ERGODIC CONTROL OF A CLASS OF JUMP DIFFUSIONS WITH FINITE LÉVY MEASURES AND ROUGH KERNELS

ARI ARAPOSTATHIS*, LUIS CAFFARELLI†, GUODONG PANG‡, AND YI ZHENG‡

Abstract. We study the ergodic control problem for a class of jump diffusions in $\mathbb{R}^d$, which are controlled through the drift with bounded controls. The Lévy measure is finite, but has no particular structure—it can be anisotropic and singular. Moreover, there is no blanket ergodicity assumption for the controlled process. Unstable behavior is 'discouraged' by the running cost which satisfies a mild coercive hypothesis, often referred in the literature as a near-monotone condition. We first study the problem in its weak formulation as an optimization problem on the space of infinitesimal ergodic occupation measures, and derive the Hamilton–Jacobi–Bellman equation under minimal assumptions on the parameters, including verification of optimality results, using only analytical arguments. We also examine the regularity of invariant measures. Then, we address the jump diffusion model, and obtain a complete characterization of optimality.

Key words. controlled jump diffusions; compound Poisson process; Lévy process; ergodic control; Hamilton–Jacobi–Bellman equation

AMS subject classifications. Primary, 93E20, 60J75, 35Q93; Secondary, 60J60, 35F21, 93E15

1. Introduction. Optimal control of jump diffusions has recently attracted much attention from the control community, primarily due to its applicability to queueing networks, mathematical finance [1], image processing [2], etc. Many results for the discounted problem are available in [3], including the game theoretic setting, and different applications are discussed. However, studies of the ergodic control problem are rather scarce. Ergodic control of reflected jump diffusions over a bounded domain can be found in [4]. The ergodic control problem in $\mathbb{R}^d$ is studied in [5], albeit under very strong blanket stability assumptions. We should also mention here the treatment of the impulse control problem in [6, 7, 8].

Our work in this paper is motivated from ergodic control problems for multiclass stochastic networks in the Halfin–Whitt regime, under service interruptions. For this model, the pure jump process driving the limiting queueing process is compound Poisson (see Theorem 3.2 in [9]), with a Lévy measure that is anisotropic, and in general, singular with respect to the Lebesgue measure. In fact, the jumps are biased towards a given direction, and thus the Lévy measure has no symmetry whatsoever. We assume that the running cost is coercive, also known as near-monotone (see (2.3)), and do not impose any blanket stability hypotheses on the controlled jump diffusion. We treat a general class of jump diffusions which is abstracted from diffusion approximations of stochastic networks, and whose controlled infinitesimal generator has the form

$$\begin{aligned}\mathcal{A}u(x,z) \coloneqq \sum_{i,j} a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_i b^i(x,z)\frac{\partial u}{\partial x_i}(x) \\ + \int_{\mathbb{R}^d}\bigl(u(x+y) - u(x) - \mathbb{1}_{\{|y|\le 1\}}\langle y, \nabla u(x)\rangle\bigr)\,\nu(x,\mathrm{d}y)\,. \end{aligned} \tag{1.1}$$

*Department of Electrical and Computer Engineering, The University of Texas at Austin, 2501 Speedway, EER 7.824, Austin, TX 78712 (ari@ece.utexas.edu).

†Department of Mathematics, The University of Texas at Austin, 2515 Speedway, RLM 10.150, Austin, TX 78712 (caffarel@math.utexas.edu).

‡The Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, College of Engineering, Pennsylvania State University, University Park, PA 16802 (gup3@psu.edu, yxz282@psu.edu).

Here, z is a control parameter that lives in a compact metric space $\mathcal{Z}$, and $\nu(x, \mathrm{d}y)$ is a nonnegative finite Borel measure on $\mathbb{R}^d$ for each x, while $x \mapsto \nu(x, A)$ is a Borel measurable function for each Borel set A. We let $K \pm x$ denote the Minkowski summation/subtraction for $K \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$. Throughout the paper, we assume that $d \geq 2$. The coefficients of $\mathcal{A}$ are assumed to satisfy the following.

(A1) (a) The matrix $a = [a^{ij}]$ is symmetric, positive definite, and locally Lipschitz continuous. The drift $b\colon \mathbb{R}^d \times \mathcal{Z} \to \mathbb{R}^d$ is continuous.

(b) The map

$$x \mapsto \int_{\mathbb{R}^d} \bigl(f(x+y) - f(x) - \mathbb{1}_{\{|y|\leq 1\}} \langle y, \nabla f(x)\rangle\bigr)\, \nu(x, \mathrm{d}y)$$

is continuous and bounded for all $f \in \mathcal{C}^2_c(\mathbb{R}^d)$, the space of functions having compact support and continuous second order partial derivatives.

(b) The map $x \mapsto \bar{\nu}(x) \coloneqq \nu(x, \mathbb{R}^d)$ is locally bounded.

(c) the map $x \mapsto \nu(x, K - x)$ is bounded on $\mathbb{R}^d$ for any fixed compact set $K \subset \mathbb{R}^d$.

We compare (A1) to the following weaker hypothesis.

(A1′) Assume part (a) of (A1), but part (b) is replaced by:

(i) The map $x \mapsto \bar{\nu}(x) \coloneqq \nu(x, \mathbb{R}^d)$ is locally bounded.

(ii) the map $x \mapsto \nu(x, K - x)$ is bounded on $\mathbb{R}^d$ for any fixed compact set $K \subset \mathbb{R}^d$.

The structural assumption on the coefficients in (A1) is in effect by default throughout the paper. However, Assuming (A1′) only, we establish existence and uniqueness of solutions to the ergodic HJB equation in Theorem 3.3. Note that under (A1′), the operator $\mathcal{A}$ in (1.1) maps $\mathcal{C}^2_c(\mathbb{R}^d)$ to bounded Borel measurable functions in $\mathbb{R}^d \times \mathcal{Z}$, while under (A1) it maps the same space to continuous and bounded functions. Continuity is essential for the standard convex analytic argument which shows the existence of an optimal stationary Markov control, and this is the only place where the stronger hypothesis (A1) need be used (see Theorem 2.3).

On the other hand, concerning solutions to the martingale problem associated with $\mathcal{A}$, fairly general results can be found in [10]. It can be seen from these results that the kernel ν may be rough enough so as not to satisfy (A1), yet the martingale problem has a solution. If this is the case, then even though we cannot follow the convex analytic argument [11], the existence of an optimal stationary Markov control could be asserted from the HJB equation. If, for example, the solution of the HJB is inf-compact, or if the drift b has at most linear growth, then a control which is a measurable selector from the minimizer renders the process ergodic, and it is standard to show that it is optimal. Therefore, we mention at various places how the weaker hypotheses of (A1′) can be used to establish the results.

The generator $\mathcal{A}$ in (1.1) covers a variety of models of jump diffusions which appear in the literature [12, 13, 14, 15, 16]. Note also that the 'jump rate' $\bar{\nu}(x) \coloneqq \nu(x, \mathbb{R}^d)$ is allowed to be state dependent as in [17]. The hypotheses in (A1) are quite general, and do not imply the existence of a controlled process with generator $\mathcal{A}$. Our main goal in this paper is to establish general results for ergodic control of jump diffusions governed for this class of operators. To accomplish this, we first state the ergodic control problem for the operator $\mathcal{A}$ as a convex optimization problem over the set of infinitesimal ergodic occupation measures. We then proceed to study the ergodic Hamilton–Jacobi–Bellman (HJB) equation via analytical methods, without assuming that the martingale problem for $\mathcal{A}$ is well posed. This of course precludes arguments that utilize stochastic representations of solutions of elliptic equations. Later, in section 4, we specialize these results to a fairly general model of controlled jump diffusions with finite Lévy measure.

It is well known that the standard method of deriving the ergodic HJB on $\mathbb{R}^d$ is based on the vanishing discount approach, and relies crucially on structural properties that permit uniform estimates for the gradient (e.g., viscous equations in $\mathbb{R}^d$), or the Harnack property. Recent work on nonlocal equations has resulted in important regularity results [18, 19, 20, 21] that should prove very valuable in studying control problems. However, most of this work concerns Lévy jump processes whose kernel has a 'nice' density resembling that of a fractional Laplacian. For the problem at hand, even though the Lévy measure $\nu(x,\cdot\,)$ is finite, and there is a nondegenerate Wiener process component, the Lévy measure is anisotropic, and could be singular [9, Section 3.2]. As a result, there is no hope for the Harnack property for positive solutions to hold as the following example shows.

Example 1.1. Consider an operator $\mathcal{A}$ in $\mathbb{R}^2$, with a the identity matrix, $b = (3,0)$, and $\nu(x,\cdot\,)$ a Dirac mass at $\tilde{x} = (3,0)$. Let $f_\epsilon \in \mathcal{C}^2(\mathbb{R}^2)$, with $\epsilon \in (0,1)$, be defined in polar coordinates by

$$f_\epsilon(r,\theta) \;\coloneqq\; -\log(r)\,\mathbb{1}_{\{r\ge\epsilon\}} + \Bigl(\tfrac{3}{4} - \tfrac{r^2}{\epsilon^2} + \tfrac{r^4}{4\epsilon^4} - \log(\epsilon)\Bigr)\,\mathbb{1}_{\{r<\epsilon\}}\,.$$

This function is used in [22, p. 111] to exhibit a family of positive superharmonic functions for the Laplacian that violates the Harnack property. Let u_ϵ be a function which agrees with f_ϵ on the unit ball B_1 centered at 0, and takes the values $u_\epsilon(\tilde{r},\tilde{\theta}) = \bigl(\frac{4}{\epsilon^2} - \frac{4\tilde{r}^2}{\epsilon^4} + f_\epsilon(\tilde{r},\tilde{\theta})\bigr)\,\mathbb{1}_{\{\tilde{r}<\epsilon\}}$ on the unit ball $B_1(\tilde{x})$ centered at $\tilde{x}$, when expressed in polar coordinates $(\tilde{r},\tilde{\theta})$ which are centered at $\tilde{x}$. Let u_ϵ take any nonnegative value elsewhere in $\mathbb{R}^2$. Then u_ϵ is nonnegative on $\mathbb{R}^2$ and satisfies $\mathcal{A}u_\epsilon = 0$ in B. However, $\frac{u_\epsilon(0,\theta)}{u_\epsilon(\mathrm{e}^{-1},\theta)} = -\log(\epsilon)$, and thus the family violates the Harnack property for $\mathcal{A}$.

Under the general hypotheses of (A1), even if the operator $\mathcal{A}$ is the generator of a Markov process, the process might not be regular, or, in case it is positive recurrent, the mean hitting times to an open ball might not be locally bounded. In the latter case, it is futile to search for solutions to the ergodic HJB equation, even in a viscosity sense. In section 3, we add two hypotheses to address these pathologies. The first (see (H1)), is the Feller–Has′minskiĭ criterion for a diffusion process with generator $\mathcal{A}$ to be *regular* (or *conservative*, or *non-explosive*), which requires that the equation $\mathcal{A}u - u = 0$ has no bounded positive solutions on $\mathbb{R}^d$. This property is equivalent to regularity, and it is clear from the proof of this equivalence in [23, Theorem 4.1] that the equation can be replaced by $\mathcal{A}u - \alpha u = 0$ for $\alpha > 0$. The second hypothesis, (H2), states that under some stationary Markov control there exists a nonnegative solution $\mathcal{V}$ to the Lyapunov equation $\mathcal{A}\mathcal{V} \le C\mathbb{1}_{\mathcal{B}} - \mathcal{R}$, where $\mathcal{R}$ is the running cost, $\mathcal{B}$ is a ball, and C is a constant. We relax (H2), after imposing additional assumptions on ν, and establish solutions of the Poisson equation and the HJB (see Theorems 3.6 and 3.8).

The paper is organized as follows. In subsection 1.1 we summarize the notation we use. Section 2 states the ergodic control problem, in a weak sense, as a convex optimization problem over the set of infinitesimal ergodic occupation measures for the operator $\mathcal{A}$, and shows that optimality is attained. Regularity properties of infinitesimal invariant measures are in subsection 2.3. Section 3 is devoted to the study of the HJB equation under (H1)–(H2) mentioned above. In Section 4 we study a class of jump diffusions, which is abstracted from the limiting diffusions encountered in stochastic networks under service interruptions.

1.1. Notation. The standard Euclidean norm in $\mathbb{R}^d$ is denoted by $|\cdot|$, and $\langle\cdot\,,\cdot\rangle$ denotes the inner product. Given two real numbers a and b, the minimum (maximum) is denoted by $a \wedge b$ ($a \vee b$), respectively. The closure, boundary, complement, and the indicator function of a set $A \subset \mathbb{R}^d$ are denoted by $\bar{A}$, ∂A, A^c, and $\mathbb{1}_A$, respectively. We denote by $\uptau(A)$ the *first exit time* of the process X from a set $A \subset \mathbb{R}^d$, defined by $\uptau(A) := \inf\{t > 0\colon X_t \notin A\}$. The open ball of radius R in $\mathbb{R}^d$, centered at the origin, is denoted by B_R, and we let $\uptau_R := \uptau(B_R)$, and $\breve{\uptau}_R := \uptau(B_R^c)$. The Borel σ-field of a topological space E is denoted by $\mathfrak{B}(E)$, and $\mathcal{P}(E)$ denotes the set of probability measures on $\mathfrak{B}(E)$.

For a domain $Q \subset \mathbb{R}^d$, the space $\mathcal{C}^k(Q)$ ($\mathcal{C}^\infty(Q)$), $k \ge 0$, refers to the class of all real-valued functions on Q whose partial derivatives up to order k (of any order) exist and are continuous, while $\mathcal{C}^k_c(Q)$ ($\mathcal{C}^k_b(Q)$) denote the subsets of $\mathcal{C}^k(Q)$, consisting of functions that have compact support (whose partial derivatives are bounded in Q). The space $L^p(Q)$, $p \in [1,\infty)$, stands for the Banach space of (equivalence classes of) measurable functions f satisfying $\int_Q |f(x)|^p\,\mathrm{d}x < \infty$, and $L^\infty(Q)$ is the Banach space of functions that are essentially bounded in Q. We denote the usual norm on this space by $\|f\|_{L^p(Q)}$, $p \in [1,\infty]$. The standard Sobolev space of functions on Q whose generalized derivatives up to order k are in $L^p(Q)$, equipped with its natural norm, is denoted by $\mathscr{W}^{k,p}(Q)$, $k \ge 0$, $p \ge 1$. In general, if $\mathcal{X}$ is a space of real-valued functions on Q, $\mathcal{X}_{\mathrm{loc}}$ consists of all functions f such that $f\varphi \in \mathcal{X}$ for every $\varphi \in \mathcal{C}^\infty_c(Q)$. In this manner we obtain, for example, the space $\mathscr{W}^{2,p}_{\mathrm{loc}}(Q)$.

We adopt the notation $\partial_i := \frac{\partial}{\partial x_i}$ and $\partial_{ij} := \frac{\partial^2}{\partial x_i \partial x_j}$ for $i, j \in \{1,\dots,d\}$, and we often use the standard summation rule that repeated subscripts and superscripts are summed from 1 through d.

2. The convex analytic formulation. Define $\mathcal{L}\colon \mathcal{C}^2(\mathbb{R}^d) \to \mathcal{C}(\mathbb{R}^d \times \mathcal{Z})$ by

$$\mathcal{L}u(x,z) \;:=\; a^{ij}(x)\partial_{ij}u(x) + \hat{b}^i(x,z)\partial_i u(x)\,,$$

with $\hat{b}(x,z) := b(x,z) + \int_{\mathbb{R}^d} y\,\mathbb{1}_{\{|y|\le 1\}}\nu(x,\mathrm{d}y)$, and let

$$\mathcal{I}u(x) \;:=\; \int_{\mathbb{R}^d} \bigl(u(x+y) - u(x)\bigr)\,\nu(x,\mathrm{d}y)\,,$$

provided that the integral is finite. Thus $\mathcal{A}u(x,z) = \mathcal{L}u(x,z) + \mathcal{I}u(x)$. With $z \in \mathcal{Z}$ treated as a parameter, we define $\mathcal{L}_z u(x) := \mathcal{L}u(x,z)$, and $\mathcal{A}_z u(x) := \mathcal{A}u(x,z)$.

Let $\mathcal{B}(\mathbb{R}^d,\mathcal{Z})$ denote the set of Borel measurable maps $v\colon \mathbb{R}^d \to \mathcal{Z}$. Such a map v is called a *stationary Markov control*, and we use the symbol $\mathfrak{U}_{\mathsf{sm}}$ to denote this class of controls. For $v \in \mathfrak{U}_{\mathsf{sm}}$, we use the simplified notation $b_v(x) := b\bigl(x, v(x)\bigr)$, and define $\mathcal{A}_v$, $\mathcal{R}_v$ and ϱ_v analogously.

We augment the class $\mathfrak{U}_{\mathsf{sm}}$ by adopting the well known *relaxed control* framework [24, Section 2.3]. According to this relaxation, controls take values in $\mathcal{P}(\mathcal{Z})$, the latter denoting the set of probability measures on $\mathcal{Z}$ under the Prokhorov topology. Thus, a control $v \in \mathfrak{U}_{\mathsf{sm}}$ may be viewed as a kernel on $\mathcal{P}(\mathcal{Z}) \times \mathbb{R}^d$, which we write as $v(\mathrm{d}z\,|\,x)$. We extend the definition of b and $\mathcal{R}$, without changing the notation, that is, we let $b_v(x) := \int_{\mathcal{Z}} b(x,z)\,v(\mathrm{d}z\,|\,x)$, and analogously for $\mathcal{R}_v$. We endow $\mathfrak{U}_{\mathsf{sm}}$ with the topology that renders it a compact metric space, referred to as the *topology of Markov controls* [24, Section 2.4]. A control is said to be *precise* if it is a measurable map from $\mathbb{R}^d$ to $\mathcal{Z}$, that is, if it agrees with the definition in the preceding paragraph. It is easy to see that this relaxation preserves (A1).

2.1. The ergodic control problem for the operator $\mathcal{A}$. We fix a countable dense subset $\mathcal{C}$ of $\mathcal{C}_0^2(\mathbb{R}^d)$ consisting of functions with compact supports. Here, $\mathcal{C}_0^2(\mathbb{R}^d)$ denotes the Banach space of functions $f\colon \mathbb{R}^d \to \mathbb{R}$ that are twice continuously differentiable and their derivatives up to second order vanish at infinity.

Definition 2.1. *A probability measure $\mu_v \in \mathcal{P}(\mathbb{R}^d)$, $v \in \mathfrak{V}_{\mathsf{sm}}$, is called* infinitesimally invariant *under $\mathcal{A}_v$ if*

$$\int_{\mathbb{R}^d} \mathcal{A}_v f(x)\, \mu_v(\mathrm{d}x) \;=\; 0 \quad \forall\, f \in \mathcal{C}\,. \tag{2.1}$$

If such a μ_v exists, then we say that v is a stable *control, and define the* (infinitesimal) ergodic occupation measure *$\uppi_v \in \mathcal{P}(\mathbb{R}^d \times \mathcal{Z})$ by $\uppi_v(\mathrm{d}x, \mathrm{d}z) \coloneqq \mu_v(\mathrm{d}x)\, v(\mathrm{d}z \,|\, x)$. We denote by $\mathfrak{V}_{\mathsf{ssm}}$, $\mathcal{M}$, and $\mathcal{G}$, the sets of stable controls, infinitesimal invariant probability measures, and ergodic occupation measures, respectively.*

Remark 2.2. In Definition 2.1 we select $\mathcal{C}$ as the function space, deviating from common practice, where this is selected as $\mathcal{C}_0^\infty(\mathbb{R}^d)$, the space of smooth functions vanishing at infinity. In general, there is no uniqueness of solutions to (2.1) [25]. For the relation between infinitesimally invariant measures and invariant probability measures for diffusions we refer the reader to [26]. Note also, that as shown in [27], in order to assert that μ_v is an invariant probability measure for a Markov process with generator $\mathcal{A}_v$, it suffices to verify (2.1) for a dense subclass of the domain of $\mathcal{A}_v$ consisting of functions such that the martingale problem is well posed.

It follows from Definition 2.1 that $\uppi \in \mathcal{P}(\mathbb{R}^d \times \mathcal{Z})$ is an ergodic occupation measure if and only if $\int_{\mathbb{R}^d\times\mathcal{Z}} \mathcal{A}_z f(x)\, \uppi(\mathrm{d}x, \mathrm{d}z) = 0$ for all $f \in \mathcal{C}$. It is also easy to show that the set of ergodic occupation measures $\mathcal{G}$ is a closed and convex subset of $\mathcal{P}(\mathbb{R}^d \times \mathcal{Z})$ (see [24, Lemma 3.2.3]).

Let $\mathcal{R}\colon \mathbb{R}^d \times \mathcal{Z} \mapsto \mathbb{R}_+$ be a continuous function, which we refer to as the *running cost* function. The ergodic control problem for $\mathcal{A}$ seeks to minimize $\uppi(\mathcal{R}) = \int \mathcal{R}\, \mathrm{d}\uppi$ over $\uppi \in \mathcal{G}$. Thus, the optimization problem is an infinite dimensional LP. We define

$$\underline{\varrho} \;\coloneqq\; \inf_{\uppi\in\mathcal{G}}\, \uppi(\mathcal{R})\,, \tag{2.2}$$

and assume, of course, that this is finite. Also for $v \in \mathfrak{V}_{\mathsf{ssm}}$, we let $\varrho_v \coloneqq \uppi_v(\mathcal{R})$, and we say that v is *optimal* if $\varrho_v = \underline{\varrho}$. We seek to obtain a full characterization of optimal controls via the study of the dual problem, and this leads to the HJB equation. For more details on this LP formulation see Section 4 in [11].

2.2. Well posedness of the control problem. We impose a structural assumption on the running cost which renders the optimization problem well posed. We say that a function $h : \mathbb{R}^d \times \mathcal{Z} \to \mathbb{R}_+$ is *coercive relative to a constant* $c \in \mathbb{R}$, if there exists a constant $\epsilon > 0$, such that the set $\{x \in \mathbb{R}^d : \inf_{z\in\mathcal{Z}}\, h(x, z) \le c + \epsilon\}$ is bounded (or empty).

Throughout the paper, we assume that the running cost is coercive relative to $\underline{\varrho}$, and we fix a ball $\mathcal{B}_\circ$ and a constant $\epsilon_\circ$ such that $\mathcal{R}(x, z) > \underline{\varrho} + 2\epsilon_\circ$ on $\mathcal{B}_\circ^c$. Naturally, this property depends on $\underline{\varrho}$, but note that, since $\underline{\varrho} < \infty$, it is always satisfied if the running cost is inf-compact on $\mathbb{R}^d \times \mathcal{Z}$. Coerciveness of $\mathcal{R}$ relative to $\underline{\varrho}$ is also known as *near-monotonicity* in the literature, and it is often written as

$$\liminf_{|y|\to\infty}\; \inf_{z\in\mathcal{Z}}\, \mathcal{R}(y, z) \;>\; \underline{\varrho}\,. \tag{2.3}$$

We state the following theorem, which follows easily by mimicking the proofs of Lemma 3.2.11 and Theorem 3.4.5 in [24], since the map $(x,z)\mapsto \mathcal{A}f(x,z)$ is continuous and bounded for each $f\in\mathcal{C}^\infty(\mathbb{R}^d)$ by (A1).

THEOREM 2.3. *The map* $\uppi\mapsto\uppi(\mathcal{R})$ *attains its minimum in* $\mathcal{G}$.

2.3. Regularity properties of infinitesimal invariant measures. In this section we establish regularity properties of the densities of infinitesimal invariant probability measures. Recall the notation $\bar\nu(x)=\nu(x,\mathbb{R}^d)$ introduced in (A1). We say that ν is *translation invariant* if $\nu(x,\cdot\,)$ does not depend on x, in which case we denote it simply as $\nu(\mathrm{d}y)$. We need the following definition.

DEFINITION 2.4. *We decompose* $\mathcal{A}_z=\widetilde{\mathcal{L}}_z+\widetilde{\mathcal{I}}$, *with*

$$\widetilde{\mathcal{L}}_z u(x) \coloneqq \mathcal{L}_z u(x)-\bar\nu(x)u(x)\,,\qquad \text{and}\quad \widetilde{\mathcal{I}}u(x)\coloneqq\int_{\mathbb{R}^d}u(x+y)\,\nu(x,\mathrm{d}y)\,.$$

THEOREM 2.5. *Every* $\mu\in\mathcal{M}$ *has a density* $\upphi=\upphi[\mu]$ *which belongs to* $L^p_{loc}(\mathbb{R}^d)$ *for any* $p\in\bigl[1,\frac{d}{d-2}\bigr)$, *and is strictly positive. In addition, if* ν *is translation invariant and has compact support, then, for any* $\beta\in(0,1)$, *there exists a constant* $\bar C=\bar C(\beta,R)$, *such that*

$$|\upphi(x)-\upphi(y)|\le\bar C\,|x-y|^\beta\qquad\forall\,x,y\in B_R\,. \tag{2.4}$$

Proof. As shown in [28, Theorem 2.1], if in some domain $Q\subset\mathbb{R}^d$, a probability measure μ satisfies

$$\int_Q a^{ij}\partial_{ij}f\,\mathrm{d}\mu\;\le\;C\,\sup_Q\bigl(|f|+|\nabla f|\bigr)\qquad\forall f\,\in\mathcal{C}^\infty_c(Q) \tag{2.5}$$

for some constant C, then μ has a density which belongs to $L^p_{\mathrm{loc}}(Q)$ for every $p\in[1,d')$, where $d'=\frac{d}{d-1}$. It is straightforward to verify, using only (A1′), that a bound of the form (2.5) holds for any $\mu\in\mathcal{M}$ on any bounded domain Q. It follows that the density $\upphi$ of μ is in $L^p_{\mathrm{loc}}(\mathbb{R}^d)$ for any $p\in[1,d')$, and that it is a generalized solution to the equation

$$\begin{aligned}\sum_{i,j}\int_{\mathbb{R}^d}\Bigl(a^{ij}(x)\partial_j\upphi(x)+\bigl(\partial_j a^{ij}(x)-\hat b^i_v(x)\bigr)\upphi(x)\Bigr)\partial_i f(x)\,\mathrm{d}x&\\ -\int_{\mathbb{R}^d}\bar\nu(x)\upphi(x)f(x)\,\mathrm{d}x\;=\;-\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}f(x+y)\nu(x,\mathrm{d}y)\,\upphi(x)\,\mathrm{d}x\,,&\end{aligned} \tag{2.6}$$

for $f\in\mathcal{C}^\infty_c(\mathbb{R}^d)$. By (2.6), $\upphi$ is a supersolution to

$$\widetilde{\mathcal{L}}^*_v\upphi(x)\coloneqq\partial_i\Bigl(a^{ij}(x)\partial_j\upphi(x)+\bigl(\partial_j a^{ij}(x)-\hat b^i_v(x)\bigr)\upphi(x)\Bigr)-\bar\nu(x)\upphi(x)\;=\;0\,. \tag{2.7}$$

Therefore, by the estimate for supersolutions in [29, Theorem 8.18], we deduce that $\upphi\in L^p_{\mathrm{loc}}(\mathbb{R}^d)$ for any $p\in\bigl[1,\frac{d}{d-2}\bigr)$, and that it is strictly positive. Note that this theorem assumes that the supersolution is in $\mathcal{W}^{1,2}_{\mathrm{loc}}(\mathbb{R}^d)$, but this is unnecessary. The theorem is valid for functions in $\mathcal{W}^{1,p}_{\mathrm{loc}}(\mathbb{R}^d)$ for any $p>1$, as seen from the results in Section 5.5 of [30], or one can use the mollifying technique in [24, Theorem 5.3.4] to show this.

Now suppose that ν is translation invariant and has compact support. Let $\widehat{\mathcal{I}}\upphi(x)\coloneqq\int_{\mathbb{R}^d}\upphi(x-y)\,\nu(\mathrm{d}y)$. Then (2.6) takes the form $\widetilde{\mathcal{L}}^*_v\upphi(x)=-\widehat{\mathcal{I}}\upphi(x)$. The operator $\widetilde{\mathcal{L}}_v^*$ satisfies the hypotheses of Theorem 5.5.5′ in [30], which asserts that ϕ satisfies

$$\|\phi\|_{\mathcal{W}^{1,q}(B_R)} \le \kappa(p,R)\bigl(\|\widehat{\mathcal{I}}\phi\|_{L^p(B_{2R})} + \|\phi\|_{L^1(B_{2R})}\bigr) \qquad \forall\, p>1\,, \tag{2.8}$$

with $q = q(p) \coloneqq \frac{dp}{d-p}$, and a constant $\kappa(p,R)$ that depends also on d, $\bar{\nu}$, and the bounds in (A1′). Without loss of generality, suppose that ν is supported on a ball $B_{R_\circ}$. By Minkowski's integral inequality we have

$$\|\widehat{\mathcal{I}}\phi\|_{L^p(B_{2R})} \le \bar{\nu}\,\|\phi\|_{L^p(B_{2R+R_\circ})}\,. \tag{2.9}$$

On the other hand, by the Sobolev embedding theorem, $\mathcal{W}^{1,q}(B_R) \hookrightarrow L^r(B_R)$ is a continuous embedding for $q \le r \le \frac{qd}{d-q}$ and $q<d$, and $\mathcal{W}^{1,q}(B_R) \hookrightarrow \mathcal{C}^{0,r}(\overline{B}_R)$ is compact for $r < 1-\frac{d}{q}$ and $q>d$. Therefore, starting say from $p=\frac{d}{d-1}$, we deduce by repeated applications of (2.8)–(2.9), and Sobolev embedding, that $\phi \in \mathcal{W}^{1,q}_{\mathrm{loc}}(\mathbb{R}^d)$ for any $q>1$, which implies (2.4). ∎

Remark 2.6. Consider a jump diffusion with $\upsigma = \sqrt{2}$, $b(x) = x\mathbb{1}_{\bar{B}_1}(x)$, and $\nu(x,\mathrm{d}y) = \delta_{-x}$, where δ_{-x} denotes the Dirac mass at $-x$. Then $\mathcal{A} = \Delta - 1 + \delta_0$. It can be easily verified that the diffusion is geometrically ergodic by employing the Lyapunov function $\mathcal{V}(x) = |x|^2$. The density of the invariant measure ϕ satisfies $\int\sum_{ij}(\partial_i\phi)(\partial_j f) + \int \phi f = f(0)$ for all $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, and thus it is a solution of $-\Delta\phi + \phi = \delta_0$ (viewed in the sense of distributions $\mathcal{D}'(\mathbb{R}^d)$). However, as shown in [31], every positive solution ϕ of this equation, which vanishes at infinity, satisfies $\phi(x) \sim \Gamma(x)$ as $x \to 0$, where Γ denotes the fundamental solution of $-\Delta$ in $\mathbb{R}^d$. Thus the density of the invariant measure in the vicinity of $x=0$ is not any better than what is claimed in the first step in the proof, which shows that it belongs to $L^p_{\mathrm{loc}}(\mathbb{R}^d)$ for $p < \frac{d}{d-2}$. One can select the jumps to induce multiple such singularities, and generate very pathological examples. Thus, in general, the hypothesis that ν is translation invariant cannot be relaxed in Theorem 2.5, unless we assume that ν has a suitable density as shown in Corollary 2.8 below.

Definition 2.7. *We say that ν has* locally compact support *if there exists an increasing map $\gamma\colon(0,\infty)\to(0,\infty)$ such that $\nu(x, x+B^c_{\gamma(R)}) = 0$ for all $x\in B_R$. Let $\widehat{\gamma}(R) \coloneqq R+\gamma(R)$. It follows from this definition that $B_{\widehat{\gamma}(R)}$ contains the support of ν for all $x \in B_R$.*

Corollary 2.8. *Assume that ν has locally compact support, and that it has a density $\psi_x \in L^{p_1}_{loc}(\mathbb{R}^d)$ for some $p_1 > \frac{d}{2}$, satisfying the following: for some $p_2 \in \bigl(1,\frac{d}{d-2}\bigr)$, it holds that*

$$\int_{B_{\gamma(R)}}\biggl(\int_{B_{\widehat{\gamma}(R)}} |\psi_x(y)|^{p_i}\,\mathrm{d}y\biggr)^{\frac{1}{p_i-1}}\mathrm{d}x < \infty\,,\quad i=1,2\,,\quad \forall R>0\,.$$

Then (2.4) holds.

Proof. Note that

$$\begin{aligned}\int_{\mathbb{R}^d}\biggl(\int_{\mathbb{R}^d} f(x+y)\psi_x(y)\,\mathrm{d}y\biggr)\phi(x)\,\mathrm{d}x &= \int_{\mathbb{R}^d} f(z)\biggl(\int_{\mathbb{R}^d}\psi_{z-y}(y)\,\phi(z-y)\,\mathrm{d}y\biggr)\mathrm{d}z\\ &= \int_{\mathbb{R}^d} f(z)\biggl(\int_{\mathbb{R}^d}\psi_z(z-a)\,\phi(a)\,\mathrm{d}a\biggr)\mathrm{d}z\,.\end{aligned}$$

Therefore, $\widehat{\mathcal{I}}\phi(x) = \int_{\mathbb{R}^d} \psi_a(x-a)\,\phi(a)\,\mathrm{d}a$. By the Minkowski integral inequality and the Hölder inequality, we obtain

$$\begin{aligned}
\bigl\|\widehat{\mathcal{I}}h(z)\bigr\|_{L^p(B_R)} &= \biggl(\int_{B_R}\biggl(\int_{B_{\gamma(R)}} \psi_a(z-a)\,|h(a)|\,\mathrm{d}a\biggr)^p \mathrm{d}z\biggr)^{1/p}\\
&\le \int_{B_{\gamma(R)}} |h(a)|\biggl(\int_{B_R} |\psi_a(z-a)|^p\,\mathrm{d}z\biggr)^{1/p}\mathrm{d}a\\
&\le \|h\|_{L^p(B_{\gamma(R)})}\biggl(\int_{B_{\gamma(R)}}\biggl(\int_{B_R} |\psi_a(z-a)|^p\,\mathrm{d}z\biggr)^{1/(p-1)}\mathrm{d}a\biggr)^{(p-1)/p}\\
&\le \|h\|_{L^p(B_{\gamma(R)})}\biggl(\int_{B_{\gamma(R)}} \|\psi_a\|^{p/(p-1)}_{L^p(B_{\widehat{\gamma}(R)})}\,\mathrm{d}a\biggr)^{(p-1)/p}.
\end{aligned}$$

Therefore, the map $\widehat{\mathcal{I}}h$ is a linear mapping from $L^{p_1}(B_{\gamma(R)}) \cup L^{p_2}(B_{\gamma(R)})$ into $L^{p_1}(B_R) \cup L^{p_2}(B_R)$ and satisfies

$$\bigl|\bigl\{x\in B_R : |\widehat{\mathcal{I}}h(x)| > t\bigr\}\bigr| \le C\,\frac{\|h\|_{L^{p_i}(B_{\gamma(R)})}}{t^{\,p_i}}$$

for some constant C, for all $h \in L^{p_i}(B_R)$, $i = 1,2$. Here, $|A|$ denotes the Lebesgue measure of a set A. Thus, by the Marcinkiewicz interpolation theorem, it extends to a bounded linear map from $L^p(B_{\gamma(R)})$ into $L^p(B_R)$ for any $p \in (p_1,p_2)$. The result then follows as in the proof of Theorem 2.5. □

Remark 2.9. It is evident from Corollary 2.8 that if ν has locally compact support and a density $\psi_x \in L^p(\mathbb{R}^d)$ for some $p > \frac{d}{2}$, such that $x \mapsto \|\psi_x\|_{L^p(\mathbb{R}^d)}$ is locally bounded, then the density of an infinitesimal invariant measure is Hölder continuous.

3. The HJB equations. We first discuss the relationship between infinitesimal invariant probability measures and Foster–Lyapunov equations. Next, we derive the α-discounted HJB equation, and proceed to study the ergodic HJB equation using the vanishing discount approach. The treatment is analytical, and we refrain from using any stochastic representations of solutions. We state hypothesis (H1) which was discussed in section 1.

(H1) For any $v \in \mathfrak{V}_{\mathsf{sm}}$, and $\alpha > 0$, the equation $\mathcal{A}_v u - \alpha u = 0$ has no bounded positive solution $u \in \mathcal{W}^{2,d}_{\mathrm{loc}}(\mathbb{R}^d)$.

3.1. On the Foster–Lyapunov equation. Consider the hypothesis:

(H2) There exist $\hat{v} \in \mathfrak{V}_{\mathsf{sm}}$, a nonnegative $\mathcal{V} \in \mathcal{C}^2(\mathbb{R}^d)$, and a positive constant κ_0 such that

$$\mathcal{A}_{\hat{v}}\mathcal{V}(x) \le \kappa_0 \mathbb{1}_{\mathcal{B}_\circ}(x) - \mathcal{R}_{\hat{v}}(x) \quad \forall\, x\in\mathbb{R}^d\,, \tag{3.1}$$

where without loss of generality, and in the interest of notational simplicity, we use the same ball $\mathcal{B}_\circ$ as the one introduced in subsection 2.2.

On the other hand, $\underline{\varrho}$ is finite if and only if

(H3) There exist $\hat{v} \in \mathfrak{V}_{\mathsf{ssm}}$, and a probability measure $\mu_{\hat{v}}$ which solves (2.1), and $\mu_{\hat{v}}(\mathcal{R}_{\hat{v}}) = \int \mathcal{R}_{\hat{v}}\,\mathrm{d}\mu_{\hat{v}} < \infty$.

For continuous diffusions, equivalence of (H2) and (H3) is a celebrated result of Has′minskiĭ [32]. It is pretty straightforward to show, using probabilistic arguments, that (H2) ⇒ (H3), and this is in fact true for a large class of Markov processes. An analytical argument for continuous diffusions can be found in the work of Bogachev and Röckner [33], under the hypothesis that $\mathcal{R}_{\hat{v}}$ is inf-compact. The argument offered by Has′minskiĭ in the proof that (H3) ⇒ (H2) relies crucially on the Harnack property, and therefore is not applicable for the jump diffusions considered here. In the context of general Markov processes, existence of a solution to (3.1) is related to the f-regularity of the process. For recent work on this, see [34].

In some sense, (H2) is a very mild assumption, since in any application one would first need to establish that $\underline{\varrho}$ is finite, and the natural venue for this is via the Foster–Lyapunov equation in (3.1). A typical example is when ν is translation invariant, a has sublinear growth, and for some $\theta \in [1,2]$, $\int_{\mathbb{R}^d} |y|^\theta \nu(\mathrm{d}y) < \infty$, $\mathcal{R}_{\hat{v}}$ grows at most as $|x|^{2(\theta-1)}$, and there exist a positive definite symmetric matrix S, and positive constants c_0 and c_1 such that $\langle b_{\hat{v}}(x), Sx\rangle \le c_0 - c_1|x|^\theta$. Then (3.1) holds with $\mathcal{V}(x) = \langle x, Sx\rangle^{\theta/2}$. For other examples, see [9, Corollary 5.1].

Consider the class of ν that are either translation invariant and have compact support, or satisfy the hypotheses of Corollary 2.8, and denote it by $\mathfrak{N}_0$ for convenience. For $\nu \in \mathfrak{N}_0$, we bridge the gap between (H2) and (H3) in Theorem 3.6 by establishing the existence of a solution to the Poisson equation, and thus showing that (H3) ⇒ (H2), albeit for a function $\mathcal{V} \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$. This however is enough to relax (H2) in asserting the existence of a solution to the ergodic HJB for $\nu \in \mathfrak{N}_0$ (Theorem 3.8). Moreover, the proof of Theorem 3.8 contains an analytical argument which shows that (H2) ⇒ (H3), provided that $\nu \in \mathfrak{N}_0$, and $\mathcal{R}_{\hat{v}}$ is inf-compact.

We need the following simple assertion.

LEMMA 3.1. *Let μ_v be an infinitesimal invariant measure under $v \in \mathfrak{V}_{\mathsf{ssm}}$. Then (2.1) holds for all $\varphi \in \mathcal{W}^{2,p}_{loc}(\mathbb{R}^d) \cap \mathcal{C}_c(\mathbb{R}^d)$, $p > d$. In addition, if $\varphi \in \mathcal{W}^{2,p}_{loc}$, $p > d$, is inf-compact, and such that $\mathcal{A}_v\varphi \in L^d_{loc}(\mathbb{R}^d)$, and is nonpositive a.e. on the complement of some ball $\mathcal{B} \subset \mathbb{R}^d$, then $\mu_v\bigl(|\mathcal{A}_v\varphi|\bigr) < \infty$.*

Proof. In the interest of simplicity, we drop the explicit dependence on v in the notation. Suppose $\varphi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d) \cap \mathcal{C}_c(\mathbb{R}^d)$, $p > d$. Let ρ be a symmetric nonnegative mollifier supported on the unit ball centered at the origin, and for $\epsilon > 0$, let $\rho_\epsilon(x) \coloneqq r^{-d}\rho(\frac{x}{\epsilon})$, and $\varphi_\epsilon \coloneqq \rho_\epsilon * \varphi$, where '$*$' denotes convolution. Then, $\mu(\mathcal{A}\varphi_\epsilon) = 0$ by (2.1). Since $\partial_{ij}\varphi_\epsilon$ converges to $\partial_{ij}\varphi$ as $\epsilon \searrow 0$ in $L^p(B_R)$ for any $p > 1$ and $R > 0$, and since μ has a density $\phi \in L^{\frac{d}{d-1}}_{\mathrm{loc}}(\mathbb{R}^d)$ by Theorem 2.5, it follows by Hölder's inequality that $\int_{\mathbb{R}^d} |a^{ij}||\partial_{ij}\varphi - \partial_{ij}\varphi_\epsilon|\,\mathrm{d}\mu \to 0$ as $\epsilon \searrow 0$. Also, since $\partial_i\varphi - \partial_i\varphi_\epsilon$ converges uniformly to 0, and in view of (A1′) (b) and (c), we obtain $\mu(\hat{b}^i\partial_i\varphi_\epsilon) \to \mu(\hat{b}^i\partial_i\varphi)$, and $\mu(\mathcal{I}\varphi_\epsilon) \to \mu(\mathcal{I}\varphi)$ as $\epsilon \searrow 0$. This shows that $\mu(\mathcal{A}\varphi) = 0$.

We now turn to the second statement of the lemma. Let χ be a concave $\mathcal{C}^2(\mathbb{R}^d)$ function such that $\chi(x) = x$ for $x \le 0$, and $\chi(x) = 1$ for $x \ge 2$. Then χ' and $-\chi''$ are nonnegative on $(0,1)$. Define $\chi_R(x) \coloneqq R + \chi(x - R)$ for $R > 0$, and observe that $\chi_R(\varphi) - R - 1$ is compactly supported by construction. We have

$$\mathcal{A}\chi_R(\varphi) \;=\; \chi'_R(\varphi)\,\mathcal{A}\varphi + \chi''_R(\varphi)\,\langle\nabla\varphi, a\nabla\varphi\rangle - \bigl(\chi'_R(\varphi)\mathcal{I}\varphi - \mathcal{I}\chi_R(\varphi)\bigr)\,. \tag{3.2}$$

Note that the second and third terms on the right-hand side of (3.2) are nonpositive. Also, since $\varphi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$ and $\mathcal{A}_v\varphi \in L^d_{\mathrm{loc}}(\mathbb{R}^d)$, then $\mathcal{L}\varphi$ and $\mathcal{I}\varphi$ are both in $L^d_{\mathrm{loc}}(\mathbb{R}^d)$, and hence are locally integrable with respect to μ. Let $\mathcal{B}$ be the ball with the stated property and select any R such that $B_R \supset \mathcal{B}$. Therefore, integrating (3.2) with respect to μ and applying (2.1), and using also the nonpositivity of $\mathcal{A}\varphi$ on $\mathcal{B}^c$, we obtain

$$\int_{\mathcal{B}} \mathcal{A}\varphi \,\mathrm{d}\mu - \int_{\mathbb{R}^d\setminus\mathcal{B}} \chi'_R(\varphi)\,|\mathcal{A}\varphi|\,\mathrm{d}\mu \,\ge\, 0\,.$$

Hence, letting $R\to\infty$, we obtain

$$\int_{\mathbb{R}^d\setminus\mathcal{B}} |\mathcal{A}\varphi|\,\mathrm{d}\mu \,\le\, \int_{\mathcal{B}} \mathcal{A}\varphi\,\mathrm{d}\mu \,<\, \infty$$

by monotone convergence, from which the result follows. □

3.2. The α-discounted HJB equation. We have the following theorem.

THEOREM 3.2. *Assume (H1)–(H2). For any $\alpha\in(0,1)$, there exists a minimal nonnegative solution $V_\alpha \in \mathcal{W}^{2,p}_{loc}(\mathbb{R}^d)$, for any $p>1$, to the HJB equation*

$$\min_{z\in\mathcal{Z}} \bigl[\mathcal{A}_z\, V_\alpha(x) + \mathcal{R}(x,z)\bigr] \,=\, \alpha V_\alpha(x)\,. \tag{3.3}$$

Moreover, $\inf_{\mathbb{R}^d} \alpha V_\alpha \le \underline{\varrho}$, and this infimum is attained in the set

$$\mathit{\Gamma}_\circ \,:=\, \Bigl\{x\in\mathbb{R}^d : \sup_{z\in\mathcal{Z}} \mathcal{R}(x,z) \le \underline{\varrho}\Bigr\}\,.$$

Proof. Establishing the existence of a solution is quite standard. One starts by exhibiting a solution $\psi_{\alpha,R}\in\mathcal{W}^{2,p}(B_R)\cap\mathcal{C}(\mathbb{R}^d)$ to the Dirichlet problem

$$\begin{cases}\min_{z\in\mathcal{Z}} \bigl[\mathcal{A}_z\psi_{\alpha,R}(x) + \mathcal{R}(x,z)\bigr] \,=\, \alpha\psi_{\alpha,R}(x) & x\in B_R\,,\\ \psi_{\alpha,R}(x) \,=\, 0 & x\in B_R^c\,,\end{cases} \tag{3.4}$$

for any $\alpha\in(0,1)$ and $R>0$.

We use Definition 2.4 to write $\mathcal{A} = \widetilde{\mathcal{L}} + \widetilde{\mathcal{I}}$. Applying the well known interior estimate in [29, Theorem 9.11], for any fixed $r>0$, we obtain

$$\bigl\|\psi_{\alpha,R}\bigr\|_{\mathcal{W}^{2,p}(B_r)} \le C\Bigl(\bigl\|\psi_{\alpha,R}\bigr\|_{L^p(B_{2r})} + \bigl\|\mathcal{R}_{v_\alpha} + \widetilde{\mathcal{I}}\,\psi_{\alpha,R}\bigr\|_{L^p(B_{2r})}\Bigr)$$

for some constant $C=C(r,p)$. Here, v_α is a measurable selector from the minimizer of the α-discounted HJB in (3.3). Let $\tilde{\mathcal{V}} := \frac{\kappa_0}{\alpha} + \mathcal{V} - \psi_{\alpha,R}$. By (3.1) and (3.4), the function $\tilde{\mathcal{V}}$ satisfies $\mathcal{A}_{\hat{v}}\tilde{\mathcal{V}} - \alpha\tilde{\mathcal{V}} \le 0$ on B_R, and is positive on B_R^c. Thus it is nonnegative on B_R by the strong maximum principle. This of course implies that $\psi_{\alpha,R} \le \frac{\kappa_0}{\alpha} + \mathcal{V}$ on $\mathbb{R}^d$. Thus $\{\psi_{\alpha,R}\}$ is bounded in $\mathcal{W}^{2,p}(B_r)$, uniformly in R. We then take limits as $R\to\infty$ to obtain a function $V_\alpha\in\mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$ which solves (3.3).

Let $m_\alpha := \inf_{\mathbb{R}^d} V_\alpha$. We claim that $\alpha m_\alpha \le \underline{\varrho}$. Suppose on the contrary that $\alpha m_\alpha > \underline{\varrho}$. Let $v\in\mathfrak{U}_{\mathsf{ssm}}$. Recall the function χ in the proof of Lemma 3.1, and let $\tilde{\chi}(x) := -\chi(\frac{\underline{\varrho}}{2} + 2 - x)$. Note that $\tilde{\chi}'' \ge 0$, and $\tilde{\chi}'(\psi_{\alpha,R})\mathcal{I}\psi_{\alpha,R} - \mathcal{I}\tilde{\chi}(\psi_{\alpha,R}) \le 0$. Thus, using (3.4) and repeating the calculation in (3.2) we obtain

$$\mathcal{A}_v\tilde{\chi}(\psi_{\alpha,R}) \,\ge\, \tilde{\chi}'(\psi_{\alpha,R})\,\mathcal{A}_v\psi_{\alpha,R} \,\ge\, \tilde{\chi}'(\psi_{\alpha,R})\,\bigl(\alpha\psi_{\alpha,R} - \mathcal{R}_v\bigr)\,. \tag{3.5}$$

It is clear that $\tilde{\chi}(\psi_{\alpha,R}) \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)\cap\mathcal{C}_c(\mathbb{R}^d)$, for any $p>1$. Hence, integrating (3.5) with respect to μ_v, applying Lemma 3.1, and taking limits as $R\to\infty$, using monotone convergence, we obtain $\alpha m_\alpha \le \mu_v(\alpha V_\alpha) \le \mu_v(\mathcal{R}_v)$. Taking the infimum over $v\in\mathfrak{U}_{\mathsf{ssm}}$ contradicts the hypothesis that $\alpha m_\alpha > \underline{\varrho}$, and thus proves the claim.

Recall the definition $\epsilon_\circ$ in subsection 2.2. Let $\tilde{v} \in \mathfrak{U}_{\mathsf{sm}}$ be a measurable selector from the minimizer of (3.4) and consider the Dirichlet problem

$$\begin{cases} \mathcal{A}_{\tilde{v}}\tilde{\psi}_{\alpha,R}(x) + \mathcal{R}_{\tilde{v}}(x) = \alpha\tilde{\psi}_{\alpha,R}(x) & x \in B_R\,, \\ \tilde{\psi}_{\alpha,R}(x) = \alpha^{-1}(\underline{\varrho} + \epsilon_\circ) & x \in B_R^c\,, \end{cases} \tag{3.6}$$

for $\alpha \in (0,1)$ and $R > 0$. Arguing as in the derivation of (3.4), it follows that $\tilde{\psi}_{\alpha,R}$ converges, as $R \to \infty$, to some $\tilde{V}_\alpha \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$ which solves $\mathcal{A}_{\tilde{v}}\tilde{V}_\alpha + \mathcal{R}_{\tilde{v}}(x) = \alpha\tilde{V}_\alpha$ on $\mathbb{R}^d$. It is straightforward to show that $u = \tilde{V}_\alpha - V_\alpha$ is nonnegative and bounded. Indeed, since $u_R \coloneqq \tilde{\psi}_{\alpha,R} - \psi_{\alpha,R}$ satisfies $\mathcal{A}_{\tilde{v}}u_R - \alpha u_R = 0$ on B_R and $u_R = \alpha^{-1}(\underline{\varrho} + \epsilon_\circ)$ on B_R^c, an application of the strong maximum principle shows that u_R is nonnegative and bounded above by $\alpha^{-1}(\underline{\varrho} + \epsilon_\circ)$ on B_R. Continuing, since $\mathcal{A}_{\tilde{v}}u - \alpha u = 0$ on $\mathbb{R}^d$, it follows by (H1) that u cannot be strictly positive, and, in turn, by the strong maximum principle it has to be identically zero. Thus, given $\epsilon < \epsilon_\circ$ there exists R_ϵ such that $\min_{B_R} \alpha\tilde{\psi}_{\alpha,R} < \underline{\varrho} + \epsilon$ for all $R > R_\epsilon$. It follows by (3.6) that $\tilde{\psi}_{\alpha,R}$ attains its minimum in the set $\mathit{\Gamma}_\epsilon \coloneqq \{x \in \mathbb{R}^d : \mathcal{R}_{\tilde{v}}(x) \le \underline{\varrho} + \epsilon\}$ for all $R > R_\epsilon$, and therefore, the same applies to $\tilde{V}_\alpha$. Since $\epsilon > 0$ is arbitrary, we conclude that $\tilde{V}_\alpha$ attains its infimum in the set $\{x \in \mathbb{R}^d : \mathcal{R}_{\tilde{v}}(x) \le \underline{\varrho}\} \subset \mathit{\Gamma}_\circ$, and this completes the proof. ∎

3.3. The ergodic HJB equation. We start with the main convergence result of the paper which establishes solutions to the ergodic HJB via the vanishing discount method. To guide the reader, the technique of the proof consists of writing the operator in the form $\widetilde{\mathcal{L}} + \widetilde{\mathcal{I}}$, and obtaining estimates for supersolutions of the local operator $\widetilde{\mathcal{L}}$ using the results in [35, Corollary 2.2]. Recall the definition of the ball $\mathcal{B}_\circ$ in subsection 2.2.

THEOREM 3.3. *Grant the hypotheses of Theorem 3.2, and let V_α, $\alpha \in (0,1)$, be the family of solutions in that theorem. Then, as $\alpha \searrow 0$, $V_\alpha - V_\alpha(0)$ converges in $\mathcal{C}^{1,r}(\overline{B}_R)$ for any $r \in (0,1)$ and $R > 0$, to a function $V \in \mathcal{W}^{2,p}_{loc}(\mathbb{R}^d)$ for any $p > 1$, which is bounded from below in $\mathbb{R}^d$ and solves*

$$\min_{z \in \mathcal{Z}} \bigl[\mathcal{A}_z V(x) + \mathcal{R}(x,z)\bigr] = \varrho\,, \tag{3.7}$$

with $\varrho = \underline{\varrho}$. Also $\alpha V_\alpha(x) \to \underline{\varrho}$ uniformly on compact sets, and $V - \sup_{\mathcal{B}_\circ} V \le \mathcal{V}$ on $\mathbb{R}^d$. In addition, the solution of (3.7) with $\varrho = \underline{\varrho}$ is unique in the class of functions $V \in \mathcal{W}^{2,d}_{loc}(\mathbb{R}^d)$, satisfying $V(0) = 0$, which are bounded from below in $\mathbb{R}^d$. For $\varrho < \underline{\varrho}$, there is no such solution.

Proof. Recall the definitions of $\mathcal{B}_\circ$ and $\epsilon_\circ$ in subsection 2.2, and recall that $\mathcal{B}_\circ$ is also the ball used in (H2). Fix an arbitrary ball $\mathcal{B} \subset \mathbb{R}^d$ such that $\mathcal{B}_\circ \subset \mathcal{B}$. Since $\mathcal{V}$ and V_α are a supersolution and subsolution of $\mathcal{A}_{\hat{v}}u - \alpha u = -\mathcal{R}_{\hat{v}}$ on $\mathcal{B}^c$ by (3.1), respectively, it is straightforward to establish using the comparison principle that the solution V_α of (3.3) satisfies

$$V_\alpha(x) \le \sup_{\mathcal{B}} V_\alpha + \mathcal{V}(x) \qquad \forall\, x \in \mathbb{R}^d\,. \tag{3.8}$$

Indeed, choose any R such that $\mathcal{B} \subset B_R$, and with $\psi_{\alpha,R}$ denoting the solution of (3.4), define $\hat{\psi}_{\alpha,R}(x) \coloneqq \sup_{\mathcal{B}} \psi_{\alpha,R} + \mathcal{V}(x) - \psi_{\alpha,R}(x)$. Then $\hat{\psi}_{\alpha,R}$ satisfies $\mathcal{A}_{\hat{v}}\hat{\psi}_{\alpha,R} \le 0$ on $B_R \setminus \bar{\mathcal{B}}$, and $\hat{\psi}_{\alpha,R} \ge 0$ on $B_R^c \cup \bar{\mathcal{B}}$. By the strong maximum principle, we obtain $\psi_{\alpha,R}(x) \le \sup_{\mathcal{B}} \psi_{\alpha,R} + \mathcal{V}(x)$ for $x \in B_R$. Since this inequality holds for all such

$R > 0$, and $\psi_{\alpha,R}$ converges to V_α as $R \to \infty$, uniformly on compacta, as shown in the proof of Theorem 3.2, the inequality in (3.8) follows.

By Theorem 3.2 we have $\inf_{\mathbb{R}^d} V_\alpha = \min_{\mathcal{B}_\circ} V_\alpha$ for all $\alpha \in (0,1)$. For each $\alpha \in (0,1)$, we fix some point $\hat{x}_\alpha \in \operatorname{Arg\,min} V_\alpha \subset \mathcal{B}_\circ$. Consider the function $\varphi_\alpha := V_\alpha - V_\alpha(\hat{x}_\alpha)$. Then (3.8) implies that

$$\varphi_\alpha(x) \;\le\; \|\varphi_\alpha\|_{L^\infty(\mathcal{B})} + \mathcal{V}(x) \qquad \forall\, x \in \mathbb{R}^d\,. \tag{3.9}$$

We have

$$\min_{z\in\mathcal{Z}} \bigl[\mathcal{A}_z\varphi_\alpha(x) - \alpha\varphi_\alpha(x) + \mathcal{R}(x,z)\bigr] \;=\; \alpha V_\alpha(\hat{x}_\alpha) \;\le\; \underline{\varrho}\,,$$

where the last inequality follows by Theorem 3.2. We claim that for each $R > 0$ there exists a constant κ_R such that

$$\|\varphi_\alpha\|_{L^\infty(B_R)} \;\le\; \kappa_R \qquad \forall\, \alpha \in (0,1)\,. \tag{3.10}$$

To prove the claim, let $\mathcal{B} \equiv B_R$, and D_1, D_2 be balls satisfying $\mathcal{B} \Subset D_1 \Subset D_2$. Recall Definition 2.4. For $p > 0$, let $\|u\|_{p;Q} := \bigl(\int_Q |u(x)|\,\mathrm{d}x\bigr)^{1/p}$. Of course, this is not a norm unless $p \ge 1$, so there is a slight abuse of notation involved in this definition. Since $\mathcal{V} \in \mathcal{C}^2(\mathbb{R}^d)$, hypothesis (H2) implies that $\widetilde{\mathcal{I}}\mathcal{V} \in L^\infty_{\mathrm{loc}}(\mathbb{R}^d)$, and the same of course holds for $\widetilde{\mathcal{I}}\varphi_\alpha$ by (3.9). By the local maximum principle [29, Theorem 9.20], for any $p > 0$, there exists a constant $\tilde{C}_1(p) > 0$ such that

$$\|\varphi_\alpha\|_{L^\infty(\mathcal{B})} \;\le\; \tilde{C}_1(p)\bigl(\|\varphi_\alpha\|_{p;D_1} + \|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^d(D_1)} + \|\mathcal{R}_{v_\alpha}\|_{L^d(D_1)}\bigr)\,,$$

and by the supersolution estimate [29, Theorem 9.22], and since φ_α is nonnegative, there exist some $p > 0$ and $\tilde{C}_2 > 0$ such that $\|\varphi_\alpha\|_{p;D_1} \;\le\; \tilde{C}_2\,\underline{\varrho}\,|D_2|^{1/d}$. Combining these inequalities, we obtain

$$\|\varphi_\alpha\|_{L^\infty(\mathcal{B})} \;\le\; \tilde{C}_1(p)\bigl(\tilde{C}_2\,\underline{\varrho}\,|D_2|^{1/d} + \|\mathcal{R}_{v_\alpha}\|_{L^d(D_1)}\bigr) + \tilde{C}_1(p)\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^d(D_1)}\,. \tag{3.11}$$

Denote the first term on the right-hand side of (3.11) by κ_1. By (3.9) and (3.11) we have

$$\begin{aligned}
\|\varphi_\alpha\|_{L^\infty(D_2)} \;&\le\; \|\mathcal{V}\|_{L^\infty(D_2)} + \|\varphi_\alpha\|_{L^\infty(\mathcal{B})}\\
&\le\; \kappa_1 + \|\mathcal{V}\|_{L^\infty(D_2)} + \tilde{C}_1(p)\,\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^d(D_1)}\,.
\end{aligned}$$

This implies that, either $\|\varphi_\alpha\|_{L^\infty(D_2)} \le 2\bigl(\kappa_1 + \|\mathcal{V}\|_{L^\infty(D_2)}\bigr)$, in which case (3.10) holds with this bound, or

$$\|\varphi_\alpha\|_{L^\infty(D_2)} \;\le\; 2\tilde{C}_1(p)\,\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^d(D_1)}\,. \tag{3.12}$$

If (3.12) holds, then we write $\widetilde{\mathcal{I}}\,\varphi_\alpha = \widetilde{\mathcal{I}}(\mathbb{1}_{D_2}\varphi_\alpha) + \widetilde{\mathcal{I}}(\mathbb{1}_{D_2^c}\varphi_\alpha)$, and use the estimate

$$\widetilde{\mathcal{I}}(\mathbb{1}_{D_2^c}\varphi_\alpha)(x) \;\le\; \|\varphi_\alpha\|_{L^\infty(\mathcal{B})}\left(\sup_{x\in D_1}\,\nu(x, D_2^c - x)\right) + \widetilde{\mathcal{I}}(\mathbb{1}_{D_2^c}\mathcal{V})(x) \qquad \forall\, x \in D_1\,,$$

which holds by (3.9), together with (3.11) and (3.12), to obtain

$$\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^\infty(D_1)} \;\le\; 2\tilde{C}_1(p)\,\|\bar{\nu}\|_{L^\infty(D_1)}\,\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^d(D_1)} \tag{3.13}$$

$$+ \kappa_1 \|\bar{\nu}\|_{L^\infty(D_1)} + \|\widetilde{\mathcal{I}}(\mathbb{1}_{D_2^c}\mathcal{V})\|_{L^\infty(D_1)}\,.$$

We distinguish two cases from (3.13):
Case 1. Suppose that

$$\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^\infty(D_1)} \;\le\; 4\tilde{C}_1(p)\,\|\bar{\nu}\|_{L^\infty(D_1)}\,\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^d(D_1)}\,. \tag{3.14}$$

Let ψ_α be the solution of the Dirichlet problem

$$\widetilde{\mathcal{L}}_{v_\alpha}\psi_\alpha - \alpha\psi_\alpha \;=\; -\widetilde{\mathcal{I}}\,\varphi_\alpha \quad \text{in } D_1\,, \qquad \text{and} \quad \psi_\alpha = \varphi_\alpha \quad \text{on } \partial D_1\,,$$

with v_α a measurable selector from the minimizer in (3.3). Then ψ_α is nonnegative in D_1 by the strong maximum principle, and thus (3.14) together with [35, Corollary 2.2], implies that for some constant C_{H} we have

$$\psi_\alpha(x) \;\le\; C_{\mathsf{H}}\,\psi_\alpha(\hat{x}_\alpha) \qquad \forall\, x \in \mathcal{B}\,, \quad \forall\, \alpha \in (0,1)\,. \tag{3.15}$$

On the other hand, $\varphi_\alpha - \psi_\alpha$ satisfies

$$\widetilde{\mathcal{L}}_{v_\alpha}(\varphi_\alpha - \psi_\alpha) - \alpha(\varphi_\alpha - \psi_\alpha) \;=\; \alpha V_\alpha(\hat{x}_\alpha) - \mathcal{R}_{v_\alpha} \quad \text{in } D_1\,, \tag{3.16}$$

and $\varphi_\alpha - \psi_\alpha = 0$ on ∂D_1. Thus, by the ABP weak maximum principle [29, Theorem 9.1], and since $\alpha V_\alpha(\hat{x}_\alpha) \le \underline{\varrho}$, we obtain from (3.16) that

$$\|\varphi_\alpha - \psi_\alpha\|_{L^\infty(D_1)} \;\le\; C_\circ \qquad \forall\, \alpha \in (0,1)\,, \tag{3.17}$$

for some constant $C_\circ$. Equation (3.17) implies that $\psi_\alpha(\hat{x}_\alpha) \le C_\circ$. Combining (3.15) and (3.17) in the standard manner, we obtain

$$\begin{aligned} \varphi_\alpha(x) \;&\le\; \|\varphi_\alpha - \psi_\alpha\|_{L^\infty(D_1)} + \psi_\alpha(x) \\ &\le\; C_\circ + C_{\mathsf{H}}\,\psi_\alpha(\hat{x}_\alpha) \;\le\; C_\circ(1 + C_{\mathsf{H}}) \qquad \forall\, x \in \mathcal{B}\,, \quad \forall\, \alpha \in (0,1)\,. \end{aligned} \tag{3.18}$$

Case 2. Suppose that

$$\|\widetilde{\mathcal{I}}\,\varphi_\alpha\|_{L^\infty(D_1)} \;\le\; 2\kappa_1\|\bar{\nu}\|_{L^\infty(D_1)} + 2\,\|\widetilde{\mathcal{I}}(\mathbb{1}_{D_2^c}\mathcal{V})\|_{L^\infty(D_1)}\,.$$

In this case, we consider the solution $\tilde{\psi}_\alpha$ of the Dirichlet problem

$$\widetilde{\mathcal{L}}_{v_\alpha}\tilde{\psi}_\alpha - \alpha\tilde{\psi}_\alpha \;=\; 0 \quad \text{in } D_1\,, \qquad \text{and} \quad \tilde{\psi}_\alpha = \varphi_\alpha \quad \text{on } \partial D_1\,.$$

We have $\tilde{\psi}_\alpha(x) \le \tilde{C}_{\mathsf{H}}\,\tilde{\psi}_\alpha(\hat{x}_\alpha)$ for all $x \in \mathcal{B}$ and $\alpha \in (0,1)$, for some constant $\tilde{C}_{\mathsf{H}}$. Also,

$$\widetilde{\mathcal{L}}_{v_\alpha}(\varphi_\alpha - \tilde{\psi}_\alpha) - \alpha(\varphi_\alpha - \psi_\alpha) \;=\; -\widetilde{\mathcal{I}}\,\varphi_\alpha + \alpha V_\alpha(\hat{x}_\alpha) - \mathcal{R}_{v_\alpha} \quad \text{in } D_1\,, \tag{3.19}$$

and $\varphi_\alpha - \tilde{\psi}_\alpha = 0$ on ∂D_1. By the ABP weak maximum principle, we obtain from (3.19) that $\|\varphi_\alpha - \tilde{\psi}_\alpha\|_{L^\infty(D_1)} \le \tilde{C}_\circ$ for all $\alpha \in (0,1)$ and for some constant $\tilde{C}_\circ$. Thus again we obtain (3.18) with constants $\tilde{C}_\circ$ and $\tilde{C}_{\mathsf{H}}$. This establishes (3.10).

It follows by (3.10) that $\overline{V}_\alpha \coloneqq V_\alpha - V_\alpha(0) = \varphi_\alpha(x) - \varphi_\alpha(0)$ is locally bounded, uniformly in $\alpha \in (0,1)$. The same applies to $\widetilde{\mathcal{I}}\,\overline{V}_\alpha$ by (3.9) and (H2). Note that

$$\widetilde{\mathcal{L}}_{v_\alpha}\overline{V}_\alpha - \alpha\overline{V}_\alpha \;=\; \alpha V_\alpha(0) - \mathcal{R}_{v_\alpha} - \widetilde{\mathcal{I}}\,\overline{V}_\alpha \quad \text{on } \mathbb{R}^d\,.$$

Thus, by the interior estimate in [29, Theorem 9.11], there exists a constant $C = C(R,p)$ such that

$$\|\overline{V}_\alpha\|_{\mathscr{W}^{2,p}(B_R)} \le C\Big(\|\overline{V}_\alpha\|_{L^p(B_{2R})} + \|\alpha V_\alpha(0) - \mathcal{R}_{v_\alpha} - \widetilde{\mathcal{I}}\,\overline{V}_\alpha\|_{L^p(B_{2R})}\Big)\,.$$

Hence $\{\overline{V}_\alpha\}$ is bounded in $\mathscr{W}^{2,p}(B_R)$ for any $R>0$. A standard argument then shows that given any sequence $\alpha_n \searrow 0$, $\{\overline{V}_{\alpha_n}\}$ contains a subsequence which converges in $\mathcal{C}^{1,r}(\overline{B}_R)$ for any $r < 1 - \frac{d}{p}$ (see, for example, Lemma 3.5.4 in [24]). Taking limits in

$$\min_{z\in\mathcal{Z}} \big[\mathcal{A}_z\overline{V}_\alpha(x) - \alpha\overline{V}_\alpha(x) + \mathcal{R}(x,z)\big] \;=\; \alpha V_\alpha(0) \tag{3.20}$$

along this subsequence we obtain (3.7), as claimed in the statement of the theorem, for some $\varrho\in\mathbb{R}$. Since $\limsup_{\alpha\searrow 0}\,\alpha V_\alpha(\hat{x}_\alpha) \le \underline{\varrho}$, we have $\varrho \le \underline{\varrho}$. On the other hand, from the theory of infinite dimensional LP [36] it is well known that the value of the dual problem cannot be smaller than the value of the primal, hence $\varrho \ge \underline{\varrho}$, and we have equality (see also Section 4 in [11]). That $V - \sup_{\mathcal{B}_\circ} V \le \mathcal{V}$ on $\mathbb{R}^d$ follows by (3.9) with $\mathcal{B} = \mathcal{B}_\circ$.

Suppose now that $\widetilde{V} \in \mathscr{W}^{2,d}_{\mathrm{loc}}(\mathbb{R}^d)$ is bounded from below in $\mathbb{R}^d$, and satisfies

$$\min_{z\in\mathcal{Z}} \big[\mathcal{A}_z\,\widetilde{V}(x) + \mathcal{R}(x,z)\big] \;=\; \tilde{\varrho} \tag{3.21}$$

with $\tilde{\varrho} \le \underline{\varrho}$. Let $\tilde{v} \in \mathfrak{V}_{\mathsf{sm}}$ be an a.e. measurable selector from the minimizer of (3.21). Using the equation $\mathcal{A}_{\tilde{v}}\widetilde{V} = \tilde{\varrho} - \mathcal{R}_{\tilde{v}}$, in lieu of (3.1) in (H2), then, as we have already shown, the solution V derived as the limit of $V_\alpha - V_\alpha(0)$ satisfies $V - \sup_{\mathcal{B}} V \le \widetilde{V}$ on $\mathbb{R}^d$ for some ball $\mathcal{B}$. It is then clear that if we translate $\widetilde{V}$ by an additive constant until it first touches V at some point from above, it has to touch it at some point in $\overline{\mathcal{B}}$. Thus the function $\phi \coloneqq \widetilde{V} - V - \inf_{\mathcal{B}}(\widetilde{V} - V)$ is nonnegative on $\mathbb{R}^d$, satisfies $\mathcal{A}_{\tilde{v}}\phi \le \tilde{\varrho} - \underline{\varrho}$, and $\phi(\tilde{x}) = 0$ for some $\tilde{x}\in\overline{\mathcal{B}}$. By the strong maximum principle we must have $\tilde{\varrho} = \underline{\varrho}$ and $\phi \equiv 0$ on $\mathbb{R}^d$. It is evident from the uniqueness of the solution, that the limit of (3.20) is independent of the subsequence $\alpha_n \searrow 0$ chosen. It is also clear that $\alpha V_\alpha(x) \to \underline{\varrho}$ as $\alpha \searrow 0$, uniformly on compact sets. This completes the proof. □

Remark 3.4. If ν is translation invariant and has compact support, and $\mathcal{R}$ and b are locally Hölder continuous in x, then $\widetilde{\mathcal{I}}V$ is locally Hölder continuous, and thus the solution V in Theorem 3.3 is in $\mathcal{C}^{2,r}(\mathbb{R}^d)$ for some $r\in(0,1)$ by elliptic regularity [29, Theorem 9.19].

3.3.1. Verification of optimality.

We start with the necessity part.

THEOREM 3.5. *Assume the hypotheses of Theorem 3.3. If $v \in \mathfrak{V}_{\mathsf{ssm}}$ is optimal, then it satisfies*

$$b^i_v(x)\,\partial_i V(x) + \mathcal{R}_v(x) \;=\; \inf_{z\in\mathcal{Z}} \big[b^i(x,z)\partial_i V(x) + \mathcal{R}(x,z)\big] \qquad a.e.\ x\in\mathbb{R}^d\,. \tag{3.22}$$

Proof. Suppose not. Then there exists some ball $\mathcal{B}$ such that

$$h(x) \;\coloneqq\; \Big(b^i_v(x)\,\partial_i V(x) + \mathcal{R}_v(x) - \inf_{z\in\mathcal{Z}} \big[b^i(x,z)\partial_i V(x) + \mathcal{R}(x,z)\big]\Big)\,\mathbb{1}_{\mathcal{B}}(x) \tag{3.23}$$

is a nontrivial nonnegative function. Since $\partial_i V_\alpha$ converges uniformly to $\partial_i V$ as $\alpha\searrow 0$ on compact sets by Theorem 3.3, it follows that if we define h_α as the right-hand side of (3.23), but with V replaced by V_α, then $h - h_\alpha$ converges to 0 a.e. in $\mathcal{B}$, and also $\mu_v(|h-h_\alpha|) \to 0$ as $\alpha \searrow 0$, since μ_v has a density in $L^p_{\mathrm{loc}}(\mathbb{R}^d)$ for some $p>1$. We have $\mathcal{A}_v V_\alpha \ge \alpha V_\alpha + h_\alpha - \mathcal{R}_v$ a.e. on $\mathbb{R}^d$ by the definition of h_α. Repeating the same argument using the solution in (3.4) $\psi_{\alpha,R}$ instead of V in (3.23), we deduce that there exists $h_{\alpha,R}$ supported on $\mathcal{B}$ such that $\mu_v(|h_{\alpha,R}-h_\alpha|) \to 0$ as $R\to\infty$, and $\mathcal{A}_v\psi_{\alpha,R} \ge \psi_{\alpha,R} + h_{\alpha,R} - \mathcal{R}_v$. Thus, as in the derivation of (3.5) using the function $\breve\chi(x) \coloneqq -\chi(\frac{\varrho}{2}+2-x)$, with χ as defined in the proof of Lemma 3.1, we obtain

$$\mathcal{A}_v\tilde\chi(\psi_{\alpha,R}) \;\ge\; \tilde\chi'(\psi_{\alpha,R})\,\mathcal{A}_v\psi_{\alpha,R} \;\ge\; \tilde\chi'(\psi_{\alpha,R})\bigl(\alpha\psi_{\alpha,R} + h_{\alpha,R} - \mathcal{R}_v\bigr)\,. \tag{3.24}$$

Hence, integrating (3.24) with respect to μ_v, applying Lemma 3.1, and taking limits as $R\to\infty$, using the property that that $\mu_v(|h_{\alpha,R}-h_\alpha|) \to 0$ as $R\to\infty$, we obtain $\mu_v(\mathcal{R}_v) \ge \mu_v(\alpha V_\alpha) + \mu_v(h_\alpha)$. By the proof of Theorem 3.3 $\inf_{\mathbb{R}^d}\,\alpha V_\alpha \to \underline{\varrho}$ as $\alpha\searrow 0$. Thus, taking limits as $\alpha\searrow 0$, we obtain $\mu_v(h)\le 0$, and since μ_v has everywhere positive density, this implies $h=0$ a.e. □

Concerning the sufficiency part of the verification of optimality, or in other words, that any $v\in\mathfrak{V}_{\mathsf{sm}}$ which satisfies (3.22) is necessarily optimal, the probabilistic argument has a clear advantage here. With v_* an a.e. measurable selector from the minimizer of (3.22), the HJB takes the form of the Foster–Lyapunov equation $\mathcal{A}_{v_*}V = \underline{\varrho} - \mathcal{R}_{v_*}$, which shows that the controlled process is ergodic (provided that the martingale problem has a solution under v_*). It then follows by a straightforward application of Itô's formula and Birkhoff's ergodic theorem that v_* is optimal.

3.4. On waiving hypothesis (H2). In this section we do not assume (H2). Recall Definition 2.7. We impose additional assumptions on ν to establish existence of solutions to the Poisson equation.

THEOREM 3.6. *We assume* (H1) *and one of the following:*

(*a*) *$\nu=\nu$ is translation invariant and has compact support.*

(*b*) *ν has locally compact support and satisfies the hypotheses of* Corollary 2.8.

Let $\hat v\in\mathfrak{V}_{\mathsf{ssm}}$ be such that $\mathcal{R}_{\hat v}$ is coercive relative to $\varrho_{\hat v}$. Then, up to an additive constant, there exists a unique $\widehat V\in\mathcal{W}^{2;d}_{loc}(\mathbb{R}^d)$ which is bounded from below in $\mathbb{R}^d$, and satisfies

$$\mathcal{A}_{\hat v}\,\widehat V(x) + \mathcal{R}_{\hat v}(x) \;=\; \beta \qquad \forall\, x\in\mathbb{R}^d\,, \tag{3.25}$$

for some $\beta=\varrho_{\hat v}$. For $\beta<\varrho_{\hat v}$, there is no such solution.

Proof. For $n\in\mathbb{N}$, let $\mathcal{R}^n = n\wedge\mathcal{R}$ denote the n-truncation of the running cost. It is clear that $\mathcal{R}^n$ is coercive relative to $\varrho_{\hat v}$ for all $n>\varrho_{\hat v}$. Let $\hat\psi^n_{\alpha,R}\in\mathcal{W}^{2,p}(B_R)\cap\mathcal{W}^{1,p}_0(B_R)$ be the unique solution of the Dirichlet problem

$$\begin{cases}\mathcal{A}_{\hat v}\hat\psi^n_{\alpha,R}(x) + \mathcal{R}^n_{\hat v}(x) \;=\; \alpha\hat\psi^n_{\alpha,R}(x) & x\in B_R\,,\\ \hat\psi^n_{\alpha,R}(x) \;=\; 0 & x\in B_R^c\,.\end{cases}$$

It is clear that $\|\hat\psi^n_{\alpha,R}\|_{L^\infty(\mathbb{R}^d)} \le \frac{n}{\alpha}$, and this is inherited by the function $\widehat V^n_\alpha$ at the limit $R\to\infty$. Thus, by the proof of Theorem 3.2, $\widehat V^n_\alpha$ is in $\mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$ for any $p\ge1$, and satisfies $\mathcal{A}_{\hat v}\,\widehat V^n_\alpha + \mathcal{R}^n_{\hat v} = \alpha\widehat V^n_\alpha$. Repeating the argument in the proof of Theorem 3.3, the infimum of $\widehat V^n_\alpha$ over $\mathbb{R}^d$ is attained in a ball $\mathcal{B}_\circ$ as defined in subsection 2.2 (relative to $\varrho_{\hat v}$), and if $\hat x^n_\alpha\in\mathcal{B}_\circ$ denotes a point where the infimum is attained, then $\alpha\widehat V^n_\alpha(\hat x^n_\alpha)\le\varrho_{\hat v}$. With $\varphi^n_\alpha \coloneqq \widehat V^n_\alpha - \widehat V^n_\alpha(\hat x^n_\alpha)$, we write the equation as

$$\widetilde{\mathcal{L}}_{\hat v}\varphi^n_\alpha(x) - \alpha\varphi^n_\alpha(x) \;=\; \alpha\widehat V^n_\alpha(\hat x^n_\alpha) - \mathcal{R}^n_{\hat v}(x) - \widetilde{\mathcal{I}}\varphi^n_\alpha(x) \tag{3.26}$$

$$\leq \varrho_{\hat{v}} - \mathcal{R}^n(x) - \widetilde{\mathcal{I}}\varphi_\alpha^n(x) \qquad \text{a.e. } x \in \mathbb{R}^d\,.$$

We express (3.26) in divergence form as

$$\partial_j\bigl(a^{ij}\partial_i\varphi_\alpha^n\bigr) + \bigl(\hat{b}^i - \partial_i a^{ij}\bigr)\partial_i\varphi_\alpha^n - \bar{\nu}\varphi_\alpha^n \;\le\; \varrho_{\hat{v}} - \mathcal{R}_{\hat{v}} - \widetilde{\mathcal{I}}\varphi_\alpha^n\,,$$

and apply [29, Theorem 8.18] to obtain $\bigl\|\varphi_\alpha^n\bigr\|_{L^p(B_{2R}(x_0))} \le \varrho_{\hat{v}}\,\kappa_{p,R}$ for some constant $\kappa_{p,R}$, for any $p \in \bigl(1, \frac{d}{d-2}\bigr)$. Therefore, $\inf_{B_{2R}(x_0)\setminus B_R(x_0)} \varphi_\alpha^n$ is bounded over $\alpha \in (0,1)$ and $n \ge \varrho_{\hat{v}}$. Thus, we can select some $x_0' \in B_{2R}(x_0)\setminus B_R(x_0)$ satisfying $\sup_n \varphi_\alpha^n(x_0') < \infty$, and repeat the procedure to show by induction that φ_α^n is locally bounded in L^p for any $p \in \bigl(1, \frac{d}{d-2}\bigr)$, uniformly over $\alpha \in (0,1)$ and $n \ge \varrho_{\hat{v}}$.

Next, we apply successively the Calderón–Zygmund estimate [29, Theorem 9.11] to the non-divergence form of the equation in (3.26) which states that

$$\bigl\|\varphi_\alpha^n\bigr\|_{\mathcal{W}^{2,p}(B_R)} \;\le\; C\Bigl(\bigl\|\varphi_\alpha^n\bigr\|_{L^p(B_{2R})} + \bigl\|\alpha V_\alpha(\hat{x}_\alpha^n) - \mathcal{R}_{\hat{v}}^n - \widetilde{\mathcal{I}}\,\varphi_\alpha^n\bigr\|_{L^p(B_{2R})}\Bigr)\,.$$

We start with the L^p estimate, say with $p = \frac{d}{d-r}$ for $r \in (1,2)$. If (a) holds, then $\|\widetilde{I}\varphi_\alpha^n\|_{L^p(B_R(x))} \;\le\; \bar{\nu}\,\|\varphi_\alpha^n\|_{L^p(B_{R+R_\circ}(x))}$ by the Minkowski integral inequality, where $R_\circ$ is such that the support of ν is contained in $B_{R_\circ}$, while in case (b) we use the technique in the proof of Corollary 2.8. Using the compactness of the embedding $\mathcal{W}^{2,p}(B_R) \hookrightarrow L^q(B_R)$ for $p \le q < \frac{pd}{d-2p}$, we choose $q = \frac{pd}{d-rp}$ to improve the estimate to a new $p = \frac{d}{d-2r}$. Continuing in this manner, in at most $d-1$ steps we obtain

$$\sup_{n\ge\varrho_{\hat{v}}}\;\sup_{\alpha\in(0,1)}\;\|\varphi_\alpha^n\|_{\mathcal{W}^{2,p}(B_R)} \;<\; \infty$$

for any $p > d$ and $R > 0$. Letting first $n \to \infty$, and then $\alpha \searrow 0$, along an appropriate subsequence, we obtain a solution to (3.25) as claimed. The rest follow as in the proof of Theorem 3.3. ◻

Corollary 3.7. *Grant the hypotheses of Theorem 3.6. Then the conclusions of Theorems 3.3 and 3.5 hold.*

Proof. Note that the only place we use the assumption $\mathcal{V} \in \mathcal{C}^2(\mathbb{R}^d)$ in the proof of Theorem 3.3 is to assert that $\widetilde{\mathcal{I}}\mathcal{V} \in L^\infty_{\text{loc}}(\mathbb{R}^d)$. Thus, under (a), or (b) of Theorem 3.6, if we select $\hat{v} \in \mathfrak{V}_{\mathsf{ssm}}$ such that $\varrho_{\hat{v}} \le \underline{\varrho} + \epsilon_\circ$, then the Poisson equation in (3.25) can be used in lieu (H2), and the conclusions of Theorems 3.3 and 3.5 follow. ◻

In the next theorem, under the hypothesis that V is inf-compact, we show that any $v \in \mathfrak{V}_{\mathsf{sm}}$ satisfying (3.22) is stable by constructing a density for the associated infinitesimal invariant measure.

Theorem 3.8. *Grant the hypotheses of Theorem 3.6, and suppose that V is inf-compact. Then any $v \in \mathfrak{V}_{\mathsf{sm}}$ satisfying (3.22) is stable and optimal.*

Proof. We adapt the technique which is used in [33, Theorem 1.2] for a local operator, to construct an infinitesimal invariant measure μ_v. Let $\widetilde{\mathcal{L}}_v^*$ be the operator in (2.7), and set $\widehat{\mathcal{I}}u(x) := \int_{\mathbb{R}^d} u(x-y)\nu(\mathrm{d}y)$ if ν is translation invariant; otherwise, under hypothesis (b) of Theorem 3.6, we define $\widehat{\mathcal{I}}u(x) := \int_{\mathbb{R}^d} \psi_{x-y}(y)\,u(x-y)\,\mathrm{d}y$. Consider the solution ϕ_k of the Dirichlet problem $\widetilde{\mathcal{L}}_v^*\phi_k + \widehat{\mathcal{I}}\phi_k = 0$ on B_k, with ϕ_k equal to a positive constant c_k on B_k^c.

Concerning the solvability of the Dirichlet problem, note that for $f \in L^2(B_k)$, the problem $\widetilde{\mathcal{L}}_v^* u = -\widehat{\mathcal{I}}f$ on B_k, with $u = c_k$ on B_k^c, has a unique solution $u \in \mathcal{W}^{2,2}(B_k)$, which obeys the estimate $\|u\|_{\mathcal{W}^{2,2}(B_k)} \le \kappa(1 + \|u\|_{L^2(B_k)} + \|\widehat{\mathcal{I}}f\|_{L^2(B_k)})$ for some constant κ. Thus we can combine Corollary 2.8, the compactness of the embedding $\mathcal{W}^{2,2}(B_R) \hookrightarrow L^q(B_R)$ for $q = \frac{2d}{d-1}$, and the Leray–Schauder fixed point theorem to assert the existence of a solution $\phi_k \in \mathcal{W}^{2,2}(B_k)$ as claimed in the preceding paragraph. The solutions ϕ_k are nonnegative by the weak maximum principle [29, Theorem 8.1]. We choose the constant c_k so that $\int_{B_k} \phi_k(x)\,\mathrm{d}x = 1$.

We improve the regularity of ϕ_k by following the proofs of Theorem 2.5 and Corollary 2.8, and show that for any $n > 0$, there exists $N(n) \in \mathbb{N}$ such that the sequence $\{\phi_k : k > N(n)\}$ is Hölder equicontinuous on the ball B_n. Since $\int_{B_k} \phi_k(x)\,\mathrm{d}x = 1$, it follows that the sequence is bounded on each ball B_n uniformly over $k > N(n)$, and thus by the Arzelà–Ascoli theorem combined with Fatou's lemma, converges along a subsequence to some nonnegative, locally Hölder continuous $\phi \in L^1(\mathbb{R}^d)$ uniformly on compact sets. It is clear that ϕ is a generalized solution of (2.6). Let $R = R(n) > 0$ be such that $V(x) > R + 1$ on B_n^c. It is always possible to select such $R(n)$ in a manner that $R(n) \to \infty$ as $n \to \infty$ by the assumption that V is inf-compact. Employing the function $\chi_R(V)$ as in the proof of Lemma 3.1 and using (3.7), it follows that $\int_{B_{R(n)}} \mathcal{R}_v(x)\,\phi_k(x)\,\mathrm{d}x \le \underline{\varrho}$ for all $k > N(n)$ and $n \in \mathbb{N}$. This implies that $\int_{\mathcal{B}_\circ} \phi_k(x)\,\mathrm{d}x \ge \frac{2\epsilon_\circ}{\underline{\varrho}+2\epsilon_\circ}$ for all large enough k, and the same must hold for the limit ϕ by uniform convergence. This implies that ϕ is a nontrivial nonnegative function, and being a generalized solution of (2.6), it satisfies $\int_{\mathbb{R}^d} \mathcal{A}_v f(x)\phi(x)\,\mathrm{d}x = 0$ for all $f \in \mathcal{C}$. Thus, after normalization, ϕ is the density of an infinitesimal invariant measure. Therefore, $v \in \mathfrak{V}_{\mathsf{ssm}}$.

Optimality of v is easily established by the argument in the proof of Lemma 3.1, using the function χ_R. □

4. A jump diffusion model. In this section, we consider a jump diffusion process $X = \{X_t \colon t \ge 0\}$ in $\mathbb{R}^d$, $d \ge 2$, defined by the Itô equation

$$\mathrm{d}X_t \;=\; b(X_t, Z_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t + \mathrm{d}L_t\,, \quad X_0 = x \in \mathbb{R}^d\,. \tag{4.1}$$

Here, $W = \{W_t,\, t \ge 0\}$ is a d-dimensional standard Wiener process, and $L = \{L_t,\, t \ge 0\}$ is a Lévy process such that $\mathrm{d}L_t = \int_{\mathbb{R}^m_*} g(X_{t-}, \xi)\,\widetilde{\mathcal{N}}(\mathrm{d}t, \mathrm{d}\xi)$, where $\widetilde{\mathcal{N}}$ is a martingale measure in $\mathbb{R}^m_* = \mathbb{R}^m \setminus \{0\}$, $m \ge 1$, corresponding to a standard Poisson random measure $\mathcal{N}$. In other words, $\widetilde{\mathcal{N}}(t, A) = \mathcal{N}(t, A) - t\mathit{\Pi}(A)$ with $\mathbb{E}[\mathcal{N}(t, A)] = t\mathit{\Pi}(A)$ for any $A \in \mathfrak{B}(\mathbb{R}^m)$, with $\mathit{\Pi}$ a σ-finite measure on $\mathbb{R}^m_*$, and g a measurable function.

The processes W and $\mathcal{N}$ are defined on a complete probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. Assume that the initial condition X_0, W_0, and $\mathcal{N}(0, \cdot)$ are mutually independent. The control process $Z = \{Z_t,\, t \ge 0\}$ takes values in a compact, metrizable space $\mathcal{Z}$, is $\mathfrak{F}_t$-adapted, and *non-anticipative*: for $s < t$, $\bigl(W_t - W_s,\, \mathcal{N}(t, \cdot) - \mathcal{N}(s, \cdot)\bigr)$ is independent of

$$\mathfrak{F}_s \;\coloneqq\; \text{the completion of } \sigma\{X_0, Z_r, W_r, \mathcal{N}(r, \cdot) : \; r \le s\} \text{ relative to } (\mathfrak{F}, \mathbb{P})\,.$$

Such a process Z is called an *admissible control* and we denote the set of admissible controls by $\mathfrak{Z}$.

4.1. The ergodic control problem for the jump diffusion. Let $\mathcal{R}\colon \mathbb{R}^d \times \mathcal{Z} \mapsto \mathbb{R}_+$ denote the running cost function, which is assumed to satisfy (2.3).

For an admissible control process $Z \in \mathfrak{Z}$, we consider the *ergodic cost* defined by

$$\tilde{\varrho}_Z(x) \;\coloneqq\; \limsup_{T\to\infty}\, \frac{1}{T}\, \mathbb{E}^Z_x\biggl[\int_0^T \mathcal{R}(X_t, Z_t)\,\mathrm{d}t\biggr]\,.$$

Here $\mathbb{E}_x^Z$ denotes the expectation operator corresponding to the process controlled under Z, with initial condition $X_0 = x \in \mathbb{R}^d$. The ergodic control problem seeks to minimize the ergodic cost over all admissible controls. We define $\tilde{\varrho}(x) \coloneqq \inf_{Z\in\mathfrak{Z}} \tilde{\varrho}_Z(x)$. As we show in Theorem 4.4, this infimum is realized with a stationary Markov control, and $\tilde{\varrho}(x) = \underline{\varrho}$, with $\underline{\varrho}$ as defined in subsection 2.1, so it does not depend on x.

4.2. Assumptions on the parameters and the running cost. We impose the following set of assumptions on the data which guarantee the existence of a solution to the Itô equation (4.1) (see, for example, [24, 15]). These replace (A1), and are assumed throughout this section by default. In these hypotheses, C_R is a positive constant, depending on $R \in (0,\infty)$. Also $a \coloneqq \frac{1}{2}\upsigma\upsigma'$, $\mathbb{R}_*^m \coloneqq \mathbb{R}^m \setminus \{0\}$, and $\|M\| \coloneqq \bigl(\operatorname{trace} MM'\bigr)^{1/2}$ denotes the Hilbert–Schmidt norm of a $d\times k$ matrix M for $d,k\in\mathbb{N}$.

$$\begin{aligned}
|b(x,z)-b(y,z)|^2 + \|\upsigma(x)-\upsigma(y)\|^2 &+ \int_{\mathbb{R}_*^m} |g(x,\xi)-g(y,\xi)|^2 \mathit{\Pi}(\mathrm{d}\xi) \\
&+ |\mathcal{R}(x,z)-\mathcal{R}(y,z)|^2 \,\le\, C_R|x-y|^2 \qquad \forall\, x,y\in B_R\,,\ \forall\, z\in\mathcal{Z}\,,\\
\bigl\langle x, b(x,z)\bigr\rangle^+ + \|\upsigma(x)\|^2 + \int_{\mathbb{R}_*^m} |g(x,\xi)|^2 \mathit{\Pi}(\mathrm{d}\xi) &\,\le\, C_1(1+|x|^2) \qquad \forall\,(x,z)\in\mathbb{R}^d\times\mathcal{Z}\,,\\
\sum_{i,j} a^{ij}(x)\zeta_i\zeta_j \,\ge\, (C_R)^{-1}|\zeta|^2 &\quad \forall \zeta\in\mathbb{R}^d\,,\ \forall\, x\in B_R\,.
\end{aligned}$$

The measure ν in (1.1) then takes the form $\nu(x,A) = \mathit{\Pi}\bigl(\{\xi\in\mathbb{R}_*^m : g(x,\xi)\in A\}\bigr)$, and it clearly satisfies $\int_{\mathbb{R}^d}|y|^2\,\nu(x,\mathrm{d}y) < C_R|x|^2$. Note that for this model $\bar{\nu} = \nu(x,\mathbb{R}^d)$ is constant. It is evident that if $g(x,\xi)$ does not depend on x, then ν is translation invariant.

4.3. Existence of solutions. For any admissible control Z_t, the Itô equation in (4.1) has a unique strong solution [15], is right-continuous w.p.1, and is a strong Feller process. On the other hand, if Z_t is a Markov control, that is, if it takes the form $Z_t = v(t,X_t)$ for some Borel measurable function $v\colon \mathbb{R}_+\times\mathbb{R}^d$, then it follows from the results in [37] that, under the assumptions in subsection 4.2, the diffusion

$$\mathrm{d}\tilde{X}_t \,=\, b(\tilde{X}_t, v(t,\tilde{X}_t))\,\mathrm{d}t + \upsigma(\tilde{X}_t)\,\mathrm{d}W_t\,, \quad X_0 = x\in\mathbb{R}^d \tag{4.2}$$

has a unique strong solution. As shown in [16], since the the Lévy measure is finite, the solution of (4.1) can be constructed in a piecewise fashion using the solution of (4.2) (see also [38]). It thus follows that, under a Markov control, (4.2) has a unique strong solution. In addition, its transition probability has positive mass.

Of fundamental importance in the study of functionals of X is Itô's formula. For $f\in\mathcal{C}_b^2(\mathbb{R}^d)$ and Z_s an admissible control, it holds that

$$f(X_t) \,=\, f(X_0) + \int_0^t \mathcal{A}f(X_s,Z_s)\,\mathrm{d}s + \mathcal{M}_t \quad \text{a.s.,} \tag{4.3}$$

with $\mathcal{A}$ as in (1.1), and

$$\begin{aligned}
\mathcal{M}_t \,\coloneqq\, &\int_0^t \bigl\langle \nabla f(X_s), \upsigma(X_s)\,\mathrm{d}W_s\bigr\rangle \\
&+ \int_0^t \int_{R_*^m} \Bigl(f\bigl(X_{s-}+g(X_{s-},\xi)\bigr) - f(X_{s-})\Bigr)\,\widetilde{\mathcal{N}}(\mathrm{d}s,\mathrm{d}\xi)
\end{aligned} \tag{4.4}$$

is a local martingale. In the lemma which follows, we show that Krylov's extension of the Itô formula [39, p. 122] is valid for functions f in the local Sobolev space $\mathcal{W}^{2,d}_{\mathrm{loc}}(\mathbb{R}^d)$ satisfying $\mathcal{A}_v f \in L^d_{\mathrm{loc}}(\mathbb{R}^d)$. This is stated for Markov controls $v \in \mathfrak{V}_{\mathsf{sm}}$, which suits our framework, and proved in Appendix A. However, it can be easily extended to admissible controls with a simple variation of the proof.

Lemma 4.1. *Let D be a smooth $\mathcal{C}^{1,1}$ domain, and $\uptau(D) \coloneqq \inf\{t > 0\colon X_t \notin D\}$. For any $f \in \mathcal{W}^{2,d}_{loc}(\mathbb{R}^d)$ such that $\mathcal{A}_v f \in L^d_{loc}(\mathbb{R}^d)$, we have*

$$\mathbb{E}^v_x\bigl[f(X_{t\wedge\uptau(D)})\bigr] \;=\; f(x) + \mathbb{E}^v_x\left[\int_0^{t\wedge\uptau(D)} \mathcal{A}_v f(X_{s-})\,\mathrm{d}s\right] \tag{4.5}$$

for $t \in [0,\infty]$, $x \in D$ and $v \in \mathfrak{V}_{\mathsf{sm}}$, and the right-hand side of (4.5) is finite.

Recall that, in the context of diffusions, a control $v \in \mathfrak{V}_{\mathsf{sm}}$ is called *stable* if the process X under v is positive Harris recurrent. This is of course equivalent to the existence of an invariant probability measure for X, and it follows by the Theorem in [27] that μ_v is an invariant probability measure for the diffusion if and only if it is infinitesimally invariant for the operator $\mathcal{A}$ in the sense of (2.1). Thus the two notions of stable controls agree.

4.4. Existence of an optimal stationary Markov control.

Definition 4.2. *For $Z \in \mathfrak{Z}$ and $x \in \mathbb{R}^d$, we define the mean empirical measures $\{\bar\zeta^Z_{x,t} : t > 0\}$, and (random) empirical measures $\{\zeta^Z_{x,t} : t > 0\}$, by*

$$\bar\zeta^Z_{x,t}(f) \;=\; \int_{\mathbb{R}^d\times\mathcal{Z}} f(x,z)\,\bar\zeta^Z_{x,t}(\mathrm{d}x,\mathrm{d}z) \;\coloneqq\; \frac{1}{t}\int_0^t \mathbb{E}^Z_x\left[\int_{\mathcal{Z}} f(X_s,z)\,Z_s(\mathrm{d}z)\right]\mathrm{d}s\,, \tag{4.6}$$

and $\zeta^Z_{x,t}$ as in (4.6) but without the expectation $\mathbb{E}^Z_x$, respectively, for all $f \in \mathcal{C}_b(\mathbb{R}^d\times\mathcal{Z})$.

We let $\overline{\mathbb{R}}^d$ denote the one-point compactification of $\mathbb{R}^d$, and we view $\mathbb{R}^d \subset \overline{\mathbb{R}}^d$ via the natural imbedding. As a result, $\mathcal{P}(\mathbb{R}^d \times \mathcal{Z})$ is viewed as a subset of $\mathcal{P}(\overline{\mathbb{R}}^d \times \mathcal{Z})$. Let $\bar{\mathcal{G}}$ denote the closure of $\mathcal{G}$ in $\mathcal{P}(\overline{\mathbb{R}}^d \times \mathcal{Z})$.

Lemma 4.3. *Almost surely, every limit $\hat\zeta \in \mathcal{P}(\overline{\mathbb{R}}^d \times \mathcal{Z})$ of $\zeta^Z_{x,t}$ as $t \to \infty$ takes the form $\hat\zeta = \delta\zeta' + (1-\delta)\zeta''$ for some $\delta \in [0,1]$, with $\zeta' \in \mathcal{G}$ and $\zeta''(\{\infty\}\times\mathcal{Z}) = 1$. The same claim holds for the mean empirical measures, without the qualifier 'almost surely'.*

Proof. Write $\hat\zeta = \delta\zeta' + (1-\delta)\zeta''$ for some $\zeta' \in \mathcal{P}(\mathbb{R}^d\times\mathcal{Z})$, and $\zeta''(\{\infty\}\times\mathcal{Z}) = 1$. Recall $\mathcal{C}$ defined in the beginning of subsection 2.1. For $f \in \mathcal{C}$, applying Itô's formula, we obtain

$$\frac{f(X_t) - f(X_0)}{t} \;=\; \frac{1}{t}\int_0^t \mathcal{A}_{Z_s} f(X_s)\,\mathrm{d}s + \frac{1}{t}\,\mathcal{M}_t\,,$$

where $\mathcal{M}_t$ is given in (4.4). As shown in the proof of [24, Lemma 3.4.6], we have $\frac{1}{t}\int_0^t \bigl\langle\nabla f(X_s), \upsigma(X_s)\,\mathrm{d}W_s\bigr\rangle \to 0$ a.s. as $t \to \infty$.

Define

$$M_{1,t} \;\coloneqq\; \int_0^t \int_{R^m_*} \Bigl(f\bigl(X_{s-} + g(X_{s-},\xi)\bigr) - f(X_{s-})\Bigr)\,\mathcal{N}(\mathrm{d}s,\mathrm{d}\xi)\,, \tag{4.7}$$

and $M_{2,t}$ analogously by replacing $\mathcal{N}(\mathrm{d}s,\mathrm{d}\xi)$ by $\mathit{\Pi}(\mathrm{d}\xi)\,\mathrm{d}s$ in (4.7). Note that the second integral in (4.4), denoted as M_t, is a square integrable martingale, and takes the form $M_t = M_{1,t} - M_{2,t}$. Since f is bounded on $\mathbb{R}^d$ and $\mathit{\Pi}$ is a finite measure, we have $\langle M_1\rangle_t \le C_1\,\mathcal{N}(t,\mathbb{R}^m_*)$, and $\langle M_2\rangle_t \le C_2 t$ for some positive constants C_1 and C_2. Since $\langle M\rangle_t \le \langle M_1\rangle_t + \langle M_2\rangle_t$, then by Proposition 7.1 in [40] we obtain $\limsup_{t\to\infty}\frac{\langle M\rangle_t}{t} < \infty$ a.s. For the discrete parameter square-integrable martingale $\{M_n : n\in\mathbb{N}\}$, it is well known that $\lim_{n\to\infty}\frac{M_n}{\langle M\rangle_n} = 0$ a.s. on the event $\{\langle M\rangle_\infty = \infty\}$. Thus, we obtain

$$\lim_{n\to\infty}\frac{M_n}{n} \,=\, 0 \quad \text{a.s.} \tag{4.8}$$

on the event $\{\langle M\rangle_\infty = \infty\}$. Since f is bounded, then for some constant $C>0$, we have

$$\sup_{t\in[n,n+1]}\frac{|M_t - M_n|}{n} \,\le\, \frac{C}{n}\bigl(\mathcal{N}(n+1,\mathbb{R}^m_*) - \mathcal{N}(n,\mathbb{R}^m_*) + 1\bigr) \xrightarrow[n\to\infty]{} 0\,, \tag{4.9}$$

and (4.8)–(4.9) imply that $\lim_{t\to\infty}\frac{1}{t}\,M_t \to 0$ a.s. on the event $\{\langle M\rangle_\infty = \infty\}$.

Next, we examine convergence on the event $\{\langle M\rangle_\infty < \infty\}$. It is well known that a square-integrable martingale $\{M_n : n\in\mathbb{N}\}$ with quadratic variation $\langle M\rangle$ satisfies $\{\langle M\rangle_\infty < \infty\} \subset \{M_n \to\ \}$ a.s., where we write $\{M_n\to\ \}$ for the event on which (M_n) converges to a real-valued limit [41, Theorem 2.15]. Thus (4.8) holds on the event $\{\langle M\rangle_\infty < \infty\}$, and it then follows by (4.9) that $\lim_{t\to\infty}\frac{1}{t}\,M_t \to 0$ a.s.

Thus we have shown that $\lim_{t\to\infty}\frac{1}{t}\,\mathcal{M}_t \to 0$ a.s., and the claims of the lemma then follow as in the proof of [24, Theorem 3.4.7]. □

THEOREM 4.4. *There exists an optimal control $v\in\mathfrak{U}_{\mathsf{ssm}}$ for the ergodic problem. In addition, every stationary Markov optimal control v_* is in $\mathfrak{U}_{\mathsf{ssm}}$, and is pathwise optimal in somewhat stronger sense, that is, it satisfies*

$$\liminf_{T\to\infty}\,\frac{1}{T}\biggl[\int_0^T \mathcal{R}(X_t,Z_t)\,\mathrm{d}t\biggr] \,\ge\, \limsup_{T\to\infty}\,\frac{1}{T}\biggl[\int_0^T \mathcal{R}\bigl(X_t,v_*(X_t)\bigr)\,\mathrm{d}t\biggr] \,=\, \underline{\varrho} \tag{4.10}$$

a.s. for any admissible control Z_t.

Proof. Define $\underline{\varrho} \coloneqq \inf_{\pi\in\mathcal{G}}\pi(\mathcal{R})$. Following the proof of [24, Theorem 3.4.5], we have $\underline{\varrho} = \pi_{v_*}(\mathcal{R})$ for some $v_*\in\mathfrak{U}_{\mathsf{ssm}}$. Also, (4.10) holds by Lemma 4.3 and the proof in [24, Theorem 3.4.7]. □

4.5. The ergodic HJB equation. We summarize the results in the following theorem.

THEOREM 4.5. *We assume (H2) for some $\hat{v}\in\mathfrak{U}_{\mathsf{sm}}$. Then we have the following:*

(*a*) *There exists a unique function $V\in\mathscr{W}^{2,p}_{loc}(\mathbb{R}^d)$, $p>d$, with $V(0)=0$, which is bounded from below in $\mathbb{R}^d$ and solves $\min_{z\in\mathcal{Z}}\bigl[\mathcal{A}_z V(x) + \mathcal{R}(x,z)\bigr] = \varrho$, with $\varrho = \underline{\varrho}$. For $\varrho < \underline{\varrho}$, there is no such solution. Moreover, if ν has locally compact support* (*see Definition 2.7*)*, then $V\in\mathcal{C}^2(\mathbb{R}^d)$.*

(*b*) *A control $v\in\mathfrak{U}_{\mathsf{sm}}$ is optimal if and only if it satisfies*

$$b^i_v(x)\,\partial_i V(x) + \mathcal{R}_v(x) \,=\, \inf_{z\in\mathcal{Z}}\bigl[b^i(x,z)\partial_i V(x) + \mathcal{R}(x,z)\bigr] \quad \textit{a.e. } x\in\mathbb{R}^d\,. \tag{4.11}$$

(*c*) *The solution V has the stochastic representation*

$$V(x) \,=\, \lim_{r\searrow 0}\,\inf_{v\in\mathfrak{U}_{\mathsf{ssm}}}\,\mathbb{E}^v_x\biggl[\int_0^{\breve{\tau}_r}\bigl(\mathcal{R}_v(X_t) - \underline{\varrho}\bigr)\,\mathrm{d}t\biggr]\,.$$

Proof. Under the assumptions in subsection 4.2, it is straightforward to establish Theorem 3.2. Thus, part (a) follows from Theorem 3.3 and Remark 3.4. Using the Itô formula in Lemma 4.1, one can readily show that any v which satisfies (4.11) is stable and optimal. The necessity part of (b) follows by Theorem 3.5. Part (c) can be established by following the proof of Lemma 3.6.9 in [24].

5. Concluding remarks. The results in this paper extend naturally to models under uniform stability, in which case, of course, we do not need to assume that $\mathcal{R}$ is coercive. Suppose that there exist nonnegative functions $\Psi \in \mathcal{C}^2(\mathbb{R}^d)$, and $h\colon \mathbb{R}^d \times \mathcal{Z}$, with $h \ge 1$ and locally bounded, satisfying

$$\mathcal{A}_z \Psi(x) \;\le\; \kappa\,\mathbb{1}_{\mathcal{B}}(c) - h(x,z) \qquad \forall\,(x,z) \in \mathbb{R}^d \times \mathcal{Z}\,, \tag{5.1}$$

for some constant κ and a ball $\mathcal{B} \subset \mathbb{R}^d$. In addition, suppose that either $\mathcal{R}$ is bounded, or that $|\mathcal{R}|$ grows slower than h. Under (5.1), the jump diffusion is positive recurrent under any stationary Markov control, and the collection of ergodic occupation measures is tight. Using Ψ as a barrier, all the results in section 4 can be readily obtained, and moreover, for any $v \in \mathfrak{V}_{\mathsf{sm}}$, the Poisson equation $\mathcal{A}_v \Phi = \mathcal{R}_v - \varrho_v$ has a solution in $\mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, for any $p > 1$, which is unique, up to an additive constant, in the class of functions Φ which satisfy $|\Phi| \le C(1 + h_v)$ for some constant C.

We have not considered allowing the jumps to be control dependent, primarily because this is not manifested in the queueing network model motivating this work, but also because this would require us to introduce various assumptions on the regularity of the jumps and the Lévy measure (see, for example, [5]). This, however, is an interesting problem for future work.

In conclusion, what we aimed for in this work, was to study the ergodic control problem for jump diffusions controlled through the drift via analytical methods, and under minimal assumptions on the (finite) Lévy measure and the parameters.

Acknowledgments. This work was supported in part by the National Science Foundation through grants DMS-1540162, DMS-1715210, CMMI-1538149, and DMS-1715875, in part by the Army Research Office through grant number W911NF-17-1-0019, and in part Office of Naval Research through grant number N00014-16-1-2956.

Appendix A. Proof of Lemma 4.1.

Proof of Lemma 4.1. The hypothesis implies that $\mathcal{L}_v f$ and $\mathcal{I} f$ are in $L^d(\mathbb{R}^d)$. Without loss of generality we may suppose that f is nonnegative. Recall the function $\chi_n\colon \mathbb{R} \to \mathbb{R}$ in the proof of Lemma 3.1, which satisfies $\chi_n(x) = x$ for $x \le n$ and $\chi_n(x) = n+1$ for $x \ge n+2$. Let $f_n \coloneqq \chi_n \circ f = \chi_n(f)$, and recall (3.2). Then for any $n > \sup_D\, f$, applying the Itô–Krylov formula we have

$$\begin{aligned}
\mathbb{E}^v_x\bigl[f_n(X_{t\wedge\uptau(D)})\bigr] \;&=\; f(x) + \mathbb{E}^v_x\biggl[\int_0^{t\wedge\uptau(D)} \widetilde{\mathcal{L}}_v f(X_{s-})\,\mathrm{d}s\biggr]\\
&\quad + \mathbb{E}^v_x\biggl[\int_0^{t\wedge\uptau(D)}\int_{\mathbb{R}^m_*} \chi_n\bigl(f(X_{s-} + g(X_{s-},\xi))\bigr)\,\mathit{\Pi}(\mathrm{d}\xi)\,\mathrm{d}s\biggr]\\
&\quad + \mathbb{E}^v_x\biggl[\int_0^{t\wedge\uptau(D)} \bigl\langle \nabla f(X_{s-}), \upsigma(X_s)\,\mathrm{d}W_s\bigr\rangle\biggr]\\
&\quad + \mathbb{E}^v_x\biggl[\int_0^{t\wedge\uptau(D)}\int_{\mathbb{R}^m_*} \bigl(\chi_n(f(X_{s-} + g(X_{s-},\xi))) - f(X_{s-})\bigr)\,\widetilde{\mathcal{N}}(\mathrm{d}s,\mathrm{d}\xi)\biggr]\,.
\end{aligned} \tag{A.1}$$

Since $\{X_s\colon 0\le s\le t\}$ has countably many of jumps, it follows by using the martingale property that the fourth term on the right-hand side of (A.1) is equal to 0 (see, for example, [42, Lemma 7.3.2]). Also, the last term on the right-hand side of (A.1) is equal to 0 by [43, Claim 1, page 6]. It is clear that the left side of (A.1) converges to $\mathbb{E}_x^v\bigl[f(X_{t\wedge\uptau(D)})\bigr]$ as $n\to\infty$, while the third term on the right-hand side converges to

$$\mathbb{E}_x^v\biggl[\int_0^{t\wedge\uptau(D)}\widetilde{\mathcal{I}}f(X_{s-})\,\mathrm{d}s\biggr] \;=\; \mathbb{E}_x^v\biggl[\int_0^{t\wedge\uptau(D)}\int_{\mathbb{R}^m_*} f(X_{s-}+g(X_{s-},\xi))\,\mathit{\Pi}(\mathrm{d}\xi)\,\mathrm{d}s\biggr] \tag{A.2}$$

by monotone convergence. To estimate the integral in (A.2), consider the solution $\varphi\in\mathscr{W}^{2,p}(D)\cap\mathscr{W}^{1,p}_0(\overline{D})$ to the Dirichlet problem

$$\mathcal{A}_v\varphi=-h\quad\text{in } D\,,\qquad \varphi=0\quad\text{in } D^c\,,$$

and then let $h=\widetilde{\mathcal{I}}f$. By [44, Theorem 3.1.22] the problem has a unique solution in $\mathscr{W}^{2,d}(D)$ for each $h\in L^d(D)$, and satisfies

$$\bigl\|\varphi\bigr\|_{\mathscr{W}^{2,d}(D)}\;\le\;\kappa_D\,\bigl\|h\bigr\|_{L^d(D)} \tag{A.3}$$

for some constant κ_D. Now since $\mathit{\Pi}(\mathbb{R}^m_*)$ is constant, we have

$$\bigl\|\widetilde{\mathcal{I}}\varphi\bigr\|_{L^d(D)}\;\le\;\tilde\kappa_D\,\bigl\|\varphi\bigr\|_{L^d(D)} \tag{A.4}$$

for some constant $\tilde\kappa_D$. Since $\mathcal{A}_v=\widetilde{\mathcal{L}}_v+\widetilde{\mathcal{I}}$, we have $\widetilde{\mathcal{L}}_v\varphi=-h-\widetilde{\mathcal{I}}\varphi$. Thus, invoking the Alexandroff–Bakelman–Pucci weak maximum principle, together with (A.3) and (A.4), we obtain

$$\begin{aligned}\sup_D\,\lvert\varphi\rvert\;&\le\;C_D\,\bigl\|\widetilde{\mathcal{I}}\varphi+h\bigr\|_{L^d(D)}\\ &\le\;C_D(1+\kappa_D\tilde\kappa_D)\bigl\|h\bigr\|_{L^d(D)}\end{aligned} \tag{A.5}$$

for some constant C_D. This shows that for each $x\in D$, the map $h\mapsto\varphi(x)$ defines a bounded linear functional on $L^d(D)$. By the Riesz representation theorem, we have

$$\varphi(x)\;=\;\int_{B_D} G_D(x,y)h(y)\,\mathrm{d}y$$

for some function $G_D(x,\cdot\,)\in L^{\frac{d}{d-1}}(D)$ (this is nothing else but the Green's function).

By the Krylov–Itô formula, which can be applied since $\varphi=0$ on D^c, the function φ has the stochastic representation

$$\varphi(x)\;=\;\mathbb{E}_x^v\biggl[\int_0^{\uptau(D)}\widetilde{\mathcal{I}}f(X_{s-})\,\mathrm{d}s\biggr]\,. \tag{A.6}$$

Applying this as a bound to (A.2), and using Hölder's inequality, we obtain

$$\begin{aligned}\mathbb{E}_x^v\biggl[\int_0^{t\wedge\uptau(D)}\widetilde{\mathcal{I}}f(X_{s-})\,\mathrm{d}s\biggr]\;&\le\;\mathbb{E}_x^v\biggl[\int_0^{\uptau(D)}\widetilde{\mathcal{I}}f(X_{s-})\,\mathrm{d}s\biggr]\\ &=\;\int_D G_D(x,y)\,\widetilde{\mathcal{I}}f(y)\,\mathrm{d}y\\ &\le\;\bigl\|G_D(x,\cdot)\bigr\|_{L^{\frac{d}{d-1}}(D)}\,\bigl\|\widetilde{\mathcal{I}}f\bigr\|_{L^d(D)}\,.\end{aligned}$$

Thus the integral in (A.2) is finite, and this completes the proof. □

REFERENCES

[1] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[2] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.*, 7(3):1005–1028, 2008.

[3] A. Bensoussan and J.-L. Lions. *Impulse control and quasivariational inequalities.* μ. Gauthier-Villars, Montrouge; Heyden & Son, Inc., Philadelphia, PA, 1984. Translated from the French by J. M. Cole.

[4] J.-L. Menaldi and M. Robin. Ergodic control of reflected diffusions with jumps. *Appl. Math. Optim.*, 35(2):117–137, 1997.

[5] J.-L. Menaldi and M. Robin. On optimal ergodic control of diffusions with jumps. In *Stochastic analysis, control, optimization and applications*, Systems Control Found. Appl., pages 439–456. Birkhäuser Boston, Boston, MA, 1999.

[6] E. Bayraktar, T. Emmerling, and J.-L. Menaldi. On the impulse control of jump diffusions. *SIAM J. Control Optim.*, 51(3):2612–2637, 2013.

[7] M. H. A. Davis, X. Guo, and G. Wu. Impulse control of multidimensional jump diffusions. *SIAM J. Control Optim.*, 48(8):5276–5293, 2010.

[8] J. Liu, K. F. C. Yiu, and A. Bensoussan. Optimal inventory control with jump diffusion and nonlinear dynamics in the demand. *SIAM J. Control Optim.*, 56(1):53–74, 2018.

[9] A. Arapostathis, G. Pang, and N. Sandrić. Ergodicity of Lévy-driven SDEs arising from multiclass many-server queues. *ArXiv e-prints*, 1707.09674, 2017.

[10] T. Komatsu. Markov processes associated with certain integro-differential operators. *Osaka J. Math.*, 10:271–303, 1973.

[11] A. G. Bhatt and V. S. Borkar. Occupation measures for controlled Markov processes: characterization and optimality. *Ann. Probab.*, 24(3):1531–1562, 1996.

[12] R. F. Bass. Stochastic differential equations with jumps. *Probab. Surv.*, 1:1–19 (electronic), 2004.

[13] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov. On parabolic inequalities for generators of diffusions with jumps. *Probab. Theory Related Fields*, 158(1-2):465–476, 2014.

[14] M. Foondun. Harmonic functions for a class of integro-differential operators. *Potential Anal.*, 31(1):21–44, 2009.

[15] I. I. Gihman and A. V. Skorohod. *Stochastic differential equations*, volume 72 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer-Verlag, Berlin, 1972.

[16] A. V. Skorokhod. *Asymptotic methods in the theory of stochastic differential equations*, volume 78 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1989. Translated from the Russian by H. H. McFaden.

[17] E. Löcherbach and V. Rabiet. Ergodicity for multidimensional jump diffusions with position dependent jump rate. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(3):1136–1163, 2017.

[18] R. F. Bass. Regularity results for stable-like operators. *J. Funct. Anal.*, 257(8):2693–2722, 2009.

[19] C. Bjorland, L. Caffarelli, and A. Figalli. Non-local gradient dependent operators. *Adv. Math.*, 230(4-6):1859–1894, 2012.

[20] L. Caffarelli and L. Silvestre. Regularity theory for fully nonlinear integro-differential equations. *Comm. Pure Appl. Math.*, 62(5):597–638, 2009.

[21] L. Caffarelli and L. Silvestre. Regularity results for nonlocal equations by approximation. *Arch. Ration. Mech. Anal.*, 200(1):59–88, 2011.

[22] M. H. Protter and H. F. Weinberger. *Maximum principles in differential equations*. Springer-Verlag, New York, 1984. Corrected reprint of the 1967 original.

[23] R. Z. Has′minskiĭ. Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations. *Theory Probab. Appl.*, 5(2):179–196, 1960.

[24] A. Arapostathis, V. S. Borkar, and M. K. Ghosh. *Ergodic control of diffusion processes*, volume 143 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2012.

[25] S. V. Shaposhnikov. On nonuniqueness of solutions to elliptic equations for probability measures. *J. Funct. Anal.*, 254(10):2690–2705, 2008.

[26] V. I. Bogachev, M. Rökner, and V. Stannat. Uniqueness of solutions of elliptic equations and uniqueness of invariant measures of diffusions. *Mat. Sb.*, 193(7):3–36, 2002.

[27] P. Echeverría. A criterion for invariant measures of Markov processes. *Z. Wahrscheinlichkeitstheorie verw Gebiete*, 61(1):1–16, 1982.

[28] V. I. Bogachev, N. V. Krylov, and M. Röckner. On regularity of transition probabilities and

invariant measures of singular diffusions under minimal conditions. *Comm. Partial Differential Equations*, 26(11-12):2037–2080, 2001.

[29] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 1983.

[30] C. B. Morrey, Jr. *Multiple integrals in the calculus of variations*. Die Grundlehren der mathematischen Wissenschaften, Band 130. Springer-Verlag New York, Inc., New York, 1966.

[31] T. Sato. Positive solutions with weak isolated singularities to some semilinear elliptic equations. *Tohoku Math. J. (2)*, 47(1):55–80, 1995.

[32] R. Z. Has′minskiĭ. *Stochastic stability of differential equations*. Sijthoff & Noordhoff, The Netherlands, 1980.

[33] V. I. Bogachev and M. Röckner. A generalization of Khas′minskiĭ's theorem on the existence of invariant measures for locally integrable drifts. *Theory Probab. Appl.*, 45(3):363–378, 2001.

[34] I. Kontoyiannis and S. P. Meyn. On the f-norm ergodicity of Markov processes in continuous time. *Electron. Commun. Probab.*, 21:Paper No. 77, 10, 2016.

[35] A. Arapostathis, M. K. Ghosh, and S. I. Marcus. Harnack's inequality for cooperative weakly coupled elliptic systems. *Comm. Partial Differential Equations*, 24(9-10):1555–1571, 1999.

[36] E. J. Anderson and P. Nash. *Linear programming in infinite-dimensional spaces*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Ltd., Chichester, 1987.

[37] I. Gyöngy and N. Krylov. Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Theory Related Fields*, 105(2):143–158, 1996.

[38] C. W. Li. Lyapunov exponents of nonlinear stochastic differential equations with jumps. In *Stochastic inequalities and applications*, volume 56 of *Progr. Probab.*, pages 339–351. Birkhäuser, Basel, 2003.

[39] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Applications of Mathematics*. Springer-Verlag, New York, 1980.

[40] S. M. Ross. *Introduction to probability models*. Elsevier/Academic Press, Amsterdam, 2014. Eleventh edition.

[41] P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980. Probability and Mathematical Statistics.

[42] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[43] B. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions*. Universitext. Springer-Verlag, Berlin, second edition, 2007.

[44] M. G. Garroni and J. L. Menaldi. *Second order elliptic integro-differential problems*, volume 430 of *Chapman & Hall/CRC Research Notes in Mathematics*. Chapman & Hall/CRC, Boca Raton, FL, 2002.