A Selection Procedure for Extracting the Unique Feller Weak Solution of Degenerate Diffusions

Sumith Reddy Anugu · Vivek S. Borkar

Accepted: 23 December 2022 / Published online: 13 March 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract
In this work, we show that for the martingale problem for a class of degenerate diffusions with bounded continuous drift and diffusion coefficients, the small noise limit of non-degenerate approximations leads to a unique Feller limit. The proof uses the theory of viscosity solutions applied to the associated backward Kolmogorov equations. Under appropriate conditions on drift and diffusion coefficients, we will establish a comparison principle and a one-one correspondence between Feller solutions to the martingale problem and continuous viscosity solutions of the associated Kolmogorov equation. This work can be considered as an extension to the work in Borkar and Kumar in (J Theor Probab 23(3): 729–747, 2010).

Keywords Degenerate diffusions · Markov selection · Feller solution · Small noise limit · Non-uniqueness of weak solution · Backward Kolmogorov equation · Uniqueness of continuous viscosity solutions

Mathematics Subject Classification 60H10 · 60J25 · 34F05 · 35K65 · 35D40 · 49L25

The work of ASR was funded by Institute Postdoctoral Fellowship, Indian Institute of Technology Bombay. The work of VSB was supported in part by an S. S. Bhatnagar Fellowship from the Council for Scientific and Industrial Research, Government of India.

Sumith Reddy Anugu
anugu.reddy@rice.edu

Vivek S. Borkar
borkar.vs@gmail.com

1 Computational Applied Mathematics and Operation Research Department, Rice University, Houston, TX 77005-1827, USA

2 Department of Electrical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India
1 Introduction

Let $C \equiv C([0, T], \mathbb{R}^P)$ denote the space of continuous $\mathbb{R}^P$-valued functions endowed with supremum norm and let $B \equiv B(C)$ denote the Borel $\sigma$-algebra of $C$. Consider a stochastic differential equation on $\mathbb{R}^n$ given by

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 \sim \mu,$$

where, $X_t \in \mathbb{R}^n$, $b(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma(\cdot) : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ with $W(\cdot)$ being a Brownian motion on $\mathbb{R}^m$. We define $\mathbb{P}_X$ to be the law of process $X$ and $\mathbb{E}_X$ to be the corresponding expectation. We just write $\mathbb{E}$ when it is clear from the context.

If we assume that the coefficients of (1.1) are bounded continuous, then it is well known that the notion of weak solution and the notion of solution to the corresponding martingale problem (due to Stroock and Varadhan) of (1.1) are equivalent [17, Proposition 4.11] and that (1.1) has a solution to the corresponding martingale problem [23, Theorem 6.1.7]. However, there can be more than one weak solution which can also be non-Markov (see Section 12.3 of [23]). In this work, under additional conditions on coefficients, we give a way of selecting weak solutions to (1.1) that are Feller. This problem was initially studied in [4] where the authors give a procedure to select a Feller solution under appropriate assumptions, one of which is difficult to verify in practice. This work is an extension of [4]. The main contribution of this work is proving Lemma 2 below under the relatively general explicit assumptions on $b(\cdot)$ and $\sigma(\cdot)$ (see Assumption 2). We also show that under these assumptions, there is a unique Feller solution which is given by the selection procedure (see Theorem 4).

A classical approach to selecting a Markov solution of the martingale problem in face of non-uniqueness is due to Krylov [20] (see also Section 12.2 of [23]). This is based on successive minimization of a countable family of functionals on the solution measures, each one being minimized over the set of minimizers of the preceding one, and then arguing that the intersection of the nested family of minimizers thus obtained leads to a Markov solution. The problem with this approach is that no uniqueness is claimed, nor can the possibility of dependence on the specific choice of functionals to be minimized and the order thereof be ruled out.

Usually, whenever there is an instance of non-uniqueness of solutions of any equation, the natural question that arises is: What is the most relevant or physical/natural solution among the available ones? Kolmogorov (quoted in [9, Pg. 626]) had suggested a way to answer this question viz., perturb the equation with small noise and if this perturbed equation has a unique solution, then taking the noise to zero will give us the physical solution, if the corresponding limit exists and is unique (see [11, Chapter 2] for a detailed discussion). This point of view is adopted in various works in the literature. Here we give a very incomplete collection of examples: In [18], the author identified special invariant measures of the limiting smooth dynamical systems. For $n = 1$ without diffusion, this problem is thoroughly studied under various set of assumptions in [1, 7, 12, 14, 19, 24]. In [14], a large deviation principle is established for the one-dimensional case without diffusion and $b$ of the form $\text{sgn}(x)|x|^\gamma$, for $1 > \gamma > 0$ (a generalisation of this result can be found in [15]). For multidimensional case without diffusion, see [8, 26]. In [5], the authors studied the case with bounded...
measurable drift and showed that the solution to the limiting case lies in the set of associated Fillipov solutions (see [10], for a definition). The selection procedure in [4] is inspired by this philosophy and is adopted in this work as well.

The paper is organised as follows: In Sect. 2, we set up the notation and give a few important definitions required for the rest of the paper. Selection procedure, along with the main results of the paper, are given in Sect. 3. The proofs of these results will be given in Sects. 4 and 5. Finally, we conclude by discussing the assumptions along with examples and counterexamples in Sect. 6.

**Remark 1** In what follows, by a solution of a stochastic differential equation we always mean a solution to its corresponding martingale problem.

### 2 Problem Formulation

Let $C_b(\mathbb{R}^n)$ and $C^u_b(\mathbb{R}^n)$ as the space of bounded continuous functions and space of bounded uniformly continuous functions on $\mathbb{R}^n$ endowed with the topology of uniform convergence on compact sets, respectively.

Since we are trying to study the Markovian nature of solutions of (1.1), it will be fruitful to consider the corresponding backward Kolmogorov equation i.e.,

\[
\begin{align*}
\partial_t u - \mathcal{L}(D^2 u, Du, x) &= 0 \text{ on } \mathcal{O} = \mathbb{R}^n \times (0, T), \\
u(x, 0) &= f(x) \in C^u_b(\mathbb{R}^n),
\end{align*}
\]

where

\[\mathcal{L}(M, p, x) \doteq \frac{1}{2} \text{Tr} \left\{ \sigma(x) \sigma(x)^\dagger M \right\} + b(x).p, \text{ for } (M, p, x) \in S^n \times \mathbb{R}^n \times \mathbb{R}^n\]

with $S^n$ the space of symmetric $n \times n$ matrices, $\text{Tr}(A)$ the trace of $n \times n$ matrix $A$, and $p, q$ the usual Euclidean inner product. Denote $\mathcal{O} \cup (\mathbb{R}^n \times \{0\})$ as $\mathcal{O}^*$.

**Remark 2** 1. Note that the (2.1) and (2.2) is written as an equivalent initial value problem rather than final value problem i.e.,

\[u(x, t) = \mathbb{E} \left[ f(X_T) | X_{T-t} = x \right]\]

(see Theorem 2 below). It is clear to see that these are equivalent ways of studying the above equation.

2. The $f$ above is chosen to be bounded uniformly continuous functions instead of just a bounded continuous function. As we shall in Proposition 2, this choice does not cost us any generality.

Following [9, Pg. 626] and [4], for $\varepsilon > 0$, consider the following stochastic differential equation:

\[
dX_t^\varepsilon = b^\varepsilon(X_t^\varepsilon)dt + \sigma^\varepsilon(X_t^\varepsilon)dW_t, \quad X_0^\varepsilon \sim \mu,
\]

\[\varepsilon > 0,\text{ for }\]
where, $b^\varepsilon$ and $\sigma^\varepsilon$ satisfy the following assumption:

**Assumption 1** $b$, $b^\varepsilon$, $\sigma$ and $\sigma^\varepsilon$ are bounded continuous such that

$$b^\varepsilon \to b \text{ and } \sigma^\varepsilon \to \sigma,$$

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

and $\sigma^\varepsilon(\sigma^\varepsilon) \dagger$ is uniformly positive definite, for every $\varepsilon > 0$.

**Remark 3** In [4], the authors considered $b^\varepsilon \equiv b$ and $\sigma^\varepsilon$ to be of the form

$$\sigma^\varepsilon(\cdot)(\sigma^\varepsilon(\cdot)) \dagger = \sigma(\cdot)\sigma(\cdot) \dagger + \varepsilon I.$$

Thus the selection procedure considered in [4] is a special case of the above setup.

It is well known that there exists a unique solution of $(2.3)$ [23, Theorem 7.2.1] and additionally, that this solution is a Feller process. Consider the corresponding backward Kolmogorov equation:

$$\partial_t u^\varepsilon - \mathcal{L}^\varepsilon(D^2 u^\varepsilon, Du^\varepsilon, x) = 0, \text{ on } \mathcal{O},$$

$$u^\varepsilon(x, 0) = f(x) \in C^b_b(\mathbb{R}^n)$$

where,

$$\mathcal{L}^\varepsilon(M, p, x) \doteq \frac{1}{2} \text{Tr} \left\{ \sigma^\varepsilon(x)(\sigma^\varepsilon(x)) \dagger M \right\} + b^\varepsilon(x).p, \text{ for } (M, p, x) \in S^n \times \mathbb{R}^n \times \mathbb{R}^n.$$

**Remark 4** As mentioned above, $(2.4)$ and $(2.5)$ are written as the initial value problem rather than final value problem. This is for sake of the ease of invoking already existing results about continuous viscosity solutions, which are usually stated for parabolic p.d.e.s in this form.

**Definition 1** The set of all points $x \in \mathbb{R}^n$ such that there exists a neighborhood $U$ of $x$ such that

$$\|b(y) - b(z)\| + \|\sigma(y) - \sigma(z)\| \leq K_U(x)\|y - z\|, \text{ for every } y, z \in U.$$

is denoted by $L_{b,\sigma}$ and its complement by $NL_{b,\sigma} \doteq (L_{b,\sigma})^c$.

Following is the main assumption of this work.

**Assumption 2** $b(\cdot): \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma(\cdot): \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$ (allowed to be degenerate) are bounded and satisfy the following: There are constants $C_b$ and $C_\sigma > 0$ such that

1. For $x, y \in \mathbb{R}^n$, $\|b(x) - b(y)\| \leq C_b\|x - y\|^\alpha$ and $\|\sigma(x) - \sigma(y)\| \leq C_\sigma\|x - y\|^\beta$, where $\alpha$ and $\beta$ are chosen as below.

2. For every $x \in NL_{b,\sigma}$, there is $r > 0$ (depending on $x$) such that $b(x) = 0$, $\sigma(x) = 0$ and

$$C_\sigma\|x - y\|^{2\beta}\|v\|^2 \geq v^\dagger \sigma(y)(\sigma(y))^\dagger v \geq C_\sigma^{-1}\|x - y\|^{2\beta}\|v\|^2,$$

for $y \in Br(x) \cap L_{b,\sigma}$. Here, $B_r(x)$ is the open ball of radius $r$ around $x$. 

\[\square\] Springer
3. $\alpha$ and $\beta$ are such that $1 + \alpha - 2\beta > 0$ and $\beta > \frac{1}{2}$.

We will show later that the above assumption implies that there is a unique Feller solution of (1.1) which is a corollary of the one of the main results of this work (See Theorem 2). The assumptions are discussed in Section 6.

To proceed further, we introduce the notion of continuous viscosity solutions (see [6] for an excellent survey). For a locally bounded function $w : \mathcal{O} \rightarrow \mathbb{R}$ and $(x, t) \in \mathcal{O}$, we define

$$P_{\mathcal{O}}^{2,+} w(x, t) = \{ (a, p, X) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n : \text{for } \mathcal{O} \ni (y, s) \rightarrow (x, t), \ w(y, s) \leq w(x, t) + a(s - t) + p.(y - x) + \frac{1}{2} (y - x)^\dagger X(y - x) + o(|t - s| + \|x - y\|^2) \}$$

and $P_{\mathcal{O}}^{2,-} w = -P_{\mathcal{O}}^{2,+} (-w)$. Also,

$$\overline{P}_{\mathcal{O}}^{2,+} w(x, t) = \{ (a, p, X) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n : \exists (a_n, p_n, X_n, (x_n, t_n)) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n \times \mathcal{O} \text{ such that } (a_n, p_n, X_n) \in P_{\mathcal{O}}^{2,+} w(x_n, t_n) \text{ and } (a_n, p_n, X_n, (x_n, t_n), u(x_n, t_n)) \overset{n \rightarrow \infty}{\longrightarrow} (a, p, X, (x, t), u(x, t)) \}$$

and $\overline{P}_{\mathcal{O}}^{2,-} w(x, t)$ is defined analogously.

**Definition 2** An upper semicontinuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is called a viscosity subsolution of (2.1) and (2.2) if

$$a - \mathcal{L}(X, p, x) \leq 0, \text{ for } (x, t) \in \mathcal{O} \text{ and } (a, p, X) \in P_{\mathcal{O}}^{2,+} u(x, t), \quad (2.6)$$

and $u(x, 0) \leq f(x)$.

A lower semicontinuous function $u : \mathcal{O} \rightarrow \mathbb{R}$ is called a viscosity supersolution of (2.1) and (2.2) if

$$a - \mathcal{L}(X, p, x) \geq 0, \text{ for } (x, t) \in \mathcal{O} \text{ and } (a, p, X) \in P_{\mathcal{O}}^{2,-} u(x, t), \quad (2.7)$$

and $u(x, 0) \geq f(x)$.

A function $u : \mathcal{O} \rightarrow \mathbb{R}$ is called a continuous viscosity solution of (2.1) and (2.2) if it is both a viscosity subsolution and a viscosity supersolution.

**Remark 5**

1. Similar definitions hold for (2.4) and (2.5).

2. Equivalent definitions of viscosity subsolution and supersolution can be found in [21, Pg. 1237-1238].

3. In what follows, sometimes we refer to viscosity subsolutions (viscosity supersolutions, respectively) as just subsolutions (supersolutions, respectively). Also, we
refer to an upper semi-continuous function (a lower semi-continuous function, respectively) on \( O \) as a subsolution (supersolution, respectively) even if it satisfies only (2.6) ((2.7), respectively).

We will later see that under the Assumptions 1 and 2, a Feller solution corresponds to a continuous viscosity solution and vice versa.

As mentioned already, we are interested in studying the limiting behavior of \( X^\varepsilon \) as \( \varepsilon \to 0 \). Since \( X^\varepsilon \) is a Feller process, this can be done by studying the limiting behavior of \( u^\varepsilon = E[f(X^\varepsilon_T) | X^\varepsilon_T = x] \) as \( \varepsilon \to 0 \). In general, we do not have strong uniform estimates (in \( \varepsilon \)) on \( u^\varepsilon \) that guarantee convergence. The best a priori estimate that we have in general is the following:

\[
\|u^\varepsilon(\cdot, t)\|_\infty \leq \|f\|_\infty, \text{ for } 0 \leq t < T.
\]

**Remark 6** From the boundedness of \( b \) and \( \sigma \), one can easily get the following continuity estimate uniform in \( \varepsilon \):

\[
\|u^\varepsilon(x, t) - u^\varepsilon(x, s)\| \leq K\|t - s\|, \text{ for } x \in \mathbb{R}^n \text{ and } s, t \in [0, T).
\]

Additionally, if \( b \) and \( \sigma \) are globally Lipschitz and \( f \) is smooth enough, then there exists \( K^* > 0 \) such that for any \( t \in [0, T) \), we have

\[
\|u^\varepsilon(x, t) - u^\varepsilon(y, t)\| \leq K^*\|x - y\|^{1/2}, \text{ for } x, y \in \mathbb{R}^n.
\]

(2.8)

Indeed, we first observe that for any \( \varepsilon > 0 \), there is a unique strong solution to (2.3) for every initial condition \( X^\varepsilon_0 = x \) and hence, for any Brownian motion, a simple application of Ito’s formula and Ito’s isometry to \( f(X^\varepsilon_t, x) - f(X^\varepsilon_t, y) \) proves (2.8). Here, \( X^\varepsilon_{\varepsilon, z} \) is written in place of \( X^\varepsilon_t \) to emphasize the dependence of initial condition \( z \).

Define functions \( u^* \) and \( u_* \) on \( O^* \) as

\[
u^*(x, t) = \limsup_{\delta \downarrow 0, \varepsilon > 0} \left\{ u^\varepsilon(y, s) : \|x - y\| + |t - s| < \delta, \varepsilon < \delta \right\}
\]

and

\[
u_*(x, t) = \liminf_{\delta \downarrow 0, \varepsilon > 0} \left\{ u^\varepsilon(y, s) : \|x - y\| + |t - s| < \delta, \varepsilon < \delta \right\}.
\]

Clearly, \( u^* \geq u_* \) and

\[
\|u^*(\cdot, t)\|_\infty, \|u_*(\cdot, t)\|_\infty \leq \|f\|_\infty, \text{ for } 0 \leq t < T.
\]

(2.9)

From the previous remark and the definition of \( u^* \) and \( u_* \), it is easy to conclude that for any fixed \( x \in \mathbb{R}^n \), \( u^*(x, \cdot) \) and \( u_*(x, \cdot) \) are continuous functions.
3 Selection Procedure and the Main Results

In this section, we provide a procedure that selects a unique Feller solution of (1.1) under mild conditions. More precisely, under Assumptions 1 and 2, we consider a family of processes \( \{X^\varepsilon\} \) such that they are tight and converge weakly to a limit. This limit will turn out to be the unique Feller solution of (1.1).

The selection procedure goes as follows: For \( \{X^\varepsilon\}_{\varepsilon > 0} \) as defined in Sect. 2:

1. Establish that \( \{X^\varepsilon\}_{\varepsilon > 0} \) is tight in \( C \).
2. Show that one of the limit points in law, \( X^* \) (say), is a Feller process.
3. Show that the limit point \( X^* \) is in fact a solution of (1.1).
4. Finally, show that there is a unique Feller solution of (1.1).

Through the above, we would have established that the given selection procedure picks out the unique Feller process from the set of solutions of (1.1). We again remind the reader that the above selection procedure was already described in [4], but under the assumptions that are not all easily verifiable. Also, in step 2, the authors show the Feller property for a sequential limit \( X^* \), which raises the question concerning what other sequential limit points might be like. The goal of this paper is to perform steps 1, 2, 3 and 4, where step 2 is as follows:

2. Show that one of the limit point \( X^* \) is Feller process and that any limit point shares the same finite dimensional distribution as that of this Feller process. From the continuity of paths, this in turn, means that the processes are identical in terms of distribution. (This is because ‘cylinder sets’ of the type \( \{x(\cdot) : x(t_i) \in A_i, 1 \leq i \leq n\} \) for \( n \geq 1, t_i \in [0, T], A_i \) open in \( \mathbb{R}^n \), generate the full Borel \( \sigma \)-field of \( C ([0, T]; \mathbb{R}^n) \).)

With the intention of making this work self-contained, we present the results of [4] that are used in this work.

**Theorem 1** Under Assumption 1, \( \{X^\varepsilon\}_{\varepsilon > 0} \) is tight in \( C \).

**Proof** From the definition of \( X^\varepsilon \), for \( 0 \leq s \leq t \leq T \), we have

\[
X^\varepsilon_t - X^\varepsilon_s = \int_s^t b^\varepsilon(X^\varepsilon_u)du + \int_s^t \sigma^\varepsilon(X^\varepsilon_u)dW_u
\]

Now consider

\[
\|X^\varepsilon_t - X^\varepsilon_s\|^4 \leq 8\|\int_s^t b^\varepsilon(X^\varepsilon_u)du\|^4 + 8\|\int_s^t \sigma^\varepsilon(X^\varepsilon_u)dW_u\|^4
\]

\[
\leq 8K^4|t-s|^4 + 24K^4|t-s|^2, \text{ from [22, Lemma 4.1]},
\]

where, \( K \doteq \max\{\|b\|_\infty, \|\sigma\|_\infty\} \).

Therefore, using [3, Theorem 12.3], we conclude the desired tightness. \( \square \)

Before we proceed, we recall the following result from [4]:

\( \square \) Springer
Theorem 2 If there is a Feller solution to (1.1), then there is a continuous viscosity solution to (2.1) and (2.2). The continuous viscosity solution is given by

\[ u(x, t) = E[f(X_T)|X_{T-t} = x]. \]

Remark 7 Analogous statement also holds for (2.3), (2.4) and (2.5).

Using the theorem above, we can establish the existence of \( u^\varepsilon \) using a probabilistic argument. Thus we have:

Corollary 1 Under Assumption 1, there exists a continuous viscosity solution to (2.4) and (2.5).

Proof To see this, we note that (2.3) has a unique weak solution (denoted by \( X^\varepsilon \)) and hence it is a Feller solution. Therefore from the above remark, there is a continuous viscosity solution given by

\[ u^\varepsilon(x, t) = E[f(X_{T}^\varepsilon)|X_{T-t}^\varepsilon = x]. \]

Proposition 1 Under Assumption 1, \( u^* \) and \( u_* \) are subsolution and supersolutions of (2.1) and (2.2), respectively.

Remark 8 The proof below is almost the same as the proof of [2, Theorem V.1.7] and is given here for the sake of completeness. We only show that \( u^* \) is a subsolution of (2.1) and (2.2), as showing that \( u_* \) is a supersolution can be done in a similar way.

Proof From an equivalent definition of viscosity subsolution (see [21, Remark I.9]), it suffices to show that if \((\bar{x}, \bar{t})\) is a strict local maximum of \( u^* - \phi \) for a twice continuously differentiable \( \phi \) on \([0, T] \times \mathbb{R}^n\), then

\[ \partial_t \phi(\bar{x}, \bar{t}) - \mathcal{L}(D^2 \phi(\bar{x}, \bar{t}), D\phi(\bar{x}, \bar{t}), \bar{x}) \leq 0. \]

As in [2, Lemma V.1.6], we can show that there is a subsequence (still denoted by superscript \( \varepsilon \)) such that \((x^\varepsilon, t^\varepsilon)\) is a local maximum of \( u^\varepsilon - \phi \) and \((x^\varepsilon, t^\varepsilon) \to (\bar{x}, \bar{t})\). Using the fact that \( u^\varepsilon \) is a subsolution, we get

\[ \partial_t \phi(x^\varepsilon, t^\varepsilon) - \mathcal{L}^\varepsilon(D^2 \phi(x^\varepsilon, t^\varepsilon), D\phi(x^\varepsilon, t^\varepsilon), x^\varepsilon) \leq 0. \] (3.1)

Now taking \( \varepsilon \to 0 \) and using uniform convergence on compact sets of \( \mathcal{L}^\varepsilon \) to \( \mathcal{L} \), we get the result. Finally, from [13, Proposition VII.5.1] we have

\[ u^*(x, 0) = f(x), \text{ for } x \in \mathbb{R}^n. \]

This proves that \( u^* \) is a viscosity subsolution of (2.1) and (2.2). \( \square \)
The functions $u^*$ and $u_*$ are the possible candidates for continuous viscosity solution of (2.1) and (2.2). If $u^* = u_* = \bar{u}$ on $O$, then clearly, $\bar{u}$ is the desired continuous viscosity solution. Even though $u^* \geq u_*$ trivially, we cannot say that $u^* \leq u_*$ (and hence $u^* = u_*$) in general without a comparison principle which says that under appropriate conditions, for a subsolution $u$ and supersolution $v$ of (2.1), we have $u(x, t) \leq v(x, t)$ on $O$ whenever $u(x, 0) \leq v(x, 0)$ on $\mathbb{R}^n$. Before we proceed to establish the aforementioned comparison principle, we give a very important consequence of equality of $u^*$ and $u_*$ on $O$. The following lemma follows as in [2, Lemma 1.9].

**Lemma 1** Suppose $u^* = u_* = \bar{u}$. Then

$$u^\varepsilon \to \bar{u}, \text{ uniformly on the compact sets of } O^*.$$

The comparison principle given below is in greater generality than what is required for our purposes.

**Lemma 2** (Comparison principle) Let $u$ and $v$ be bounded sub and supersolutions of (2.1), respectively and that either $u(\cdot, 0) \in C^u_b(\mathbb{R}^n)$ or $v(\cdot, 0) \in C^u_b(\mathbb{R}^n)$. Suppose Assumption 2 holds. Then

$$\sup_{O^*} [u - v] = \sup_{x \in \mathbb{R}^n} \{ [u(x, 0) - v(x, 0)] \lor 0 \}. \quad (3.2)$$

**Remark 9** 1. The proof (which will be given in Sect. 4) adapts the techniques of [13, Theorem II.9.1] and [25, Theorem 1.3.2].

2. $u^*$ and $u_*$ satisfy the hypothesis of this lemma. Indeed, using [13, Proposition VII.5.1] and the fact that $f \in C^u_b(\mathbb{R}^n)$, we have $u^*(x, 0) = u_*(x, 0) = f(x)$, for $x \in \mathbb{R}^n$. Therefore, this lemma implies that $u^* = u_*$. To summarize, until now we have proved that $u^* (= u_*)$ is a continuous viscosity solution whenever $f \in C^u_b(\mathbb{R}^n)$. In the following proposition, we extend this result to allow for $f$ to lie in $C_b(\mathbb{R}^n)$.

**Proposition 2** Suppose $u^f$ stands for $u^*$ when $u^*(\cdot, 0) = f \in C_b(\mathbb{R}^n)$. Then $u^f$ is the continuous viscosity solution of (2.1).

**Proof** First, we show that for $f_1, f_2 \in C^u_b(\mathbb{R}^n)$,

$$u^{f_1 + f_2} = u^{f_1} + u^{f_2}. \quad (3.3)$$

From Lemma 2 and 1, we know that for $0 \leq t < T$ as $\varepsilon \to 0$,

$$\|u^{\varepsilon, f_1}(\cdot, t) - u^{f_1}(\cdot, t)\|_\infty, \|u^{\varepsilon, f_2}(\cdot, t) - u^{f_2}(\cdot, t)\|_\infty, \|u^{\varepsilon, f_1 + f_2}(\cdot, t) - u^{f_1 + f_2}(\cdot, t)\|_\infty \to 0. \quad (3.4)$$

Here, $u^{\varepsilon, g}$ denotes the continuous viscosity solution of (2.4) whenever $u^{\varepsilon, g}(\cdot, 0) = g$. It is clear from the proof of Corollary 1 that for any $\varepsilon > 0$, $u^{\varepsilon, g}$ is linear in $g$, for $g \in C_b(\mathbb{R}^n)$. Together with (3.4), the aforementioned linearity of $u^{\varepsilon, g}$ gives us (3.3).
To keep the notation standard, let us write $u^f(x, t)$ as $(P_t f)(x)$. From (2.9), we have

$$\|P_t f\|_\infty \leq \|f\|_\infty, \text{ for } 0 \leq t < T.$$ 

To prove the statement of the proposition, we note that $C^u_b(\mathbb{R}^n)$ is dense in $C_b(\mathbb{R}^n)$ under the topology of uniform convergence on compact sets of $\mathbb{R}^n$. Hence, $P_t$ can be extended continuously to $C_b(\mathbb{R}^n)$ viz., for $f \in C_b(\mathbb{R}^n)$ and $0 \leq t < T$

$$P_t f \doteq \lim_{n \to \infty} P_t f_n,$$

whenever \( \{f_n\}_{n \geq 1} \subset C^u_b(\mathbb{R}^n) \) such that $f_n \to f$ as $n \to \infty$ uniformly on compact sets of $\mathbb{R}^n$. In fact, the above limit is uniform over any compact set of $[0, T)$.

From Theorem 2, we know that every Feller process gives a corresponding continuous viscosity solution. The following result from [4] gives the converse of this statement viz., a continuous viscosity solution of (2.1) gives a corresponding Feller process.

**Theorem 3** [4, Theorem 3.2] Suppose Assumption 1 and the statement of Lemma 2 hold. Then all the limit points (in the sense of weak convergence) are solutions to the martingale problem corresponding to (1.1) and one of the limit points is Feller. Also, if $\tilde{X}$ and $\bar{X}$ are two limit points, then for $0 \leq s \leq t \leq T$, $(\tilde{X}_s, \tilde{X}_t)$ and $(\bar{X}_s, \bar{X}_t)$ has the same laws.

The above theorem does not prevent the limit points from having different finite dimensional distributions. The following result says that all limit points share the same finite dimensional distribution and consequently, the same laws.

**Theorem 4** Under Assumptions 1 and 2, $X^\varepsilon$ converges weakly to $X^*$. Here, $X^*$ is the Feller process obtained in Theorem 3.

Proof of this result is given in Sect. 5.

**Remark 10** In [4], the authors showed that this result holds for $k \leq 2$ i.e., for any $f_1, f_2 \in C_b(\mathbb{R}^n)$, we have

$$\lim_{\varepsilon \to 0} \mathbb{E}_{X^\varepsilon} \left[ f_1(X^\varepsilon_s) f_2(X^\varepsilon_t) \right] = \mathbb{E}_\tilde{X} \left[ f_1(\tilde{X}_s) f_2(\tilde{X}_t) \right], \ s \leq t, \quad (3.5)$$

for every (weak) limit point $\tilde{X}$ of $\{X^\varepsilon\}_{\varepsilon > 0}$. Here, we are considering the possibility of distinct limit points of $X^\varepsilon$ as we do not apriori know that $X^\varepsilon$ is weakly convergent.

**Remark 11** From Lemma 2, uniqueness of continuous viscosity solution of (2.1) and (2.2) follows. This means that among all the solutions of (1.1), there exists a unique Feller solution.
Remark 12 In the following, we make the distinction between the classical results of [1] and ours. The article [1] consider the following system: $n = 1$ and $\sigma \equiv 0$. The assumptions in this article involves appropriate kind of the local integrability on $b^{-1}$ at say, $x = 0$, for simplicity. This is relatively weaker than the Assumption 2 for the case $n = 1$ and $\sigma \equiv 0$. This is because the Assumption 2 involves some kind of uniform continuity of $b$ at $x = 0$ (in particular, Hölder continuity) which is stronger than the integrability of $b^{-1}$ that is considered in [1]. This is clearly evident from the the discussion in [4, Pg. 746], where authors discuss the following two examples considered in [1] and show that there does not exist a Feller-Markov solution.

1. For $0 < \alpha < \beta < 1$, $b(x) = x^\alpha \log x$, for $x \geq 0$ and $= -|x|^\beta$, for $x < 0$.
2. For $\alpha < 1$, $b(x) = x^\alpha$, for $x \geq 0$ and $= -3|x|^\alpha$, for $x < 0$.

4 Proof of Lemma 2

We split the proof into various lemmas. The proof we give is by contradiction. Before we proceed, we make the following transformation: $u(x, t) \mapsto e^{-\gamma t}u(x, t)$ and $v(x, t) \mapsto e^{-\gamma t}v(x, t)$ with $\gamma > 0$. This helps us later in establishing the contradiction. With abuse of notation, we denote the transformed $u$ and $v$ also by $u$ and $v$. By assuming that $u$ and $v$ are smooth, we can guess the equation that $u$ and $v$ satisfy to be

$$\partial_t u + \gamma u - \mathcal{L}(D^2 u, D u, x) = 0.$$ 

The fact that this indeed is the case is given by [13, Pg. 98, Lemma II.9.1] ([13, Lemma II.9.1] is applicable only for the first order case, proof for second order case is exactly along the same lines). Since $\gamma > 0$ is arbitrary, it suffices to show (3.2) for the transformed $u$ and $v$.

We consider two cases: Recall that $\mathcal{O}^* = \mathbb{R}^n \times [0, T)$.

1. (Case 1)

$$\sup_{\mathcal{O}^*} [u - v] > u(x, t) - v(x, t), \forall (x, t) \in NL_{b,\sigma} \times [0, T).$$

2. (Case 2)

$$\sup_{\mathcal{O}^*} [u - v] = u(x_*, t_*) - v(x_*, t_*), \text{ for some } (x_*, t_*) \in NL_{b,\sigma} \times [0, T).$$

In the following, we prove that the assertion of Lemma 2 holds in both the above cases.

Proposition 3 Suppose that the assumptions of Lemma 2 hold and that

$$\sup_{\mathcal{O}^*} [u - v] > u(x, t) - v(x, t), \forall (x, t) \in NL_{b,\sigma} \times [0, T).$$

Then (3.2) holds.
Proof We assume the contrary to (3.2) i.e.,

$$\sup_{O^*} [u - v] - \delta > \sup_{x \in \mathbb{R}^n} \{[u(x, 0) - v(x, 0)] \vee 0\},$$  

(4.1)

for some $\delta > 0$.

We define $u^\rho$ as $u - \rho T - t$. Now consider the following auxiliary function: For $\alpha, \rho, \beta > 0$, define

$$M_{\alpha, \rho, \beta} : \mathbb{R}^n \times \mathbb{R}^n \times [0, T) \to \mathbb{R}$$

by

$$M_{\alpha, \rho, \beta}(x, y, t) = u^\rho(x, t) - v(y, t) - \alpha \|x - y\|^2 - \beta \|x\|^2,$$  

(4.2)

From the behavior of $-\beta \|x\|^2$, boundedness of $u, v$ and upper semicontinuity of $M_{\alpha, \beta, \rho}$, we know that there exists $(\hat{x}, \hat{y}, \hat{t}) \in \mathbb{R}^n \times \mathbb{R}^n \times [0, T)$ which is a maximizer of $M_{\alpha, \beta, \rho}$ on $\mathbb{R}^n \times \mathbb{R}^n \times [0, T)$.

Using Lemma 6 in the Appendix, we can ensure that, for small enough $\rho$ and $\beta$

$$\sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0, T)} M_{\alpha, \beta, \rho} > \eta > 0,$$

for some $\eta > 0$. From the arguments of [16, Pg. 916, Eq. 3.11], we have

$$\lim_{\alpha \to \infty} \lim_{\beta, \rho \to 0} \alpha \|\hat{x} - \hat{y}\|^2 = 0.$$  

(4.3)

From Equation (A.5) in the Appendix, we have $(x_0, y_0, t_0)$ such that

$$M_{\alpha, \beta, \rho}(x_0, y_0, t_0) \geq \sup_{O^*} [u(x, t) - v(x, t)] - 5 \epsilon, \text{ for small enough } \epsilon$$

(4.4)

$$\implies \sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0, T)} [M_{\alpha, \beta, \rho}] \geq \sup_{O^*} [u(x, t) - v(x, t)] - 5 \epsilon.$$  

(4.5)

From Equation (4.3), we have

$$\lim_{\alpha \to \infty} \lim_{\beta, \rho \to 0} \sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0, T)} [M_{\alpha, \beta, \rho}] = \sup_{O^*} [u - v] > u(x, t) - v(x, t), \forall (x, t) \in NL_{b, \sigma} \times [0, T).$$

The inequality above follows from the arbitrariness of $\epsilon$.

Therefore, we have shown that for large enough $\alpha$ and small enough $\beta$ and $\rho$, there is a $r > 0$ such that $\hat{x}, \hat{y} \notin B_r(NL_{b, \sigma})$.

Since from the assumption of the contrary of (3.2) and choice of $\alpha, \beta, \rho, (\hat{x}, \hat{y}, \hat{t})$ are interior points of $O^*$, using [6, Theorem 8.3], we know that there exist $a, b, X, Y$ such that

$$(a, \alpha(\hat{x} - \hat{y}), X) \in \overline{D}^{2,+}(\hat{u}(\hat{x}, \hat{t}) - \beta \|\hat{x}\|^2),$$
such that \( a - b = 0 \) and
\[
-3\alpha \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq 3\alpha \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}, \tag{4.6}
\]
where \( I \) is the identity matrix in \( \mathbb{R}^{n \times n} \). We note here that the hypothesis of the \[6, \text{Theorem 8.3}\] requires only upper semicontinuity of \( u \) and taking the trace gives us
\[
\begin{align*}
\langle b, \alpha(\hat{x} - \hat{y}) - Y \rangle & \in \overline{P^2} \cdot v(\hat{y}, \hat{i}), \\
\text{where } \alpha & \text{ is the identity matrix in } \mathbb{R}^{n \times n}.
\end{align*}
\]
In the above, we used the fact that left multiplying \( (4.6) \) with \( b \) and small enough \( \alpha \) and hence \[6, \text{Theorem 8.3}\] can be applied. Using the fact that \( u \) and \( v \) are sub and supersolution respectively, we get
\[
\begin{align*}
a + \gamma \left( u^\rho(\hat{x}, \hat{i}) - \beta \|\hat{x}\|^2 \right) - \mathcal{L}(\hat{x}, \alpha(\hat{x} - \hat{y}) + 2\beta \hat{x}, X + 2\beta I) & \leq -\frac{\rho}{T^2}, \\
b + \gamma v(\hat{y}, \hat{i}) - \mathcal{L}(\hat{x}, \alpha(\hat{x} - \hat{y}), Y) & \geq 0.
\end{align*}
\]
From above, we have
\[
\begin{align*}
\gamma \left( u(\hat{x}, \hat{y}) - v(\hat{y}, \hat{i}) \right) - \frac{\gamma \rho}{T - \hat{i}} - \gamma \beta \|\hat{x}\|^2 - \beta \langle b(\hat{x}), \hat{x} \rangle - \beta \text{Tr} \left\{ \sigma(\hat{x})\sigma(\hat{x})^\dagger \right\} \\
- \langle b(\hat{x}) - b(\hat{y}), \alpha(\hat{x} - \hat{y}) \rangle - \text{Tr} \left\{ \sigma(\hat{x})\sigma(\hat{x})^\dagger X - \sigma(\hat{y})\sigma(\hat{y})^\dagger Y \right\} & \leq -\frac{\rho}{T^2}.
\end{align*}
\]
From \( (4.3) \) and the fact that there is a \( r > 0 \) such that \( \hat{x}, \hat{y} \notin B_r(NL_b, \sigma) \) (for large enough \( \alpha \) and small enough \( \beta \) and \( \rho \)), we have a \( K_r > 0 \) such that
\[
\|b(\hat{x}) - b(\hat{y})\| + \|\sigma(\hat{x}) - \sigma(\hat{y})\| \leq K_r\|\hat{x} - \hat{y}\|.
\]
This gives us
\[
\begin{align*}
\gamma \left( u(\hat{x}, \hat{y}) - v(\hat{y}, \hat{i}) \right) - \frac{\gamma \rho}{T - \hat{i}} - \gamma \beta \|\hat{x}\|^2 - \beta \langle b(\hat{x}), \hat{x} \rangle \\
- \text{Tr} \left\{ \sigma(\hat{x})\sigma(\hat{x})^\dagger \right\} + \frac{\rho}{T^2} & \leq 3\alpha \|\sigma(\hat{x}) - \sigma(\hat{y})\|^2 + K_r\|\hat{x} - \hat{y}\|^2 \\
& \leq 3\alpha K_r\|\hat{x} - \hat{y}\|^2. \tag{4.7}
\end{align*}
\]
In the above, we used the fact that left multiplying \( (4.6) \) with
\[
\Sigma = \begin{pmatrix} \sigma(\hat{x})\sigma(\hat{x})^\dagger & \sigma(\hat{y})\sigma(\hat{x})^\dagger \\ \sigma(\hat{x})\sigma(\hat{y})^\dagger & \sigma(\hat{y})\sigma(\hat{y})^\dagger \end{pmatrix}
\]
and taking the trace gives us
\[
\text{Tr} \left\{ \sigma(\hat{x})\sigma(\hat{x})^\dagger X - \sigma(\hat{y})\sigma(\hat{y})^\dagger Y \right\} \leq 3\alpha \text{Tr} \left\{ ((\sigma(\hat{x}) - \sigma(\hat{y}))(\sigma(\hat{x}) - \sigma(\hat{y}))^\dagger \right\} \\
\leq 3\alpha \|\sigma(\hat{x}) - \sigma(\hat{y})\|^2.
\]
Finally, taking $\beta, \rho \downarrow 0$ and then $\alpha \uparrow \infty$ in (4.7), we have a contradiction, since it leads to

$$0 < \gamma \left( \sup_{O^*} [u - v] \right) \leq 0.$$  

□

**Proposition 4** Suppose that the assumptions of Lemma 2 hold and that

$$\sup_{O^*} [u - v] = u(x_*, t_*) - v(x_*, t_*),$$  

(4.8)

for some $(x_*, t_*) \in NL_{b,\sigma} \times [0, T)$. Then (3.2) holds.

The proof of this result is given later and it proceeds in the following way: we select a very special supersolution of (2.1) denoted by $\psi$ in order to understand the behavior of $u - v$ around $(x_*, t_*)$. This special subsolution $\psi$ is such that

1. $u - \psi$ and $v + \psi$ are also subsolutions and supersolutions, respectively.
2. $\psi(x, t) = h(\|x - x_*\|)$, for some modulus of continuity $h$.

A few remarks are in order before we proceed. It is important to note that it is not apriori clear that $u + v$ is a subsolution whenever $u$ and $v$ are subsolutions. Hence, asking for $u + \psi$ ($v - \psi$, respectively) to be a subsolution (supersolution, respectively) is a non-trivial requirement. It will be clear from the proof that if we desired to relax the Assumption 2 to a much weaker assumption which says that $b$ and $\sigma$ are merely bounded uniformly continuous, then the current result will continue to hold as long as one can find a very special subsolution that satisfies the aforementioned properties. As constructing the subsolutions or supersolutions is much easier than constructing viscosity solutions, this approach will hopefully be helpful in studying cases where $b$ and $\sigma$ are merely bounded uniformly continuous.

In the following lemma, we establish the existence of $\psi$.

**Lemma 3** For $K > 0$ and $x_* \in NL_{b,\sigma}$,

$$\psi = K \|x - x_*\|^\gamma$$

is a supersolution of (2.1) over a neighbourhood around $x_*$ whenever $\gamma > \max\{2(1 - \beta), 1 - \alpha\}$. Additionally, $u - \psi$ and $v + \psi$ are subsolution and supersolutions of (2.1), respectively.

**Proof** From the linearity of (2.1), it is clear that without loss in generality we can take $K = 1$. We first prove that for $\theta > 0$

$$\psi_\theta(x, t) \doteq \left( \|x - x_*\|^2 + \theta \right)^{\frac{\gamma}{2}}$$

is a supersolution of (4.9) below. Before we do that, note that as $\theta \downarrow 0$, $\psi_\theta \to \psi$ uniformly on the compact sets of $\mathbb{R}^n$. Since $\psi_\theta$ is twice differentiable in $x$ and smooth
Assumption 2, i.e., supersolution for some $C$ whenever $x_\xi > 0$ for small $\xi > 0$ and sufficiently small $\theta$ (depending on $\xi$). Firstly, note that from Assumption 2, $b(x) = 0$ and $A(x) = 0$, whenever $x \in N L_{b,\sigma}$. Therefore, $L_\theta(x) = 0$ whenever $x \in N L_{b,\sigma}$. This means that we only have to show that $L_\theta(x) \geq 0$ for $x \in L_{b,\sigma}$.

$\partial_t \psi_\theta - b(x).\nabla \psi_\theta - Tr\{A(x)\Delta \psi_\theta\} = -y \frac{\gamma}{\|x - x_*\|^2 + \theta} b(x). (x - x_*)$

$\geq -\gamma C_b \|x - x_*\|^\alpha \|x - x_*\|$

from Assumption 2.

In the above, we used

\[ \frac{\partial}{\partial x_j} \frac{\gamma (x - x_*)_i}{\|x - x_*\|^2 + \theta} = \frac{\gamma \delta_{ij}}{\|x - x_*\|^2 + \theta} - \frac{2\gamma (1 - \frac{\gamma}{2}) (x - x_*)_i (x - x_*)_j}{\|x - x_*\|^2 + \theta} \]

From Assumption 2, it is clear that

\[ C_\sigma^{-1} \|x - x_*\|^{2\beta} \leq Tr\{A(x)\} \leq T(x) \leq C_\sigma n \|x - x_*\|^2 \]

for some $C_\sigma > 0$. Thus we have

\[ L_\theta(x) \geq \frac{-\gamma C_b}{\|x - x_*\|^2 + \theta} \|x - x_*\|^{1+\alpha} \]

\[ \geq -\gamma C_\sigma \|x - x_*\|^{2\beta} + \frac{2\gamma (1 - \frac{\gamma}{2}) \|x - x_*\|^{2\beta} \|x - x_*\|^2}{n C_\sigma \|x - x_*\|^2 + \theta} \]

\[ \]
\[
\geq \frac{1}{\{\|x - x_\ast\|^2 + \theta\}^{1 - \frac{\gamma}{2}}} \left[ -\gamma C_b \|x - x_\ast\|^\alpha + \frac{\gamma C_\sigma}{n} \|x - x_\ast\|^{2\beta} + \frac{2\gamma(1 - \frac{\gamma}{2})}{n C_\sigma \{\|x - x_\ast\|^2 + \theta\}^{1 - \frac{\gamma}{2}}} \|x - x_\ast\|^{2\beta + 2} \right].
\]

(4.10)

It is clear that if \( \gamma > \max\{2(1 - \beta), 1 - \alpha\} \), then \( \alpha + \gamma - 1 > 0, 2\beta + \gamma - 1 > 0 \) and \( 2\beta + \gamma - 2 > 0 \). This means that as \( \theta \to 0 \),

\[
\|x - x_\ast\|^{\alpha + 1} \to \|x - x_\ast\|^{\alpha + \gamma - 1},
\]

(4.11)

\[
\|x - x_\ast\|^{2\beta} \to \|x - x_\ast\|^{2\beta + \gamma - 1},
\]

and

\[
\|x - x_\ast\|^{2\beta + 2} \to \|x - x_\ast\|^{2\beta + \gamma - 2}
\]

(4.13)

uniformly on compact sets of \( \mathbb{R}^n \). Therefore for small \( \xi > 0 \), small enough \( \theta \) and on a large closed ball of \( \mathbb{R}^n \),

\[
\mathcal{L}_\theta(x) \geq -\gamma C_b \|x - x_\ast\|^\alpha + \frac{\gamma C_\sigma}{n} \|x - x_\ast\|^{2\beta + \gamma - 1} + \frac{2\gamma(1 - \frac{\gamma}{2})}{n C_\sigma \{\|x - x_\ast\|^2 + \theta\}^{1 - \frac{\gamma}{2}}} \|x - x_\ast\|^{2\beta + \gamma - 2} - 3\xi.
\]

(4.14)

Since \( 1 + \alpha - 2\beta > 0 \) from Assumption 2, we can conclude that on a neighbourhood around \( x_\ast \), the sum of the first three terms in the above equations is non-negative and this implies that \( \psi_\theta \) is a supersolution of (4.9). Since \( v \) is a supersolution of (2.1), it is automatically a supersolution of (4.9). This together with the smoothness of \( \psi_\theta \) concludes that \( v + \psi_\theta \) is a supersolution of (4.9) (see Lemma 5).

Noting that \( v + \psi_\theta \to v + \psi \), uniformly on compact sets, we can conclude that \( v + \psi \) is a supersolution of (2.1) (from [13, Lemma II.6.2]). We can conclude that \( u - \psi \) is a subsolution of (2.1) by considering \( \xi < 0 \) in (4.9) and arguing similarly as above. This proves the lemma.

\[\Box\]

Now that we have proved that a \( \psi \) with the desired properties exists, we are in a position to give the proof of Proposition 4.

### 4.1 Proof of Proposition 4

Let us assume the contrary to (3.2) i.e.,

\[
u(x_\ast, t_\ast) - v(x_\ast, t_\ast) - \delta > \sup_{x \in \mathbb{R}^n} \{\|u(x, 0) - v(x, 0)\| \vee 0\}, \tag{4.15}\]

\[\bigcirc\] Springer
for some $\delta > 0$. In the following lemma, we take $h : \mathbb{R}^{+} \to \mathbb{R}^{+}$ as $h(r) = Kr^\gamma$, where $\gamma$ is as in Lemma 3 and $K > 0$.

**Lemma 4** For some $r > 0$,

$$u(x, t) - v(y, t) - h(\|x - x_*\|) - h(\|y - x_*\|) \leq 0,$$

whenever $\|x - x_*\| < r$, $\|y - x_*\| < r$ and $t \in [0, T)$.

**Proof** From Lemma 3, we know that there exists $r > 0$ such that $h(\|x - x_*\|)$ is a supersolution of (2.1) and $r$ can be chosen uniformly with respect to $K$. From now on, we fix such an $r > 0$. Choose

$$K \equiv \sup_{0 \leq t < T} \left\{ \|u(\cdot, t)\|_\infty + \|v(\cdot, t)\|_\infty \right\}^{1/\gamma}.$$ 

Define

$$\Phi(x, y, t) \equiv u(x, t) - v(y, t) - h(\|x - x_*\|) - h(\|y - x_*\|) - \beta \left( \|x\|^2 + \|y\|^2 \right) - \frac{\rho}{T - t}$$

and

$$\Delta \equiv \{(x, y, t) \in \mathbb{R}^n \times \mathbb{R}^n \times [0, T) : t > 0, \|x - x_*\| < r \text{ and } \|y - x_*\| < r \}.$$

From (4.8), (4.15) it is clear that there is $(\hat{x}, \hat{y}, \hat{t}) \in \Delta$ such that

$$u(\hat{x}, \hat{t}) - v(\hat{y}, \hat{t}) - h(\|\hat{x} - x_*\|) - h(\|\hat{y} - x_*\|) > \eta > 0. \quad (4.16)$$

From the definition of $\Phi$, we know that there is a maximum $(\bar{x}_*, \bar{y}_*, \bar{t}_*)$ (depending on $\beta$ and $\rho$) of $\Phi$ on $\bar{\Delta}$. For small enough $\beta$ and $\rho$ (Lemma 6), we can ensure

$$\Phi(\bar{x}_*, \bar{y}_*, \bar{t}_*) > 0.$$ 

From boundedness of $u$ and $v$ and sub-linearity of $h$, it is clear that

$$\beta(\|\bar{x}_*\|^2 + \|\bar{y}_*\|^2) < \infty \text{ and } \lim_{\beta \to 0} \beta(\|\bar{x}_*\| + \|\bar{y}_*\|) \to 0.$$ 

It is clear that $0 < \bar{t}_* < T$, for small enough $\beta$ and $\rho$. To see that $\|\bar{x}_* - x_*\| < r$ and $\|\bar{y}_* - x_*\| < r$, for small enough $\beta$ and $\rho$, observe that from the choice of $K$, we have

$$h(\|x - x_*\|) + h(\|y - x_*\|) \geq \sup_{0 \leq t < T} \left\{ \|u(\cdot, t)\|_\infty + \|v(\cdot, t)\|_\infty \right\},$$

whenever $\|x - x_*\| = r$ or $\|y - x_*\| = r$. From this, it follows that

$$\Phi(\bar{x}_*, \bar{y}_*, \bar{t}_*) \leq -\beta(\|\bar{x}_*\|^2 + \|\bar{y}_*\|^2) - \frac{\rho}{T - \bar{t}_*}.$$
This is contradictory to the assumption as \((\bar{x}_*, \bar{y}_*, \bar{t}_*)\) is the maximum of \(\Phi\) and
\[
\Phi(\bar{x}_*, \bar{y}_*, \bar{t}_*) > 0.
\]

Therefore, for small enough \(\beta\) and \(\rho\), \((\bar{x}_*, \bar{y}_*, \bar{t}_*) \in \Delta\).

Now using [6, Theorem 8.3], we know that there exist \(a, b, X, Y\) such that
\[
(a, \beta \bar{x}_*, X) \in \overline{\mathcal{P}}^{2,+}(u^0(\bar{x}_*, \bar{t}_*) - h(\|\bar{x}_* - x_*\|)) \\
(b, \beta \bar{y}_*, -Y) \in \overline{\mathcal{P}}^{2,-}(v(\bar{y}_*, \bar{t}_*) + h(\|\bar{y}_* - x_*\|))
\]
such that \(a = b\) and for every \(\epsilon = \beta^{-1}\),
\[
-(\frac{1}{\epsilon} + \beta) \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq (\beta + \epsilon \beta^2) \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix},
\]
where \(\mathbb{I}\) is the identity matrix in \(\mathbb{R}^{n \times n}\). We again remind the reader that the hypothesis of the [6, Theorem 8.3] requires only upper semicontinuity of \(u^0\) and lower semicontinuity of \(v\) and hence [6, Theorem 8.3] can be applied. Using the fact that \(u - h(\|\cdot - x_*\|)\) and \(v + h(\|\cdot - x_*\|)\) are sub and supersolution, respectively, we get
\[
a + \gamma \left( u^0(\bar{x}_*, \bar{t}_*) - h(\|\bar{x}_* - x_*\|) \right) - \mathcal{L}(\bar{x}_*, 2\beta \bar{x}_*, X) \leq -\frac{\rho}{T^2} \\
b + \gamma \left( v(\bar{y}_*, \bar{t}_*) + h(\|\bar{y}_* - x_*\|) \right) - \mathcal{L}(\bar{y}_*, 2\beta \bar{y}_*, Y) \geq 0.
\]

From above, we have
\[
\gamma \left( u(\bar{x}_*, \bar{t}_*) - v(\bar{y}_*, \bar{t}_*) - h(\|\bar{x}_* - x_*\|) - h(\|\bar{y}_* - x_*\|) \right) - \beta \langle b(\bar{x}_*), \bar{x}_* \rangle \\
+ \beta \langle b(\bar{y}_*), \bar{y}_* \rangle - Tr \left[ \sigma(\bar{x}_*) \sigma(\bar{x}_*)^\dagger X - \sigma(\bar{y}_*) \sigma(\bar{y}_*)^\dagger Y \right] \leq -\frac{\rho}{T^2},
\]
Taking \(\beta\) and \(\rho\) to 0, we have a contradiction, from the boundedness of \(b\) and \(\sigma\). This concludes the proof. \(\square\)

From the lemma, we know that
\[
u(x, t) - v(y, t) - h(\|x - x_*\|) - h(\|y - x_*\|) \leq 0,
\]
whenever \(\|x - x_*\| < r, \|y - x_*\| < r\) and \(t \in [0, T]\). In particular,
\[
u(x_*, t_*) - v(y_*, t_*) \leq 0.
\]
This contradicts (4.15). Hence, we have
\[
sup_{\mathcal{O}^*} [u - v] = \sup_{x \in \mathbb{R}^n} \{ [u(x, 0) - v(x, 0)] \vee 0 \}.
\]
This concludes the proof of Lemma 2.
5 Proof of Theorem 4

Let \( u^\varepsilon, f \) (and \( u^0, f \)) be the continuous viscosity solutions of (2.4) with initial condition \( f \in C_b(\mathbb{R}^n) \) (and (2.1) with initial condition \( f \in C_b(\mathbb{R}^n) \)).

The only obstacle in proving the desired result for \( k > 2 \) is to show that

\[
\lim_{\varepsilon \to 0} \mathbb{E}_{X^\varepsilon} \left[ \prod_{i=1}^{k} f_i(X^\varepsilon_{t_i}) \right]
\]

exists for any \( \{f_i\} \in C_b(\mathbb{R}^n) \) (set of bounded continuous functions on \( \mathbb{R}^n \)) and \( 0 \leq t_1 < t_2 < t_3 < \ldots < t_k \leq T \). We will only show that the above limit exists and equals

\[
\mathbb{E}_{X^*} \left[ f_1(X^*_1) f_2(X^*_2) f_3(X^*_3) \right].
\]

for \( k = 3 \). Extension to \( k > 3 \) case follows along the similar lines.

Let \( P^\varepsilon(t, A, y) \) (with \( 0 \leq t < T \), \( y \in \mathbb{R}^n \) and \( A \in \mathcal{B}(\mathbb{R}^n) \)) be the corresponding transition kernel for \( X^\varepsilon \). Using Lemma 2 and 1 together, we know that for any \( f \in C_b(\mathbb{R}^n) \),

\[
u^\varepsilon, f (x, t) = \int_{\mathbb{R}^n} f(y) P^\varepsilon(t, dy, x) \to \tilde{u}(x, t)
\]

uniformly on compact sets of \( \mathcal{O}^* \),

where \( \tilde{u} \) is defined in Lemma 1.

For \( 0 \leq t < T \), [4, Theorem 3.1] gives us the existence of measure \( P^0(t, dy, x) \) such that

\[
u^0, f (x, t) = \int_{\mathbb{R}^n} f(y) P^0(t, dy, x).
\]

For \( t_1 > t_2 > t_3 \), consider

\[
\mathbb{E} \left[ f_1(X^\varepsilon_{t_1}) f_2(X^\varepsilon_{t_2}) f_3(X^\varepsilon_{t_3}) \right] = \int f_1(x_1) P^\varepsilon(t_2 - t_1, dx_1, x_2) f_2(x_2) P^\varepsilon(t_3 - t_2, dx_2, x_3) f_3(x_3) \times P^\varepsilon(t_1, dx_1, x) \mu(dx).
\]

Consider

\[
h^\varepsilon_1(x_2, t_2) \doteq \int_{x_1 \in \mathbb{R}^n} f_1(x_1) P^\varepsilon(t_2 - t_1, x_1, x_2) = u^\varepsilon, f_1(x_2, t_2).
\]

From Lemma 1, we know that

\[
h^\varepsilon_1 \to h_1^0 \doteq \int_{x_1 \in \mathbb{R}^n} f_1(x_1) P^0(\cdot - t_1, dx_1, \cdot), \text{ uniformly on compact sets of } \mathcal{O}^*.
\]
Now define

\[
h^\varepsilon_2(x_3, t_3) = \int_{x_2 \in \mathbb{R}^n} h^\varepsilon_1(x_2, t_2) f_2(x_2) P^\varepsilon(t_3 - t_2, dx_2, x_3) = u^{\varepsilon, (h^\varepsilon_1, f_2)}(x_3, t_3)
\]

We note that \((h^\varepsilon_1, f_2)\) converges uniformly on compact sets of \(\mathbb{R}^n\) to \(h^0_1 f_2\) and that for every \(\delta > 0\) and \(0 \leq t < T\), there is a compact set \(K\) (depending on \(t\)) such that \(P^\varepsilon(t, K, x) > 1 - \delta\). Noting that \(\|f_i(\cdot, t)\|_{\infty} \leq 1\) and \(\|h_i(\cdot, t)\|_{\infty} \leq 1\), we have

\[
\left| u^{\varepsilon, h^\varepsilon_1 f_2}(x_3, t_3) - u^{0, h^0_1 f_2}(x_3, t_3) \right| \leq |J_1 - J_2| + 2\delta
\]

\[
\leq |J_3 - J_4| + |J_5 - J_6| + 2\delta,
\]

where,

\[
J_1 = \int_{x_2 \in K} h^\varepsilon_1(x_2, t_2) f_2(x_2) P^\varepsilon(t_3 - t_2, dx_2, x_3),
\]

\[
J_2 = \int_{x_2 \in K} h^0_1(x_2, t_2) f_2(x_2) P^0(t_3 - t_2, dx_2, x_3),
\]

\[
J_3 = \int_{x_2 \in K} h^\varepsilon_1(x_2, t_2) f_2(x_2) P^\varepsilon(t_3 - t_2, dx_2, x_3),
\]

\[
J_4 = \int_{x_2 \in K} h^0_1(x_2, t_2) f_2(x_2) P^\varepsilon(t_3 - t_2, dx_2, x_3),
\]

\[
J_5 = \int_{x_2 \in K} h^0_1(x_2, t_2) f_2(x_2) P^\varepsilon(t_3 - t_2, dx_2, x_3),
\]

\[
J_6 = \int_{x_2 \in K} h^0_1(x_2, t_2) f_2(x_2) P^0(t_3 - t_2, dx_2, x_3).
\]

The first and second terms above go to zero as \(\varepsilon \to 0\) and as \(\delta\) is arbitrary, we have

\[
h^\varepsilon_2(x_3, t_3) \to h^0_2(x_3, t_3) = \int_{x_2 \in \mathbb{R}^n} h^0_1(x_2, t_2) f_2(x_2) P^0(t_3 - t_2, dx_2, x_3).
\]

Since \(t_1 > t_2 > t_3\) are fixed throughout and all the functions are evaluated only at one of these time instants, all convergence claims can be assumed to be uniform in them. So we only consider uniform convergence in \(x\), the first argument. From the above calculation, only pointwise convergence of \(h^\varepsilon_2(\cdot, t_3)\) can be inferred. We now show that

the convergence of \(h^\varepsilon_2\) to \(h^0_2\) is uniform on compact sets of \(\mathbb{R}^n\). Recall that

\[
u^{\varepsilon, f} \to u^{0, f}, \text{ uniformly on compact sets of } \mathbb{R}^n \times [0, T),
\]

\(\square\) Springer
for every $f \in C_b(\mathbb{R}^n)$. In other words, for every $\rho > 0$ and compact $\mathcal{J} \subset \mathbb{R}^n$, there is an $\varepsilon_0 > 0$ such that for $\varepsilon < \varepsilon_0$, we have

$$\left| \int_{\mathbb{R}^n} f(y) P^\varepsilon(t, dy, x) - \int_{\mathbb{R}^n} f(y) P^0(t, dy, x) \right| < \rho, \text{ for } \forall x \in \mathcal{J}.$$ 

The following set is relatively compact in $\mathcal{P}(\mathbb{R}^n)$:

$$\mathcal{P} = \{ P^\varepsilon(t, \cdot, x) : \varepsilon > 0, x \in \mathcal{J} \}.$$ 

Indeed, consider a sequence $(\varepsilon_n, x_n)$ such that $(\varepsilon_n > 0)$ and $\{x_n \in \mathcal{J}\}$. Then there exists a subsequence (still denoted by $n$) $(\varepsilon_n, x_n)$ such that $x_n \to x^* \in \mathcal{J}$ and there exists a further sub-subsequence (still denoted by $n$) such that $P^{\varepsilon_n}(t, \cdot, x^*)$ converges weakly to $P^0(t, \cdot, x^*)$. Along the subsequence $(\varepsilon_n, x_n)$ and for $g \in C_b(\mathbb{R}^n)$, consider

$$\left| \int_{\mathbb{R}^n} g(y) P^{\varepsilon_n}(t, dy, x_n) - \int_{\mathbb{R}^n} g(y) P^0(t, dy, x) \right| \leq \left| \int_{\mathbb{R}^n} g(y) P^{\varepsilon_n}(t, dy, x_n) - \int_{\mathbb{R}^n} g(y) P^0(t, dy, x_n) \right| + \left| \int_{\mathbb{R}^n} g(y) P^0(t, dy, x_n) - \int_{\mathbb{R}^n} g(y) P^0(t, dy, x^*) \right|.$$ 

As $n \to \infty$, second term on the right hand side goes to zero due to Feller continuity of $X^*$ and the first term goes to zero due to the uniform convergence on compact sets of $\mathbb{R}^n$. Hence $\mathcal{P}$ is relatively compact. Thus for any $\delta > 0$, there exists a compact set $\mathcal{K}_{\delta, \mathcal{J}}$ (only depending on $\delta$ and $\mathcal{J}$) such that

$$P^\varepsilon(t, \mathcal{K}_{\delta, \mathcal{J}}, x) > 1 - \delta, \text{ for every } \varepsilon > 0 \text{ and } x \in \mathcal{J}. \quad (5.1)$$

We are now in a position to prove the desired uniform convergence of $h^\varepsilon_2$ to $h^0_2$. We follow the method of [4, Pg. 737-738]. To that end, fix a $\delta > 0$ and a compact set $\mathcal{J} \subset \mathbb{R}^n$ and choose the corresponding compact set $\mathcal{K}_{\delta, \mathcal{J}} \subset \mathbb{R}^n$. For $x \in \mathcal{J}$,

$$|h^\varepsilon_2(x, t_3) - h^0_2(x, t_3)| \leq 2\delta + \left| \int_{\mathcal{K}_{\delta, \mathcal{J}}} h^\varepsilon_1(x_2, t_2) f_2(x_2) P^\varepsilon(t_3 - t_2, dx_2, x) \right. \left. - \int_{\mathcal{K}_{\delta, \mathcal{J}}} h^0_1(x_2, t_2) f_2(x_2) P^0(t_3 - t_2, dx_2, x) \right| \leq 4\delta,$$

for some small enough $\varepsilon$, chosen uniformly on $\mathcal{J}$. To get the final inequality, we used the fact that $h^\varepsilon_1(\cdot, t_3)$ converges to $h^0_1(\cdot, t_3)$ and also the fact that $P^\varepsilon(t_3 - t_2, \cdot, x)$ converges weakly to $P^0(t_3 - t_2, \cdot, x)$, uniformly in $x \in \mathcal{J}$ (this follows from (5.1)).
This finally shows that $h_2^\varepsilon \to h_2^0$ uniformly on compact sets. Similarly, defining

$$h_3^\varepsilon(x) \doteq \int_{x \in \mathbb{R}^n} h_2^\varepsilon(x_3, t_3) f_3(x_3) P^\varepsilon(t_3, d x_3, x)$$

and proceeding as above, we can conclude that

$$h_3^\varepsilon \to h_3^0 \doteq \int_{x \in \mathbb{R}^n} h_2^0(x_3, t_3) f_3(x_3) P^0(t_3, d x_3, x)$$

uniformly on compact sets of $\mathbb{R}^n$. We conclude that

$$\lim_{\varepsilon \to 0} \mathbb{E} X^\varepsilon \left[ f_1(X_{t_1}^\varepsilon) f_2(X_{t_2}^\varepsilon) f_3(X_{t_3}^\varepsilon) \right] = \mathbb{E} X^* \left[ f_1(X_{t_1}^*) f_2(X_{t_2}^*) f_3(X_{t_3}^*) \right]$$

$$= \mathbb{E} \bar{X} \left[ f_1(\bar{X}_{t_1}) f_2(\bar{X}_{t_2}) f_3(\bar{X}_{t_3}) \right],$$

where $\bar{X}$ is any other limit point of $X^\varepsilon$. Since $\bar{X}$ and $X^*$ are $C$-valued processes, we can conclude that they have the same distribution, as already observed.

This proves Theorem 4.

6 Discussion on Assumptions

From the analysis thus far, two things are clear - the set $L_{b, \sigma}$ and the Assumption 2 are the most crucial to this work. Also, Assumption 1 is a very weak assumption once we have Assumption 2. Therefore we discuss the latter in detail in what follows.

Before that, we explain why the set $L_{b, \sigma}$ is important. Techniques in existing literature [6, 13] to establish a comparison principle similar to Lemma 2 in the case degenerate $\sigma$ assume Lipschitz property of $b$, $\sigma$. A closer look at the proof of [6, Theorem 8.2] and trying to apply the same technique to prove Lemma 2 tells us that the Lipschitz property of either $b$ or $\sigma$ (in fact, local Lipschitz continuity of $b$ and $\sigma$ at $\hat{x}$ and $\hat{y}$) is invoked only in Equation (4.7). Recall that $(\hat{x}, \hat{y}, \hat{t})$ is a maximizer of $M_{\alpha, \beta, \rho} (\cdot, \cdot, \cdot)$ for large enough $\alpha$ and small enough $\beta$ and $\rho$. Therefore if we can ensure that $b$ and $\sigma$ are locally Lipschitz continuous at $\hat{x}$ and $\hat{y}$, then we can establish Lemma 2. Unfortunately, since we do not a priori know if this is possible, we have to consider two cases: 1.) $b$ and $\sigma$ are locally Lipschitz at $\hat{x}$ and $\hat{y}$ and 2.) it is otherwise. This motivates us to consider the set $L_{b, \sigma}$.

Assumption 2 covers a large class of Hölder continuous drifts and diffusion coefficients, going beyond what was achieved in [4]. It allows us to construct a class of viscosity supersolutions (see Definition (2)) with desired properties (Lemma 3) in the following manner:

- 1. of Assumption 2 is mainly used because derivatives of function $\|x - x_*\|^{\gamma}$, for some $0 < \gamma < 1$ and $x_* \in \mathbb{R}^n$ are again of a similar form viz., $\|x - x_*\|^{\gamma - 1}$ at $x \neq x_*$. This fact is used in comparing the terms in (4.14).
- 2. of Assumption 2 is used to arrive at the estimate (4.10).
In 3. of Assumption 2, $\beta > \frac{1}{2}$ is assumed to ensure that the convergence in (4.12) and (4.13) holds. Assuming $1 + \alpha - 2\beta > 0$ ensures that the third term in (4.14) dominates the first two terms, thereby making the sum of the first three terms of (4.14) non-negative in a neighbourhood of any $x_* \in NL_{b, \sigma}$.

We give a few examples of pairs $(b, \sigma)$ satisfying Assumption 2 where (1.1) can be shown to have non-unique weak solutions. The examples that we give are for $n = 1$, but one can easily construct examples for $n > 1$. However, showing non-uniqueness of weak solutions may be difficult.

1. $b(x) = |x|^\alpha$, $\sigma(x) = 0$ and $0 < \alpha < 1$. It can be easily seen there are multiple solutions for $\mu = \delta_0$, where, $\delta_x$ denotes the Dirac measure at $x \in \mathbb{R}^n$.

2. [17, Exercise 5.2.17] $b(x) = 3x^{\frac{1}{3}}$, $\sigma(x) = 3x^{\frac{2}{3}}$. For $\mu = \delta_0$, it is known that (1.1) has multiple strong solutions in this case. This pair of $(b, \sigma)$ satisfy both 1. and 2. of Assumption 2. However, $\alpha + 1 - 2\beta = 0$ in this case. This will not be an issue as Lemma 3 can still be shown to hold by considering the explicit forms of $b$, $\sigma$ and optimum values (which are easy to find in this case) of all the relevant constants in the proof.

Even though we have assumed appropriate Hölder continuity of $b$ and $\sigma$, we expect that just uniform continuity of $b$ and $\sigma$ might suffice for our results to hold. We however, could not provide a proof of this statement. Relaxing this condition further gives rise to cases where no Feller solution exists for the corresponding martingale problem. Here is one such counterexample.

**Counterexample 1** [23, Example 12.4.2] For $n = 1$, $\sigma \equiv 0$ and

$$b(x) = \begin{cases} 
(\text{sgn}(x)|x|^\frac{1}{2}, & \text{for } |x| \leq 1 \\
1, & \text{for } x \geq 1 \\
-1, & \text{for } x \leq -1
\end{cases}$$

(6.1)

It can be easily seen that (1.1) with the above coefficients does not have a Feller solution. In particular, the transition kernel fails to be continuous at $x = 0$. For more such examples, see [1] and [4, Pg. 746].

Since the case with $\sigma \equiv 0$ has received significant attention in the literature, we discuss it for a very special example: $n = 1$, $b(x) = \sqrt{|x|}$ and $\sigma \equiv 0$ i.e.,

$$\dot{X}_t = \sqrt{|X_t|}, \quad X_0 = x \in \mathbb{R}.$$  

(6.2)

It is clear that for $x \neq 0$, there is a unique local solution i.e., the solution $X_t$ satisfies (6.2) only in an open interval $(0, T_x)$, for some $T_x > 0$ depending on $x \neq 0$. For $x = 0$, no such claim can be made and in fact, it is easy to see that there are two solutions, only one of which is a Feller solution.

**Declarations**

**Conflict of interest** The authors declare that they have no conflict of interest.
Appendix

Lemma 5 If $u$ and $v$ are viscosity subsolutions of (2.1) and $v$ is $C^{1,2}((0, T) \times \mathbb{R}^n)$, then $u + v$ is also a subsolution.

Proof From the definition of viscosity subsolution, we know that when $\mathcal{P}_O^{2,+}u(x, t) \neq \phi$,

\[
a - \mathcal{L}(X, p, x) \leq 0, \text{ for } (t, x) \in (0, T) \times \mathbb{R}^n \text{ and } (a, p, X) \in \mathcal{P}_O^{2,+}u(x, t). \quad (A.1)
\]

\[
b - \mathcal{L}(Y, q, x) \leq 0, \text{ for } (t, x) \in (0, T) \times \mathbb{R}^n \text{ and } (b, q, Y) \in \mathcal{P}_O^{2,+}v(x, t). \quad (A.2)
\]

From the differentiability of $v$ and definition of $\mathcal{P}_O^{2,+}$, it is clear that $\mathcal{P}_O^{2,+}v(x, t) \neq \emptyset$, for every $(x, t) \in (0, T) \times \mathbb{R}^n$ and more importantly, is a singleton. Using Taylor’s theorem and the definition of $\mathcal{P}_O^{2,+}v(x, t)$ (on Page 5), we have

\[
v(y, s) = v(x, t) + b(s - t) + q.(y - x) + \frac{1}{2}(y - x)^\top Y(y - x) + o \left( |t - s| + \|x - y\|^2 \right)
\]

(A.3)

Note that the ‘$\leq$’ in the definition of $\mathcal{P}_O^{2,+}$ became ‘$=$’ because of the differentiability of $v$. Let $(c, r, Z) \in \mathcal{P}_O^{2,+}(u + v)(x, t)$. Then we have $(c - b, r - q, Z - Y) \in \mathcal{P}_O^{2,+}u(x, t)$. To see this, we use the definition of $\mathcal{P}_O^{2,+}(u + v)(x, t)$:

\[
u(y, s) = u(x, t) + v(x, t) + c(s - t) + r.(y - x) + \frac{1}{2}(y - x)^\top Z(y - x) + o \left( |t - s| + \|x - y\|^2 \right).
\]

Now subtracting Equation (A.3) from the above inequality gives us the following:

\[
u(y, s) \leq u(x, t) + (c - b)(s - t) + (r - q).(y - x) + \frac{1}{2}(y - x)^\top (Z - Y)(y - x) + o \left( |t - s| + \|x - y\|^2 \right)
\]

and this impies that $(c - b, r - q, Z - Y) \in \mathcal{P}_O^{2,+}u(x, t)$. Since $u$ is a subsolution, we have

\[
c - b - \mathcal{L}(Z - Y, r - q, x) \leq 0 \implies c - \mathcal{L}(Z, r, x) \leq b - \mathcal{L}(Y, q, x) \leq 0,
\]

from (2.6) and the linearity of $\mathcal{L}$. Finally, from the definition of viscosity subsolution, we have the result. \hfill \Box

In the rest of the appendix, we assume the conditions of the statement of Lemma 2. Let $M_{\alpha, \beta, \rho}$ be as defined in (4.2) and $(\hat{x}, \hat{y}, \hat{t})$ is the maximizer of $M_{\alpha, \beta, \rho}$ on $\mathbb{R}^n \times \mathbb{R}^n \times [0, T]$. \hfill \(\heartsuit\) Springer
Lemma 6 [25, Pg. 29] Suppose (4.1) holds. Then there exists $\beta_0$ and $\rho_0$ such that for every $\forall \alpha > 0$, $\beta < \beta_0$ and $\rho < \rho_0$, we have the following:

$$\sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0,T]} M_{\alpha, \beta, \rho} > \eta > 0.$$ 

Proof Let

$$\gamma \doteq \lim_{r \to 0} \sup_{(x, y, t) \in \mathbb{R}^n \times [0,T]} \left\{ u(x, t) - v(y, t) : \|x - y\| < r \right\}.$$

From (3.2), clearly we have,

$$\gamma \geq \sup_{\mathbb{R}^n \times [0,T]} [u(x, t) - v(x, t)] > \sup_{\mathbb{R}^n} [(u(x, 0) - v(x, 0)] \vee 0 \doteq M_b.$$

For $\epsilon > 0$, there is a $r_0$ such that for $r < r_0$, we have

$$\sup_{(x, y, t) \in \mathbb{R}^n \times [0,T]} [u(x, t) - v(y, t) : \|x - y\| < r] > \gamma - \epsilon.$$

To ensure that $\alpha \|x_0 - y_0\|^2$ is small, choose $r = \min\{\sqrt{\frac{\epsilon}{2\alpha}}, r_0\}$. Now there exists $(x_0, y_0, t_0) \in \mathbb{R}^n \times [0,T]$ such that

$$u(x_0, t_0) - v(y_0, t_0) + \epsilon > \gamma - \epsilon$$

and $\alpha \|x_0 - y_0\|^2 < \epsilon$.

It is also clear that there exists $\beta_0$ and $\rho_0$ such that $\frac{\rho}{T - t_0} < \epsilon$ and $\beta \|x_0\|^2 < \epsilon$, for every $\beta < \beta_0$ and $\rho < \rho_0$. To summarize, we have shown that

$$M_{\alpha, \beta, \rho} (x_0, y_0, t_0) = u(x_0, t_0) - v(y_0, t_0) - \alpha \|x_0 - y_0\|^2 - \frac{\rho}{T - t_0} - \beta \|x_0\|^2$$

(A.4)  

$$> \gamma - 5 \epsilon$$

(A.5)  

$$> M_b, \text{ for small enough } \epsilon.$$  

(A.6)

This concludes the proof. \square

Note that $\hat{x}$, $\hat{y}$ and $\hat{t}$ depend on $\alpha$, $\beta$ and $\rho$. We then have:

Lemma 7

$$\lim_{\alpha \to \infty} \lim_{\beta, \rho \to 0} \alpha \|\hat{x} - \hat{y}\|^2 = 0.$$  

Proof From Lemma 6, $\sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0,T]} M_{\alpha, \beta, \rho} > 0$ and this along with the boundedness of $u$, $v$ means that

$$\alpha \|\hat{x} - \hat{y}\|^2 + \beta \|\hat{x}\|^2 + \frac{\rho}{T - \hat{t}} \leq \|u\|_{\infty} + \|v\|_{\infty} \doteq M.$$  

$\Box$  

Springer
This immediately gives us
\[ \| \hat{x} - \hat{y} \| \leq \sqrt{\frac{M}{\alpha}}, \quad \beta \| \hat{x} \|^2 \leq M \] and \( \frac{\rho}{T - \hat{t}} \leq M \).

then
\[ \sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0,T]} M_{\alpha,\beta,\rho} - \sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0,T]} M_{\alpha,\beta,\rho} \leq M_{\alpha,\beta,\rho}(\hat{x}, \hat{y}, \hat{t}) - M_{\alpha,\beta,\rho}(\hat{x}, \hat{y}, \hat{t}) = \frac{\alpha}{2} \| \hat{x} - \hat{y} \|^2 - \frac{\beta}{2} \| \hat{x} \|^2 - \frac{\rho}{2(T - \hat{t})} \]

From Lemma 6, we know that \( M_{\alpha,\beta,\rho} \) is bounded away from zero from below uniformly for small enough \( \beta \) and \( \rho \) and all \( \alpha \). From the monotone convergence theorem,
\[ \lim_{\alpha \to \infty} \lim_{\beta,\rho \to 0} \left( \alpha \| \hat{x} - \hat{y} \|^2 + \beta \| \hat{x} \|^2 + \frac{\rho}{T - \hat{t}} \right) = 0 \]
and we are done. \( \square \)

Lemma 8 For \( u, v \) as in Lemma 2, there exists \( \alpha_0, \beta_0 \) and \( \rho_0 \) such that for every \( \forall \alpha > \alpha_0, \beta < \beta_0 \) and \( \rho < \rho_0 \), we have the following:
\[ \sup_{\mathbb{R}^n \times \mathbb{R}^n \times [0,T]} M_{\alpha,\beta,\rho} > \sup_{\mathbb{R}^n} \{ [u(x, 0) - v(x, 0)] \lor 0 \}. \]

Proof Suppose \( v \) is uniformly continuous. Assume the contrary to the statement of the lemma. Then there exists a sequence \( (\alpha_n, \beta_n, \rho_n) \to (\infty, 0, 0) \) such that the corresponding \( \hat{t} = 0 \) (as \( \hat{t} \neq T \) because of the term involving \( \rho \)). It is easy to see from definition of \( M_{\alpha,\beta,\rho} \), Lemmas 6 and 7 that
\[ M_b < M_{\alpha_n,\beta_n,\rho_n}(\hat{x}, \hat{y}, 0) = u(\hat{x}, 0) - v(\hat{y}, 0) \leq v(\hat{x}, 0) - v(\hat{y}, 0) + \sup_{\mathbb{R}^n} \{ [u(\hat{x}, 0) - v(\hat{x}, 0)] \lor 0 \} \]
Letting \( n \to \infty \) and using uniform continuity of \( v \), we have \( M_b < M_b \), a contradiction. \( \square \)

References
1. Bafico, R., Baldi, P.: Small random perturbations of Peano phenomena. Stochastics 6(3–4), 279–292 (1982)
2. Bardi, M., Dolcetta, I.C.: Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations, vol. 12. Springer, New York (1997)
3. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
4. Borkar, V., Kumar, K.S.: A new Markov selection procedure for degenerate diffusions. J. Theor. Probab. 23(3), 729–747 (2010)
5. Buckdahn, R., Ouknine, Y., Quincampoix, M.: On limiting values of stochastic differential equations with small noise intensity tending to zero. Bull. Sci. Math. 133(3), 229–237 (2009)
6. Crandall, M.G., Ishii, H., Lions, P.L.: User’s guide to viscosity solutions of second order partial differential equations. Bull. Am. Math. Soc. 27(1), 1–67 (1992)
7. Delarue, F., Flandoli, F.: The transition point in the zero noise limit for a 1D Peano example. Discrete Contin. Dyn. Syst. 34(10), 4071 (2014)
8. Delarue, F., Maurelli, M.: Zero noise limit for multidimensional SDEs driven by a pointy gradient. arXiv preprint arXiv:1909.08702 (2019)
9. Eckmann, J.P., Ruelle, D.: Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. 57, 617–656 (1985)
10. Filippov, A.F.: Differential Equations with Discontinuous Righthand Sides: Control Systems, vol. 18. Springer, New York (2013)
11. Flandoli, F.: Random Perturbation of PDEs and Fluid Dynamic Models: École d’été de Probabilités de Saint-Flour XL–2010, Vol. 2015. Springer, New York (2011)
12. Flandoli, F., Högele, M.: A solution selection problem with small symmetric stable perturbations. arXiv preprint arXiv:1407.3469 (2014)
13. Fleming, W.H., Soner, H.M.: Controlled Markov Processes and Viscosity Solutions, vol. 25. Springer, New York (2006)
14. Gradinaru, M., Herrmann, S., Roynette, B.: A singular large deviations phenomenon. Ann. l’Inst. Henri Poincare B 37(5), 555–580 (2001)
15. Herrmann, S.: Phénomène de Peano et grandes déviations. C. R. Acad. Sci. Ser. I 332(11), 1019–1024 (2001)
16. Ishii, H., Kobayasi, K.: On the uniqueness and existence of solutions of fully nonlinear parabolic PDEs under the Osgood type condition. Differ. Integral Equ. 7(3–4), 909–920 (1994)
17. Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus, vol. 113. Springer, Berlin (2012)
18. Kifer, J.I.: On small random perturbations of some smooth dynamical systems. Math. USSR Izvestiya 8(5), 1083 (1974)
19. Krykun, I.G., Makhno, S.Y.: The Peano phenomenon for Itô equations. J. Math. Sci. 192(4), 441–458 (2013)
20. Krylov, N.V.: The selection of a Markov process from a Markov system of processes. Izv. Acad. Nauk 37, 691–708 (1973). ((in Russian))
21. Lions, P.L.: Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations part 2: Viscosity solutions and uniqueness. Commun. Partial Differ. Equ. 8(11), 1229–1276 (1983)
22. Liptser, R.S., Shiryaev, A.N.: Statistics of Random Processes: General Theory, vol. 394. Springer, Moskow (1977)
23. Stroock, D.W., Varadhan, S.S.: Multidimensional Diffusion Processes, vol. 233. Springer, New York (1979)
24. Veretennikov, A.Y.: Approximation of ordinary differential equations by stochastic differential equations. Math. Notes 33(6), 476–477 (1983)
25. Zhan, Y.: Viscosity solutions of nonlinear degenerate parabolic equations and several applications. Ph.D. thesis, University of Toronto (2000)
26. Zhang, L.: Random perturbation of some multi-dimensional non-Lipschitz ordinary differential equations. arXiv preprint arXiv:1202.4131 (2012)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.