

CHAPTER 6

Stochastic Representations for Nonlinear Parabolic PDEs

H. Mete Soner*

Koç University, Istanbul, Turkey
E-mail: msoner@ku.edu.tr

Contents

1. Introduction 3
Notation 4
2. Linear case: Feynman–Kac representation 4
2.1. Linear monotone semigroups 5
2.2. Zeroth-order term and forcing 8
2.3. Diffusions and parabolic PDEs 9
2.4. Initial value problems 11
2.5. Discrete Markov processes and simply coupled equations 12
2.6. Jump Markov processes and integro-differential equations 14
2.7. Dirichlet boundary conditions 16
2.8. Neumann condition and the Skorokhod problem 17
2.9. Stationary problems 18
3. Representation via controlled processes 19
3.1. Viscosity solutions 22
3.2. Optimal control of diffusion processes 26
3.3. Smooth value function and verification 31
3.4. Optimal control of jump Markov processes 33
3.5. Other type of control problems 34
3.6. Stochastic target and geometric problems 35
4. Backward representations 38
4.1. Backward stochastic differential equations 38
4.2. Second-order backward stochastic equations 39
5. Monte Carlo methods 45
5.1. The linear case 45
5.2. The semilinear case 46
5.3. The fully nonlinear case 47
Acknowledgement 48
References 48

*Member of the Turkish Academy of Sciences.

Abstract

We discuss several different representations of nonlinear parabolic partial differential equations in terms of Markov processes. After a brief introduction of the linear case, different representations for nonlinear equations are discussed. One class of representations is in terms of stochastic control and differential games. An extension to geometric equations is also discussed. All of these representations are through the appropriate expected values of the data. Different type of representations are also available through backward stochastic differential equations. A recent extension to second-order backward stochastic differential equations allow us to represent all fully nonlinear scalar parabolic equations.

Keywords: Second-order backward stochastic differential equations, Fully nonlinear parabolic partial differential equations, Viscosity solutions, Superdiffusions, Feynman–Kac formula, BSDE, 2BSDE

MSC: 60H10, 35K55, 60H30, 60H35

1		1
2		2
3		3
4		4
5		5
6		6
7		7
8		8
9		9
10		10
11		11
12		12
13		13
14		14
15		15
16		16
17		17
18		18
19		19
20		20
21		21
22		22
23		23
24		24
25		25
26		26
27		27
28		28
29		29
30		30
31		31
32		32
33		33
34		34
35		35
36		36
37		37
38		38
39		39
40		40
41		41
42		42
43		43
44		44
45		45

1. Introduction

In this chapter we outline several connections between partial differential equations (PDE) and stochastic processes. These are extensions of Feynman–Kac-type representation of solutions to PDEs as the expected value of certain stochastic processes. Possible numerical implications of these connections are discussed as well. Although we restrict the scope of this paper to representation formulae, stochastic analysis provides much more analytical tools for PDEs. In particular, superdiffusions as developed by Dynkin [28–31] are related to nonlinear PDEs with a power-type nonlinearity, for reaction-diffusion equations interesting connections were used by Freidlin [38,39] to prove deep analytical results for these equations, also Barlow and Bass [7] study equations on fractals using random processes. Other important issues such as Martin boundaries, hypoellipticity and Malliavin calculus is not covered in these notes. Moreover, the theory of partial differential equation with stochastic forcing terms is not included. Interested readers may consult the papers [50,51] by Lions and Souganidis and [17] by Buckhadam and Ma and the references therein.

The starting point of most of our analysis is the celebrated Feynman–Kac formula [35, 43] which states that any solution of the linear heat equation

$$\frac{\partial u}{\partial t}(t, x) = \Delta u(t, x), \quad t > 0, x \in \mathbb{R}^d,$$

with initial condition $u(0, x) = f(x)$ with certain growth conditions (see Section 2) is given by

$$u(t, x) := E[f(x + \sqrt{2}W(t))],$$

where $W(\cdot)$ is the standard d -dimensional Brownian motion. This well-understood connection can be explained in several different ways. We will employ the semigroups to motivate this connection and to generalize it to more general stochastic processes. In that section we will briefly state Feynman–Kac-type formulae for several class of linear equations. The most general class of equations we will consider are second-order parabolic type integro-differential equations. Boundary value problems of Dirichlet and Neumann type are also discussed.

In Section 3 we extend these results to nonlinear equations of same type by using controlled stochastic processes. Since these equations do not always admit classical (or smooth) solutions, we will employ the theory of viscosity solutions to prove the representation formulae rigorously. The chief tool in this analysis is the dynamic programming principle which was first observed by Bellman [8]. The infinitesimal version of the dynamic programming principle is in fact gives the related partial differential equation. In this context, semigroup motivation plays an important role, as the theory of viscosity solutions is best explained through semigroups and the dynamic programming principle is in fact the semigroup property. We refer to the books by Bensoussan and Lions [9,10], Krylov [46] and Fleming and Soner [36] for more references and the historical development of the theory.

In that section we also provide a more recent representation formulae for geometric type equations. This is achieved by using a nonclassical control problem called target problems [67,68].

In this chapter we restrict our attention to only optimal control. However, these methods extend naturally to stochastic differential games. For this extension, we refer to the paper by Fleming and Souganidis [37] and Chapter 11 in the second edition of [36].

Another type of connection between PDEs and stochastic processes is given by backward stochastic differential equations (BSDEs in short). These formulae is analogous to method of characteristics for first-order equations. Indeed, initially BSDEs were studied by Bismut [11,12] then by Peng [59] as an extension of Pontryagin maximum principle which itself is an extension of characteristics. We provide a brief introduction to BSDEs and then outline a recent result of Cheredito, Soner, Touzi and Victoir [21]. This result extends the representation to formulae to all fully nonlinear, parabolic, second-order partial differential equations.

Last section is devoted to possible numerical implications of these formulae.

Notation

Let $d \geq 1$ be a natural number. We denote by $\mathcal{M}_{d,k}$ the set of all $d \times k$ matrices with real components, $\mathcal{M}^d = \mathcal{M}_{d,d}$. B' is the transpose of a matrix $B \in \mathcal{M}^d$ and $\text{Tr}[B]$ its trace. By $\mathcal{M}_{\text{inv}}^d$ we denote the set of all invertible matrices in \mathcal{M}^d , by \mathcal{S}^d all symmetric matrices in \mathcal{M}^d , and by \mathcal{S}_+^d all positive semidefinite matrices in \mathcal{M}^d . For $B, C \in \mathcal{M}^d$, we write $B \geq C$ if $B - C \in \mathcal{S}_+^d$. For $x \in \mathbb{R}^d$, we set

$$|x| := \sqrt{x_1^2 + \cdots + x_d^2}$$

and for $B \in \mathcal{M}^d$,

$$|B| := \sup_{x \in \mathbb{R}^d, |x| \leq 1} Bx.$$

Equalities and inequalities between random variables are always understood in the almost sure sense. $W(\cdot)$ is a multidimensional Brownian motion on a complete probability space (Ω, \mathcal{F}, P) . For $t \geq 0$, we denote by $(\mathcal{F}_t)_{t \geq 0}$ a filtration satisfying the usual conditions and containing the filtration generated by $\{W(s)\}_{s \in [0, T]}$.

2. Linear case: Feynman–Kac representation

Let $W(\cdot)$ be the standard d -dimensional Brownian motion, a be positive constant, and f be a scalar-valued, continuous function f on \mathbb{R}^d satisfying the growth condition

$$|f(x)| \leq C[1 + |x|^\alpha] \quad \forall x \in \mathbb{R}^d,$$

for some constants $C, \alpha \geq 0$. Then, the Feynman–Kac formula [35,43] states that

$$u(t, x) := E[f(x + a\sqrt{2}W(t))], \quad (2.1)$$

is the unique solution of the heat equation

$$\frac{\partial u}{\partial t}(t, x) := u_t(t, x) = a \Delta u(t, x), \quad t > 0, x \in \mathbb{R}^d,$$

together with the initial condition

$$u(0, x) = f(x), \quad x \in \mathbb{R}^d.$$

Indeed, once we know either if the function defined by the expected value is $C^{1,2}$, or if the heat equation has a smooth solution, then the above representation is an direct application of the Itô formula, Theorem 3.3 in [44]. The above growth condition is sufficient for either one of these conditions; see for instance Section 4.4, Remark 4.4 in [44]. Moreover, in this special case, polynomial growth can be weakened.

The above formula generalizes to a large class of Markov processes, linear equations and boundary problems. In this section we briefly and formally describe all these generalizations.

2.1. Linear monotone semigroups

The connection between the Markov processes and certain linear equations is now well understood and can be explained in many ways. In this chapter we will utilize semigroups to motivate this connection. The semigroup approach has the advantage that it generalizes to the nonlinear setting and it is well adapted to the theory of viscosity solutions. However, we use this approach only to motivate the connection and therefore our discussion of semigroups is only formal. In particular, we will not be precise about the domains of the operators.

In the initial discussion, we assume that the equations are defined on a metric space \mathbb{D} which is equal to either \mathbb{R}^d or to $\mathbb{R}^d \times \{1, 2, \dots, N\}$. Problems on bounded subsets of \mathbb{R}^d are, of course, common and similar representation results are available for these equations as well. But, in this subsection, we restrict our analysis to problems defined on all of \mathbb{R}^d or $\mathbb{R}^d \times \{1, 2, \dots, N\}$. Boundary problems will be discussed in the Sections 2.7 and 2.8.

Also, to simplify the presentation, we will consider PDEs that are backward in time. For these equations a terminal data at a given time T , instead of an initial data, is given. Then a solution is constructed for all times prior to T . Of course, there is a direct connection between terminal value problems and initial value problems through a simple time reversal. We perform this change for diffusion processes in the Section 2.4.

For all $t \geq 0$, let $\widehat{\mathcal{L}}_t$ be a linear operator on a subset $\widehat{\mathcal{D}}$ of $C_b(\mathbb{D})$ – bounded, scalar-valued, continuous functions on \mathbb{R}^d . Let $\varphi \in C_b(\mathbb{D})$ be a given function. For $T \geq 0$, consider the linear equation

$$-u_t(t, z) = (\widehat{\mathcal{L}}_t u(t, \cdot))(z), \quad \forall (t, z) \in (-\infty, T) \times \mathbb{D}, \quad (2.2)$$

1 together with final data

$$2 \quad u(T, z) = \varphi(z) \quad \forall z \in \mathbb{D}. \quad (2.3) \quad 3$$

4
5 Assume that for every T and $\varphi \in C_b(\mathbb{D})$ this equation has a unique smooth solution and let
6 u be this unique solution. Clearly this solution depends on T and φ , but this dependence is
7 always suppressed. Now, define a two-parameter family of operators

$$8 \quad (\widehat{\mathcal{T}}_{t,T}\varphi)(z) := u(t, z), \quad \forall t \leq T, z \in \mathbb{D}. \quad 9$$

10
11 By uniqueness, this family is a linear semigroup, i.e.,

$$12 \quad \widehat{\mathcal{T}}_{t,T}\varphi = \widehat{\mathcal{T}}_{t,r}(\widehat{\mathcal{T}}_{r,T}\varphi) \quad \forall t \leq r \leq T. \quad (2.4) \quad 13$$

14
15 Moreover, it is clear that the infinitesimal generator of this semigroup is the operator $\widehat{\mathcal{L}}_t$,

$$16 \quad \lim_{h \downarrow 0} \frac{\widehat{\mathcal{T}}_{t,t+h}\varphi - \varphi}{h} = \widehat{\mathcal{L}}_t\varphi \quad 17$$

18
19 for every $\varphi \in \widehat{\mathcal{D}}$.

20
21 We continue by constructing a similar semigroup using Markov processes. For this pur-
22 pose, let (Ω, P, \mathcal{F}) be probability space and $\{\mathcal{F}_r\}_{r \geq 0}$ be a filtration. Let $\{Z^{t,z}(r)\}_{r \geq t}$ be
23 a \mathbb{D} -valued Markov process on this probabilistic structure starting from $Z^{t,z}(t) = z$. For a
24 continuous bounded function φ , define a two parameter semigroup by

$$25 \quad \mathcal{T}_{t,T}\varphi(z) := E[\varphi(Z^{t,z}(T))] \quad \forall t \leq T. \quad 26$$

27
28 Formally, the Markov property of Z implies that $\mathcal{T}_{t,T}$ satisfies (2.4). Indeed, for $t \leq r \leq T$,

$$29 \quad \mathcal{T}_{t,r}(\mathcal{T}_{r,T}\varphi)(z) = E[(\mathcal{T}_{r,T}\varphi)(Z^{t,z}(r))] = E(E[\varphi(Z^{r,Z^{t,z}(r)}(T))]). \quad 30$$

31
32 Since, by the Markov property of the process,

$$33 \quad Z^{r,Z^{t,z}(r)}(T) = Z^{t,z}(T), \quad (2.5) \quad 34$$

35
36 we have

$$37 \quad \mathcal{T}_{t,r}(\mathcal{T}_{r,T}\varphi)(z) = E(E[\varphi(Z^{t,z}(T))]) = \mathcal{T}_{t,T}\varphi(z). \quad 38$$

39
40 Hence, $\mathcal{T}_{t,T}$ is a two-parameter semigroup. Following the theory of semigroups, the infi-
41 nitesimal generator

$$42 \quad \mathcal{L}_t\varphi := \lim_{t' \rightarrow t, h \downarrow 0} \frac{\mathcal{T}_{t,t+h}\varphi - \varphi}{h} \quad (2.6) \quad 43$$

exists for $\varphi \in \mathcal{D} \subset C_b(\mathbb{D})$, for some subset \mathcal{D} (see Section 5.1 in [44]).

Now, it is clear that to have a Feynman–Kac representation for the equations (2.2) and (2.3), we need construct a Markov process whose infinitesimal generator \mathcal{L}_t agrees with the linear operator $\widehat{\mathcal{L}}_t$ that appears in (2.2). To see this connection formally, suppose $\mathcal{L}_t = \widehat{\mathcal{L}}_t$. Fix T and φ , and set

$$u(t, z) := \mathcal{T}_{t,T}\varphi(z) = E[\varphi(Z^{t,z}(T))]. \quad (2.7)$$

We formally claim that u solves (2.2). Indeed, assume that u is smooth in the sense that it has a continuous time derivative and $u(t, \cdot) \in \mathcal{D}$ for all t . Then, by the semigroup property,

$$\begin{aligned} -u_t(t, z) &= \lim_{h \downarrow 0} \frac{u(t-h, z) - u(t, z)}{h} \\ &= \lim_{h \downarrow 0} \frac{\mathcal{T}_{t-h,T}\varphi(x) - u(t, z)}{h} \\ &= \lim_{h \downarrow 0} \frac{\mathcal{T}_{t-h,t}(u(t, \cdot))(z) - u(t, z)}{h} \\ &= \mathcal{L}_t(u(t, \cdot))(z). \end{aligned}$$

Hence, u defined by (2.7) solves (2.2). The terminal value (2.3) follows from the definition of u . Hence, if equation (2.2) together with (2.3) has a unique “smooth” solution, then it must be given by (2.7). However, in most cases we can prove this representation directly and obtain uniqueness as a by product of the representation.

In the example of a Brownian motion, it is well known that the infinitesimal generator is the Laplacian and therefore we have the representation (2.1) for the heat equation. In the same spirit, we can prove representation results for a large class of linear equations including equations with nonlocal terms. However, the semigroup generated by Markov processes are monotone and this puts a certain restriction on the operators \mathcal{L}_t that are the infinitesimal generators of the semigroups constructed by Markov processes. Indeed, the monotonicity of the stochastic semigroup (2.7) is a direct consequence of the definition and stated as

$$\varphi \leq \psi \implies \mathcal{T}_{t,r}\varphi \leq \mathcal{T}_{t,r}\psi. \quad (2.8)$$

Suppose that $\varphi, \psi \in \mathcal{D}$ and there exists $z_0 \in \mathbb{D}$ such that

$$0 = (\varphi - \psi)(z_0) = \max_{\mathbb{D}}(\varphi - \psi). \quad (2.9)$$

Then $\varphi(z_0) = \psi(z_0)$, $\varphi \leq \psi$ and by monotonicity, $\mathcal{T}_{t,r}\varphi \leq \mathcal{T}_{t,r}\psi$ for every $t \leq r$. By (2.6),

$$\mathcal{L}_t\varphi(z_0) = \lim_{h \downarrow 0} \frac{\mathcal{T}_{t,t+h}\varphi(z_0) - \varphi(z_0)}{h} \leq \lim_{h \downarrow 0} \frac{\mathcal{T}_{t,t+h}\psi(z_0) - \psi(z_0)}{h} = \mathcal{L}_t\psi(z_0).$$

Hence we proved that for every t , any infinitesimal generator \mathcal{L}_t of a Markov process satisfies the *maximum principle*: For any functions φ, ψ in the domain of \mathcal{L}_t , and $z_0 \in \mathbb{D}$ satisfying (2.9), we have

$$\mathcal{L}_t \varphi(z_0) \leq \mathcal{L}_t \psi(z_0). \quad (2.10)$$

This is essentially the only important restriction for equation (2.2) to have a Feynman–Kac-type representation. Also, it is important to note that the maximum principle is the crucial property for the development of viscosity solutions as well. Moreover, the maximum principle can be directly extended to nonlinear operators and this extension will be discussed in the preceding section.

To see the importance of the maximum principle, let us consider the example of a partial differential operator. So suppose that $\mathbb{D} = \mathbb{R}^d$ and \mathcal{L}_t be given by

$$\mathcal{L}_t \varphi(x) = H(t, x, \varphi(x), D\varphi(x), \dots, D^k \varphi(x))$$

for some given function H . Then, by calculus, we see that \mathcal{L}_t has maximum principle if and only if $k = 2$ and

$$H(t, x, u, p, B + B') \leq H(t, x, u, p, B) \quad \forall B' \geq 0. \quad (2.11)$$

(Here and in the rest of the chapter, for symmetric matrices inequalities are understood in the sense of quadratic forms.) This property means that the corresponding equation is a second-order (possibly degenerate) parabolic equation. These equations are related to diffusion processes that will be discussed in Section 2.3. There are nonlocal operators that have maximum principle and some examples will be discussed in Sections 2.5 and 2.6.

Also, the infinitesimal generators of Markov processes, again by definition, are translation invariant. Indeed, for any φ in the domain of \mathcal{L}_t and a constant β , $T_{t,T}(\varphi + \beta) = (T_{t,T}\varphi) + \beta$. Hence

$$\mathcal{L}_t(\varphi + \beta) = \lim_{h \downarrow 0} \frac{T_{t,t+h}(\varphi + \beta) - (\varphi + \beta)}{h} = \lim_{h \downarrow 0} \frac{T_{t,t+h}\varphi - \varphi}{h} = \mathcal{L}_t\varphi,$$

and therefore, the infinitesimal operators of Markov processes do not contain any zeroth-order terms. However, with a minor modification in the definition of the semigroup, a zeroth-order term and a forcing function can be included in the theory. This is the subject of the next subsection.

2.2. Zeroth-order term and forcing

Let $\{Z^{t,z}(s)\}_{s \geq t}$ be as in the previous subsection. To include a term $r(t, z)u(t, z) + h(t, z)$ to equation (2.2), we modify the Markov semigroup as follows. Define random variables,

$$B(t, T; z) := \exp\left(-\int_t^T r(s, Z^{t,z}(s)) ds\right)$$

1 and

$$2 \quad H(t, T; z) := \int_t^T B(t, s; z) h(s, Z^{t,z}(s)) ds. \quad 3$$

4 For a continuous bounded function φ , define a two parameter semigroup by

$$5 \quad \mathcal{T}_{t,T}\varphi(z) := E[B(t, T; z)\varphi(Z^{t,z}(T)) + H(t, T; z)] \quad \forall t \leq T. \quad 6$$

7 To prove the semigroup property, for $t \leq r \leq T$, observe that

$$8 \quad \begin{aligned} 9 \quad \mathcal{T}_{t,r}(\mathcal{T}_{r,T}\varphi)(z) &= E[B(t, r; z)(\mathcal{T}_{r,T}\varphi)(Z^{t,z}(r)) + H(t, r; z)] \\ 10 &= E(B(t, r; z)E[B(r, T; Z^{t,z}(r))\varphi(Z^{r,Z^{t,z}(r)}(T)) \\ 11 &\quad + H(r, T; Z^{t,z}(r))] + H(t, r; z)). \end{aligned} \quad 12$$

13 By the Markov property of $Z(\cdot)$ (or equivalently (2.5)),

$$14 \quad \begin{aligned} 15 \quad B(t, T; z) &= B(t, r; z)B(r, T; Z^{t,z}(r)), \\ 16 \quad H(t, T; z) &= B(t, r; z)H(r, T; Z^{t,z}(r)) + H(t, r; z). \end{aligned} \quad 17$$

18 Hence, we have

$$19 \quad \mathcal{T}_{t,r}(\mathcal{T}_{r,T}\varphi)(z) = E(E[B(t, T; z)\varphi(Z^{t,z}(T)) + H(t, r; z)]) = \mathcal{T}_{t,T}\varphi(z). \quad 20$$

21 Now the infinitesimal generator of this semigroup is given by

$$22 \quad \begin{aligned} 23 \quad \tilde{\mathcal{L}}_t\varphi(z) &= \lim_{h \downarrow 0} \frac{\mathcal{T}_{t+h,t}\varphi(z) - \varphi(z)}{h} \\ 24 &= \lim_{h \downarrow 0} \frac{1}{h} E[B(t, t+h; z)\varphi(Z^{t,z}(t+h)) - \varphi(Z^{t,z}(t+h))] \\ 25 &\quad + \lim_{h \downarrow 0} \frac{E[H(t, t+h; z)]}{h} + \lim_{h \downarrow 0} \frac{1}{h} (E[\varphi(Z^{t,z}(t+h))] - \varphi(z)) \\ 26 &= r(t, z)\varphi(z) + h(t, z) + \mathcal{L}_t\varphi(z). \end{aligned} \quad 27$$

28 2.3. Diffusions and parabolic PDEs

29 In this section, $\mathbb{D} = \mathbb{R}^d$ and $Z = X$ is diffusion process satisfying the stochastic differential

30 equation (SDE),

$$31 \quad dX(s) = \mu(s, X(s)) ds + \sigma(s, X(s)) dW(s), \quad \forall s > t, \quad 32 \quad (2.12) \quad 33$$

1 with initial condition

$$2 \quad X(t) = x. \quad (2.13) \quad 3$$

4 We assume that

$$5 \quad \mu : \mathbb{R}^+ \times \mathbb{R}^d \mapsto \mathbb{R}^d, \quad \sigma : \mathbb{R}^+ \times \mathbb{R}^d \mapsto \mathcal{M}^{d,k} \quad 6$$

7 are given functions satisfying usual conditions (cf. [44]) and $W(\cdot)$ is a \mathbb{R}^k -valued Brownian
8 motion. Then the infinitesimal generator of this process is

$$9 \quad \mathcal{L}_t = \mu(t, x) \cdot \nabla + \frac{1}{2} a(t, x) : D^2, \quad (2.14) \quad 10$$

11 where for two $d \times d$ symmetric matrices A and B ,

$$12 \quad A : B := \text{trace}[AB] = \sum_{i,j=1}^d A_{i,j} B_{i,j}, \quad 13$$

$$14 \quad a_{i,j}(t, x) = \sum_{l=1}^k \sigma_{i,l}(t, x) \sigma_{j,l}(t, x), \quad 15$$

16 and ∇ , D^2 are respectively the gradient and the Hessian with respect to the spatial vari-
17 able x . Hence, the related partial differential equation is

$$18 \quad -u_t(t, x) = \mathcal{L}_t u(t, x) \quad 19$$

$$20 \quad = \mu(t, x) \cdot \nabla u(t, x) + \frac{1}{2} a(t, x) : D^2 u(t, x) \quad 21$$

$$22 \quad = \sum_{i=1}^d \mu_i(t, x) u_{x_i}(t, x) \quad 23$$

$$24 \quad + \frac{1}{2} \sum_{i,j=1}^d a_{i,j}(t, x) u_{x_i x_j}(t, x) \quad \text{on } (-\infty, T) \times \mathbb{R}^d. \quad (2.15) \quad 25$$

26 The connection between the diffusion processes and the above equation can be proved di-
27 rectly by using the Itô calculus as well. Indeed, suppose that (2.15) together with the final
28 data (2.3) has a smooth solution u . Fix $t < T$ and $x \in \mathbb{R}^d$ and let $\{X(s) = X^{t,x}(s)\}_{s \geq t}$ be
29 the solution of the stochastic differential equation (2.12), (2.13). By Itô formula (Theo-
30 rem 3.3.6 in [44]),

$$31 \quad d(u(s, X(s))) = [u_s(s, X(s)) + (\mathcal{L}_s u(s, \cdot))(X(s))] ds \quad 32$$

$$33 \quad + \nabla u(s, X(s)) \cdot \sigma(s, X(s)) dW(s). \quad 34$$

By (2.15), the ds term in the above equation is zero. Hence, the process $Y(s) := u(s, X(s))$ is a local martingale. Under suitable growth conditions on u or on ∇u , we can show that

$$Y(t) = E[Y(T)] \implies u(t, x) = E[u(T, X^{t,x}(T))] = E[\varphi(X^{t,x}(T))].$$

Note that the above proof of representation using the Itô calculus has the advantage that it also proves uniqueness under some growth conditions.

A linear term and a forcing function to equation (2.15) can be added by the technique developed in Section 2.2.

2.4. Initial value problems

In this subsection, we will briefly discuss how we may translate the above results to initial value problems. Consider an initial value problem

$$v_t(t, x) = (\tilde{\mathcal{L}}_t(v(t, \cdot)))(x) \quad \text{on } (0, \infty) \times \mathbb{R}^d,$$

together with

$$v(0, x) = \varphi(x),$$

where

$$\tilde{\mathcal{L}}_t = \tilde{\mu}(t, x) \cdot \nabla + \frac{1}{2} \tilde{a}(t, x) : D^2.$$

Fix $(t, x) \in (0, \infty) \times \mathbb{R}^d$ and set $\tilde{X}(s)$ be the solution of

$$d\tilde{X}(s) = \tilde{\mu}(t-s, \tilde{X}(s)) ds + \tilde{\sigma}(t-s, \tilde{X}(s)) dW(s),$$

with initial data $\tilde{X}(0) = x$. Apply the Itô rule to the process $Y(s) := v(t-s, \tilde{X}(s))$. The result is

$$dY(s) = [-v_t(t-s, \tilde{X}(s)) + (\tilde{\mathcal{L}}_s(v(t-s, \cdot)))(\tilde{X}(s))] ds + (\dots) dW(s).$$

Again ds term is zero by the equation and the stochastic term is a local martingale. Under suitable growth assumptions,

$$Y(0) = E[Y(t)] \implies v(t, x) = E[v(0, \tilde{X}(t))] = E[\varphi(\tilde{X}(t))].$$

This result can also be *directly* derived from the results of Section 2.3 by a time reversal. Indeed, for a given T and a solution v of the above initial value problem, set

$$u(t, x) := v(T-t, x) \quad \forall t \leq T, x \in \mathbb{R}^d,$$

so that u solves (2.15) with

$$\mu(t, x) = \tilde{\mu}(T - t, x), \quad a(t, x) = \tilde{a}(T - t, x).$$

Let $X(s) = X^{T-t, x}(s)$ be the solution of (2.12) with initial data $X^{T-t, x}(T - t) = x$. Then, $X(s) = \tilde{X}(s - T + t)$. In particular, $X(T) = \tilde{X}(t)$, and in view of the representation proved in Section 2.3,

$$v(t, x) = u(T - t, x) = E[\varphi(X(T))] = E[\varphi(\tilde{X}(t))].$$

2.5. Discrete Markov processes and simply coupled equations

In this subsection, we will first consider Markov processes on a discrete set $\{1, 2, \dots, N\}$ and then couple these processes with diffusion processes of Section 2.3.

Let $v(\cdot)$ be a Markov process on a discrete set $\Sigma := \{1, 2, \dots, N\}$. For $(i, j) \in \Sigma$ and $r \geq t \geq 0$, let

$$P_{i,j}(t, r) := P(v(r) = j | v(t) = i),$$

be the transition probabilities. Assume that the rate functions

$$p_{i,j}(t) := \frac{\partial}{\partial r} P_{i,j}(t, r)$$

exist. Since

$$\sum_j P_{i,j}(t, r) = 1, \quad P_{i,j}(t, r) \geq 0 = P_{i,j}(t, t) \quad \forall i \neq j, r \geq t,$$

we conclude that

$$p_{i,j}(t) \geq 0 \quad \text{for } i \neq j \quad \text{and} \quad p_{i,i}(t) = - \sum_{j \neq i} p_{i,j}(t).$$

Moreover, the infinitesimal generator is given by

$$\mathcal{L}_t \varphi(i) = \sum_{j=1}^N p_{i,j}(t) \varphi(j) = \sum_{j \neq i} p_{i,j}(t) [\varphi(j) - \varphi(i)].$$

Therefore, the simple difference equation

$$-u_t(t, i) = \sum_{j \neq i} p_{i,j}(t) [u(t, j) - u(t, i)] \quad \forall t \leq T, i = 1, \dots, N, \quad (2.16)$$

with terminal data

$$u(T, \cdot) = \varphi(\cdot) \quad \text{on } \Sigma,$$

has the representation

$$u(t, i) = E[\varphi(v(t)) | v(t) = i].$$

Also a linear term and a forcing term can be added to this equation by the exponential discounting technique developed in Section 2.2. Hence, the representation result covers all equations of the form

$$\begin{aligned} -u_t(t, i) &= \sum_{j=1}^N p_{i,j}(t) [u(t, j) - u(t, i)] \\ &\quad + r_i(t)u(t, i) + h_i(t) \quad \forall t \leq T, i = 1, \dots, N, \end{aligned} \quad (2.17)$$

for given functions $p_{i,j}(t) \geq 0$ for $i \neq j$ and general functions $r_i(t), h_i(t)$ without sign restrictions. Indeed, any solution of (2.17) with terminal data (2.16), has the stochastic representation,

$$u(t, i) = E[H(t, T; i) + B(t, T; i)\varphi(v(T)) | v(t) = i],$$

where as in Section 2.2,

$$\begin{aligned} B(t, T; i) &:= \exp\left(\int_t^T r_{v(s)}(s) ds\right), \\ H(t, T; i) &:= \int_t^T B(t, s; i)h_{v(s)}(s) ds. \end{aligned}$$

We will now combine the above representation with the results of Section 2.3 to obtain a representation for a simply couple system of parabolic equations as well. Indeed, let $Z := (X, v) \in \mathbb{R}^d \times \{1, 2, \dots, N\}$ be a Markov process constructed as follows. Let $v \in \{1, 2, \dots, N\}$ be a discrete Markov process as above, and for each $i \in \{1, 2, \dots, N\}$, $X_i^{t,x}$ be a independent diffusion processes solving the SDE

$$dX_i^{t,x}(s) = \mu_i(s, X_i^{t,x}(s)) ds + \sigma_i(s, X_i^{t,x}(s)) dW(s),$$

with initial condition (2.13). Given an initial condition $t, z := (x, i)$, we start the process

$$Z^{t,z}(t) = (X(t), v(t)) = z.$$

Then, there are strictly increasing stopping times $t < \tau_1 < \tau_2 < \dots$, so that

$$\begin{aligned} v(s) &= i, \quad s \in [t, \tau_1), \\ v(s) &= i_1, \quad s \in [\tau_1, \tau_2), \quad \dots, \quad v(s) = i_N, \quad s \in [\tau_N, \tau_{N+1}). \end{aligned}$$

Given these stopping times, we define a continuous X process recursively by

$$\begin{aligned} X(s) &= X_i^{t,x}(s), & s \in [t, \tau_1], \\ X(s) &= X_{i_N}^{\tau_N, X(\tau_N)}(s), & s \in (\tau_N, \tau_{N+1}], N = 1, 2, \dots \end{aligned}$$

It is clear that $Z = Z^{t,x,i}(\cdot)$ is a Markov process with an infinitesimal generator

$$\begin{aligned} \mathcal{L}_t \varphi(x, i) &= \mu_i(t, x) \cdot \nabla \varphi(x, i) \\ &\quad + \frac{1}{2} a_i(t, x) : D^2 \varphi(x, i) + \sum_{j \neq i} p_{i,j} [\varphi(x, j) - \varphi(x, i)]. \end{aligned}$$

Therefore

$$u(t, x, i) = E[\varphi(X(T), v(T)) | Z(t) = (X(t), v(t)) = (x, i)],$$

is a solution of the coupled parabolic equation

$$-u_t(t, x, i) = (\mathcal{L}u(t, \cdot, \cdot))(x, i) \quad \forall t < T, (x, i) \in \mathbb{R}^d \times \{1, 2, \dots, N\},$$

with final data

$$u(T, x, i) = \varphi(x, i).$$

Notice that one may see the above equation as a system of coupled parabolic equations with a solution

$$v(t, \cdot) := (u(t, \cdot, 1), \dots, u(t, \cdot, N)) : \mathbb{R}^d \mapsto \mathbb{R}^N$$

for all t . However, this coupling is only through the zeroth-order terms and the coupling constants $p_{i,j}$'s are all nonnegative. For that reason, we would like to view the above equation as a scalar valued function

$$u(t, \cdot, \cdot) : \mathbb{R}^d \times \{1, 2, \dots, N\} \mapsto \mathbb{R}^1.$$

These two different point of views have been effectively used by Freidlin in his pioneering work [38] on the analysis of some reaction–diffusion equations by stochastic methods.

2.6. Jump Markov processes and integro-differential equations

In this subsection, we will consider Markov processes which solve a stochastic differential equation which is more general than the one considered in the Section 2.3. This is done by adding a stochastic integral to the standard diffusion equation (2.12). This stochastic

1 integral is generally independent of the Brownian motion and it is driven by a random
 2 martingale measure. In particular, this measure contains jump terms and as such they gen-
 3 eralize all the processes considered in the previous subsections.

4 Precisely, let π a positive Borel measure on \mathbb{R}^d , called the *compensator*, satisfying

$$5 \int_{\mathbb{R}^d} [1 \wedge |\xi|^2] \pi(d\xi) < \infty. \quad 6$$

7
 8
 9 Given this compensator measure π , there exists a random counting measure p on the Borel
 10 subsets of $\mathbb{R}^+ \times \mathbb{R}^d$ so that, for any Borel set $A \subset \mathbb{R}^+ \times \mathbb{R}^d$, $p(A)$ has a Poisson distribution
 11 with mean

$$12 \lambda(A) := \int_A \pi(d\xi) dt. \quad 13$$

14
 15
 16 Moreover, $\tilde{p} := p - \lambda$ is a martingale measure. We refer to a manuscript of Skorokhod [62]
 17 for a construction of such measures, or to a recent book by Oksendall and Sulem [55].

18 In the manuscript of Skorokhod [62] and in the paper of Fujiwara and Kunita [40] exist-
 19 ence and uniqueness of stochastic differential equations are also proved. Indeed, let μ, σ
 20 be as in Section 2.3 and let

$$21 f : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}^d \quad 22$$

23
 24 be a function satisfying

$$25 \int_{\mathbb{R}^d} [1 \wedge |f(t, x, \xi)|^2] \pi(d\xi) < \infty. \quad 26$$

27
 28
 29 We assume the standard Lipschitz condition

$$30 \begin{aligned} & |\mu(t, x) - \mu(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \\ & + \int_{\mathbb{R}^d} |f(t, x, \xi) - f(t, y, \xi)|^2 \pi(d\xi) \leq C|x - y|^2 \end{aligned} \quad 31$$

32
 33
 34 for all $t \in \mathbb{R}^1$, $x, y \in \mathbb{R}^d$ for some constant C . Then there exists a unique solution to

$$35 \begin{aligned} X(r) = x + \int_t^r \mu(s, X(s)) ds + \int_t^r \sigma(s, X(s)) dW(s) \\ + \int_t^r \int_{\mathbb{R}^d} f(s, X(s), \xi) \tilde{p}(ds \times d\xi) \quad \forall r \geq t, \end{aligned} \quad 36$$

(2.18) 37

38
 39
 40
 41
 42
 43
 44 for any initial condition (t, x) , a random measure \tilde{p} constructed as above and an independ-
 45 ent standard Brownian motion $W(\cdot)$.

Moreover, the process X is a Markov process with an infinitesimal generator,

$$\mathcal{L}_t \varphi(x) = \mu(t, x) \cdot \nabla \varphi(x) + \frac{1}{2} a(t, x) : D^2 \varphi(x) + \tilde{\mathcal{L}}_t \varphi(x),$$

where the part corresponding to the random measure is given by

$$\tilde{\mathcal{L}}_t \varphi(x) = \int_{\mathbb{R}^d} [\varphi(x + f(t, x, \xi)) - \varphi(x) - f(t, x, \xi) \cdot \nabla \varphi(x)] \pi(d\xi). \quad (2.19)$$

See for instance, p. 94 of [62] or [55] for a proof. Hence, we have the stochastic representation discussed earlier for the integro-differential equation,

$$-u_t(t, x) - \mu(t, x) \cdot \nabla u(t, x) - \frac{1}{2} a(t, x) : D^2 u(t, x) - (\tilde{\mathcal{L}}_t u(t, \cdot))(x) = 0.$$

A linear term and forcing can be added as before and also with further coupling with a discrete Markov process would yield a system of integro-differential equations.

2.7. Dirichlet boundary conditions

In the previous subsections, for the ease of exposition, we restricted our discussion to problems on all of \mathbb{R}^d . However, with a simple absorption rule at the boundary of a given region, we can include all Dirichlet problems into this theory. In this subsection, we outline the main tools that can be used for almost all processes. Indeed, let O be an open set with smooth boundary and \mathcal{L}_t be as in (2.14) and consider the boundary value problem

$$-u_t - (\mathcal{L}_t u(t, \cdot))(x) = 0 \quad \forall t < T, x \in O,$$

together with the terminal data

$$u(T, x) = \varphi(x) \quad \forall x \in O, \quad (2.20)$$

and the boundary condition,

$$u(t, x) = g(t, x) \quad \forall t < T, x \in \partial O, \quad (2.21)$$

for some given function g . Usually, we require a compatibility condition, $g(T, x) = \varphi(x)$ for all x on the boundary of O . We can now view the solution $u(t, \cdot)$ as the value of the semigroup $\mathcal{T}_{t, T}$ applied to the terminal data φ . We include the boundary conditions in the definition of the domain of this semigroup, which can be taken as

$$C_g(t, \cdot) := \{v: C(\bar{O}) \mid v(x) = g(t, x) \forall x \in \partial O\}.$$

The stochastic semigroup is defined as follows: given φ , g , T as above, and an initial condition $x \in \mathbb{R}^d$, $t < T$, let $X^{t,x}(\cdot)$ be the solution of the SDE (2.12), (2.13). Let θ be the exit time from the domain $O \times [t, T]$,

$$\theta := \inf\{s \geq t: X^{t,x}(s) \notin O\} \wedge T.$$

Set

$$\mathcal{I}_{t,T}\varphi(x) := E[\varphi(X^{t,x}(T))\chi_{\{\theta=T\}} + g(\theta, X^{t,x}(\theta))\chi_{\{\theta < T\}}].$$

Under suitable growth and regularity conditions, we can show that the two semigroups are equal to each other. Thus, we have a stochastic representation for the boundary value problem. Integro-differential equations can be dealt with similarly. However, boundary data on all of $\mathbb{R}^d \setminus O$ is needed instead of data only on ∂O .

2.8. Neumann condition and the Skorokhod problem

Neumann-type boundary conditions are included into the theory with some care. Consider the same parabolic equation with boundary condition

$$-u_\nu(t, x) := -v(t, x) \cdot \nabla u(t, x) = g(t, x) \quad \forall t < T, x \in \partial O, \quad (2.22)$$

for some given function g and a given unit vector field $v(t, x)$. We require that

$$v(t, x) \cdot n(x) > 0, \quad (2.23)$$

where $n(x)$ is unit *inward normal* to the boundary ∂O at $x \in \partial O$. To obtain a representation we use the local time on ∂O . Indeed, we modify the SDE (2.12) in the following way. Given an initial condition $x \in O$ and $t < T$, we look for continuous processes $X^{t,x}(\cdot)$ and $l(\cdot)$ satisfying (2.13) and for $s \in [t, T]$,

$$\begin{aligned} X^{t,x}(s) &= x + \int_t^s \mu(r, X^{t,x}(r)) dr + \int_t^s \sigma(r, X^{t,x}(r)) dW(r) \\ &\quad + \int_t^r v(r, X^{t,x}(r)) dl(r) \in \bar{O}, \end{aligned} \quad (2.24)$$

$$l(s) = \int_t^s \chi_{\{X^{t,x}(r) \in \partial O\}} dl(r),$$

$$l(0) = 0 \quad \text{and} \quad l \text{ is nondecreasing and continuous.} \quad (2.25)$$

In the literature the solution $X^{t,x}(\cdot)$ called the *reflected* diffusion process, l is the *local time* and the above set of equations are called the *Skorokhod problem*. Under the usual Lipschitz conditions on μ , σ , v , and smoothness assumption on the boundary ∂O , the Skorokhod problem has a unique solution. This and more was proved by Lions and Sznitman [52].

1 We also refer to Lions [48] for the connection to partial differential equations and viscosity 1
2 solutions. 2

3 The only difference between (2.24) and (2.12) is the last dl integral and the important 3
4 requirement that $X^{t,x}(s) \in \bar{O}$ for all s . Notice that (2.25) ensures that dl increases only 4
5 when the diffusion processes $X^{t,x}$ is on the boundary ∂O . Hence, formally $X^{t,x}$ processes 5
6 is “reflected” on the boundary ∂O in the direction $\nu(t, x)$. In view of the condition (2.23) 6
7 and the fact that n is the inward normal, the reflection direction $\nu(t, x)$ points inward from 7
8 $x \in \partial O$. These guarantee that the process $X^{t,x}$ takes values in \bar{O} . 8

9 We now define Markov stochastic semigroup, 9

$$10 \quad u(t, x) = E \left[\varphi(X^{t,x}(T)) + \int_t^T g(s, X^{t,x}(s)) dl(s) \right]. \quad 10$$

11 We claim that any smooth solution $v \in C^{1,2}((0, T) \times \bar{O})$ of (2.15), (2.3) and (2.22) is 11
12 equal to u . Indeed, let $X^{t,x}, l$ be a solution of the Skorokhod problem with initial data 12
13 $X^{t,x}(t) = x$. Since l is a monotone function, X is a semimartingale and with the use of 13
14 Itô’s rule we obtain 14

$$15 \quad \begin{aligned} 15 \quad \varphi(X^{t,x}(T)) &= v(T, X^{t,x}(T)) \\ 16 &= v(t, x) + M(T) + \int_t^T [v_t + \mathcal{L}v](s, X^{t,x}(s)) ds \\ 17 &\quad + \int_t^T \nabla v(s, X^{t,x}(s)) \cdot \nu(s, X^{t,x}(s)) dl(s), \end{aligned} \quad 17$$

18 where M is a local martingale with $M(t) = 0$. By (2.15), the first integrand is zero, and 18
19 by (2.22) the second integrand is equal to $-g(s, X^{t,x}(s))$. Also under some suitable growth 19
20 conditions $E[M(T)] = 0$. We use these observations and then take the expected value. The 20
21 result is $v = u$. 21

22 Once again, a linear term and a forcing function can be added into the theory as in the 22
23 Section 2.2. 23

24 2.9. Stationary problems 24

25 Time homogeneous linear problems also have similar stochastic representations. Indeed, 25
26 let $r(x) \geq \beta > 0$ be a given function. Let $Z^z(\cdot)$ be a time homogeneous Markov process 26
27 with infinitesimal generator \mathcal{L} and initial condition $Z^z(0) = z$. Given an open set O with 27
28 smooth boundary, consider the boundary value problem 28

$$29 \quad r(z)u(z) - \mathcal{L}u(z) = h(z) \quad \forall z \in O, \quad 29$$

30 together with the boundary condition, 30

$$31 \quad u(z) = g(z) \quad \forall z \in \mathbb{R}^d \setminus O, \quad (2.26) \quad 31$$

1 for some given function g .

2 To obtain a stochastic representation, let θ be the exit time from the domain O ,

$$3 \quad \theta^z := \inf\{s \geq 0: Z^z(s) \notin O\}.$$

6 Set

$$7 \quad u(z) := E \left[\int_0^{\theta^z} B(s)h(Z^z(s)) ds + B(\theta^z)g(Z^z(\theta^z))\chi_{\{\theta^z < \infty\}} \right],$$

10 where

$$12 \quad B(t) := \exp\left(-\int_0^t r(s, Z^z(s)) ds\right).$$

15 Notice that due our strict positivity assumption on r , $B(s) \geq e^{-\beta s}$ for all $s \geq 0$. Therefore, the integral term in the above expression is integrable under reasonable growth assumptions on h .

18 For $T > 0$, define a stochastic semigroup by

$$20 \quad \mathcal{T}_T \varphi(z) := E \left[\int_0^{T \wedge \theta^z} B(s)L(Z^z(s)) ds \right. \\ 22 \quad \left. + B(\theta^z)\varphi(Z^z(\theta^z))\chi_{\{\theta^z < T\}} + B(T)g(Z^z(T))\chi_{\{\theta^z \geq T\}} \right].$$

26 Time homogeneity of the Markov process implies the semigroup property,

$$28 \quad \mathcal{T}_{T+S}\varphi = \mathcal{T}_T(\mathcal{T}_S\varphi) \quad \forall T, S \geq 0.$$

30 Also, u is a fixed point of this semigroup for every T ; that is, $u = \mathcal{T}_T u$ for all $T \geq 0$. Then, under suitable growth and regularity conditions, we can show that u is the unique solution of the linear equation. Thus, we have a stochastic representation for the stationary boundary value problem as well. For partial differential equation, the boundary condition is needed only on ∂O . Neumann boundary conditions are handled as in the previous subsection.

38 3. Representation via controlled processes

39 In this section we will consider nonlinear equations of the form

$$40 \quad -u_t(t, z) + \mathcal{H}(t, z, u(t, \cdot)) = 0 \quad \forall (t, x) \in (-\infty, T) \times \mathbb{D}, \quad (3.1)$$

42 where \mathbb{D} is as before. The most general form of the nonlinearity \mathcal{H} is of the form

$$44 \quad \mathcal{H}(t, z, \varphi(\cdot)) := \inf_{\beta \in B} \sup_{\alpha \in A} \{-\mathcal{L}_t^{\alpha, \beta} \varphi(z) - r(t, z, \alpha, \beta)\varphi(z) - L(t, z, \alpha, \beta)\}, \quad (3.2)$$

where for a set of parameters α, β in some control sets A and B , $\mathcal{L}_t^{\alpha, \beta}$ is the infinitesimal generator of Markov process on \mathbb{D} , as in Section 2.1, and r and L are given functions. These equations are related to stochastic differential games. An interesting and an important class of equations related to stochastic optimal control is obtained by taking B to be a singleton. In this case, the nonlinearity reduces to

$$\mathcal{H}(t, z, \varphi(\cdot)) := \sup_{\alpha \in A} \{-\mathcal{L}_t^\alpha \varphi(z) - r(t, z, \alpha)\varphi(z) - L(t, z, \alpha)\}. \quad (3.3)$$

In this chapter we only discuss operators of the above form. For differential games, we refer the interested reader to Chapter 11 in the *second edition* of [36].

Finally, note that the nonlinearity in (3.2) has the maximum principle as defined in Section 2.1 (cf. (2.10)). Recall that as a consequence of maximum principle all local operators \mathcal{H} on $\mathbb{D} = \mathbb{R}^d$ of the above form must be given by

$$\mathcal{H}(t, x, \varphi(\cdot)) = H(t, x, \varphi(x), \nabla \varphi(x), D^2 \varphi(x))$$

for some given function H satisfying (2.11).

Consider the nonlinear equation (3.1). If this equation together with the terminal data (2.3) has a unique solution, then we define a nonlinear semigroup acting on the terminal data φ by

$$\widehat{T}_{t, T}(\varphi)(z) := u(t, z).$$

By uniqueness, this is a semigroup. To obtain the related stochastic semigroup, we consider \mathcal{H} as in (3.3).

Let \mathcal{A} be the set of all bounded, progressively measurable processes $\alpha(t) \in A$. Again as in the linear case, given an initial condition t, z and a processes $\alpha(\cdot) \in \mathcal{A}$, consider a class of processes $Z^{t, z, \alpha(\cdot)}$. We assume that, for every fixed $\alpha \in A$, the infinitesimal generator of the processes $Z^{t, z, \bar{\alpha}}$ is equal to \mathcal{L}_t^α , where $\bar{\alpha}$ is process which is equal to the constant α everywhere. Define the *value function* v by

$$v(t, z) := \inf_{\alpha \in \mathcal{A}} (\mathcal{J}_{t, T}^{\alpha(\cdot)} \varphi)(z), \quad (3.4)$$

where with $Z = Z^{t, z, \alpha(\cdot)}$,

$$(\mathcal{J}_{t, T}^{\alpha(\cdot)} \varphi)(z) := E \left[\int_t^T B(s) L(s, Z(s), \alpha(s)) ds + B(T) \varphi(Z(T)) \right],$$

$$B(T) = B(t, T; z, \alpha(\cdot)) = \exp \left(- \int_t^T r(s, Z(s), \alpha(s)) ds \right).$$

Bellman's dynamic programming in this context states that, for any stopping time

$$\theta \in [t, T],$$

$$\begin{aligned} v(t, z) &= \inf_{\alpha(\cdot)} \left\{ E \left[\int_t^\theta B(s) L(s, Z(s), \alpha(s)) ds + B(\theta) v(\theta, Z(\theta)) \right] \right\} \\ &= \inf_{\alpha(\cdot)} \{ (\mathcal{J}_{t, \theta}^{\alpha(\cdot)} v(\theta, \cdot))(z) \}. \end{aligned}$$

We refer to [66] for a general abstract proof of dynamic programming under some structural assumptions. The crucial structure needed to prove the above result is the additive structure given by the Markov assumption on the process Z and the fact that

$$B(t, T; z, \alpha(\cdot)) = B(t, \theta; z, \alpha(\cdot)) B(\theta, T; Z^{t, z, \alpha(\cdot)}(\theta), \alpha(\cdot)).$$

Now we define the stochastic semigroup by

$$(\mathcal{T}_{t, T} \varphi) = v(t, \cdot) = \inf_{\alpha \in \mathcal{A}} (\mathcal{J}_{t, T}^{\alpha(\cdot)} \varphi),$$

so that the dynamic programming principle implies that, for any stopping time $\theta \in [t, T]$,

$$(\mathcal{T}_{t, T} \varphi) = (\mathcal{T}_{t, \theta} u(\theta, \cdot)) = (\mathcal{T}_{t, \theta} [\mathcal{T}_{\theta, T} \varphi]).$$

Hence, dynamic programming implies that $\mathcal{T}_{t, T}$ is a semigroup. Indeed, the dynamic programming principle and the semigroup properties are essentially equivalent.

It now remains to show that the infinitesimal generator of this stochastic semigroup agrees with \mathcal{H} given in (3.3). We show this connection only formally here. We then introduce the theory of viscosity solutions of Crandall and Lions to prove it rigorously. Indeed, we need to compute the following limit,

$$\begin{aligned} &\lim_{\theta \downarrow t} \frac{(\mathcal{T}_{t, \theta} \varphi)(z) - \varphi(z)}{\theta - t} \\ &= \lim_{\theta \downarrow t} \inf_{\alpha(\cdot) \in \mathcal{A}} \left[\frac{E \int_t^\theta B(s) L(s, Z(s), \alpha(s)) ds}{\theta - t} + \frac{E(B(\theta) \varphi(Z(\theta)) - \varphi(z))}{\theta - t} \right]. \end{aligned}$$

Now, formally, assume that we may interchange the order of limit and infimum. Also, again formally, assume that the infimum can be taken only over all $\bar{\alpha}$, where $\bar{\alpha}(s) = \alpha$ for all s . Then, formally,

$$\lim_{\theta \downarrow t} \frac{(\mathcal{T}_{t, \theta} \varphi)(z) - \varphi(z)}{\theta - t} = \inf_{\alpha \in \mathcal{A}} \lim_{\theta \downarrow t} [\mathcal{J}_1(t, \theta, \alpha)(z) + (\mathcal{J}_2(t, \theta, \alpha) \varphi)(z)],$$

where

$$\mathcal{J}_1(t, \theta, \alpha)(z) = \frac{E \int_t^\theta B(s) L(s, Z(s), \alpha) ds}{\theta - t}$$

1 and

$$2 \quad (\mathcal{J}_2(t, \theta, \alpha)\varphi)(z) = \frac{E[B(\theta)\varphi(Z(\theta)) - \varphi(z)]}{\theta - t}.$$

3
4
5 Since the infinitesimal generator of the controlled process Z with control process $\bar{\alpha}$ is assumed to be \mathcal{L}_t^α , the limit of \mathcal{J}_2 is equal to $\mathcal{L}_t^\alpha\varphi(z) + r(t, z, \alpha)\varphi(z)$. Also, by the continuity of the processes, the limit of \mathcal{J}_1 is equal to $L(t, z, \alpha)$. Hence, formally, we compute the infinitesimal generator is

$$6 \quad \lim_{\theta \downarrow t} \frac{(\mathcal{T}_{t, \theta}\varphi)(z) - \varphi(z)}{\theta - t} = \inf_{\alpha \in A} \{ \mathcal{L}_t^\alpha\varphi(z) + r(t, z, \alpha)\varphi(z) + L(t, z, \alpha) \}$$

$$7 \quad = -\mathcal{H}(t, z, \varphi(\cdot)).$$

8 Hence, we have shown that the generators of the semigroup $\widehat{\mathcal{T}}_{t, T}$ related to the PDE and the stochastic semigroup $\mathcal{T}_{t, T}$ have the same infinitesimal generator. Therefore, as in the linear case, if the PDE has unique solution in a certain class containing v then, the value function v is the unique solution of the PDE.

9 The main focus of the preceding subsections, is to make the above calculations rigorous and to extend these results to general nonlinearities.

10 3.1. Viscosity solutions

11 This subsection follows very closely [36].

12 Let \mathbb{D} be closed subset of a Banach space and \mathcal{C} be a collection of functions on \mathbb{D} which is closed under addition, i.e.,

$$13 \quad \phi, \psi \in \mathcal{C} \Rightarrow \phi + \psi \in \mathcal{C}.$$

14 As in the previous sections, the main object of our analysis is a two parameter family of operators $\{\mathcal{T}_{tr} : t \leq r \leq T\}$ with the common domain \mathcal{C} . In the applications the exact choice of \mathcal{C} is not important. However, when \mathbb{D} is compact, we will require that \mathcal{C} contains $C(\mathbb{D})$. For noncompact Σ , additional conditions are often imposed. Indeed, in most of our examples, we will require that \mathcal{C} contains $\mathcal{M}(\mathbb{D}) \cap C_p(\mathbb{D})$ ($\mathcal{M}(\mathbb{D})$ is set of all real-valued functions which are bounded from below, $C_p(\mathbb{D})$ is set of all continuous, real-valued functions which are polynomially growing). We assume that \mathcal{T}_{tt} is the identity.

15 Next we want to state the semigroup property. When, $\mathcal{T}_{rT}\varphi$ belongs to \mathcal{C} , the semigroup property is (2.4). However, $\mathcal{T}_{rT}\varphi$ may not be in the domain. So, in general we assume that, for all $\phi, \psi \in \mathcal{C}$ and $t \leq r \leq s \leq T$,

$$16 \quad \mathcal{T}_{tr}\phi \leq \mathcal{T}_{ts}\psi \quad \text{if } \phi \leq \mathcal{T}_{rs}\psi, \tag{3.5}$$

$$17 \quad \mathcal{T}_{tr}\phi \geq \mathcal{T}_{ts}\psi \quad \text{if } \phi \geq \mathcal{T}_{rs}\psi. \tag{3.6}$$

1 By taking $r = s$ in (3.2) we conclude that the above conditions imply \mathcal{T}_{Tr} is monotone, (2.8). 1
 2 Moreover, if $\mathcal{T}_{r,s}\psi \in \mathcal{C}$, by taking $\phi = \mathcal{T}_{r,s}\psi$, we obtain (2.4). So, in general, (3.5), (3.6) is 2
 3 a convenient way of stating monotonicity and the semigroup properties. 3

4 All the linear semigroups introduced in the previous section satisfy the above condi- 4
 5 tions. We will now give the example of a semigroup generated by stochastic optimal. This 5
 6 example will then be studied in detail in the next subsection. 6

7
 8 **EXAMPLE 3.1** (Controlled diffusion processes). We follow a construction as in (3.4). Let 8
 9 A be a control set, O be an open subset of \mathbb{R}^d . Set $\mathbb{D} = \bar{O}$ and $\mathcal{C} = \mathcal{M}(\Sigma)$, set of all mea- 9
 10 surable functions bounded from below. Let μ, σ, L, g be functions satisfying the standard 10
 11 Lipschitz conditions (see [36]), i.e., for any function 11

$$12 \quad \phi : (-\infty, T] \times \mathbb{D} \times A \mapsto \mathcal{M}, \quad 12$$

13 where \mathcal{M} is any normed space (in our applications \mathcal{M} is either \mathbb{R}^d or the set of real 13
 14 matrices with usual norm), we say that ϕ satisfies the standard Lipschitz condition if ϕ is 14
 15 continuous in the (t, x) variables and 15
 16 16

$$17 \quad \|\phi(t, x, \alpha) - \phi(t, x', \alpha)\|_{\mathcal{M}} \leq C|x - x'| \quad \forall t \in (-\infty, T], x, x' \in \bar{O}, \alpha \in A, \quad 17$$

18 (3.7) 18

19 with a constant independent of all variables. Let (Ω, P, \mathcal{F}) be a probability space, $W(\cdot)$ 19
 20 be a standard \mathbb{R}^k Brownian motion and $\{\mathcal{F}_t\}$ be the filtration satisfying the usual conditions 20
 21 as in [44]. 21

22 Let \mathcal{A} be all bounded, progressively measurable, A -valued random processes. We all 22
 23 \mathcal{A} the set of *admissible controls*. In some applications further restrictions on the controls 23
 24 are needed. These can be modeled easily by introducing (t, x) depended subsets of \mathcal{A} . 24
 25 However, in that case certain conditions must be satisfied as discussed in [36,66]. 25

26 Given a process $\alpha(\cdot) \in \mathcal{A}$ and an initial condition (2.13), we consider the controlled 26
 27 stochastic differential equation 27
 28 28

$$29 \quad dX(s) = \mu(s, X(s), \alpha(s)) ds + \sigma(s, X(s), \alpha(s)) dW(s). \quad 29$$

30 (3.8) 30

31 For a given boundary function g , a running cost function L and a function ψ , set 31
 32 32

$$33 \quad (\mathcal{J}_{t,T}^{\alpha(\cdot)} \varphi)(x) = E \left[\int_t^{\theta \wedge r} B(s)L(s, X(s), \alpha(s)) ds \right. \quad 33$$

$$34 \quad \left. + B(\theta)g(\theta, X(\theta))\chi_{\theta < r} + B(r)\psi(X(r))\chi_{\theta \geq r} \right], \quad 34$$

35 (3.9) 35

36 where 36
 37 37

$$38 \quad B(T) = B(t, T; z, \alpha(\cdot)) = \exp\left(-\int_t^T r(s, Z(s), \alpha(s)) ds\right), \quad 38$$

39 39

40 where 40
 41 41

$$42 \quad B(T) = B(t, T; z, \alpha(\cdot)) = \exp\left(-\int_t^T r(s, Z(s), \alpha(s)) ds\right), \quad 42$$

43 43

44 where 44
 45 45

$$45 \quad B(T) = B(t, T; z, \alpha(\cdot)) = \exp\left(-\int_t^T r(s, Z(s), \alpha(s)) ds\right), \quad 45$$

θ is the exit time of $(s, X(s))$ from $\bar{Q} = [t, T] \times \bar{O}$. The nonlinear semigroup is given by

$$(\mathcal{T}_{t,r}\psi)(x) := \inf_{\alpha(\cdot) \in \mathcal{A}} \mathcal{J}_{t,r}^{\alpha(\cdot)} \psi(x).$$

Assume that L, g and ψ are all bounded from below. Then, $\mathcal{T}_{t,r}\psi$ is also bounded from below and therefore, for every $\psi \in \mathcal{C}$, $\mathcal{T}_{t,r}\psi$ is well defined and belongs to \mathcal{C} . Clearly $\mathcal{T}_{t,r}$ is monotone (2.8). Also, dynamic programming for optimal control (cf. [36,66]) implies the semigroup property (2.4).

Notice that the infinitesimal generator of the controlled process is

$$\mathcal{L}_t^\alpha = \mu(t, x, \alpha) \cdot \nabla + \frac{1}{2} a(t, x, \alpha) : D^2, \quad (3.10)$$

where “:” is as before and

$$a_{i,j}(t, x, \alpha) = \sum_{l=1}^k \sigma_{i,l}(t, x, \alpha) \sigma_{j,l}(t, x, \alpha).$$

Hence, in view of the formal argument given in the introduction of this section, the related partial differential equation is (3.1) with \mathcal{H} as in (3.3) with the above \mathcal{L}_t^α .

In the next section, we will rigorously prove the connection between the dynamic programming equation and the above semigroup.

For $\psi \in \mathcal{C}$, $t \leq T$, $x \in \mathbb{D}$, set

$$v(t, x) = (\mathcal{T}_{t,T}\psi)(x). \quad (3.11)$$

In analogy with control problems, we call $v(t, x)$ the *value function*. Using the semigroup property, we conclude that the value function satisfies

$$v(t, x) = (\mathcal{T}_{t,r}v(r, \cdot))(x) \quad \forall x \in \mathbb{D}, t \leq r \leq T, \quad (3.12)$$

provided that $v(r, \cdot) \in \mathcal{C}$. This identity is just a restatement of the dynamic programming principle when the semigroup is related to an optimal control problem. Hence, we refer to (3.12) as the *(abstract) dynamic programming principle*.

Having formulated the dynamic programming principle abstractly, we proceed to derive the corresponding dynamic programming equation. Let $r = t + h$ in (3.12) for some $h > 0$ and small. Assume that $v(t + h, \cdot) \in \mathcal{C}$. Then

$$-\frac{1}{h} [(\mathcal{T}_{t,t+h}v(t+h, \cdot))(x) - v(t, x)] = 0, \quad (3.13)$$

for all $x \in \mathbb{D}$ and $t < t + h \leq T$. To continue even formally, we need to assume that the above quantity has a limit as $h \downarrow 0$, when v is “smooth”. So we assume that there exist an open set $\mathbb{D}' \subset \mathbb{D}$, a set of smooth functions $\mathcal{D} \subset C((-\infty, T) \times \mathbb{D}')$ and a one-parameter

1 family of nonlinear operators $\{\mathcal{G}_t\}_{t \leq T}$ of functions of \mathbb{D} , satisfying the following conditions
 2 with

$$3 \quad Q = (-\infty, T) \times \mathbb{D}', \quad 4$$

$$5 \quad \varphi_t(t, x), (\mathcal{G}_t \varphi(t, \cdot))(x) \in C(Q) \quad \text{and} \quad \varphi(t, \cdot) \in \mathcal{C} \quad \forall t \leq T, \quad (3.14) \quad 5$$

$$6 \quad \varphi, \tilde{\varphi} \in \mathcal{D}, \lambda \geq 0 \implies \varphi + \tilde{\varphi} \in \mathcal{D}, \lambda \varphi \in \mathcal{D}, \quad (3.15) \quad 6$$

$$7 \quad \lim_{h \downarrow 0} \frac{1}{h} [(\mathcal{T}_{t, t+h} \varphi(t+h, \cdot))(x) - \varphi(t, x)] = \varphi_t(t, x) - (\mathcal{G}_t \varphi(t, \cdot))(x) \quad (3.16) \quad 7$$

8 for all $\varphi \in \mathcal{D}$, $(t, x) \in Q$. We refer to the elements of \mathcal{D} as *test functions* and \mathcal{G}_t as the
 9 *infinitesimal generator* of the semigroup $\{\mathcal{T}_{t,r}\}$. Note that, if φ is any test function, then
 10 $\varphi_t(t, x)$ is defined for all $(t, x) \in (-\infty, T] \times \mathbb{D}$ even though (3.16) is required to hold only
 11 for $(t, x) \in Q$.

12 Like the choice of \mathcal{C} , the exact choice of \mathcal{D} is not important. One should think of \mathcal{D}
 13 as the set of “smooth” functions. For example, if $\mathbb{D}' = O$ is a bounded subset of \mathbb{R}^d and
 14 $\mathbb{D} = \overline{O}$, then we require that \mathcal{D} contains $C^\infty(\overline{Q})$. Indeed, this requirement will be typical
 15 when \mathcal{G}_t is a partial differential operator.

16 In most applications, \mathbb{D}' is simply the interior of \mathbb{D} . However, in the case of a controlled
 17 jump Markov process which is stopped after the exit from an open set $O \subset \mathbb{R}^d$, we have
 18 $\mathbb{D}' = O$, while \mathbb{D} is the closure of the set that can be reached from O .

19 Now suppose that $v \in \mathcal{D}$ and let h go to zero in (3.13). The result is

$$20 \quad -v_t(t, x) + (\mathcal{G}_t v(t, \cdot))(x) = 0, \quad (t, x) \in Q. \quad (3.17) \quad 20$$

21 In analogy with optimal control, the above equation is called the *(abstract) dynamic*
 22 *programming equation*.

23 In general, the value function is not in \mathcal{D} and therefore it is *not* a classical solution
 24 of (3.17). In that case the equation (3.17) has to be interpreted in a weaker sense. This will
 25 be the subject of viscosity solutions.

26 We are now in a position to give the definition of viscosity solutions in the abstract
 27 setting. This is a straightforward generalization of the original definition given by Crandall
 28 and Lions [25]. Also see Crandall, Evans and Lions [23]. Let $Q = (-\infty, T) \times \mathbb{D}'$, \mathcal{D} and \mathcal{C}
 29 as before.

30 In the below definition, we assume continuity to simplify the presentation. However, for
 31 the definition we only need the solution to be locally bounded, see [4,36].

32 DEFINITION 3.2 (Viscosity solutions). Let $w \in C((-\infty, T] \times \mathbb{D})$. Then

33 (i) w is a *viscosity subsolution* of (3.17) in Q if for each $\varphi \in \mathcal{D}$,

$$34 \quad -\varphi_t(\bar{t}, \bar{x}) + (\mathcal{G}_{\bar{t}} \varphi(\bar{t}, \cdot))(\bar{x}) \leq 0, \quad (3.18) \quad 34$$

35 at every $(\bar{t}, \bar{x}) \in Q$ which is a maximizer of $w - \varphi$ on $(-\infty, T] \times \mathbb{D}$ with $w(\bar{t}, \bar{x}) = \varphi(\bar{t}, \bar{x})$.

36 (ii) w is a *viscosity supersolution* of (3.17) in Q if for each $\varphi \in \mathcal{D}$,

$$37 \quad -\varphi_t(\bar{t}, \bar{x}) + (\mathcal{G}_{\bar{t}} \varphi(\bar{t}, \cdot))(\bar{x}) \geq 0, \quad (3.19) \quad 37$$

at every $(\bar{t}, \bar{x}) \in Q$ which is a minimizer of $w - \varphi$ on $(-\infty, T] \times \mathbb{D}$ with $w(\bar{t}, \bar{x}) = \varphi(\bar{t}, \bar{x})$.
 (iii) w is a viscosity solution of (3.17) in Q if it is both a viscosity subsolution and a viscosity supersolution of (3.17) in Q .

It follows from the monotonicity and the semigroup properties and the definitions that any classical solution of (3.17) is also a viscosity solution, see for instance [36]. Another immediate consequence is the following.

THEOREM 3.3. *Assume (3.5), (3.6), (3.14)–(3.16). Suppose that the value function v defined by (3.11) is continuous. Then, v is a viscosity solution of (3.17) in Q .*

PROOF. Let $\varphi \in \mathcal{D}$ and $(\bar{t}, \bar{x}) \in Q$ be a maximizer of the difference $v - \varphi$ on \bar{Q} satisfying $v(\bar{t}, \bar{x}) = \varphi(\bar{t}, \bar{x})$. Then, $\varphi \geq v$. Using (3.6) with $\phi = \varphi(r, \cdot)$ and $s = T$, we obtain for every $r \in [\bar{t}, T]$,

$$(\mathcal{T}_{\bar{t}, r} \varphi(r, \cdot))(\bar{x}) \geq (\mathcal{T}_{\bar{t}, T} \psi)(\bar{x}) = v(\bar{t}, \bar{x}) = \varphi(\bar{t}, \bar{x}).$$

Recall that by (3.14), $\varphi(r, \cdot)$ is in the domain of $\mathcal{T}_{\bar{t}, r}$. Take $r = \bar{t} + h$ and use (3.16) to arrive at

$$-\varphi_{\bar{t}}(\bar{t}, \bar{x}) + (\mathcal{G}_{\bar{t}} \varphi(\bar{t}, \cdot))(\bar{x}) = -\lim_{h \downarrow 0} \frac{1}{h} [(\mathcal{T}_{\bar{t}, \bar{t}+h} \varphi(\bar{t} + h, \cdot))(\bar{x}) - \varphi(\bar{t}, \bar{x})] \leq 0.$$

Hence (3.18) is satisfied and consequently v is a viscosity subsolution of (3.17) in Q . The supersolution property of v is proved *exactly* the same way as the subsolution property. \square

3.2. Optimal control of diffusion processes

In this subsection, we will prove that the value function of an optimal control problem is the unique viscosity solution of the dynamic programming equation (3.1) with \mathcal{H} as in (3.3) with \mathcal{L}^α given by (3.10). In order to achieve this, we will define a stochastic nonlinear semigroup as in the introduction of this section and in Example 3.1. Then, we will verify the assumptions of Theorem 3.3 to show that the value function of the stochastic optimal control problem is the viscosity solution.

If the controlled Markov processes are uniformly parabolic, then there are classical solutions to the dynamic programming equation (3.17) and uniqueness is standard under natural conditions well known in the PDE literature. Combined with these PDE results, the viscosity property of the value function provides a stochastic representation. However, we do not, in general, assume the uniform parabolicity and therefore we only expect the value function to be a viscosity solution. Still in this case, there are uniqueness results for viscosity solutions (see [24,36]) and a representation result follows. The main difference between the smooth (or equivalently the uniformly elliptic) case and the nonsmooth case is that, as for the linear problems, smooth solutions with certain growth conditions can be directly shown to be the value function. This point is further developed in the next subsection. In

1 this section, we consider the boundary value problems with no exponential discounting 1
 2 (i.e., equation with no linear term: $r \equiv 0$). However, our results easily extend to the other 2
 3 cases as described in Section 2. 3

4 Now let $\mathcal{T}_{t,r}$ be the semigroup defined in Example 3.1. To simplify the presentation, we 4
 5 take $r \equiv 0$. So that the nonlinear operators on \mathcal{C} are given by 5
 6

$$\begin{aligned}
 (\mathcal{T}_{t,r}\varphi)(x) = \inf_{\alpha(\cdot) \in \mathcal{A}} E \left\{ \int_t^{\theta \wedge r} L(s, X(s), \alpha(s)) ds \right. \\
 \left. + g(\theta, X(\theta))\chi_{\theta < r} + \varphi(X(r))\chi_{\theta \geq r} \right\},
 \end{aligned}$$

7 where $t \leq r \leq T$, $\varphi \in \mathcal{C}$, θ is the exit time of $(s, X(s))$ from $Q = (-\infty, T] \times \bar{O}$ and 7
 8 $g \in C(\bar{Q})$ is a given function, which we call the *lateral boundary data*. Clearly $\mathcal{T}_{t,r}$ satisfies 8
 9 (3.5) and (3.6). The semigroup property however, is equivalent to the dynamic program- 9
 10 ming principle. We refer to Chapter 5 in [36] or [66] for the general structure of dynamic 10
 11 programming in a certain context. 11

12 To apply the results of Theorem 3.3, we also have to verify (3.16). Indeed we shall prove 12
 13 that (3.16) holds with $\mathbb{D} = \bar{O}$, $\mathbb{D}' = O$ (hence $Q = (-\infty, T] \times O$), $\mathcal{D} = C^{1,2}(\bar{Q})$ and the 13
 14 infinitesimal generator, 14
 15

$$(\mathcal{G}_t\varphi)(x) = \mathcal{H}(t, x, \varphi(\cdot)), \quad (t, x) \in Q,$$

16 where \mathcal{H} is as in (3.3) and \mathcal{L}_t^α as in (3.10). In view of Theorem 3.3, this result will im- 16
 17 ply that the value function is a viscosity solution of the dynamic programming equation 17
 18 provided that it is continuous. 18

19 Recall O is assumed to be bounded. For the unbounded case, we refer to [36]. 19
 20

21 **THEOREM 3.4.** *Suppose that f, σ satisfy (3.7), A is compact and g, L are continuous. 21
 22 Then, for every $w \in \mathcal{D}$ and $(t, x) \in Q$, we have 22
 23*

$$\lim_{h \downarrow 0} \frac{1}{h} [(\mathcal{T}_{t,t+h}w(t+h, \cdot))(x) - w(t, x)] = w_t(t, x) - (\mathcal{G}_t w(t, \cdot))(x).$$

24 **PROOF.** We start with a probabilistic estimate. Let $X(\cdot)$ be the solution of (3.8) with con- 24
 25 trol $\alpha(\cdot)$ and initial condition $X(t) = x \in O$. Since Q is bounded, f and σ are bounded, 25
 26 for any positive integer m and $h \in (0, 1]$, we have 26
 27

$$\begin{aligned}
 E \sup_{t \leq \rho \leq t+h} |X(\rho) - x|^{2m} \\
 = E \sup_{t \leq \rho \leq t+h} \left| \int_t^\rho f(s, X(s), \alpha(s)) ds \right. \\
 \left. + \int_t^\rho \sigma(s, X(s), \alpha(s)) dW(s) \right|^{2m}
 \end{aligned}$$

$$\begin{aligned}
&\leq \widehat{C}_m E \left(\int_t^{t+h} |f(s, X(s), \alpha(s))| ds \right)^{2m} \\
&\quad + \widehat{C}_m E \sup_{t \leq \rho \leq t+h} \left| \int_t^\rho \sigma(s, X(s), \alpha(s)) dW(s) \right|^{2m} \\
&\leq \widehat{C}_m \|f\|^{2m} h^{2m} + \widetilde{C}_m \|\sigma\|^{2m} h^m \leq C_m h^m,
\end{aligned} \tag{3.20}$$

where $\|\cdot\|$ denotes the sup-norm on Q and $C_m, \widehat{C}_m, \widetilde{C}_m$ are suitable constants. Set $d(x) = \text{dist}(x, \partial O)$ and recall that θ is the exit time from Q . Then for $t+h \leq T$,

$$\begin{aligned}
P(\tau \leq t+h) &\leq P\left(\sup_{t \leq \rho \leq t+h} |X(\rho) - x| \geq d(x) \right) \\
&\leq \left(E \sup_{t \leq \rho \leq t+h} |X(\rho) - x|^{2m} \right) (d(x))^{-2m} \\
&\leq \frac{C_m h^m}{(d(x))^{2m}}.
\end{aligned} \tag{3.21}$$

Fix $\alpha \in A$ and let $\bar{\alpha}(s) \equiv \alpha$. Then the definition of $\mathcal{T}_{t,t+h}$ yields

$$\begin{aligned}
I(h) &= \frac{1}{h} \left[(\mathcal{T}_{t,t+h} w(t+h, \cdot))(x) - w(t, x) \right] \\
&\leq \frac{1}{h} E \int_t^{(t+h) \wedge \theta} L(s, X(s), \alpha) ds \\
&\quad + \frac{1}{h} E [w(t+h, X(t+h)) - w(t, x)] \chi_{\theta \geq t+h} \\
&\quad + \frac{1}{h} E [g(\theta, X(\theta)) - w(t, x)] \chi_{\theta < t+h}.
\end{aligned} \tag{3.22}$$

The estimate (3.21) with $m = 2$ yields

$$\lim_{h \downarrow 0} \frac{1}{h} P_{t,x}(\tau \leq t+h) = 0$$

for every $(t, x) \in Q$. Hence

$$\lim_{h \downarrow 0} \frac{1}{h} E \int_t^{(t+h) \wedge \theta} L(s, X(s), \alpha) ds = L(t, x, \alpha)$$

and

$$\lim_{h \downarrow 0} \frac{1}{h} E [g(\theta, X(\theta)) - w(t, x)] \chi_{\theta < t+h} = 0.$$

Also, by Itô's formula

$$\begin{aligned}
& \lim_{h \downarrow 0} \frac{1}{h} E[w(t+h, x(t+h)) - w(t, x)] \chi_{\theta \geq t+h} \\
&= \lim_{h \downarrow 0} \frac{1}{h} E[w((t+h) \wedge \theta, X((t+h) \wedge \theta)) - w(t, x)] \\
&= \lim_{h \downarrow 0} \frac{1}{h} E \int_t^{(t+h) \wedge \theta} [w_t(s, X(s)) + (\mathcal{L}_s^\alpha w(s, \cdot))(X(s))] ds \\
&= (\mathcal{L}_t^\alpha w(t, \cdot))(x).
\end{aligned}$$

Substitute the above into (3.21) to obtain

$$\limsup_{h \downarrow 0} I(h) \leq L(t, x, \alpha) + w_t(t, x) + (\mathcal{L}_t^\alpha w(t, \cdot))(x)$$

for all $\alpha \in A$. We take the infimum over α . The result is

$$\limsup_{h \downarrow 0} I(h) \leq w_t(t, x) - (\mathcal{G}_t w(t, \cdot))(x).$$

For any sequence $h_n \downarrow 0$, there exists $\alpha_n(\cdot)$ satisfying

$$\begin{aligned}
& (\mathcal{T}_{t, t_n} w(t_n, \cdot))(x) \\
& \geq E \left[\int_t^{\theta_n} L(s, x_n(s), \alpha_n(s)) ds + g(\theta_n, X_n(\theta_n)) \chi_{\theta_n < t_n} \right. \\
& \quad \left. + w(t_n, X_n(t_n)) \chi_{\theta_n = t_n} \right] - (h_n)^2,
\end{aligned}$$

where $t_n = t + h_n$, $\theta_n = \hat{\theta}_n \wedge t_n$, $X_n(\cdot)$ is the solution of (3.8), (2.13) with control α_n , and $\hat{\theta}_n$ is the exit time of $(s, X_n(s))$ from Q . Therefore

$$\begin{aligned}
I(h_n) & \geq \frac{1}{h_n} E \int_t^{\theta_n} L(s, X_n(s), \alpha_n(s)) ds \\
& \quad + \frac{1}{h_n} E[w(t_n, X(t_n)) - w(t, x)] \chi_{\theta_n = t_n} \\
& \quad + \frac{1}{h_n} E[g(\theta_n, X_n(\theta_n)) - w(t, x)] \chi_{\theta_n < t_n} - h_n. \tag{3.23}
\end{aligned}$$

The probabilistic estimate (3.21) with $m = 2$ implies that the limit of the third term is zero

and

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \left| \frac{1}{h_n} E \left(\int_t^{t_n} L(t, x, \alpha_n(s)) ds - \int_t^{\theta_n} L(s, X_n(s), \alpha_n(s)) ds \right) \right| \\
& \leq \lim_{n \rightarrow \infty} \frac{1}{h_n} \left[\|L\|_\infty E(t_n - \theta_n) \right. \\
& \quad \left. + E \int_t^{t_n} |L(t, x, \alpha_n(s)) - L(s, X_n(s), \alpha_n(s))| ds \right]. \quad (3.24)
\end{aligned}$$

Since $\bar{Q} \times A$ is compact, L is uniformly continuous. Also (3.20) implies that, for every $\delta > 0$,

$$\lim_{n \rightarrow \infty} P \left(\sup_{t \leq \rho \leq t+h_n} |X_n(\rho) - x| \geq \delta \right) = 0.$$

Therefore the uniform continuity of L and (3.20) imply that the limits in (3.24) are zero. We now use (3.21) and Itô formula to obtain

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \frac{1}{h_n} \left| E \left\{ [w(t_n, X(t_n)) - w(t, x)] \chi_{\theta_n=t_n} \right. \right. \\
& \quad \left. \left. - \int_t^{t_n} w_t(t, x) + (\mathcal{L}_t^{\alpha_n(s)} w(t, \cdot))(x) ds \right\} \right| \\
& \leq \lim_{n \rightarrow \infty} \frac{1}{h_n} \sup_\alpha |w_t(t, x) + (\mathcal{L}_t^\alpha w(t, \cdot))(x)| E(t_n - \theta_n) \\
& \quad + \lim_{n \rightarrow \infty} \frac{1}{h_n} E \int_t^{\theta_n} |(\mathcal{L}_t^{\alpha_n(s)} w(s, \cdot))(X_n(s)) - (\mathcal{L}_s^{\alpha_n(s)} w(t, \cdot))(x)| ds \\
& \quad + \lim_{n \rightarrow \infty} \frac{1}{h_n} E \int_t^{\theta_n} |w_t(s, X_n(s)) - w_t(t, x)| ds.
\end{aligned}$$

Since $w \in C^{1,2}(\bar{Q})$, $\mathcal{L}_s^\alpha w(s, y)$ is a uniformly continuous function of \bar{Q} . As in (3.24), the dominated convergence theorem and (3.20) imply that the above limit is zero. Combine this with (3.23) and (3.24) to obtain

$$I(h_n) \geq L^n + G^n - e(n),$$

where

$$\begin{aligned}
(L^n, G^n) := & \frac{1}{h_n} \left(E \int_t^{t+h_n} L(t, x, \alpha_n(s)) ds, \right. \\
& \left. E \int_t^{t+h_n} w_t(t, x) + (\mathcal{L}_t^{\alpha_n(s)} w(t, \cdot))(x) ds \right),
\end{aligned}$$

and the error term $e(n)$ converges to zero as $n \rightarrow \infty$. Define a set

$$\begin{aligned} \widehat{A} = \{ (L, G) \in \mathbb{R}^2 : & L = L(t, x, \alpha), \\ & G = w_t(t, x) + (\mathcal{L}_t^\alpha w(t, \cdot))(x) \text{ for some } \alpha \in A \}. \end{aligned}$$

Then $(L^n, G^n) \in \overline{\text{co}}(\widehat{A})$, where $\overline{\text{co}}$ denotes the convex, closed hull of \widehat{A} . Also,

$$\begin{aligned} L^n + G^n &\geq \inf\{L + G : (L, G) \in \overline{\text{co}}(\widehat{U})\} \\ &= \inf\{L + G : (L, G) \in \widehat{U}\} \\ &= w_t(t, x) - (\mathcal{G}_t w(t, \cdot))(x). \end{aligned} \quad \square$$

As in Section 3.1, let v be the value function. Then, in view of Theorem 3.3, we have the following representation result for (3.1) with nonlinearity \mathcal{H} given in (3.3) with \mathcal{L}^α given in (3.10) and $r \equiv 0$. However, this restriction that $r \equiv 0$ can easily be removed by the techniques developed in Section 2. For that reason we state the result including the linear term $r(t, z, \alpha)v(t, x)$. For uniqueness we need the boundary conditions. It is clear that v satisfies the terminal condition (2.20). Also, under some conditions, the value function satisfies the Dirichlet boundary condition (2.21) (see Chapter 5 in [36]). Also, in degenerate cases, (2.21) may hold only in the viscosity sense. We refer the interested reader to the book of Barles [4], or Section 7.6 in [36].

COROLLARY 3.5 (Control representation for (3.1)). *Suppose that $v \in C(\overline{Q})$. Then v , a viscosity solution of the dynamic programming equation (3.1), with nonlinearity \mathcal{H} given in (3.3) with the infinitesimal generator as in (3.10), i.e.,*

$$-v_t(t, x) + H(t, x, v(t, x), D(t, x), D^2v(t, x)) = 0 \quad \text{on } (-\infty, T) \times O, \quad (3.25)$$

where

$$\begin{aligned} H(t, x, v, p, B) &= \sup_{\alpha \in A} \left\{ -r(t, x, \alpha)v - \mu(t, x, \alpha) \cdot p - \frac{1}{2}a(t, x, \alpha) : B - L(t, x, \alpha) \right\}. \end{aligned}$$

In particular, if v satisfies (2.21) and if there is only one continuous viscosity solution of (3.25) together with (2.21), (2.20), then this solution is given as the value function of the stochastic optimal control problem.

3.3. Smooth value function and verification

In this subsection, we assume that there exists a $u \in C^{1,2}((-\infty, T] \times \overline{O})$ that solves the dynamic programming equation (3.25) together with boundary conditions (2.21), (2.20).

Then, we will show by a direct application of Itô calculus that this solution must be equal to the value of the stochastic semigroup defined in Example 3.1. This, in particular, proves uniqueness of $C^{1,2}$ solutions to equations (3.25), (2.20), (2.21).

THEOREM 3.6 (Verification). *Let $u \in C^{1,2}((-\infty, T] \times \bar{O})$ be a solution of (3.25), (2.20), (2.21). Then, for every $\alpha(\cdot) \in \mathcal{A}$,*

$$u(t, x) \leq (\mathcal{J}_{t,T}^{\alpha(\cdot)} \varphi)(x).$$

In addition, suppose that there exists an optimal control $\alpha^(\cdot) \in \mathcal{A}$ so that, for Lebesgue almost all $s \in [t, T]$,*

$$\alpha^*(s) \in \arg \min_{\alpha \in \mathcal{A}} \left\{ \begin{aligned} & -r(s, X^*(s), \alpha^*(s))v(s, X^*(s)) \\ & -\mu(s, X^*(s), \alpha^*(s)) \cdot Dv(s, X^*(s)) \\ & -\frac{1}{2}a(s, X^*(s), \alpha^*(s)) : D^2v(s, X^*(s)) \\ & -L(s, X^*(s), \alpha^*(s)) \end{aligned} \right\},$$

where $X^*(s)$ is the solution of (3.8), with initial data (2.13) and control $\alpha^*(\cdot)$. Then

$$u(t, x) = (\mathcal{J}_{t,T}^{\alpha^*(\cdot)} \varphi)(x).$$

PROOF. Fix (t, x) and $\alpha(\cdot)$. Let $X(\cdot)$ be the corresponding state process. Apply the Itô rule to $Y(s) := B(s)u(s, X(s))$. The result is

$$\begin{aligned} dY(s) = & B(s) \left[r(s, X(s), \alpha(s))u(s, X(s)) + \mu(s, X(s), \alpha(s)) \cdot Du(s, X(s)) \right. \\ & \left. + \frac{1}{2}a(s, X(s), \alpha(s)) : D^2u(s, X(s)) + u_t(s, X(s)) \right] ds \\ & + B(s)Du(s, X(s))dW(s). \end{aligned}$$

We integrate the above on $[t, \theta]$, take the expected value and then use the equations (3.25), (2.21), (2.20). The result is

$$\begin{aligned} Y(t) &= u(t, x) \\ &= E[B(\theta)u(\theta, X(\theta))] \\ &\quad - E \left(\int_t^\theta B(s) \left[u_t(s, X(s)) + r(s, X(s), \alpha(s))u(s, X(s)) \right. \right. \\ &\quad \left. \left. + \mu(s, X(s), \alpha(s)) \cdot Du(s, X(s)) \right] ds \right) \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{2} a(s, X(s), \alpha(s) : D^2 u(s, X(s)) \Big] ds \Big) \\
 & \geq E[B(\theta)u(\theta, X(\theta))] + E\left(\int_t^\theta B(s)L(s, X(s), \alpha(s)) ds\right) \\
 & = (\mathcal{J}_{t,T}^{\alpha(\cdot)} \varphi)(x).
 \end{aligned}$$

This proves the first part of the statement. To prove the second part, we repeat the above calculations with the control $\alpha^*(\cdot)$ and $X^*(\cdot)$. The inequality in the above calculation is now an equality and the optimality of $\alpha^*(\cdot)$ follows. \square

3.4. Optimal control of jump Markov processes

In this subsection, we briefly discuss the extension of the representation results for the integro-differential equations discussed in Section 2.6. We refer to the books [36,55] and the paper [2] for more information.

As in the diffusion case, we will introduce controlled diffusion equations driven by a Markov process and a random measure. Indeed, let A be a control set, and μ, σ, L be as in Section 3.2. Further, let π be a compensator measure on \mathbb{R}^d and f be function satisfying

$$\sup_{\alpha \in A} \int_{\mathbb{R}^d} [1 \wedge |f(t, x, \xi, \alpha)|^2] \pi(d\xi) < \infty.$$

We also assume the standard Lipschitz condition

$$\int_{\mathbb{R}^d} |f(t, x, \xi, \alpha) - f(t, y, \xi, \alpha)|^2 \pi(d\xi) \leq C|x - y|^2$$

for all $t \in \mathbb{R}^1, x, y \in \mathbb{R}^d$ for some constant C . As before, let \mathcal{A} be all bounded, progressively measurable, A -valued random processes. Then, given a control process $\alpha(\cdot) \in \mathcal{A}$ and initial point (t, x) , there exists a unique solution to

$$\begin{aligned}
 X(T) = x & + \int_t^T \mu(s, X(s), \alpha(s)) ds + \int_t^T \sigma(s, X(s), \alpha(s)) dW(s) \\
 & + \int_t^T \int_{\mathbb{R}^d} f(s, X(s), \xi, \alpha(s)) \tilde{p}(ds \times d\xi),
 \end{aligned} \tag{3.26}$$

where \tilde{p} is a martingale random measure with compensator π and $W(\cdot)$ is an independent standard Brownian motion.

For a constant control $\alpha(\cdot) \equiv \alpha$, the infinitesimal generator is given by

$$\mathcal{L}_t^\alpha \varphi(x) = \mu(t, x, \alpha) \cdot \nabla \varphi(x) + \frac{1}{2} a(t, x, \alpha) : D^2 \varphi(x) + \tilde{\mathcal{L}}_t^\alpha \varphi(x),$$

where the part corresponding to the random measure is as in (2.19),

$$\tilde{\mathcal{L}}_t \alpha \varphi(x) = \int_{\mathbb{R}^d} [\varphi(x + f(t, x, \xi, \alpha)) - \varphi(x) - f(t, x, \xi, \alpha) \cdot \nabla \varphi(x)] \pi(d\xi).$$

Now we define the pay-off functional and the value function as in the diffusion case (see (3.9)),

$$\begin{aligned} (\mathcal{J}_{t,T}^{\alpha(\cdot)} \psi)(x) = E & \left[\int_t^T B(s) L(s, X(s), \alpha(s)) ds \right. \\ & \left. + B(\theta) g(\theta, X(\theta)) \chi_{\theta < T} + B(T) \psi(X(T)) \chi_{\theta \geq T} \right], \end{aligned}$$

where

$$B(T) = B(t, T; z, \alpha(\cdot)) = \exp\left(-\int_t^T r(s, Z(s), \alpha(s)) ds\right),$$

and θ is the exit time of $(s, X(s))$ from $\bar{Q} = [t, T] \times \bar{O}$. The nonlinear semigroup is given by

$$(\mathcal{T}_{t,r} \psi)(x) := \inf_{\alpha(\cdot) \in \mathcal{A}} \mathcal{J}_{t,r}^{\alpha(\cdot)} \psi(x).$$

Again it follows that $\mathcal{T}_{t,r}$ is monotone, semigroup by dynamic programming for optimal control.

We can then show that the value function is a viscosity solution of the corresponding dynamic programming equation. Such a result for jump Markov processes was first proved in [63], and then by Sayah [61]. We refer to the recent book of Oksendall and Sulem [55] for more information.

Given the form of the generator \mathcal{L}_t^α and the formal discussion given in the beginning of this section, the related dynamic programming equation is (3.1), with \mathcal{H} in (3.3) and \mathcal{L}_t^α as above.

The boundary conditions are (2.20) and (2.21). But the important point to emphasize is that (2.21) holds for all x not only on ∂O but in all of $\mathbb{R}^d \setminus O$ as $X(\theta) \in \mathbb{R}^d \setminus O$.

3.5. Other type of control problems

Several other types of control problems have been studied in the literature. These problems are related to so-called quasivariational inequalities. Indeed, stopping time problems are related to obstacle problems [9]. Impulse or switching controls yield quasivariational inequalities [10]. Singular control problems allow the state processes to be discontinuities [36]. Dynamic programming equations for singular control problem are again quasivariational inequalities but with constraints on the first derivative of the solution. Equations with constraints on the second derivatives are much rare and obtained only in [20].

3.6. Stochastic target and geometric problems

In this section, we consider a special class of nonlinear parabolic equations. These equations are related to geometric flows of manifolds embedded in \mathbb{R}^d . The nonlinearities H that appear in these equation are, in addition to being parabolic (2.11), also geometric, i.e.,

$$\mathcal{H}(t, x, \lambda p, B + \mu p \otimes p) = \lambda \mathcal{H}(t, x, p, B) \quad \forall \lambda \geq 0, \mu \in \mathbb{R}^1. \quad (3.27)$$

It was shown in [67,68] that a large subclass of above nonlinearities have a stochastic representation similar to that discussed in Section 3.2. In this representation, however, a new class of control problems called *stochastic target problems* are used [65,66]. A stochastic target problem is a nonclassical control problem in which the controller tries to steer a controlled stochastic process into a given target set G by judicious choices of controls. The chief object of study is the set of all initial positions from which the controlled process can be steered into G with *probability one* in an allowed time interval. Clearly these *reachability sets* depend on the allowed time. Thus, they can be characterized by an evolution equation which is the analogue of the dynamic programming equation of stochastic optimal control.

Geometric equations express the velocity of the boundary as a possibly nonlinear function of the normal and the curvature vectors. In [66,67] it was shown that smooth solutions of these geometric equations, when exist, are equal to the reachability sets. However, as a Cauchy problem, these equations in general do not admit classical smooth solutions and a weak formulation is needed. Several such formulations were given starting with the pioneering work of Brakke [15]. Here we consider the viscosity formulation given independently by Chen, Giga and Goto [19] and by Evans and Spruck [34]. The main idea of this approach is to characterize the geometric solution as the zero level set of a continuous function. Then, this function solves a partial differential equation (3.25) with a geometric H satisfying (3.27).

The chief goal of this subsection is to give a stochastic characterization of the unique level set solutions of [19,34] in terms of the target problem. The stochastic semigroup is given by

$$v(t, x) := \inf_{\alpha(\cdot) \in \mathcal{A}} \operatorname{ess\,sup}_{\omega \in \Omega} \varphi(X_{t,x}^{\alpha(\cdot)}(T, \omega)), \quad (3.28)$$

where for initial data $(t, x) \in (-\infty, T) \times \mathbb{R}^d$, control process $\alpha(\cdot) \in \mathcal{A}$, the controlled process $\{X(s) := X_{t,x}^{\alpha(\cdot)}(s)\}_{s \geq t}$ is the solution of (3.8) and (2.13).

The following representation result is proved in [66].

THEOREM 3.7. *Suppose that the standard Lipschitz assumption (3.7) holds and that H is locally Lipschitz on $\{p \neq 0\}$. Then, v defined in (3.28) satisfies (2.20) pointwise and it is a discontinuous viscosity solution of (3.25) with*

$$H(t, x, p, B) := \sup_{v \in \mathcal{N}(t, x, p)} \left\{ -\mu(t, x, \alpha) \cdot p - \frac{1}{2} a(t, x, \alpha) : B \right\}, \quad (3.29)$$

The proof of this theorem is a straightforward application of Theorem 3.7 and the results of [6]. Observe that the boundedness of φ is not a restriction, as one can replace φ by $\varphi(1 + |\varphi|)^{-1}$.

The stochastic target problems with jump-diffusion processes are discussed by Bouchard [13]. Also, target problems are related to forward-backward stochastic differential equations (FBSDEs) discussed in Section 4. A similar representation theorem for the special case of the codimension-one mean curvature flow was also obtained by Buckdahn, Cardaliaguet and Quincampoix [16].

We close this subsection by the important example of mean curvature flow.

EXAMPLE 3.9. Consider the example with $A = \mathcal{P}^{k,d}$ be the set of all projection matrices on \mathbb{R}^d onto a hyperplane of dimension k , $\mu \equiv 0$ and $\sigma(t, x, \alpha) = \sqrt{2}\alpha$. Then, the state equation (3.8) reduces to

$$dX(s) = \sqrt{2}\alpha(s) dW(s),$$

where $W(\cdot)$ is the standard d -dimensional Brownian motion. Hence, at each time s , the controller decides on which k -dimensional space $X(\cdot)$ should diffuse. Then the related PDE has the form

$$\mathcal{H}^k(p, B) = \sup\{\alpha : B \mid \alpha \in \mathcal{P}^{k,d} \text{ and } \alpha p = 0\}.$$

This is exactly the same nonlinear function used by Ambrosio and Sonner [1] to describe the weak flow of codimension $d - k$ mean curvature flow. In the special case of $k = d - 1$, any $\alpha \in \mathcal{P}^{d-1,d}$ is given by $\alpha = I - \nu \otimes \nu$ for some unit vector $\nu \in \mathbb{R}^d$. Also for such a matrix α and $p \neq 0$,

$$\alpha p = 0 \implies [I - \nu \otimes \nu]p = 0 \implies \nu = \pm \frac{p}{|p|} := \bar{p}.$$

Hence,

$$\mathcal{H}^{d-1}(p, B) = \sup\{[I - \nu \otimes \nu] : B \mid \nu = \pm \bar{p}\} = \text{trace}(B) - B\bar{p} \cdot \bar{p},$$

and equation (3.25) has the form

$$-v_t(t, x) - \Delta v(t, x) + \frac{D^2 v(t, x) \nabla v(t, x) \cdot \nabla v(t, x)}{|\nabla v(t, x)|^2} = 0.$$

This is the level set equation (in reversed time) for the mean curvature flow [19,34]. Then, in the special case of this example, results of this subsection can be stated as follows. The unique viscosity solution of the above level set equation of the mean curvature flow has the stochastic representation

$$v(t, x) := \inf_{\alpha(s) \in \mathcal{P}^{k,d}} \text{ess sup}_{\omega \in \Omega} \varphi \left(x + \int_t^T \alpha(s) dW(s) \right).$$

4. Backward representations

In this section we outline a different connection between PDEs and stochastic processes. Vaguely, this connection is analogous to the connection between ordinary differential equations and first-order PDEs through the method of characteristics. Indeed, it is first observed by Bismut [11] in his seminal work on the extension of Pontryagin maximum principle to stochastic optimal control. Pontryagin's maximum principle itself is the extension of the Hamilton–Jacobi theory of classical mechanics to deterministic optimal control and provides conditions for maximality through a set of ordinary differential equations. For stochastic optimal control, Bismut achieved this using stochastic processes. As well known, the method of characteristics and its mentioned generalizations have both initial and terminal boundary data to be satisfied. In the stochastic context, due to the adaptability conditions, this makes the problem harder. However, a deep theory is now developed through the recent works of Peng, Pardoux and others [32,54,56–58,60]. This theory known as *Backward stochastic differential equations* (BSDEs) will be outlined in the next subsection. BSDEs have a natural connection with PDEs and several numerical methods have been developed [27,53,69]. However, the PDEs connected to BSDEs are always quasilinear. Recently, Cheridito, Soner, Touzi and Victoir [21] extended this theory to cover all fully nonlinear, parabolic, second-order PDEs. This extension and the possible numerical implications are outlined below. For a more complete introduction to BSDEs we refer to the survey paper of El-Karoui, Peng and Quenez [32].

4.1. Backward stochastic differential equations

Let $X(\cdot) := X^{t,x}(\cdot)$ be the solution of (2.12), (2.13). Given real-valued, nonlinear function f and terminal data φ consider the equation

$$dY(s) = f(s, X(s), Y(s)) ds + Z(s) \cdot \sigma(s, X(s)) dW, \quad (4.1)$$

with terminal data

$$Y(T) = \varphi(X(T)). \quad (4.2)$$

The problem is to find processes $Y(\cdot)$ and $Z(\cdot)$ that are integrable and adapted to the filtration \mathcal{F}_t . Adaptedness condition is a serious technical condition as the given data for $Y(\cdot)$ is specified at the terminal time T . In the probabilistic literature the BSDE is defined more generally. Random f and more general $X(\cdot)$ process with $Y(\cdot)$ dependence are also considered. Here we restrict ourselves to the above framework to simplify the presentation.

Let us assume that the solution $Y(s)$ is given as a deterministic function of time and $X(s)$, i.e., assume that there is a deterministic function v so that

$$Y(s) = v(s, X(s)) \quad \forall s \in [t, T].$$

If we also assume that v is smooth, then by the Itô formula, we have

$$\begin{aligned} d[v(s, X(s))] &= [v_t(s, X(s)) + (\mathcal{L}_s v(s, \cdot))(X(s))] ds \\ &\quad + \nabla v(s, X(s)) \cdot \sigma(s, X(s)) dW(s), \end{aligned}$$

where \mathcal{L}_t is as in (2.14). If σ has full rank, then equating the above equation to (4.1) yields

$$Z(s) = \nabla v(s, X(s)),$$

and v must solve

$$-v_t - \mathcal{L}_t v + f(t, x, v, \nabla v) = 0 \quad \text{on } (-\infty, T] \times \mathbb{R}^d.$$

Hence, smooth solutions of the above semilinear PDE has the representation in terms of the BSDE (4.1). Numerical implication of this connection is discussed in Section 5.2. Also the rigorous connection between the PDE and the BSDE is given in the references cited before.

4.2. Second-order backward stochastic equations

In the BSDE literature it has not been possible to consider PDEs with a nonlinear second-order term. Only quasilinear PDEs were shown to have connection with the BSDEs. This is achieved by introducing a Y, Z dependence in the dynamics of X .

In recent work [21], BSDEs were generalized by restricting the Z process to be a semimartingale. Precisely a *second-order backward stochastic differential equation* (2BSDE in short) has the X and Y equations, (2.12), (2.13) and (4.1), (4.2), and an additional equation

$$dZ(s) = a(s) ds + \Gamma(s)\sigma(s, X(s)) dW(s) \tag{4.3}$$

for some processes $a(\cdot)$ and $\Gamma(\cdot)$.

For simplicity, let us assume that $\sigma \equiv I_d$. Then, we rewrite the 2BSDE as

$$\begin{aligned} dX(s) &= \mu(s, X(s)) ds + dW(s), \\ dY(s) &= H(s, X(s), Y(s), Z(s), \Gamma(s)) ds + Z(s) \circ dX(s), \\ dZ(s) &= A(s) ds + \Gamma(s) dX(s), \\ Y(T) &= \varphi(X(T)), \quad X(t) = x, \end{aligned} \tag{4.4}$$

where H is a given function and the Fisk–Stratonovich integral \circ is given by

$$Z(s) \circ dW(s) = Z(s) dW(s) + \frac{1}{2} \text{trace}[\Gamma(s)] ds.$$

Below, we will give the precise function spaces in which we look for the solutions. However, to establish the connection between the PDEs let us formally assume that there is a solution and is given by $Y(s) = v(s, X(s))$. Then, by the Itô formula (using the definition of the Fisk–Stratonovich integral),

$$d[v(s, X(s))] = v_t(s, X(s)) ds + \nabla v(s, X(s)) \circ dX(s).$$

Comparing this to the dY equation in (4.4), again we conclude that $Z(s) = \nabla v(s, X(s))$ and

$$-v_t(s, X(s)) + H(s, X(s), Y(s), Z(s), \Gamma(s)) = 0 \quad \forall s \in [t, T]. \quad (4.5)$$

Now we apply the Itô rule to $\nabla v(s, X(s))$. The result is

$$d[\nabla v(s, X(s))] = [\nabla v_t(s, X(s)) + \mathcal{L}_t \nabla v(s, X(s))] ds + D^2 v(s, X(s)) dW(s).$$

We compare this to the dZ equation in (4.4) to conclude that $\Gamma(s) = D^2 v(s, X(s))$. We substitute these into (4.5) to obtain

$$\begin{aligned} & -v_t(s, X(s)) \\ & + H(s, X(s), v(s, X(s)), \nabla v(s, X(s)), D^2 v(s, X(s))) = 0 \quad \forall s \in [t, T]. \end{aligned}$$

If the X process has full support, then we conclude that

$$-v_t(t, x) + H(t, x, \nabla v(t, x), D^2 v(t, x)) = 0 \quad \forall (t, x) \in (-\infty, T] \times \mathbb{R}^d. \quad (4.6)$$

Notice that we did not make any assumptions on H . In particular, no parabolicity is assumed. Although, the above formal calculations do not require parabolicity, the existence of a solution to 2BSDE fails without parabolicity. In the remainder of this subsection, we will give the precise definitions of the functions spaces and the assumptions needed. Then, we will state the main representation result of [21] without proof.

In addition to usual local Lipschitz conditions, we assume there are constants $C \geq 0$ and $p_1 \in [0, 1]$ such that

$$|\mu(t, x)| \leq C(1 + |x|^{p_1}), \quad (t, x) \in (-\infty, T] \times \mathbb{R}^d.$$

DEFINITION 4.1. Let $(t, x) \in (-\infty, T) \times \mathbb{R}^d$ and $(Y(\cdot), Z(\cdot), \Gamma(\cdot), A(\cdot))$ be a quadruple of \mathcal{F} -progressively measurable processes on $[t, T]$ with values in \mathbb{R} , \mathbb{R}^d , \mathcal{S}^d and \mathbb{R}^d , respectively. Then we call (Y, Z, Γ, A) a solution to the second-order backward stochastic differential equation (2BSDE) corresponding to $(X^{t,x}, H, \varphi)$ if they solve (4.4).

Equations (4.4) can be viewed as a whole family of 2BSDEs indexed by $(t, x) \in [0, T) \times \mathbb{R}^d$. We have formally argued that the solution of these equations are related to the fully nonlinear partial differential equation (4.6).

Since Z is a semimartingale, the use of the Fisk–Stratonovich integral in (4.4) means no loss of generality, but it simplifies the notation in the PDE (4.6). Alternatively, (4.4) could be written in terms of the Itô integral as

$$dY(s) = \tilde{H}(s, X^{t,x}(s), Y(s), Z(s), \Gamma(s)) ds + Z(s) \cdot dX^{t,x}(s),$$

where (recalling that σ in the X equation is taken to be the identity)

$$\tilde{H}(t, x, y, z, \gamma) = H(t, x, y, z, \gamma) + \frac{1}{2} \text{trace}[\gamma].$$

In terms of \tilde{H} , the PDE (4.6) can be rewritten as

$$-v_t(t, x) + \tilde{H}(t, x, v(t, x), Dv(t, x), D^2v(t, x)) - \frac{1}{2} \Delta v(t, x) = 0.$$

Finally, notice that the form of the PDE (4.6) does not depend on the functions the dynamics of the X process. So, we could restrict our attention to the case where $\mu \equiv 0$ and $\sigma \equiv I_d$, the $d \times d$ identity matrix. But the freedom to choose the dynamics of X from a more general class of diffusions provides additional flexibility in the design of the Monte Carlo schemes discussed in Section 5.

From a solution of the PDE to a solution of the 2BSDE. Assume $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a continuous function such that

$$v_t, Dv, D^2v, \mathcal{L}Dv \text{ exist and are continuous on } [0, T) \times \mathbb{R}^d,$$

and v solves the PDE (4.6) with terminal condition (2.20). Then it follows directly from Itô's formula that for each pair $(t, x) \in (-\infty, T) \times \mathbb{R}^d$, the processes

$$Y(s) = v(s, X^{t,x}(s)), \quad s \in [t, T],$$

$$Z(s) = Dv(s, X^{t,x}(s)), \quad s \in [t, T],$$

$$\Gamma(s) = D^2v(s, X^{t,x}(s)), \quad s \in [t, T],$$

$$A(s) = \mathcal{L}Dv(s, X^{t,x}(s)), \quad s \in [t, T],$$

solve the 2BSDE corresponding to $(X^{t,x}, H, \varphi)$.

From a solution of the 2BSDE to a solution of the PDE. In all of this subsection, we assume that

$$H : (-\infty, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d \rightarrow \mathbb{R} \quad \text{and} \quad \varphi : \mathbb{R}^d \rightarrow \mathbb{R}$$

are continuous functions that satisfy the following Lipschitz and growth assumptions:

(A1) For every $N \geq 1$ there exists a constant F_N such that

$$|H(t, x, y, z, \gamma) - H(t, x, \tilde{y}, z, \gamma)| \leq F_N |y - \tilde{y}|$$

for all $t \in (-\infty, T]$, $x, z \in \mathbb{R}^d$, $y, \tilde{y} \in \mathbb{R}^1$, $\gamma \in \mathcal{S}^d$ with $\max\{|x|, |y|, |\tilde{y}|, |z|, |\gamma|\} \leq N$.

(A2) There exist constants F and $p_2 \geq 0$ such that

$$|H(t, x, y, z, \gamma)| \leq F(1 + |x|^{p_2} + |y| + |z|^{p_2} + |\gamma|^{p_2})$$

for all $(t, x, y, z, \gamma) \in (-\infty, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}^d$.

(A3) There exist constants G and $p_3 \geq 0$ such that

$$|\varphi(x)| \leq G(1 + |x|^{p_3}) \quad \text{for all } x \in \mathbb{R}^d.$$

Admissible strategies. We fix constants $p_4, p_5 \geq 0$ and denote for all $(t, x) \in (-\infty, T] \times \mathbb{R}^d$ and $m \geq 0$ by $\mathcal{A}_m^{t,x}$ the class of all processes of the form

$$Z(s) = z + \int_t^s A(r) dr + \int_t^s \Gamma(r) dX^{t,x}(r), \quad s \in [t, T],$$

where $z \in \mathbb{R}^d$, $(A(\cdot), \Gamma(\cdot)) \in \mathbb{R}^d \times \mathcal{S}^d$ progressively measurable processes satisfying

$$\max\{|Z(s)|, |A(s)|, |\Gamma(s)|\} \leq m(1 + |X^{t,x}(s)|^{p_4}) \quad \forall s \in [t, T], \quad (4.7)$$

and

$$\begin{aligned} |\Gamma(r) - \Gamma(s)| &\leq m(1 + |X^{t,x}(r)|^{p_5} + |X^{t,x}(s)|^{p_5}) \\ &\times (|r - s| + |X^{t,x}(r) - X^{t,x}(s)|) \quad \forall r, s \in [t, T]. \end{aligned} \quad (4.8)$$

Set $\mathcal{A}^{t,x} := \bigcup_{m \geq 0} \mathcal{A}_m^{t,x}$. It follows from the assumptions (A1) and (A2) on H and the condition (4.7) on Z that for all $y \in \mathbb{R}$ and $Z \in \mathcal{A}^{t,x}$, the forward SDE

$$dY(s) = f(s, X^{t,x}(s), Y(s), Z(s), \Gamma(s)) ds + Z(s) \circ dX^{t,x}(s), \quad s \in [t, T],$$

with $Y(t) = y$, has a unique strong solution $Y^{t,x,y,Z}(\cdot)$ (this can, for instance, be shown with the arguments in the proofs of Theorems 2.3, 2.4 and 3.1 in Chapter IV of Ikeda and Watanabe [41]).

Notice that $Z \in \mathcal{A}^{t,x}$ is a solution of the 2BSDE if $Y^{t,x,y,Z}(T) = \varphi(X^{t,x}(T))$.

We will show that solutions to 2BSDE in the class $\mathcal{A}^{t,x}$ has at most one solution.

1 *Uniqueness of 2BSDE in $\mathcal{A}^{t,x}$.* For our last assumption and the statement of Theorem 4.3, 1
 2 we need the following definition. 2

3
 4 DEFINITION 4.2. Let $q \geq 0$. 4

5 (1) We call a function $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ a viscosity solution with growth q of the 5
 6 PDE (4.6) with terminal condition (2.20) if v is a viscosity solution of (4.6) on $(-\infty, 6$
 7 $T) \times \mathbb{R}^d$ such that $v^*(T, x) = v_*(T, x) = g(x)$ for all $x \in \mathbb{R}^d$ and there exists a constant C 7
 8 such that $|v(t, x)| \leq C(1 + |x|^q)$ for all $(t, x) \in (-\infty, T] \times \mathbb{R}^d$. 8

9 (2) We say that the PDE (4.6) with terminal condition (2.20) has *comparison with 9*
 10 *growth q* if the following holds: 10

11 If $w : (-\infty, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ is lower semicontinuous and a viscosity supersolution of (4.6) 11
 12 on $(-\infty, T) \times \mathbb{R}^d$ and $u : (-\infty, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ upper semicontinuous and a viscosity sub- 12
 13 solution of (4.6) on $(-\infty, T) \times \mathbb{R}^d$ such that 13

$$14 \quad w(T, x) \geq g(x) \geq u(T, x) \quad \text{for all } x \in \mathbb{R}^d \quad 15$$

16 and there exists a constant $C \geq 0$ with 17

$$18 \quad w(t, x) \geq -C(1 + |x|^p) \quad \text{and} \quad u(t, x) \leq C(1 + |x|^p) \quad 19$$

$$20 \quad \text{for all } (t, x) \in (-\infty, T) \times \mathbb{R}^d, \quad 21$$

22 then $w \geq u$ on $(-\infty, T] \times \mathbb{R}^d$. 23

24
 25 With this definition our last assumption on H and φ is 25

26 (A4) The PDE (4.6) with terminal condition (2.20) has comparison with growth $p =$ 26
 27 $\max\{p_2, p_3, p_2 p_4, p_4 + 2p_1\}$. 27

28 The following result is proved in [20]. 28

29
 30 THEOREM 4.3 (Uniqueness of 2BSDE). Assume (A1)–(A4) and that H is parabo- 30
 31 lic (2.11). For $x_0 \in \mathbb{R}^d$ suppose that the 2BSDE corresponding to (X^{0,x_0}, H, φ) has a 31
 32 solution with $Z^{0,x_0} \in \mathcal{A}^{0,x_0}$. Then 32

33 (i) The associated PDE (4.6) with terminal condition (2.20) has a unique viscos- 33
 34 ity solution v with growth $p = \max\{p_2, p_3, p_2 p_4, p_4 + 2p_1\}$, and v is continuous on 34
 35 $[0, T] \times \mathbb{R}^d$. 35

36 (ii) For all $(t, x) \in [0, T) \times \mathbb{R}^d$, there exists exactly one solution $(Y^{t,x}, Z^{t,x}, \Gamma^{t,x}, A^{t,x})$ 36
 37 to the 2BSDE corresponding to $(X^{t,x}, H, \varphi)$ such that $Z^{t,x} \in \mathcal{A}^{t,x}$ and 37
 38 38

$$39 \quad Y^{t,x}(s) = v(s, X^{t,x}(s)), \quad s \in [t, T], \quad 40$$

41 where v is the unique continuous viscosity solution with growth p of (4.6) and (2.20). 42

43
 44 REMARK 4.4. 1. Under the hypothesis of the above theorem, the solution of the 2BSDE 44
 45 satisfies $Y^{t,x}(t) = v(t, x)$. Hence, $v(t, x)$ can be approximated by backward simulation 45

of the process $(Y^{t,x}(s))_{s \in [t, T]}$. If v is C^2 , it follows from Itô's lemma that $Z^{t,x}(s) = Dv(s, X^{t,x}(s))$, $s \in [t, T]$. Then $Dv(t, x)$ can also be approximated by backward simulation. Moreover, for v is C^3 , $\Gamma^{t,x}(s) = D^2v(s, X^{t,x}(s))$ can be simulated in this way. A formal discussion of a potential numerical scheme for the backward simulation of the processes $Y^{t,x}$, $Z^{t,x}$ and $\Gamma^{t,x}$ is provided in Section 5.3.

2. We have already shown that a classical solution v of (4.6) and (2.20) and its derivatives provide a solution of the 2BSDE.

3. The parabolicity assumption (2.11) is natural from the PDE viewpoint. If H is uniformly elliptic: there exists a constant $C > 0$ such that

$$H(t, x, y, z, \gamma - B) \geq H(t, x, y, z, \gamma) + C \operatorname{Tr}[B] \quad \forall B \geq 0.$$

Then the PDE (4.6) is uniformly parabolic, and there exist general results on existence, uniqueness and smoothness of solutions, see for instance, [45]. When H is linear in the γ variable (in particular, for the semi- and quasilinear equations discussed in Section 5.2), uniform ellipticity essentially guarantees existence, uniqueness and smoothness of solutions to the PDE (4.6) and (2.20); see for instance, Section 5.4 in [47].

4. Condition (A4) is an implicit assumption on the functions H and φ as we find it more convenient to assume comparison directly in the form (A4) instead of placing technical assumptions on H and φ which guarantee that the PDE (4.6) with terminal condition (2.20) has comparison. However, several comparison results for nonlinear PDEs are available in the literature; see for example, Crandall, Ishii and Lions [24], Fleming and Soner [36], Cabre and Caffarelli [18]. However, most results are stated for equations in bounded domains. For equations in the whole space, the critical issue is the interplay between the growth of solutions at infinity and the growth of the nonlinearity. We list some typical situations where comparison holds:

(a) *Comparison with growth 1.* Assume (A1)–(A3) and there exists a function $h : [0, \infty) \rightarrow [0, \infty)$ with $\lim_{x \rightarrow 0} h(x) = 0$ such that

$$|H(t, x, y, \alpha(x - \tilde{x}), A) - H(t, \tilde{x}, y, \alpha(x - \tilde{x}), B)| \leq h(\alpha|x - \tilde{x}|^2 + |x - \tilde{x}|)$$

for all (t, x, \tilde{x}, y) , $\alpha > 0$ and A, B satisfying

$$-\alpha \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \leq \begin{bmatrix} A & 0 \\ 0 & -B \end{bmatrix} \leq \alpha \begin{bmatrix} I & -I \\ -I & I \end{bmatrix}.$$

Then it follows from Theorem 8.2 in [24] that equations of the form (4.6), (2.20) have comparison with growth 0 if the domain is bounded. If the domain is unbounded, it follows from the modifications outlined in Section 5.D of Crandall et al. [24] that (4.6) and (2.20) have comparison with growth 1.

(b) For the dynamic programming equation (3.25) related to a stochastic optimal control problem, a comparison theorem for bounded solutions is given in [36], Section 5.9, Theorem V.9.1.

(c) Many techniques in dealing with unbounded solutions were developed by Ishii [42] for first-order equations (that is, when f is independent of γ). These techniques can be

extended to second-order equations. Some related results can be found in [4,5]. In [5], in addition to comparison results for PDEs, one can also find BSDEs based on jump Markov processes.

5. Monte Carlo methods

In this section we provide a formal discussion of the numerical implications of our representation results. We start by recalling some well-known facts in the linear case. We then review some results for the semilinear and the quasilinear cases. Then, we conclude with the fully nonlinear case related to Theorem 4.3.

5.1. The linear case

In this subsection we assume that the function H is of the form

$$H(t, x, y, z, \gamma) = -\alpha(t, x) - \beta(t, x)y - \mu(x) \cdot z - \frac{1}{2}a(x) : \gamma.$$

Then (4.6) is a linear parabolic equation and we discussed already that the Feynman–Kac representation has the form

$$v(t, x) = E \left[\int_t^T B_{t,s} \alpha(s, X^{t,x}(s)) ds + B_{t,T} g(X^{t,x}(T)) \right],$$

where

$$B_{t,s} := \exp \left(\int_t^s \beta(r, X^{t,x}(r)) dr \right)$$

(see, for instance, Theorem 5.7.6 in [44]). This representation suggests a numerical approximation of the function v by means of the so-called Monte Carlo method.

(i) Given J independent copies $\{X^j, 1 \leq j \leq J\}$ of the process $X^{t,x}$, set

$$\hat{v}^{(J)}(t, x) := \frac{1}{J} \sum_{j=1}^J \int_t^T B_{t,s}^j \alpha(s, X^j(s)) ds + B_{t,T}^j g(X^j(T)),$$

where $B_{t,s}^j := \exp(\int_t^s \beta(r, X^j(r)) dr)$. Then, it follows from the law of large numbers and the central limit theorem that

$$\begin{aligned} \hat{v}^{(J)}(t, x) &\rightarrow v(t, x) \quad \text{a.s. and} \\ \sqrt{J}(\hat{v}^{(J)}(t, x) - v(t, x)) &\rightarrow N(0, \rho) \quad \text{in distribution,} \end{aligned}$$

where ρ is the variance of the random variable $\int_t^T B_{t,s}\alpha(s, X^{t,x}(s)) ds + B_{t,T}g(X^{t,x}(T))$. Hence, $\hat{v}^{(J)}(t, x)$ is a consistent approximation of $v(t, x)$. Moreover, in contrast to finite differences or finite elements methods, the error estimate is of order $J^{-1/2}$, independent of the dimension d .

(ii) In practice, it is not possible to produce independent copies $\{X^j, 1 \leq j \leq J\}$ of the process $X^{t,x}$, except in trivial cases. In most cases, the above Monte Carlo approximation is performed by replacing the process $X^{t,x}$ by a suitable discrete-time approximation X^N with time step of order N^{-1} for which independent copies $\{X^{N,j}, 1 \leq j \leq J\}$ can be produced. The simplest discrete-time approximation is the following discrete Euler scheme: Set $X_t^N = x$ and for $1 \leq n \leq N$,

$$X_{t_n}^N = X_{t_{n-1}}^N + \mu(X_{t_{n-1}}^N)(t_n - t_{n-1}) + \sigma(X_{t_{n-1}}^N)(W_{t_n} - W_{t_{n-1}}),$$

where $t_n := t + n(T - t)/N$. We refer to [69] for a survey of the main results in this area.

5.2. The semilinear case

We next consider the case where H is given by

$$H(t, x, y, z, \gamma) = \varphi(t, x, y, z) - \mu(x) \cdot z - \frac{1}{2}a(x) : \gamma.$$

Then the PDE (4.6) is semilinear. We assume that the assumptions of Theorem 4.3 are satisfied. In view of the connection between Fisk–Stratonovich and Itô integration, the 2BSDE (4.4) reduces to an uncoupled forward–backward SDE (FBSDE) of the form

$$dY(s) = \varphi(s, X^{t,x}(s), Y(s), Z(s)) ds + Z(s) \cdot \sigma(X^{t,x}(s)) dW(s),$$

with terminal data $Y(T) = g(X^{t,x}(T))$ (compare to Peng [60], Pardoux and Peng [57]). For $N \geq 1$, we denote $t_n := t + n(T - t)/N$, $n = 0, \dots, N$, and we define the discrete-time approximation Y^N of Y by the backward scheme

$$Y_T^N := g(X_T^{t,x})$$

and, for $n = 1, \dots, N$,

$$Y_{t_{n-1}}^N := E[Y_{t_n}^N | X_{t_{n-1}}^{t,x}] - \varphi(t_{n-1}, X_{t_{n-1}}^{t,x}, Y_{t_{n-1}}^N, Z_{t_{n-1}}^N)(t_n - t_{n-1}), \quad (5.9)$$

$$Z_{t_{n-1}}^N := \frac{1}{t_n - t_{n-1}} (\sigma(X_{t_{n-1}}^{t,x}))^{-1} E[(W_{t_n} - W_{t_{n-1}}) Y_{t_n}^N | X_{t_{n-1}}^{t,x}]. \quad (5.10)$$

Then, we have

$$\limsup_{N \rightarrow \infty} \sqrt{N} |Y_t^N - v(t, x)| < \infty,$$

and in case that v is C^2 also,

$$\limsup_{N \rightarrow \infty} \sqrt{N} |Z_t^N - Dv(t, x)| < \infty,$$

see for instance, Bally and Pagès [3], Bouchard and Touzi [14]. The practical implementation of this backward scheme requires the computation of the conditional expectations appearing in (5.9) and (5.10). This suggests the use of a Monte Carlo approximation, as in the linear case. But at every time step, we need to compute conditional expectations based on J independent copies $\{X^j, 1 \leq j \leq J\}$ of the process $X^{t,x}$. Recently, several approaches to this problem have been developed. We refer to Bally and Pagès [3], Bouchard and Touzi [14], Lions and Regnier [49] and the references therein for the methodology and the analysis of such nonlinear Monte Carlo methods.

We refer to Chevance [22] and to a recent article by Delarue and Menozzi [26] for Monte Carlo simulations for the quasilinear case using forward–backward stochastic equations.

5.3. The fully nonlinear case

We now discuss the case of a general H as in the previous section. Let μ, σ be as in the dynamics (2.12)

$$\tilde{H}(t, x, y, z, \gamma) = H(t, x, y, z, \gamma) + \mu(t, x) \cdot z + \frac{1}{2} a(t, x) : \gamma.$$

Then, for all (t, x) the 2BSDE corresponding to $(X^{t,x}, H, \varphi)$ can be written as

$$\begin{aligned} dY(s) &= \tilde{H}(s, X^{t,x}(s), Y(s), Z(s), \Gamma(s)) ds \\ &\quad + Z(s) \cdot \sigma(s, X^{t,x}(s)) dW(s), \quad s \in [t, T], \\ dZ(s) &= A(s) ds + \Gamma(s) dX^{t,x}(s), \quad s \in [t, T], \\ Y(T) &= g(X^{t,x}(T)). \end{aligned} \tag{5.11}$$

We assume that the conditions of Theorem 4.3 hold true, so that the PDE (4.6) has a unique viscosity solution v with growth $p = \max\{p_2, p_3, p_2 p_4, p_4 + 2p_1\}$, and there exists a unique solution $(Y^{t,x}, Z^{t,x}, \Gamma^{t,x}, A^{t,x})$ to the 2BSDE (5.11) with $Z^{t,x} \in \mathcal{A}^{t,x}$.

Comparing with the backward scheme (5.9), (5.10) in the semilinear case, we suggest the following discrete-time approximation of the processes $Y^{t,x}, Z^{t,x}$ and $\Gamma^{t,x}$:

$$Y_T^N := g(X_T^{t,x}), \quad Z_T^N := Dg(X_T^{t,x}),$$

and, for $n = 1, \dots, N$,

$$\begin{aligned} Y_{t_{n-1}}^N &:= E[Y_{t_n}^N | X_{t_{n-1}}^{t,x}] - \varphi(t_{n-1}, X_{t_{n-1}}^{t,x}, Y_{t_{n-1}}^N, Z_{t_{n-1}}^N, \Gamma_{t_{n-1}}^N)(t_n - t_{n-1}) \\ Z_{t_{n-1}}^N &:= \frac{1}{t_n - t_{n-1}} (\sigma(X_{t_{n-1}}^{t,x}))^{-1} E[(W_{t_n} - W_{t_{n-1}}) Y_{t_n}^N | X_{t_{n-1}}^{t,x}], \end{aligned}$$

and

$$\Gamma_{t_{n-1}}^N := \frac{1}{t_n - t_{n-1}} E[Z_{t_n}^N (W_{t_n} - W_{t_{n-1}})' | X_{t_{n-1}}^{t,x}] \sigma(X_{t_{n-1}}^{t,x})^{-1}.$$

We expect that

$$(Y_t^N, Z_t^N, \Gamma_t^N) \rightarrow (v(t, x), Dv(t, x), D^2v(t, x)) \quad \text{as } N \rightarrow \infty.$$

However, proof of the above assertion is not yet available.

Acknowledgement

This work was partly supported by the Turkish Academy of Sciences.

References

- [1] L. Ambrosio and H.M. Soner, *Level set approach to mean curvature flow in arbitrary codimension*, J. Differential Geom. **43** (1996), 693–737.
- [2] O. Alvarez and A. Tourin, *Viscosity solutions of nonlinear integro-differential equations*, Ann. Inst. H. Poincaré **13** (1996), 293–317.
- [3] V. Bally, G. Pagès, *Error analysis of the quantization algorithm for obstacle problems*, Stochastic Process. Appl. **106** (1) (2003), 1–40.
- [4] G. Barles, *Solutions de viscosité des équations de Hamilton–Jacobi*, Math. Appl., Vol. 17, Springer-Verlag, Paris (1994).
- [5] G. Barles, R. Buckdahn and E. Pardoux, *Backward stochastic differential equations and integral-partial differential equations*, Stoch. Stoch. Rep. **60** (1/2) (1997), 57–83.
- [6] G. Barles, H.M. Soner and P.E. Souganidis, *Front propagation and phase field theory*, SIAM. J. Control Optim. **31** (2) (1993), 439–469.
- [7] M.T. Barlow and R.F. Bass, *Brownian motion and harmonic analysis on Sierpinski carpets*, J. Canad. Math. **51** (4) (1999), 673–744.
- [8] R. Bellman, *Dynamic Programming*, Princeton Univ. Press, Princeton, NJ (1957).
- [9] A. Bensoussan and J.-L. Lions, *Applications des inéquations variationnelles en contrôle stochastique*, Dunod, Paris (1978).
- [10] A. Bensoussan and J.-L. Lions, *Contrôle impulsionnel et inéquations quasi-variationnelles*, Dunod, Paris (1981).
- [11] J.M. Bismut, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Appl. **44** (1973), 384–404.
- [12] J.M. Bismut, *Contrôle des système linéaire quadratiques: Applications de l’intégrale stochastique*, Sémin. Probab. XII, Lecture. Notes in Math., Vol. 649, Springer-Verlag (1978), 180–264.
- [13] B. Bouchard, *Stochastic target problem with jump-diffusion process*, Stochastic Process. Appl. (2000).
- [14] B. Bouchard and N. Touzi, *Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations*, Stochastic Process. Appl. **111** (2004), 175–206.
- [15] K.A. Brakke, *The Motion of a Surface by Its Mean Curvature*, Princeton Univ. Press, Princeton, NJ (1978).
- [16] R. Buckdahn, P. Cardaliaguet and M. Quimcampois, *A representation formula for the mean curvature motion*, SIAM J. Math. Anal. **33** (4) (2002), 827–846.
- [17] R. Buckdahn and J. Ma, *Pathwise stochastic Taylor expansions and stochastic viscosity solutions for fully nonlinear stochastic PDEs*, Ann. Probab. **30** (2002), 1131–1171.
- [18] X. Cabre and L. Caffarelli, *Fully Nonlinear Elliptic Equations*, Amer. Math. Soc., Providence, RI (1995).

- [19] Y.-G. Chen, Y. Giga and S. Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*, J. Differential Geom. **33** (1991), 749–786.
- [20] P. Cheridito, H.M. Soner and N. Touzi, *The multi-dimensional super-replication problem under gamma constraints*, Ann. Inst. H. Poincaré Anal. Non Linéaire, to appear (2005).
- [21] P. Cheridito, H.M. Soner, N. Touzi and N. Victoir, *Second order backward stochastic differential equations and fully non-linear parabolic PDEs*, Preprint (2005).
- [22] D. Chevance, *Numerical Methods for Backward Stochastic Differential Equations*, Publ. Newton Inst., Cambridge University Press (1997).
- [23] M.G. Crandall, L.C. Evans and P.-L. Lions, *Some properties of viscosity solutions of Hamilton–Jacobi equations*, Trans. Amer. Math. Soc. **282** (1984), 487–502.
- [24] M.G. Crandall, H. Ishii and P.-L. Lions, *A user’s guide to viscosity solutions*, Bull. Amer. Math. Soc. **27** (1992), 1–67.
- [25] M.G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton–Jacobi equations*, Trans. Amer. Math. Soc. **277** (1984), 1–42.
- [26] F. Delarue and F. Menozzi, *A forward–backward stochastic algorithm for quasi-linear PDEs*, Preprint 932, University of Paris VI & VII (2004).
- [27] J. Douglas, J. Ma and P. Protter, *Numerical methods for forward–backward stochastic differential equations*, Ann. Appl. Probab. **6** (1996), 940–968.
- [28] E.B. Dynkin, *Superprocesses and partial differential equations*, Ann. Probab. **21** (3) (1993), 1185–1262.
- [29] E.B. Dynkin, *An Introduction to Branching Measure-Valued Processes*, CRM Monogr. Ser., Vol. 6, Amer. Math. Soc., Providence, RI (1994).
- [30] E.B. Dynkin, (1998), *A probabilistic approach to a nonlinear differential equation on a Riemannian manifold*, Teor. Veroyatnost. i Primenen. **42** (2) (1997), 336–341 (in Russian); English transl.: Theory Probab. Appl. no. 2, 289–294.
- [31] E.B. Dynkin, *Superdiffusions and Positive Solutions of Nonlinear Partial Differential Equations*, Univ. Lecture Ser., Vol. 34, Amer. Math. Soc., Providence, RI (2004).
- [32] N. El Karoui, S. Peng and M.C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1) (1997), 1–71.
- [33] L.C. Evans and P.E. Souganidis, *Differential games and representation formulas for solutions of Hamilton–Jacobi equations*, Indiana Univ. Math. J. **33** (1984), 773–797.
- [34] L.C. Evans and L. Spruck, *Motion of level sets by mean curvature*, J. Differential Geom. **33** (1991), 635–681.
- [35] R.P. Feynman, *Space–time approach to nonrelativistic quantum mechanics*, Rev. Modern Phys. **20** (1948), 367–387.
- [36] W.H. Fleming and H.M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Appl. Math., Vol. 25, Springer-Verlag, New York (1993).
- [37] W.H. Fleming and P.E. Souganidis, *On the existence of value function of two-player, zero-sum stochastic differential games*, Indiana Univ. Math. J. **38** (1989), 293–314.
- [38] M. Freidlin, *Markov Processes and Differential Equations: Asymptotic Problems*, Lectures Math. ETH Zürich, Birkhäuser, Basel (1996).
- [39] M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag, New York (1984).
- [40] T. Fujiwara and H. Kunita, *Stochastic differential equations of jump type and Levy processes in diffeomorphism groups*, J. Math. Kyoto Univ. **25** (1989), 71–106.
- [41] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd Edition, North-Holland (1989).
- [42] H. Ishii, *Uniqueness of unbounded viscosity solutions of Hamilton–Jacobi equations*, Indiana U. Math. J. **33** (1984), 721–748.
- [43] M. Kac, *On distributions of certain Wiener functional*, Trans. Amer. Math. Soc. (1949), 1–13.
- [44] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd Edition, Springer-Verlag (1991).
- [45] N.V. Krylov, *Nonlinear elliptic and Parabolic Partial Differential Equations of Second Order*, Math. Appl., Reider (1987).
- [46] N.V. Krylov, *Controlled Diffusion Processes*, Springer-Verlag, New York (1980).
- [47] O.A. Ladyzenskaya, V.A. Solonnikov and N.N. Uralseva, *Linear and Quasilinear Equations of Parabolic Type*, Amer. Math. Soc., Providence, RI (1967).

- [48] P.-L. Lions, *Neumann type boundary conditions for Hamilton–Jacobi equations*, Duke J. Math. **52** (1985), 793–820.
- [49] P.-L. Lions and H. Regnier, *Calcul du prix et des sensibilités d’une option américaine par une méthode de Monte Carlo*, Preprint (2001).
- [50] P.-L. Lions and P.E. Souganidis, *Fully nonlinear stochastic partial differential equations: Non-smooth equations and applications*, CRAS **327** (1998), 735–741.
- [51] P.-L. Lions and P.E. Souganidis, *Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations*, CRAS **331** (2000), 783–790.
- [52] P.-L. Lions and A.S. Sznitman, *Stochastic differential equations with reflecting boundary conditions*, Comm. Pure Appl. Math. **37** (1984), 511–537.
- [53] J. Ma, P. Protter, J. San Martin and S. Torres, *Numerical methods for backward stochastic differential equations*, Ann. Appl. Probab. **12** (2002), 302–316.
- [54] J. Ma, P. Protter and J. Yong, *Solving backward stochastic differential equations explicitly – A four step scheme*, Probab. Theory Related Fields **98** (1994), 339–359.
- [55] B. Oksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer-Verlag, New York (2005).
- [56] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14** (1990), 55–61.
- [57] E. Pardoux and S. Peng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, Lecture Notes in Control and Inform. Sci., Vol. 176, Springer-Verlag (1992), 200–217.
- [58] E. Pardoux and S. Peng, *Backward doubly stochastic differential equations and systems of quasilinear parabolic SPDEs*, Probab. Theory Related Fields **98** (1994), 209–227.
- [59] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim. **28** (1990), 966–979.
- [60] S. Peng, *Probabilistic interpretation for systems of quasilinear parabolic partial differential equations*, Stochastics **37** (1991), 61–74.
- [61] A. Sayah, *Equations d’Hamilton–Jacobi du premier ordre avec termes intégrés différentiels*. Partes I and II, Comm. Partial Differential Equations **16** (1991), 1057–1093.
- [62] A.V. Skorokhod, *Studies in the Theory of Random Processes*, Dover, New York (1980); 1st Edition (1965).
- [63] H.M. Soner, *Jump Markov processes and viscosity solutions*, IMA Vol. Math. Appl., Vol. 10, Springer-Verlag, New York (1986), 501–511.
- [64] H.M. Soner, *Motion of a set by the curvature of its boundary*, J. Differential Equations **101** (1993), 313–372.
- [65] H.M. Soner and N. Touzi, *Stochastic target problems, dynamic programming and viscosity solutions*, SIAM J. Control Optim. **41** (2002), 404–424.
- [66] H.M. Soner and N. Touzi, *Dynamic programming for stochastic target problems and geometric flows*, J. Eur. Math. Soc. **4** (2002), 201–236.
- [67] H.M. Soner and N. Touzi, *A stochastic representation for mean curvature type geometric flows*, Ann. Probab. **31** (2002), 1145–1165.
- [68] H.M. Soner and N. Touzi, *Level set characterization of stochastic target problems*, Comm. Partial Differential Equation **27** (9/10) (2002), 2031–2053.
- [69] D. Talay, *Probabilistic numerical methods for partial differential equations: Elements of analysis*, Probabilistic Models for Nonlinear Partial Differential Equations, D. Talay and L. Tubaro, eds, Lecture Notes in Math., Vol. 1627 (1996), 48–196.