates of the generalized principal eigenvalue for superlinear viscous hamilton–jacobi equations with inward drift*, *Nonlinear Differential Equations and Applications NoDEA* **29** (2022).
5. Wing Fung Chong, Ying Hu, Gechun Liang, and Thaleia Zariphopoulou, *An ergodic BSDE approach to forward entropic risk measures: Representation and large-maturity behavior*, *Finance and Stochastics* **23** (2019), 239–273.
6. Andrea Cosso, Marco Fuhrman, and Huyên Pham, *Long time asymptotics for fully nonlinear Bellman equations: A backward SDE approach*, *Stochastic Processes and their Applications* **126** (2016), 1932–1973.
7. Arnaud Debussche, Ying Hu, and Gianmario Tessitore, *Ergodic BSDEs under weak dissipative assumptions*, *Stochastic Processes and their Applications* **121** (2010), 407–426.
8. Freddy Delbaen, Ying Hu, and Xiaobo Bao, *Backward sdes with superquadratic growth*, *Probability Theory and Related Fields* **150** (2009).
9. N. El Karoui, S. Peng, and M. C. Quenez, *Backward stochastic differential equations in finance*, *Mathematical Finance* **7** (1997), 1–71.
10. Wendell H. Fleming and William M. McEneaney, *Risk-sensitive control on an infinite time horizon*, *SIAM J. Control. Optim.* **33** (1995), 1881–1915.
11. Marco Fuhrman, Ying Hu, and Gianmario Tessitore, *Ergodic BSDEs and optimal ergodic control in banach spaces*, *SIAM J. Control. Optim.* **48** (2009), 1542–1566.
12. David Gilbarg and Neil S. Trudinger, *Elliptic partial differential equations of second order*, Springer, 1977.
13. Ying Hu, Gechun Liang, and Shanjian Tang, *Systems of ergodic BSDEs arising in regime switching forward performance processes*, *SIAM J. Control. Optim.* **58** (2020), 2503–2534.
14. Ying Hu, Pierre-Yves Madec, and Adrien Richou, *A probabilistic approach to large time behavior of mild solutions of hjb equations in infinite dimension*, *SIAM J. Control. Optim.* **53** (2015), 378–398.
15. Naoyuki Ichihara and Shuenn-Jyi Sheu, *Large time behavior of solutions of hamilton-jacobi-bellman equations with quadratic nonlinearity in gradients*, *SIAM J. Math. Anal.* **45** (2013), 279–306.
16. Hidehiro Kaise and Shuenn-Jyi Sheu, *On the structure of solutions of ergodic type bellman equation related to risk-sensitive control*, *Annals of Probability* **34** (2006), 284–320.
17. Gechun Liang and Thaleia Zariphopoulou, *Representation of homothetic forward performance processes in stochastic factor models via ergodic and infinite horizon BSDE*, *SIAM J. Financial Math.* **8** (2017), 344–372.
18. Adrien Richou, *Ergodic BSDEs and related PDEs with Neumann boundary conditions*, *Stochastic Processes and their Applications* **119** (2008), 2945–2969.
19. Fujita Yasuhiro, Hitoshi Ishii, and Paola Loreti, *Asymptotic solutions of hamilton-jacobi equations in euclidean n space*, *Indiana University Mathematics Journal* **55** (2006), 1671–1700.
20. J. Zhang, *Backward Stochastic Differential Equations*, Springer, 2017.
21. Xicheng Zhang, *Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients*, *Stochastic Processes and their Applications* **115** (2005), 1805–1818.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT AUSTIN
 Email address: jjacks01@utexas.edu

DEPARTMENT OF STATISTICS, UNIVERSITY OF WARWICK
 Email address: g.liang@warwick.ac.uk