Stabilization of a quasi-linear parabolic Cauchy problem associated with ergodic control of diffusions

Ari Arapostathis, Fellow, IEEE
Electrical and Computer Engineering
The University of Texas at Austin
Austin, Texas 78703
Email: ari@utexas.edu

Vivek S. Borkar, Fellow, IEEE
Department of Electrical Engineering
Indian Institute of Technology
Powai, Mumbai 400076, India
Email: borkar.vs@gmail.com

K. Suresh Kumar,
Department of Mathematics
Indian Institute of Technology
Powai, Mumbai 400076, India
Email: suresh@math.iitb.ac.in

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.

I. 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:

\[
dX_t = b(X_t, U_t) dt + \sigma(X_t) dW_t.
\]

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). These are described in \( \text{§III-A} \). Let \( r: \mathbb{R}^d \times U \to \mathbb{R} \) be a continuous function bounded from below, which without loss of generality we assume 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

\[
\limsup_{t \to \infty} \frac{1}{t} \int_0^t r(X_s, U_s) \, ds.
\]

A weaker, average formulation seeks to minimize

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t \mathbb{E}^U [r(X_s, U_s)] \, ds.
\]

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

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

i.e., the infimum of (3) over all admissible controls (for the definition of admissible controls see §III-A). Under suitable hypotheses solutions to the ergodic control problem can be constructed via the Hamilton–Jacobi–Bellman (HJB) equation

\[
a^{ij}(x) \partial_{ij} V + H(x, \nabla V) = \varrho,
\]

where \( a = [a^{ij}] \) is the symmetric matrix \( \frac{1}{2}\sigma\sigma^T \) 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

\[
H \left( x, \nabla V(x) \right) = b(x, v^*(x)) \cdot \nabla V(x) + r(x, v^*(x))
\]

a.e. in \( \mathbb{R}^d \). Obtaining solutions to (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 [1] 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) [2].

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

\[
\partial_t \varphi(t, x) = a^{ij}(x) \partial_{ij} \varphi(t, x) + H(x, \nabla \varphi) - \varphi(t, 0),
\]

\[
\varphi(0, x) = \varphi_0(x).
\]

Under a uniform (geometric) ergodicity condition that ensures the well-posedness of the associated HJB equation we showed in [3] that the solution of (7) converges as \( t \to \infty \) to a solution of (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 [4].

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 that 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:
\[
\lim_{|x| \to \infty} \min_{u \in U} r(x, u) > \varrho. \tag{8}
\]

In this paper we relax the blanket geometric ergodicity assumption and study the relative value iteration in (7) under the near-monotone hypothesis (8). It is well known that for near-monotone costs the HJB equation (5) possesses a unique up to a constant solution \( V \) which is bounded below in \( \mathbb{R}^d \). However, this uniqueness result is restricted. In general, for \( \beta > \varrho \) the equation
\[
a^{ij}(x) \partial_{ij} V + H(x, \nabla V) = \beta \tag{9}
\]
can have a multitude of solutions which are bounded below [5, Section 3.8.1].

The proof of convergence of (7) is facilitated by the study of the value iteration (VI) equation
\[
\partial_t \varphi(t, x) = a^{ij}(x) \partial_{ij} \varphi(t, x) + H(x, \nabla \varphi) - \varrho,
\varphi(0, x) = \varphi_0(x).
\tag{10}
\]
The initial condition is the same as in (7). Also \( \varrho \) is as in (4), so it is assumed known. Note that if \( \varphi \) is a solution of (7), then
\[
\varphi(t, x) = \varphi(x) - \varrho t + \int_0^t \varphi(s, 0) \, ds \tag{11}
\]
solves (10). We have in particular that
\[
\varphi(t, x) - \varphi(t, 0) = \varphi(x) - \varphi(0)
\tag{12}
\]
It follows that the function \( f \triangleq \varphi - \varphi \) does not depend on \( x \in \mathbb{R}^d \) and satisfies
\[
\frac{df}{dt} + f = \varrho - \varphi(t, 0).
\tag{13}
\]
Conversely, if \( \varphi \) is a solution of (10), then solving (13) one obtains a corresponding solution of (7) that takes the form [3, Lemma 4.4]:
\[
\varphi(t, x) = \varphi(x) - \int_0^t e^{\varrho s} \varphi(s, 0) \, ds + \varrho (1 - e^{-\varrho t}). \tag{14}
\]

It also follows from (14) that if \( t \to \varphi(t, x) \) is bounded for each \( x \in \mathbb{R}^d \), then so is the map \( t \to \varphi(t, x) \), and if the former converges as \( t \to \infty \), pointwise in \( x \), then so does the latter.

We note here that we study solutions of the VI equation that have the stochastic representation
\[
\varphi(t, x) = \inf_{U} \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) \, ds + \varphi_0(X_t) \right] - \varrho t,
\tag{15}
\]
where the infimum is over all admissible controls. These are called canonical solutions (see Definition 3.4). The first term in (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 \( t \to \varphi(t, x) \) is locally bounded in \( x \in \mathbb{R}^d \). In contrast, under the near-monotone hypothesis alone, \( t \to \varphi(t, x) \) 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 (10) and then show that all \( \omega \)-limit points of the semiflow of (7) lie in this region.

While we prefer to think of (7) as a continuous time and space relative value iteration, it can also be viewed as a ‘stabilization of a quasilinear parabolic PDE problem’ analogous to the celebrated result of Has’minskii (see [6]). 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 (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 (5) for the ergodic control problem is integrable with respect to some optimal invariant probability distribution. Then for any bounded initial condition \( \varphi_0 \in C^2(\mathbb{R}^d) \) it holds that
\[
\lim_{t \to \infty} \varphi(t, x) = V(x) - V(0) + \varrho,
\]
on 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 1.1:** Under the assumptions of Theorem 1.1 the value function \( V \) of the HJB for the ergodic control problem has the stochastic representation:
\[
V(x) - V(y) = \lim_{t \to \infty} \left( \inf_{U} \mathbb{E}_x^U \left[ \int_0^t r(X_s, U_s) \, ds \right] - \inf_{U} \mathbb{E}_y^U \left[ \int_0^t r(X_s, U_s) \, ds \right] \right)
\]
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., \( \min_{u} r(x, u) \to \infty \) as \( |x| \to \infty \). The initial condition \( \varphi_0 \) is chosen as a 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 (5) is also integrable with respect to \( \mu_{v^*} \).

Work related to this paper has appeared in [8]–[10]. In [8] the author considers a \( d \)-dimensional controlled diffusion
governed by
\[ dX_t = U_t \, dt + dW_t, \]
where the control process also lives in \( \mathbb{R}^d \) and the running cost satisfies certain structural conditions and has polynomial growth. It is shown that (9) admits a unique solution \( (V^*, \beta) \) with \( V^* \) having polynomial growth and satisfying
\[ \min_{i \in \mathbb{N}} V^* = 1. \]
Moreover, \( \beta = \varphi \). Provided that the running cost satisfies a certain growth condition, and the initial condition \( \varphi_0 \) is bounded below and has at most polynomial growth, it is shown that the solution of (10) converges uniformly on compacta and that \( \varphi(t, \cdot) - V^*(\cdot) \) tends to a constant as \( t \to \infty \).

In [9] the authors consider the Cauchy problem
\[ \partial_t f - \frac{1}{2} \Delta f + H(x, Df) = 0 \quad \text{in} \quad (0, \infty) \times \mathbb{R}^d, \]
\[ f(0, \cdot) = f_0 \quad \text{in} \quad \mathbb{R}^d. \]
They assume that the Hamiltonian \( H(x, p) \) has at most polynomial growth with respect to \( x \), and that it is convex and has at most quadratic growth in \( p \). They also assume that the Hessian of \( H \) with respect to \( p \) is strictly positive definite and bounded. Additional assumptions which result in ergodicity are also employed. Then provided that the initial condition \( \varphi_0 \) has a certain minimal rate of growth, and grows at most at a polynomial rate, convergence of \( \varphi(t, \cdot) - V^*(\cdot) \) to a constant as \( t \to \infty \) is established.

The paper is organized as follows. The next section introduces the notation used in the paper. Section III starts by describing in detail the model and the assumptions imposed. In §III-B 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 §III-C we address the issue of existence and uniqueness of solutions to (7) and (10) and describe some basic properties of these solutions. In §III-D we exhibit a region of attraction for the solutions of the VI. In §IV 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 V is dedicated to the proof of convergence of the solutions of the RVI, while §VI concludes with some pointers to future work.

Due to space limitations most of the results are stated without proofs. For detailed proofs we refer the reader to [11]. We do however demonstrate the proof of the main result using the lemmas in §IV and §V.

II. 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 \( \emptyset \) denotes the indicator function. We denote by \( \tau(A) \) the first exit time of a process \( \{X_t, \ t \geq 0\} \) from a set \( A \subset \mathbb{R}^d \), defined by
\[ \tau(A) \triangleq \inf \{ t > 0 : X_t \not\in A \}. \]

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

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 \( C^k(D) (C^\infty(D)) \) refers to the class of all real-valued functions on \( D \) whose partial derivatives up to order \( k \) (of any order) exist and are continuous, and \( C_b(D) \) denotes the set of all bounded continuous real-valued functions on \( D \).

We adopt the notation \( \partial_i \triangleq \frac{\partial}{\partial x_i} \) and for \( i,j \in \mathbb{N} \), \( \partial_{ij} \triangleq \frac{\partial^2}{\partial x_i \partial x_j} \). We often use the standard summation rule that repeated subscripts and superscripts are summed from 1 through \( d \). For example,
\[ a^{ij} \partial_{ij} \varphi + b^i \partial_i \varphi \triangleq \sum_{i,j=1}^d a^{ij} \partial^2 \varphi_{x_i x_j} + \sum_{i=1}^d b^i \partial_i \varphi. \]

For a nonnegative multi-index \( \alpha = (\alpha_1, \ldots, \alpha_d) \) we let \( D^\alpha \triangleq \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d} \). Let \( Q \) be a domain in \( \mathbb{R}_+ \times \mathbb{R}^d \). Recall that \( C^{r,k+2r}(Q) \) stands for the set of bounded continuous real-valued functions \( \varphi(t,x) \) defined on \( Q \) such that the derivatives \( D^\alpha \partial_t^r \varphi \) are bounded and continuous in \( Q \) for
\[ |\alpha| + 2r \leq k + 2r, \quad r \leq r. \]
By a slight abuse of notation, whenever the whole space \( \mathbb{R}^d \) is concerned, we write \( f \in C^{r,k+2r}(I \times \mathbb{R}^d) \) where \( I \) is an interval in \( \mathbb{R}_+ \), whenever \( f \in C^{r,k+2r}_c(Q) \) for all bounded domains \( Q \subset I \times \mathbb{R}^d \).

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 \( \varphi \in C_{\text{loc}}^\infty(Q) \), the space of smooth functions on \( Q \) with compact support. In this manner we obtain for example the space \( W^{2,p}_{\text{loc}}(Q) \).

III. Problem Statement and Preliminary Results

A. 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). All random processes in (1) live in a complete probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). 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 \( \mathcal{U} \), 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 the completion of \( \sigma\{X_0, U_r, W_r, r \leq s\} \) relative to \( (\mathcal{F}, \mathbb{P}) \). 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).

(A1) Local Lipschitz continuity: The functions
\[ b = [b^1, \ldots, b^d]^T \quad \text{and} \quad \sigma = [\sigma^{ij}] \]
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 \( u \in \mathbb{U} \),

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

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

(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 \( \|\sigma\|^2 \triangleq \text{trace}(\sigma \sigma^T) \).

(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 \( \xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d \).

In integral form, (1) is written as

\[
X_t = X_0 + \int_0^t b(X_s, U_s) \, ds + \int_0^t \sigma(X_s) \, dW_s. \tag{16}
\]

The second term on the right hand side of (16) is an Itô stochastic integral. We say that a process \( X = \{X_t(\omega)\} \) is a solution of (1), if it is \( \mathbb{F}_t \)-adapted, continuous in \( t \), defined for all \( \omega \in \Omega \) and \( t \in [0, \infty) \), and satisfies (16) for all \( t \in [0, \infty) \) at once a.s.

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

\[
L^u f(x) = a^{ij}(x) \partial_{ij} f(x) + b^i(x, u) \partial_i f(x). \tag{17}
\]

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 \( f \in C^2(\mathbb{R}^d) \) and with \( L^u \) as defined in (17), it holds that

\[
f(X_t) = f(X_0) + \int_0^t L^{U_s} f(X_s) \, ds + M_t, \quad \text{a.s.,} \quad \tag{18}
\]

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 Itô’s formula [12, p. 122] extends (18) to functions \( f \) in the local Sobolev space \( W^{2,p}_{\text{loc}}(\mathbb{R}^d) \), \( p \geq d \).

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

\[
X_t = x_0 + \int_0^t b(X_s, v(s, X_s)) \, ds + \int_0^t \sigma(X_s) \, dW_s. \tag{19}
\]

is said to have a **strong solution** if given a Wiener process \((W_t, \mathbb{F}_t)\) on a complete probability space \((\Omega, \mathbb{F}, \mathbb{P})\), there exists a process \( X \) on \((\Omega, \mathbb{F}, \mathbb{P})\), with \( X_0 = x_0 \in \mathbb{R}^d \), which is continuous, \( \mathbb{F}_t \)-adapted, and satisfies (19) 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 \( C([0, \infty), \mathbb{R}^d) \). It is well known that under Assumptions (A1)–(A3), for any Markov control \( v \), (19) has a unique strong solution [13].

Let \( \mathbb{U}_{\text{SM}} \) denote the set of stationary Markov controls. Under \( v \in \mathbb{U}_{\text{SM}} \), the process \( X \) is strong Markov, and we denote its transition function by \( P_t^v(x, \cdot) \). It also follows from the work of [14], [15] that under \( v \in \mathbb{U}_{\text{SM}} \), the transition probabilities of \( X \) have densities which are locally Hölder continuous. Thus \( L^v \) defined by

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

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

**B. The ergodic control problem**

We assume that the running cost function \( r(x, u) \) is nonnegative, 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 \( r(\cdot, u) \) over \( B_R \). In summary, we assume that

(A4) \( r : \mathbb{R}^d \times \mathbb{U} \to \mathbb{R}_+ \) is continuous and satisfies, for some constant \( \kappa_R > 0 \)

\[
|r(x, u) - r(y, u)| \leq \kappa_R |x - y| \quad \forall x, y \in B_R,
\]

for all \( u \in \mathbb{U} \), and all \( R > 0 \).

As mentioned in §3, 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. Throughout this paper the near-monotone hypothesis (8) is in effect.

The ergodic control problem for near-monotone cost functions is characterized by the following theorem which combines Theorems 3.4.7, 3.6.6 and 3.6.10, and Lemmas 3.6.8 and 3.6.9 in [5]. 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 \( V^* \in C^2(\mathbb{R}^d) \) which solves the HJB equation (5), and satisfies \( \min_{\mathbb{U}_{\text{SM}}} V^* = 1 \). Also, a control \( v^* \in \mathbb{U}_{\text{SM}} \) is optimal with respect to the criteria (2) and (3) if and only if it satisfies (6) a.e. in \( \mathbb{R}^d \). Moreover, recalling that \( \bar{\tau}_R = \tau(B_R^*), R > 0 \), we have

\[
V^*(x) = \inf_{v \in \mathbb{U}_{\text{SM}}} \mathbb{E}^v_x \left[ \int_0^{\bar{\tau}_R} (r(X_t, v(X_t)) - \varrho) \, dt + V^*(X_{\bar{\tau}_R}) \right] \tag{20}
\]

for all \( x \in B_R^* \), for all \( R > 0 \).

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

It follows that the optimal control \( v^* \) in Theorem 3.1 is stable.

The technical assumption in Theorem 1.1 is the following:

**Assumption 3.1:** The value function \( V^\ast \) is integrable with respect to some optimal invariant probability distribution \( \mu_{v^*} \).

However, many results in this paper do not rely on Assumption 3.1.

**Remark 3.1:** Assumption 3.1 is equivalent to the following [5, Lemma 3.3.4]: there exists an optimal stationary control \( v^\ast \), an inf-compact function \( V \in C^2(\mathbb{R}^d) \), and an open ball \( B \subset \mathbb{R}^d \) such that
\[
L^\varepsilon V(x) \leq -V^\ast(x) \quad \forall x \in B^c. \tag{21}
\]

For the rest of the paper \( v^\ast \in \mathcal{U}_SSM \) denotes some fixed control satisfying (6) and (21).

**Remark 3.2:** Assumption 3.1 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 \( \lim_{x \to \infty} \frac{1}{\sigma(x)} < -\frac{1}{2} \), the mean hitting time at \( 0 \) is bounded above by a second-degree polynomial in \( x \) [16, Theorem 5.6]. Therefore, in this case, the existence of second moments for \( \mu_{v^*} \) implies Assumption 3.1.

Assumption 3.1: An example of this case, borrowed from [5, Section 3.8.1], is the one-dimensional controlled diffusion
\[
dX_t = U_t dt + dW_t, \quad X_0 = x,
\]
where \( U_t \in [-1, 1] \) is the control variable. Let \( r(x, u) = 1 - e^{-|x|} \) be the running cost function. Clearly \( r \) is near-monotone. An optimal stationary Markov control is given by \( v^\ast(x) = -\text{sign}(x) \), and the corresponding invariant probability measure is
\[
\mu_{v^*}(dx) = e^{-2|x|} dx.
\]

Also \( \varrho = 1/3 \). Solving the HJB we obtain
\[
V^\ast(x) = \frac{2}{3} \left(e^{-|x|} + |x| - 1\right),
\]
which is clearly integrable with respect to \( \mu_{v^*} \).

Another class of problems for which Assumption 3.1 holds are linear controlled diffusions of the form
\[
dX_t = (AX_t + BU_t) dt + dW_t, \quad X_0 = x,
\]
where \( X_t \in \mathbb{R}^d \), \( A \in \mathbb{R}^{d \times d} \), \( B \in \mathbb{R}^{d \times m} \), and \( U_t \in \mathbb{R}^m \). Let \( r(x, u) = x^T R x + u^T S u \), with \( R \) and \( S \) positive definite square matrices. Suppose that the pair \( (A, B) \) is controllable. Then under any constant feedback control \( \psi(x) = Z x, Z \in \mathbb{R}^{m \times d} \), such that the matrix \( A + B Z \) is Hurwitz, the diffusion is positive recurrent and the average cost finite. Therefore \( \varrho < \infty \) and \( r \) is clearly near-monotone. Since the action space is not compact this problem does not fit our model exactly. So we modify the model as follows. Let \( Z \subset \mathbb{R}^{m \times d} \) be a compact set that contains the optimal gain corresponding to the optimal linear feedback control for the linear quadratic problem above. We use the transformation \( U_t = Z_t X_t \), with \( Z_t \) denoting the new control variable which lives in \( Z \). It is well known that the optimal invariant distribution is Gaussian, and that \( V^\ast \) is quadratic in \( x \). Integrability of \( V^\ast \) follows.

Assumption 3.1 is also implied by Assumption 3.2 in §III-D, under which we obtain convergence of the VI algorithm (see Theorem 3.6).

We need the following lemma.

**Lemma 3.1:** Under Assumption 3.1,
\[
\mathbb{E}^x_{\mu_{v^*}} \left[V^\ast(X_t) \right] \xrightarrow{t \to \infty} \mu_{v^*}[V^\ast] \triangleq \int_{\mathbb{R}^d} V^\ast(x) \mu_{v^*}(dx),
\]
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
\[
\sup_{t \geq 0} \mathbb{E}^x_{\mu_{v^*}} \left[V^\ast(X_t) \right] \leq m_r V^\ast(x) + 1 \quad \forall x \in \mathbb{R}^d. \tag{22}
\]

**Definition 3.1:** We let \( C_{v^\ast} (\mathbb{R}^d) \) denote the Banach space of functions \( f \in C(\mathbb{R}^d) \) with norm
\[
\|f\|_{v^\ast} \triangleq \sup_{x \in \mathbb{R}^d} |f(x)|.
\]

We also define
\[
\mathcal{O}_{v^\ast} \triangleq \left\{ f \in C_{v^\ast} (\mathbb{R}^d) \cap C^2(\mathbb{R}^d) : f \geq 0 \right\}.
\]

### C. The relative value iteration

The RVI and VI equations in (7) and (10) can also be written in the form
\[
\partial_t \varphi(t, x) = \min_{u \in U} \left[L^u \varphi(t, x) + r(x, u) \right] - \varphi(t, 0), \tag{23}
\]
\[
\partial_t \overline{\varphi}(t, x) = \min_{u \in U} \left[L^u \overline{\varphi}(t, x) + r(x, u) \right] - \overline{\varphi}(t, 0). \tag{24}
\]
with \( \varphi(0, x) = \overline{\varphi}(0, x) = \varphi_0(x) \).

**Definition 3.2:** Let \( \hat{v} = \{\hat{v}_t \in \mathbb{R}_+ : t \geq 0\} \) denote a measurable selector from the minimizer in (24) corresponding to a solution \( \overline{\varphi} \in C([0, \infty) \times \mathbb{R}^d) \cap C^{1,2}((0, \infty) \times \mathbb{R}^d) \). This is also a measurable selector from the minimizer in (23), provided \( \overline{\varphi} \) and \( \varphi \) are related by (11) and (14), and vice-versa. Note that the Markov control associated with \( \hat{v} \) is computed ‘backward’ in time (see (15)). Hence, for each \( t \geq 0 \) we define the (nonstationary) Markov control
\[
\hat{v}_t \triangleq \{\hat{v}_s \in \hat{v}_{t-s} : s \in [0, t]\}.
\]

Also, we adopt the simplifying notation
\[
\overline{\varphi}(x, u) \triangleq r(x, u) - \varrho.
\]

In most of the statements of intermediary results the initial data \( \varphi_0(x) \) is assumed without loss of generality to be nonnegative. We start with a theorem that proves the existence of a solution to (24) that admits the stochastic representation in (15). This does not require Assumption 3.1. First we need the following definition.
Definition 3.3: 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 $C_T^\infty(\mathbb{R}^d)$ denote the Banach space of functions in $C(\overline{\mathbb{R}_T^d})$ with norm

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

Theorem 3.2: Provided $\varphi_0 \in \mathcal{O}_{V^*}$, then

$$
\varphi(t, x) = \inf_{\nu \in \mathcal{U}} \mathbb{E}_x^t \left[ \int_0^t \mathcal{L}(X_s, \nu_s) \, ds + \varphi_0(X_t) \right],
$$

(25a)

is the minimal solution of (24) in $C([0, \infty) \times \mathbb{R}^d)$ which is bounded below on $\mathbb{R}_T^d$, for any $T > 0$. With $\varphi^t$ as in Definition 3.2, it admits the representation

$$
\varphi(t, x) = \mathbb{E}_x^t \left[ \int_0^t \mathcal{L}(X_s, \nu_s^t) \, ds + \varphi_0(X_t) \right],
$$

(25b)

and it holds that

$$
\mathbb{E}_x^t \left[ \varphi(t - \tau_R \wedge t, X_{\tau_R}) \right] \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty
$$

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

$$
\|\varphi\|_{V^*, T} \leq (1 + \rho T) \max \{1, \|\varphi_0\|_{V^*}\}.
$$

In the interest of economy of language we refer to the solution in (25a) as canonical. This is detailed in the following definition.

Definition 3.4: Given an initial condition $\varphi_0 \in \mathcal{O}_{V^*}$ we define the canonical solution to the VI in (24) as the solution which was constructed in the proof of Theorem 3.2 and was shown to admit the stochastic representation in (25a). In other words, this is the minimal solution of (24) in $C([0, \infty) \times \mathbb{R}^d)$ 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 (23) via (14).

For the rest of the paper a solution to the RVI or VI is always meant to be a canonical solution.

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

Lemma 3.2: Provided $\varphi_0 \in \mathcal{C}_{V^*}(\mathbb{R}^d) \cap C^2(\mathbb{R}^d)$, then the canonical solution $\varphi \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ of (24) satisfies the bound

$$
\mathbb{E}_x^t \left[ \varphi_0(X_t) - V^*(X_t) \right] \leq \mathbb{E}_x^t \left[ \varphi(t, x) - V^*(x) \right] \leq m_V \|\varphi_0 - V^*\|_{V^*} (V^*(x) + 1).
$$

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

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). 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 3.3: 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 $\varphi \in C(\overline{\mathbb{R}_T^d}) \cap C^{1,2}(\mathbb{R}_T^d)$ of (24), for some $T > 0$, which is bounded below in $\overline{\mathbb{R}_T^d}$ and satisfies $\|\varphi\|_{V^*, T} < \infty$ agrees with the canonical solution $\varphi$ on $\overline{\mathbb{R}_T^d}$.

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 3.4 below, whose proof uses the technique in [17].

We define the following class of functions:

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

Theorem 3.4: Suppose $V^* \in \mathcal{G}$ and that $\|\sigma\|$ is bounded in $\mathbb{R}^d$. Then, provided $\varphi_0 \in \mathcal{O}_{V^*}$, there exists a unique solution $\psi$ to (24) such that $\max_{t \in [0, T]} \psi(t, \cdot) \in \mathcal{G}$ for each $T > 0$.

We do not enforce any of the assumptions of Theorems 3.3 or 3.4 for the main results of the paper. Rather our analysis is based on the canonical solution to the VI and RVI which is well defined (see Definition 3.4).

D. A region of attraction for the VI algorithm

In this section we describe a region of attraction for the VI algorithm. This is a subset of $C^2(\mathbb{R}^d)$ which is positively invariant under the semiflow defined by (24) and all its points are convergent, i.e., converge to a solution of (5).

Definition 3.5: We let $\mathcal{E}_t[\varphi_0] : C^2(\mathbb{R}^d) \rightarrow C^2(\mathbb{R}^d)$, $t \in [0, \infty]$, denote the canonical solution (semiflow) of the VI in (24) starting from $\varphi_0$, and $\Phi_t[\varphi_0]$ denote the corresponding canonical solution (semiflow) of the RVI in (23). Let $\mathcal{E}$ denote the set of solutions of the HJB in (5), i.e.,

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

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

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

Let Assumption 3.1 hold. We claim that for each $c \in \mathbb{R}$, $\mathcal{G}_c$ is positively invariant under the semiflow $\mathcal{E}_t$. Indeed by (22) and (26), if $\varphi_0 \in \mathcal{G}_c$, then we have that

$$
c \leq \mathcal{E}_t[\varphi_0](x) - V^*(x) \leq \mathcal{E}_t[\varphi_0](X_t) - V^*(X_t) \leq m_V \|\varphi_0 - V^*\|_{V^*} (V^*(x) + 1).
$$

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

Theorem 3.5: Under Assumption 3.1, for each $\varphi_0 \in \mathcal{G}_0$, the semiflow $\mathcal{E}_t[\varphi_0]$ converges to $c_0 + V^* \in \mathcal{E}$ as $t \rightarrow \infty$, for some $c_0 \in \mathbb{R}$ that satisfies

$$
0 \leq c_0 \leq \int_{\mathbb{R}^d} (\varphi_0(x) - V^*(x)) \mu_{V^*}(dx).
$$

Also $\Phi_t[\varphi_0]$ converges to $V^*(\cdot) - V^*(0) + \varphi$ as $t \rightarrow \infty$. 

Remark 3.3: It follows from Theorem 3.5 that the map

$$
t \mapsto \int_{\mathbb{R}^d} \mathcal{E}_t[\varphi_0](x) \mu_{V^*}(dx)
$$

(27)
is strictly decreasing along the semiflow \( \Phi_t \) unless \( \varphi_0 \in E \).
This is because if the map in (27) is constant on some interval \([t_0, t_1]\), with \( t_1 > t_0 \), then we must have
\[
\Phi_t[\varphi_0](x) = \mathbb{E}^x \left[ \int_0^{t-t_0} \tau(X_s, v^*(X_s)) \, ds + \Phi_{t_0}[\varphi_0](X_{t-t_0}) \right] \quad \forall t \in [t_0, t_1]. \tag{28}
\]

But (28) implies that, for some constant \( C_0 \), we must have \( \Phi_t[\varphi_0](x) = C_0 + V^*(x) \) for all \( t \in [t_0, t_1] \). As a result of this monotone property of the map in (27), if \( A \) is a bounded subset of \( C_V(\mathbb{R}^d) \), then the only subsets of \( G_c \cap A \), with \( c \in \mathbb{R} \), which are invariant under the semiflow \( \Phi_t \) are the subsets of \( E \cap G_c \cap A \). Similarly, the only subset of \( G_c \cap A \) which is invariant under the semiflow \( \Phi_t \) is the singleton \( \{V^*(\cdot) - V^*(0) + g\} \), assuming of course that it is contained in \( G_c \cap A \). These facts are used later in the proof of Theorem 1.1.

We also have the following result which does not require Assumption 3.1.

**Corollary 3.1:** Suppose \( \varphi_0 \in C^2(\mathbb{R}^d) \) is such that \( \varphi_0 - V^* \) is bounded. Then \( \Phi_t[\varphi_0] \) converges as \( t \to \infty \) to a point in \( E \).

An interesting class of problems are those for which \( V^* \) does not grow faster than \( \min_{u \in U} r(\cdot, u) \). More precisely, we consider the following property:

**Assumption 3.2:** There exist positive constants \( \theta_1 \) and \( \theta_2 \) such that
\[
\min_{u \in U} r(x, u) \geq \theta_1 V^*(x) - \theta_2 \quad \forall x \in \mathbb{R}^d.
\]
Under Assumption 3.2 we have that
\[
L^* V^*(x) = \varrho - r(x, v^*(x)) \leq \varrho + \theta_2 - \theta_1 V^*(x),
\]
and it follows that under the control \( v^* \) the diffusion is geometrically ergodic with a Lyapunov function \( V^* \). In particular Assumption 3.2 implies Assumption 3.1.

We have the following theorem.

**Theorem 3.6:** Suppose Assumption 3.2 and the hypotheses of either Theorem 3.3 or 3.4 hold. Then, provided \( \varphi_0 \in \mathcal{O}_V \), the semiflow \( \Phi_t[\varphi_0] \) converges, as \( t \to \infty \), to a point \( c_0 + V^* \in E \) satisfying
\[
-\frac{\varrho \theta_2}{\varrho} \leq c_0 \leq \frac{\varrho + \theta_2}{\varrho} \|\varphi_0\|_{V^*}.
\]
Therefore \( \Phi_t[\varphi_0] \) converges to \( V^*(\cdot) - V^*(0) + \varrho \) as \( t \to \infty \).

**IV. GROWTH ESTIMATES FOR SOLUTIONS OF THE VALUE ITERATION**

Most of the results of this section do not require Assumption 3.1. It is only needed for Lemma 4.5. Throughout this section and also in \( S_V \) a solution \( \varphi \) (or \( \varphi(t) \)) always refers to the canonical solution of the VI (RVI) without further mention (see Definition 3.4).

**Lemma 4.1:** Suppose \( \varphi_0 \in \mathcal{O}_V \). Then
\[
\frac{1}{t} \varphi(t, x) \to 0 \quad \text{as} \quad t \to \infty.
\]

**Lemma 4.2:** Provided \( \|\varphi_0\|_{\infty} < \infty \), it holds that for all \( t \geq 0 \)
\[
\varphi(t, x) - \varphi(t, x) \leq \varrho \tau + \text{osc} \varphi \quad \forall x \in \mathbb{R}^d, \quad \forall \tau \in [0, t].
\]

**Definition 4.1:** We define:
\[
\mathcal{K} \triangleq \{ x \in \mathbb{R}^d : \min_{u \in U} r(x, u) \leq \varrho \}.
\]

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

**Lemma 4.3:** Suppose \( \varphi_0 \in \mathcal{O}_V \). Then for any \( t > 0 \) we have
\[
\varphi(t, x) > \min_{\partial B_0^c} \varphi(t, \cdot) \geq 2 \|\varphi_0\|_{\infty} + (1 + \varrho \delta_0)^{-1} V^*(x) \quad \forall x \in \mathbb{R}^d \setminus B_0^c.
\]

Using the parabolic Harnack inequality [18, Theorem 4.1] we obtain the following bound:

**Lemma 4.4:** Provided \( \varphi_0 \in C^2(\mathbb{R}^d) \) is nonnegative and bounded, we have
\[
\varphi(t, x) - \max_{\partial B_0} \varphi(t, \cdot) \leq 2 \|\varphi_0\|_{\infty} + (1 + \varrho \delta_0)^{-1} V^*(x) \quad \forall x \in B_0^c.
\]

**V. CONVERGENCE OF THE RELATIVE VALUE ITERATION**

In the next lemma we use the variable
\[
\Psi(t, x) \triangleq \varphi(t, x) - \varphi(t, 0).
\]

**Lemma 5.1:** Let Assumption 3.1 hold and also suppose that the initial condition \( \varphi_0 \in C^2(\mathbb{R}^d) \) is nonnegative and bounded. Then
\[
\Psi(t, x) \leq C_0 + 2 \|\varphi_0\|_{\infty} + (1 + \varrho \delta_0)^{-1} V^*(x), \quad \tag{29a}
\]
where \( C_0 \) is the constant in Lemma 4.5, and there exists a constant \( \tilde{M}_0 \) such that
\[
\tilde{\tau}(t) - \tilde{\tau}(t', 0) \leq \tilde{M}_0 \quad \forall t > t'. \tag{29b}
\]

The following corollary now follows by Lemmas 4.2, 4.5 and 5.1.

**Corollary 5.1:** Under the hypotheses of Lemma 5.1, it holds that
\[
\text{osc} \varphi \leq 2 C_0 + \text{osc} \varphi + \tilde{M}_0 + \varrho (t' - t) \quad \forall t' > t \geq 0.
\]

**Lemma 5.2:** Under the hypotheses of Lemma 5.1,
\[
\varphi(t, 0) \mathbb{P}_x^d(\{t > t\}) \to 0 \quad \text{as} \quad t \to \infty,
\]
uniformly on \( x \) in compact sets of \( \mathbb{R}^d \).

**Lemma 5.3:** Let Assumption 3.1 hold and also suppose the initial condition \( \varphi_0 \in C^2(\mathbb{R}^d) \) is nonnegative and bounded. Then the map \( t \mapsto \varphi(t, 0) \) is bounded on \([0, \infty)\), and it holds that
\[
-\text{osc} \varphi_0 \leq \liminf_{t \to \infty} \varphi(t, 0) \leq \limsup_{t \to \infty} \varphi(t, 0) \leq \tilde{M}_0 + \varrho.
\]
Combining Corollary 5.1, the boundedness of \( t \mapsto \psi(t,0) \) asserted in Lemma 5.3, and (12), it follows that \( x \mapsto \psi(t,x) \) is locally bounded in \( \mathbb{R}^d \), uniformly in \( t \geq 0 \). Recall Definition 3.5. The standard interior estimates of the solutions of (23) provide us with the following regularity result:

**Theorem 5.1:** Under the hypotheses of Lemma 5.3 the closure of the orbit \( \{ \varphi_t[\varphi_0], \ t \in \mathbb{R}_+ \} \) is locally compact in \( C^2(\mathbb{R}^d) \).

We now turn to the proof of our main result.

**Proof of Theorem 1.1.** Let \( \{ t_n \} \) be any sequence tending to \( \infty \) and let \( f \) be any limit in the topology of Markov controls (see [5, Section 2.4]) of \( \{ \varphi_{t_n} \} \) 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

\[
\liminf_{n \to \infty} \mathbb{E}_x^\varphi_{t_n} \left[ \int_0^{t_n \wedge t} \tau(X_s, \varphi_{s}^n(X_s)) \, ds \right] \\
\geq \mathbb{E}_x^\varphi \left[ \int_0^{t} \tau(X_s, f_s(X_s)) \, ds \right] \\
\geq \inf_{v \in \mathcal{C}_b(\mathcal{X})} \mathbb{E}_x^\varphi \left[ \int_0^{t} \tau(X_s, v(X_s)) \, ds \right] \\
\geq V^*(x) - \| V^* \|_{\infty, B_0} \quad \forall x \in B_0^c. \tag{30}
\]

Therefore, by (12), (20), (29a), (30) and Lemmas 5.2 and 5.3 there exist constants \( c_1 < 0 \) and \( c_2 > 0 \) (which depend on \( \varphi_0 \)) such that

\[
\liminf_{t \to \infty} \varphi(t, x) \geq V^*(x) + c_1 \quad \forall x \in B_0^c, \tag{31}
\]

\[
\limsup_{t \to \infty} \varphi(t, x) \leq c_2 + (1 + 9 \delta_0^{-1}) V^*(x) \tag{32}
\]

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

Hence, by (31)–(32) if we define

\[
A \triangleq \{ h \in C^2(\mathbb{R}^d) : \| h \|_{V^*} \leq c_2 + (1 + 9 \delta_0^{-1}) \},
\]

then any \( \omega \)-limit point of \( \varphi(t, x) \) as \( t \to \infty \) lies in \( G_c \cap A \) (see Definition 3.5). Since the \( \omega \)-limit set of \( \varphi_0 \) is invariant under the semiflow \( \Phi_t \), and by Remark 3.3 the only invariant subset of \( G_c \cap A \) is the singleton \( \{ V^* - V^*(0) + g \} \), the result follows.

### VI. 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.

### VII. ACKNOWLEDGEMENT

Ari Arapostathis’ work was supported in part by the Office of Naval Research under the Electric Ship Research and Development Consortium. Vivek Borkar’s work was supported in part by by grant 11RCCSG014 from IIT Bombay and a J. C. Bose Fellowship from the Department of Science and Technology, Government of India.