CrossMark

Proceedings of the Royal Society of Edinburgh, **151**, 1305–1330, *2021*
DOI:10.1017/prm.2020.61

On the policy improvement algorithm for ergodic risk-sensitive control

Ari Arapostathis
Department of Electrical and Computer Engineering, The University of Texas at Austin, EER 7.824, Austin, TX 78712 (ari@utexas.edu)

Anup Biswas and Somnath Pradhan
Department of Mathematics, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pune 411008, India
(anup@iiserpune.ac.in; somnath@iiserpune.ac.in)

(MS received 17 December 2019; accepted 29 July 2020)

In this article we consider the ergodic risk-sensitive control problem for a large class of multidimensional controlled diffusions on the whole space. We study the minimization and maximization problems under either a blanket stability hypothesis, or a near-monotone assumption on the running cost. We establish the convergence of the policy improvement algorithm for these models. We also present a more general result concerning the region of attraction of the equilibrium of the algorithm.

Keywords: Principal eigenvalue; semilinear differential equations; stochastic representation; policy improvement

2010 *Mathematics subject classification:* 35P30; 93E20; 60J60

1. Introduction

Given controlled dynamics governed by the Itô equation

$$\mathrm{d}X_t = b(X_t, U_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t$$

for some suitable functions b, and σ, where U is an admissible control taking values in a compact metric space $\mathbb{U}$, and W is a standard Wiener process, we consider the problem of minimizing or maximizing the functional

$$\mathcal{E}_x(c, U) := \limsup_{T\to\infty} \frac{1}{T}\,\log \mathbb{E}_x^U\big[\mathrm{e}^{\int_0^T c(X_s,U_s)\,\mathrm{d}s}\big]$$

over all admissible controls U, where c is a suitable *running cost* function. This is of course known as the (ergodic) risk-sensitive control problem. The presence of the exponential in the functional reduces the possibility of rare but devastating large excursions of the state process. Risk-sensitive control has attracted a lot of interest because of its applications in large deviations [**25**], mathematical finance [**16**, **22**], missile guidance [**30**] and other fields. For a book-length exposition of this topic see

© The Author(s), 2020. Published by Cambridge University Press on behalf of The Royal Society of Edinburgh

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

[33]. Though this topic has been around for quite sometime, some of major issues such as uniqueness of value functions, verification results, variational representations, etc., for the (ergodic) risk-sensitive control problem for a controlled diffusion were resolved fairly recently [4, 6–8]. The goal of this article is to establish the convergence of the policy improvement algorithm (PIA) for this problem. We are interested in both the minimization and maximization problems. Due to the nonlinear nature of the criterion (w.r.t. c), these two problems behave quite differently. This article complements the results of [19], where a policy improvement algorithm is considered for skip-free countable state controlled Markov chains with a finite action set, under a near-monotone cost assumption.

To describe our methodology, let us consider the minimization problem which seeks to determine

$$\mathcal{E}^*_{\min} \coloneqq \inf_{x\in\mathbb{R}^d} \inf_{U\in\mathfrak{U}} \mathcal{E}_x(c,U),$$

where $\mathfrak{U}$ denotes the set of all admissible control. As well known, under suitable hypotheses, $\mathcal{E}^*_{\min}$ is the principal eigenvalue of a certain semilinear PDE. More precisely, there exists a positive $V\in C^2(\mathbb{R}^d)$ satisfying

$$\operatorname{Tr}\bigl(a(x)\nabla^2 V(x)\bigr) + \min_{\zeta\in\mathbb{U}} \bigl[b(x,\zeta)\cdot\nabla V(x) + c(x,\zeta)V(x)\bigr] = \mathcal{E}^*_{\min} V(x),$$

where $a \coloneqq \frac{1}{2}\,\sigma\sigma^{\mathsf{T}}$. Existence, uniqueness and verification of optimality of stationary Markov controls are studied in [4, 8]. For an early account of this problem we refer to [17, 18, 21, 27, 29]. The PIA can be described as follows: (1) we start with some stationary Markov control v_0, (2) we determine the principal eigenpair (λ_k, V_k), $k\geqslant 0$, solving the eigenvalue problem in $\mathbb{R}^d$ with drift $b(x,v_k)$ and running cost $c(x,v_k)$, and (3) we improve the control by selecting v_{k+1} from

$$\operatorname*{Arg\,min}_{\zeta\in\mathbb{U}} \bigl[b(x,\zeta)\cdot\nabla V_k(x) + c(x,\zeta)V_k(x)\bigr].$$

We say that the PIA converges if (λ_k, V_k) converges to $(\mathcal{E}^*_{\min}, V)$, in a suitable sense, as $k\to\infty$. Similar problems have been studied for the ergodic control problem with an additive criterion; see for instance [3, 26] and references therein. Recently, [1, 2] study the PIA for ergodic control problems in dimension 1, where the diffusion matrix is also allowed to depend on the control parameter. The analysis of the PIA for the risk-sensitive problem is very different from the ergodic control problems with additive criteria. Our analysis heavily uses the monotonicity property of the principal eigenvalue, and the recurrence properties of the *ground state process* [8]. The main results are theorem 3.3 in § 3 which studies the minimization problem under uniform ergodicity hypotheses (see assumption 2.3), and theorem 4.3 in § 4 which deals with maximization problem under the near-monotone hypothesis in assumption 2.5. Also, theorem 5.7 in § 5 which is devoted to a more abstract study of the convergence of the PIA, and finally theorem 5.11 for the minimization problem under a near-monotone hypothesis on the running cost.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

1.1. Notation

We denote by $\tau(A)$ the *first exit time* of the process $\{X_t\}$ from the set $A \subset \mathbb{R}^d$, defined by

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

The open ball of radius r centred at $x \in \mathbb{R}^d$ is denoted by $\mathcal{B}_r(x)$, and $\mathcal{B}_r$ without an argument denotes the ball centred at 0. We let $\tau_r \coloneqq \tau(\mathcal{B}_r)$, and $\breve{\tau}_r \coloneqq \tau(\mathcal{B}_r^c)$.

The complement and closure of a set $A \subset \mathbb{R}^d$ are denoted by A^c and $\bar{A}$, respectively, and $\mathbb{1}_A$ denotes its indicator function. Given $a, b \in \mathbb{R}$, the minimum (maximum) is denoted by $a \wedge b$ ($a \vee b$), respectively, and $a^{\pm} \coloneqq (\pm a) \vee 0$. The inner product of two vectors x and y in $\mathbb{R}^d$ is denoted as $x \cdot y$, or $\langle x, y\rangle$, $|\cdot|$ denotes the Euclidean norm, x^{T} stands for the transpose of x, and $\operatorname{Tr} S$ denotes the trace of a square matrix S.

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 $C^k(D)$ ($C_b^k(D)$) $k \geqslant 0$, refers to the class of all real-valued functions on D whose partial derivatives up to order k exist and are continuous (and bounded), $C_{\mathrm{c}}^k(D)$ denotes its subset consisting of functions that have compact support, and $C_0^k(D)$ the closure of $C_{\mathrm{c}}^k(D)$. The space $L^p(D)$, $p \in [1, \infty)$, stands for the Banach space of (equivalence classes of) measurable functions f satisfying $\int_D |f(x)|^p \,\mathrm{d}x < \infty$, and $L^\infty(D)$ is the Banach space of functions that are essentially bounded in D. The standard Sobolev space of functions on D whose generalized derivatives up to order k are in $L^p(D)$, equipped with its natural norm, is denoted by $\mathcal{W}^{k,p}(D)$, $k \geqslant 0$, $p \geqslant 1$. In general, if $\mathcal{X}$ is a space of real-valued functions on Q, $\mathcal{X}_{\mathrm{loc}}$ consists of all functions f such that $f\varphi \in \mathcal{X}$ for every $\varphi \in C_{\mathrm{c}}(Q)$. Likewise, we define $\mathcal{W}^{k,p}_{\mathrm{loc}}(D)$.

The constants $\kappa_1, \kappa_2, \ldots$ are generic and their values might differ from place to place.

2. Model and assumptions

The controlled diffusion process $\{X_t\}_{t \geqslant 0}$ in $\mathbb{R}^d$ is governed by the Itô equation

$$\mathrm{d}X_t = b(X_t, U_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \quad X_0 = x \in \mathbb{R}^d. \tag{2.1}$$

Here, W is a d-dimensional standard Wiener process defined on a complete probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. The control process $\{U_t\}_{t \geqslant 0}$ takes values in a compact metric space $\mathbb{U}$, is progressively measurable with respect to $\mathfrak{F}_t$, and is *non-anticipative*: for $s < t$, $W_t - W_s$ is independent of

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

The process U is called an *admissible* control, and the set of all admissible control is denoted by $\mathfrak{U}$.

Let $a = \frac{1}{2}\sigma\sigma^{\mathsf{T}}$. We impose the following assumptions to guarantee existence of solution of (2.1).

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

(A1) *Local Lipschitz continuity:* for some constant $C_R > 0$ depending on $R > 0$, the function $b\colon \mathbb{R}^d \times \mathbb{U} \to \mathbb{R}^d$, satisfies

$$|b(x,\zeta) - b(y,\zeta)|^2 \leqslant C_R\,|x-y|^2$$

for all $x, y \in \mathcal{B}_R$, and $\zeta \in \mathbb{U}$. Furthermore, σ is assumed to be bounded and globally Lipschitz.

(A2) *Growth condition:* For some constant $C_0 > 0$, we have

$$\sup_{\zeta\in\mathbb{U}} |b(x,\zeta)| \leqslant C_0\big(1+|x|\big) \qquad \forall\, x \in \mathbb{R}^d. \tag{2.2}$$

(A3) *Nondegeneracy:* For some positive constant C, it holds that

$$\sum_{i,j=1}^{d} a^{ij}(x)\eta_i\eta_j \geqslant C|\eta|^2 \qquad \forall\, x \in \mathbb{R}^d,\ \forall\, \eta = (\eta_1,\dots,\eta_d)^{\mathsf{T}} \in \mathbb{R}^d.$$

It is well known that under hypotheses (A1)–(A2), the stochastic differential equation in (2.1) has a unique strong solution for every admissible control (see for example, [**9**, theorem 2.2.4]). By a Markov control, we mean an admissible control of the form $v(t, X_t)$ for some Borel measurable function $v\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{U}$. If v is independent of t, we call it a stationary Markov control, and the set of all stationary Markov controls is denoted by $\mathfrak{U}_{\mathrm{sm}}$. The hypotheses in (A1)–(A3) imply the existence of unique strong solutions under Markov controls [**24**, theorem 2.8]. We also mention here that for bounded measurable coefficients Veretennikov [**31**] established the existence of a strong solution by proving the pathwise uniqueness (see also theorem 5 in [**32**]).

Definition 2.1. Let $c\colon \mathbb{R}^d \times \mathbb{U} \to \mathbb{R}$ be a continuous function which represents the *running cost*. We assume that c is locally Lipschitz continuous in its first argument uniformly with respect to the second. Given a control $U \in \mathfrak{U}$ the *risk-sensitive criterion* is given by

$$\mathcal{E}_x(c,U) \coloneqq \limsup_{T\to\infty} \frac{1}{T}\,\log \mathbb{E}_x^U\Big[\mathrm{e}^{\int_0^T c(X_s,U_s)\,\mathrm{d}s}\Big],$$

and the optimal values are defined as

$$\mathcal{E}^*_{\min} \coloneqq \inf_{x\in\mathbb{R}^d}\ \inf_{U\in\mathfrak{U}}\ \mathcal{E}_x(c,U),$$

and

$$\mathcal{E}^*_{\max} \coloneqq \sup_{x\in\mathbb{R}^d}\ \sup_{U\in\mathfrak{U}}\ \mathcal{E}_x(c,U).$$

These optimal values are extensively studied in [**4**, **6**–**8**, **17**, **18**, **21**, **27**]. In this article, we impose the following assumption on the growth of c.

(A4) *Growth of the running cost:* For some constant C, we have

$$\sup_{\zeta\in\mathbb{U}} |c(x,\zeta)| \leqslant C\big(1+|x|^2\big) \qquad \forall\, x \in \mathbb{R}^d.$$

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

Hypotheses (A1)–(A4) are assumed throughout the paper, unless explicitly indicated otherwise.

DEFINITION 2.2. We define the operators $\mathcal{A}$ and $\mathcal{A}^c$ mapping $C^2(\mathbb{R}^d)$ to $C(\mathbb{R}^d \times \mathbb{U})$ by

$$\mathcal{A}u(x,\zeta) \coloneqq \operatorname{Tr}\big(a(x)\nabla^2 u(x)\big) + b(x,\zeta)\cdot\nabla u(x),$$
$$\mathcal{A}^c u(x,\zeta) \coloneqq \mathcal{A}u(x,\zeta) + c(x,\zeta)u(x),$$

and for $v \in \mathfrak{U}_{\mathrm{sm}}$, we often use the simplifying notation

$$b_v(x) \coloneqq b\big(x, v(x)\big), \quad \text{and} \quad c_v(x) \coloneqq c\big(x, v(x)\big). \tag{2.3}$$

In the study of $\mathcal{E}^*_{\min}$ we enforce the following Foster–Lyapunov condition on the dynamics.

ASSUMPTION 2.3. In (a) and (b) below, $\mathcal{V} \in C^2(\mathbb{R}^d)$ is some function taking values in $[1,\infty)$, c is nonnegative, $\widehat{C}$ is a positive constant, and $\mathcal{K} \subset \mathbb{R}^d$ is a compact set.

(a) If c is bounded, we assume without loss of generality that $\inf_{\mathbb{R}^d\times\mathbb{U}} c = 0$, and that there exists some constant $\gamma > \|c\|_\infty$ satisfying

$$\mathcal{A}\mathcal{V}(x,\zeta) \leqslant \widehat{C}\mathbb{1}_{\mathcal{K}}(x) - \gamma\mathcal{V}(x) \qquad \forall\,(x,\zeta) \in \mathbb{R}^d \times \mathbb{U}. \tag{2.4}$$

(b) If c is not bounded, we assume that there exists an inf-compact function ℓ (i.e., the sublevel sets $\{\ell \leqslant \kappa\}$ are compact, or empty, in $\mathbb{R}^d$ for each $\kappa \in \mathbb{R}$) such that $x \mapsto \ell(x) - \max_{\zeta\in\mathbb{U}} c(x,\zeta)$ is also inf-compact, and

$$\mathcal{A}\mathcal{V}(x,\zeta) \leqslant \widehat{C}\mathbb{1}_{\mathcal{K}}(x) - \ell(x)\mathcal{V}(x) \qquad \forall\,(x,\zeta) \in \mathbb{R}^d \times \mathbb{U}. \tag{2.5}$$

As well known (see [**8**]), if a and b are bounded, it might not be possible to find an unbounded function ℓ satisfying (2.5). This is the reason for (2.4). Also, due to (2.2), it is known from [**9**, pg. 65] that $x \mapsto \mathbb{E}^U_x[\tau(\mathcal{K}^c)]$ is inf-compact for any admissible control U, and therefore, the Lyapunov function $\mathcal{V}$ in (2.4) and (2.5) are also inf-compact.

Before we proceed further, let us exhibit a class of dynamics satisfying assumption 2.3.

EXAMPLE 2.4. Suppose that $\sup_{\zeta\in\mathbb{U}} b(x,\zeta)\cdot x \leqslant -\kappa|x|^\alpha$ outside a compact set for some $\alpha \in [1,2]$, and a is bounded. Define $\mathcal{V}(x) \coloneqq \exp(\delta\sqrt{|x|^2+1})$. Then an easy calculation shows that

$$\mathcal{A}\mathcal{V}(x) \leqslant \kappa_1\left(\mathbb{1}_{\mathcal{K}_1}(x) + \delta\frac{1}{\sqrt{|x|^2+1}} + \delta^2\frac{|x|^2}{|x|^2+1}\right)\mathcal{V}(x) - \delta\frac{|x|^\alpha}{\sqrt{|x|^2+1}}\mathcal{V}(x)$$

for some constant κ_1, and a compact set $\mathcal{K}_1$. Thus, if $\alpha > 1$, and we choose $\ell \sim |x|^{\alpha-1}$, (2.5) is satisfied. For $\alpha = 1$, choosing δ suitably small we obtain (2.4).

For the maximization problem, we use the near-monotone hypothesis in assumption 2.5 below, which is somewhat weaker than [**7**, assumption 3.1 (iii)].

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

ASSUMPTION 2.5. It holds that

$$\lim_{r\to\infty} \max_{(x,\zeta)\in \mathcal{B}_r^c\times\mathbb{U}} c(x,\zeta) < \mathcal{E}^*_{\max} < \infty. \tag{2.6}$$

We let ϱ_* denote the principal eigenvalue of the maximal operator, defined as follows

$$\varrho_* \coloneqq \inf\Big\{\lambda : \exists\, \psi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d),\ p > d,\ \psi > 0,\\ \text{satisfying } \max_{\zeta\in\mathbb{U}} \mathcal{A}(x,\zeta)\psi \leqslant \lambda\psi \text{ in } \mathbb{R}^d\Big\}.$$

Theorem 2.5 implies that ϱ_* is finite. To see this, note first that $\varrho_* > -\infty$, since the Dirichlet eigenvalue on the unit ball is finite. Also, there exists some $r_0 \in (0,\infty)$ such that

$$\max_{(x,\zeta)\in \mathcal{B}_{r_0}^c\times\mathbb{U}} c(x,\zeta) < \mathcal{E}^*_{\max}$$

by (2.6). Hence, we have

$$\begin{aligned}\varrho_* &\leqslant \max_{(x,\zeta)\in\mathbb{R}^d\times\mathbb{U}} c(x,\zeta)\\ &< \max_{(x,\zeta)\in \mathcal{B}_{r_0}^c\times\mathbb{U}} c(x,\zeta) + \max_{(x,\zeta)\in \mathcal{B}_{r_0}\times\mathbb{U}} c(x,\zeta)\\ &< \mathcal{E}^*_{\max} + \max_{(x,\zeta)\in \mathcal{B}_{r_0}\times\mathbb{U}} c(x,\zeta) < \infty.\end{aligned}$$

REMARK 2.6. One can also study a maximization problem under assumption 2.3; the results and proofs are similar to the minimization problem. Therefore, we do not discuss the maximization problem under assumption 2.3.

2.1. Principal eigenvalues of linear operators

In this section we recall some facts about principal eigenvalues which are needed later. Let $b, f\colon \mathbb{R}^d \to \mathbb{R}$ be locally bounded Borel measurable functions. We also assume that $\langle x, b(x)\rangle^+ \leqslant C(1+|x|^2)$ for $x \in \mathbb{R}^d$. Consider the linear operator

$$\mathcal{L}^f u(x) \coloneqq \mathrm{Tr}\big(a(x)\nabla^2 u(x)\big) + b(x)\cdot\nabla u(x) + f(x)u(x).$$

The principal eigenvalue $\lambda^*(\mathcal{L}^f)$ of $\mathcal{L}^f$ is defined as follows [15]:

$$\lambda^*(\mathcal{L}^f) \coloneqq \inf\big\{\lambda\in\mathbb{R} : \exists \text{ positive } \psi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d),\ p > d,\\ \text{satisfying } \mathcal{L}^f\psi \leqslant \lambda\psi \text{ in } \mathbb{R}^d\big\}. \tag{2.7}$$

The reader should have noticed that the convention in this definition is the opposite of what is normally used in the pde literature. See for example the definition in [14, eq. (1.10)]. The same convention is used in [4, 6, 8] which contain results that we cite in this paper.

When we want to emphasize the dependence of $\lambda^*(\mathcal{L}^f)$ on f, we write this as $\lambda^*(f)$. It is also known from [15, theorem 1.4], that for any $\lambda \in [\lambda^*(f),\infty)$, there

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

exists a positive $\Psi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, $p > d$, satisfying $\mathcal{L}^f \Psi = \lambda \Psi$ in $\mathbb{R}^d$. We denote the principal eigenpair by (λ^*, Ψ^*). We now recall the following definition of strict monotonicity from [8]

DEFINITION 2.7. We say that λ^* is *strictly monotone at* f, if for some non-trivial nonnegative function h, vanishing at infinity, we have $\lambda^*(f-h) < \lambda^*(f)$. We also say that λ^* is *strictly monotone at* f *on the right*, if for any non-trivial nonnegative function h, we have $\lambda^*(f) < \lambda^*(f+h)$.

Note that strictly monotonicity, implies strictly monotonicity on the right, since the eigenvalue is a convex function of f.

Given an eigenpair (λ, Ψ) we define the *twisted process* $\{Y_t\}_{t \geqslant 0}$ as a solution to the SDE

$$\mathrm{d}Y_t = b(Y_t)\,\mathrm{d}t + 2a(Y_t)\nabla\psi(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t, \tag{2.8}$$

where $\psi = \log \Psi$. The twisted process corresponding to a principal eigenpair (λ^*, Ψ^*) is called a *ground state process*. Then we have the following result from [8, theorem 2.1].

THEOREM 2.8. *Suppose that* $\inf_{\mathbb{R}^d} f > -\infty$ *and* $\lambda^*(\mathcal{L}^f)$ *is finite. Then, the following hold:*

(i) *For* $\lambda > \lambda^*$, *the twisted process* (2.8) *corresponding to any eigenpair* (λ, Ψ) *is transient.*

(ii) *A ground state process is exponentially ergodic if and only if* λ^* *is strictly monotone at* f.

REMARK 2.9. The results of theorem 2.8 also hold if $\sup_{\mathbb{R}^d} f < \infty$.

3. Policy improvement for the minimization problem

In this section we prove the convergence of the policy improvement algorithm (PIA) for $\mathcal{E}^*_{\min}$ under assumption 2.3. We begin with the following result which is an extension of [8, theorem 4.1].

THEOREM 3.1. *Grant assumption 2.3. There exists a positive solution* $V \in C^2(\mathbb{R}^d)$ *satisfying*

$$\min_{\zeta \in \mathbb{U}} \mathcal{A}^c V(x,\zeta) = \mathcal{E}^*_{\min} V(x) \quad \text{in } \mathbb{R}^d, \quad \text{and} \quad V(0) = 1. \tag{3.1}$$

In addition, with $\overline{\mathfrak{U}}_{\mathrm{sm}} \subset \mathfrak{U}_{\mathrm{sm}}$ *denoting the class of Markov controls* v *which satisfy*

$$\begin{aligned} &b\big(x, v(x)\big) \cdot \nabla V(x) + c\big(x, v(x)\big) V(x) \\ &\quad = \min_{\zeta \in \mathbb{U}} \big[b(x,\zeta) \cdot \nabla V(x) + c(x,\zeta) V(x) \big] \ \text{a.e. in } \mathbb{R}^d, \end{aligned}$$

the following hold:

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

(a) $\overline{\mathfrak{U}}_{\rm sm} \subset \mathfrak{U}^*_{\rm sm}$ *and* $\lambda^*(c_v) = \mathcal{E}^*_{\min}$ *for all* $v \in \overline{\mathfrak{U}}_{\rm sm}$. *Here,* $\mathfrak{U}^*_{\rm sm}$ *denotes the set of all optimal stationary Markov controls.*

(b) $\mathfrak{U}^*_{\rm sm} \subset \overline{\mathfrak{U}}_{\rm sm}$.

(c) *Equation* (3.1) *has a unique solution in* $C^2(\mathbb{R}^d)$.

Next, we describe the PIA.

Algorithm 3.2. Policy iteration.

1. Initialization. Set $k = 0$ and select any $v_0 \in \mathfrak{U}_{\rm sm}$.
2. Value determination. Let $V_k \in \mathcal{W}^{2,p}_{\rm loc}(\mathbb{R}^d)$, $p > d$, be the unique principal eigenfunction satisfying $V_k(0) = 1$ and

$$\mathrm{Tr}\big(a(x)\nabla^2 V_k(x)\big) + b(x, v_k(x)) \cdot \nabla V_k(x) + c(x, v_k(x))V_k(x) = \lambda^*(c_{v_k})V_k(x) \tag{3.2}$$

for $x \in \mathbb{R}^d$. Existence of a unique principal eigenfunction in (3.2) follows from [**8**, § 3]. We let $\lambda_k \coloneqq \lambda^*(c_{v_k})$.

3. Policy improvement. Choose any $v_{k+1} \in \mathfrak{U}_{\rm sm}$ satisfying

$$v_{k+1}(x) \in \operatorname*{Arg\,min}_{\zeta \in \mathbb{U}} \left[b(x, \zeta) \cdot \nabla V_k(x) + c(x, \zeta)V_k(x) \right], \quad x \in \mathbb{R}^d.$$

The main result of this section is the following.

Theorem 3.3. *Under assumption* 2.3*, the following hold:*

(i) *For all* $k \in \mathbb{N}$, *we have* $\lambda_{k+1} \leqslant \lambda_k$, *and* $\lim_{k\to\infty} \lambda_k = \mathcal{E}^*_{\min}$.

(ii) *The sequence* $\{V_k\}$ *converges weakly in* $\mathcal{W}^{2,p}_{\rm loc}(\mathbb{R}^d)$, $p > d$, *to the unique solution* V *of* (3.1).

In the sequel, we use the notation

$$\begin{aligned} b_k(x) &\coloneqq b\big(x, v_k(x)\big), \qquad c_k(x) \coloneqq c\big(x, v_k(x)\big), \\ \mathcal{L}_k f(x) &\coloneqq \mathrm{Tr}\big(a(x)\nabla^2 f(x)\big) + b_k(x) \cdot \nabla f(x) + c_k(x) f(x). \end{aligned} \tag{3.3}$$

Let us start with following lemma.

Lemma 3.4. *We have* $\lambda_{k+1} \leqslant \lambda_k$ *for all* $k \geqslant 1$.

Proof. From the policy improvement algorithm, we have

$$\begin{aligned} \mathcal{A}V_k\big(x, v_{k+1}(x)\big) + c\big(x, v_{k+1}(x)\big)V_k(x) &= \min_{\zeta \in \mathbb{U}} \mathcal{A}^c V_k(x, \zeta) \\ &\leqslant \mathcal{A}V_k\big(x, v_k(x)\big) + c\big(x, v_k(x)\big)V_k(x) \\ &= \lambda_k V_k(x) \quad \text{a.e. in } \mathbb{R}^d. \end{aligned} \tag{3.4}$$

It then follows from the definition of the principal eigenvalue that $\lambda_{k+1} \leqslant \lambda_k$. □

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

Next we recall the following global estimate from [**7**, lemma 4.1].

LEMMA 3.5. *Suppose that a is bounded and has a uniform modulus of continuity in $\mathbb{R}^d$, and $b\colon \mathbb{R}^d \to \mathbb{R}^d$, $c\colon \mathbb{R}^d \to \mathbb{R}$ are locally bounded. Then there exists a constant $\widetilde{C}$, dependent on a, such that for any strong positive solution $\phi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, $p > d$, of $\mathcal{A}^c\phi = 0$ we have*

$$\frac{|\nabla\phi(x)|}{\phi(x)} \leqslant \widetilde{C}\left(1 + \sup_{y\in\mathcal{B}_1(x)}\left(|b(y)| + \sqrt{|c(y)|}\right)\right), \quad x \in \mathbb{R}^d.$$

Continuing, we define

$$\begin{aligned}\psi_k(x) &\coloneqq -\frac{1}{V_{k-1}(x)}\mathcal{A}V_{k-1}\big(x, v_k(x)\big) - c\big(x, v_k(x)\big) + \lambda_{k-1}\\ &= -\frac{1}{V_{k-1}(x)}\min_{\zeta\in\mathbb{U}}\mathcal{A}^cV_{k-1}(x,\zeta) + \lambda_{k-1}\\ &= \frac{1}{V_{k-1}(x)}\Big(b\big(x, v_{k-1}(x)\big)\nabla V_{k-1}(x) + c\big(x, v_{k-1}(x)\big)V_{k-1}(x)\\ &\qquad - \min_{\zeta\in\mathbb{U}}\big[b(x,\zeta)\cdot\nabla V_{k-1}(x) + c(x,\zeta)V_{k-1}(x)\big]\Big), \quad k \in \mathbb{N}.\end{aligned} \tag{3.5}$$

Note that ψ_k is a nonnegative function. We write (3.4) in the form

$$\min_{\zeta\in\mathbb{U}}\mathcal{A}^cV_k(x,\zeta) - \lambda_kV_k(x) = -\psi_{k+1}(x)V_k(x) \leqslant 0 \quad \text{a.e. in } \mathbb{R}^d. \tag{3.6}$$

Applying lemma 3.5 to the equation

$$\mathcal{L}_kV_k(x) = \lambda_kV_k(x) \quad \text{a.e. in } \mathbb{R}^d, \tag{3.7}$$

we get that

$$\begin{aligned}\frac{|\nabla V_k(x)|}{V_k(x)} &\leqslant \widetilde{C}\left(1 + \sup_{y\in\mathcal{B}_1(x)}\left(|b_k(y)| + \sqrt{|c_k(y) - \lambda_k|}\right)\right)\\ &\leqslant \widetilde{C}_1(1 + |x|), \quad x \in \mathbb{R}^d,\end{aligned} \tag{3.8}$$

for some constant $\widetilde{C}_1$, where in the last inequality we use (A2) and (A4). Again, using (3.7) together with (3.8), we find that

$$\frac{1}{V_k(x)}\big|\mathrm{Tr}\big(a(x)\nabla^2V_k(x)\big)\big| \leqslant |b_k(x)|\frac{|\nabla V_k(x)|}{V_k(x)} + |c_k(x) - \lambda_k| \leqslant \widetilde{C}_2(1 + |x|^2).$$

Thus, from (3.6), we obtain

$$\psi_{k+1}(x) \leqslant \widetilde{C}_3(1 + |x|^2), \quad \text{for all } x \in \mathbb{R}^d, \tag{3.9}$$

for some constant $\widetilde{C}_3$.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

Let $\{Y_t^k\}_{t\geqslant 0}$ be the twisted process (2.8) corresponding to the principal eigenpair (λ_k, V_k), that is, $\{Y_t^k\}_{t\geqslant 0}$ is the unique solution to the following SDE

$$\mathrm{d}Y_t^k = b_k(Y_t^k)\,\mathrm{d}t + 2a(Y_t^k)\nabla \log V_k(Y_t^k)\,\mathrm{d}t + \sigma(Y_t^k)\,\mathrm{d}W_t, \tag{3.10}$$

and let

$$\widetilde{\mathcal{L}}_k f \coloneqq (\mathcal{L}_k - c_k(x))f(x) + 2a(x)\nabla \log V_k(x)\cdot \nabla f(x) \tag{3.11}$$

denote the extended generator of (3.10). It follows by (3.8) that the drift of (3.10) has at most linear growth, and, therefore, $\{Y_t^k\}_{t\geqslant 0}$ is non-explosive. An easy calculation shows that, for any positive $\phi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, we have the following identity

$$\widetilde{\mathcal{L}}_k\left(\frac{\phi}{V_k}\right) = \left(\frac{\mathcal{A}\phi}{\phi} - \frac{\mathcal{A}V_k}{V_k}\right)\left(\frac{\phi}{V_k}\right) \quad \text{a.e. in } \mathbb{R}^d. \tag{3.12}$$

We let $\widetilde{P}_t^k(x,\mathrm{d}y)$ denote the transition probability of $\{Y_t^k\}_{t\geqslant 0}$.

Lemma 3.6. *Grant assumption 2.3. Then the family of invariant probability measures $\{\tilde{\mu}_k\}_{k\in\mathbb{N}}$ corresponding to the operators $\{\widetilde{\mathcal{L}}_k\}_{k\in\mathbb{N}}$ is tight. In addition, there exist positive constants $\gamma_\circ$ and $C_{\gamma_\circ}$ such that*

$$\left\|\widetilde{P}_t^k(x,\cdot) - \tilde{\mu}_k(\cdot)\right\|_{\mathrm{TV}} \leqslant C_{\gamma_\circ}\frac{\mathcal{V}(x)}{V_k(x)}\,\mathrm{e}^{-\gamma_\circ t} \qquad \forall\,(t,x)\in\mathbb{R}_+\times\mathbb{R}^d,\ \forall\,k\in\mathbb{N}. \tag{3.13}$$

Proof. Consider assumption 2.3 (a), and let $\epsilon_\circ \leqslant \gamma - \|c\|_\infty$. By (3.7) and (3.12), we have

$$\begin{aligned}\widetilde{\mathcal{L}}_k\left(\frac{\mathcal{V}}{V_k}\right)(x) &\leqslant \left(\widehat{C}\mathcal{V}^{-1}(x)\mathbb{1}_{\mathcal{K}}(x) - \gamma + c_k(x) - \lambda_k\right)\left(\frac{\mathcal{V}}{V_k}\right)(x)\\ &\leqslant \left(\widehat{C}\mathbb{1}_{\mathcal{K}}(x) - \epsilon_\circ - \lambda_k\right)\left(\frac{\mathcal{V}}{V_k}\right)(x) \quad \forall\,k\in\mathbb{N},\end{aligned} \tag{3.14}$$

since $\mathcal{V}^{-1}\leqslant 1$. Equation (3.14) implies that

$$\int_{\mathbb{R}^d}\frac{\mathcal{V}(x)}{V_k(x)}\,\tilde{\mu}_k(\mathrm{d}x) \leqslant M \quad \forall\,k\in\mathbb{N}, \tag{3.15}$$

for some constant M. Note that here $\mathcal{V}/V_k$ is, in general, not in $C^2(\mathbb{R}^d)$, so to derive (3.15), we apply [**10**, lemma 3.1]. It is well known that the linear growth bound in (3.8) implies that $x\mapsto \inf_{k\in\mathbb{N}}\widetilde{\mathbb{E}}_x^k[\tau(\mathcal{K}^c)]$ is inf-compact. Recall the definition $\breve{\tau}_r = \tau(\mathcal{B}_r^c)$ in § 1.1. Let $\{X_t^k\}$ be the process satisfying (2.1) with the Markov control v_k and $\mathbb{E}^k[\cdot]$ be the corresponding expectation operator. It is straightforward

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

to show, using the stochastic representation

$$V_k(x) = \mathbb{E}_x^k \left[\mathrm{e}^{\int_0^{\breve{\tau}_r} (c_k(X_t^k) - \lambda_k)\, \mathrm{d}t}\, V_k\big(X_{\breve{\tau}_r}^k\big) \right],$$

which holds for any $r > 0$ by [**8**, theorem 3.2], and the inequality

$$\mathcal{V}(x) \geqslant \mathbb{E}_x^k \left[\mathrm{e}^{\gamma \tau(\mathcal{K}^c)}\, \mathcal{V}\left(X_{\tau(\mathcal{K}^c)}^k\right) \right],$$

which follows from (2.4), that $\inf_{(k,x)\in\mathbb{N}\times\mathbb{R}^d} \frac{\mathcal{V}(x)}{V_k(x)} > 0$. On the other hand, applying the Itô–Krylov formula to §(3.14), followed by Fatou's lemma and the Jensen inequality, we obtain

$$\frac{\mathcal{V}(x)}{V_k(x)} \geqslant \epsilon_\circ\, \widetilde{\mathbb{E}}_x^k\big[\tau(\mathcal{K}^c)\big] \inf_{(k,x)\in\mathbb{N}\times\mathbb{R}^d} \frac{\mathcal{V}(x)}{V_k(x)},$$

which implies that $x \mapsto \inf_{k\in\mathbb{N}} \frac{\mathcal{V}(x)}{V_k(x)}$ is inf-compact. Hence, tightness of $\{\tilde{\mu}_k\}_{k\in\mathbb{N}}$ follows by (3.15). The proof under assumption 2.3 (b) is identical.

It is well known that (3.14) implies that $\{Y_t^k\}_{t\geqslant 0}$ is exponentially ergodic. It remains to show that the bound $\gamma_\circ$ on the rate convergence, and the constant $C_{\gamma_\circ}$ can be chosen independently of $k \in \mathbb{N}$. For this, we employ the same exact argument as in the proof of [**12**, theorem 2.1 (b)], which utilizes the estimates in [**28**, theorem 2.3]. This completes the proof. □

LEMMA 3.7. *Under assumption 2.3, ψ_k converges to 0 in $L^p_{loc}(\mathbb{R}^d)$ for any $p \geqslant 1$.*

Proof. Applying the Itô–Krylov formula to $\mathcal{L}_k V_{k-1} + (\psi_k - \lambda_{k-1})V_{k-1} = 0$, we obtain

$$\int_0^T \mathbb{E}_x^k \left[\mathrm{e}^{\int_0^t (c_k(X_s^k) - \lambda_{k-1})\, \mathrm{d}s}\, \psi_k(X_t^k) V_{k-1}(X_t^k) \right] \mathrm{d}t \leqslant V_{k-1}(x) \tag{3.16}$$

for all $k \in \mathbb{N}$, $T > 0$.

Let

$$h_k := \psi_k \frac{V_{k-1}}{V_k}, \quad k \in \mathbb{N}, \tag{3.17}$$

and $\tilde{\lambda}_k := \lambda_{k-1} - \lambda_k$. By [**8**, lemma 2.3], we have

$$\begin{aligned} &\mathbb{E}_x^k \left[\mathrm{e}^{\int_0^t (c_k(X_s^k) - \lambda_{k-1})\, \mathrm{d}s}\, \psi_k(X_t^k) V_{k-1}(X_t^k) \right] \\ &\quad = \mathrm{e}^{-\tilde{\lambda}_k t}\, V_k(x)\, \widetilde{\mathbb{E}}_x^k\big[h_k(Y_t^k)\big] \\ &\quad = \mathrm{e}^{-\tilde{\lambda}_k t}\, V_k(x) \int_{\mathbb{R}^d} \widetilde{P}_t^k(x, \mathrm{d}y) h_k(y). \end{aligned} \tag{3.18}$$

It is clear, by Sobolev imbedding, that the families of functions $\{V_k\}_{k\in\mathbb{N}}$, $\{V_k^{-1}\}_{k\in\mathbb{N}}$ and $\{\nabla V_k\}_{k\in\mathbb{N}}$ are locally Hölder equicontinuous. Hence, $\{\psi_k\}_{k\in\mathbb{N}}$ and $\{h_k\}_{k\in\mathbb{N}}$ defined in (3.5) and (3.17), respectively, are locally bounded.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

If $\tilde{\lambda}_k = 0$, then we have $\tilde{\mathcal{L}}_k(V_{k-1}/V_k) \leqslant 0$ by (3.12), and then by lemma 3.6 we obtain $V_{k-1} = V_k$ implying $\psi_k = 0$. So we suppose, without loss of generality, that $\tilde{\lambda}_k > 0$. Let $D \subset \mathbb{R}^d$ be an arbitrary ball. By (3.13), we have

$$\int_D \widetilde{P}_t^k(x, \mathrm{d}y) h_k(y) \geqslant \int_D h_k(y)\, \tilde{\mu}_k(\mathrm{d}y) - \|h_k\|_{\infty,D}\, C_{\gamma_\circ} \frac{\mathcal{V}(x)}{V_k(x)}\, \mathrm{e}^{-\gamma_\circ t}, \tag{3.19}$$

where $\|\cdot\|_{\infty,D}$ denotes the infinity norm of the restriction of a function to D. Thus, by (3.16), (3.18), and (3.19) we obtain

$$\frac{1}{\tilde{\lambda}_k} V_k(x) \int_D h_k(y)\, \tilde{\mu}_k(\mathrm{d}y) - \frac{C_{\gamma_\circ}}{\gamma_\circ} \|h_k\|_{\infty,D}\, \mathcal{V}(x) \leqslant V_{k-1}(x) \quad \forall\, k \in \mathbb{N},$$

which implies that

$$\int_D h_k(y)\, \tilde{\mu}_k(\mathrm{d}y) \leqslant \frac{\tilde{\lambda}_k}{V_k(x)} \Big(V_{k-1}(x) + \frac{C_{\gamma_\circ}}{\gamma_\circ} \|h_k\|_{\infty,D}\, \mathcal{V}(x) \Big) \quad \forall\, k \in \mathbb{N}. \tag{3.20}$$

Since $\{\tilde{\mu}_k\}_{k\in\mathbb{N}}$ is a tight family, and a is uniformly elliptic, and the drifts of (3.11) are bounded in D uniformly in $k \in \mathbb{N}$, the corresponding densities ρ_k are bounded above and away from 0 uniformly in $k \in \mathbb{N}$ (see [**9**, lemma 3.2.4 (b)]). Thus, the result follows by (3.20), since for any fixed x, $\{V_k(x)\}_{k\in\mathbb{N}}$, $\{V_k^{-1}(x)\}_{k\in\mathbb{N}}$ and $\big\{\|h_k\|_{\infty,D}\big\}_{k\in\mathbb{N}}$ are bounded, whereas $\tilde{\lambda}_k \searrow 0$ as $k \to \infty$. This completes the proof. □

Now we are ready to prove theorem 3.3.

Proof of theorem 3.3. From (3.6) we have

$$\begin{aligned} &\min_{\zeta\in\mathbb{U}} \mathcal{A}^c V_k(x,\zeta) + \psi_{k+1}(x) V_k(x) - \lambda_k V_k(x) \\ &\quad = \mathcal{L}_{k+1} V_k(x) + \psi_{k+1}(x) V_k(x) - \lambda_k V_k(x) \\ &\quad = 0. \end{aligned} \tag{3.21}$$

Since V_k solves (3.2) and $V_k(0) = 1$, and λ_k converges by lemma 3.4, it follows from the Harnack inequality that V_k is bounded. Thus, applying the well-known a priori estimate [**20**, lemma 5.3], it follows that $\{V_k\}_{k\geqslant 0}$ is locally bounded in $\mathcal{W}^{2,p}(\mathbb{R}^d)$, for $p > d$. Therefore, employing the Cantor diagonal argument we can extract a subsequence $\{V_{n_k}\}$ that converges weakly to some $\widetilde{V} \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, $p > d$. Furthermore, we have $V_{n_k} \to \widetilde{V}$ in $C^1_{\mathrm{loc}}(\mathbb{R}^d)$ by Sobolev embedding, and $\widetilde{V} \in \mathfrak{o}(\mathcal{V})$ from the proof of lemma 3.6. Let $\Lambda = \lim_{k\to\infty} \lambda_k \geqslant 0$. Then passing to the limit in (3.21) and using lemma 3.7 we obtain

$$\min_{\zeta\in\mathbb{U}} \mathcal{A}^c \widetilde{V}(x,\zeta) - \Lambda \widetilde{V}(x) = 0 \quad \text{in } \mathbb{R}^d. \tag{3.22}$$

It follows from standard elliptic regularity theory (see [**23**, theorem 9.19]) that $\widetilde{V} \in C^2(\mathbb{R}^d)$. More precisely, since $\widetilde{V} \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, for any $p > d$, it follows from

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

Sobolev embedding theorem that $\widetilde{V} \in C^{1,\alpha}_{\mathrm{loc}}(\mathbb{R}^d)$ for any $\alpha \in (0,1)$. Now using the locally Lipschitz property of b and c we can write (3.22) as

$$\operatorname{Tr}\bigl(a(x)\nabla^2\widetilde{V}(x)\bigr) = f(x)$$

for some locally α-Hölder continuous function f. Applying [**23**, theorem 9.19], we then obtain that $\widetilde{V} \in C^{2,\alpha}(\mathcal{B}_R)$ for any ball $\mathcal{B}_R$.

Next we show that $\mathcal{E}^*_{\min} = \Lambda$. It is obvious that $\mathcal{E}^*_{\min} \leqslant \Lambda$. Suppose that $\mathcal{E}^*_{\min} < \Lambda$. Let $v \in \overline{\mathfrak{U}}_{\mathrm{sm}}$ (see theorem 3.1) and $\mathcal{B}_r \ni \mathcal{K}$. Then, as shown [**8**], the solution V of (3.1) has the stochastic representation

$$V(x) = \mathbb{E}^v_x\Bigl[\mathrm{e}^{\int_0^{\check{\tau}_r}(c_v(X_t)-\mathcal{E}^*_{\min})\,\mathrm{d}t}\,V(X_{\check{\tau}_r})\Bigr], \quad x \in \mathcal{B}_r^c,\ v \in \overline{\mathfrak{U}}_{\mathrm{sm}}. \tag{3.23}$$

Since $\widetilde{V} \in \mathfrak{o}(\mathcal{V})$, it follows from (3.22) and [**6**, lemma 3.2] that

$$\begin{aligned}\widetilde{V}(x) &\leqslant \mathbb{E}^v_x[\mathrm{e}^{\int_0^{\check{\tau}_r}(c_v(X_t)-\Lambda)\,\mathrm{d}t}\,\widetilde{V}(X_{\check{\tau}_r})]\\ &\leqslant \mathbb{E}^v_x[\mathrm{e}^{\int_0^{\check{\tau}_r}(c_v(X_t)-\mathcal{E}^*_{\min})\,\mathrm{d}t}\,\widetilde{V}(X_{\check{\tau}_r})], \quad x \in \mathcal{B}_r^c.\end{aligned} \tag{3.24}$$

Let $\kappa = \inf_{\mathcal{B}_r} V/\widetilde{V}$. Then $V - \kappa\widetilde{V} \geqslant 0$ by (3.23) and (3.24), and it vanishes at some point in $\bar{\mathcal{B}}_r$. On the other hand, we have

$$\mathcal{A}(V-\kappa\widetilde{V})\bigl(x,v(x)\bigr) - \mathcal{E}^*_{\min}(V-\kappa\widetilde{V}) \leqslant 0 \quad \text{in } \mathbb{R}^d.$$

Hence, by the strong maximum principle we must have $V = \kappa\widetilde{V}$, which implies that $\mathcal{E}^*_{\min} = \Lambda$. Thus we reach a contradiction, and we conclude that $\mathcal{E}^*_{\min} = \Lambda$. Finally, from theorem 3.1 (c), it follows that $V = \widetilde{V}$. This also implies that the sequence $\{V_k\}$ converges to V. This completes the proof. □

4. Policy improvement for the maximization problem

In this section, we study the maximization problem, under theorem 2.5. We start the following theorem.

THEOREM 4.1. *Grant assumption 2.5. There exists a positive solution $\widehat{V} \in C^2(\mathbb{R}^d)$ satisfying*

$$\max_{\zeta\in\mathbb{U}} \mathcal{A}^c\widehat{V}(x,\zeta) = \varrho_*\widehat{V}(x) \quad \text{in } \mathbb{R}^d, \quad \text{and} \quad \widehat{V}(0) = 1. \tag{4.1}$$

In addition, if $\overline{\mathfrak{U}}_{\mathrm{sm}} \subset \mathfrak{U}_{\mathrm{sm}}$ denotes the class of Markov controls v which satisfy

$$\begin{aligned}&b\bigl(x,v(x)\bigr)\cdot\nabla\widehat{V}(x) + c\bigl(x,v(x)\bigr)\widehat{V}(x)\\ &\quad= \max_{\zeta\in\mathbb{U}}\bigl[b(x,\zeta)\cdot\nabla\widehat{V}(x) + c(x,\zeta)\widehat{V}(x)\bigr] \quad \text{a.e. in } \mathbb{R}^d,\end{aligned}$$

then the following hold:

(a) $\overline{\mathfrak{U}}_{\mathrm{sm}} = \mathfrak{U}^*_{\mathrm{sm}}$, *and* $\lambda^*(c_v) = \mathcal{E}^*_{\max} = \varrho_*$ *for all* $v \in \overline{\mathfrak{U}}_{\mathrm{sm}}$. *Here,* $\mathfrak{U}^*_{\mathrm{sm}}$ *denotes the set of all optimal stationary Markov controls.*

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

(b) *Equation* (4.1) *has a unique solution in* $\widehat{V} \in C^2(\mathbb{R}^d)$, *and this vanishes at infinity.*

(c) *For any* $v \in \overline{\mathfrak{U}}_{\rm sm}$ *we have*

$$\widehat{V}(x) = \mathbb{E}_x^v \left[e^{\int_0^{\check{\tau}_r} (c_v(X_t) - \varrho_*)\,\mathrm{d}t}\, \widehat{V}(X_{\check{\tau}_r}) \mathbb{1}_{\{\check{\tau}_r < \infty\}} \right] \quad \forall\, x \in \bar{\mathcal{B}}_r^c,\ \forall r > 0. \tag{4.2}$$

Proof. The statement of the theorem is the same as [7, theorem 3.1]. However, assumption 2.5 differs from [7, assumption 3.1 (iii)]. In (2.6) if we replace $\mathcal{E}^*_{\max}$ with ϱ_*, then the assertions follow from [7, theorem 3.1], since any limit of Neumann eigenvalues cannot be less than the principal eigenvalue. Suppose then that

$$\varrho_* \leqslant \lim_{r\to\infty} \max_{(x,\zeta)\in \mathcal{B}_r^c \times \mathbb{U}} c(x,\zeta) < \mathcal{E}^*_{\max} < \infty. \tag{4.3}$$

We use the fact that ϱ_* is a convex, and nondecreasing function of $c\colon \mathbb{R}^d \times \mathbb{U} \to \mathbb{R}$. Indeed, this follows from the convexity of the maximal operator, and since the principal eigenvalue $\lambda^*(\mathcal{L}^f)$ is a convex function of f. We add the dependence of ϱ_* on the coefficient c of the operator explicitly in the notation by denoting the eigenvalue as $\varrho_*(c)$. It is also clear from the remark following assumption 2.5 that $\varrho_*(c + \delta \mathbb{1}_{\mathcal{B}_1})$ is finite for each $\delta > 0$, and therefore, being convex, it is a continuous function of $\delta \in (0,\infty)$. Therefore, since ϱ_* is the limit of Dirichlet eigenvalues, which are strictly increasing as a function of c, it follows from (4.3) that there exists some $\delta > 0$ such that

$$\lim_{r\to\infty} \max_{(x,\zeta)\in \mathcal{B}_r^c \times \mathbb{U}} c(x,\zeta) < \varrho_*(c + \delta \mathbb{1}_{\mathcal{B}_1}) < \mathcal{E}^*_{\max}.$$

But then, as argued earlier, the assertions of the theorem hold, and these then imply the first equality in

$$\mathcal{E}^*_{\max}(c + \delta \mathbb{1}_{\mathcal{B}_1}) = \varrho_*(c + \delta \mathbb{1}_{\mathcal{B}_1}) < \mathcal{E}^*_{\max}.$$

Thus we are led to a contradiction. This precludes (4.3) as a possibility, and completes the proof. □

Next we state the PIA.

Algorithm 4.2. Policy iteration.

1. Initialization. Set $k = 0$ and select any $v_0 \in \mathfrak{U}_{\rm sm}$ which satisfies

$$\lambda^*(c_{v_0}) > \lim_{r\to\infty} \max_{(x,\zeta)\in \mathcal{B}_r^c \times \mathbb{U}} c(x,\zeta).$$

2. Value determination. Let $\widehat{V}_k \in \mathcal{W}^{2,p}_{\rm loc}(\mathbb{R}^d)$, $p > d$, be the unique principal eigenfunction satisfying

$$\mathrm{Tr}\big(a(x)\nabla^2 \widehat{V}_k(x)\big) + b(x,v_k)\cdot \nabla \widehat{V}_k(x) + c(x,v_k)\widehat{V}_k(x) = \lambda^*(c_{v_k})\widehat{V}_k(x) \tag{4.4}$$

a.e. $x \in \mathbb{R}^d$, with $\widehat{V}_k(0) = 1$. Existence of unique eigenfunction in (4.4) follows from [7]. Let $\widehat{\lambda}_k = \lambda^*(c_{v_k})$.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

3. Policy improvement. Choose any $v_{k+1} \in \mathfrak{U}_{\rm sm}$ satisfying

$$v_{k+1}(x) \in \operatorname*{Arg\,max}_{\zeta\in\mathbb{U}} \Big[b(x,\zeta)\cdot\nabla\widehat{V}_k(x) + c(x,\zeta)\widehat{V}_k(x) \Big], \quad x\in\mathbb{R}^d.$$

The main result of this section is the following.

THEOREM 4.3. *Grant assumption 2.5. Then, the following hold:*

(i) *For all $k\in\mathbb{N}$, we have $\widehat{\lambda}_k \geqslant \widehat{\lambda}_{k-1}$, and $\lim_{k\to\infty}\widehat{\lambda}_k = \mathcal{E}^*_{\max}$.*

(ii) *The sequence $\{\widehat{V}_k\}$ converges weakly in $\mathcal{W}^{2,p}_{\rm loc}(\mathbb{R}^d)$, $p>d$, to the unique solution $\widehat{V}$ of (4.1).*

We divide the proof in several lemmas. We adopt the notation in (3.3).

LEMMA 4.4. *We have $\widehat{\lambda}_{k+1}\geqslant\widehat{\lambda}_k$.*

Proof. We assume that $\widehat{\lambda}_k > \lim_{r\to\infty}\max_{(x,\zeta)\in\mathcal{B}_r^c\times\mathbb{U}} c(x,\zeta)$, and establish that $\widehat{\lambda}_{k+1}\geqslant\widehat{\lambda}_k$. We employ the method of induction. This holds for $k=0$. First, we show that

$$\mathcal{E}_x(c_{k+1},v_{k+1}) \geqslant \widehat{\lambda}_k \quad \forall\, x\in\mathbb{R}^d. \tag{4.5}$$

Note that

$$\operatorname{Tr}\big(a(x)\nabla^2\widehat{V}_k(x)\big) + b_k(x)\cdot\nabla\widehat{V}_k(x) + c_k(x)\widehat{V}_k(x) = \widehat{\lambda}_k\widehat{V}_k(x),$$

and $\lim_{|x|\to\infty}\widehat{V}_k(x)=0$ by theorem 4.1 (see also [**7**, theorem 3.2]). In particular, $\widehat{V}_k\in C_b(\mathbb{R}^d)$. Moreover,

$$\begin{aligned}
&\operatorname{Tr}\big(a(x)\nabla^2\widehat{V}_k(x)\big) + b_{k+1}(x)\cdot\nabla\widehat{V}_k(x) + \big(c_{k+1}(x)-\widehat{\lambda}_k\big)\widehat{V}_k(x)\\
&\quad= \max_{\zeta\in\mathbb{U}} \mathcal{A}^c\widehat{V}_k(x,\zeta) - \widehat{\lambda}_k\widehat{V}_k(x)\\
&\quad\geqslant \operatorname{Tr}\big(a(x)\nabla^2\widehat{V}_k(x)\big) + b_k(x)\cdot\nabla\widehat{V}_k(x) + (c_k-\widehat{\lambda}_k)\widehat{V}_k(x)\\
&\quad= 0.
\end{aligned} \tag{4.6}$$

Recall the definition $\uptau_n=\uptau(\mathcal{B}_n)$ in Section 1.1. Applying the Itô–Krylov formula to (4.6), we find that

$$\mathbb{E}_x^{v_{k+1}}\Big[\mathrm{e}^{\int_0^{T\wedge\uptau_n}(c_{k+1}(X_t)-\widehat{\lambda}_k)\,\mathrm{d}t}\,\widehat{V}_k(X_{T\wedge\uptau_n})\Big] \geqslant \widehat{V}_k(x) \qquad \forall\, x\in\mathcal{B}_n.$$

Since $\|c^+\|_\infty<\infty$, letting $n\to\infty$ above, and applying the dominated convergence theorem, we obtain

$$\mathbb{E}_x^{v_{k+1}}\Big[\mathrm{e}^{\int_0^{T}(c_{k+1}(X_t)-\widehat{\lambda}_k)\,\mathrm{d}t}\,\widehat{V}_k(X_T)\Big] \geqslant \widehat{V}_k(x).$$

Thus

$$\log\widehat{V}_k(x) \leqslant -\widehat{\lambda}_k T + \log\mathbb{E}_x^{v_{k+1}}\Big[\mathrm{e}^{\int_0^T c_{k+1}(X_t)\mathrm{d}t}\Big] + \log\|\widehat{V}_k\|_\infty.$$

Now dividing by T on both sides and letting $T\to\infty$ we have (4.5).

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

To complete the proof, we show that $\widehat{\lambda}_{k+1} = \sup_{x\in\mathbb{R}^d} \mathcal{E}_x(c_{k+1}, v_{k+1})$. In view of (4.6), the calculations above, and [**15**, theorem 1.7], we note that $\widehat{\lambda}_{k+1} \leqslant \sup_{x\in\mathbb{R}^d} \mathcal{E}_x(c_{k+1}, v_{k+1})$. Arguing as in the proof of theorem 4.1, if

$$\widehat{\lambda}_{k+1} > \lim_{r\to\infty} \max_{(x,\zeta)\in\mathcal{B}_r^c\times\mathbb{U}} c(x,\zeta), \tag{4.7}$$

then we have $\widehat{\lambda}_{k+1} = \mathcal{E}_x(c_{k+1}, v_{k+1})$ for all x. Now suppose that

$$\widehat{\lambda}_{k+1} \leqslant \lim_{r\to\infty} \max_{(x,\zeta)\in\mathcal{B}_r^c\times\mathbb{U}} c(x,\zeta) < \sup_{x\in\mathbb{R}^d} \mathcal{E}_x(c_{k+1}, v_{k+1}),$$

where the last inequality follows from (4.5). We know from [**15**] that $\delta \mapsto \lambda^*(c_{k+1} + \delta\mathbb{1}_{\mathcal{B}_1})$ is a convex function. Again, since $\lambda^*(c_{k+1} + \delta\mathbb{1}_{\mathcal{B}_1})$ are obtained as a increasing limit of Dirichlet principal eigenvalues, it follows that

$$\lim_{\delta\to\infty} \lambda^*(c_{k+1} + \delta\mathbb{1}_{\mathcal{B}_1}) \geqslant \lim_{\delta\to\infty} \lambda_1(c_{k+1} + \delta) = \infty,$$

where $\lambda_1(c_{k+1} + \delta)$ denotes the Dirichlet principal eigenvalue in the unit ball. Thus, we can find a $\delta_\circ$ satisfying

$$\lim_{r\to\infty} \max_{(x,\zeta)\in\mathcal{B}_r^c\times\mathbb{U}} c(x,\zeta) < \lambda^*(c_{k+1} + \delta_\circ\mathbb{1}_{\mathcal{B}_1}) < \sup_{x\in\mathbb{R}^d} \mathcal{E}_x(c_{k+1}, v_{k+1}).$$

Therefore, as argued earlier, it follows that

$$\sup_{x\in\mathbb{R}^d} \mathcal{E}(c_{k+1} + \delta_\circ\mathbb{1}_{\mathcal{B}_1}, v_{k+1}) = \lambda^*(c_{k+1} + \delta_\circ\mathbb{1}_{\mathcal{B}_1}) < \sup_{x\in\mathbb{R}^d} \mathcal{E}_x(c_{k+1}, v_{k+1}),$$

which is a contradiction. Thus, (4.7) must hold, and this completes the proof. □

Next we establish the strict monotonicity of the eigenvalue at every c_k (see definition 2.7). Recall the definition of $\mathcal{L}_k$ in (3.3).

LEMMA 4.5. *For every $k \geqslant 0$, the principal eigenvalue of $\mathcal{L}_k$ is strictly monotone with respect to the potential c_k.*

Proof. Let $h \neq 0$ be a nonnegative, compactly supported continuous function. From the definition of the principal eigenvalue it is clear that $\lambda_1 := \lambda^*(\mathcal{L}_k - h) \leqslant \lambda^*(\mathcal{L}_k) := \lambda_2$. Suppose that $\lambda_1 = \lambda_2$. Let $\Psi_i \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$ be the eigenfunction corresponding to the eigenvalue λ_i for $i = 1, 2$. Now, using lemma 4.4 it follows that

$$\lambda_1 = \lambda_2 > \lim_{r\to\infty} \max_{(x,\zeta)\in\mathcal{B}_r^c\times\mathbb{U}} c(x,\zeta),$$

and therefore, $\lim_{|x|\to\infty} \Psi_i(x) = 0$ and both the eigenfunctions have stochastic representation (4.2). Since h is compactly supported, by choosing r large enough we

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

obtain

$$\Psi_i(x) = \mathbb{E}_x^{v_k}\Big[e^{\int_0^{\check{\tau}_r}(c_k(X_t)-\lambda_1)\mathrm{d}t}\,\Psi_i(X_{\check{\tau}_r})\mathbb{1}_{\{\check{\tau}_r<\infty\}}\Big], \quad x\in\bar{\mathcal{B}}_r^c,\ i=1,2.$$

Now choose a constant $\kappa>0$ so that $\kappa\Psi_2\geqslant\Psi_1$, and equality holds at some point in $\bar{\mathcal{B}}_r$. On the other hand, we also have

$$\mathcal{L}_k(\kappa\Psi_2-\Psi_1) = -h\Psi_1\leqslant 0.$$

Thus, by the strong maximum principle we must have $\kappa\Psi_2=\Psi_1$ in $\mathbb{R}^d$, which contradicts the fact that $h\neq 0$. Therefore, $\lambda_1<\lambda_2$, thus completing the proof. □

Recall that the twisted process is given by

$$\mathrm{d}Y_t^k = b_k(Y_t^k)\,\mathrm{d}t + 2a(Y_t^k)\nabla\log\widehat{V}_k(Y_t^k)\,\mathrm{d}t + \sigma(Y_t^k)\,\mathrm{d}W_t. \tag{4.8}$$

From theorem 2.8, remark 2.9, lemma 4.5 we note that (4.8) is exponentially ergodic. Let $\tilde{\mu}_k$ be the corresponding unique invariant measure. In the lemma which follows, we show that $\{\tilde{\mu}_k\}_{k\geqslant 0}$ is tight.

LEMMA 4.6. *The family of invariant measures* $\{\tilde{\mu}_k\}_{k\geqslant 0}$ *is tight.*

Proof. Let $\epsilon,r>0$ be such that $\max_{\zeta\in\mathbb{U}}c(x,\zeta)-\widehat{\lambda}_0<-\epsilon$ for all $x\in\mathcal{B}_r^c$. It follows by lemma 4.4, that $\max_{\zeta\in\mathbb{U}}c(x,\zeta)-\widehat{\lambda}_k<-\epsilon$ for all $x\in\mathcal{B}_r^c$ and $k\geqslant 0$. Let $\widetilde{\mathcal{L}}_k$ denote the extended generator of (4.8). Then an easy calculation reveals that, with $\check{V}_k\coloneqq(\widehat{V}_k)^{-1}$, we have

$$\widetilde{\mathcal{L}}_k\check{V}_k + (\widehat{\lambda}_k - c_k)\check{V}_k = 0 \quad \text{in } \mathbb{R}^d, \tag{4.9}$$

and moreover, $\check{V}_k$ is inf-compact. As done earlier, denote by $\widetilde{\mathbb{E}}^k$ the expectation operator on the canonical space of the process (4.8). Then applying the Itô–Krylov formula and Fatou's lemma, we obtain

$$\begin{aligned}
\check{V}_k(x) &\geqslant \widetilde{\mathbb{E}}_x^k\Big[\mathrm{e}^{\int_0^{\check{\tau}_r}(\widehat{\lambda}_k-c_k(Y_t^k))\,\mathrm{d}t}\,\check{V}_k(Y_{\check{\tau}_r}^k)\Big]\\
&\geqslant \Big(\max_{\bar{\mathcal{B}}_r}\widehat{V}_k\Big)^{-1}\widetilde{\mathbb{E}}_x^k\big[\mathrm{e}^{\epsilon\check{\tau}_r}\big]\\
&\geqslant \Big(\max_{\bar{\mathcal{B}}_r}\widehat{V}_k\Big)^{-1}\epsilon\,\exp\big(\widetilde{\mathbb{E}}_x^k[\check{\tau}_r]\big) \qquad \forall\, x\in\bar{\mathcal{B}}_r^c.
\end{aligned}$$

Since, by lemma 3.5, $|\nabla\log\widehat{V}_k|\leqslant\kappa(1+|x|)$ for all $x\in\mathbb{R}^d$ and $k\in\mathbb{N}$, it follows from (A2) that $\check{V}_k$ is inf-compact, uniformly in k (see for instance, [**5**, lemma 5.1]), that is, $\inf_k\check{V}_k$ is inf-compact. Then, the result follows from (4.9). □

Next, we present the proof of theorem 4.3.

Proof of theorem 4.3. Let $\widehat{\Lambda}=\lim_{k\to\infty}\widehat{\lambda}_k$. Existence follows from lemma 4.4. In view of the proof of theorem 3.3, it is enough to show that $\widehat{\Lambda}=\mathcal{E}^*_{\max}$. As earlier,

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

we define

$$\max_{\zeta\in\mathbb{U}} \mathcal{A}^c\widehat{V}_k(x,\zeta) - \widehat{\lambda}_k\widehat{V}_k(x) = \widehat{\psi}_{k+1}(x)\widehat{V}_k(x).$$

Then $\widehat{\psi}_k$ satisfies (3.9), and using lemma 4.6, we obtain

$$\lim_{k\to\infty} \int_{\mathcal{B}_R} \widehat{\psi}_k(x)\,\mathrm{d}x = 0 \quad \forall\, R > 0.$$

Also, the uniform estimate on $\check{V}_k$ (see the proof of lemma 4.6) shows that any limit of $\{\widehat{V}_k\}$ must vanish at infinity. Therefore, we can follow the arguments in theorem 3.3 together with theorem 4.1 to complete the proof. □

5. A general result on convergence

In this section, we provide sufficient conditions for the PIA to converge, without assuming blanket stability hypotheses or near-monotonicity of the running cost. We apply these to the minimization problem under a near-monotone cost hypothesis in § 5.1. We address the minimization problem. Let

$$\mathcal{G}f(x) \coloneqq \mathrm{Tr}\big(a(x)\nabla^2 f(x)\big) + \min_{\zeta\in\mathbb{U}} \big[b(x,\zeta)\cdot\nabla f(x) + c(x,\zeta)f(x)\big],$$

and denote by $\lambda^*(\mathcal{G})$ the generalized principal eigenvalue of the operator $\mathcal{G}$ on $\mathbb{R}^d$, which is defined by (2.7) with $\mathcal{L}^f$ replaced by $\mathcal{G}$, and which is assumed to be finite. In this section, the coefficients a, b and c are not restricted to satisfy (A1)–(A4). Rather, we assume that they satisfy the more general hypotheses in [**6**, § 2.1]. That is, we replace (A1)–(A4) with the following:

(B1) The functions $b\colon \mathbb{R}^d\times\mathbb{U}\to\mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d\to\mathbb{R}^{d\times d}$ are continuous and satisfy

$$|b(x,u)-b(y,u)| + \|\sigma(x)-\sigma(y)\| \leqslant C_R\,|x-y| \qquad \forall\, x,y\in B_R,\ \forall\, u\in\mathbb{U},$$

for some constant $C_R > 0$ depending on $R > 0$.

(B2) For some $C_0 > 0$, we have

$$\sup_{\zeta\in\mathbb{U}} \langle b(x,\zeta),x\rangle^+ + \|\sigma(x)\|^2 \leqslant C_0\big(1+|x|^2\big) \qquad \forall\, x\in\mathbb{R}^d.$$

(B3) For each $R > 0$, it holds that

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

and for all $\xi = (\xi_1,\dots,\xi_d)^{\mathsf{T}}\in\mathbb{R}^d$, with C_R the constant in (B1).

(B4) The running cost $c(x,\zeta)$ is in $C(\mathbb{R}^d\times\mathbb{U},\mathbb{R})$, is bounded below in $\mathbb{R}^d$, and is locally Lipschitz in x uniformly with respect to $\zeta\in\mathbb{U}$.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

Hypotheses (B1)–(B4) are assumed throughout this section.

DEFINITION 5.1. We adopt the notation

$$\mathcal{A}_v^c f(x) \coloneqq \operatorname{Tr}\bigl(a(x)\nabla^2 f(x)\bigr) + b_v(x)\cdot\nabla f(x) + c_v(x) f(x), \quad v \in \mathfrak{U}_{\mathrm{sm}},$$

where b_v and c_v are as in (2.3), and let λ_v^* denote the principal eigenvalue $\lambda^*(\mathcal{A}_v^c)$ of $\mathcal{A}_v^c$ on $\mathbb{R}^d$ defined in (2.7). We let

$$\Lambda^* \coloneqq \inf_{v\in\mathfrak{U}_{\mathrm{sm}}} \lambda_v^*,$$

and

$$\mathcal{U}^* \coloneqq \bigl\{v \in \mathfrak{U}_{\mathrm{sm}} \colon \lambda_v^* = \Lambda^*\bigr\}.$$

We denote $\lambda^*(\mathcal{G})$ also as $\lambda^*(\mathcal{G}, c)$ when we need its dependence on the coefficient c of the operator to be explicitly captured in the notation. The following theorem extends the results in [4].

THEOREM 5.2. *Suppose that the eigenvalue $\lambda^*(\mathcal{G}, c)$ of the semilinear operator $\mathcal{G}$ is strictly monotone at c on the right, that is, for any non-trivial nonnegative function $h\colon \mathbb{R}^d \to \mathbb{R}$, we have $\lambda^*(\mathcal{G}, c) < \lambda^*(\mathcal{G}, c+h)$ (compare with definition 2.7). Then there exists a unique positive $\Psi_* \in C^2(\mathbb{R}^d)$, which satisfies $\Psi_*(0) = 1$, and*

$$\operatorname{Tr}\bigl(a(x)\nabla^2\Psi_*(x)\bigr) + \min_{\zeta\in\mathbb{U}}\bigl[b(x,\zeta)\cdot\nabla\Psi_*(x) + c(x,\zeta)\Psi_*(x)\bigr] = \lambda^*(\mathcal{G})\Psi_*(x). \tag{5.1}$$

In addition, if $\overline{\mathfrak{U}}_{\mathrm{sm}} \subset \mathfrak{U}_{\mathrm{sm}}$ denotes the class of Markov controls v which satisfy

$$b_v(x)\cdot\nabla\Psi_*(x) + c_v(x)\Psi_*(x) = \min_{\zeta\in\mathbb{U}}\bigl[b(x,\zeta)\cdot\nabla\Psi_*(x) + c(x,\zeta)\Psi_*(x)\bigr] \quad \text{a.e. in } \mathbb{R}^d,$$

then $\overline{\mathfrak{U}}_{\mathrm{sm}} = \mathcal{U}^*$.

Proof. First note that, under the hypothesis that $\lambda^*(\mathcal{G})$ is finite, existence of a positive solution Ψ_* is standard, and can be constructed as a limit of Dirichlet eigenvalue problems on an increasing sequence of balls in $\mathbb{R}^d$ as in [4, lemma 3.1]. Indeed, by [4, lemma 3.1], there exists a unique pair $(\widehat{\Psi}_n, \hat{\lambda}_n) \in \bigl(C^2(\mathcal{B}_n)\cap C(\bar{\mathcal{B}}_n)\bigr)\times\mathbb{R}$, $n\in\mathbb{N}$, satisfying $\widehat{\Psi}_n > 0$ on $\mathcal{B}_n$, $\widehat{\Psi}_n = 0$ on $\partial\mathcal{B}_n$, and $\widehat{\Psi}_n(0) = 1$, which solves

$$\min_{\zeta\in\mathbb{U}}\bigl[\mathcal{A}\widehat{\Psi}_n(x,\zeta) + c(x,\zeta)\,\widehat{\Psi}_n(x)\bigr] = \hat{\lambda}_n\,\widehat{\Psi}_n(x), \qquad x\in\mathcal{B}_n,$$

and $\hat{\lambda}_n$ is an increasing sequence. Any limit point $\widehat{\Psi}_n$ as $n\to\infty$ satisfies (5.1). However, here we use a different construction for Ψ_*. Since $\hat{\lambda}_n < \lambda^*(\mathcal{G})$ for all $n\in\mathbb{N}$, for any $\alpha_n > 0$, the Dirichlet problem

$$\min_{\zeta\in\mathbb{U}}\bigl[\mathcal{A}\varphi_n(x,\zeta) + \bigl(c(x,\zeta) - \lambda^*(\mathcal{G})\bigr)\varphi_n(x)\bigr] = -\alpha_n\,\mathbb{1}_{\mathcal{B}}(x) \quad \text{a.e. } x\in B_n, \tag{5.2}$$

with $\varphi_n = 0$ on ∂B_n, has a unique positive solution $\varphi_n \in \mathcal{W}^{2,p}_{\mathrm{loc}}(B_n)\cap C(\bar{B}_n)$, for any $p\geqslant 1$, by [13, theorem 2.3 (ii)] (see also [34, theorem 1.1 (ii)]). It is also easy to

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

see that $\varphi_n(0)$ is a continuous increasing function of the constant α_n. We choose the constant α_n as follows: if $\tilde\alpha_n > 0$ is such that the solution φ_n of (5.2) with $\alpha_n = \tilde\alpha_n$ satisfies $\varphi_n(0) = 1$, we set $\alpha_n = \min(1, \tilde\alpha_n)$. Passing to the limit as $n \to \infty$ along some subsequence in (5.2), we obtain a positive solution $\Psi_* \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, for $p > d$, of

$$\min_{\zeta\in\mathbb{U}} \left[\mathcal{A}\Psi_*(x,\zeta) + \big(c(x,\zeta) - \lambda^*(\mathcal{G})\big)\,\Psi_*(x)\right] = -\alpha\,\mathbb{1}_{\mathcal{B}}(x), \qquad x \in \mathbb{R}^d, \tag{5.3}$$

for some $\alpha \geqslant 0$. Writing (5.3) as

$$\min_{\zeta\in\mathbb{U}} \left[\mathcal{A}\Psi_*(x,\zeta) + \left(c(x,\zeta) + \alpha\tfrac{\mathbb{1}_{\mathcal{B}}(x)}{\Psi_*(x)} - \lambda^*(\mathcal{G})\right)\Psi_*(x)\right] = 0, \qquad x \in \mathbb{R}^d,$$

it is clear that the strict right-monotonicity of $\lambda^*(\mathcal{G}, c)$ at c implies that $\alpha = 0$. Thus Ψ_* solves (5.1), and is in $C^2(\mathbb{R}^d)$ by elliptic regularity [**23**, theorem 9.19].

If $\bar v \in \overline{\mathfrak{U}}_{\mathrm{sm}}$, then we have

$$\mathcal{A}^c_{\bar v}\Psi_* = \lambda^*(\mathcal{G})\Psi_* \quad \text{a.e. on } \mathbb{R}^d. \tag{5.4}$$

It follows from the definition of the principal eigenvalue $\lambda^*(\mathcal{G})$ that $\lambda^*_v = \lambda^*(\mathcal{A}^c_v) \geqslant \lambda^*(\mathcal{G})$ for any $v \in \mathfrak{U}_{\mathrm{sm}}$. Therefore, $\lambda^*_{\bar v} = \lambda^*(\mathcal{G})$ for any $\bar v \in \overline{\mathfrak{U}}_{\mathrm{sm}}$ by (5.4). This shows that $\Lambda^* = \lambda^*(\mathcal{G})$ and that $\overline{\mathfrak{U}}_{\mathrm{sm}} \subset \mathcal{U}^*$.

Let $v \in \mathcal{U}^*$. By Itô's formula applied to (5.3), with $\check\tau_r = \tau(\mathcal{B}^c_r)$ as defined in § 1.1 for $r < n$, we obtain

$$\begin{aligned}\varphi_n(x) &\leqslant \mathbb{E}^v_x\left[\mathrm{e}^{\int_0^{\check\tau_r}[c_v(X_s)-\Lambda^*]\,\mathrm{d}s}\,\varphi_n(X_{\check\tau_r})\,\mathbb{1}_{\{\check\tau_r<T\wedge\tau_n\}}\right]\\ &\quad + \mathbb{E}^v_x\left[\mathrm{e}^{\int_0^{T}[c_v(X_s)-\Lambda^*]\,\mathrm{d}s}\,\varphi_n(X_T)\,\mathbb{1}_{\{T<\check\tau_r\wedge\tau_n\}}\right] \qquad \forall\, x \in \mathcal{B}_n \setminus \mathcal{B}_r,\ \forall\, T > 0.\end{aligned}$$

Then we use the argument in the proof of [**4**, lemma 2.11], by replacing f with c_v, and $\Lambda(f)$ with Λ^* in [**4**, (2.34)–(2.36)] to obtain

$$\Psi_*(x) \leqslant \mathbb{E}^v_x\left[\mathrm{e}^{\int_0^{\check\tau_r}[c_v(X_s)-\Lambda^*]\,\mathrm{d}s}\,\Psi_*(X_{\check\tau_r})\,\mathbb{1}_{\{\check\tau_r<\infty\}}\right] \qquad \forall\, x \in \mathcal{B}^c_r, \tag{5.5}$$

and for all $v \in \mathcal{U}^*$. We alert the reader to the fact that this part of the argument in the proof of [**4**, lemma 2.11] does not rely on the near-monotonicity of c.

Next, suppose $v^* \in \mathcal{U}^*$, and let $\Phi_* \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, $p > d$, be an eigenfunction for the operator $\mathcal{A}^c_{v^*}$ corresponding to the eigenvalue $\lambda^*_{v^*} = \Lambda^*$, and such that $\Phi_*(0) = 1$. It also follows from definition 2.7 that the strict monotonicity of $\lambda^*(\mathcal{G}, c)$ at c on the right implies the same property for the eigenvalue $\lambda^*_{v^*}$ of the linear operator $\mathcal{A}^c_{v^*}$ for any $v^* \in \mathcal{U}^*$. This implies that $\lambda^*_{v^*}$ is a simple eigenvalue of the linear operator $\mathcal{A}^c_{v^*}$ for $v^* \in \mathcal{U}^*$, and that Φ_* has the stochastic representation

$$\Phi_*(x) \;=\; \mathbb{E}^{v^*}_x\left[\mathrm{e}^{\int_0^{\check\tau_r}[c_{v^*}(X_s)-\Lambda^*]\,\mathrm{d}s}\,\Phi_*(X_{\check\tau})\,\mathbb{1}_{\{\check\tau_r<\infty\}}\right] \qquad \forall\, x \in \bar{\mathcal{B}}^c_r, \quad \forall\, r > 0, \tag{5.6}$$

by [**8**, theorem 2.1]. A standard application of the strong maximum principle as in the proof of [**4**, lemma 3.5] using (5.5) and (5.6) then shows that $\Phi_* = \Psi_*$. This shows that $\mathcal{U}^* \subset \overline{\mathfrak{U}}_{\mathrm{sm}}$, and thus we must have equality. Uniqueness of the solution

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

Ψ_* of (5.1) clearly then follows from the simplicity of $\lambda^*_{v^*}$ for any $v^* \in \mathcal{U}^*$. This completes the proof. □

Let $\mathcal{M}_*$ denote the class of invariant probability measures, corresponding to the ground state processes with $v \in \overline{\mathfrak{U}}_{\rm sm}$. More precisely, $\mathcal{M}_*$ is the collection of the invariant probability measures corresponding to the extended generators

$$\widetilde{\mathcal{L}}_v f(x) \coloneqq \operatorname{Tr}\big(a(x)\nabla^2 f(x)\big) + \big(b\big(x, v(x)\big) + 2a(x)\nabla \log \Psi_*(x)\big) \cdot \nabla f(x), \quad v \in \overline{\mathfrak{U}}_{\rm sm}.$$

Note that for $\mathcal{M}_* \neq \varnothing$, it is sufficient that the eigenvalue λ^*_v of the operator $\mathcal{A}^c_v$ be strictly monotone at c_v for some $v \in \overline{\mathfrak{U}}_{\rm sm}$ [**8**, theorem 2.2].

The following assumption is enforced throughout this section, without further mention.

Assumption 5.3. The eigenvalue $\lambda^*(\mathcal{G})$ of the semilinear operator $\mathcal{G}$ is strictly monotone at c on the right, and $\mathcal{M}_* \neq \varnothing$.

We let $\mathbf{\Psi}_v$ denote the set of eigenfunctions Ψ_v obtained as limits of Dirichlet eigenvalue problems for the operator $\mathcal{A}^c_v$, normalized so that $\Psi_v(0) = 1$, and for $\rho \in \mathbb{R}$, we define the space of functions

$$\mathbf{\Psi}(\rho) \coloneqq \big\{V \in \mathbf{\Psi}_v : \lambda^*_v \leqslant \rho,\ v \in \mathfrak{U}_{\rm sm}\big\},$$

that is, the set of eigenfunctions corresponding to eigenvalues not exceeding ρ. It is clear that $\mathbf{\Psi}(\rho) = \varnothing$ if $\rho < \lambda^*(\mathcal{G})$, and, under assumption 5.3, $\mathbf{\Psi}\big(\lambda^*(\mathcal{G})\big) = \{\Psi_*\}$ by theorem 5.2.

We let

$$\mathcal{U}(\rho) \coloneqq \{v \in \mathfrak{U}_{\rm sm} : \lambda^*_v \leqslant \rho\}.$$

We need the following definition.

Definition 5.4. Let $\rho \in (\lambda^*(\mathcal{G}), \infty)$.

(i) We say that ρ has *Property A* if

$$\int_{\mathbb{R}^d} \left(\sup_{V \in \mathbf{\Psi}(\rho)} \frac{V(x)}{\Psi_*(x)} \right) \mu(\mathrm{d}x) < \infty \tag{5.7}$$

for some $\mu \in \mathcal{M}_*$.

(ii) We say that ρ has *Property B* if the ground state diffusions with generators $\widetilde{\mathcal{L}}_v$, $v \in \mathcal{U}(\rho)$, are positive recurrent and the corresponding invariant probability measures $\big\{\tilde{\mu}_v : v \in \mathcal{U}(\rho)\big\}$ are tight.

Property B implies that $\mathbf{\Psi}_v$ is a singleton for all $v \in \mathcal{U}(\rho)$ [**8**, lemma 2.7 and theorem 2.3]. Recall assumption 3.2. We have the following convergence result.

Lemma 5.5. *Suppose that $\rho \in (\lambda^*(\mathcal{G}), \infty)$ has Property B. Then, if the control $v_0 \in \mathfrak{U}_{\rm sm}$ in the initialization of assumption 3.2 is such that $\lambda^*_{v_0} \leqslant \rho$, then $\lambda_k \to \hat{\lambda} \in$*

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

$[\lambda^*(\mathcal{G}), \rho]$, *and the sequence* $\{V_k\}$ *converges uniformly on compact sets to some* $\Phi \in C^2(\mathbb{R}^d)$ *satisfying*

$$\min_{\zeta \in \mathbb{U}} \mathcal{A}^c \Phi(x, \zeta) = \hat{\lambda} \Phi(x). \tag{5.8}$$

Proof. We follow the proof of lemma 3.7. It is clear that $\lambda_k \leqslant \lambda_{k-1}$ for all $k \in \mathbb{N}$, so that λ_k converges to some $\hat{\lambda} \in [\lambda^*(\mathcal{G}), \rho]$. Using (3.18) in (3.16), and evaluating at $x = 0$, and $T = \infty$, we have

$$\int_0^\infty \mathrm{e}^{-\tilde{\lambda}_k t} \, \widetilde{\mathbb{E}}_0^k \big[h_k(Y_t^k) \big] \mathrm{d}t \leqslant 1, \tag{5.9}$$

and we know that $\tilde{\lambda}_k \searrow 0$. Let $R > 0$, which is used in the proof as a parameter. Define

$$J_k(x) \coloneqq \int_0^\infty \mathrm{e}^{-\tilde{\lambda}_k t} \, \widetilde{\mathbb{E}}_0^{v_k} \big[h_k(Y_t) \, \mathbb{1}_{\mathcal{B}_R}(Y_t) \big] \mathrm{d}t.$$

Let $\tilde{\mu}_k$ denote the invariant measure of the ground state process $\{Y_t^k\}_{t \geqslant 0}$. These are tight by the hypothesis in Property B. Thus, by invariance, the Harnack property of the densities of the invariant measures together with tightness, and the fact that h_k is bounded on each ball $\mathcal{B}_R$ uniformly in k, there exists a positive constant ϵ_1 depending only on $R > 0$, such that

$$\int_{\mathbb{R}^d} \tilde{\lambda}_k J_k(x) \, \tilde{\mu}_k(\mathrm{d}x) = \int_{\mathcal{B}_R} h_k(x) \, \tilde{\mu}_k(\mathrm{d}x) \geqslant \epsilon_1 \| h_k \|_{L^1(\mathcal{B}_R)}. \tag{5.10}$$

It is well known that the supremum of $J_k(x)$ on $\mathbb{R}^d$ is realized at $\bar{\mathcal{B}}_R$ [**9**, lemma 3.6.1]. Therefore

$$\int_{\mathbb{R}^d} J_k(x) \, \tilde{\mu}_k(\mathrm{d}x) \leqslant \sup_{\mathcal{B}_R} J_k. \tag{5.11}$$

As shown in [**9**, (3.7.24)], by employing [**11**]*corollary 2.2, the resolvent J_k has the Harnack property. Thus, there exists a constant C_{H}, independent of k, such that

$$C_{\mathrm{H}} \inf_{\mathcal{B}_R} J_k \geqslant \sup_{\mathcal{B}_R} J_k. \tag{5.12}$$

Combining (5.10)–(5.12), we have

$$\begin{aligned} \| h_k \|_{L^1(\mathcal{B}_R)} &\leqslant \frac{\tilde{\lambda}_k}{\epsilon_1} \sup_{\mathcal{B}_R} J_k \\ &\leqslant \frac{\tilde{\lambda}_k}{\epsilon_1} C_{\mathrm{H}} \inf_{\mathcal{B}_R} J_k \\ &\leqslant \frac{\tilde{\lambda}_k}{\epsilon_1} C_{\mathrm{H}} \, J_k(0) \\ &\leqslant \frac{\tilde{\lambda}_k}{\epsilon_1} C_{\mathrm{H}}, \end{aligned} \tag{5.13}$$

where in the last inequality we use (5.9). Since $\tilde{\lambda}_k \to 0$ as $k \to \infty$, it follows from (5.13) that h_k converges to 0 in L^1 on every ball $\mathcal{B}_R$, and thus also converges in L^p

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

for any $p \geqslant 1$, since it is uniformly bounded on each ball. The rest follows exactly as in lemma 3.7. □

LEMMA 5.6. *In addition to the hypotheses of lemma 5.5, suppose that* $\rho \in (\lambda^*(\mathcal{G}), \infty)$ *has Property A. Then,* $(\hat{\lambda}, \Phi) = (\lambda^*(\mathcal{G}), \Psi_*)$ *is the unique solution of (5.8).*

Proof. Suppose that $\Phi \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$, $p > d$, solves (5.8) for some $\hat{\lambda} \in [\lambda^*(\mathcal{G}), \rho]$. Since Φ is a limit of a sequence V_k of assumption 3.2, it is clear that $\Phi \leqslant \sup_{V \in \boldsymbol{\Psi}(\rho)} V$. Let $v \in \overline{\mathfrak{U}}_{\mathrm{sm}}$ be such that the corresponding ground process $\{Y_t\}_{t \geqslant 0}$ has an invariant probability measure $\tilde{\mu}_v \in \mathcal{M}_*$, satisfying (5.7). We have

$$\widetilde{\mathcal{L}}_v \left(\frac{\Phi}{\Psi_*} \right) \geqslant \big(\hat{\lambda} - \lambda^*(\mathcal{G}) \big) (\Phi / \Psi_*).$$

Therefore, $\{(\Phi/\Psi_*)(Y_t)\}_{t \geqslant 0}$ is a nonnegative submartingale, and is integrable under the invariant probability measure $\tilde{\mu}_v$. Then since $\widetilde{\mathbb{E}}^v_x \big[(\Phi/\Psi_*)(Y_t) \big] \geqslant (\Phi/\Psi_*)(x)$, we obtain $(\Phi/\Psi_*)(x) \leqslant \tilde{\mu}_v \big((\Phi/\Psi_*) \big)$, so that (Φ/Ψ_*) is bounded. Thus $\big\{ (\Phi/\Psi_*)(Y_t) \big\}_{t \geqslant 0}$ converges a.s., and since the process is recurrent it must converge to a constant. This shows that $\Phi = \kappa \Psi_*$ for some $\kappa > 0$. In turn, this implies that $\hat{\lambda} = \lambda^*(\mathcal{G})$. □

We are ready to state the main convergence result.

THEOREM 5.7. *Grant assumption 5.3, and suppose that* $\rho \in (\lambda^*(\mathcal{G}), \infty)$ *has properties A and B. Then, provided that the control* $v_0 \in \mathfrak{U}_{\mathrm{sm}}$ *in the initialization of assumption 3.2 is such that* $\lambda^*_{v_0} \leqslant \rho$*, the following hold:*

(a) $\lambda_k < \lambda_{k-1}$ *for all* $k \in \mathbb{N}$*, unless* $\lambda_k = \lambda^*(\mathcal{G})$.

(b) *The sequence* $\{\lambda_k\}$ *converges to* $\lambda^*(\mathcal{G})$ *as* $k \to \infty$.

(c) *The sequence* $\{V_k\}$ *converges as* $k \to \infty$*, uniformly on compact sets, to* Ψ_* *satisfying (5.1).*

Proof. The assertions (a)–(c) follow as in the proof of theorem 5.10, using lemmas 5.5, 5.6. □

REMARK 5.8. Theorem 5.7 describes a region of 'stability' of the PIA using the abstract properties in definition 5.4. We have not considered in this section the equality $\lambda^*(\mathcal{G}) = \mathcal{E}^*_{\mathrm{min}}$. This is considered in § 5.1 which follows next.

5.1. The minimization problem for near-monotone running costs

We borrow the results in [**7**, Proposition 5.1] which addresses the near-monotone running cost case without imposing any blanket stability assumptions, and improves [**4**, Proposition 1.1].

In general, we say that a function $f \colon \mathcal{X} \to \mathbb{R}$ defined on a locally compact space is *coercive, or near-monotone, relative to a constant* $\beta \in \mathbb{R}$ if there exists a compact set K such that $\inf_{K^c} f > \beta$.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

We start with the hypotheses in [**7**, § 5] which we quote here as follows.

ASSUMPTION 5.9. In addition to the regularity hypotheses on the data in (A1)–(A3), we assume the following:

(i) The drift b and running cost c satisfy, for some $\theta \in [0,1)$ and a constant κ_0, the bound

$$|b(x,\zeta)| \leqslant \kappa_0\bigl(1+|x|^{\theta}\bigr), \quad \text{and} \quad |c(x,\zeta)| \leqslant \kappa_0\bigl(1+|x|^{2\theta}\bigr) \quad \forall\, (x,\zeta) \in \mathbb{R}^d \times \mathbb{U}.$$

(ii) The drift b satisfies

$$\frac{1}{|x|^{1-\theta}} \max_{\zeta\in\mathbb{U}} \bigl\langle b(x,\zeta),\, x\bigr\rangle^{+} \xrightarrow[|x|\to\infty]{} 0.$$

(iii) The running cost is coercive relative to $\mathcal{E}^*_{\min}$, that is,

$$\mathcal{E}^*_{\min} < \lim_{r\to\infty} \min_{(x,\zeta)\in\mathcal{B}_r^c\times\mathbb{U}} c(x,\zeta).$$

For a full characterization of optimality under the above hypotheses see [**7**, Proposition 5.1]. We state this here in the following form.

THEOREM 5.10. *Grant assumptions 5.3, 5.9. Then there exists a unique solution $V \in C^2(\mathbb{R}^d)$ of*

$$\min_{\zeta\in\mathbb{U}} \mathcal{A}^c V(x,\zeta) = \mathcal{E}^*_{\min} V(x) \quad \text{in } \mathbb{R}^d, \quad \text{and} \quad V(0)=1.$$

We also have $\mathcal{E}^*_{\min} = \lambda^*(\mathcal{G})$, and (a)–(c) of theorem 3.1 hold. In addition, if c is near-monotone with respect to λ^*_v for some $v \in \mathfrak{U}_{\mathrm{sm}}$, then $\mathcal{E}_x(c,v) = \lambda^*_v$, and there exists an inf-compact eigenfunction $\Psi_v \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$ for λ^*_v. In particular, the diffusion controlled by such a v is exponentially ergodic.

We now state the results concerning the PIA algorithm for this model, whose proof follows directly from theorems 5.7, 5.10.

THEOREM 5.11. *Grant assumptions 5.3, 5.9, suppose that c is near-monotone relative to $\rho \in (\lambda^*(\mathcal{G}),\infty)$, and the latter has Properties A and B in definition 5.4. Then the conclusions of theorem 5.7 follow.*

REMARK 5.12. We compare the assumptions in theorem 5.11 to the ones in [**19**, theorem 5.4]. Assumption 5.9 is used to guarantee the existence of a solution to the HJB equation, together with the standard verification of optimality results, in theorem 5.10, so as to provide a complete set of results for this model. As remarked in that paper, [**19**, theorem 3.6] does not assert the existence of a solution to the dynamic programming equation, but rather a dynamic programming inequality. Existence of a solution is imposed as an assumption in the study of the PIA. Then, assumption 5.3 agrees with [**19**, (A4)], Property A matches [**19**, theorem 5.4 (i)], and Property B is essentially the same as [**19**, theorem 5.4 (ii)].

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

Acknowledgments

We wish to thank the anonymous referee for the careful reading of the manuscript and proposed improvements. The research of Ari Arapostathis was supported in part by the National Science Foundation through grant DMS-1715210, and in part by the Army Research Office through grant W911NF-17-1-001. The research of Anup Biswas was supported in part by DST-SERB grants EMR/2016/004810, and MTR/2018/000028.

References

1 S. Anulova, H. Mai and A. Veretennikov. On averaged expected cost control as reliability for 1d ergodic diffusions. *Reliab. Theory Appl.* **12** (2017), 31–38.

2 S. Anulova, H. Mai and A. Veretennikov. Yet again on iteration improvement for averaged expected cost control for 1d ergodic diffusions. *ArXiv e-prints*, 1812.10665, 2018.

3 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.*, eds. Daniel Hernández-Hernández and J. Adolfo Minjárez-Sosa, pp. 1–12 (New York: Birkhäuser/Springer, 2012). doi:10.1007/978-0-8176-8337-5_1.

4 A. Arapostathis and A. Biswas. Infinite horizon risk-sensitive control of diffusions without any blanket stability assumptions. *Stochastic Process. Appl.* **128** (2018), 1485–1524, ISSN0304-4149. doi:10.1016/j.spa.2017.08.001.

5 A. Arapostathis and A. Biswas. Risk-sensitive control for a class of diffusions with jumps. *ArXiv e-prints*, 1910.05004, 2019.

6 A. Arapostathis, A. Biswas. A variational formula for risk-sensitive control of diffusions in $\mathbb{R}^d$. *SIAM J. Control Optim.* **58** (2020), 85–103, ISSN 0363-0129. doi:10.1137/18M1218704.

7 A. Arapostathis, A. Biswas, V. S. Borkar and K. Suresh Kumar. A variational characterization of the risk-sensitive average reward for controlled diffusions in $\mathbb{R}^d$. *ArXiv e-prints*, 1903.08346, 2019a.

8 A. Arapostathis, A. Biswas and S. Saha. Strict monotonicity of principal eigenvalues of elliptic operators in $\mathbb{R}^d$ and risk-sensitive control. *J. Math. Pures Appl.* **124** 2019b, 169–219. doi:10.1016/j.matpur.2018.05.008.

9 A. Arapostathis. V. S. Borkar and M. K. Ghosh. *Ergodic control of diffusion processes.* Encyclopedia of Mathematics and its Applications vol. 143. (Cambridge: Cambridge University Press, 2012).

10 A. Arapostathis, L. Caffarelli, G. Pang and Y. Zheng. Ergodic control of a class of jump diffusions with finite Lévy measures and rough kernels. *SIAM J. Control Optim.* **57** (2019c), 1516–1540. ISSN0363-0129. doi:10.1137/18M1166717.

11 A. Arapostathis, M. K. Ghosh and S. I. Marcus. Harnack's inequality for cooperative weakly coupled elliptic systems. *Comm. Partial Differ. Equ.* **24** (1999), 1555–1571. doi:10.1080/03605309908821475.

12 A. Arapostathis, H. Hmedi and G. Pang. On uniform exponential ergodicity of Markovian multiclass many-server queues in the Halfin–Whitt regime. *Math. Oper. Res. (to appear)*, (2020).

13 S. N. Armstrong. Principal eigenvalues and anti-maximum principle for homogeneous fully nonlinear elliptic equations. *J. Diff. Equ.* **246** (2009), 2958–2987. ISSN 0022-0396. doi:10.1016/j.jde.2008.10.026.

14 H. Berestycki, L. Nirenberg and S. R. S. Varadhan. The principal eigenvalue and maximum principle for second-order elliptic operators in general domains. *Comm. Pure Appl. Math.* **47** (1994), 47–92. ISSN 0010-3640. doi:10.1002/cpa.3160470105.

15 H. Berestycki and L. Rossi. Generalizations and properties of the principal eigenvalue of elliptic operators in unbounded domains. *Comm. Pure Appl. Math.* **68** (2015), 1014–1065. doi:10.1002/cpa.21536.

16 T. R. Bielecki and S. R. Pliska. Risk-sensitive dynamic asset management. *Appl. Math. Optim.* **39** (1999), 337–360. ISSN 0095-4616. doi:10.1007/s002459900110.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press

17 A. Biswas. An eigenvalue approach to the risk sensitive control problem in near monotone case. *Systems Control Lett.* **60** (2011a), 181–184. doi:10.1016/j.sysconle.2010.12.002.

18 A. Biswas. Risk sensitive control of diffusions with small running cost. *Appl. Math. Optim.* **64** (2011b), 1–12. doi:10.1007/s00245-010-9127-4.

19 V. S. Borkar and S. P. Meyn. Risk-sensitive optimal control for Markov decision processes with monotone cost. *Math. Oper. Res.* **27** (2002), 192–209. ISSN 0364-765X. doi:10.1287/moor.27.1.192.334.

20 Y-Z. Chen and L-C. Wu. *Second order elliptic equations and elliptic systems*. Translations of Mathematical Monographs vol. 174. (Providence, RI: American Mathematical Society, 1998). ISBN 0-8218-0970-9. Translated from the 1991 Chinese original by Bei Hu.

21 W. H. Fleming, W. M. McEneaney. Risk-sensitive control on an infinite time horizon. *SIAM J. Control Optim.* **33** (1995), 1881–1915. ISSN 0363-0129. doi:10.1137/S0363012993258720.

22 W. H. Fleming and S. J. Sheu. Risk-sensitive control and an optimal investment model. *Math. Finance*, **10** (2000), 197–213. ISSN 0960-1627. doi:10.1111/1467-9965.00089. INFORMS Applied Probability Conference (Ulm, 1999).

23 D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften*. (Berlin: Springer-Verlag, 1983), 2nd edn. doi:10.1007/978-3-642-61798-0.

24 I. Gyöngy and N. Krylov. Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Theory Relat. Fields* **105** (1996), 143–158. doi:10.1007/BF01203833.

25 H. Kaise and S-J. Sheu. On the structure of solutions of ergodic type Bellman equation related to risk-sensitive control. *Ann. Probab.* **34** (2006), 284–320. doi:10.1214/009117905000000431.

26 N. V. Krylov. *Controlled diffusion processes*. Applications of Mathematics vol. 14. (New York-Berlin: Springer-Verlag, 1980). ISBN 0-387-90461-1.

27 J-L. Menaldi and M. Robin. Remarks on risk-sensitive control problems. *Appl. Math. Optim.* **52** (2005), 297–310. doi:10.1007/s00245-005-0829-y.

28 S. P. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.* **4** (1994), 981–1011.

29 H. Nagai. Bellman equations of risk-sensitive control. *SIAM J. Control Optim.* **34** (1996), 74–101. ISSN 0363-0129. doi:10.1137/S0363012993255302.

30 J. Speyer. An adaptive terminal guidance scheme based on an exponential cost criterion with application to homing missile guidance. *IEEE Trans. Autom. Control*, **21** (1976), 371–375. ISSN 2334-3303. doi:10.1109/TAC.1976.1101206.

31 A. Yu. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S.)* **111** (1980), 434–452, 480. ISSN 0368–8666.

32 A. Yu. Veretennikov. Parabolic equations and stochastic equations of Itô with coefficients that are discontinuous with respect to time. *Mat. Zametki*, **31** (1982), 549–557, 654 . ISSN 0025-567X.

33 P. Whittle. *Risk-sensitive optimal control*. Wiley-Interscience Series in Systems and Optimization. (Chichester: John Wiley & Sons, Ltd., 1990). ISBN 0-471-92622-1.

34 Y. Yoshimura. A note on demi-eigenvalues for uniformly elliptic Isaacs operators. *Viscosity Solution Theory of Differential Equations and its Developments*, pp. 106–114, 2006.

https://doi.org/10.1017/prm.2020.61 Published online by Cambridge University Press