av_multi_control_c.tex

On iteration improvement for averaged control for multidimensional ergodic diffusions

S.V. Anulova[*,†] A.Yu. Veretennikov[‡,§]

December 31, 2019

Abstract

Ergodic Bellman's (HJB) equation is proved for a special case of a uniformly ergodic multidimensional controlled diffusion with variable diffusion and drift coefficients both depending on control; convergence of Howard's iterative reward improvement algorithm to the unique solution of Bellman's equation is established.

1 Introduction

The paper is a continuation of [1] and [2] where the dimension was equal to one. We consider a d-dimensional stochastic differential equation (SDE) on the probability

[*]Institute for Information Transmission Problems, Moscow, Russia; email: anulova @ mail.ru

[†]For the first author this research has been supported by the Russian Foundation for Basic Research grant no. 17-01-00633_a

[‡]University of Leeds, UK, & National Research University Higher School of Economics, & Institute for Information Transmission Problems, Moscow, Russia; email: a.veretennikov @ leeds.ac.uk

[§]The third author is grateful to the financial support by the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications" at Bielefeld University during his stay there in August 2018; also, for this author this study has been funded by the Russian Academic Excellence Project '5-100' and by the Russian Foundation for Basic Research grant no. 17-01-00633_a.

space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ with a one-dimensional $(\mathcal{F}_t)$ Wiener process $B = (B_t)_{t\ge 0}$ with coefficients b and σ, and with a feedback control function α (called strategy in the sequel)

$$\begin{aligned} dX_t^\alpha &= b(\alpha(X_t^\alpha), X_t^\alpha)\, dt + \sigma(\alpha(X_t^\alpha), X_t^\alpha)\, dW_t, \quad t \ge 0, \\ X_0^\alpha &= x, \end{aligned} \tag{1}$$

with the function $\sigma \in \mathbb{R}^1$.

Let a compact set $U \subset \mathbb{R}$ be a set where any strategy takes its values. The functions b and σ on $U \times \mathbb{R}$ are assumed Borel; later on some further conditions will be imposed, but we note straight away that σ will be assumed non-degenerate and that a weak solution of the equation (1) always exists, see [17]. Denote the class of all Borel functions α with values in U by $\mathcal{A}$. For $u \in U$ and $\alpha(\cdot) \in \mathcal{A}$ denote

$$L^u(x) = b(u, x)\nabla_x + \frac{1}{2}\sigma^2(u, x) \sum_i \frac{\partial^2}{\partial x_i^2},$$

and

$$L^\alpha(x) = b(\alpha(x), x)\nabla_x + \frac{1}{2}\sigma^2(\alpha(x), x) \sum_i \frac{\partial^2}{\partial x_i^2}, \quad x \in \mathbb{R}^d.$$

This is not the most general form of the (non-divergent) second order differential operator. However, this restriction on L, namely, the scalar multiplier σ^2 with the Laplacian, makes sense because in most media the Einstein heat coefficient is, indeed, scalar, although, of course, there exist anisotropic media where the scalar description does not suffice.

A bit more general case may be considered,

$$L^u(x) = b(u, x)\nabla_x + \frac{1}{2}\sigma^2(u, x) \sum_{ij} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j},$$

where $a(x) = (a_{ij}(x))$ is a non-degenerate bounded Hölder-continuous matrix which does not depend on control u. An additional restriction will be imposed on the drift b so as to guarantee the uniform recurrence.

Denote by $\mathcal{K}$ the class of function on $U \times \mathbb{R}^d$ (also just on $\mathbb{R}^d$) growing no faster than some polynomial. The *running cost* function f will always be chosen from this class. The *averaged cost* function corresponding to the strategy $\alpha \in \mathcal{A}$ is then defined as

$$\rho^\alpha(x) := \limsup_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}_x f(\alpha(X_t^\alpha), X_t^\alpha)\, dt. \tag{2}$$

For a strategy $\alpha \in \mathcal{A}$ the function $f^\alpha : \mathbb{R}^d \to \mathbb{R},\ f^\alpha(x) = f(\alpha(x), x),\ x \in \mathbb{R}^d$, is defined. Then (2) has an equivalent form

$$\rho^\alpha(x) = \limsup_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}_x f^\alpha(X_t^\alpha)\, dt. \tag{3}$$

Now, the *cost function* for the model under consideration is defined as

$$\rho(x) := \inf_{\alpha\in\mathcal{A}} \limsup_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}_x f^\alpha(X_t^\alpha)\, dt. \tag{4}$$

It will be assumed that for every $\alpha \in \mathcal{A}$ the solution of the equation (1) X^α is Markov ergodic, i.e., there exists a limiting in total variation distribution μ^α of X_t^α, $t \to \infty$, this distribution μ^α does not depend on the initial condition $X_0 = x \in \mathbb{R}^d$, is unique and is invariant for the generator L^α. The cost function ρ^α then does not depend on x and can be rewritten as

$$\rho^\alpha(x) \equiv \rho^\alpha := \int f^\alpha(x)\, \mu^\alpha(dx) =: \langle f^\alpha, \mu^\alpha \rangle. \tag{5}$$

Then what we want to find (compute) is the value

$$\rho := \inf_{\alpha\in\mathcal{A}} \int f^\alpha(x)\, \mu^\alpha(dx) = \inf_{\alpha\in\mathcal{A}} \langle f^\alpha, \mu^\alpha \rangle. \tag{6}$$

For any strategy $\alpha \in \mathcal{A}$ let us also define an auxiliary function

$$v^\alpha(x) := \int_0^\infty \mathbb{E}_x (f^\alpha(X_t^\alpha) - \rho^\alpha)\, dt.$$

The convergence of this integral will follow from the assumptions.

One goal of this paper is to show the *ergodic HJB or Bellman's equation* on tghe pair (V, ρ)

$$\inf_{u\in U} [L^u V(x) + f^u(x) - \rho] = 0, \quad x \in \mathbb{R}^d, \tag{7}$$

where ρ is a constant and $V : \mathbb{R}^d \mapsto \mathbb{R}$, or, equivalently,

$$\inf_{u\in U} \left[\frac{1}{\sigma^2(u,x)} L^u V(x) + \frac{f^u(x)}{\sigma^2(u,x)} - \frac{\rho}{\sigma^2(u,x)}\right] = 0, \quad x \in \mathbb{R}^d, \tag{8}$$

with solution $(V(\cdot), \rho)$, to show uniqueness of the second component and that it coincides with the cost from (6). The meaning of the first component V will be explained later. The uniqueness of V will be shown up to an additive constant.

The class where the solution (V, ρ) will be studied is the family of all Borel functions V and constants $\rho \in \mathbb{R}$ such that V has two Sobolev derivatives locally integrable in any power. Respectively, the equation (7) is to be understood almost everywhere; yet, in the 1D situation and under our assumptions it will follow straightforward that this equation, actually, is satisfied for all $x \in \mathbb{R}$. Note that the first derivative can be considered continuous (due to the embedding theorems), and the second derivative will be always taken Borel, as one of the Borel representatives of Lebesgue's measurable function.

One more goal of the paper is to show how approach the solution ρ of the main problem by some successive approximation procedure called Reward Improvement Algorithm (RIA). It is interesting that under our minimal assumptions on regularity of strategies for the weak SDE solution setting it is yet possible to justify a *monotonic convergence* of the "exact" RIA; compare to [16, ch.1, §4] where it was necessary to work with "approximate" RIA (called Bellman–Howard's iteration procedure there) and with regularized Lipschitz strategies.

Concerning a full uniqueness for the solution of (7), note that with any solution (V, ρ) and for any constant C, the couple $(V + C, \rho)$ is also a solution. There are two close enough options how to tackle this fact: either accept that uniqueness will be established up to a constant, or to choose a certain "natural" constant satisfying some "centering condition" as will be done below.

To guarantee ergodicity, we will assume so called "blanket" recurrence conditions (see below), which provide in some sense a uniform recurrence for *any* strategy. Conditions of this type are sometimes considered as too restrictive; however, they do allow to include models and cases not covered earlier in this theory and by this reason we regard this restriction as a reasonable price for the time being. It is likely that such restrictions may be relaxed so as to include the "near monotonicity" type conditions (cf. [6]).

Let us say just a few words about the history of the problem. More can be found in the references provided below. Earlier results on ergodic control in continuous time were obtained in [21], [24], [7], et al. In his book [21] Mandl established apparently first results on ergodic (averaged) control for controlled 1D diffusion on a finite interval with boundary conditions including jumps from the boundary. He established the HJB equation and proved uniqueness of the couple (up to a constant for the first component). Improvement of control was discussed, too, however, without convergence.

Discrete time controlled models were considered in the monographs [8], [12], [13], [26], and others, and in the papers [3], [22], [27], etc. Note that the paper [22] contains a brief section on continuous time models as well; however, it does not touch technical difficulties which we deal with in the present paper.

Continuous time controlled processes were also treated in the 80s in a chapter of the monograph [7] where ergodic control for stable diffusions was considered. Arapostathis and Borkar [5], Arapostathis [4], Arapostathis, Borkar and Ghosh [6] treated diffusion with the "relaxed control" and with the diffusion coefficient not depending on the control, under weaker recurrence assumptions (i.e., under two types of condition, stable or near-monotone). In this setting, they establish Bellman's equation, existence, uniqueness, and RIA convergence. In the present paper we allow diffusion coefficient to depend on control, even though in a reduced form, and we do not use relaxed control.

The latest works include [4], [6], [27], see also the references therein. In the very first papers and books compact cases with some auxiliary boundary conditions – so as to simplify ergodicity – were studied; convergence of the improvement control algorithms were studied only partially. In the later investigations noncompact spaces are allowed; however, apparently, *ergodic* control in the diffusion coefficient σ of the process was not tackled earlier. About controlled diffusion processes on a finite horizon, or on infinite horizon with discount (technically equivalent to killing) the reader may consult [7] and [16].

The paper consists of three sections: 1 – Introduction, 2 – Assumptions and some auxiliaries, 3 – Main result and its proof. We will be using the convention that arbitrary constants C in the calculus may change from line to line.

2 Assumptions and some auxiliaries

To ensure ergodicity of X^α under any feedback control strategy $\alpha \in \mathcal{A}$, we make the following assumptions on the drift and diffusion coefficients.

(A1) (boundedness, non-degeneracy, continuity) The functions b and σ are Borel measurable and bounded in their variables; $|b(u,x)| \le C_b$, $|\sigma(u,x)| \le C_\sigma$, σ is uniformly non-degenerate, $|\sigma(u,x)|^{-1} \le C_\sigma$; the functions $\sigma(u,x)$ and $f^u(x)$ are continuous in u for every x.

(A2) (recurrence)

$$\lim_{|x|\to\infty} \sup_u \langle x, b(u,x)\rangle = -\infty. \tag{9}$$

(A3) (running cost) The function f belongs to the class $\mathcal{K}$ of functions which are Borel measurable in x for each u and admit a uniform in u polynomial bound: there exist constants $C_1, m_1 > 0$ such that for any x,

$$\sup_{u\in U} |f^u(x)| \le C_1(1+|x|^{m_1}).$$

(A4) (compactness of U) The set U is compact.

We will need the following three lemmata and two corollaries from them.

Lemma 1. *Let the assumptions (A1) – (A3) hold true. Then*

- *For any $C_1, m_1 > 0$ there exist $C, m > 0$ such that for any strategy $\alpha \in \mathcal{A}$ and for any function g growing no faster than $C_1(1+|x|^{m_1})$,*

$$\sup_{t\ge 0} |\mathbb{E}_x g(X_t^\alpha)| \le C(1+|x|^m). \tag{10}$$

- *For any $\alpha \in \mathcal{A}$, the (unique) invariant measure μ^α integrates any polynomial and*

$$\sup_{\alpha\in\mathcal{A}} \int |x|^k \, \mu^\alpha(dx) < \infty, \quad \forall\, k > 0. \tag{11}$$

- *For any strategy $\alpha \in \mathcal{A}$ the function ρ^α is a constant, and*

$$\sup_{\alpha\in\mathcal{A}} |\rho^\alpha| \le C < \infty; \tag{12}$$

moreover, for any $k > 0$ and $f \in \mathcal{K}$ there exist $C, k_0 > 0$ such that

$$\sup_{\alpha\in\mathcal{A}} |\mathbb{E}_x f^\alpha(X_t^\alpha) - \rho^\alpha| \le C\frac{1+|x|^{k_0}}{1+t^k}, \tag{13}$$

and

$$\sup_{\alpha\in\mathcal{A}} \left|\frac{1}{T}\int_0^T \mathbb{E}_x f^\alpha(X_t^\alpha)\, dt - \rho^\alpha\right| \to 0, \quad T\to\infty. \tag{14}$$

Proof. Follows from the calculus similar to that in [29] and [25].

Corollary 1. *Under the same assumptions,*

$$\sup_{t\ge 0} \mathbb{E}_x 1(|X_t^\alpha| > N)| \le \sup_{t\ge 0} \mathbb{E}_x \frac{|X_t^\alpha|^m}{N^m} \le \frac{C(1+|x|^m)}{N^m}. \tag{15}$$

Proof is straightforward by Bienaymé – Chebyshev – Markov's inequality.

Lemma 2. *Let the assumptions (A1) – (A3) be satisfied. Then for any strategy $\alpha \in \mathcal{A}$ the cost function v^α has the following properties:*

1. *The function v^α is continuous as well as (∇v^α), and there exist $C, m > 0$ both depending only on the constants in (A1)–(A3) such that*

$$\sup_\alpha (|v^\alpha(x)| + |\nabla v^\alpha(x)|) + |\Delta v^\alpha(x)|) \le C(1+|x|^m). \tag{16}$$

2. *$v^\alpha \in W^2_{p,loc}$ for any $p \ge 1$.*

3. *$v^\alpha \in C^{1,Lip}$ (i.e., (∇v^α) is locally Lipschitz).*

4. *v^α satisfies a Poisson equation in the whole space,*

$$L^\alpha v^\alpha + f^\alpha - \langle f^\alpha, \mu^\alpha \rangle = 0, \tag{17}$$

in the Sobolev sense; in particular, for almost every $x \in \mathbb{R}$

$$L^\alpha(x) v^\alpha(x) + f^\alpha(x) - \langle f^\alpha, \mu^\alpha \rangle = 0, \tag{18}$$

or, equivalently,

$$\frac{1}{\sigma^2(\alpha)} \left(L^\alpha v^\alpha + f^\alpha - \langle f^\alpha, \mu^\alpha \rangle \right) = 0. \tag{19}$$

5. *Solution of the equation (17) is unique up to an additive constant in the class of Sobolev solutions $W^2_{p,loc}$ with any $p > 1$ with a no more than some (any) polynomial growth of the solution v^α.*

6. *$\langle v^\alpha, \mu^\alpha \rangle = 0$.*

Proof. Denote

$$\bar{L}^u = \frac{1}{\sigma^2(u,x,z)} L^u, \quad \bar{L}^\alpha = \frac{1}{\sigma^2(\alpha(x,z),x,z)} L^\alpha,$$

Let $t'(t)$ denote the inverse function for the mapping $t \mapsto \displaystyle\int_0^t \sigma^2(\alpha(X_s^\alpha), X_s^\alpha)\, ds$, and let

$$\bar X_t := X_{t'(t)}.$$

This random time change $t \mapsto t'(t)$ – see, for example, ([9, Theorem 15.5]), – tells us that the process $\bar X_t^\alpha$ satisfies the following SDE

$$d\bar X_t^\alpha = \bar b(\alpha(\bar X_t), \bar X_t)\, dt + d\bar W_t, \quad \bar X_0^\alpha = x,$$

with a new Wiener process $\bar W_t = \displaystyle\int_0^{t'(t)} \sigma(\alpha(X_s^\alpha), X_s^\alpha)\, dW_s$ and with $\bar b^\alpha(x) = b^\alpha(x)\sigma^{-2}(\alpha(x), x)$. The process $(\bar X^\alpha)$ is unique in distribution (more than that, it is, actually, also pathwise unique), ergodic with a unique invariant measure $\bar\mu^\alpha(dx)$ and there is a convergence better than any polynomial to it in total variation. In fact, the Lemma 1 is applicable straighforwarrdly to the process $\bar X$, so that the statements similar to (10) – (14) hold true for the process $\bar X$ as well.

This time change allows to rewrite the definition of v^α as follows,

$$v^\alpha(x) = \int_0^\infty E_x(f^\alpha(X_t^\alpha) - \rho^\alpha)\, dt = \int_0^\infty E_x \bar f^\alpha(\bar X_t^\alpha)\, dt, \tag{20}$$

with

$$\bar f^\alpha(x) = \frac{f^\alpha(x) - \rho^\alpha}{\sigma^2(\alpha(x), x)}.$$

Hence, the only option for the integral in (20) to converge is that $\bar f^\alpha$ satisfies a centering condition

$$\langle \bar f^\alpha, \bar\mu^\alpha \rangle = 0.$$

Now it follows from (20) in a standard way (see, for example, [25]) that the function v^α is a solution of the Poisson equation

$$\bar L^\alpha v^\alpha + \bar f^\alpha = 0; \tag{21}$$

recall that here

$$\bar L^u g(x) := \bar b(u, x)\nabla g(x) + \frac{1}{2}\Delta g(x), \quad x \in \mathbb{R}^d,$$

and

$$\bar L^\alpha g(x) := \bar b(\alpha(x), x)\nabla g(x) + \frac{1}{2}\Delta g(x).$$

Let us show existence of derivatives v_x^α and v_{xx}^α in a Sobolev sense in the spaces $W^2_{p,loc}$ with any $p>1$ (see notations in [19]). For this aim let us denote

$$v^{(T)}(s,x,z) := \int_0^{T-s} \mathbb{E}_{x,z}\bar f^\alpha(\bar X_t^\alpha)\,dt, \qquad 0\le s\le T,$$

for any $T>0$. In fact, it turns out to be a bit easier to prove a formally stronger claim than just existence of v_x^α and v_{xx}^α, namely, that

$$(v_x^{(T)}, v_{xx}^{(T)})|_{t=0} \underset{T\to\infty}{\to} (v_x^\alpha, v_{xx}^\alpha) \quad \text{in the } L_{p,loc} \text{ sense},$$

or, a bit more precisely, that the left hand side here converges in $L_{p,loc}$, and the limit turns out to be $(v_x^\alpha, v_{xx}^\alpha)$. The functions v^T have been introduced for this aim. In the sequel the following short notations will be used: $\bar T_t \bar f^\alpha(x) := \mathbb{E}_x(\bar f^\alpha(\bar X_t^\alpha))$. (On the one hand, there is a slight abuse of notations here: T is time and $\bar T_t$ is a semigroup; on the other hand, the semigroup is never used without a lower index, so there is no danger to get one for the other.) We have,

$$\begin{aligned}
v_x^{(T)}(0,x) = \partial_x \int_0^T \bar T_t \bar f^\alpha(x))dt &= \partial_x \int_0^1 \bar T_t \bar f^\alpha(x)\,dt + \partial_x \int_1^T \bar T_t \bar f^\alpha(x))dt \\
&= \partial_x \int_0^1 \bar T_t \bar f^\alpha(x)dt + \partial_x \int_1^T \bar T_1 \bar T_{t-1}\bar f^\alpha(x)\,dt \\
&= \partial_x \int_0^1 \bar T_t \bar f^\alpha(x)dt + \int_1^T \partial_x \bar T_1(\bar T_{t-1}\bar f^\alpha)(x)dt. \qquad (22)
\end{aligned}$$

The first term here does not change with time (as far as $T\ge 1$) and is well-defined in $L_{p,loc}$ Sobolev sense along with $\partial_x^2 \displaystyle\int_0^1 \bar T_t \bar f^\alpha(x,z))dt$, due to [28, Theorem 5.5, 5.7].

The integrand in the second term admits the bound (see [28, Theorems 5.7 & 5.5] with $T_1=\epsilon>0$),

$$\|\bar T_\cdot(\bar T_{t-1}\bar f^\alpha)\|_{W_p^2((\epsilon,1)\times B_R)} \le C_\epsilon\big(\|\bar T_\cdot(\bar T_{t-1}\bar f^\alpha)\|_{L_p((\epsilon,1)\times B_{R+1})}\big)$$

It follows from ergodic bounds of the SDE solutions similar to those in [29] that the right hand side here decreases to zero faster than any polynomial in time, and, hence, the second integral in the representation (22) converges as $T\to\infty$; this convergence is locally uniform with respect to the initial value x and, of course, uniform with respect to z, as the latter variable takes values from a finite set. The same is also true for any partial derivative of the second order $\left(\partial_x^2 \int_1^\infty \bar T_t \bar f^\alpha(x))dt\right)$. Thus, we obtain,

$$\nabla_x \int_0^T \bar T_t \bar f^\alpha(x))dt \to \bar v_1(x), \quad \partial_x^2 \int_0^T \bar T_t \bar f^\alpha(x))dt \to \bar v_2(x)$$

as $T \to \infty$, both locally uniformly in x and, hence, also in the $L_{p,loc}$ sense. In a standard manner by integration over x it follows that $\bar v_1(x)$ and $\bar v_2(x)$ serve as ∇v^α and $(v^\alpha_{x_i x_j})$, respectively. Namely, for any x_1, x_1' we have by the first theorem of the calculus (also known as Newton – Leibnitz formula),

$$v^T(0, x_1, x_2, \ldots) - v^T(0, x_1', x_2, \ldots) = \int_{x_1}^{x_1'} v^T_{x_1}(0, e, x_2, \ldots)\, de$$

and in the limit as $T \to \infty$ we obtain

$$v^\alpha(x_1, x_2, \ldots) - v^\alpha(x_1', x_2, \ldots) = \int_{x_1}^{x_1'} \bar v_1^1(e, x_2, \ldots)\, de,$$

which means exactly that $\bar v_1^1 = v^\alpha_{x_1}$. Similarly for any other variable x_i, and similarly for the second order derivatives: for example,

$$v^T_{x_1}(0, x_1, x_2, \ldots) - v^T_x(0, x_1', x_2, \ldots) = \int_{x_1}^{x_1'} v^T_{x_1 x_1}(0, e, x_2, \ldots)\, de$$

and in the limit as $T \to \infty$ we obtain

$$v^\alpha_{x_1}(x_1, x_2, \ldots) - v^\alpha_{x_1}(x_1', x_2, \ldots) = \int_{x_1}^{x_1'} \bar v_2^{11}(e, x_2, \ldots)\, de,$$

which means that $\bar v_2^{11} = v^\alpha_{x_1 x_1}$, as required. So, indeed, $v^\alpha \in W^2_{p,loc}$.

Now let us show that in the generalised sense the function u from the previous step satisfies the equation (21) and, hence, also to the equivalent one (17). Indeed, for any smooth test function $g(x)$ with a compact support we have,

$$\langle \bar L^\alpha v^\alpha, g\rangle = \langle v^\alpha, \bar L^* g\rangle = \lim_{T\to\infty} \langle v^{(T)}|_{t=0}, (\bar L^\alpha)^* g\rangle = \lim_{T\to\infty} \langle \bar L^\alpha v^{(T)}|_{t=0}, g\rangle$$

$$= \lim_{T\to\infty} \langle \int_0^T \bar L^\alpha E_x \bar f^\alpha(\bar X_t^\alpha) dt, g\rangle$$

$$= \lim_{T\to\infty} \langle \int_0^T \partial_t E_x \bar f^\alpha(\bar X_t^\alpha) dt, g\rangle$$

$$= \lim_{T\to\infty} \langle E_x \bar f^\alpha(\bar X_T^\alpha) - \bar f^\alpha(x), g\rangle$$

$$= -\langle \bar f^\alpha, g\rangle,$$

which means that u is a generalised solution of (21); the last equality in this calculus is because of the ergodic properties of the process $(\bar X_t^\alpha)$ from the Lemma 1 and due to the centering property of $\bar f^\alpha$. Now, since v^α, actually, possesses two Sobolev derivatives in x, this function is not just a generalised, but a true Sobolev solution to the equation (21), or, equivalently, (17).

The bound

$$\sup_\alpha |v^\alpha(x)| \le C(1+|x|^m) \tag{23}$$

follows from the bounds (14) and (11).

The bound

$$\sup_\alpha (|v^\alpha(x)| + |\nabla v^\alpha(x)|) \le C(1+|x|^m) \tag{24}$$

follows from (23) along with the a priori bound

$$\|v\|_{W_p^2()} \le C(\|v\|_{L_p()} + \|\bar f^\alpha\|_{L_p()})$$

[11, Theorem 9.11] and embedding theorems (see, e.g., [19]).

The bound (16)

$$\sup_\alpha (|v^\alpha(x)| + |\nabla v^\alpha(x)|) + |\Delta v^\alpha(x)|) \le C(1+|x|^m)$$

now follows from (24) and from the equation (21).

Now let us show uniqueness of solution of the linear Poisson equation (system)

$$L^\alpha v(x) + f^\alpha(x) - \rho^\alpha = 0$$

in the Sobolev sense up to a constant in the class of functions growing no faster than some polynomial in x. (Clearly, if v is a solution, then $v + C$ is also a solution for any constant C.) Suppose there is a solution v and let us add a constant to it so that to make it centered, i.e.,

$$\langle v, \mu^\alpha \rangle = 0.$$

Let us use Itô – Krylov's formula:

$$\mathbb{E}_x v(X_T^\alpha) - v(x) = \mathbb{E}_x \int_0^T L^\alpha v(X_t^\alpha) dt.$$

Since $\mathbb{E}_x v(X_T^\alpha) \to \langle v, \mu^\alpha \rangle = 0$ as $T \to \infty$, then in the limit we get

$$v(x) = \int_0^\infty \mathbb{E}_x (f^\alpha(X_t^\alpha) - \rho^\alpha)\, dt. \tag{25}$$

This shows uniqueness of solution of the Poisson system (17) in the described class of functions.

Now when it has been established that the solution v of the equation (17) equals v^α, since this right hand side satisfies polynomial growth restrictions, the property (16) follows from a similar lemma for an equation in the 1D case (see [2, Lemma 1]).

We have already seen that the centered version of the solution to (17) is equal to v^α; so, v^α itself is μ^α-centered. Equivalently the equality $\langle v^\alpha, \mu^\alpha \rangle = 0$ follows from integration of the right hand side in the definition of v^α with respect to μ^α, due to the centering property of the function $\bar{f}^\alpha$. The Lemma 2 follows.

□

Lemma 3. *Let the assumptions (A1) – (A2) hold true. Then*

$$\Lambda \ll \bar{\mu}^\alpha$$

(Λ is absolutely continuous with respect to μ^α), where Λ is the Lebesgue's measure in $\mathbb{R}^d$.

Moreover, if all functions $g \ge 0$ from some (abstract) family possess a uniform polynomial bound $g(x) \le C(1 + |x|^\ell)$ and the value $\langle g, \bar{\mu}^\alpha \rangle$ is uniformly small, then for any $R > 0$ the L_d-norm $\|1(|\cdot| \le R) g(\cdot)\|_{L_d}$ is uniformly small. More precisely, for any $\delta > 0$ and any $R > 0$ there exists $\epsilon > 0$ such that if $\langle g, \bar{\mu}^\alpha \rangle < \epsilon$, then

$$\|1(|\cdot| \le R) g(\cdot)\|_{L_d} \le \delta.$$

Proof. We have, by Chapman – Kolmogorov's equation, where $Q^\alpha(x, dx')$ is the transition kernel of the process $\bar X$,

$$\langle g, \bar\mu^\alpha\rangle = \int g(x)\mu^\alpha(dx) = \iint g(x')\bar\mu^\alpha(dx)Q^\alpha(x, dx')$$

$$= \int \mathbb{E}_x g(\bar X_1)\bar\mu^\alpha(dx).$$

Let

$$\rho := \exp(-\int_0^1 b(\alpha(\bar X_s), \bar X_s)dW_s - \frac{1}{2}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds).$$

Since b is bounded, by Girsanov's theorem ρ is a probability density [10]. Consider a new probability measure equivalent to $\mathbb{P}$,

$$\tilde{\mathbb{P}}(A) := \mathbb{E}\rho 1(A).$$

We have,

$$\mathbb{E}_x g(\bar X_1) = \tilde{\mathbb{E}}_x g(\bar X_1)\rho^{-1} = \tilde{\mathbb{E}}_x g(\bar X_1)\exp(\int_0^1 b(\alpha(\bar X_s), \bar X_s)dW_s + \frac{1}{2}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)$$

$$= \tilde{\mathbb{E}}_x g(\bar X_1)\exp(\int_0^1 b(\alpha(\bar X_s), \bar X_s)d\tilde W_s - \frac{1}{2}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds),$$

where

$$\tilde W_t := W_t + \int_0^t b(\alpha(\bar X_s), \bar X_s)ds$$

is a new Wiener process on the $[0, 1]$ time interval ith respect to the measure $\tilde{\mathbb{P}}$.

Further, by virtue of the Cauchy – Bouniakovskii – Schwarz inequality ($\mathbb{E}\xi\eta \geq (\mathbb{E}\sqrt{\xi})^2(\mathbb{E}\eta^{-1/2})^{-1}$) (if only $\xi, \eta \geq 0$ and $\mathbb{E}\eta^{-1/2} \neq 0$), we estimate for any $R > 0$,

$$\mathbb{E}_x g(\bar X_1) = \tilde{\mathbb{E}}_x g(\bar X_1)\exp(\int_0^1 b(\alpha(\bar X_s), \bar X_s)d\tilde W_s - \frac{1}{2}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)$$

$$\geq \left(\tilde{\mathbb{E}}_x g^{1/2}(\bar X_1)\right)^2 \left(\tilde{\mathbb{E}}_x \exp(-\frac{1}{2}\int_0^1 b(\alpha(\bar X_s), \bar X_s)d\tilde W_s + \frac{1}{4}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)\right)^{-1}$$

$$\geq \left(\tilde{\mathbb{E}}_x g^{1/2}(\bar X_1)1(|\bar X_1| \leq R)\right)^2 \left(\tilde{\mathbb{E}}_x \exp(-\frac{1}{2}\int_0^1 b(\alpha(\bar X_s), \bar X_s)d\tilde W_s + \frac{1}{4}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)\right)^{-1}.$$

(Recall that only the functions $g \ge 0$ are considered.) Here again due to the CBS inequality

$$\left(\tilde{\mathbb{E}}_x \exp(-\frac{1}{2}\int_0^1 b(\alpha(\bar X_s), \bar X_s) d\tilde W_s + \frac{1}{4}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)\right)$$

$$\le \left(\tilde{\mathbb{E}}_x \exp(-\int_0^1 b(\alpha(\bar X_s), \bar X_s) d\tilde W_s + \frac{1}{2}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)\right)^{1/2}$$

$$\times \left(\tilde{\mathbb{E}}_x \exp(\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)\right)^{1/2} \le \exp(\|b\|/2).$$

So,

$$\mathbb{E}_x g(\bar X_1) = \tilde{\mathbb{E}}_x g(\bar X_1) \exp(\int_0^1 b(\alpha(\bar X_s), \bar X_s) d\tilde W_s - \frac{1}{2}\int_0^1 |b(\alpha(\bar X_s), \bar X_s)|^2 ds)$$

$$\ge \left(\tilde{\mathbb{E}}_x g^{1/2}(\bar X_1) 1(|\bar X_1| \le R)\right)^2 \exp(-\|b\|/2),$$

or, equivalently,

$$\left(\tilde{\mathbb{E}}_x g^{1/2}(\bar X_1) 1(|\bar X_1| \le R)\right)^2 \le \mathbb{E}_x g(\bar X_1) \exp(\|b\|/2).$$

Therefore,

$$\langle g, \bar\mu^\alpha\rangle = \int \mathbb{E}_x g(\bar X_1)\bar\mu^\alpha(dx) \ge \exp(-\|b\|/2)\int \left(\tilde{\mathbb{E}}_x g^{1/2}(\bar X_1) 1(|\bar X_1| \le R)\right)^2 \mu^\alpha(dx)$$

$$\ge \exp(-\|b\|/2)\int \left(\tilde{\mathbb{E}}_x g^{1/2}(\bar X_1) 1(|\bar X_1| \le R)\right)^2 1(|x|\le R)\mu^\alpha(dx)$$

$$\ge \exp(-\|b\|/2)\int \left(\int_{|x'|\le R} g^{1/2}(x')\frac{1}{(2\pi)^{d/2}}\exp(-(x'+x)^2/2)\right)^2 1(|x|\le R)\mu^\alpha(dx)$$

$$\ge \exp(-\|b\|/2)\int \left(\int_{|x'|\le R} g^{1/2}(x')\frac{1}{(2\pi)^{d/2}}\exp(-(2R)^2/2)dx'\right)^2 1(|x|\le R)\mu^\alpha(dx)$$

$$= \exp(-\|b\|/2)\frac{1}{(2\pi)^d}\exp(-(2R)^2)\left(\int_{|x'|\le R} g^{1/2}(x')dx'\right)^2 \int 1(|x|\le R)\mu^\alpha(dx).$$

In other words,

$$\left(\int_{|x'|\le R} g^{1/2}(x')dx'\right)^2 \le \langle g, \bar\mu^\alpha\rangle \exp(\|b\|/2)\frac{1}{(2\pi)^d}\exp((2R)^2)\mu^\alpha(|x|\le R)^{-1}$$

Choose $R>0$ so large that such that $\mu^\alpha(|x|\le R)\ge 1/2$ for any α. Then

$$\left(\int_{|x'|\le R} g^{1/2}(x')dx'\right)^2 \le 2\langle g, \bar\mu^\alpha\rangle \exp(\|b\|/2)\frac{1}{(2\pi)^d}\exp((2R)^2).$$

In particular, this implies the claim $\Lambda \ll \bar\mu^\alpha$. The Lemma is proved.

Remark 1. *Note that a similar upper bound can be established for the expression $\langle g, \bar\mu^\alpha\rangle$, which implies that, in fact, $\Lambda \sim \bar\mu^\alpha$. However, this will not be used in the sequel.*

Corollary 2. *Under the assumptions of the Lemma 3, let the sequence $(g_n \ge 0, n \ge 1)$ satisfy the bound $\sup_n g_n(x) \le C(1+|x|^\ell)$ and*

$$\langle g_n, \bar\mu^\alpha\rangle \to 0, \quad n\to\infty.$$

Then for any $R > 0$ and for any $m \geq 1$,

$$\int_{|x'| \leq R} g_n^m(x')dx' \to 0, \quad n \to \infty. \tag{26}$$

Proof. We have, uniformly in n,

$$\int_{|x'| \leq R} g_n^m(x')dx' \leq \left(2\langle g_n^{2m}, \bar{\mu}^\alpha\rangle \exp(\|b\|/2) \frac{1}{(2\pi)^{d/2}} \exp((2R)^2)\right)^{1/2}. \tag{27}$$

Note that under the assumptions of the Corollary,

$$\langle g_n^{2m}, \bar{\mu}^\alpha\rangle \to 0, \quad n \to \infty. \tag{28}$$

Indeed, for any $\delta > 0$ the value $R_0 > 0$ can be chosen so large that for any $R \geq R_0$,

$$\int 1(|x|) > R)C^{2m}(1 + |x|)^{2m\ell}\bar{\mu}^\alpha(dx) < \delta, \tag{29}$$

where C is the constant from the assumption $|g_n(x)| \leq C(1+|x|)^\ell$. Hence, it implies

$$\sup_n \int 1(|x|) > R)g_n^{2m}(x)\bar{\mu}^\alpha(dx) < \delta. \tag{30}$$

Then, for the R fixed we have,

$$\int 1(|x| < R)g_n^{2m}(x)\bar{\mu}^\alpha(dx) \leq C(1 + R)^{(2m-1)\ell}\langle g_n, \bar{\mu}^\alpha\rangle \to 0. \tag{31}$$

From (30) and (31) the claim (26) follows. The Corollary 2 is proved.

As a family of such functions (g) the sequence ψ_n (or even ψ_{n_j}) for n large enough (or n_j large enough) will be used in the proof of the Theorem: firstly, we will show that $\langle \psi_n, \mu^\alpha\rangle$ is small enough for n large (see (43) below), and then we deduce from the Corollary 2 that the value $\sup_{n\geq n(\epsilon)} \int_{|x|\leq R} \psi_n^d(x)\, dx$ is also small (see (45) below). The latter will be used for Krylov's estimate.

3 Main results

We accept in this section that solution of the SDE with any Markov strategy exists and is a *weak* solution; we also want it to be unique in distribution, strong Markov and Markov ergodic. All of these follow from [17] and from the assumptions (A1) and (A2) (see [30] about ergodicity).

For any pair (v, ρ) : $v \in \bigcap_{p>1} W^2_{p,loc}$, $\rho \in \mathbb{R}$, define

$$F[v,\rho](x) := \inf_{u\in U} \left[L^u v(x) + f^u(x) - \rho\right], \quad G[v](x) := \inf_{u\in U} \left[L^u v(x) + f^u(x)\right],$$

and

$$F_1[v',\rho](x) := \inf_{u\in U} [\hat{b}^u v' + \hat{f}^u - \hat{\rho}](x),$$

where

$$a^u(x) = \frac{1}{2}(\sigma^u(x))^2, \quad \hat{b}^u(x) = b^u(x)/a^u(x),$$
$$\hat{f}^u(x) = f^u(x)/a^u(x), \quad \hat{\rho}^u(x) = \rho/a^u(x).$$

The functions v and ∇v may be regarded as continuous and absolutely continuous due to the embedding theorems [19]. The function $F[v,\rho](\cdot)$ is defined by the formula above as a function of the class $L_{p,loc}$ for any $p > 1$; in particular, it is Lebesgue measurable and as such it is defined only a.e. in x. We may and will use a (any) Borel measurable version of it, which existence follows, e.g., from Luzin's Theorem [20]). It will be shown in the sequel that the function $F_1[v', \rho](x)$ is continuous in x and locally Lipschitz in the two other variables.

Let us recall what a reward improvement algorithm (RIA) is. We start with some (any) feedback strategy $\alpha_0 \in \mathcal{A}$. Denote the corresponding cost, the invariant measure, and the auxiliary function $\rho_0 = \rho^{\alpha_0} = \langle f^{\alpha_0}, \mu^{\alpha_0}\rangle$, and $v_0 = v^{\alpha_0}$. If for some $n = 0, 1, \ldots$ the triple (α_n, ρ_n, v_n) is determined, then the strategy α_{n+1} is defined as follows: for a.e. x the function α_{n+1} is chosen so that for each x

$$L^{\alpha_{n+1}} v_n(x) + f^{\alpha_{n+1}}(x) = G[v_n](x), \tag{32}$$

or, in other words,

$$\alpha_{n+1}(x) \in \mathrm{Argmin}_{u\in U} \left[L^u v_n(x) + f^u(x)\right].$$

We assume that a Borel measurable version of such strategy may be chosen; see the reference in the Appendix. To this strategy α_{n+1} there correspond the unique invariant measure $\mu^{\alpha_{n+1}}$, the value $\rho_{n+1} := \langle f^{\alpha_{n+1}}, \mu^{\alpha_{n+1}}\rangle$, and the function $v_{n+1} = v^{\alpha_{n+1}}$. Note that the value ρ_{n+1} and the function v^{n+1} do not depend on a particular choice of Borel measurable versions of F and α_{n+1}.

Theorem 1. *Let the assumptions (A1) – (A4) be satisfied. Then:*

1. For any n, $\rho_{n+1} \le \rho_n$, and there is a limit $\rho_n \downarrow \tilde{\rho}$.

2. The sequence (v_n) is compact in $C^1[-N, N]$ for each $N > 0$, and there exists a bounded sequence of constants β_n such that there is a limit $\lim_n (v_n(x) + \beta_n) =: \tilde{v}(x)$.

3. The couple $(\tilde{v}, \tilde{\rho})$ solves the equation (7).

4. This solution $(\tilde{v}, \tilde{\rho})$ is unique – up to an additive constant for $\tilde{v}$ – in the class of functions growing no faster than some (any) polynomial and belonging to the class $W^2_{p,loc}$ for any $p > 0$ for the first component and for $\tilde{\rho} \in \mathbb{R}$.

5. The component $\tilde{\rho}$ in the couple $(\tilde{v}, \tilde{\rho})$ coincides with ρ.

Proof. **1**. Due to (32) and (17), for almost every (a.e.) $x \in \mathbb{R}$,

$$\rho_n = L^{\alpha_n} v_n(x) + f^{\alpha_n}(x) \ge G[v_n](x) = L^{\alpha_{n+1}} v_n(x) + f^{\alpha_{n+1}}(x)$$

and also for a.e. $x \in \mathbb{R}$,

$$\rho_{n+1} = L^{\alpha_{n+1}} v_{n+1}(x) + f^{\alpha_{n+1}}(x)$$

So,

$$\begin{aligned} \rho_n - \rho_{n+1} &\overset{a.e.}{\ge} (L^{\alpha_{n+1}} v_n + f^{\alpha_{n+1}})(x) - (L^{\alpha_{n+1}} v_{n+1} + f^{\alpha_{n+1}})(x) \\ &= (L^{\alpha_{n+1}} v_n - L^{\alpha_{n+1}} v_{n+1})(x). \end{aligned} \tag{33}$$

Let us apply Ito – Krylov's formula (see [16]) with expectations (also known as Dynkin's formula) to $(v_n - v_{n+1})(X^{\alpha_{n+1}}_t)$: we have for any $x \in \mathbb{R}$,

$$\begin{aligned} &\mathbb{E}_x \left(v_n(X^{\alpha_{n+1}}_t) - v_{n+1}(X^{\alpha_{n+1}}_t) \right) - (v_n - v_{n+1})\,(x) \\ &= \mathbb{E}_x \int_0^t (L^{\alpha_{n+1}} v_n - L^{\alpha_{n+1}} v_{n+1})(X^{\alpha_{n+1}}_s)\, ds \le \mathbb{E}_x \int_0^t (\rho_n - \rho_{n+1})\, ds = (\rho_n - \rho_{n+1})\, t. \end{aligned} \tag{34}$$

Why for any x: since the functions $v_n \in C$ due to the embedding theorems [19] as Sobolev solutions of Poisson equations, and because $\mathbb{E}_x v_n(X^{\alpha_{n+1}}_t)$ and $\mathbb{E}_x v_{n+1}(X^{\alpha_{n+1}}_t)$ as functions of x for each $t > 0$ are both Hölder continuous, being solutions of non-degenerate parabolic equations [18]. We also used the fact that the distribution of $X^{\alpha_{n+1}}_s$ for almost all $s > 0$ is absolutely continuous with respect to the Lebesgue measure due to the non-degeneracy and by virtue of Krylov's estimates [16]; due to this reason and because $v_n, v_{n+1} \in C$, the a.e. inequality (33) implies (34) for every x. Further, since the left hand side in (34) is bounded for a fixed x by virtue of the Lemma 2, we divide all terms of the latter inequality by t and let $t \to \infty$ to get,

$$0 \le \rho_n - \rho_{n+1},$$

as required. Thus, $\rho_n \ge \rho_{n+1}$, so that $\rho_n \downarrow \tilde{\rho}$ (since the sequence ρ_n is bounded for $f \in \mathcal{K}$, see (12) in the Lemma 1) with some $\tilde{\rho}$. So, the RIA does converge.

Note that clearly $\tilde{\rho} \ge \rho$, since ρ is the inf over all Markov strategies, while $\tilde{\rho}$ is the inf over some countable subset of them. *Later we shall show that they do coincide.*

Now we want to show that there exists a bounded sequence of real values (non-random!) $\{\beta_n\}$ such that $v_n + \beta_n \to \tilde{v}$, so that the couple $(\tilde{v}, \tilde{\rho})$ satisfies the equation (7), and that $\tilde{\rho}$ here is unique, as well as $\tilde{v}$ in some sense. In the first instance we will do it for some subsequence n_j; eventually the convergence of the whole sequence v_n will follow from the uniqueness of the solution of Bellman's equation, although, it is not important for the proof of the Theorem.

2. Let us show local compactness of the family of functions (v_n) in C^1. (Follows from embedding theorems! Even for a general non-degenerate σ!) Note that the equation (7) is equivalent to the following:

$$\frac{1}{2}\sum_i \frac{d^2}{dx_i^2} v(x) + \inf_{u\in U}\left[\frac{b(u,x)}{a(u,x)}\nabla_x v(x) + \frac{f(u,x)}{a(u,x)} - \frac{\rho}{a(u,x)}\right] = 0, \tag{35}$$

while the equation

$$L^{\alpha_{n+1}} v_{n+1}(x) + f^{\alpha_{n+1}}(x) - \rho_{n+1} \stackrel{a.e.}{=} 0, \tag{36}$$

is equivalent to

$$\frac{1}{2}\sum_i \frac{d^2}{dx_i^2} v_{n+1}(x) + \frac{b(\alpha_{n+1}(x),x)}{a(\alpha_{n+1}(x),x)}\nabla_x v_{n+1}(x) + \frac{f(\alpha_{n+1}(x),(x))}{a(\alpha_{n+1}(x),x)} - \frac{\rho_{n+1}}{a(\alpha_{n+1}(x),x)} = 0.$$

According to the Lemma 2, the functions ∇v_{n+1} are uniformly locally bounded. Since the sequence ρ_{n+1} is bounded and due to the uniform local boundedness of the functions $f(\alpha_{n+1}(x), x)$ and uniform nondegeneracy of a, it follows that (Δv_n) are locally uniformly bounded and satisfy the uniform in n growth bounds similar to (16) for the function itself and for its first derivative due to the equation (e.g., due to (35)). This guarantees pre-compactness of (v_n) in C^1 locally.

3. Due to the (local) compactness property showed in the previous step, by the diagonal procedure from any infinite sub-family of functions v_n it is possible to choose a converging in C^1_{loc} subsequence. We want to show that up to a constant the limit is unique. For this aim, first of all we shall see in a minute that if some $v_{n_j}(x)$ has a limit, say, $\tilde v(x)$ (locally in C) then $v_{n_j+1}(x) + \beta_{n_j}$ has the same limit, where β_n is some bounded sequence of real values. (In fact, what will be established is a little bit more complicated but still enough for our purposes.) We have,

$$L^{\alpha_n} v_n(x) + f^{\alpha_n}(x) - \rho_n \stackrel{a.e.}{=} 0,$$

and

$$L^{\alpha_{n+1}} v_{n+1}(x) + f^{\alpha_{n+1}}(x) - \rho_{n+1} \stackrel{a.e.}{=} 0,$$

and

$$L^{\alpha_{n+1}} v_n(x) + f^{\alpha_{n+1}}(x) - \rho_n =: -\psi_{n+1}(x) \stackrel{a.e.}{\le} 0. \tag{37}$$

Let us rewrite it as follows,

$$L^{\alpha_{n+1}} v_n(x) + f^{\alpha_{n+1}}(x) - \rho_n + \psi_{n+1}(x) \stackrel{a.e.}{=} 0.$$

In other words, the function v_n solves the Poisson equation with the second order operator $L^{\alpha_{n+1}}$ and the "right hand side" $-(f^{\alpha_{n+1}}(x) + \psi_{n+1}(x) - \rho_n)$. This is only possible if the expression $f^{\alpha_{n+1}}(x) + \psi_{n+1}(x) - \rho_n$ is centered with respect to the invariant measure μ^{n+1} because Poisson equations in the whole space have no solutions for non-centered right hand sides (cf., e.g., [25]). This implies that

$$\langle f^{\alpha_{n+1}}(x) + \psi_{n+1} - \rho_n, \mu^{n+1}\rangle = 0$$

So,

$$\langle \psi_{n+1}, \mu^{n+1}\rangle = \rho_n - \rho_{n+1}. \tag{38}$$

By virtue of the Lemma 3 and from the definition (37),

$$\psi_n(x) \le C(1 + |x|^m). \tag{39}$$

2. Denote

$$w_n(x) := v_n(x) - v_{n+1}(x).$$

We have,

$$L^{\alpha_{n+1}} w_n(x) + \psi_{n+1}(x) - (\rho_n - \rho_{n+1}) \stackrel{a.e.}{=} 0.$$

So, there exists a constant $\beta_n = \langle w_n, \mu^{n+1}\rangle$ such that

$$w_n(x) - \beta_n = \int_0^\infty \mathbb{E}_x(\psi_{n+1}(X_t^{n+1}) - (\rho_n - \rho_{n+1}))\,dt. \tag{40}$$

Using the time change we have,

$$\int_0^\infty \mathbb{E}_x(\psi_{n+1}(X_t^{n+1}) - (\rho_n - \rho_{n+1}))\,dt$$

$$= \int_0^\infty \mathbb{E}_x(\psi_{n+1}(\bar X_s^{n+1}) - (\rho_n - \rho_{n+1}))\,\sigma^{-2}(\alpha_{n+1}(\bar X_s^{n+1}), \bar X_s^{n+1})ds$$

$$=: \int_0^\infty \mathbb{E}_x \phi_{n+1}(\bar X_s^{n+1})ds,$$

with

$$\phi_{n+1}(z) := (\psi_{n+1}(z) - (\rho_n - \rho_{n+1}))\,\sigma^{-2}(\alpha_{n+1}(z), z).$$

So, we obtain

$$\langle \psi_{n+1}(\cdot), \bar\mu_{n+1}\rangle - (\rho_n - \rho_{n+1})\langle \sigma^{-2}(\alpha_{n+1}(\cdot), \cdot), \bar\mu_{n+1}\rangle = \langle \phi_{n+1}, \bar\mu_{n+1}\rangle = 0.$$

3. Let us show that for any $N > 0$,

$$\int_{|x|\le N} \psi_n^{2d}(x)\,dx \to 0, \quad n \to \infty. \tag{41}$$

First of all, note that all functions ψ_n and, hence, ψ_n^2 are uniformly locally bounded and may only grow polynomially fast,

$$(0 \le\ )\ \psi_n(x) \le C(1 + |x|^m), \tag{42}$$

with some constants C, m which are the same for all values of n, see the Lemma 3 and Corollary 2. Also, recall that

$$\langle \psi_{n+1}, \mu^{n+1}\rangle = \rho_n - \rho_{n+1} \to 0, \quad n \to \infty. \tag{43}$$

Now let us rewrite the equation (40) via a stationary version of our diffusion, say, $\tilde X_t^{n+1}$:

$$w_n(x) - \beta_n = \int_0^\infty \mathbb{E}_x(\psi_{n+1}(X_t^{n+1}) - E_{\mu^{n+1}}(\psi_{n+1}(\tilde X_t^{n+1}))\,dt.$$

(Note that if we knew that w_n were centered with respect to the invariant measure μ^{n+1} then we would have $\beta_n = 0$; however, the functions v_n and v_{n+1} are both centered with respect to two different measures, and this is the reason why their difference is not just small, but small up to some additive constant; this very constant is denoted by β_n.

4. Using the coupling idea (cf., e.g., [29]), let us consider the independent processes X_t^{n+1} and $\tilde X_t^{n+1}$ on the same (new!) probability space and denote the moment of the first meeting

$$\tau := \inf(t \ge 0 : \, X_t^{n+1} = \tilde X_t^{n+1}).$$

It is known (see [29]) that under our recurrence assumptions for any $k > 0$ there are some constants C_k, m such that uniformly with respect to n,

$$\mathbb{E}_{x,\mu^{n+1}} \tau^k \le C_k(1 + |x|^m).$$

Denote

$$\hat X_t^{n+1} := 1(t < \tau) X_t^{n+1} + 1(t \ge \tau) \tilde X_t^{n+1}.$$

Since τ is a stopping time and because the couple $(X_t^{n+1}, \tilde X_t^{n+1})$ is strong Markov (see [15]), the process $(\hat X_t^{n+1})$ is also strong Markov equivalent to (X_t^{n+1}). Therefore, it is possible to rewrite,

$$w_n(x) - \beta_n = \int_0^\infty \mathbb{E}_{x,\mu} (\psi_{n+1}(\hat X_t^{n+1}) - \psi_{n+1}(\tilde X_t^{n+1}))\, dt.$$

Hence, using the fact that after τ the processes $\hat X_t^{n+1}$ and $\tilde X_t^{n+1}$ coincide, we obtain

$$w_n(x) - \beta_n = \int_0^\infty \mathbb{E}_{x,\mu} 1(t < \tau)(\psi_{n+1}(\hat X_t^{n+1}) - \psi_{n+1}(\tilde X_t^{n+1}))\, dt$$

$$= \int_0^\infty \mathbb{E}_{x,\mu} \sum_{i=0}^\infty 1(i \le \tau < i+1) 1(t < \tau)(\psi_{n+1}(\hat X_t^{n+1}) - \psi_{n+1}(\tilde X_t^{n+1}))\, dt$$

$$= \sum_{i=0}^\infty \mathbb{E}_{x,\mu} \int_0^\infty 1(i \le \tau < i+1) 1(t < \tau)(\psi_{n+1}(\hat X_t^{n+1}) - \psi_{n+1}(\tilde X_t^{n+1}))\, dt.$$

Thus, using Cauchy-Buniakovsky-Schwarz inequality and Fubini Theorem, we have,

$$|w_n(x) - \beta_n| \leq \sum_{i=0}^{\infty} \mathbb{E}_{x,\mu} \int_0^{i+1} 1(\tau > i)|\psi_{n+1}(\hat{X}_t^{n+1}) - \psi_{n+1}(\tilde{X}_t^{n+1})|\, dt$$

$$\leq \sum_{i=0}^{\infty} \int_0^{i+1} \mathbb{E}_{x,\mu} 1(\tau > i)(|\psi_{n+1}(\hat{X}_t^{n+1})| + |\psi_{n+1}(\tilde{X}_t^{n+1})|)\, dt$$

$$\leq \sum_{i=0}^{\infty} \int_0^{i+1} (\mathbb{E}_{x,\mu} 1(\tau > i))^{1/2} (\mathbb{E}_{x,\mu}(|\psi_{n+1}(\hat{X}_t^{n+1})| + |\psi_{n+1}(\tilde{X}_t^{n+1})|)^2)^{1/2}\, dt$$

$$\leq 2 \sum_{i=0}^{\infty} (\mathbb{E}_{x,\mu} 1(\tau > i))^{1/2} \int_0^{i+1} (\mathbb{E}_{x,\mu}|\psi_{n+1}(\hat{X}_t^{n+1})|^2 + \mathbb{E}_{x,\mu}|\psi_{n+1}(\tilde{X}_t^{n+1})|^2)^{1/2}\, dt$$

$$\leq 2 \sum_{i=0}^{\infty} (\mathbb{E}_{x,\mu} 1(\tau > i))^{1/2} \int_0^{i+1} [(\mathbb{E}_{x,\mu}(\psi_{n+1}(\hat{X}_t^{n+1}))^2)^{1/2} + (\mathbb{E}_{x,\mu}(\psi_{n+1}(\tilde{X}_t^{n+1}))^2)^{1/2}]\, dt.$$

Now, let us take any $\epsilon > 0$ and use the inequality $\sqrt{a} \leq \frac{\epsilon}{2} + \frac{a}{2\epsilon}$. We estimate,

$$\int_0^{i+1} [(\mathbb{E}_{x,\mu}(\psi_{n+1}(\hat{X}_t^{n+1}))^2)^{1/2} + (\mathbb{E}_{x,\mu}(\psi_{n+1}(\tilde{X}_t^{n+1}))^2)^{1/2}]\, dt$$

$$\leq \epsilon(i+1) + \frac{1}{2\epsilon} \int_0^{i+1} [\mathbb{E}_{x,\mu}\psi_{n+1}^2(\hat{X}_t^{n+1}) + \mathbb{E}_{x,\mu}\psi_{n+1}^2(\tilde{X}_t^{n+1})]\, dt.$$

Let us first consider the stationary term. We have,

$$\frac{1}{2\epsilon} \int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2(\tilde{X}_t^{n+1})\, dt$$

$$= \frac{1}{2\epsilon} \int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{|x| \leq N}(\tilde{X}_t^{n+1})\, dt + \frac{1}{2\epsilon} \int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{\mathbb{R}^d \setminus B_N}(\tilde{X}_t^{n+1})\, dt$$

$$+ \frac{1}{2\epsilon} \int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{B_N}(\hat{X}_t^{n+1})\, dt + \frac{1}{2\epsilon} \int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{\mathbb{R}^d \setminus B_N}(\hat{X}_t^{n+1})\, dt.$$

Given (42) and because any stationary measure integrates uniformly any power function, let us find such N that uniformly with respect to n,

$$\langle C(1+|x|^{2m})1_{\mathbb{R}^d\setminus B_N}, \mu^{n+1}\rangle < \epsilon^2/2, \tag{44}$$

which is possible due to the Lemmata 1 and 3, and also such that $N > \epsilon^{-2}$. Then choose $n(\epsilon)$ such that

$$\sup_{n\ge n(\epsilon)} \int_{|x|\le N} \psi_n^{2d}(x)\, dx < \epsilon^2/2. \tag{45}$$

Now we evaluate with $n \ge n(\epsilon)$ due to Krylov's estimate [16, 17],

$$\frac{1}{2\epsilon}\int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{B_N}(\tilde X_t^{n+1})\, dt$$

$$= \frac{1}{2\epsilon}\sum_{k=0}^{i}\mathbb{E}_x\int_k^{k+1} \psi_{n+1}^2 1_{B_N}(\tilde X_t^{n+1})\, dt \le \frac{i+1}{2\epsilon}K\|\psi_{n+1}^2 1_{|x|\le N}\|_{L^d} \le \frac{(i+1)K\epsilon}{2}.$$

This argument works for the non-stationary process as well: due to Krylov's estimate,

$$\frac{1}{2\epsilon}\int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{B_N}(\hat X_t^{n+1})\, dt$$

$$= \frac{1}{2\epsilon}\sum_{k=0}^{i}\mathbb{E}\int_k^{k+1} \psi_{n+1}^2 1_{B_N}(\hat X_t^{n+1})\, dt \le \frac{i+1}{2\epsilon}K\|\psi_{n+1}^2 1_{B_N}\|)_{L^d} \le \frac{(i+1)K\epsilon}{2}.$$

Further,

$$\frac{1}{2\epsilon}\int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{\mathbb{R}^d\setminus B_N}(\tilde X_t^{n+1})\, dt \le \frac{i+1}{2\epsilon}\times\frac{\epsilon^2}{2} = \frac{(i+1)\epsilon}{4}.$$

Finally, using (15), we obtain with some m,

$$\frac{1}{2\epsilon}\int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{\mathbb{R}^d\setminus B_N}(\hat X_t^{n+1})\, dt = \frac{1}{2\epsilon}\int_0^{i+1} \mathbb{E}_{x,\mu}\psi_{n+1}^2 1_{\mathbb{R}^d\setminus B_N}(X_t^{n+1})\, dt$$

$$\le C\frac{i+1}{2\epsilon}\frac{(1+|x|^m)}{N} \le C(i+1)(1+|x|^m)\epsilon.$$

Overall, this shows that with the appropriately chosen (uniformly bounded) β_n,

$$|w_n(x) - \beta_n| \le C(1+|x|^{2m})\epsilon \sum_{i=0}^{\infty}(i+1)(\mathbb{E}_{x,\mu}1(\tau > i))^{1/2}, \quad n \ge n(\epsilon). \tag{46}$$

By virtue of the results in [29], for any $k > 0$ there are $C, m > 0$ such that

$$\mathbb{P}_{x,\mu}1(\tau > i) \le C\frac{1+|x|^m}{1+i^k}.$$

Therefore, taking any $k > 1$, we have that the series in (46) converges providing us an estimate

$$|w_n(x) - \beta_n| \le C(1+|x|^{3m})\epsilon, \quad n \ge n(\epsilon). \tag{47}$$

In other words, the difference $w_n(x) - \beta_n = v_n - v_{n+1} - \beta_n$ is locally uniformly converging to zero as $n \to \infty$. Naturally, it also implies that for any subsequence n_j such that v_{n_j} converges locally uniformly in C^1 we have that for any $1 \le i \le d$, $\partial_{x_i} v_{n_j}$ and $\partial_{x_i} v_{n_j+1}$ may only converge to the same limit, i.e., derivatives $\partial_{x_i} v_{n_j} - \partial_{x_i} v_{n_j+1} \to 0$ (locally uniformly) as $j \to \infty$. Indeed, otherwise we just integrate to show that the limits of v_{n_j} and $v_{n_j+1} + \beta_{n_j}$ are different, which contradicts to what was established earlier. This contradiction shows that $\nabla v_{n_j} - \nabla v_{n_j+1} \to 0$ locally uniformly as $j \to \infty$.

5. What we want to do now is to pass to the limit as $j \to \infty$ in the equations

$$L^{\alpha_{n_j+1}} v_{n_j+1}(x) + f^{\alpha_{n_j+1}}(x) - \rho_{n_j+1} \stackrel{a.e.}{=} 0, \quad \& \quad G[v_{n_j}](x) - \rho_{n_j} \le 0,$$

where $(n_j, j \to \infty)$ is any sequence such that v_{n_j} converges (locally uniformly) in C^1. From

$$G[v_{n_j}](x) - \rho_{n_j} = L^{\alpha_{n_j+1}} v_{n_j}(x) + f^{\alpha_{n_j+1}}(x) - \rho_{n_j}$$

$$(= \inf_{u \in U}[L^u v_{n_j}(x) + f^u(x) - \rho_{n_j}] \stackrel{a.e.}{\le} 0),$$

by subtracting zero a.e. (36), we obtain a.e.,

$$G[v_{n_j}](x) - \rho_{n_j} = L^{\alpha_{n_j+1}}(v_{n_j}(x) - v_{n_j+1}(x)) - (\rho_{n_j} - \rho_{n_j+1}). \tag{48}$$

Now we want to show that (48) implies

$$\frac{1}{2}\Delta\tilde{v}(x) + F_1[x, \nabla\tilde{v}(x), \tilde{\rho}] = 0, \tag{49}$$

where

$$F_1[x, v_1, r] := \inf_{u \in U} \left[\frac{b(u,x)}{a(u,x)} v_1 + \frac{f(u,x)}{a(u,x)} - \frac{r}{a(u,x)} \right]$$

for any x, where $v_1 \in \mathbb{R}^d$ and bv_1 is a scalar product.

Let us show that (48) indeed, implies (49). Note that $G[v_{n_j}](x) - \rho_{n_j} \le 0$ (a.e.). Let us divide (48) by $a_{n_j+1} = a^{\alpha_{n_j+1}}$ and use $\delta := \inf_{u,x} a^u(x) > 0$: we get a.e. with some $K > 0$,

$$0 \ge \frac{(G[v_{n_j}](x) - \rho_n)}{a_{n_j+1}}$$

$$= (\frac{1}{2}\Delta v_{n_j}(x) - \frac{1}{2}\Delta v_{n_j+1}(x)) + (\hat{b}^{\alpha_{n_j+1}}(\nabla v_{n_j} - \nabla v_{n_j+1})) - \frac{(\rho_{n_j} - \rho_{n_j+1})}{a_{n_j+1}}$$

$$\ge (\frac{1}{2}\sum_j (D^2_{x_i} v_{n_j}(x) - D^2_{x_i} v_{n_j+1})(x)) - \frac{K}{\delta}|\nabla v_{n_j}(x) - \nabla v_{n_j+1}(x)| - \frac{1}{\delta}(\rho_{n_j} - \rho_{n_j+1}). \tag{50}$$

So, we have just shown that a.e.,

$$0 \ge (\frac{1}{2}\Delta v_{n_j}(x) - \frac{1}{2}\Delta v_{n_j+1}(x)) - \frac{K}{\delta}|\nabla v_{n_j}(x) - \nabla v_{n_j+1}(x)| - \frac{\rho_{n_j} - \rho_{n_j+1}}{\delta}. \tag{51}$$

The next trick is to note that again due to (50) and $\rho_{n_j} \ge \rho_{n_j+1}$, and since $\delta \le a \le C$,

$$0 \overset{a.e.}{\ge} G[v_{n_j}](x) - \rho_{n_j} \ge a_{n_j+1}(\frac{1}{2}\Delta v_{n_j} - \frac{1}{2}\Delta v_{n_j+1})(x) - C'|v'_{n_j} - v'_{n_j+1}|(x) - (\rho_{n_j} - \rho_{n_j+1}),$$

which implies that with some $C, c > 0$,

$$0 \overset{a.e.}{\ge} \frac{1}{2}\Delta v_{n_j} + F_1[v'_{n_j}, \rho_{n_j}] \ge ((\frac{1}{2}\Delta v_{n_j} - \frac{1}{2}\Delta v_{n_j+1}) - C|v'_{n_j} - v'_{n_j+1}|) - c(\rho_{n_j} - \rho_{n_j+1}). \tag{52}$$

That is,

$$\begin{aligned} 0 &\overset{a.e.}{\ge} \frac{1}{2}\Delta v_{n_j} + \inf_{u \in U} \left[\frac{b(u,x)}{a(u,x)} \nabla_x v_{n_j}(x) + \frac{f(u,x)}{a(u,x)} - \frac{\rho_{n_j}}{a(u,x)} \right] \\ &\ge (\frac{1}{2}\Delta v_{n_j} - \frac{1}{2}\Delta v_{n_j+1})(x) - C|\nabla v_{n_j} - \nabla v_{n_j+1}|(x) - c(\rho_{n_j} - \rho_{n_j+1}). \end{aligned} \tag{53}$$

Since $C|\nabla v_{n_j} - \nabla v_{n_j+1}|) - c(\rho_{n_j} - \rho_{n_j+1}) \to 0$ locally uniformly, and because $\nabla v_{n_j} \to \nabla \tilde{v}$ and $\nabla v_{n_j+1} \to \nabla \tilde{v}$ also locally uniformly, then (53) prompts that there exists a (locally uniform) limit

$$\lim_{n_j\to\infty} \frac{1}{2}\Delta v_{n_j+1}(x) = -F_1[x, \nabla\tilde{v}(x), \tilde{\rho}(x)],$$

or, for short,

$$\lim_{n_j\to\infty} \frac{1}{2}\Delta v_{n_j+1} = -F_1[\nabla\tilde{v}, \tilde{\rho}].$$

It is plausible that in the limit we obtain the equation for the function $\tilde{v}$:

$$\frac{1}{2}\Delta\tilde{v} = -F_1[\nabla\tilde{v}, \tilde{\rho}]. \tag{54}$$

6. The equation (54) requires some justification because a priori it is not known whether or not the function $\tilde{v}$ possesses the second order derivatives. For this justification let us rewrite the double inequality (53) as a double *equality* ((55) & (56)),

$$\frac{1}{2}\Delta v_{n_j+1}(x) + \inf_{u\in U}\left[\frac{b(u,x)}{a(u,x)}\nabla_x v_{n_j}(x) + \frac{f(u,x)}{a(u,x)} - \frac{\rho_{n_j}}{a(u,x)}\right] = \zeta_{n_j}(x), \tag{55}$$

and

$$\frac{1}{2}\Delta v_{n_j}(x) + \inf_{u\in U}\left[\frac{b(u,x)}{a(u,x)}\nabla_x v_{n_j}(x) + \frac{f(u,x)}{a(u,x)} - \frac{\rho_{n_j}}{a(u,x)}\right] = \eta_{n_j}(x), \tag{56}$$

where

$$\zeta_{n_j}(x) \ge -C|\nabla v_{n_j} - \nabla v_{n_j+1}|(x) - c(\rho_{n_j} - \rho_{n_j+1})$$

and

$$\eta_{n_j}(x) \le 0.$$

In other words,

$$\frac{1}{2}\Delta v_{n_j+1}(x) + F_1\left[\nabla_x v_{n_j}, \rho_{n_j}\right] = \zeta_{n_j}(x), \tag{57}$$

and

$$\frac{1}{2}\Delta v_{n_j}(x) + F_1\left[\nabla_x v_{n_j}, \rho_{n_j}\right] = \eta_{n_j}(x). \tag{58}$$

Let B_r be any open ball of a radius r, and

$$\tau_r := \inf(t \geq 0 :\ x + W_t \notin B_r).$$

Note that the stopping time τ_r has some finite exponential moment along with all polynomial moments. From (57) and (58) it follows that for $x \in B_r$ we have,

$$\begin{aligned} v_{n_j}(x) &= \mathbb{E}\left(\int_0^{\tau_r} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s) - \eta_{n_j}(x+W_s)\right) ds + v_{n_j}(x+W_{\tau_r}) \right) \\ &\geq \mathbb{E}\left(\int_0^{\tau_r} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s)\right) ds + v_{n_j}(x+W_{\tau_r}) \right), \end{aligned} \tag{59}$$

and

$$\begin{aligned} v_{n_j+1}(x) &= \mathbb{E}\left(\int_0^{\tau_r} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s) - \zeta_{n_j}(x+W_s)\right) ds + v_{n_j}(x+W_{\tau_r}) \right) \\ &\leq \mathbb{E}\left(\int_0^{\tau_r} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s) + (C|\nabla v_{n_j} - \nabla v_{n_j+1}|(x+W_s) \right.\right. \\ &\qquad \left.\left. + c(\rho_{n_j} - \rho_{n_j+1}))\right) ds + v_{n_j}(x+W_{\tau_r}) \right). \end{aligned} \tag{60}$$

Indeed, let $T_N := \inf(t \geq 0 :\ x + W_t \notin B_{r-1/N})$ for $N > 1/r$. Then we have,

$$T_N \uparrow \tau_r, \quad N \to \infty,$$

and

$$\begin{aligned} v_{n_j}(x) &= \mathbb{E}\left(\int_0^{T_N} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s) - \eta_{n_j}(x+W_s)\right) ds + v_{n_j}(x+W_{T_N}) \right) \\ &\geq \mathbb{E}\left(\int_0^{T_N} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s)\right) ds + v_{n_j}(x+W_{T_N}) \right), \end{aligned} \tag{61}$$

and

$$v_{n_j+1}(x) = \mathbb{E}\left(\int_0^{T_N} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s) - \zeta_{n_j}(x+W_s)\right) ds + v_{n_j}(x+W_{T_N})\right)$$

$$\leq \mathbb{E}\left(\int_0^{T_N} \left(F_1\left[\nabla_x v_{n_j}\right](x+W_s) + (C|\nabla v_{n_j} - \nabla v_{n_j+1}|(x+W_s) \right.\right.$$

$$\left.\left. + c(\rho_{n_j} - \rho_{n_j+1}))\right) ds + v_{n_j}(x+W_{T_N})\right). \tag{62}$$

As $N \to \infty$, we get the desired inequalities (59) and (60).

Further, as $n_j \to \infty$, we obtain from (59) and (60) in the limit,

$$\tilde{v}(x) \geq \mathbb{E}\left(\int_0^{\tau_r} \left(F_1\left[\nabla_x \tilde{v}\right](x+W_s)\right) ds + \tilde{v}(x+W_{\tau_r})\right),$$

and

$$\tilde{v}(x) \leq \mathbb{E}\left(\int_0^{\tau_r} \left(F_1\left[\nabla_x \tilde{v}\right](x+W_s)\right) ds + \tilde{v}(x+W_{\tau_r})\right).$$

This means that actually for each $x \in B_r$

$$\tilde{v}(x) = \mathbb{E}\left(\int_0^{\tau_r} \left(F_1\left[\nabla_x \tilde{v}\right](x+W_s)\right) ds + \tilde{v}(x+W_{\tau_r})\right). \tag{63}$$

However, if we denote by $\bar{v}(x)$ the right hand side of (63), then due to the standard probabilistic arguments the function $\bar{v}$ is a unique solution of the elliptic PDE (see [11, Corollary 9.18])

$$\frac{1}{2}\Delta\bar{v}(x) + F_1\left[\nabla_x \tilde{v}\right](x) = 0, \quad \bar{v}|_{\partial B_r} = \tilde{v}|_{\partial B_r}, \tag{64}$$

in the Sobolev classes $W^2_{p,loc}(B_r) \cap C(\bar{B}_r)$ with any $p > 1$.

Indeed, let us take the (unique) solution $\bar{v}$ of the equation (64) from this class and apply Itô-Krylov's formula to $\bar{v}(x+W_t)$ for $t < \tau_r$:

$$d\bar{v}(x+W_t) = \frac{1}{2}\Delta\bar{v}(x+W_t)dt + \nabla\bar{v}(x+W_t)dW_t.$$

Using the same stopping time sequence $(T_N, N \geq 1)$ as above, we have in the integral form,

$$\bar{v}(x + W_{T_N}) - \bar{v}(x) = \int_0^{T_N} \frac{1}{2}\Delta\bar{v}(x + W_t)dt + \int_0^{T_N} \nabla\bar{v}(x + W_t)dW_t.$$

Taking expectations of both sides, we obtain

$$\mathbb{E}\bar{v}(x + W_{T_N}) - \bar{v}(x) = \mathbb{E}\int_0^{T_N} \frac{1}{2}\Delta\bar{v}(x + W_t)dt = \mathbb{E}\int_0^{T_N} F_1\left[\nabla_x\tilde{v}\right](x + W_t)dt.$$

Here due to the properties of $\bar{v}, F_1 \in C$, we get in the limit as $N \to \infty$,

$$\mathbb{E}\bar{v}(x + W_{\tau_r}) - \bar{v}(x) = \mathbb{E}\int_0^{\tau_r} F_1\left[\nabla_x\tilde{v}\right](x + W_t)dt. \tag{65}$$

Comparing it to (63), we conclude that the function $\tilde{v}$ itself in B_r satisfies the equation (64) and belongs to the same class, $\tilde{v} \in W^2_{p,loc}(B_r) \cap C(\bar{B}_r)$ with any $p > 1$ (and for any ball B_r), and the equation (54) holds true, as it was promised. It is equivalent to (49).

In the sequel it will follow from the uniqueness of solution of Bellman's equation that actually the whole sequence v_n converges up to an additive constant sequence locally uniformly in C^1 to a single limit. However, it is not needed for our proof.

7. *Uniqueness for ρ in (7).* Assume that there are two solutions of the (HJB) equation, (v^1, ρ^1) and (v^2, ρ^2) with $v^i \in \mathcal{K}$, $i = 1, 2$:

$$\inf_{u\in U}(L^u v^1(x) + f^u(x) - \rho^1) = \inf_{u\in U}(L^u v^2(x) + f^u(x) - \rho^2) = 0.$$

Earlier it was shown that both v^1 and v^2 are classical solutions with locally Lipschitz second derivatives. Let $w(x) := v^1(x) - v^2(x)$ and consider two strategies $\alpha_1, \alpha_2 \in \mathcal{A}$ such that $\alpha_1(x) \in \text{Argmax}_{u\in U}(L^u w(x))$ and $\alpha_2(x) \in \text{Argmin}_{u\in U}(L^u w(x))$, and let X^1_t, X^2_t be solutions of the SDEs corresponding to each strategy α_i, $i = 1, 2$. Note that due to the measurable choice arguments – see the Appendix – such Borel strategies exist; corresponding weak solutions also exist. Let us denote

$$h_1(x) := \sup_{u\in U}(L^u w(x) - \rho^1 + \rho^2), \quad h_2(x) := \inf_{u\in U}(L^u w(x) - \rho^1 + \rho^2).$$

Then,

$$h_2(x) = \inf_{u\in U}(L^u v^1(x) + f^u(x) - \rho^1 - (L^u v^2(x) + f^u(x) - \rho^2))$$

$$\leq \inf_{u\in U}(L^u v^1(x) + f^u(x) - \rho^1) - \inf_{u\in U}(L^u v^2(x) + f^u(x) - \rho^2) = 0.$$

Similarly,

$$h_1(x) = -\inf_u(L^u(-v^2)(x) - \rho^2 + \rho^1)$$

$$= -\inf_u(L^u v^2(x) + f^u(x) + \rho^2 - (L^u v^1(x) + f^u(x) + \rho^1))$$

$$\geq -[\inf_u(L^u v^2(x) + f^u(x) - \rho^2) - \inf_u(L^u v^1(x) + f^u(x) - \rho^1)] = 0.$$

We have,

$$L^{\alpha_2} w(x) = h_2(x) - \rho^2 + \rho^1,$$

and

$$L^{\alpha_1} w(x) = h_1(x) - \rho^2 + \rho^1.$$

Due to Dynkin's formula we have,

$$\mathbb{E}_x w(X^1_t) - w(x) = \mathbb{E}_x \int_0^t L^{\alpha_1} w(X^1_s)\, ds$$

$$= \mathbb{E}_x \int_0^t h_1(X^1_s)\, ds + (\rho^1 - \rho^2)\, t \overset{(h_1\geq 0)}{\geq} (\rho^1 - \rho^2)\, t.$$

Since the left hand side here is bounded for a fixed x, due to the Lemma 1 we get,

$$\rho^1 - \rho^2 \leq 0.$$

Similarly, considering α_2 we conclude that

$$\mathbb{E}_x w(X^2_t) - w(x) = \mathbb{E}_x \int_0^t L^{\alpha_2} w(X^2_s)\, ds$$

$$= \mathbb{E}_x \int_0^t h_2(X^2_s)\, ds + (\rho^1 - \rho^2)\, t.$$

From here, due to the boundedness of the left hand side (Lemma 1) we get,

$$\rho^2 - \rho^1 = \liminf_{t\to 0}(t^{-1}\mathbb{E}_x \int_0^t h_2(X^2_s)\, ds) \overset{(h_2\leq 0)}{\leq} 0.$$

Thus, $\rho^1 - \rho^2 \geq 0$ and, hence,
$$\rho^1 = \rho^2.$$

8. *Proof of $\rho = \tilde\rho$.* Recall that for any initial $\alpha_0 \in \mathcal{A}$, the sequence ρ_n converges to the same value $\tilde\rho$, which is a unique component of solution of the equation (7). Let us take any $\epsilon > 0$ and consider a strategy α_0 such that

$$\rho_0 = \rho^{\alpha_0} < \rho + \epsilon.$$

Since the sequence (ρ_n) decreases, the limit $\tilde\rho$ must satisfy the same inequality,

$$\tilde\rho = \lim_{n\to\infty} \rho_n < \rho + \epsilon.$$

Due to uniqueness of $\tilde\rho$ as a component of solution of the equation (7) and since $\epsilon > 0$ is arbitrary, we find that
$$\tilde\rho \leq \rho.$$
But also $\tilde\rho \geq \rho$ since $\tilde\rho$ is the infimum of the cost function values over a smaller – just countable – family of strategies. So, in fact,

$$\tilde\rho = \rho.$$

9. *Uniqueness for V.* Let us have another look at the earlier equations in the step 6, replacing $\rho^2 - \rho^1$ by zero as we already know that the second component in the solution is unique:

$$\mathbb{E}_x w(X^1_t) - w(x) = \mathbb{E}_x \int_0^t h_1(X^1_s)\, ds \leq \mathbb{E}_x \int_0^{Ct} h_1(\tilde X^1_s)\, ds$$

Clearly, $h_1 \geq 0$ with $h_1 \not\equiv 0$ – i.e., with $\tilde\mu_1(x : \, h_1(x) > 0) > 0$ – would imply that $\langle h_1, \tilde\mu_1\rangle > 0$, which contradicts to the zero left hand side (after division by t with $t \to \infty$). So, we conclude that

$$h_1 = 0, \quad \tilde\mu_1 - \text{a.s.}$$

Since $\Lambda \ll \tilde\mu_1$ (see the Lemma 3), we have by virtue of Krylov's estimate that $0 \leq \mathbb{E}_x \int_0^{Ct} h_1(\tilde X^1_s)\, ds \leq N\|h_1\|_{L_d} = 0$. So, in fact,

$$\mathbb{E}_x w(X^1_t) - w(x) = 0. \tag{66}$$

Further, from (66) and due to the last statement of the Lemma 1 it follows that

$$w(x) = \lim_{t\to\infty} \mathbb{E}_x w(X_t^1) = \langle w, \mu_1 \rangle.$$

Hence, $w(x)$ is a constant. Recall that uniqueness of the first component V is stated up to a constant, and it was just established that

$$v^1(x) - v^2(x) = \text{const}.$$

10. Returning to the second statement of the Theorem, note that due to uniqueness of solution of the HJB equation, convergence of the whole sequence (v_n) up to additive constants depending only on n is to the unique limit v.

$\square$

References

[1] S.V. Anulova, H. Mai, A.Yu. Veretennikov, On averaged expected cost control as reliability for 1D ergodic diffusions, *Reliability: Theory & Applications (RT&A)*, 12, 4(47), 31-38, 2017.

[2] S.V. Anulova, H. Mai, A.Yu. Veretennikov, On averaged expected cost control for 1D ergodic diffusions, https://arxiv.org/abs/1812.10665.

[3] A. Arapostatis, V.S. Borkar, E. Fernãndes-Gaucherand, M.K. Ghosh, and S.I. Markus, Discrete-time controlled Markov processes with average cost criterion: a survey, *SIAM J. Control and Optimization*, 31(2), 282-344, 1993.

[4] 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., 1–12. Birkhäuser/Springer, New York, 2012.

[5] A. Arapostathis, V.S. Borkar, A relative value iteration algorithm for nondegenerate controlled diffusions, *SIAM Journal on Control and Optimization*, 50(4), 1886-1902, 2012.

[6] A. Arapostathis, V.S. Borkar, M.K. Ghosh. *Ergodic control of diffusion processes.* Encyclopedia of Mathem. and its Appl. 143. Cambridge: CUP, 2012.

[7] V.S. Borkar. *Optimal control of diffusion processes.* Harlow: Longman Scientific & Technical; New York: John Wiley & Sons, 1989.

[8] E.B. Dynkin and A.A. Yushkevich. *Upravlyaemye markovskie protsessy i ikh prilozheniya*, Moskva: "Nauka", 1975 (in Russian).

[9] I.I. Gikhman, A.V. Skorokhod, *Stochastic Differential Equations.* Springer, 1972.

[10] I.V. Girsanov, On transforming a class of stochastic processes by absolutely continuous substitution of measures, *Theory Probab. Appl.* 1960, 3(5), 285-301.

[11] D. Gilbarg, N.S. Trudinger, *Elliptic partial differential equations of second order, 2nd edition,* Springer: Berlin et al., 1983.

[12] R.A. Howard. *Dynamic programming and Markov processes.* New York – London: John Wiley & Sons, Inc. and the Technology Press of the MIT, 1960.

[13] R.A. Howard. *Dynamic probabilistic systems. Vol. II: Semi-Markov and decision processes. Reprint of the 1971 original ed.* Mineola, NY: Dover Publ., 577-1108, 2007.

[14] R. Khasminskii. *Stochastic stability of differential equations. With contributions by G.N. Milstein and M.B. Nevelson. 2nd completely revised and enlarged ed.* Berlin: Springer, 2012.

[15] N.V. Krylov. On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes. *Math. USSR Izv.* 7: 691–709, 1973.

[16] N.V. Krylov. *Controlled diffusion processes, 2nd ed.* Berlin, Springer, 2009.

[17] N.V. Krylov. On Ito's stochastic integral equations. *Theory Probab. Appl.*, 14: 330–336, 1969.

[18] N.V. Krylov, M.V. Safonov. A certain property of solutions of parabolic equations with measurable coefficients. *Math. USSR Izv.,* 16:1, 151–164, 1981.

[19] O.A. Ladyzhenskaja, V.A. Solonnikov, N.N. Ural'ceva, *Linear and Quasi-linear Equations of Parabolic Type*, AMS, 1968.

[20] N. Lusin. Sur les propriétés des fonctions mesurables, *Comptes rendus de l'Académie des Sciences de Paris*, 154, 1688-1690, 1912.

[21] P. Mandl. *Analytical treatment of one-dimensional Markov processes.* Berlin-Heidelberg-New York: Springer, 1968.

[22] S. Meyn, The Policy Iteration Algorithm for Average Reward Markov Decision Processes with General State Space, *IEEE Transactions on Automatic Control,* 42(12), 1663-1680, 1997.

[23] H. Mine and S. Osaki. *Markovian decision processes.* New York: American Elsevier Publishing Company, Inc., 1970.

[24] R. Morton. On the optimal control of stationary diffusion processes with inaccessible boundaries and no discounting. *J. Appl. Probab.*, 8: 551–560, 1971.

[25] É. Pardoux and A.Yu. Veretennikov. On Poisson equation and diffusion approximation. I. *Ann. Probab.*, 29(3): 1061–1085, 2001.

[26] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, 2005.

[27] V.V. Rykov, Controllable Queueing Systems: From the Very Beginning up to Nowadays, *Reliability: Theory & Applications (RT&A)*, vol. 12, 2(45), 39-61, 2017.

[28] V. A. Solonnikov, On boundary value problems for linear parabolic systems of differential equations of general form, *Proc. Steklov Inst. Math.,* 83 (1965), 1-184.

[29] A.Yu. Veretennikov. On Polynomial Mixing and Convergence Rate for Stochastic Difference and Differential Equations, *Theory of Probability & Its Applications,* 44(2): 361-374, 2000.

[30] A.Yu. Veretennikov. On polynomial mixing for SDEs with a gradient-type drift. *Theory Probab. Appl.*, 45(1): 160–164, 2001.

[31] A.Yu. Veretennikov. On Sobolev Solutions of Poisson Equations in $\mathbb{R}^d$ with a Parameter, *Journal of Mathematical Sciences*, 179(1), 48-79, 2011.