

# CONVERGENCE OF THE RELATIVE VALUE ITERATION FOR THE ERGODIC CONTROL PROBLEM OF NONDEGENERATE DIFFUSIONS UNDER NEAR-MONOTONE COSTS

ARI ARAPOSTATHIS\*, VIVEK S. BORKAR†, AND K. SURESH KUMAR‡

**Abstract.** We study the relative value iteration for the ergodic control problem under a near-monotone running cost structure for a nondegenerate diffusion controlled through its drift. This algorithm takes the form of a quasilinear parabolic Cauchy initial value problem in  $\mathbb{R}^d$ . We show that this Cauchy problem stabilizes, or in other words, that the solution of the quasilinear parabolic equation converges for every bounded initial condition in  $C^2(\mathbb{R}^d)$  to the solution of the Hamilton–Jacobi–Bellman (HJB) equation associated with the ergodic control problem.

**Key words.** controlled diffusions; ergodic control; Hamilton–Jacobi–Bellman equation; relative value iteration; reverse martingales

**AMS subject classifications.** Primary, 93E15, 93E20; Secondary, 60J25, 60J60, 90C40

**1. Introduction.** This paper is concerned with the time-asymptotic behavior of an optimal control problem for a nondegenerate diffusion controlled through its drift and described by an Itô stochastic differential equation (SDE) in  $\mathbb{R}^d$  having the following form:

$$(1.1) \quad dX_t = b(X_t, U_t) dt + \sigma(X_t) dW_t.$$

Here  $U_t$  is the control variable that takes values in some compact metric space. We impose standard assumptions on the data to guarantee the existence and uniqueness of solutions to (1.1). These are described in §3.1. Let  $r: \mathbb{R}^d \times \mathbb{U} \rightarrow \mathbb{R}$  be a continuous function bounded from below, which without loss of generality we assume it is nonnegative, referred to as the *running cost*. As is well known, the ergodic control problem, in its *almost sure* (or *pathwise*) formulation, seeks to a.s. minimize over all admissible controls  $U$  the functional

$$(1.2) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t r(X_s, U_s) ds.$$

A weaker, *average* formulation seeks to minimize

$$(1.3) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbb{E}^U [r(X_s, U_s)] ds.$$

---

\*Department of Electrical and Computer Engineering, The University of Texas at Austin, 1 University Station, Austin, TX 78712 (ari@mail.utexas.edu). This author’s work was supported in part by the Office of Naval Research through the Electric Ship Research and Development Consortium.

†Department of Electrical Engineering, Indian Institute of Technology, Powai, Mumbai 400076, India. (borkar.vs@gmail.com). This author’s work was supported in part by the J. C. Bose Fellowship from the Government of India and grant 11IRCCSG014 from IRCC, IIT Bombay.

‡Department of Mathematics, Indian Institute of Technology, Powai, Mumbai 400076, India. (suresh@math.iitb.ac.in).

Here  $\mathbb{E}^U$  denotes the expectation operator associated with the probability measure on the canonical space of the process under the control  $U$ . We let  $\varrho$  be defined as

$$(1.4) \quad \varrho \triangleq \inf_U \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbb{E}^U [r(X_s, U_s)] ds,$$

i.e., the infimum of (1.3) over all admissible controls (for the definition of admissible controls see §3.1). Under suitable hypotheses solutions to the ergodic control problem can be synthesized via the Hamilton–Jacobi–Bellman (HJB) equation

$$(1.5) \quad a^{ij}(x) \partial_{ij} V + H(x, \nabla V) = \varrho,$$

where  $a = [a^{ij}]$  is the symmetric matrix  $\frac{1}{2} \sigma \sigma^\top$  and

$$H(x, p) \triangleq \min_u \{b(x, u) \cdot p + r(x, u)\}.$$

The desired characterization is that a stationary Markov control  $v$  is optimal for the ergodic control problem if and only if it satisfies

$$(1.6) \quad H(x, \nabla V(x)) = b(x, v(x)) \cdot \nabla V(x) + r(x, v(x))$$

a.e. in  $\mathbb{R}^d$ . Obtaining solutions to (1.5) is further complicated by the fact that  $\varrho$  is unknown. For controlled Markov chains the *relative value iteration* originating in the work of White [20] provides an algorithm for solving the ergodic dynamic programming equation for the finite state finite action case. Moreover its ramifications have given rise to popular learning algorithms (*Q-learning*) [1].

In [3] we introduced a continuous time, continuous state space analog of White’s relative value iteration (RVI) given by the quasilinear parabolic evolution equation

$$(1.7) \quad \partial_t \varphi(t, x) = a^{ij}(x) \partial_{ij} \varphi(t, x) + H(x, \nabla \varphi) - \varphi(t, 0), \quad \varphi(0, x) = \varphi_0(x).$$

Under a uniform (geometric) ergodicity condition that ensures the well-posedness of the associated HJB equation we showed in [3] that the solution of (1.7) converges as  $t \rightarrow \infty$  to a solution of (1.5), the limit being independent of the initial condition  $\varphi_0$ . In a related work we extended these results to zero-sum stochastic differential games and controlled diffusions under the risk sensitive criterion [5].

Even though the work in [3] was probably the first such study of convergence of a relative iteration scheme for continuous time and space Markov processes, the blanket stability hypothesis imposed weakens these results. Models of controlled diffusions enjoying a uniform geometric ergodicity do not arise often in applications. Rather, what we frequently encounter is a running cost which has a structure which penalizes unstable behavior and thus renders all stationary optimal controls stable. Such is the case for quadratic costs typically used in linear control models. A fairly general class of running costs of this type, which includes ‘norm-like’ costs, consists of costs satisfying the *near-monotone* condition:

$$(1.8) \quad \{x \in \mathbb{R}^d : \min_u r(x, u) \leq \varrho\} \text{ is a compact set.}$$

In this paper we relax the blanket geometric ergodicity assumption and study the relative value iteration in (1.7) under the near-monotone hypothesis (1.8). It is well known that for near-monotone costs the HJB equation (1.5) possesses a unique up to a constant solution  $V$  which is bounded below in  $\mathbb{R}^d$  [4]. However, this uniqueness result is restricted. In general, for  $\beta > \rho$  the equation

$$(1.9) \quad a^{ij}(x) \partial_{ij} V + H(x, \nabla V) = \beta$$

can have a multitude of solutions which are bounded below [4]. As a result, the policy iteration algorithm (PIA) may fail to converge to the optimal value [2, 17]. In order to guarantee convergence of the PIA to an optimal control, in addition to the near-monotone assumption, a blanket Lyapunov condition is imposed in [17, Theorem 5.2] which renders all stationary Markov controls stable. In contrast, the RVI algorithm always converges to the optimal value function when initialized with some bounded initial value  $\varphi_0$ . The reason behind the difference in performance of the two algorithms can be explained as follows: First, recall that the PIA algorithm consists of the following steps:

1. Initialization. Set  $k = 0$  and select some stationary Markov control  $v_0$  which yields a finite average cost.
2. Value determination. Determine the average cost  $\rho_{v_k}$  under the control  $v_k$  and obtain a solution  $V_k$  to the Poisson equation

$$a^{ij}(x) \partial_{ij} V_k + b^i(x, v_k(x)) \partial_i V_k(x) + r(x, v_k(x)) = \rho_{v_k}, \quad x \in \mathbb{R}^d.$$

3. Termination. If

$$H(x, \nabla V_k) = [b(x, v_k(x)) \cdot \nabla V_k(x) + r(x, v_k(x))] \quad \text{a.e.},$$

then return  $v_k$ .

4. Policy improvement. Select  $v_{k+1} \in \mathfrak{U}_{\text{SM}}$  which satisfies

$$v_{k+1}(x) \in \underset{u \in \mathfrak{U}}{\text{Arg min}} [b(x, u) \cdot \nabla V_k(x) + r(x, u)], \quad x \in \mathbb{R}^d.$$

It is straightforward to show that if  $\widehat{V}$  is a solution to (1.9) whose growth rate does not exceed the growth rate of an optimal value function  $V$  from (1.5), or in other words the weighted norm  $\|\widehat{V}\|_V$  is finite, then  $\beta = \rho$  and  $\widehat{V}$  is an optimal value function. It turns out that if the value function  $\varphi_0$  determined at the first step  $k = 0$  does not grow faster than an optimal value function  $V$  then the algorithm will converge to an optimal value function. Otherwise, it might converge to a solution of (1.9) that is not optimal. However, the growth rate of an optimal value function is not known, and there is no simple way of selecting the initial control  $v_0$  that will result in the right growth rate for  $\varphi_0$ . To do so one must solve a HJB-type equation, which is precisely what the PIA algorithm tries to avoid. In contrast, as we show in this paper, the solution of the RVI algorithm has the property that  $x \mapsto \varphi(t, x)$  has the

same growth rate as the optimal value function  $V$ , asymptotically in  $t$ . This is an essential ingredient of the mechanism responsible for convergence.

The proof of convergence of (1.7) is facilitated by the study of the *value iteration* (VI) equation

$$(1.10) \quad \partial_t \bar{\varphi}(t, x) = a^{ij}(x) \partial_{ij} \bar{\varphi}(t, x) + H(x, \nabla \bar{\varphi}) - \varrho, \quad \bar{\varphi}(0, x) = \varphi_0(x).$$

The initial condition is the same as in (1.7). Also  $\varrho$  is as in (1.4), so it is assumed known. Note that if  $\varphi$  is a solution of (1.7), then

$$(1.11) \quad \bar{\varphi}(t, x) = \varphi(t, x) - \varrho t + \int_0^t \varphi(s, 0) ds, \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d.$$

solves (1.10). We have in particular that

$$(1.12) \quad \bar{\varphi}(t, x) - \bar{\varphi}(t, 0) = \varphi(t, x) - \varphi(t, 0) \quad \forall x \in \mathbb{R}^d, \quad \forall t \geq 0.$$

It follows that the function  $f \triangleq \varphi - \bar{\varphi}$  does not depend on  $x \in \mathbb{R}^d$  and satisfies

$$(1.13) \quad \frac{df}{dt} + f = \varrho - \bar{\varphi}(t, 0).$$

Conversely, if  $\bar{\varphi}$  is a solution of (1.10) then solving (1.13) one obtains a corresponding solution of (1.7) that takes the form [3, Lemma 4.4]:

$$(1.14) \quad \varphi(t, x) = \bar{\varphi}(t, x) - \int_0^t e^{s-t} \bar{\varphi}(s, 0) ds + \varrho(1 - e^{-t}), \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d.$$

It also follows from (1.14) that if  $t \mapsto \bar{\varphi}(t, x)$  is bounded for each  $x \in \mathbb{R}^d$  then so is the map  $t \mapsto \varphi(t, x)$ , and if the former converges as  $t \rightarrow \infty$ , pointwise in  $x$ , then so does the latter.

We note here that we study solutions of the VI equation that have the stochastic representation

$$(1.15) \quad \bar{\varphi}(t, x) = \inf_U \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) ds + \varphi_0(X_t) \right] - \varrho t,$$

where the infimum is over all admissible controls. These are called *canonical solutions* (see Definition 3.10). The first term in (1.15) is the total cost over the finite horizon  $[0, t]$  with terminal penalty  $\varphi_0$ . Under the uniform geometric ergodicity hypothesis used in [3] it is straightforward to show that  $t \mapsto \bar{\varphi}(t, x)$  is locally bounded in  $x \in \mathbb{R}^d$ . In contrast, under the near-monotone hypothesis alone,  $t \mapsto \bar{\varphi}(t, x)$  may diverge for each  $x \in \mathbb{R}^d$ . To show convergence, we first identify a suitable region of attraction of the solutions of the HJB under the dynamics of (1.7) and then show that all  $\omega$ -limit points of the semiflow of (1.7) lie in this region.

While we prefer to think of (1.7) as a continuous time and space relative value iteration, it can also be viewed as a ‘stabilization of a quasilinear parabolic PDE problem’ in analogy to the celebrated result of Has’minskiĭ (see [11]). Thus, the results in this paper are also likely to be of independent interest to the PDE community.

We summarize below the main result of the paper. We make one mild assumption: let  $v^*$  be some optimal stationary Markov control, i.e., a measurable function that satisfies (1.6). It is well known that under the near-monotone hypothesis the diffusion under the control  $v^*$  is positive recurrent. Let  $\mu_{v^*}$  denote the unique invariant probability measure of the diffusion under the control  $v^*$ . We assume that the value function  $V$  in the HJB is integrable under  $\mu_{v^*}$ .

**THEOREM 1.1.** *Suppose that the running cost is near-monotone and that the value function  $V$  of the HJB equation (1.5) for the ergodic control problem is integrable with respect to some optimal invariant probability distribution. Then for any bounded initial condition  $\varphi_0 \in \mathcal{C}^2(\mathbb{R}^d)$  it holds that*

$$\lim_{t \rightarrow \infty} \varphi(t, x) = V(x) - V(0) + \varrho,$$

uniformly on compact sets of  $\mathbb{R}^d$ .

We also obtain a new stochastic representation for the value function of the HJB under near-monotone costs which we state as a corollary. This result is known to hold under uniform geometric ergodicity, but under the near-monotone cost hypothesis alone it is completely new.

**COROLLARY 1.2.** *Under the assumptions of Theorem 1.1 the value function  $V$  of the HJB for the ergodic control problem has the stochastic representation:*

$$V(x) - V(y) = \lim_{t \rightarrow \infty} \left( \inf_U \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) ds \right] - \inf_U \mathbb{E}_y^U \left[ \int_0^t r(X_s, U_s) ds \right] \right)$$

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

We would like to note here that in [7] the authors study the value iteration algorithm for countable state controlled Markov chains, with ‘norm-like’ running costs, i.e.,  $\min_u r(x, u) \rightarrow \infty$  as  $|x| \rightarrow \infty$ . The initial condition  $\varphi_0$  is chosen as some Lyapunov function corresponding to some stable control  $v_0$ . We leave it to the reader to verify that under these hypotheses  $\|V\|_{\varphi_0} < \infty$ . Moreover they assume that  $\varphi_0$  is integrable with respect to the invariant probability distribution  $\mu_{v^*}$  (see the earlier discussion concerning the PIA algorithm). Thus their hypotheses imply that the optimal value function  $V$  from (1.5) is also integrable with respect to  $\mu_{v^*}$ .

The paper is organized as follows. The next section introduces the notation used in the paper. Section 3 starts by describing in detail the model and the assumptions imposed. In §3.2 we discuss some basic properties of the HJB equation for the ergodic control problem under near-monotone costs and the implications of the integrability of the value function under some optimal invariant distribution. In §3.3 we address the issue of existence and uniqueness of solutions to (1.7) and (1.10) and describe some basic properties of these solutions. In §3.4 we exhibit a region of attraction for the solutions of the VI. In §4 we derive some essential growth estimates for the solutions of the VI and show that these solutions have locally bounded oscillation in  $\mathbb{R}^d$ , uniformly in  $t \geq 0$ . Section 5 is dedicated to the proof of convergence of the solutions of the RVI, while §6 concludes with some pointers to future work.

**2. Notation.** The standard Euclidean norm in  $\mathbb{R}^d$  is denoted by  $|\cdot|$ . The set of nonnegative real numbers is denoted by  $\mathbb{R}_+$ ,  $\mathbb{N}$  stands for the set of natural numbers, and  $\mathbb{I}$  denotes the indicator function. We denote by  $\tau(A)$  the *first exit time* of a process  $\{X_t, t \in \mathbb{R}_+\}$  from a set  $A \subset \mathbb{R}^d$ , defined by

$$\tau(A) \triangleq \inf \{t > 0 : X_t \notin A\}.$$

The closure and the boundary of a set  $A \subset \mathbb{R}^d$  are denoted by  $\bar{A}$  and  $\partial A$ , respectively. The open ball of radius  $R$  in  $\mathbb{R}^d$ , centered at the origin, is denoted by  $B_R$ , and we let  $\tau_R \triangleq \tau(B_R)$ , and  $\check{\tau}_R \triangleq \tau(B_R^c)$ .

The term *domain* in  $\mathbb{R}^d$  refers to a nonempty, connected open subset of the Euclidean space  $\mathbb{R}^d$ . For a domain  $D \subset \mathbb{R}^d$ , the space  $\mathcal{C}^k(D)$  ( $\mathcal{C}^\infty(D)$ ) refers to the class of all functions whose partial derivatives up to order  $k$  (of any order) exist and are continuous.

We adopt the notation  $\partial_t \triangleq \frac{\partial}{\partial t}$ , and for  $i, j \in \mathbb{N}$ ,  $\partial_i \triangleq \frac{\partial}{\partial x_i}$  and  $\partial_{ij} \triangleq \frac{\partial^2}{\partial x_i \partial x_j}$ . We often use the standard summation rule that repeated subscripts and superscripts are summed from 1 through  $d$ . For example,

$$a^{ij} \partial_{ij} \varphi + b^i \partial_i \varphi \triangleq \sum_{i,j=1}^d a^{ij} \frac{\partial^2 \varphi}{\partial x_i \partial x_j} + \sum_{i=1}^d b^i \frac{\partial \varphi}{\partial x_i}.$$

For a nonnegative multi-index  $\alpha = (\alpha_1, \dots, \alpha_d)$  we let  $D^\alpha \triangleq \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d}$ . Let  $Q$  be a domain in  $\mathbb{R}_+ \times \mathbb{R}^d$ . Recall that  $\mathcal{C}^{r,k+2r}(Q)$  stands for the set of bounded continuous functions  $\varphi(t, x)$  defined on  $Q$  such that the derivatives  $D^\alpha \partial_t^\ell \varphi$  are bounded and continuous in  $Q$  for

$$|\alpha| + 2\ell \leq k + 2r, \quad \ell \leq r.$$

In general if  $\mathcal{X}$  is a space of real-valued functions on  $Q$ ,  $\mathcal{X}_{\text{loc}}$  consists of all functions  $f$  such that  $f\varphi \in \mathcal{X}$  for every  $\varphi \in \mathcal{C}_c^\infty(Q)$ , the space of smooth functions on  $Q$  with compact support. In this manner we obtain for example the spaces  $\mathcal{C}_{\text{loc}}^{1,2}(\mathbb{R}^d)$  and  $\mathcal{W}_{\text{loc}}^{2,p}(Q)$ .

We won't introduce here the parabolic Sobolev space  $\mathcal{W}^{r,k+2r,p}(Q)$ , since the solutions of (1.7) and (1.10) are in  $\mathcal{C}_{\text{loc}}^{1,2}(\mathbb{R}^d)$ . The only exception is the function  $\psi$  in Theorem 4.7 and the function  $\psi_T$  used in the proof of Lemma 4.8. We refer the reader to [13] for definitions and properties of the parabolic Sobolev space.

### 3. Problem Statement and Preliminary Results.

**3.1. The model.** The dynamics are modeled by a controlled diffusion process  $X = \{X_t, t \geq 0\}$  taking values in the  $d$ -dimensional Euclidean space  $\mathbb{R}^d$ , and governed by the Itô stochastic differential equation in (1.1). All random processes in (1.1) live in a complete probability space  $(\Omega, \mathfrak{F}, \mathbb{P})$ . The process  $W$  is a  $d$ -dimensional standard Wiener process independent of the initial condition  $X_0$ . The control process  $U$  takes values in a compact, metrizable set  $\mathbb{U}$ , and  $U_t(\omega)$  is jointly measurable in

$(t, \omega) \in [0, \infty) \times \Omega$ . Moreover, it is *non-anticipative*: for  $s < t$ ,  $W_t - W_s$  is independent of

$$\mathfrak{F}_s \triangleq \text{the completion of } \sigma\{X_0, U_r, W_r, r \leq s\} \text{ relative to } (\mathfrak{F}, \mathbb{P}).$$

Such a process  $U$  is called an *admissible control*, and we let  $\mathfrak{U}$  denote the set of all admissible controls.

We impose the following standard assumptions on the drift  $b$  and the diffusion matrix  $\sigma$  to guarantee existence and uniqueness of solutions to (1.1).

**(A1)** *Local Lipschitz continuity*: The functions

$$b = [b^1, \dots, b^d]^\top : \mathbb{R}^d \times \mathbb{U} \mapsto \mathbb{R}^d \quad \text{and} \quad \sigma = [\sigma^{ij}] : \mathbb{R}^d \mapsto \mathbb{R}^{d \times d}$$

are locally Lipschitz in  $x$  with a Lipschitz constant  $\kappa_R > 0$  depending on  $R > 0$ . In other words, for all  $x, y \in B_R$  and  $u \in \mathbb{U}$ ,

$$|b(x, u) - b(y, u)| + \|\sigma(x) - \sigma(y)\| \leq \kappa_R |x - y|.$$

**(A2)** *Affine growth condition*:  $b$  and  $\sigma$  satisfy a global growth condition of the form

$$|b(x, u)|^2 + \|\sigma(x)\|^2 \leq \kappa_1(1 + |x|^2) \quad \forall (x, u) \in \mathbb{R}^d \times \mathbb{U},$$

where  $\|\sigma\|^2 \triangleq \text{trace}(\sigma\sigma^\top)$ .

**(A3)** *Local nondegeneracy*: For each  $R > 0$ , we have

$$\sum_{i,j=1}^d a^{ij}(x) \xi_i \xi_j \geq \kappa_R^{-1} |\xi|^2 \quad \forall x \in B_R,$$

for all  $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$ .

We also assume that  $b$  is continuous in  $(x, u)$ .

In integral form, (1.1) is written as

$$(3.1) \quad X_t = X_0 + \int_0^t b(X_s, U_s) ds + \int_0^t \sigma(X_s) dW_s.$$

The second term on the right hand side of (3.1) is an Itô stochastic integral. We say that a process  $X = \{X_t(\omega)\}$  is a solution of (1.1), if it is  $\mathfrak{F}_t$ -adapted, continuous in  $t$ , defined for all  $\omega \in \Omega$  and  $t \in [0, \infty)$ , and satisfies (3.1) for all  $t \in [0, \infty)$  at once a.s.

We define the family of operators  $L^u : \mathcal{C}^2(\mathbb{R}^d) \mapsto \mathcal{C}(\mathbb{R}^d)$ , where  $u \in \mathbb{U}$  plays the role of a parameter, by

$$(3.2) \quad L^u f(x) = a^{ij}(x) \partial_{ij} f(x) + b^i(x, u) \partial_i f(x), \quad u \in \mathbb{U}.$$

We refer to  $L^u$  as the *controlled extended generator* of the diffusion.

Of fundamental importance in the study of functionals of  $X$  is Itô's formula. For  $f \in \mathcal{C}^2(\mathbb{R}^d)$  and with  $L^u$  as defined in (3.2), it holds that

$$(3.3) \quad f(X_t) = f(X_0) + \int_0^t L^{U_s} f(X_s) ds + M_t, \quad \text{a.s.},$$

where

$$M_t \triangleq \int_0^t \langle \nabla f(X_s), \sigma(X_s) dW_s \rangle$$

is a local martingale. Krylov's extension of the Itô formula [12, p. 122] extends (3.3) to functions  $f$  in the local Sobolev space  $W_{\text{loc}}^{2,p}(\mathbb{R}^d)$ ,  $p \geq d$ .

Recall that a control is called *Markov* if  $U_t = v(t, X_t)$  for a measurable map  $v : \mathbb{R}_+ \times \mathbb{R}^d \mapsto \mathbb{U}$ , and it is called *stationary Markov* if  $v$  does not depend on  $t$ , i.e.,  $v : \mathbb{R}^d \mapsto \mathbb{U}$ . Correspondingly, the equation

$$(3.4) \quad X_t = x_0 + \int_0^t b(X_s, v(s, X_s)) ds + \int_0^t \sigma(X_s) dW_s$$

is said to have a *strong solution* if given a Wiener process  $(W_t, \mathfrak{F}_t)$  on a complete probability space  $(\Omega, \mathfrak{F}, \mathbb{P})$ , there exists a process  $X$  on  $(\Omega, \mathfrak{F}, \mathbb{P})$ , with  $X_0 = x_0 \in \mathbb{R}^d$ , which is continuous,  $\mathfrak{F}_t$ -adapted, and satisfies (3.4) for all  $t$  at once, a.s. A strong solution is called *unique*, if any two such solutions  $X$  and  $X'$  agree  $\mathbb{P}$ -a.s., when viewed as elements of  $\mathcal{C}([0, \infty), \mathbb{R}^d)$ . It is well known that under Assumptions (A1)–(A3), for any Markov control  $v$ , (3.4) has a unique strong solution [10].

Let  $\mathfrak{U}_{\text{SM}}$  denote the set of stationary Markov controls. Under  $v \in \mathfrak{U}_{\text{SM}}$ , the process  $X$  is strong Markov, and we denote its transition function by  $P_v^t(x, \cdot)$ . It also follows from the work of [6, 19] that under  $v \in \mathfrak{U}_{\text{SM}}$ , the transition probabilities of  $X$  have densities which are locally Hölder continuous. Thus  $L^v$  defined by

$$L^v f(x) = a^{ij}(x) \partial_{ij} f(x) + b^i(x, v(x)) \partial_i f(x), \quad v \in \mathfrak{U}_{\text{SM}},$$

for  $f \in \mathcal{C}^2(\mathbb{R}^d)$ , is the generator of a strongly-continuous semigroup on  $\mathcal{C}_b(\mathbb{R}^d)$ , which is strong Feller. We let  $\mathbb{P}_x^v$  denote the probability measure and  $\mathbb{E}_x^v$  the expectation operator on the canonical space of the process under the control  $v \in \mathfrak{U}_{\text{SM}}$ , conditioned on the process  $X$  starting from  $x \in \mathbb{R}^d$  at  $t = 0$ .

**3.2. The ergodic control problem.** We assume that the running cost function  $r : \mathbb{R}^d \times \mathbb{U} \rightarrow \mathbb{R}_+$  is continuous and locally Lipschitz in its first argument uniformly in  $u \in \mathbb{U}$ . Without loss of generality we let  $\kappa_R$  be a Lipschitz constant of  $r$  over  $B_R$ . More specifically, we assume that

$$|r(x, u) - r(y, u)| \leq \kappa_R |x - y| \quad \forall x, y \in B_R, \quad \forall u \in \mathbb{U},$$

and all  $R > 0$ .

As mentioned in §1, an important class of running cost functions arising in practice for which the ergodic control problem is well behaved are the near-monotone cost functions.

The ergodic control problem for near-monotone cost functions is characterized by the following theorem which we quote from [4]. Note that we choose to normalize the value function  $V^*$  differently here, in order to facilitate the use of weighted norms.

**THEOREM 3.1.** *There exists a unique function  $V^* \in \mathcal{C}^2(\mathbb{R}^d)$  which solves the HJB equation (1.5), and satisfies  $\min_{\mathbb{R}^d} V^* = 1$ . Also, a control  $v \in \mathfrak{U}_{\text{SM}}$  is optimal*



with respect to the criteria (1.2) and (1.3) if and only if it satisfies (1.6) a.e. in  $\mathbb{R}^d$ . Moreover, recalling that  $\check{\tau}_R = \tau(B_R^c)$ ,  $R > 0$ , we have

$$(3.5) \quad V^*(x) = \inf_{v \in \mathfrak{U}_{\text{SSM}}} \mathbb{E}_x^v \left[ \int_0^{\check{\tau}_R} (r(X_t, v(X_t)) - \varrho) dt + V^*(X_{\tau(B_R^c)}) \right] \quad \forall x \in B_R^c,$$

for all  $R > 0$ .

Recall that control  $v \in \mathfrak{U}_{\text{SM}}$  is called *stable* if the associated diffusion is positive recurrent. We denote the set of such controls by  $\mathfrak{U}_{\text{SSM}}$ , and let  $\mu_v$  denote the unique invariant probability measure on  $\mathbb{R}^d$  for the diffusion under the control  $v \in \mathfrak{U}_{\text{SSM}}$ . Recall that  $v \in \mathfrak{U}_{\text{SSM}}$  if and only if there exists an inf-compact function  $\mathcal{V} \in \mathcal{C}^2(\mathbb{R}^d)$ , a bounded domain  $D \subset \mathbb{R}^d$ , and a constant  $\varepsilon > 0$  satisfying

$$L^v \mathcal{V}(x) \leq -\varepsilon \quad \forall x \in D^c.$$

It follows that the optimal control  $v^*$  in Theorem 3.1 is stable.

We make the following mild technical assumption which is in effect throughout the paper:

*Assumption 3.2.* The value function  $V^*$  is integrable with respect to some optimal invariant probability distribution  $\mu_{v^*}$ .

*Remark 3.3.* Assumption 3.2 is equivalent to the following [4, Lemma 3.3.4]: there exists an optimal stationary control  $v^*$  and an inf-compact function  $\mathcal{V} \in \mathcal{C}^2(\mathbb{R}^d)$  and an open ball  $B \subset \mathbb{R}^d$  such that

$$(3.6) \quad L^{v^*} \mathcal{V}(x) \leq -V^*(x) \quad \forall x \in B^c.$$

For the rest of the paper  $v^* \in \mathfrak{U}_{\text{SSM}}$  denotes some fixed control satisfying (1.6) and (3.6).

*Remark 3.4.* Assumption 3.2 is pretty mild. In the case that  $r$  is bounded it is equivalent to the statement that the mean hitting times to an open bounded set are integrable with respect to some optimal invariant probability distribution. In the case of one dimensional diffusions, provided  $\sigma(x) > \sigma_0$  for some constant  $\sigma_0 > 0$ , and  $\limsup_{|x| \rightarrow \infty} \frac{x b(x)}{\sigma^2(x)} < -\frac{1}{2}$ , then the mean hitting time of  $0 \in \mathbb{R}$  is bounded above by a second-degree polynomial in  $x$  [15, Theorem 5.6]. Therefore, in this case, the existence of second moments for  $\mu_{v^*}$  implies Assumption 3.2.

We need the following lemma.

LEMMA 3.5. Under Assumption 3.2,

$$\mathbb{E}_x^{v^*} [V^*(X_t)] \xrightarrow{t \rightarrow \infty} \mu_{v^*}[V^*] \triangleq \int_{\mathbb{R}^d} V^*(x) \mu_{v^*}(dx) \quad \forall x \in \mathbb{R}^d,$$

where, as defined earlier,  $\mu_{v^*}$  is the invariant probability measure of the diffusion under the control  $v^*$ . Also there exists a constant  $m_r$  depending on  $r$  such that

$$(3.7) \quad \sup_{t \geq 0} \mathbb{E}_x^{v^*} [V^*(X_t)] \leq m_r(V^*(x) + 1) \quad \forall x \in \mathbb{R}^d.$$

*Proof.* Since  $r$  is nonnegative, by Dynkin's formula we have

$$(3.8) \quad \mathbb{E}_x^{v^*} [V^*(X_t)] \leq V^*(x) + \varrho t \quad \forall t \geq 0, \forall x \in \mathbb{R}^d.$$

Therefore, since  $V^*$  is integrable with respect to  $\mu_{v^*}$  by Assumption 3.2, the first result follows by [18, Theorem 5.3 (i)]. The bound in (3.7) is the continuous time analogue of (14.5) in [16]. Recall that a skeleton of a continuous-time Markov process is a discrete-time Markov process with transition probability  $\widehat{P} = \int_0^\infty \alpha(dt)P^t$ , where  $\alpha$  is a probability measure on  $(0, \infty)$ . Since the diffusion is nondegenerate, any skeleton of the process is  $\phi$ -irreducible, with an irreducibility measure absolutely continuous with respect to the Lebesgue measure (for a definition of  $\phi$ -irreducibility we refer the reader to [16, Chapter 4]). It is also straightforward to show that compact subsets of  $\mathbb{R}^d$  are petite. Define the transition probability  $\widetilde{P}$  by

$$\widetilde{P}f(x) = \int_{\mathbb{R}^d} \widetilde{P}(x, dy) f(y) \triangleq \mathbb{E}_x^{v^*} [f(X_t)] \Big|_{t=1}, \quad x \in \mathbb{R}^d$$

for all bounded functions  $f \in \mathcal{C}(\mathbb{R}^d)$ , and

$$g_r(x) \triangleq \mathbb{E}_x^{v^*} \left[ \int_0^1 r(X_s, v^*(X_s)) ds \right], \quad x \in \mathbb{R}^d.$$

Then (1.5) translates into the discrete time Poisson equation:

$$(3.9) \quad \widetilde{P}V^*(x) - V^*(x) = \varrho - g_r(x), \quad x \in \mathbb{R}^d.$$

It easily follows from the near-monotone hypothesis (1.8) that there exists a constant  $\varepsilon_0 > 0$  and a ball  $B_{R_0} \subset \mathbb{R}^d$ ,  $R_0 > 0$ , such that  $g_r(x) - \varrho > \varepsilon_0$  for all  $x \in B_{R_0}^c$ . Since, in addition,  $\int_{\mathbb{R}^d} V^*(x) \mu_{v^*}(dx) < \infty$ , it follows by [16, Theorem 14.0.1] that there exists a constant  $\widetilde{m}$  such that

$$(3.10) \quad \sum_{n=0}^{\infty} |\widetilde{P}^n g_r(x) - \varrho| \leq \widetilde{m}(V^*(x) + 1) \quad \forall x \in \mathbb{R}^d.$$

By (3.9)–(3.10) we obtain

$$(3.11) \quad \begin{aligned} \widetilde{P}^n V^*(x) &= V^*(x) - \sum_{k=0}^{n-1} (\widetilde{P}^k g_r(x) - \varrho) \\ &\leq (\widetilde{m} + 1)(V^*(x) + 1). \end{aligned}$$

By (3.8) and (3.11), writing the arbitrary  $t \in \mathbb{R}_+$  as  $t = n + \delta$  where  $n$  is the integer

part of  $t$  and using the Markov property, we obtain

$$\begin{aligned}
\mathbb{E}_x^{v^*} [V^*(X_t)] &= \mathbb{E}_x^{v^*} \left[ \mathbb{E}_{X_\delta}^{v^*} [V^*(X_{t-\delta})] \right] \\
&= \mathbb{E}_x^{v^*} [\tilde{P}^n V^*(X_\delta)] \\
&\leq \mathbb{E}_x^{v^*} [(\tilde{m} + 1)(V^*(X_\delta) + 1)] \\
&\leq (\tilde{m} + 1)(V^*(x) + \varrho \delta + 1) \\
&\leq (\tilde{m} + 1)(V^*(x) + \varrho + 1) \quad \forall t \geq 0, \forall x \in \mathbb{R}^d,
\end{aligned}$$

thus establishing (3.7).  $\square$

DEFINITION 3.6. We let  $\mathcal{C}_{V^*}(\mathbb{R}^d)$  denote the Banach space of functions  $f \in \mathcal{C}(\mathbb{R}^d)$  with norm

$$\|f\|_{V^*} \triangleq \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{V^*(x)}.$$

We also define

$$\mathcal{O}_{V^*} \triangleq \{f \in \mathcal{C}_{V^*}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d) : f \geq 0\}.$$

**3.3. The relative value iteration.** The RVI and VI equations in (1.7) and (1.10) can also be written in the form

$$(3.12) \quad \partial_t \varphi(t, x) = \min_{u \in \mathbb{U}} [L^u \varphi(t, x) + r(x, u)] - \varphi(t, 0), \quad \varphi(0, x) = \varphi_0(x),$$

$$(3.13) \quad \partial_t \bar{\varphi}(t, x) = \min_{u \in \mathbb{U}} [L^u \bar{\varphi}(t, x) + r(x, u)] - \varrho, \quad \bar{\varphi}(0, x) = \varphi_0(x).$$

DEFINITION 3.7. Let  $\hat{v} = \{\hat{v}_t, t \in \mathbb{R}_+\}$  denote a measurable selector from the minimizer in (3.13) corresponding to a solution  $\bar{\varphi} \in \mathcal{C}_{\text{loc}}^{1,2}(\mathbb{R}^d)$ . This is also a measurable selector from the minimizer in (3.12), provided  $\bar{\varphi}$  and  $\varphi$  are related by (1.11) and (1.14), and vice-versa. Note that the Markov control associated with  $\hat{v}$  is computed ‘backward’ in time (see (1.15)). Hence, for each  $t \geq 0$  we define the (nonstationary) Markov control

$$\hat{v}^t \triangleq \{\hat{v}_s^t = \hat{v}_{t-s}, s \in [0, t]\}.$$

Also, we adopt the simplifying notation

$$\bar{r}(x, u) \triangleq r(x, u) - \varrho.$$

In most of the statements of intermediary results the initial data  $\varphi_0$  is assumed without loss of generality to be nonnegative. We start with a theorem that proves the

existence of a solution to (3.13) that admits the stochastic representation in (1.15). This does not require Assumption 3.2.

First we need the following definition.

DEFINITION 3.8. We define  $\mathbb{R}_T^d \triangleq (0, T) \times \mathbb{R}^d$ , and let  $\overline{\mathbb{R}_T^d}$  denote its closure. We also let  $\mathcal{C}_{V^*}^T(\mathbb{R}^d)$  denote the Banach space of functions in  $\mathcal{C}(\overline{\mathbb{R}_T^d})$  with norm

$$\|f\|_{V^*, T} \triangleq \sup_{(t, x) \in \overline{\mathbb{R}_T^d}} \frac{|f(t, x)|}{V^*(x)}.$$

THEOREM 3.9. Provided  $\varphi_0 \in \mathcal{O}_{V^*}$ , then

$$(3.14a) \quad \overline{\varphi}(t, x) = \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t \overline{r}(X_s, U_s) ds + \varphi_0(X_t) \right],$$

is the minimal solution of (3.13) in  $C_{\text{loc}}^{1,2}((0, \infty) \times \mathbb{R}^d) \cap \mathcal{C}([0, \infty) \times \mathbb{R}^d)$  which is bounded below on  $\mathbb{R}_T^d$ , for any  $T > 0$ . With  $\hat{v}^t$  as defined in Definition 3.7, it admits the representation

$$(3.14b) \quad \overline{\varphi}(t, x) = \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^t \overline{r}(X_s, \hat{v}_s^t(X_s)) ds + \varphi_0(X_t) \right],$$

and it holds that

$$(3.15) \quad \mathbb{E}_x^{\hat{v}^t} [\overline{\varphi}(t - \tau_R, X_{\tau_R}) \mathbb{I}\{\tau_R < t\}] \xrightarrow{R \rightarrow \infty} 0$$

for all  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ . Moreover  $\overline{\varphi}(t, \cdot) \geq -\varrho t$  and satisfies the estimate

$$(3.16) \quad \|\overline{\varphi}\|_{V^*, T} \leq (1 + \varrho T) \max(1, \|\varphi_0\|_{V^*}) \quad \forall T > 0.$$

*Proof.* Let  $r^n$  and  $\varphi_0^n$ , for  $n \in \mathbb{N}$ , be smooth truncations of  $r$  and  $\varphi_0$ , respectively, satisfying  $\|r^n\|_\infty \leq n$  and  $\|\varphi_0^n\|_\infty \leq n$  and such that  $r^n \uparrow r$  and  $\varphi_0^n \uparrow \varphi_0$  as  $n \rightarrow \infty$ . Let  $\varrho_n$  denote the optimal ergodic cost corresponding to  $r^n$ . The boundary value problem

$$(3.17) \quad \begin{aligned} \partial_t \widehat{\varphi}_n^R(t, x) &= \min_{u \in \mathfrak{U}} [L^u \widehat{\varphi}_n^R(t, x) + r^n(x, u)] && \text{in } (0, T) \times B_R \\ \widehat{\varphi}_n^R(0, x) &= \varphi_0^n(x) \quad \forall x \in \overline{B}_R, && \widehat{\varphi}_n^R(t, \cdot)|_{\partial B_R} = \varphi_0^n \quad \forall t \in [0, T], \end{aligned}$$

has a unique nonnegative solution in  $\mathcal{C}^{1,2}((0, T) \times B_R) \cap \mathcal{C}([0, T] \times \overline{B}_R)$  for all  $T > 0$  and  $R > 0$ . This solution has the stochastic representation

$$(3.18) \quad \widehat{\varphi}_n^R(t, x) = \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^{\tau_R \wedge t} r^n(X_s, U_s) ds + \varphi_0^n(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right].$$

where, as defined in §2,  $\tau_R$  denotes the first exit time from the ball  $B_R$ . By (3.18) we obtain

$$\begin{aligned} \widehat{\varphi}_n^R(t, x) &\leq \mathbb{E}_x^{v^*} \left[ \int_0^{\tau_R \wedge t} r^n(X_s, v^*(X_s)) ds + \varphi_0^n(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right] \\ &\leq \max(1, \|\varphi_0\|_{V^*}) \mathbb{E}_x^{v^*} \left[ \int_0^{\tau_R \wedge t} r(X_s, v^*(X_s)) ds + V^*(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right] \\ &\leq \max(1, \|\varphi_0\|_{V^*}) (V^*(x) + \varrho t). \end{aligned}$$

Therefore by the interior estimates of solutions of (3.17) (see [14, Theorem 5.1]) the derivatives  $\{D^\alpha \partial_t^\ell \widehat{\varphi}_n^R : |\alpha| + 2\ell \leq 2, R > 0, n \in \mathbb{N}\}$  are locally Hölder equicontinuous in  $\mathbb{R}_T^d$ . Thus passing to the limit as  $R \rightarrow \infty$  along a subsequence we obtain a nonnegative function  $\widehat{\varphi}_n \in \mathcal{C}_{\text{loc}}^{1,2}(\mathbb{R}_T^d) \cap \mathcal{C}(\overline{\mathbb{R}_T^d})$ , for all  $T > 0$ , which satisfies

$$(3.19) \quad \begin{aligned} \partial_t \widehat{\varphi}_n(t, x) &= \min_{u \in \mathbb{U}} [L^u \widehat{\varphi}_n(t, x) + r^n(x, u)] \quad \text{in } (0, \infty) \times \mathbb{R}^d \\ \widehat{\varphi}_n(0, x) &= \varphi_0^n(x) \quad \forall x \in \mathbb{R}^d. \end{aligned}$$

By using Dynkin's formula on the cylinder  $[0, t] \times B_R$ , we obtain from (3.19) that

$$(3.20) \quad \widehat{\varphi}_n(t, x) = \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^{\tau_R \wedge t} r^n(X_s, U_s) ds + \widehat{\varphi}_n(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right],$$

It follows by (3.18) that  $\|\widehat{\varphi}_n(t, \cdot)\|_\infty \leq n(t+1)$  for all  $n \in \mathbb{N}$  and  $t \geq 0$ . By (3.20) we have the inequality

$$(3.21) \quad \begin{aligned} \widehat{\varphi}_n(t, x) &\leq \mathbb{E}_x^U \left[ \int_0^{\tau_R \wedge t} r^n(X_s, U_s) ds + \widehat{\varphi}_n(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right] \\ &\leq \mathbb{E}_x^U \left[ \int_0^{\tau_R \wedge t} r^n(X_s, U_s) ds + \varphi_0^n(X_t) \mathbb{I}\{\tau_R > t\} \right] + n \mathbb{P}_x^U(\tau_R \leq t) \end{aligned}$$

for all  $U \in \mathfrak{U}$ . Taking limits as  $R \rightarrow \infty$  in (3.21), using dominated convergence, we obtain

$$(3.22) \quad \widehat{\varphi}_n(t, x) \leq \mathbb{E}_x^U \left[ \int_0^t r^n(X_s, U_s) ds + \varphi_0^n(X_t) \right] \quad U \in \mathfrak{U}.$$

Note that

$$(3.23) \quad 0 \leq \widehat{\varphi}_n(t, x) \leq \limsup_{R \rightarrow \infty} \widehat{\varphi}_n^R(t, x) \leq \max(1, \|\varphi_0\|_{V^*}) (V^*(x) + \varrho t).$$

Hence, as mentioned earlier, the derivatives  $\{D^\alpha \partial_t^\ell \widehat{\varphi}_n : |\alpha| + 2\ell \leq 2, n \in \mathbb{N}\}$  are locally Hölder equicontinuous in  $(0, \infty) \times \mathbb{R}^d$ . Also as shown in [4, p. 119] we have  $\varrho_n \rightarrow \varrho$  as  $n \rightarrow \infty$ . Let  $\{k_n\}_{n \in \mathbb{N}} \subset \mathbb{N}$  be an arbitrary sequence. Then there exists some subsequence  $\{k'_n\} \subset \{k_n\}$  such that  $\widehat{\varphi}_{k'_n} \rightarrow \widehat{\varphi} \in \mathcal{C}_{\text{loc}}^{1,2}(\mathbb{R}_T^d) \cap \mathcal{C}(\overline{\mathbb{R}_T^d})$ , for all  $T > 0$ , and  $\widehat{\varphi}$  satisfies

$$(3.24) \quad \begin{aligned} \partial_t \widehat{\varphi}(t, x) &= \min_{u \in \mathbb{U}} [L^u \widehat{\varphi}(t, x) + r(x, u)] \quad \text{in } (0, \infty) \times \mathbb{R}^d \\ \widehat{\varphi}(0, x) &= \varphi_0(x) \quad \forall x \in \mathbb{R}^d. \end{aligned}$$

Let  $\hat{v}^t$  denote a stationary Markov control associated with the minimizer in (3.24) as in Definition 3.7. By using Dynkin's formula on the cylinder  $[0, t] \times B_R$ , we obtain from (3.24)

$$(3.25a) \quad \hat{\varphi}(t, x) = \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^{\tau_R \wedge t} r(X_s, U_s) ds + \hat{\varphi}(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right],$$

$$(3.25b) \quad \hat{\varphi}(t, x) = \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^{\tau_R \wedge t} r(X_s, \hat{v}_s^t(X_s)) ds + \hat{\varphi}(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right].$$

Since  $\hat{\varphi}(t, \cdot)$  is nonnegative, letting  $R \rightarrow \infty$  in (3.25b), by Fatou's Lemma we obtain

$$(3.26) \quad \begin{aligned} \hat{\varphi}(t, x) &\geq \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^t r(X_s, \hat{v}_s^t(X_s)) ds + \varphi_0(X_t) \right] \\ &\geq \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) ds + \varphi_0(X_t) \right]. \end{aligned}$$

Taking limits as  $n \rightarrow \infty$  in (3.22), using monotone convergence for the first term on the right hand side, we obtain

$$(3.27) \quad \hat{\varphi}(t, x) \leq \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) ds + \varphi_0(X_t) \right] \quad \forall U \in \mathfrak{U}.$$

By (3.26)–(3.27) we have

$$(3.28a) \quad \hat{\varphi}(t, x) = \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) ds + \varphi_0(X_t) \right]$$

$$(3.28b) \quad \hat{\varphi}(t, x) = \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^t r(X_s, \hat{v}_s^t(X_s)) ds + \varphi_0(X_t) \right].$$

Let  $\bar{\varphi}(t, x) \triangleq \hat{\varphi}(t, x) - \rho t$ . Then  $\bar{\varphi}$  solves (3.13) and (3.14a)–(3.14b) follow by (3.28a)–(3.28b). It is also clear that  $\bar{\varphi}(t, x) \geq -\rho t$ , which together with (3.23) implies (3.16).

By (3.25a) we have

$$(3.29) \quad \begin{aligned} \hat{\varphi}(t, x) &= \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^{\tau_R \wedge t} r(X_s, \hat{v}_s^t(X_s)) ds + \varphi_0(t, X_t) \mathbb{I}\{\tau_R \geq t\} \right] \\ &\quad + \mathbb{E}_x^{\hat{v}^t} \left[ \hat{\varphi}(t - \tau_R, X_{\tau_R}) \mathbb{I}\{\tau_R < t\} \right]. \end{aligned}$$

The first term on the right hand side of (3.29) tends to the right hand side of (3.28b) by monotone convergence as  $R \uparrow \infty$ . Therefore (3.15) holds.

Suppose  $\tilde{\varphi}$  is a solution of (3.24) in  $\mathcal{C}_{\text{loc}}^{1,2}(\mathbb{R}_T^d) \cap \mathcal{C}(\overline{\mathbb{R}_T^d})$ , for some  $T > 0$ , which is bounded below, and  $\tilde{v}^t$  is an associated stationary Markov control from the minimizer of (3.24). Applying Dynkin's formula on the cylinder  $[0, t] \times B_R$  and letting  $R \rightarrow \infty$

using Fatou's lemma, we obtain

$$\begin{aligned}\tilde{\varphi}(t, x) &\geq \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^t r(X_s, \hat{v}_s^t(X_s)) ds + \varphi_0(X_t) \right] \\ &\geq \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) ds + \varphi_0(X_t) \right] \\ &\geq \hat{\varphi}(t, x).\end{aligned}$$

Therefore  $\bar{\varphi}(t, x)$  is the minimal solution of (3.13) in  $\mathcal{C}_{\text{loc}}^{1,2}((0, \infty) \times \mathbb{R}^d) \cap \mathcal{C}([0, \infty) \times \mathbb{R}^d)$  which is bounded below on  $\mathbb{R}_T^d$ , for each  $T > 0$ .  $\square$

In the interest of economy of language we refer to the solution in (3.14a) as canonical. This is detailed in the following definition.

**DEFINITION 3.10.** *Given an initial condition  $\varphi_0 \in \mathcal{O}_{V^*}$  we define the canonical solution to the VI in (3.13) as the solution which was constructed in the proof of Theorem 3.9 and was shown to admit the stochastic representation in (3.14a). In other words, this is the minimal solution of (3.13) in  $\mathcal{C}_{\text{loc}}^{1,2}((0, \infty) \times \mathbb{R}^d) \cap \mathcal{C}([0, \infty) \times \mathbb{R}^d)$  which is bounded below on  $\mathbb{R}_T^d$ , for any  $T > 0$ . The canonical solution to the VI well defines the canonical solution to the RVI in (3.12) via (1.14).*

For the rest of the paper a solution to the RVI or VI is always meant to be a canonical solution. In summary, these are characterized by:

$$\begin{aligned}(3.30) \quad \varphi(t, x) + \int_0^t \varphi(s, 0) ds &= \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) ds + \varphi_0(X_t) \right] \\ &= \int_0^t \mathbb{E}_x^{\hat{v}^t} [r(X_s, \hat{v}_s^t(X_s))] ds + \mathbb{E}_x^{\hat{v}^t} [\varphi_0(X_t)].\end{aligned}$$

Similarly

$$\begin{aligned}\bar{\varphi}(t, x) &= \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t \bar{r}(X_s, U_s) ds + \varphi_0(X_t) \right] \\ &= \int_0^t \mathbb{E}_x^{\hat{v}^t} [\bar{r}(X_s, \hat{v}_s^t(X_s))] ds + \mathbb{E}_x^{\hat{v}^t} [\varphi_0(X_t)].\end{aligned}$$

The next lemma provides an important estimate for the canonical solutions of the the VI.

**LEMMA 3.11.** *Provided  $\varphi_0 \in \mathcal{C}_{V^*}(\mathbb{R}^d) \cap \mathcal{C}^2(\mathbb{R}^d)$ , then the canonical solution  $\bar{\varphi} \in \mathcal{C}_{\text{loc}}^{1,2}((0, \infty) \times \mathbb{R}^d) \cap \mathcal{C}([0, \infty) \times \mathbb{R}^d)$  of (3.13) satisfies the bound*

$$(3.31) \quad \mathbb{E}_x^{\hat{v}^t} [\varphi_0(X_t) - V^*(X_t)] \leq \bar{\varphi}(t, x) - V^*(x) \leq \mathbb{E}_x^{v^*} [\varphi_0(X_t) - V^*(X_t)]$$

for all  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ .

*Proof.* By (1.5) and (3.13) we obtain

$$-\partial_t(V^* - \bar{\varphi}) + L^{v^*}(V^* - \bar{\varphi}) \leq 0$$

and

$$-\partial_t(V^* - \bar{\varphi}) + L^{\hat{v}^t}(V^* - \bar{\varphi}) \geq 0$$

from which, by an application of Itô's formula to  $V^*(X_s) - \bar{\varphi}(t-s, X_s)$ ,  $s \in [0, t]$ , it follows that

$$\mathbb{E}_x^{v^*} [V^*(X_t) - \varphi_0(X_t)] \leq V^*(x) - \bar{\varphi}(t, x)$$

and

$$\mathbb{E}_x^{\hat{v}^t} [V^*(X_t) - \varphi_0(X_t)] \geq V^*(x) - \bar{\varphi}(t, x),$$

respectively, and the estimate follows.  $\square$

Concerning the uniqueness of the canonical solution in a larger class of functions, this depends on the growth of  $V^*$  and the coefficients of the SDE in (1.1). Various such uniqueness results can be given based on different hypotheses on the growth of the data. The following result assumes that  $V^*$  has polynomial growth, which is the case in many applications.

**THEOREM 3.12.** *Let  $\varphi_0 \in \mathcal{O}_{V^*}$  and suppose that for some constants  $c_1, c_2$  and  $m > 0$ ,  $V^*(x) \leq c_1 + c_2|x|^m$ . Then any solution  $\bar{\varphi}' \in \mathcal{C}_{\text{loc}}^{1,2}(\mathbb{R}_T^d) \cap \mathcal{C}(\overline{\mathbb{R}_T^d})$  of (3.13), for some  $T > 0$ , which is bounded below in  $\overline{\mathbb{R}_T^d}$  and satisfies  $\|\bar{\varphi}'\|_{V^*, T} < \infty$  agrees with the canonical solution  $\bar{\varphi}$  on  $\overline{\mathbb{R}_T^d}$ .*

*Proof.* Let  $\bar{\varphi}'$  be a solution satisfying the hypothesis in the theorem, and let  $\bar{\varphi}$  be the canonical solution of (3.13) and  $\hat{v}^t$  the associated Markov control as in Definition 3.7. Let  $\bar{\varphi}_\varepsilon$ , for  $\varepsilon > 0$ , denote the canonical solution of (3.13) with initial data  $\varphi_0 + \varepsilon V^*$  and  $\hat{v}_\varepsilon$  the associated minimizer. By Theorem 3.9 for each  $\varepsilon > 0$  we obtain

$$\begin{aligned} \bar{\varphi}_\varepsilon(t, x) &= \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t \bar{r}(X_s, U_s) ds + \varphi_0(X_t) + \varepsilon V^*(X_t) \right] \\ &\geq -\rho t + \varepsilon \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t \bar{r}(X_s, U_s) ds + V^*(X_t) \right] \\ &\geq \varepsilon V^*(x) - \rho t. \end{aligned}$$

Therefore by (3.15) for each  $\varepsilon > 0$ , we have

$$\mathbb{E}_x^{\hat{v}_\varepsilon^t} [V^*(t - \tau_R, X_{\tau_R}) \mathbb{I}\{\tau_R < t\}] \xrightarrow{R \rightarrow \infty} 0 \quad \forall (t, x) \in \mathbb{R}_T^d,$$

which in turn implies, since  $\|\bar{\varphi}'\|_{V^*, T} < \infty$ , that

$$(3.34) \quad \mathbb{E}_x^{\hat{v}_\varepsilon^t} [\bar{\varphi}'(t - \tau_R, X_{\tau_R}) \mathbb{I}\{\tau_R < t\}] \xrightarrow{R \rightarrow \infty} 0 \quad \forall (t, x) \in \mathbb{R}_T^d.$$

Since  $-\partial_t \bar{\varphi}' + L^{\hat{v}_\varepsilon^t} \bar{\varphi}' + \bar{r}(x, \hat{v}_{\varepsilon, t}(x)) \geq 0$ , we have that for all  $(t, x) \in \mathbb{R}_T^d$ ,

$$(3.35) \quad \bar{\varphi}'(t, x) \leq \mathbb{E}_x^{\hat{v}_\varepsilon^t} \left[ \int_0^{\tau_R \wedge t} \bar{r}(X_s, \hat{v}_{\varepsilon, s}^t(X_s)) ds + \bar{\varphi}'(t - \tau_R \wedge t, X_{\tau_R \wedge t}) \right],$$



and taking limits as  $R \rightarrow \infty$  in (3.35), using (3.34), it follows that  $\bar{\varphi}' \leq \bar{\varphi}_\varepsilon$  on  $\overline{\mathbb{R}_T^d}$ .

The polynomial growth of  $V^*$  implies that there exists a constant  $m(x, T)$  such that  $\mathbb{E}_x^U[V^*(X_t)] \leq m(x, T)$  for all  $(t, x) \in \mathbb{R}_T^d$  and  $U \in \mathfrak{U}$  [4, Theorem 2.2.2]. Therefore, since

$$(3.36) \quad \begin{aligned} \bar{\varphi}_\varepsilon(t, x) &\leq \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^t \bar{r}(X_s, \hat{v}_s^t(X_s)) ds + \varphi_0(X_t) + \varepsilon V^*(X_t) \right] \\ &\leq \bar{\varphi}(t, x) + \varepsilon m(x, T) \quad \forall (t, x) \in \mathbb{R}_T^d, \end{aligned}$$

and  $\bar{\varphi}_\varepsilon \geq \bar{\varphi}$ , it follows by (3.36) that  $\bar{\varphi}_\varepsilon \rightarrow \bar{\varphi}$  on  $\overline{\mathbb{R}_T^d}$  as  $\varepsilon \downarrow 0$ . Thus  $\bar{\varphi}' \leq \bar{\varphi}$  on  $\overline{\mathbb{R}_T^d}$ , and by the minimality of  $\bar{\varphi}$  we must have equality.  $\square$

We can also obtain a uniqueness result on a larger class of functions that does not require  $V^*$  to have polynomial growth, but assumes that the diffusion matrix is bounded in  $\mathbb{R}^d$ . This is given in Theorem 3.13 below, whose proof uses the technique in [8].

We define the following class of functions:

$$\mathfrak{G} \triangleq \{f \in \mathcal{C}^2(\mathbb{R}^d) : \lim_{|x| \rightarrow \infty} f(x) e^{-k|x|^2} = 0, \text{ for some } k > 0\}.$$

**THEOREM 3.13.** *Suppose  $V^* \in \mathfrak{G}$  and that  $\|\sigma\|$  is bounded in  $\mathbb{R}^d$ . Then, provided  $\varphi_0 \in \mathcal{O}_{V^*}$ , there exists a unique solution  $\psi$  to (3.13) such that  $\max_{t \in [0, T]} \psi(t, \cdot) \in \mathfrak{G}$  for each  $T > 0$ .*

*Proof.* Let  $\hat{\varphi} \in \mathcal{C}_{\text{loc}}^{1,2}((0, \infty) \times \mathbb{R}^d) \cap \mathcal{C}([0, \infty) \times \mathbb{R}^d)$  be the minimal nonnegative solution of

$$(3.37) \quad \begin{aligned} \partial_t \hat{\varphi}(t, x) &= \min_{u \in \mathbb{U}} [L^u \hat{\varphi}(t, x) + r(x, u)] \quad \text{in } (0, \infty) \times \mathbb{R}^d, \\ \hat{\varphi}(0, x) &= \varphi_0(x) \quad \forall x \in \mathbb{R}^d, \end{aligned}$$

and let  $\{\hat{v}_t, t \in \mathbb{R}_+\}$  denote a measurable selector from the minimizer in (3.37). Suppose that  $\tilde{\varphi} \in \mathcal{C}_{\text{loc}}^{1,2}((0, \infty) \times \mathbb{R}^d) \cap \mathcal{C}([0, \infty) \times \mathbb{R}^d)$  is any solution of (3.37) satisfying the hypothesis of the theorem, and let  $\{\tilde{v}_t, t \in \mathbb{R}_+\}$  denote a measurable selector from the corresponding minimizer. Then  $f \triangleq \tilde{\varphi} - \hat{\varphi}$  satisfies, for any  $T > 0$ ,

$$(3.38) \quad \partial_t f - L^{\hat{v}^T} f \leq 0 \quad \text{and} \quad \partial_t f - L^{\tilde{v}^T} f \geq 0 \quad \text{in } (0, T] \times \mathbb{R}^d,$$

and  $f(0, x) = 0$  for all  $x \in \mathbb{R}^d$ . By (3.16), the hypothesis that  $V^* \in \mathfrak{G}$ , and the hypothesis on the growth of  $f$ , it follows that for some  $k = k(T) > 0$  large enough

$$(3.39) \quad \lim_{|x| \rightarrow \infty} \max_{t \in [0, T]} |f(t, x)| e^{-k|x|^2} = 0.$$

It is straightforward to verify by direct computation using the bounds on the coefficients of the SDE that there exists  $\gamma = \gamma(k) > 1$  such that  $g(t, x) \triangleq e^{(1+\gamma t)(1+k|x|^2)}$  is a supersolution of

$$(3.40) \quad \partial_t g - L^{\hat{v}^{T_0}} g \geq 0 \quad \text{in } (0, T_0] \times \mathbb{R}^d, \quad \text{with } T_0 \equiv \gamma^{-1}.$$

By (3.39), for any  $\varepsilon > 0$  we can select  $R > 0$  large enough such that  $|f(t, x)| \leq \varepsilon g(t, x)$  for all  $(t, x) \in [0, \gamma^{-1}] \times \partial B_R$ . Using (3.38), (3.40) and Dynkin's formula on the strip  $[0, \gamma^{-1}] \times \overline{B}_R$  it follows that  $|f(t, x)| \leq \varepsilon g(t, x)$  for all  $(t, x) \in [0, \gamma^{-1}] \times \overline{B}_R$ . Since  $\varepsilon > 0$  was arbitrary this implies  $f \equiv 0$ , or equivalently that  $f = \widehat{\varphi}$  on  $[0, \gamma^{-1}] \times \mathbb{R}^d$ .

Since, by (3.16),  $\widehat{\varphi}(\gamma^{-1}, \cdot) \in \mathcal{O}_{V^*}$ , we can repeat the argument to show that  $f = \widehat{\varphi}$  on  $[\gamma^{-1}, 2\gamma^{-1}] \times \mathbb{R}^d$ , and that the same holds by induction on  $[n\gamma^{-1}, (n+1)\gamma^{-1}] \times \mathbb{R}^d$ ,  $n = 2, 3, \dots$ , until we cover the interval  $[0, T]$ . This shows that  $f = \widehat{\varphi}$  on  $\overline{\mathbb{R}_T^d}$ , and since  $T > 0$  was arbitrary the same holds on  $[0, \infty) \times \mathbb{R}^d$ .  $\square$

We do not enforce any of the assumptions of Theorem 3.13 in the rest of the paper. Rather our analysis is based on the canonical solution to the VI and RVI which is well defined (see Definition 3.10).

**3.4. A region of attraction for the VI algorithm.** In this section we describe a region of attraction for the VI algorithm. This is an subset of  $\mathcal{C}^2(\mathbb{R}^d)$  which is invariant under the semiflow defined by (3.13) and all its points are convergent, i.e., converge to a solution of (1.5).

DEFINITION 3.14. *We let  $\overline{\Phi}_t[\varphi_0] : \mathcal{C}^2(\mathbb{R}^d) \rightarrow \mathcal{C}^2(\mathbb{R}^d)$ ,  $t \in [0, \infty)$ , denote the canonical solution (semiflow) of the VI in (3.13) starting from  $\varphi_0$ , and  $\Phi_t[\varphi_0]$  denote the corresponding canonical solution (semiflow) of the RVI in (3.12). Let  $\mathcal{E}$  denote the set of solutions of the HJB in (1.5), i.e.,*

$$\mathcal{E} \triangleq \{V^* + c : c \in \mathbb{R}\}.$$

Also for  $c \in \mathbb{R}$  we define the set  $\mathcal{G}_c \subset \mathcal{C}^2(\mathbb{R}^d)$  by

$$\mathcal{G}_c \triangleq \{h \in \mathcal{C}^2(\mathbb{R}^d) : h - V^* \geq c, \|h\|_{V^*} < \infty\}.$$

We claim that for each  $c \in \mathbb{R}$ ,  $\mathcal{G}_c$  is invariant under the flow  $\overline{\Phi}_t$ . Indeed by (3.7) and (3.31), if  $\varphi_0 \in \mathcal{G}_c$ , then we have that

$$\begin{aligned} c \leq \overline{\Phi}_t[\varphi_0](x) - V^*(x) &\leq \mathbb{E}_x^{v^*} [\varphi_0(X_t) - V^*(X_t)] \\ &\leq m_r \|\varphi_0 - V^*\|_{V^*} (V^*(x) + 1) \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d. \end{aligned}$$

Since translating  $\varphi_0$  by a constant simply translates the orbit  $\overline{\Phi}_t[\varphi_0]$ , without loss of generality we let  $c = 0$ , and we show that all the points of  $\mathcal{G}_0$  are convergent.

THEOREM 3.15. *Under Assumption 3.2, for each  $\varphi_0 \in \mathcal{G}_0$  the orbit  $\overline{\Phi}_t[\varphi_0]$ , and therefore also  $\Phi_t[\varphi_0]$ , converges as  $t \rightarrow \infty$  to a point in  $\mathcal{E} \cap \mathcal{G}_0$ .*

*Proof.* Since, as we showed in the paragraph preceding the theorem,  $\overline{\Phi}_t[\varphi_0] \in \mathcal{G}_0$  for all  $t \geq 0$ , by (3.14a) we have

$$(3.41) \quad \overline{\Phi}_t[\varphi_0](x) \leq \mathbb{E}_x^{v^*} \left[ \int_0^{t-\tau} \overline{r}(X_s, v^*(X_s)) ds + \overline{\Phi}_\tau[\varphi_0](X_{t-\tau}) \right] \quad \forall \tau \in [0, t],$$

Since  $\bar{\Phi}_t[\varphi_0](x) - V^*(x) \geq 0$ , and  $\int_{\mathbb{R}^d} \bar{\Phi}_t[\varphi_0](x) \mu_{v^*}(dx)$  is finite by Assumption 3.2, it follows by integrating (3.41) with respect to  $\mu_{v^*}$  that the map

$$(3.42) \quad t \mapsto \int_{\mathbb{R}^d} \bar{\Phi}_t[\varphi_0](x) \mu_{v^*}(dx)$$

is nonincreasing and bounded below. Hence it must be constant on the  $\omega$ -limit set of  $\varphi_0$  denoted by  $\omega(\varphi_0)$ . Let  $h \in \omega(\varphi_0)$  and define

$$(3.43) \quad f(t, x) \triangleq -\partial_t \bar{\Phi}_t[h](x) + L^{v^*}(\bar{\Phi}_t[h](x) - V^*(x)).$$

Then  $f(t, x) \geq 0$  for all  $(t, x)$ , and by applying Itô's formula to (3.43), we obtain

$$(3.44) \quad \bar{\Phi}_t[h](x) - V^*(x) - \mathbb{E}_x^{v^*}[h(X_t) - V^*(X_t)] = -\mathbb{E}_x^{v^*} \left[ \int_0^t f(t-s, X_s) ds \right].$$

Integrating (3.44) with respect to the invariant distribution  $\mu_{v^*}$  we obtain

$$(3.45) \quad \int_{\mathbb{R}^d} (\bar{\Phi}_t[h](x) - h(x)) \mu_{v^*}(dx) = - \int_0^t \int_{\mathbb{R}^d} f(t-s, x) \mu_{v^*}(dx) ds \quad \forall t \geq 0.$$

Since the term on the left-hand-side of (3.45) equals 0, as we argued above, it follows that  $f(t, x) = 0$ ,  $(t, x) - a.e.$ , which in turn implies that

$$\lim_{t \rightarrow \infty} \bar{\Phi}_t[h](x) = V^*(x) - \int_{\mathbb{R}^d} (V^*(x) - h(x)) \mu_{v^*}(dx).$$

It follows that  $\omega(\varphi_0) \subset \mathcal{E} \cap \mathcal{G}_0$  and since the map in (3.42) is nonincreasing, it is straightforward to verify that  $\omega(\varphi_0)$  must be a singleton.  $\square$

We also have the following result which does not require Assumption 3.2.

**COROLLARY 3.16.** *Suppose  $\varphi_0 \in \mathcal{C}^2(\mathbb{R}^d)$  is such that  $\varphi_0 - V^*$  is bounded. Then  $\bar{\Phi}_t[\varphi_0]$  converges as  $t \rightarrow \infty$  to a point in  $\mathcal{E}$ .*

*Proof.* By (3.31), under the hypothesis,  $x \mapsto \bar{\varphi}(t, x) - V^*(x)$  is bounded uniformly in  $t$ . Thus the result follows as in the proof of Theorem 3.15.  $\square$

**4. Growth Estimates for Solutions of the Value Iteration.** Most of the results of this section do not require Assumption 3.2. It is only needed for Lemma 4.10 and Corollary 4.11. Throughout this section and also in §5 a solution  $\bar{\varphi}(\varphi)$  always refers to the canonical solution of the VI (RVI) without further mention (see Definition 3.10).

**LEMMA 4.1.** *Suppose  $\varphi_0 \in \mathcal{O}_{V^*}$ . Then*

$$\frac{1}{t} \bar{\varphi}(t, x) \xrightarrow{t \rightarrow \infty} 0.$$

*Proof.* Since  $\|\varphi_0\|_{V^*} < \infty$  it follows that  $\frac{1}{t} \mathbb{E}_x^{v^*} [\varphi_0(X_t)] \rightarrow 0$  as  $t \rightarrow \infty$  (see [4, Lemma 3.7.2 (ii)]), and so we have

$$\begin{aligned} 0 &\leq \liminf_{t \rightarrow \infty} \frac{1}{t} \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t \bar{r}(X_s, U_s) ds + \varphi_0(X_t) \right] \\ &= \liminf_{t \rightarrow \infty} \frac{\bar{\varphi}(t, x)}{t} \leq \limsup_{t \rightarrow \infty} \frac{\bar{\varphi}(t, x)}{t} \\ &\leq \limsup_{t \rightarrow \infty} \frac{1}{t} \mathbb{E}_x^{v^*} \left[ \int_0^t \bar{r}(X_s, v^*(X_s)) ds + \varphi_0(X_t) \right] \\ &= 0. \end{aligned}$$

The first inequality above uses the fact that  $\varphi_0$  is bounded below and that  $\varrho$  is the optimal ergodic cost.  $\square$

LEMMA 4.2. *Provided  $\|\varphi_0\|_\infty < \infty$ , it holds that for all  $t \geq 0$*

$$\bar{\varphi}(t - \tau, x) - \bar{\varphi}(t, x) \leq \varrho \tau + \operatorname{osc}_{\mathbb{R}^d} \varphi_0 \quad \forall x \in \mathbb{R}^d, \quad \forall \tau \in [0, t].$$

*Proof.* We have

$$\begin{aligned} \bar{\varphi}(t - \tau, x) - \bar{\varphi}(t, x) &= \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^{t-\tau} \bar{r}(X_s, U_s) ds + \varphi_0(X_{t-\tau}) \right] \\ &\quad - \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t \bar{r}(X_s, U_s) ds + \varphi_0(X_t) \right] \\ &\leq - \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \varphi_0(X_t) - \varphi_0(X_{t-\tau}) + \int_{t-\tau}^t \bar{r}(X_s, U_s) ds \right] \\ &\leq \varrho \tau + \operatorname{osc}_{\mathbb{R}^d} \varphi_0. \quad \square \end{aligned}$$

DEFINITION 4.3. *We define:*

$$\mathcal{K} \triangleq \{x \in \mathbb{R}^d : \min_{u \in \mathfrak{U}} r(x, u) \leq \varrho\}.$$

Let  $B_0$  be some open bounded ball containing  $\mathcal{K}$  and define  $\check{\tau} \triangleq \tau(B_0^c)$ . Also let  $\delta_0 > 0$  be such that  $r(x, u) \geq \varrho + \delta_0$  on  $B_0^c$ .

LEMMA 4.4. *Suppose  $\varphi_0 \in \mathcal{O}_{V^*}$ . Then it holds that*

$$(4.1) \quad \bar{\varphi}(t, x) \leq \mathbb{E}_x^{v^*} \left[ \int_0^{\check{\tau} \wedge t} \bar{r}(X_s, v^*(X_s)) ds + \bar{\varphi}(t - \check{\tau} \wedge t, X_{\check{\tau} \wedge t}) \right],$$

and

$$(4.2) \quad \bar{\varphi}(t, x) \geq \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^{\check{\tau} \wedge t} \bar{r}(X_s, \hat{v}_s^t(X_s)) ds + \bar{\varphi}(t - \check{\tau} \wedge t, X_{\check{\tau} \wedge t}) \right]$$

for all  $x \in B_0^c$ .

*Proof.* Let  $B_R$  be any ball that contains  $\overline{B_0}$  and for  $n \in \mathbb{N}$ , let  $\tau_n$  denote the first exit time from  $B_{nR}$ . Using Dynkin's formula on (3.13), we obtain

$$(4.3) \quad \overline{\varphi}(t, x) = \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^{\check{\tau} \wedge \tau_n \wedge t} \overline{r}(X_s, U_s) ds + \overline{\varphi}(t - \check{\tau} \wedge \tau_n \wedge t, X_{\check{\tau} \wedge \tau_n \wedge t}) \right]$$

for  $x \in B_0^c$ . By (4.3) we have

$$(4.4) \quad \begin{aligned} \overline{\varphi}(t, x) \leq \mathbb{E}_x^{v^*} \left[ \int_0^{\check{\tau} \wedge \tau_n \wedge t} r(X_s, v^*(X_s)) ds \right] &- \varrho \mathbb{E}_x^{v^*} [\check{\tau} \wedge \tau_n \wedge t] \\ &+ \mathbb{E}_x^{v^*} [\overline{\varphi}(t - \check{\tau} \wedge \tau_n \wedge t, X_{\check{\tau} \wedge \tau_n \wedge t})]. \end{aligned}$$

We use the expansion

$$\begin{aligned} \mathbb{E}_x^{v^*} [\overline{\varphi}(t - \check{\tau} \wedge \tau_n \wedge t, X_{\check{\tau} \wedge \tau_n \wedge t})] &= \mathbb{E}_x^{v^*} [\overline{\varphi}(t - \check{\tau} \wedge t, X_{\check{\tau} \wedge t}) \mathbb{I}\{\tau_n > \check{\tau} \wedge t\}] \\ &+ \mathbb{E}_x^{v^*} [\overline{\varphi}(t - \tau_n, X_{\tau_n}) \mathbb{I}\{\tau_n \leq \check{\tau} \wedge t\}]. \end{aligned}$$

By (3.16) and the fact that, as shown in [4, Corollary 3.7.3],

$$\mathbb{E}_x^{v^*} [V^*(X_{\tau_n}) \mathbb{I}\{\tau_n \leq t\}] \xrightarrow{n \rightarrow \infty} 0,$$

we obtain

$$\mathbb{E}_x^{v^*} [\overline{\varphi}(t - \tau_n, X_{\tau_n}) \mathbb{I}\{\tau_n \leq \check{\tau} \wedge t\}] \xrightarrow{n \rightarrow \infty} 0.$$

Therefore by taking limits as  $n \rightarrow \infty$  in (4.4) and also using monotone convergence for the first two terms on the r.h.s., we obtain (4.1).

To obtain a lower bound we start from

$$(4.5) \quad \overline{\varphi}(t, x) = \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^{\check{\tau} \wedge \tau_n \wedge t} \overline{r}(X_s, \hat{v}_s^t(X_s)) ds + \overline{\varphi}(t - \check{\tau} \wedge \tau_n \wedge t, X_{\check{\tau} \wedge \tau_n \wedge t}) \right].$$

Since for any fixed  $t$  the functions  $\{\overline{\varphi}(t - s, x) : s \leq t\}$  are uniformly bounded below, taking limits in (4.5) as  $n \rightarrow \infty$ , we obtain (4.2).  $\square$

LEMMA 4.5. *Suppose  $\varphi_0 \in \mathcal{O}_{V^*}$ . Then for any  $t > 0$  we have*

$$\overline{\varphi}(t, x) > \min \left( \min_{[0, t] \times \overline{B_0}} \overline{\varphi}, \min_{\mathbb{R}^d} \varphi_0 \right) \quad \forall x \in B_0^c.$$

*Proof.* Let  $x$  be any point in the interior of  $B_0^c$ . By (4.2) we have

$$\begin{aligned} \overline{\varphi}(t, x) &\geq \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^{\check{\tau} \wedge t} \overline{r}(X_s, \hat{v}_s^t(X_s)) ds + \overline{\varphi}(t - \check{\tau} \wedge t, X_{\check{\tau} \wedge t}) \right] \\ &\geq \delta_0 \mathbb{E}_x^{\hat{v}^t} [\check{\tau} \wedge t] + \mathbb{P}_x^{\hat{v}^t}(\check{\tau} \leq t) \min_{[0, t] \times \overline{B_0}} \overline{\varphi} + \mathbb{P}_x^{\hat{v}^t}(\check{\tau} > t) \min_{\mathbb{R}^d} \varphi_0, \end{aligned}$$

and the result follows.  $\square$

*Remark 4.6.* By Lemma 4.5, if  $\inf_{[0,\infty)\times\overline{B}_0} \overline{\varphi} > -\infty$ , then  $\overline{\varphi}$  is bounded below on  $[0,\infty)\times\mathbb{R}^d$ . If this the case, the convergence of the VI and therefore also of the RVI follows as in the proof of Theorem 3.15. Therefore without loss of generality we assume for the remainder of the paper that  $\inf_{[0,\infty)\times\overline{B}_0} \overline{\varphi} = -\infty$ . By Lemma 4.5, this implies that there exists  $T_0 > 0$  such that

$$\overline{\varphi}(t, x) > \min_{[0,t]\times\overline{B}_0} \overline{\varphi} \quad \forall t \geq T_0, \forall x \in B_0^c.$$

We use the parabolic Harnack inequality which we quote in simplified form from the more general result in [9, Theorem 4.1] as follows:

**THEOREM 4.7 (Parabolic Harnack).** *Let  $B_{2R} \subset \mathbb{R}^d$  be an open ball, and  $\psi$  be a nonnegative caloric function, i.e., a nonnegative solution of*

$$\partial_t \psi(t, x) - a^{ij}(t, x) \partial_{ij} \psi(t, x) + b^i(t, x) \partial_i \psi(t, x) = 0 \quad \text{on } [0, T] \times B_{2R},$$

with  $a^{ij}(t, x)$  continuous in  $x$  and uniformly nondegenerate on  $[0, T] \times B_{2R}$ , and  $a^{ij}$  and  $b^i$  bounded on  $[0, T] \times B_{2R}$ . Then for any  $\tau \in (0, T/4]$ , there exists a constant  $C_H$  depending only on  $R, \tau$ , and the ellipticity constant (and modulus of continuity) of  $a^{ij}$  and the bound of  $a^{ij}$  and  $b^i$  on  $B_{2R}$ , such that

$$\max_{[T-3\tau, T-2\tau] \times B_R} \psi \leq C_H \min_{[T-\tau, T] \times B_R} \psi.$$

In the three lemmas that follow we apply Theorem 4.7 with  $\tau \equiv 1$  and  $B'_0 = 2B_0$ .

**LEMMA 4.8.** *There exists a constant  $M_0$  such that*

$$\max_{[T-3, T-2] \times B_0} \overline{\varphi} - \min_{[0, T] \times B_0} \overline{\varphi} \leq M_0 + C_H \left( \min_{[T-1, T] \times B_0} \overline{\varphi} - \min_{[0, T] \times B_0} \overline{\varphi} \right) \quad \forall T \geq T_0 + 4.$$

*Proof.* Let  $\psi_T(t, x)$  be the unique solution in  $\mathcal{W}_{\text{loc}}^{1,2,p}((0, T) \times B'_0) \cap \mathcal{C}([0, T] \times \overline{B'_0})$  of

$$\begin{aligned} \partial_t \psi_T(t, x) - a^{ij}(x) \partial_{ij} \psi_T(t, x) - b^i(x, \hat{v}_t^T(x)) \partial_i \psi_T(t, x) &= 0 \quad \text{on } [0, T] \times B'_0, \\ \psi_T(t, x) &= \overline{\varphi}(t, x) \quad \text{on } ([0, T] \times \partial B'_0) \cup (\{0\} \times \overline{B'_0}). \end{aligned}$$

Since  $\overline{\psi}_T \triangleq \psi_T - \overline{\varphi}$  satisfies

$$\partial_t \overline{\psi}_T(t, x) - a^{ij}(x) \partial_{ij} \overline{\psi}_T(t, x) - b^i(x, \hat{v}_t^T(x)) \partial_i \overline{\psi}_T(t, x) + \bar{r}(x, \hat{v}_t^T(x)) = 0$$

on  $[0, T] \times B'_0$ , and

$$\overline{\psi}_T(t, x) = 0 \quad \text{on } ([0, T] \times \partial B'_0) \cup (\{0\} \times \overline{B'_0}),$$

it follows that there exists a constant  $\overline{M}_0$  which depends only on  $B'_0$  (it is independent of  $T$ ) such that

$$(4.6) \quad \sup_{[0, T] \times B'_0} |\overline{\psi}_T| \leq \overline{M}_0 \quad \forall T > 0.$$

Indeed this is so because with  $\tau(B'_0)$  denoting the first exit time from  $B'_0$  and with  $\hat{v}^T$  as defined in Definition 3.7 we have

$$\begin{aligned} |\overline{\psi}_T(t, x)| &= \left| \mathbb{E}_x^{\delta^T} \left[ \int_0^{(T-t) \wedge \tau(B'_0)} \overline{r}(X_s, \hat{v}_{T-t-s}^T(X_s)) ds \right] \right| \\ &\leq \sup_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^{\tau(B'_0)} |\overline{r}(X_s, U_s)| ds \right] \\ &\leq \|\overline{r}\|_{\infty, B'_0} \sup_{x \in B'_0} \sup_{U \in \mathfrak{U}} \mathbb{E}_x^U [\tau(B'_0)] < \infty, \end{aligned}$$

since the mean exit time from  $B'_0$  is upper bounded by a constant uniformly over all initial  $x \in B'_0$  and all controls  $U \in \mathfrak{U}$  by the weak maximum principle of Alexandroff.

Let  $(\hat{t}, \hat{x})$  be a point at which  $\overline{\varphi}$  attains its minimum on  $[T-1, T] \times B_0$ . By Lemma 4.5 and Remark 4.6 the function  $(t, x) \mapsto \psi_T(t, x) - \min_{[0, T] \times B_0} \overline{\varphi}$  is nonnegative on  $[T-4, T] \times B'_0$ , and by Theorem 4.7 we have

$$(4.7) \quad \begin{aligned} \psi_T(t, x) - \min_{[0, T] \times B_0} \overline{\varphi} &\leq C_H \left( \psi_T(\hat{t}, \hat{x}) - \min_{[0, T] \times B_0} \overline{\varphi} \right) \\ &\leq C_H \left( \overline{\psi}_T(\hat{t}, \hat{x}) + \min_{[T-1, T] \times B_0} \overline{\varphi} - \min_{[0, T] \times B_0} \overline{\varphi} \right) \end{aligned}$$

for all  $t \in [T-3, T-2]$  and  $x \in B_0$ . Expressing the left hand side of (4.7) as

$$\overline{\varphi}(t, x) - \min_{[0, T] \times B_0} \overline{\varphi} + \overline{\psi}_T(t, x),$$

and using (4.6), Lemma 4.8 follows with

$$M_0 \triangleq (C_H + 1) \overline{M}_0. \quad \square$$

LEMMA 4.9. *Provided  $\varphi_0 \in \mathcal{C}^2(\mathbb{R}^d)$  is nonnegative and bounded, we have*

$$\overline{\varphi}(t, x) - \max_{\partial B_0} \overline{\varphi}(t, \cdot) \leq 2 \|\varphi_0\|_{\infty} + (1 + \varrho \delta_0^{-1}) V^*(x) \quad \forall x \in B_0^c.$$

*Proof.* By Lemma 4.2

$$(4.8) \quad \overline{\varphi}(t - \tau, x) \leq \overline{\varphi}(t, x) + \varrho \tau + \operatorname{osc}_{\mathbb{R}^d} \varphi_0 \quad \forall x \in B_0, \quad 0 \leq \tau \leq t.$$

Therefore by (4.1) and (4.8), using the fact that  $\bar{r} \geq 0$  on  $B_0^c$ , we obtain

$$\begin{aligned}
(4.9) \quad \bar{\varphi}(t, x) &\leq \mathbb{E}_x^{v^*} \left[ \int_0^{\bar{\tau} \wedge t} \bar{r}(X_s, v^*(X_s)) \, ds + \bar{\varphi}(t - \bar{\tau} \wedge t, X_{\bar{\tau} \wedge t}) \right] \\
&\leq \mathbb{E}_x^{v^*} \left[ \int_0^{\bar{\tau}} \bar{r}(X_s, v^*(X_s)) \, ds \right] + \mathbb{E}_x^{v^*} [\bar{\varphi}(t - \bar{\tau}, X_{\bar{\tau}}) \mathbb{I}\{\bar{\tau} \leq t\}] \\
&\quad + \mathbb{E}_x^{v^*} [\varphi_0(X_t) \mathbb{I}\{\bar{\tau} > t\}] \\
&\leq V^*(x) + \mathbb{E}_x^{v^*} [\bar{\varphi}(t, X_{\bar{\tau}}) \mathbb{I}\{\bar{\tau} \leq t\}] \\
&\quad + \varrho \mathbb{E}_x^{v^*} [\bar{\tau} \mathbb{I}\{\bar{\tau} \leq t\}] + \operatorname{osc}_{\mathbb{R}^d} \varphi_0 + \|\varphi_0\|_\infty \\
&\leq V^*(x) + \mathbb{P}_x^{v^*}(\{\bar{\tau} \leq t\}) \left( \max_{\partial B_0} \bar{\varphi}(t, \cdot) \right) \\
&\quad + \varrho \mathbb{E}_x^{v^*} [\bar{\tau} \mathbb{I}\{\bar{\tau} \leq t\}] + 2 \|\varphi_0\|_\infty,
\end{aligned}$$

for  $x \in B_0^c$ . Since  $-\bar{\varphi}(t, x) \leq \varrho t$ , we have

$$\begin{aligned}
(4.10) \quad -\mathbb{P}_x^{v^*}(\{\bar{\tau} > t\}) \left( \max_{\partial B_0} \bar{\varphi}(t, \cdot) \right) &\leq \varrho \mathbb{P}_x^{v^*}(\{\bar{\tau} > t\}) t \\
&\leq \varrho \mathbb{E}_x^{v^*} [\bar{\tau} \mathbb{I}\{\bar{\tau} > t\}].
\end{aligned}$$

Hence subtracting  $\max_{\partial B_0} \bar{\varphi}(t, \cdot)$  from both sides of (4.9) and using (4.10) together with the estimate  $\mathbb{E}_x^{v^*}[\bar{\tau}] \leq \delta_0^{-1} V^*(x)$ , we obtain

$$\bar{\varphi}(t, x) - \max_{\partial B_0} \bar{\varphi}(t, \cdot) \leq V^*(x) + \varrho \delta_0^{-1} V^*(x) + 2 \|\varphi_0\|_\infty. \quad \square$$

We define the set  $\mathcal{T} \subset \mathbb{R}_+$  by

$$\mathcal{T} \triangleq \left\{ t \geq T_0 + 4 : \min_{[t-1, t] \times B_0} \bar{\varphi} = \min_{[0, t] \times B_0} \bar{\varphi} \right\},$$

where  $T_0$  is as in Remark 4.6. By Remark 4.6,  $\mathcal{T} \neq \emptyset$ .

LEMMA 4.10. *Let Assumption 3.2 hold and suppose that the initial condition  $\varphi_0 \in \mathcal{C}^2(\mathbb{R}^d)$  is nonnegative and bounded. Then there exists a constant  $C_0$  such that*

$$\operatorname{osc}_{B_0} \bar{\varphi}(t, \cdot) \leq C_0 \quad \forall t \geq 0.$$

*Proof.* Suppose  $t \in \mathcal{T}$ . Then, by Lemma 4.8,

$$\max_{\partial B_0} \bar{\varphi}(t-2, \cdot) - \min_{[0, t] \times B_0} \bar{\varphi} \leq M_0.$$

Therefore, by Lemma 4.9 we have

$$(4.11) \quad \bar{\varphi}(t-2, x) - \min_{[0, t] \times B_0} \bar{\varphi} \leq M_0 + 2 \|\varphi_0\|_\infty + (1 + \varrho \delta_0^{-1}) V^*(x) \quad \forall (t, x) \in \mathcal{T} \times B_0^c.$$



Next, fix any  $t_0 \in \mathcal{T}$ . It suffices to prove the result for  $t \geq t_0$  since it trivially holds for  $t$  in the compact interval  $[0, t_0]$ . Given  $t \geq t_0$  let  $\tau \triangleq \sup \mathcal{T} \cap [t_0, t]$ . Note then that

$$(4.12) \quad \min_{[0, \tau] \times B_0} \bar{\varphi} = \min_{[0, t] \times B_0} \bar{\varphi}.$$

By (4.11)–(4.12), and since  $V^*$  is nonnegative, we obtain

$$(4.13) \quad \begin{aligned} \sup_{x \in B_0} \bar{\varphi}(t, x) &\leq \sup_{x \in B_0} \mathbb{E}_x^{v^*} \left[ \int_0^{t-\tau+2} \bar{r}(X_s, v^*(X_s)) \, ds + \bar{\varphi}(\tau - 2, X_{t-\tau+2}) \right] \\ &\leq \sup_{x \in B_0} \mathbb{E}_x^{v^*} \left[ \int_0^{t-\tau+2} \bar{r}(X_s, v^*(X_s)) \, ds + V^*(X_{t-\tau+2}) \right] \\ &\quad + \sup_{x \in B_0} \mathbb{E}_x^{v^*} [\bar{\varphi}(\tau - 2, X_{t-\tau+2})] \\ &\leq \|V^*\|_{\infty, B_0} + \min_{[0, t] \times B_0} \bar{\varphi} + M_0 + 2 \|\varphi_0\|_{\infty} + \varrho \delta_0^{-1} K_0, \end{aligned}$$

with

$$K_0 \triangleq \sup_{t \geq 0} \sup_{x \in B_0} \mathbb{E}_x^{v^*} [V^*(X_t)].$$

By Lemma 3.5,  $K_0$  is finite. Since

$$\text{osc}_{B_0} \bar{\varphi}(t, \cdot) \leq \sup_{x \in B_0} \bar{\varphi}(t, x) - \min_{[0, t] \times B_0} \bar{\varphi},$$

and  $t \geq t_0$  was arbitrary, the result follows for all  $t \geq t_0$  by (4.13).  $\square$

The following corollary now follows by Theorem 4.7 and Lemma 4.10.

**COROLLARY 4.11.** *Under the hypotheses of Lemma 4.10, for any  $\tau > 0$  there exists a constant  $\widehat{C}(\tau)$  such that*

$$\text{osc}_{[n\tau, (n+1)\tau] \times B_0} \bar{\varphi} \leq \widehat{C}(\tau) \quad \forall n \in \mathbb{N}.$$

**5. Convergence of the Relative Value Iteration.** We define the set  $\mathcal{T}_0 \subset \mathbb{R}_+$  by

$$\mathcal{T}_0 \triangleq \{t \in \mathbb{R}_+ : \bar{\varphi}(t, 0) \leq \bar{\varphi}(t', 0) \quad \forall t' \leq t\}.$$

In the next lemma we use the variable

$$\Psi(t, x) \triangleq \bar{\varphi}(t, x) - \bar{\varphi}(t, 0).$$

**LEMMA 5.1.** *Let Assumption 3.2 hold and also suppose that the initial condition  $\varphi_0 \in \mathcal{C}^2(\mathbb{R}^d)$  is nonnegative and bounded. Then*

$$(5.1a) \quad \Psi(t, x) \leq C_0 + 2 \|\varphi_0\|_{\infty} + (1 + \varrho \delta_0^{-1}) V^*(x) \quad \forall (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d,$$

and there exists a constant  $\widehat{M}_0$  such that

$$(5.1b) \quad \overline{\varphi}(t, 0) - \overline{\varphi}(t', 0) \leq \widehat{M}_0 \quad \forall t \geq t'.$$

*Proof.* The estimate in (5.1a) follows by Lemmas 4.9 and 4.10. To show (5.1b) note that

$$(5.2) \quad \overline{\varphi}(t, 0) - \overline{\varphi}(t', 0) \leq \overline{\varphi}(t, 0) - \min_{s \in [0, t]} \overline{\varphi}(s, 0) \quad \forall t \in [0, t].$$

Let  $t^* \in \text{Arg min}_{s \in [0, t]} \overline{\varphi}(s, 0)$  and define  $T \triangleq t - t^*$ . Clearly,  $t^* = t - T \in \mathcal{T}_0$ . We have

$$(5.3) \quad \begin{aligned} \overline{\varphi}(t, 0) - \overline{\varphi}(t - T, 0) &\leq \mathbb{E}_0^{v^*} \left[ \int_0^T \overline{r}(X_s, v^*(X_s)) \, ds + \overline{\varphi}(t - T, X_T) \right] \\ &\quad - \overline{\varphi}(t - T, 0) \\ &= \mathbb{E}_0^{v^*} \left[ \int_0^T \overline{r}(X_s, v^*(X_s)) \, ds + \Psi(t - T, X_T) \right] \\ &= V^*(0) - \mathbb{E}_0^{v^*} [V^*(X_T)] + \mathbb{E}_0^{v^*} [\Psi(t - T, X_T)] \\ &\leq V^*(0) + C_0 + 2 \|\varphi_0\|_\infty + \varrho \delta_0^{-1} \mathbb{E}_0^{v^*} [V^*(X_T)], \end{aligned}$$

where the last inequality follows by (5.1a). However, by Lemma 3.5 there exists a constant  $\widetilde{M}_0$  such that

$$\sup_{T \geq 0} \mathbb{E}_0^{v^*} [V^*(X_T)] \leq \widetilde{M}_0.$$

It then follows by (5.3) that  $\overline{\varphi}(t, 0) - \overline{\varphi}(t - T, 0)$  is bounded above by a constant independent of  $t$  and  $T$ . The result then follows by (5.2).  $\square$

LEMMA 5.2. *Under the hypotheses of Lemma 5.1 there exists a constant  $k_0 > 0$  such that*

$$\mathbb{E}_x^{v^t} [\check{\tau} \wedge t] \leq k_0 + 2 \delta_0^{-1} (1 + \varrho \delta_0^{-1}) V^*(x) \quad \forall x \in B_0^c.$$

*Proof.* Subtracting  $\overline{\varphi}(t, 0)$  from both sides of (4.2), we obtain

$$\begin{aligned} \Psi(t, x) &\geq \mathbb{E}_x^{v^t} \left[ \int_0^{\check{\tau} \wedge t} \overline{r}(X_s, v_s^t(X_s)) \, ds + \Psi(t - \check{\tau} \wedge t, X_{\check{\tau} \wedge t}) \mathbb{I}\{\check{\tau} \leq t\} + \varphi_0(X_t) \mathbb{I}\{\check{\tau} > t\} \right. \\ &\quad \left. - \overline{\varphi}(t, 0) \mathbb{I}\{\check{\tau} > t\} + (\overline{\varphi}(t - \check{\tau} \wedge t, 0) - \overline{\varphi}(t, 0)) \mathbb{I}\{\check{\tau} \leq t\} \right]. \end{aligned}$$

We discard the nonnegative term  $\varphi_0(X_t) \mathbb{I}\{\check{\tau} > t\}$ , and we use Lemma 4.10 and (5.1b) to write the above inequality as

$$\begin{aligned}
(5.4) \quad \Psi(t, x) &\geq \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^{\check{\tau} \wedge t} \bar{r}(X_s, \hat{v}_s^t(X_s)) \, ds \right] - \sup_{0 \leq s \leq t} \|\Psi(s, \cdot)\|_{\infty, B_0} \\
&\quad - \mathbb{E}_x^{\hat{v}^t} [\bar{\varphi}(t, 0) \mathbb{I}\{\check{\tau} > t\}] + \mathbb{E}_x^{\hat{v}^t} [(\bar{\varphi}(t - \check{\tau} \wedge t, 0) - \bar{\varphi}(t, 0)) \mathbb{I}\{\check{\tau} \leq t\}] \\
&\geq \mathbb{E}_x^{\hat{v}^t} \left[ \int_0^{\check{\tau} \wedge t} \bar{r}(X_s, \hat{v}_s^t(X_s)) \, ds \right] - C_0 - \bar{\varphi}(t, 0) \mathbb{P}_x^{\hat{v}^t}(\{\check{\tau} > t\}) - \widehat{M}_0.
\end{aligned}$$

By (5.1a) and (5.4) we obtain

$$\begin{aligned}
C_0 + 2 \|\varphi_0\|_{\infty} + (1 + \varrho \delta_0^{-1}) V^*(x) &\geq \delta_0 \mathbb{E}_x^{\hat{v}^t}[\check{\tau} \wedge t] - \bar{\varphi}(t, 0) \mathbb{P}_x^{\hat{v}^t}(\{\check{\tau} > t\}) \\
&\quad - C_0 - \widehat{M}_0 \\
&\geq \left( \delta_0 - \frac{\bar{\varphi}(t, 0)}{t} \right) \mathbb{E}_x^{\hat{v}^t}[\check{\tau} \wedge t] - C_0 - \widehat{M}_0.
\end{aligned}$$

The result then follows by Lemma 4.1.  $\square$

LEMMA 5.3. *Under the hypotheses of Lemma 5.1,*

$$\bar{\varphi}(t, 0) \mathbb{P}_x^{\hat{v}^t}(\{\check{\tau} > t\}) \xrightarrow[t \rightarrow \infty]{} 0,$$

uniformly on  $x$  in compact sets of  $\mathbb{R}^d$ .

*Proof.* By Lemma 4.1 and Lemma 5.2 we have

$$\bar{\varphi}(t, 0) \mathbb{P}_x^{\hat{v}^t}(\{\check{\tau} > t\}) \leq \frac{\bar{\varphi}(t, 0)}{t} (k_0 + 2 \delta_0^{-1} (1 + \varrho \delta_0^{-1}) V^*(x)) \xrightarrow[t \rightarrow \infty]{} 0 \quad \forall x \in B_0^c. \quad \square$$

LEMMA 5.4. *Let Assumption 3.2 hold and also suppose the initial condition  $\varphi_0 \in \mathcal{C}^2(\mathbb{R}^d)$  is nonnegative and bounded. Then the map  $t \mapsto \varphi(t, 0)$  is bounded on  $[0, \infty)$ , and it holds that*

$$-\operatorname{osc}_{\mathbb{R}^d} \varphi_0 \leq \liminf_{t \rightarrow \infty} \varphi(t, 0) \leq \limsup_{t \rightarrow \infty} \varphi(t, 0) \leq \widehat{M}_0 + \varrho.$$

*Proof.* Define

$$g(t) \triangleq \inf_{U \in \mathfrak{U}} \mathbb{E}_0^U \left[ \int_0^t r(X_s, U_s) \, ds + \varphi_0(X_t) \right].$$

By (3.30) we have

$$\begin{aligned}
(5.5) \quad \varphi(t, 0) &= g(t) - \int_0^t e^{s-t} g(s) \, ds \\
&= (1 - e^{-t})^{-1} \int_0^t e^{s-t} (g(t) - g(s)) \, ds \\
&\quad + (1 - e^{-t})^{-1} e^{-t} \int_0^t e^{s-t} g(s) \, ds,
\end{aligned}$$

for  $t > 0$ . By Lemma 5.1,  $g(t) \leq \widehat{M}_0 + \varphi_0(0) + \varrho t$ . Therefore the second term on the right hand side of (5.5) vanishes as  $t \rightarrow \infty$ . By Lemma 4.2,  $g(t) - g(s) \geq -\text{osc}_{\mathbb{R}^d} \varphi_0$  for all  $s \leq t$ . Also, by Lemma 5.1,  $g(t) - g(s) \leq \widehat{M}_0 + \varrho(t-s)$  for all  $s \leq t$ . Evaluating the first integral on the right hand side of (5.5) we obtain the bound

$$(5.6) \quad -\text{osc}_{\mathbb{R}^d} \varphi_0 \leq \int_0^t e^{s-t} (g(t) - g(s)) ds \leq \widehat{M}_0 + \varrho \quad \forall t > 0.$$

The result follows by (5.5)–(5.6).  $\square$

Combining Corollary 4.11, the boundedness of  $t \mapsto \varphi(t, 0)$  asserted in Lemma 5.4, and (1.12), it follows that  $x \mapsto \varphi(t, x)$  is locally bounded in  $\mathbb{R}^d$ , uniformly in  $t \geq 0$ . Recall Definition 3.14. The standard interior estimates of the solutions of (3.12) provide us with the following regularity result:

**THEOREM 5.5.** *Under the hypotheses of Lemma 5.4 the closure of the orbit  $\{\Phi_t[\varphi_0], t \in \mathbb{R}_+\}$  is locally compact in  $\mathcal{C}^2(\mathbb{R}^d)$ .*

*Proof.* By Corollary 4.11 and Lemma 4.10, the oscillation of  $\bar{\varphi}$  is bounded on any cylinder  $[n, n+1] \times B_R$ , uniformly over  $n \in \mathbb{N}$ . This together with Lemma 5.4 imply that  $\Phi_t[\varphi_0](x)$  is bounded on  $(t, x) \in [n, n+1] \times B_R$ , for any  $R > 0$ , uniformly in  $n \in \mathbb{N}$ . It follows that the derivatives  $\partial_{ij}\Phi_t[\varphi_0]$  are Hölder equicontinuous on every ball  $B_R$  uniformly in  $t$  [14, Theorem 5.1]. The result follows.  $\square$

We now turn to the proof of our main result.

*Proof of Theorem 1.1.* Let  $\{t_n\}$  be any diverging sequence and let  $f$  be any limit in the topology of Markov controls (see [4, Section 2.4]) of  $\{\hat{v}^{t_n}\}$  along some subsequence of  $\{t_n\}$  also denoted as  $\{t_n\}$ . By Fatou's lemma and the stochastic representation of  $V^*$  in Theorem 3.1, we have,

$$(5.7) \quad \begin{aligned} \liminf_{n \rightarrow \infty} \mathbb{E}_x^{\hat{v}^{t_n}} \left[ \int_0^{\hat{\tau} \wedge t_n} \bar{r}(X_s, \hat{v}_s^{t_n}(X_s)) ds \right] &\geq \mathbb{E}_x^f \left[ \int_0^{\hat{\tau}} \bar{r}(X_s, f_s(X_s)) ds \right] \\ &\geq \inf_{v \in \mathfrak{U}_{\text{SSM}}} \mathbb{E}_x^v \left[ \int_0^{\hat{\tau}} \bar{r}(X_s, v(X_s)) ds \right] \\ &\geq V^*(x) - \|V^*\|_{\infty, B_0} \quad \forall x \in B_0^c. \end{aligned}$$

The second inequality in (5.7) is due to the fact that the infimum of

$$\mathbb{E}_x^U \left[ \int_0^{\hat{\tau}} \bar{r}(X_s, U_s) ds \right]$$

over all  $U \in \mathfrak{U}$  is realized at some  $v \in \mathfrak{U}_{\text{SSM}}$ , while the third inequality follows by (3.5). Therefore, by (1.12), (5.4), (5.7) and Lemmas 5.3 and 5.4 we have that

$$(5.8) \quad \begin{aligned} \liminf_{t \rightarrow \infty} \varphi(t, x) &= \liminf_{t \rightarrow \infty} (\Psi(t, x) + \varphi(t, 0)) \\ &\geq V^*(x) - \|V^*\|_{\infty, B_0} - C_0 - \widehat{M}_0 - \text{osc}_{\mathbb{R}^d} \varphi_0 \quad \forall x \in B_0^c. \end{aligned}$$

Also, by (5.1a) and Lemma 5.3 we obtain

$$(5.9) \quad \begin{aligned} \limsup_{t \rightarrow \infty} \varphi(t, x) &= \limsup_{t \rightarrow \infty} (\Psi(t, x) + \varphi(t, 0))\Psi(t, x) \\ &\leq C_0 + 2 \|\varphi_0\|_\infty + (1 + \varrho \delta_0^{-1})V^*(x) + \widehat{M}_0 + \varrho \end{aligned}$$

for all  $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$ .

Hence, by (5.8)–(5.9) if we select

$$c = -(\|V^*\|_{\infty, B_0} + C_0 + \widehat{M}_0 + \operatorname{osc}_{\mathbb{R}^d} \varphi_0),$$

then any  $\omega$ -limit point of  $\varphi(t, x)$  as  $t \rightarrow \infty$  lies in  $\mathcal{G}_c$  (see Definition 3.14). By Theorem 3.15 if  $\varphi_0 \in \mathcal{G}_c$ , then  $\varphi(t, x) \rightarrow V^*(x) + \varrho$  as  $t \rightarrow \infty$ . Since the  $\omega$ -limit set of  $\varphi_0$  is invariant and the only invariant set in  $\mathcal{G}_c$  is the singleton  $\{V^* - V^*(0) + \varrho\}$  the result follows.  $\square$

**6. Concluding Remarks.** We have studied the relative value iteration algorithm for an important class of ergodic control problems wherein instability is possible, but is heavily penalized by the near-monotone structure of the running cost. The near-monotone cost structure plays a crucial role in the analysis and the proof of stabilization of the quasilinear parabolic Cauchy initial value problem that models the algorithm.

We would like to conjecture that the RVI converges starting from any initial condition  $\varphi_0 \in \mathcal{O}_{V^*}$ . It is only the estimate in Lemma 4.2 that restricts us to consider bounded initial conditions only. We want to mention here that a related such estimate can be obtained as follows:

$$\begin{aligned} \overline{\varphi}(t, x) &= \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^t \overline{r}(X_s, U_s) ds + \varphi_0(X_t) \right] \\ &= \inf_{U \in \mathfrak{U}} \mathbb{E}_x^U \left[ \int_0^\tau \overline{r}(X_s, U_s) ds + \overline{\varphi}(t - \tau, X_{t-\tau}) \right] \\ &\geq -\varrho \tau + \min_{y \in \mathbb{R}^d} \overline{\varphi}(t - \tau, y) \quad \forall \tau \in [0, t], \quad \forall x \in \mathbb{R}^d. \end{aligned}$$

In particular

$$\min_{\mathbb{R}^d} \overline{\varphi}(t - \tau, \cdot) - \min_{\mathbb{R}^d} \overline{\varphi}(t, \cdot) \leq \varrho \tau \quad \forall \tau \in [0, t],$$

and this estimate does not depend on the initial data  $\varphi_0$ . This suggests that it is probably worthwhile studying the variation of the RVI algorithm that results by replacing  $\varphi(t, 0)$  by  $\min_{\mathbb{R}^d} \varphi(t, \cdot)$  in (1.7).

Rate of convergence results and computational aspects of the algorithm are open issues.

## REFERENCES

- [1] J. ABOUNADI, D. P. BERTSEKAS, AND V. S. BORKAR, *Learning algorithms for Markov decision processes with average cost*, SIAM J. Control Optim., 40 (2001), pp. 681–698.
- [2] A. ARAPOSTATHIS, *On the policy iteration algorithm for nondegenerate controlled diffusions under the ergodic criterion*, in Optimization, control, and applications of stochastic systems, Systems Control Found. Appl., Birkhäuser/Springer, New York, 2012, pp. 1–12.
- [3] A. ARAPOSTATHIS AND V. S. BORKAR, *A relative value iteration algorithm for nondegenerate controlled diffusions*, SIAM J. Control Optim., 50 (2012), pp. 1886–1902.
- [4] A. ARAPOSTATHIS, V. S. BORKAR, AND M. K. GHOSH, *Ergodic control of diffusion processes*, vol. 143 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2011.
- [5] A. ARAPOSTATHIS, V. S. BORKAR, AND K. SURESH KUMAR, *Relative value iteration for stochastic differential games*, arXiv:1210.8188 (2012).
- [6] 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 (2001), pp. 2037–2080.
- [7] R.-R. CHEN AND S. MEYN, *Value iteration and optimization of multiclass queueing networks*, Queueing Systems Theory Appl., 32 (1999), pp. 65–97.
- [8] S. EIDELMAN, S. KAMIN, AND F. PORPER, *Uniqueness of solutions of the Cauchy problem for parabolic equations degenerating at infinity*, Asymptot. Anal., 22 (2000), pp. 349–358.
- [9] M. GRUBER, *Harnack inequalities for solutions of general second order parabolic equations and estimates of their Hölder constants*, Math. Z., 185 (1984), pp. 23–43.
- [10] I. GYÖNGY AND N. KRYLOV, *Existence of strong solutions for Itô’s stochastic equations via approximations*, Probab. Theory Related Fields, 105 (1996), pp. 143–158.
- [11] 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 (1960), pp. 179–196.
- [12] N. V. KRYLOV, *Controlled diffusion processes*, vol. 14 of Applications of Mathematics, Springer-Verlag, New York, 1980.
- [13] ———, *Lectures on elliptic and parabolic equations in Sobolev spaces*, vol. 96 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2008.
- [14] O. A. LADYŽENSKAJA, V. A. SOLONNIKOV, AND N. N. URAL’CEVA, *Linear and quasilinear equations of parabolic type*, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence, R.I., 1967.
- [15] E. LÖCHERBACH, D. LOUKIANOVA, AND O. LOUKIANOV, *Polynomial bounds in the ergodic theorem for one-dimensional diffusions and integrability of hitting times*, Ann. Inst. Henri Poincaré Probab. Stat., 47 (2011), pp. 425–449.
- [16] S. MEYN AND R. L. TWEEDIE, *Markov chains and stochastic stability*, Cambridge University Press, Cambridge, second ed., 2009.
- [17] S. P. MEYN, *The policy iteration algorithm for average reward Markov decision processes with general state space*, IEEE Trans. Automat. Control, 42 (1997), pp. 1663–1680.
- [18] S. P. MEYN AND R. L. TWEEDIE, *Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes*, Adv. in Appl. Probab., 25 (1993), pp. 518–548.
- [19] W. STANNAT, *(Nonsymmetric) Dirichlet operators on  $L^1$ : existence, uniqueness and associated Markov processes*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 28 (1999), pp. 99–140.
- [20] D. J. WHITE, *Dynamic programming, Markov chains, and the method of successive approximations*, J. Math. Anal. Appl., 6 (1963), pp. 373–376.