

ADAPTIVE ROBUST CONTROL IN CONTINUOUS TIME*

THEERAWAT BHUDISAKSANG[†] AND ÁLVARO CARTEA[‡]

Abstract. We propose a continuous-time version of the adaptive robust methodology introduced in T. R. Bielecki et al. [*SIAM J. Control Optim.*, 57 (2019), pp. 925–946]. An agent solves a stochastic control problem where the underlying uncertainty follows a jump-diffusion process and the agent does not know the drift parameters of the process. The agent considers a set of alternative measures to make the control problem robust to model misspecification and employs a continuous-time estimator to learn the value of the unknown parameters to make the control problem adaptive to the arrival of new information. We use measurable selection theorems to prove the dynamic programming principle of the adaptive robust problem and show that the value function of the agent is characterized by a nonlinear partial differential equation. As an example, we derive the optimal adaptive robust strategy for an agent who acquires a large amount of shares in an order driven market and illustrate the financial performance of the execution strategy.

Key words. adaptive robust control, model uncertainty, stochastic control, time-consistency, dynamic programming, optimal acquisition, online learning, algorithmic trading

AMS subject classifications. 93E20, 93E35, 49L20, 60J05

DOI. 10.1137/20M1336680

1. Introduction. In classical stochastic control problems, agents look for the best policy to optimize a value function that depends on a stochastic process $X = (X_t)_{t \geq 0}$. An extensively studied problem is one in which the process X is a diffusion and it is the unique solution to the stochastic differential equation (SDE)

$$dX_t = b(X_t, \theta) dt + \sigma(X_t) dB_t^{\mathbb{P}},$$

where the functions b and σ are Lipschitz, $B^{\mathbb{P}} = (B_t^{\mathbb{P}})_{t \geq 0}$ is a standard Brownian motion under a reference measure \mathbb{P} , and $\theta \in \mathbb{R}^d$. A standard method to solve the agent's problem involves two key steps. First, show that the value function admits the dynamic programming principle (DPP). Second, characterize the value function as the solution to a nonlinear partial differential equation (PDE) and derive the agent's optimal decisions; see, e.g., Pham (2009).

In the classical approach, the agent assumes that the reference measure \mathbb{P} of the process X is known, but if the reference measure is incorrectly specified, the agent will find a suboptimal policy. A common approach to making the agent's control problem robust to model misspecification is to consider a set of alternative probability measures \mathcal{P} and then specify a criterion to choose the “optimal” measure from this alternative set. The criterion is generally one where the agent adopts conservative strategies.

We highlight two shortcomings of the classical robust stochastic control approach. One, it may lead to optimal policies that are too conservative. For example, if the

*Received by the editors May 7, 2020; accepted for publication (in revised form) July 19, 2021; published electronically October 21, 2021.

<https://doi.org/10.1137/20M1336680>

Funding: The first author acknowledges support of the Development and Promotion of Science and Technology Talents Project (DPST) of the Government of Thailand and by the Mathematical Institute, University of Oxford.

[†]Mathematical Institute, University of Oxford, Oxford OX1 2JD, United Kingdom (theerawat.bhudisaksang@maths.ox.ac.uk).

[‡]Mathematical Institute, University of Oxford, Oxford OX1 2JD, United Kingdom, and Oxford-Man Institute of Quantitative Finance, Oxford OX2 6ED, United Kingdom (Alvaro.Carte@maths.ox.ac.uk).

agent's level of confidence on each measure in the set \mathcal{P} is the same, the outcome is to choose the optimal policy under the probability measure that produces the worst result with respect to the performance criterion of the agent. Two, in the classical robust approach, the agent does not update her views on the set of measures \mathcal{P} as more information is revealed when the process X evolves in time; i.e., the set of alternative measures is fixed throughout the horizon of the control problem.

In recent work, Bielecki et al. (2019) propose an “adaptive robust” framework in which the agent incorporates the evolution of the underlying stochastic process as an ingredient in a control problem robust to model misspecification. They assume that the underlying dynamics follow a discrete-time homogeneous Markov process under a reference measure, construct the set of alternative measures \mathcal{P} via a composition of probability kernels to incorporate the arrival of new information, and use a measurable selection theorem to prove that the value function of the agent satisfies the DPP.

In this paper, we assume that X is a jump-diffusion process taking values in \mathbb{R}^n and it is the unique strong solution to the SDE

$$(1.1) \quad dX_t = b(X_t, \theta^*) dt + \sigma(X_t) dB_t^{\mathbb{P}_{\theta^*}} + \xi(X_{t-}) dL_t.$$

Here, $\theta^* \in \mathbb{R}^d$ is an unknown parameter, the process $B^{\mathbb{P}_{\theta^*}} = (B_t^{\mathbb{P}_{\theta^*}})_{t \geq 0}$ is a standard Brownian motion under the probability measure \mathbb{P}_{θ^*} , and $b : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^n$, $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^m$, $\xi : \mathbb{R}^n \times \mathbb{R}^r \rightarrow \mathbb{R}^n$. The jump component of X_t is the Lévy process $L = (L_t)_{t \geq 0}$ (independent of $B^{\mathbb{P}_{\theta^*}}$), which takes values in \mathbb{R}^r and with Lévy–Khintchine triplet $(0, 0, \nu)$ such that $\int_{\mathbb{R}^r} |z| \nu(dz) < \infty$ and $\nu(\mathbb{R}^r) < \infty$, and μ is an associated Poisson measure of the process L , where $\tilde{\mu}$ denotes the compensated Poisson measure. Note that the diffusion component σ , the jump component ξ , and the jump process L do not depend on the parameter θ^* .

At time t , the agent's estimate of the parameter θ^* is denoted by $\hat{\theta}_t$, and as time evolves and new information arrives, the agent updates the estimate $\hat{\theta}_t$. We model the processes X and $\hat{\theta}$ jointly, so we let $\tilde{X} = (\hat{\theta}, X)^\top$, which satisfies the SDE

$$(1.2) \quad d\tilde{X}_t = \mathbf{b}(t, \tilde{X}_t, \theta^*) dt + \boldsymbol{\sigma}(t, \tilde{X}_t) dB_t^{\mathbb{P}_{\theta^*}} + \boldsymbol{\xi}(t, \tilde{X}_{t-}) dL_t.$$

In addition, the agent considers the controlled diffusion process $Y = (Y_t)_{t \geq 0}$ valued in $\mathbb{R}^{\bar{d}}$, which satisfies

$$(1.3) \quad dY_t = \bar{b}(Y_t, \tilde{X}_t, \alpha_t) dt + \bar{\sigma}(Y_t, \tilde{X}_t, \alpha_t) d\tilde{X}_t,$$

where $\bar{b} : \mathbb{R}^{\bar{d}} \times \mathbb{R}^{n+d} \times \mathbb{R}^k \rightarrow \mathbb{R}^{\bar{d}}$, $\bar{\sigma} : \mathbb{R}^{\bar{d}} \times \mathbb{R}^{n+d} \times \mathbb{R}^k \rightarrow \mathbb{R}^{\bar{d}} \times \mathbb{R}^{n+d}$, and the agent controls the process $\alpha = (\alpha_t)_{t \geq 0}$, which takes values in \mathbb{R}^k .

Note that the second term on the right-hand side of (1.3) is $d\tilde{X}$ whose drift depends on the parameter θ^* (see (1.2)); thus, the drift of the process Y depends on θ^* . In this paper, we focus on the case where the agent does not have full knowledge of the drift term of the process X , while the agent has complete information about the volatility and jump terms of the process X .

In the adaptive robust framework, the agent considers the set of alternative measures $\mathcal{P}(t, x, G)$, where x represents the state of the underlying stochastic process X , and G is a function of the value of x and time t . The function G specifies the model uncertainty that stems from the estimation process $\hat{\theta}$. We remark that in general it is difficult (perhaps not possible) to construct the function G to be a confidence interval for the estimator process $\hat{\theta}$. However, when the process X is a geometric Brownian motion and the estimator process is the maximum likelihood estimator, we can construct the function G that specifies the confidence interval of the estimator in a statistical framework. In the remainder of this paper, it suffices to denote $\mathcal{P}(t, x, G)$

as $\mathcal{P}(t, x)$ because the function G is defined and fixed at the initial time of the agent's control problem. The agent's performance criterion is

$$(1.4) \quad J(t, x, y, \mathbb{P}, \alpha) := \mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) ds + g(Y_T) \right],$$

where \mathbb{P} is a probability measure, f and g are continuous functions $\tilde{X}_t = x$, $Y_t = y$, and we recall that α_t is the control process. The value function of the adaptive robust control problem is

$$(1.5) \quad \begin{aligned} w(t, x, y, \alpha) &:= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} J(t, x, y, \mathbb{P}, \alpha), \\ v(t, x, y) &:= \inf_{\alpha \in \mathcal{A}_0} w(t, x, y, \alpha), \end{aligned}$$

where \mathcal{A}_0 is the set of admissible control processes.

We extend the discrete-time adaptive robust control framework in Bielecki et al. (2019). In their work, the authors solve a similar problem to that in (1.5), where the underlying discrete-time stochastic process is ergodic and Markov, while here we assume that the underlying stochastic process is driven by a continuous-time jump-diffusion process. In the discrete-time setup of Bielecki et al., it is difficult to obtain a numerical solution of the value function v because in each back-propagation of v , one needs to compute the expectation of the criterion function J in (1.4); see Chen and Ludkovski (2021). In contrast, here we show that the value function v in (1.5) satisfies the DPP, characterize the value function as the solution of a nonlinear PDE, and prove the uniqueness of the viscosity solution. The PDE we derive is not standard because it contains a nonlinear operator that depends on the state variable x (see, e.g., Pham (1998)); this is discussed in detail in section 2. Thus, the continuous-time framework for the adaptive robust control is easier to solve (numerically) because there exists extensive literature on solving non-linear PDEs. As example of the adaptive robust strategy, we derive the solution for the optimal execution problem in electronic markets; see, e.g., Cartea, Jaimungal, and Penalva (2015).

Most of the literature that considers “parameter uncertainty” in diffusion-based models assumes that the drift parameter and the volatility parameter of the diffusion process lie in a known fixed interval, which is in contrast with our adaptive robust model where both the size of the uncertainty interval and the estimates of the parameters are updated as time evolves. When the unknown parameters lie in a fixed interval, performance criteria in the form of (1.4) are time-consistent because the agent does not update the estimates of the unknown parameters. For example, Epstein and Ji (2014) consider a utility maximization problem for a controlled diffusion process in which the drift and the volatility terms of the diffusion lie in a fixed interval, and Bannör et al. (2016) investigate parameter uncertainty in energy markets; see also Denis and Kervarec (2013), Biagini and Pinar (2017), and Bergen et al. (2018), all of whom assume that one or more parameters of the model lie in a fixed set and the estimates are not updated. We also mention other literature working on robust optimal stopping problem using a similar measurable selection technique; see Bayraktar and Yao (2014) and Ekren, Touzi, and Zhang (2014).

Moreover, in the work of Ismail and Pham (2019) and Pham, Wei, and Zhou (2021), an agent maximizes a mean-variance criterion (which is not time-consistent) in a one-step optimal asset allocation problem. In both works, the authors assume that the unknown parameters lie in a fixed interval and estimates of the parameters are not updated as time evolves.

There are also a number of papers that consider Knightian uncertainty (also referred to as model uncertainty) but do not include learning as proposed in this paper. In the Knightian uncertainty approach, the agent considers a set of alternative measures, equivalent to the agent's reference measure, to make decisions that are robust to model uncertainty. There is a penalty for choosing an alternative model—the penalty is a function of the entropy between the reference measure and the measure of the alternative model. Entropic penalties are convenient because they preserve time-consistency in optimization problems, so it is straightforward to show that the DPP holds and one can employ standard control techniques to obtain the optimal strategy robust to model misspecification. See, for example, the work of Hansen and Sargent (2011), Jaimungal and Sigloch (2012), Skiadas (2013), Cartea and Jaimungal (2017), and Cartea and Sánchez-Betancourt (2021), all of which assume model uncertainty with respect to the drift of a diffusion process, while the work of Cartea, Donnelly, and Jaimungal (2017) assumes model uncertainty with respect to the drift of a diffusion and to the intensity of the arrival of jumps.

To illustrate the performance of the adaptive robust approach we analyze a classical problem in finance. We derive the optimal acquisition strategy for an agent who purchases a large block of shares over a trading window, where the drift parameter of the stock price dynamics is not known by the agent. The performance of the adaptive robust strategy is compared with that of strategies in which the agent employs a wrong value of the drift parameter or employs a robust strategy. The robust strategy assumes that the drift parameter lies in a fixed interval and there is no learning as time evolves. Our results show that when the agent has enough time to learn the value of the unknown parameter, the adaptive robust strategy we develop in this paper performs better (lower average and lower variance of acquisition costs of block of shares) than when the agent employs a robust strategy or uses the incorrect parameter estimate. The superior performance of the robust adaptive strategy stems from learning the value of the drift parameter during the trading window.

The remainder of the paper is organized as follows. Section 2 develops the set of alternative measures used by the agent and introduces the continuous-time adapted robust control problem. We show that the agent's dynamic optimization problem in (1.5) admits the DPP, and we characterize the value function as a viscosity solution of a nonlinear PDE. In section 3, we present an application of the adaptive robust control problem to financial problems. Lastly, section 4 concludes and discusses future research directions.

2. Model. To streamline the presentation of the model, we provide a few definitions and other ingredients of the adaptive robust framework.

We define the terms “drift characteristic,” “volatility characteristic,” and “jump characteristic” of a stochastic process under a probability measure. Let $V = (V_t)_{t \geq 0}$ be a stochastic process in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We call \mathbb{P} a semimartingale measure of the process V if there exists a triple $(B^{\mathbb{P}}, C, \nu^{\mathbb{P}})$ such that $B^{\mathbb{P}}$ and C are the finite-variation part and the covariation process of the continuous local martingale of $V - \sum_{0 \leq s \leq \cdot} (\Delta V_s - h(\Delta V_s))$, where h is a bounded function that satisfies $h(x) = x$ in a neighborhood of the origin and $\nu^{\mathbb{P}}$ is the predictable compensator of μ^V , where

$$(2.1) \quad \mu^V(\omega, dt, dx) = \sum_{s \geq 0} \mathbf{1}_{\{\Delta V_s(\omega) \neq 0\}} \mathbf{1}_{(s, \Delta V_s(\omega))}(dt, dx).$$

Now, assume that for the triple $(B^{\mathbb{P}}, C, \nu^{\mathbb{P}})$, there exist the processes $(\gamma^{\mathbb{P}}, a, F_{\omega,s}^{\mathbb{P}})$ such that

$$B_t^{\mathbb{P}} = \int_0^t \gamma_s^{\mathbb{P}} ds, \quad C_t = \int_0^t a_s ds, \quad \text{and} \quad \nu^{\mathbb{P}}(\omega, ds, dx) = F_{\omega,s}^{\mathbb{P}}(dx) ds$$

for $d\mathbb{P} \times dt$ almost-all $(\omega, s) \in \Omega \times (0, T]$. Then, \mathbb{P} is a semimartingale measure of the process V with absolutely continuous characteristics, in which case we say that the processes $\gamma^{\mathbb{P}}$, a , and $F_{\omega,s}^{\mathbb{P}}$ are the drift characteristic, the volatility characteristic, and the jump characteristic, respectively, of the process V under the probability measure \mathbb{P} .

The remainder of this section proceeds as follows. Subsection 2.1 uses an explicit formula of the jump-diffusion process to motivate our definition of the alternative measures. Subsection 2.2 defines the set of alternative measures $\mathcal{P}(t, x)$ and formulates the adaptive robust framework. Subsection 2.3 proves the measurability of the value function. And subsections 2.4 and 2.5 show that the adaptive robust control problem in continuous time is time-consistent, prove the DPP of the value function, and derive the nonlinear PDE it satisfies.

2.1. Set of alternative measures. Recall that the process X follows (1.1) and that the agent does not know the value of the parameter $\theta^* \in \mathbb{R}^d$. As time evolves, the agent observes the realization of the process X and updates $\hat{\theta}$ (i.e., the estimate of the parameter θ^*), which satisfies the SDE

$$(2.2) \quad d\hat{\theta}_t = \beta_t \left[\tilde{b}(X_t, \hat{\theta}_t, \theta^*) dt + \tilde{\sigma}(X_t, \hat{\theta}_t) dB_t^{\mathbb{P}_{\theta^*}} + \tilde{\xi}(X_{t-}, \hat{\theta}_t) dL_t \right].$$

Here, the functions \tilde{b} and $\tilde{\sigma}$ specify the drift and the volatility of the estimator process $\hat{\theta}$, and $\beta > 0$ is the learning rate that depends on t . The function \tilde{b} satisfies

$$(2.3) \quad \tilde{b}(x, \hat{\theta}, \theta_1) - \tilde{b}(x, \hat{\theta}, \theta_2) = \frac{(b(x, \theta_1) - b(x, \theta_2)) \tilde{\sigma}(x, \hat{\theta})}{\sigma(x)}$$

for all $x, \hat{\theta}$, and all $\theta_1, \theta_2 \in \mathbb{R}^d$.

Therefore, under each probability measure \mathbb{P}_{θ} for $\theta \in \mathbb{R}^d$, the agent considers the augmented process $\tilde{X} = (X, \hat{\theta})^{\top} \in \mathbb{R}^{n+d}$, where X follows (1.1) and $\hat{\theta}$ follows (2.2), so \tilde{X} satisfies

$$(2.4) \quad d\tilde{X}_t = \mathbf{b}(t, \tilde{X}_t, \theta) dt + \boldsymbol{\sigma}(t, \tilde{X}_t) dB_t^{\mathbb{P}_{\theta}} + \boldsymbol{\xi}(t, \tilde{X}_{t-}) dL_t,$$

where $\mathbf{b}(t, \tilde{x}, \theta) := [b(x, \theta), \beta_t \tilde{b}(x, \hat{\theta}, \theta)]^{\top}$, $\boldsymbol{\sigma}(t, \tilde{x}) := [\sigma(x), \beta_t \tilde{\sigma}(x, \hat{\theta})]^{\top}$, $\boldsymbol{\xi}(t, \tilde{x}) := [\xi(x), \beta_t \tilde{\xi}(x, \hat{\theta})]^{\top}$, and $\tilde{x} = (x, \hat{\theta})$. The coefficients in (2.4) depend on time because the learning rate β_t is time-dependent.

When X follows an arithmetic or a geometric Brownian motion, the maximum likelihood estimators for the drift of X have the form in (2.2). Also, if the process X is ergodic, the stochastic gradient descent in continuous time is also as in (2.2); see Sirignano and Spiliopoulos (2017) for diffusions and Bhudisaksang and Cartea (2021) for jump-diffusions. However, we note that in general, the estimator of the parameter in the drift term does not have the form in (2.2).

Let $G : [0, T] \times \mathbb{R}^{n+d} \rightarrow 2^{\mathbb{R}^d}$ be an exogenous function that specifies the model uncertainty of the estimation process. Assume that at time t the agent's estimate of the parameter θ^* is the progressively measurable process $\tilde{\theta}_t \in G(t, \tilde{X}_t)$. Then, by

Girsanov's theorem, there is a probability measure $\mathbb{P}_{\tilde{\theta}}$ such that $dX_t = b(X_t, \tilde{\theta}_t) dt + \sigma(X_t) dB_t^{\mathbb{P}_{\tilde{\theta}}} + \xi(X_{t-}) dL_t$, and we write the dynamics of the joint process \tilde{X} , under the probability measure $\mathbb{P}_{\tilde{\theta}}$, as

$$d\tilde{X}_t = \mathbf{b}(t, \tilde{X}_t, \tilde{\theta}_t) dt + \boldsymbol{\sigma}(t, \tilde{X}_t) dB_t^{\mathbb{P}_{\tilde{\theta}}} + \boldsymbol{\xi}(t, \tilde{X}_{t-}) dL_t.$$

This motivates our definition of the set of alternative probability measures $\mathcal{P}(t, x, G)$ that contains the measure \mathbb{P} whose drift characteristic, volatility characteristic, and jump characteristic of the process X are

$$\mathbf{b}(s, \tilde{X}_s, \tilde{\theta}_s) + \int_{\mathbb{R}^r} (\hat{h}(\boldsymbol{\xi}(t, \tilde{X}_{s-}) z) - \boldsymbol{\xi}(s, \tilde{X}_{s-}) h(z)) \nu(dz),$$

$\boldsymbol{\sigma}(s, \tilde{X}_s) \boldsymbol{\sigma}(s, \tilde{X}_s)^\top$, and $F_{\omega, t}^{\mathbb{P}}(dx) := \int_{\mathbb{R}^r} 1_{\{\boldsymbol{\xi}(s, \tilde{X}_{s-}(\omega)) z \in dx\}} \nu(dz)$, respectively. The construction of these probability measures is similar to that of a weak formulation of the control problem in Karoui and Tan (2013). We provide a formal definition of the set of alternative measures $\mathcal{P}(t, x, G)$ after the following subsection.

2.2. Model setup. In this subsection, we present the adaptive robust control problem in continuous time and define the performance criterion of the agent—we employ a weak formulation of the control problem to simplify some proofs. Denote by $\mathcal{B}(Y)$ a Borel σ -field of the Polish space Y , let $\Omega = D([0, T], \mathbb{R}^{n+d})$ be the space of all càdlàg paths $\omega = (\omega_t)_{t \geq 0}$, $\mathcal{F} = \mathcal{B}(\Omega)$, and let \tilde{X} be a canonical process, i.e., $\tilde{X}_t(\omega) = \omega_t$. Note that for any probability measure \mathbb{P} and $A \in \mathcal{F}$, we have that $\mathbb{P}(\tilde{X} \in A) = \mathbb{P}(A)$. Denote by $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ the raw filtration generated by the canonical process X . Let $\mathfrak{P}(\Omega)$ denote the set of probability measures on Ω , and denote the set of semimartingale probability measures with absolutely continuous characteristics of the process \tilde{X} by $\mathfrak{P}_{sem}^{ac}(\Omega)$. Next, we define the measure kernel

$$\bar{\mathbb{P}}(\omega, t, dx) := \int_{\mathbb{R}^r} 1_{\{\boldsymbol{\xi}(t, \tilde{X}_{t-}(\omega)) z \in dx\}} \nu(dz).$$

Note that the kernel $\bar{\mathbb{P}}$ is a Borel measurable function from $\Omega \times [0, T] \times \mathcal{B}(\mathbb{R}^{n+d})$ to \mathbb{R} .

Now, we define the learnable property.

DEFINITION 2.1. Let $t \in (0, T]$. A semimartingale probability measure $\mathbb{P} \in \mathfrak{P}_{sem}^{ac}(\Omega)$ is called learnable by G on $(t, T]$ if the drift, volatility, and jump characteristics of \mathbb{P} are

$$\gamma_s^{\mathbb{P}} \in \mathbf{b}^*(s, \tilde{X}_s), \quad a_s = \boldsymbol{\sigma}(s, \tilde{X}_s) \boldsymbol{\sigma}(s, \tilde{X}_s)^\top, \quad \text{and} \quad F_{\omega, s}^{\mathbb{P}}(dx) = \bar{\mathbb{P}}(\omega, s, dx)$$

for $d\mathbb{P} \times dt$ almost-all $(\omega, s) \in \Omega \times (t, T]$, where

$$\mathbf{b}^*(s, \tilde{X}_s) = \left\{ \mathbf{b}(s, \tilde{X}_s, \theta) + \int_{\mathbb{R}^r} (\hat{h}(\boldsymbol{\xi}(s, \tilde{X}_{s-}) z) - \boldsymbol{\xi}(s, \tilde{X}_{s-}) h(z)) \nu(dz) \mid \theta \in G(s, \tilde{X}_s) \right\}.$$

Here, the functions $h : \mathbb{R}^r \rightarrow \mathbb{R}^r$ and $\hat{h} : \mathbb{R}^{n+d} \rightarrow \mathbb{R}^{n+d}$ are bounded and satisfy $h(x) = x$, $\hat{h}(x) = x$ when $|x| < 1$ and 0 otherwise.

The mapping $\mathbf{b}^* : [0, T] \times \mathbb{R}^{n+d} \rightarrow 2^{\mathbb{R}^d}$ is continuous with respect to the Hausdorff metric by a standard calculation; see Definition B.2 in the appendix.

DEFINITION 2.2 (set of alternative measures). *The set of alternative measures $\mathcal{P}(t, x)$ consists of all probability measures \mathbb{P} that satisfy the following two properties:*

- (i) $\mathbb{P} \in \mathfrak{P}_{sem}^{ac}(\Omega)$ and $\mathbb{P}(\tilde{X}_t = x) = 1$.
- (ii) \mathbb{P} is learnable by G on the interval $(t, T]$.

The following assumption is crucial to prove measurability of the set of alternative probability measures. This assumption is standard in the literature; see, e.g., Assumption 4.1 in Nutz and van Handel (2013).

Assumption 2.3. For every $t \in \mathbb{R}_+$, we have that

$$\{(s, \omega, \gamma, \rho, F) \in [t, T] \times \Omega \times \mathbb{R}^{n+d} \times \mathbb{R}^{(n+d) \times (n+d)} \times \mathfrak{P}(\mathbb{R}^{n+d}) : \gamma \in \mathbf{b}^*(s, \omega_s), \rho = a_s(\omega_s), F = \bar{\mathbb{P}}(\omega, s, \cdot)\} \in \mathcal{B}([t, T]) \otimes \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^{n+d}) \otimes \mathcal{B}(\mathbb{R}^{(n+d) \times (n+d)}) \otimes \mathcal{B}(\mathfrak{P}(\mathbb{R}^{n+d})).$$

We denote the concatenation of a path by $\omega \otimes_t \tilde{\omega} := w_s \mathbf{1}_{s \leq t} + (w_t + \tilde{w}_s - \tilde{w}_t) \mathbf{1}_{s > t}$. Let $\mathbb{E}^{\mathbb{P}^{\tau, \tilde{\omega}}}[\xi^{\tau, \tilde{\omega}}] := \mathbb{E}^{\mathbb{P}^{\tilde{\omega}}}[\xi] = \mathbb{E}^{\mathbb{P}}[\xi | \mathcal{F}_\tau](\tilde{\omega})$, where $\xi^{\tau, \tilde{\omega}}(\omega) := \xi(\tilde{\omega} \otimes_{\tau(\tilde{\omega})} \omega)$ and $\{\mathbb{P}_\tau^{\tilde{\omega}}\}_{\tilde{\omega} \in \Omega}$ is a regular conditional probability distribution given \mathcal{F}_τ .

When the process X in (1.2) is a diffusion with no jump component, the set of alternative measures $\mathcal{P}(t, x)$ contains a semimartingale probability measure \mathbb{P} such that the drift characteristic $\gamma_s^\mathbb{P}$ of \mathbb{P} is in the set $\mathbf{b}^*(s, \tilde{X}_s) = \{\mathbf{b}(s, \tilde{X}_s, \theta) \mid \theta \in G(s, \tilde{X}_s)\}$.

Next, we define the agent's adaptive robust control problem.

The set of admissible controls, denoted by \mathcal{A}_0 , consists of all progressively measurable processes that take values in a compact set $A \in \mathbb{R}^k$. Recall from (1.4) that the performance criterion of the agent is $J : [0, T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\bar{d}} \times \mathfrak{P}(\Omega) \times \mathcal{A}_0 \rightarrow \mathbb{R}$ and is given by

$$(2.5) \quad J(t, x, y, \mathbb{P}, \alpha) := \mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) ds + g(Y_T) \right],$$

where α is a control in the admissible set \mathcal{A}_0 , $\tilde{X}_t = x$, $Y_t = y$, and f and g are continuous functions in all variables.

Assume the controls α and $\tilde{\alpha}$ are in the set \mathcal{A}_0 . We define the distance metric in the space \mathcal{A}_0 as follows:

$$\Delta(\alpha, \tilde{\alpha}) := \mathbb{E}^{\mathbb{P}_{\hat{\theta}}} \left[\int_0^T |\alpha_t - \tilde{\alpha}_t|^2 dt \right],$$

where $\mathbb{P}_{\hat{\theta}}$ is the semimartingale measure for which the drift characteristic of \tilde{X} is

$$\mathbf{b}(s, \tilde{X}_s, \hat{\theta}_s) + \int_{\mathbb{R}} \left(\hat{h}(\xi(s, \tilde{X}_{s-}) z) - \xi(s, \tilde{X}_{s-}) h(z) \right) \nu(dz).$$

Thus, we denote by $\mathcal{B}_{\mathcal{A}_0}$ the set of a Borel measurable set of the set \mathcal{A}_0 generated by the distance metric Δ . By elementary calculations, the set \mathcal{A}_0 is a Polish space. We denote the space of all probability measures on (Ω, \mathcal{F}) by $\mathfrak{P}(\Omega)$ and equip it with the weak*-topology.

The agent's value function in the adaptive robust framework is

$$(2.6) \quad \begin{aligned} w(t, x, y, \alpha) &:= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} J(t, x, y, \mathbb{P}, \alpha), \\ v(t, x, y) &:= \inf_{\alpha \in \mathcal{A}_0} w(t, x, y, \alpha). \end{aligned}$$

Assumption 2.4. The functions J , w , v in (2.6) are finite.

Assumption 2.4 is satisfied if f and g are bounded functions and if all other assumptions that guarantee that Y is integrable on $[0, T] \times \Omega$ are also satisfied. In what follows, we assume that Assumptions 2.3 and 2.4 hold.

In summary, the agent chooses a control process α to minimize the performance criterion J , but there is uncertainty about the estimate of the parameter $\hat{\theta}$. Therefore, the agent considers the worst case scenario by choosing a measure from the alternative set $\mathcal{P}(t, x)$ that depends on the state of the underlying process and the estimator process. On the other hand, if the agent chooses to maximize the performance criterion, the control problem is as in (2.6) but of inf-sup type (instead of sup-inf).

2.3. Measurability of functions and set of probability measures. Here, we show that the set $\{(\omega, t, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\}$ is Borel measurable. This allows us to use a measurable selection theorem on the supremum problem in (2.6). Then, we check the measurability of the performance criterion J , the function w , and the value function v . The measurability of these functions allows the use of measurable selection theorems for both the infimum and supremum problems in (2.6).

LEMMA 2.5. *The set $\{(\omega, t, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\}$ is a Borel measurable set.*

For a proof, see Appendix A.1.

The class of shifted control processes is constructed from the concatenation of a path as follows: For $\alpha \in \mathcal{A}_0$ and $(t, \bar{\omega}) \in [0, T] \times \Omega$, we set

$$\alpha_s^{t, \bar{\omega}}(\omega) = \alpha_s(\bar{\omega} \otimes_t \omega), \quad (s, \bar{\omega}) \in [0, T] \times \Omega.$$

For a stopping time τ and for any $\alpha \in \mathcal{A}_0$, we denote the mapping

$$\alpha^\tau : (\Omega, \mathcal{F}) \rightarrow (\mathcal{A}_0, \mathcal{B}_{\mathcal{A}_0}), \quad \bar{\omega} \mapsto \alpha^{\tau(\bar{\omega}), \bar{\omega}}.$$

We employ the shifted control processes to show the pseudo-Markov property and DPP of the problem in (2.6). Next, we show the pseudo-Markov property of the process \tilde{X} .

LEMMA 2.6 (pseudo-Markov property). *Let τ be a stopping time in $[0, T]$. We have that*

$$(2.7) \quad J(\tau, \tilde{X}_\tau, Y_\tau, \mathbb{P}^\tau, \alpha^\tau) = \mathbb{E}^\mathbb{P} \left[\int_\tau^T f(s, Y_s, \alpha_s) ds + g(Y_T) \mid \mathcal{F}_\tau \right], \quad \mathbb{P}\text{-a.s.}$$

For a proof, see Appendix A.2.

Next, we check the measurability of the functions J , w , v in (2.6), and we prove a lemma that allows us to use measurable selections for both the supremum and the infimum problems in (2.6). First, we provide the following definition. For a Borel space H , we denote by \mathcal{L}_H the smallest σ -field containing Borel subsets of H and closed under the Souslin operation.¹ See also the definitions of analytic function and universally measurable function in Appendices B.3 and B.4.

Below, Lemma 2.7 shows that there exists a universally measurable selector for the infimum problem in (2.6). In general, this is not true. If the function w is only upper semianalytic, there is no guarantee that there exists a universally measurable ϵ -optimal

¹See Definition 7.15 in Bertsekas and Tsitsiklis (1996).

selector for the infimum problem; see Nowak (2010). The function w is $\mathcal{L}_X \otimes B_Y$ -measurable because we assume that the performance criterion J is continuous in the variable y . Therefore, there exists a universally measurable ϵ -optimal selector for the infimum problem.

Now, we show the regularity of the performance criterion J with respect to all its variables and the regularity of the value function v with respect to the state (t, x, y) .

LEMMA 2.7 (regularity of the value function). *Let J, w, v be as in (2.6). The function J is Borel measurable from $[0, T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\bar{d}} \times \mathfrak{P}(\Omega) \times \mathcal{A}_0$ to \mathbb{R} and continuous with respect to t, x, y, α . The function w is upper semianalytic, and the value function v is universally measurable from $[0, T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\bar{d}}$ to \mathbb{R} .*

For a proof, see Appendix A.3.

2.4. Time-consistency and the DPP of the adaptive robust problem.

Here, we show that the set of probability measures $\mathcal{P}(t, x)$ has the stability property under conditioning and the stability property under concatenation, both of which we employ to prove the time-consistency property of problem (2.6).

First, we state two lemmas that we need for the proof of the time-consistency property for the problem in (2.6). The proofs of the following two lemmas follow directly from Lemma 4.28 in Hollender (2016).

LEMMA 2.8 (stability under conditioning). *Let $\mathbb{P} \in \mathcal{P}(t, x)$, and let τ denote a stopping time taking values in $[t, T]$. For each $\bar{\omega} \in \Omega$, the probability measure $\mathbb{P}^{\tau, \bar{\omega}}$ is in the set $\mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))$ \mathbb{P} -a.s.*

LEMMA 2.9 (stability under concatenation). *Let $\mathbb{P} \in \mathcal{P}(t, x)$, and let τ denote a stopping time taking values in $[t, T]$. Let ν be a Borel measurable kernel such that $\nu : \Omega \rightarrow \mathfrak{P}(\Omega)$ and $\nu(\bar{\omega}) \in \mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))$. Then,*

$$\bar{\mathbb{P}}(A) = \int \int 1_{\{\bar{\omega} \otimes_{\tau(\bar{\omega})} \omega \in A\}} \nu(d\omega; \bar{\omega}) \mathbb{P}(d\bar{\omega})$$

is a probability measure, and it is in the set $\mathcal{P}(t, x)$.

Now, we prove the time-consistency property of problem (2.6). This is our first main result.

THEOREM 2.10 (time-consistency property). *Let $(t, x, y) \in [0, T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\bar{d}}$, and let α be a control process in the admissible set \mathcal{A}_0 . Let τ be a stopping time that takes values in the interval $[t, T]$. The following equation holds:*

$$\begin{aligned} & \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) ds + g(Y_T) \right] \\ &= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^{\tau} f(s, Y_s, \alpha_s) ds + \sup_{\mathbb{Q} \in \mathcal{P}(\tau, \tilde{X}_{\tau})} \mathbb{E}^{\mathbb{Q}} \left[\int_{\tau}^T f(s, Y_s, \alpha_s^{\tau}) ds + g(Y_T) \right] \right]. \end{aligned}$$

Proof. Use Lemmas 2.5, 2.8, and 2.9 and the arguments in Nutz and van Handel (2013) to show time-consistency of the stochastic control problem (2.6). For each $\mathbb{P} \in \mathcal{P}(t, x)$, by conditional expectations, one obtains

$$\begin{aligned}
& \mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) ds + g(Y_T) \right] \\
&= \mathbb{E}^{\mathbb{P}} \left[\int_t^{\tau} f(s, Y_s, \alpha_s) ds + \mathbb{E}^{\mathbb{P}} \left[\int_{\tau}^T f(s, Y_s, \alpha_s) ds + g(Y_T) \mid \mathcal{F}_{\tau} \right] \right] \\
(2.8) \quad &= \mathbb{E}^{\mathbb{P}} \left[\int_t^{\tau} f(s, Y_s, \alpha_s) ds + \mathbb{E}^{\mathbb{P}^{\tau(\bar{\omega}), \bar{\omega}}} \left[\int_{\tau}^T f(s, Y_s, \alpha_s^{\tau}) ds + g(Y_T) \right] \right] \\
&\leq \mathbb{E}^{\mathbb{P}} \left[\int_t^{\tau} f(s, Y_s, \alpha_s) ds + \sup_{\mathbb{Q} \in \mathcal{P}(\tau, \tilde{X}_{\tau})} \mathbb{E}^{\mathbb{Q}} \left[\int_{\tau}^T f(s, Y_s, \alpha_s^{\tau}) ds + g(Y_T) \right] \right],
\end{aligned}$$

where the last inequality follows because $\mathbb{P}^{\tau, \bar{\omega}}$ is in the set $\mathcal{P}(\tau, \tilde{X}_{\tau})$ for all $\bar{\omega}$ \mathbb{P} -a.s.; see Lemma 2.8. Now, take the supremum on both sides of (2.8) and write

$$\begin{aligned}
(2.9) \quad & \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) ds + g(Y_T) \right] \\
&\leq \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^{\tau} f(s, Y_s, \alpha_s) ds + \sup_{\mathbb{Q} \in \mathcal{P}(\tau, \tilde{X}_{\tau})} \mathbb{E}^{\mathbb{Q}} \left[\int_{\tau}^T f(s, Y_s, \alpha_s^{\tau}) ds + g(Y_T) \right] \right].
\end{aligned}$$

Next, we prove the reverse inequality of (2.9). For any stopping time τ and $\alpha \in \mathcal{A}_0$, write the function w as

$$\begin{aligned}
& w \left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha^{\tau(\bar{\omega}), \bar{\omega}} \right) \\
&= \sup_{\mathbb{Q} \in \mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))} J \left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \mathbb{Q}, \alpha^{\tau(\bar{\omega}), \bar{\omega}} \right) \\
&= \sup_{\mathbb{Q} \in \mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))} \hat{J}(\bar{\omega}, \mathbb{Q}).
\end{aligned}$$

From the proof of Lemma 3.1 in Neufeld and Nutz (2014), the function \hat{J} is Borel measurable. From Lemma 2.5, the set $\{(\bar{\omega}, \mathbb{Q}) \mid \mathbb{Q} \in \mathcal{P}(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}))\}$ is analytic. Therefore, by a measurable selection theorem, there exists a universally measurable ν^{ϵ} from Ω to $\mathfrak{P}(\Omega)$ such that

$$w \left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha^{\tau(\bar{\omega}), \bar{\omega}} \right) < \epsilon + J \left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \nu^{\epsilon}(\bar{\omega}), \alpha^{\tau(\bar{\omega}), \bar{\omega}} \right)$$

for all $\bar{\omega} \in \Omega$. From Lemma 7.27 in Bertsekas and Shreve (1996), we have that for any probability measure $\mathbb{P} \in \mathcal{P}(t, x)$, there is a Borel measurable function $\nu^{\epsilon, \mathbb{P}}$ such that $\nu^{\epsilon, \mathbb{P}}(\bar{\omega}) = \nu^{\epsilon}(\bar{\omega})$ for \mathbb{P} -a.s. Then,

$$\begin{aligned}
& w \left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha^{\tau(\bar{\omega}), \bar{\omega}} \right) \\
&< \epsilon + J \left(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \nu^{\epsilon, \mathbb{P}}(\bar{\omega}), \alpha^{\tau(\bar{\omega}), \bar{\omega}} \right), \quad \mathbb{P}\text{-a.s.}
\end{aligned}$$

Therefore, we have the following inequality:

$$\begin{aligned}
 & \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + \sup_{\mathbb{Q} \in \mathcal{P}(\tau, \tilde{X}_\tau)} \mathbb{E}^{\mathbb{Q}} \left[\int_\tau^T f(s, Y_s, \alpha_s^\tau) ds + g(Y_T) \right] \right] \\
 &= \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + w(\tau, \tilde{X}_\tau, Y_\tau, \alpha^\tau) \right] \\
 &< \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + J(\tau, \tilde{X}_\tau, Y_\tau, \nu^{\epsilon, \mathbb{P}}, \alpha^\tau) \right] + \epsilon \\
 &= \mathbb{E}^{\mathbb{P}^{\nu^\epsilon}} \left[\int_t^T f(s, Y_s, \alpha_s) ds \right] + \epsilon,
 \end{aligned}$$

where $\mathbb{P}^{\nu^\epsilon}(A) := \int \int 1_{\{\bar{\omega} \otimes_{\tau(\bar{\omega})} \omega \in A\}} \nu^{\epsilon, \mathbb{P}}(d\omega; \bar{\omega}) \mathbb{P}(d\bar{\omega})$ is in the set $\mathcal{P}(t, x)$; see Lemma 2.9. \square

Our proof of time-consistency is standard, but the result of Theorem 2.10 is not immediately available in the literature. For example, Nutz and van Handel (2013) and Bayraktar and Yao (2014) look at the time-consistency property of problems in robust optimal control and robust stopping; however, their results cannot be applied directly in our framework to show time-consistency of the problem we study.

Next, we prove the DPP and write the value function recursively. The proof relies on the time-consistency property of the adaptive robust problem in (2.6), which we showed in Theorem 2.10. We use the measurable selection theorem in Soner and Touzi (2002) because the set of admissible controls \mathcal{A}_0 is a separable metric space.

THEOREM 2.11 (DPP). *Let $(t, x, y) \in [0, T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\bar{d}}$, and let τ be a stopping time taking values in $[t, T]$, $\tilde{X}_t = x$, and $Y_t = y$. The value function of the adaptive robust problem in (2.6) satisfies*

$$(2.10) \quad v(t, x, y) = \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + v(\tau, \tilde{X}_\tau, Y_\tau) \right].$$

Proof. Recall that $w(t, x, y, \alpha) := \sup_{\mathbb{P} \in \mathcal{P}(t, x)} J(t, x, y, \mathbb{P}, \alpha)$. From the definition of the performance criterion J , we have that

$$\begin{aligned}
 \sup_{\mathbb{P} \in \mathcal{P}(t, x)} J(t, x, y, \mathbb{P}, \alpha) &= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^T f(s, Y_s, \alpha_s) ds + g(Y_T) \right] \\
 (2.11) \quad &= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + w(\tau, \tilde{X}_\tau, Y_\tau, \alpha^\tau) \right] \\
 &\geq \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + v(\tau, \tilde{X}_\tau, Y_\tau) \right],
 \end{aligned}$$

where the second line in (2.11) follows from Theorem 2.10 and the last inequality follows from the definition of the value function v . Take infimum over $\alpha \in \mathcal{A}_0$ in (2.11) and write

$$(2.12) \quad v(t, x, y) \geq \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + v(\tau, \tilde{X}_\tau, Y_\tau) \right].$$

Now we show the reverse inequality in (2.12). Write the value function v as follows:

$$v(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega})) = \inf_{\alpha \in \mathcal{A}_0} w(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha) = \inf_{\alpha \in \mathcal{A}_0} \hat{w}(\bar{\omega}, \alpha),$$

where $\hat{w}(\bar{\omega}, \alpha) := w(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \alpha)$. From Lemma 2.7, the function w is $\mathcal{B}([0, T]) \otimes \mathcal{L}(\mathbb{R}^{n+d}) \otimes \mathcal{B}(\mathbb{R}^{\tilde{d}}) \otimes \mathcal{B}_{\mathcal{A}_0}$ -measurable. Then, the function \hat{w} is $\mathcal{L}_{\Omega} \otimes \mathcal{B}_{\mathcal{A}_0}$ -measurable because the set $[0, T] \times \mathbb{R}^{n+d} \times \mathbb{R}^{\tilde{d}} \times \mathcal{A}_0$ is Borel and the mappings $\bar{\omega} \mapsto \tau(\bar{\omega})$, $\bar{\omega} \mapsto \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega})$, and $\bar{\omega} \mapsto Y_{\tau(\bar{\omega})}(\bar{\omega})$ are Borel measurable. Therefore, from Lemma 2.7, there exists a universally measurable selection such that $\varphi^\epsilon : (\Omega, \mathcal{F}) \rightarrow (\mathcal{A}_0, \mathcal{B}_{\mathcal{A}_0})$ and

$$v(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega})) + \epsilon > w(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \varphi^\epsilon(\bar{\omega})).$$

There exists a Borel measurable function $\varphi^{\epsilon, \mathbb{P}^*}$ such that $\varphi^{\epsilon, \mathbb{P}^*}(\bar{\omega}) = \varphi^\epsilon(\bar{\omega})$ \mathbb{P}^* -a.s. for each probability measure \mathbb{P}^* on (Ω, \mathcal{F}) in the set $\mathcal{P}(t, x)$; see Lemma 7.27 in Bertsekas and Shreve (1996). For a control process $\alpha \in \mathcal{A}_0$, we define the new control process

$$(2.13) \quad \alpha^{\epsilon, \mathbb{P}^*}(t, \bar{\omega}) = \alpha(t, \bar{\omega}) 1_{\tau(\bar{\omega}) > t} + \varphi^{\epsilon, \mathbb{P}^*}(\bar{\omega}) 1_{\tau(\bar{\omega}) \leq t}.$$

Now, we check the measurability of the process $\alpha^{\epsilon, \mathbb{P}^*}$. The process $\alpha^{\epsilon, \mathbb{P}^*}$ is progressively measurable because \mathcal{A}_0 is a separable metric space and from Lemma 2.1 in Soner and Touzi (2002). By elementary calculations, the process $\alpha^{\epsilon, \mathbb{P}^*}$ is in the set \mathcal{A}_0 . Let τ be a stopping time in the interval $[t, T]$. By Theorem 2.10, we have that

$$(2.14) \quad \begin{aligned} w(t, x, y, \alpha^{\epsilon, \mathbb{P}^*}) &= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s^{\epsilon, \mathbb{P}^*}) ds + w(\tau, \tilde{X}_\tau, Y_\tau, (\alpha^{\epsilon, \mathbb{P}^*})^\tau) \right] \\ &= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + w(\tau, \tilde{X}_\tau, Y_\tau, (\alpha^{\epsilon, \mathbb{P}^*})^\tau) \right] \\ &\leq \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + v(\tau, \tilde{X}_\tau, Y_\tau) \right] + \epsilon, \end{aligned}$$

where the last inequality follows from

$$v(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega})) + \epsilon > w(\tau(\bar{\omega}), \tilde{X}_{\tau(\bar{\omega})}(\bar{\omega}), Y_{\tau(\bar{\omega})}(\bar{\omega}), \varphi^{\epsilon, \mathbb{P}^*}(\bar{\omega}))$$

for \mathbb{P}^* -a.s., and because all probability measures in the set $\mathcal{P}(t, x)$ are absolutely continuous with respect to the probability measure \mathbb{P}^* . Therefore, for an arbitrary control process $\alpha \in \mathcal{A}_0$, the value function v obeys the bound

$$v(t, x, y) \leq \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_t^\tau f(s, Y_s, \alpha_s) ds + v(\tau, \tilde{X}_\tau, Y_\tau) \right] + \epsilon.$$

This shows the reverse inequality of (2.12), which completes the proof. \square

2.5. Viscosity characterisation of the value function. In what follows, assume that Assumptions 2.3 and 2.4 hold. We show that the value function in (2.6) is the unique solution of a nonlinear PDE. We recall that the processes \tilde{X} and Y follow (1.2) and (1.3), respectively. Henceforth, the process $\tilde{Y} := [\tilde{X}, Y]^\top$ follows

$$(2.15) \quad d\tilde{Y}_t = \bar{\mathbf{b}}(t, \tilde{Y}_t, \tilde{\theta}_t, \alpha_t) dt + \bar{\boldsymbol{\sigma}}(t, \tilde{Y}_t, \alpha_t) dB_t^{\mathbb{P}^{\tilde{\theta}}} + \bar{\boldsymbol{\xi}}(t, \tilde{Y}_t, \alpha_t) dL_t,$$

where $\bar{\mathbf{b}}(t, \tilde{y}, \theta, a) := [\mathbf{b}(t, \tilde{x}, \theta), \mathbf{b}(t, \tilde{x}, \theta) \bar{\sigma}(\tilde{y}, a) + \bar{b}(\tilde{y}, a)]^\top$, $\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) := [\boldsymbol{\sigma}(t, \tilde{x}), \boldsymbol{\sigma}(t, \tilde{x}) \bar{\sigma}(\tilde{y}, a)]^\top$, $\bar{\boldsymbol{\xi}}(t, \tilde{y}, a) := [\boldsymbol{\xi}(t, \tilde{x}), \boldsymbol{\xi}(t, \tilde{x}) \bar{\sigma}(\tilde{y}, a)]^\top$, and $\tilde{y} = (y, \tilde{x})$. Denote by $|x|$ the Euclidean norm of $x \in \mathbb{R}^m$, where m is a positive integer.

Assumption 2.12. The functions $\bar{b}, \bar{\sigma}, \bar{\xi}$ are continuous with respect to a , and the function f is Lipschitz with respect to a .

Assumption 2.13. For any compact set $\Theta \subset \mathbb{R}^d$ and $\tilde{Y} \subset \mathbb{R}^{n+d+\tilde{d}}$, there exists a positive constant K_L such that for all $a \in A$

$$(2.16) \quad \begin{aligned} & \left| \bar{b}(t, \tilde{y}_1, \tilde{\theta}_1, a) - \bar{b}(t, \tilde{y}_2, \tilde{\theta}_1, a) \right| + \left| \bar{\sigma}(t, \tilde{y}_1, a) - \bar{\sigma}(t, \tilde{y}_2, a) \right| + \left| \bar{\xi}(t, \tilde{y}_1, a) - \bar{\xi}(t, \tilde{y}_2, a) \right| \\ & \leq K_L |\tilde{y}_1 - \tilde{y}_2| \end{aligned}$$

for all $t \in [0, T]$, $\tilde{y}_1, \tilde{y}_2 \in \mathbb{R}^{1+d+\tilde{d}}$, and $\tilde{\theta} \in \Theta$ and

$$(2.17) \quad \left| \bar{b}(t, \tilde{y}_1, \tilde{\theta}_1, a) - \bar{b}(t, \tilde{y}_1, \tilde{\theta}_2, a) \right| \leq K_L |\tilde{\theta}_1 - \tilde{\theta}_2|$$

for all $t \in [0, T]$, $\tilde{y}_1 \in \tilde{Y}$ and $\tilde{\theta}_1, \tilde{\theta}_2 \in \mathbb{R}^d$.

The two assumptions above are essential in the proof of the characterization of the value function v and to guarantee that the solutions to the SDEs (1.2) and (1.3) are unique.

Assumption 2.14. There exists a positive real number K_G such that for $\theta_1, \theta_2 \in G(t, x)$, we have that $|\theta_1 - \theta_2| \leq K_G |x|$.

The assumption on the size of the set G is needed in the lemma below. Note that if the set G is too large (recall that $G(t, x) \in 2^{\mathbb{R}^d}$), then

$$(2.18) \quad \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \eta} |\tilde{Y}_{t+s} - \tilde{y}|^q \right]$$

does not converge to zero quickly enough as the value of η goes to zero, in which case the value function v in (2.6) cannot be characterized as a solution of a nonlinear PDE.

Now, we use Lemma 3.4 in Fadina, Neufeld, and Schmidt (2019) to prove that the q -moment of the process \tilde{Y} is locally stable under the nonlinear expectation in (2.18).

LEMMA 2.15 (stability of moment). *If Assumptions 2.12, 2.13, 2.14 hold, then for $\alpha_t \in \mathcal{A}_0$, for $q \geq 1$ and $(t, \tilde{y}) \in [0, T] \times \mathbb{R}^{n+d+\tilde{d}}$, there exists $\epsilon(q) \in (0, 1)$ such that for all $\eta < \epsilon(q)$ we have*

$$(2.19) \quad \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q \right] \leq C (\eta^q + \eta^{q/2}),$$

where the constant C is independent of t and h , and the stopping time $\tau_\eta := \inf\{u : |\tilde{Y}_{t+u} - \tilde{y}| \geq 1\} \wedge \eta$.

Proof. Recall Hölder's inequality

$$(2.20) \quad (a + b + c + d)^q \leq c_q (a^q + b^q + c^q + d^q),$$

where $q \geq 1$ and c_q is the smallest real number such that the inequality holds for all $a, b, c, d \geq 0$. Note that the value of the constant c_q depends only on the parameter q . For $\mathbb{P} = \mathbb{P}_{\tilde{\theta}} \in \mathcal{P}(t, x)$, the process \tilde{Y} has the representation

(2.21)

$$\begin{aligned}\tilde{Y}_{t+s} &= \tilde{y} + \int_t^{t+s} \bar{\mathbf{b}}(u, \tilde{Y}_u, \tilde{\theta}_u, \alpha_u) du + \int_t^{t+s} \bar{\boldsymbol{\sigma}}(u, \tilde{Y}_u, \alpha_u) dB_u^{\mathbb{P}_{\tilde{\theta}}} + \int_t^{t+s} \bar{\boldsymbol{\xi}}(u, \tilde{Y}_u, \alpha_u) dL_u, \\ &= \tilde{y} + \int_t^{t+s} \bar{\mathbf{b}}(u, \tilde{Y}_u, \tilde{\theta}_u, \alpha_u) du + \int_t^{t+s} \int_{\mathbb{R}} \bar{\boldsymbol{\xi}}(u, \tilde{Y}_u, \alpha_u) z \nu(dz) du + \tilde{Y}_{t+s}^c + \tilde{Y}_{t+s}^d,\end{aligned}$$

where $\tilde{Y}^{c, \mathbb{P}}$ and $\tilde{Y}^{d, \mathbb{P}}$ are the continuous and discontinuous local martingale parts of \tilde{Y} , respectively, under the probability measure \mathbb{P} . Use (2.20) to write

(2.22)

$$\begin{aligned}\mathbb{E}^{\mathbb{P}} &\left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} \left| \int_t^{t+s} \bar{\mathbf{b}}(u, \tilde{Y}_u, \tilde{\theta}_u, \alpha_u) du + \int_t^{t+s} \int_{\mathbb{R}} \bar{\boldsymbol{\xi}}(u, \tilde{Y}_u, \alpha_u) z \nu(dz) du + \tilde{Y}_{t+s}^c + \tilde{Y}_{t+s}^d \right|^q \right] \\ &\leq c_q \left(\mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} \left| \int_t^{t+s} \bar{\mathbf{b}}(u, \tilde{Y}_u, \tilde{\theta}_u, \alpha_u) du \right|^q \right] + \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}^{c, \mathbb{P}}|^q \right] + \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}^{d, \mathbb{P}}|^q \right] \right. \\ &\quad \left. + \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} \left| \int_t^{t+s} \int_{\mathbb{R}} \bar{\boldsymbol{\xi}}(u, \tilde{Y}_u, \alpha_u) z \nu(dz) du \right|^q \right] \right).\end{aligned}$$

Let $\tilde{\mathbf{Y}} := \{y \in \mathbb{R}^{n+d+\tilde{d}} \mid |y - \tilde{y}| \leq 1\}$ and $\tilde{\boldsymbol{\Theta}} := \bigcup_{y \in \tilde{\mathbf{Y}}} G(t, x)$. The sets $\tilde{\mathbf{Y}}$ and $\tilde{\boldsymbol{\Theta}}$ are compact because the mapping $y \mapsto G(t, x)$ is continuous. From Assumption 2.13, there exists a positive constant \tilde{K}_L that depends on \tilde{y} such that $\bar{\mathbf{b}}(t, \tilde{y}_1, \tilde{\theta}, a) \leq \tilde{K}_L (1 + |\tilde{y}_1|)$, $\bar{\boldsymbol{\sigma}}(t, \tilde{y}_1, a) \leq \tilde{K}_L (1 + |\tilde{y}_1|)$, and $\bar{\boldsymbol{\xi}}(t, \tilde{y}_1, a) \leq \tilde{K}_L (1 + |\tilde{y}_1|)$ for all $\tilde{y}_1 \in \tilde{\mathbf{Y}}$, $\tilde{\theta} \in \tilde{\boldsymbol{\Theta}}$, and $a \in A$. Define $\hat{K} := \max\{1, K_G, \tilde{K}_L\}$, where K_G is the constant from Assumption 2.14. Consider the first term on the right-hand side in (2.22):

$$\begin{aligned}(2.23) \quad &\mathbb{E}^{\mathbb{P}} \left[\left(\int_t^{t+\tau_\eta} |\bar{\mathbf{b}}(u, \tilde{Y}_u, \tilde{\theta}_u, \alpha_u)| du \right)^q \right] \leq \\ &\leq \mathbb{E}^{\mathbb{P}} \left[\left(\int_t^{t+\tau_\eta} |\bar{\mathbf{b}}(u, \tilde{Y}_u, \hat{\theta}_u, \alpha_u)| + \hat{K} |\hat{\theta}_u - \tilde{\theta}_u| du \right)^q \right] \quad \text{from (2.17)} \\ &\leq \hat{K}^q \eta^q \mathbb{E}^{\mathbb{P}} \left[\left(1 + \hat{K} \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}| \right)^q \right] \quad \text{from (2.14)} \\ &\leq \tilde{c}_q \hat{K}^{2q} \eta^q \mathbb{E}^{\mathbb{P}} \left[1 + \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q + |\tilde{y}|^q \right], \quad \text{from (2.20)}\end{aligned}$$

where the second inequality follows from the Lipschitz assumption on $\bar{\mathbf{b}}$, the penultimate inequality follows from the linear growth property on $\bar{\mathbf{b}}$, and the last inequality follows from Hölder's inequality in (2.20). Next, consider the last term in (2.22):

$$\begin{aligned}
(2.24) \quad & \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_h} \left| \int_t^{t+s} \int_{\mathbb{R}} \bar{\xi}(u, \tilde{Y}_u, \alpha_u) z \nu(dz) du \right|^q \right] \\
& \leq \mathbb{E}^{\mathbb{P}} \left[\left| \int_t^{t+\tau_\eta} \int_{\mathbb{R}} \bar{\xi}(u, \tilde{Y}_u, \alpha_u) z \nu(dz) du \right|^q \right] \\
& \leq \mathbb{E}^{\mathbb{P}} \left[\left\{ \int_t^{t+\tau_\eta} \int_{\mathbb{R}} \left(1 + \hat{K} |\tilde{Y}_{t+u}| \right) |z| \nu(dz) du \right\}^q \right] \\
& \leq C_J \eta^q \mathbb{E}^{\mathbb{P}} \left[\left(1 + \hat{K} \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}| \right)^q \right] \\
& \leq \tilde{c}_q C_J \hat{K}^q \eta^q \mathbb{E}^{\mathbb{P}} \left[1 + \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q + |\tilde{y}|^q \right],
\end{aligned}$$

where the second inequality follows from the linear growth property on $\bar{\xi}$, and the last inequality follows from Hölder's inequality in (2.20). By the Burkholder–Davis–Gundy (BDG) inequality in Cohen and Elliott (2015), we have that

$$\begin{aligned}
(2.25) \quad & \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}^{c, \mathbb{P}}|^q \right] \\
& \leq \tilde{c}_q \mathbb{E}^{\mathbb{P}} \left[\left(\int_t^{t+\tau_\eta} \left| \bar{\sigma}(t+u, \tilde{Y}_{t+u}, \alpha_{t+u}) \bar{\sigma}(t+u, \tilde{Y}_{t+u}, \alpha_{t+u})^\top \right| du \right)^{q/2} \right] \\
& \leq \tilde{c}_q \mathbb{E}^{\mathbb{P}} \left[\hat{K}^q \eta^{q/2} \left(1 + \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}| \right)^q \right] \\
& \leq \tilde{c}_q c_q \hat{K}^q \eta^{q/2} \mathbb{E}^{\mathbb{P}} \left[1 + \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q + |\tilde{y}|^q \right],
\end{aligned}$$

where \tilde{c}_q is a constant that depends on q , the last inequality follows from (2.20), and \hat{K} does not depend on h . Next, consider the purely discontinuous local martingale; the BDG inequality and the linear growth of the function $\bar{\xi}$ imply

$$\begin{aligned}
(2.26) \quad & \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}^{d, \mathbb{P}}|^q \right] \leq \tilde{c}_q \mathbb{E}^{\mathbb{P}} \left[\left\langle \tilde{Y}_{t+\tau_\eta}^{d, \mathbb{P}} \right\rangle^{q/2} \right], \\
& \leq \tilde{c}_q \mathbb{E}^{\mathbb{P}} \left[\left| \int_t^{t+\tau_\eta} \int_{\mathbb{R}} |\bar{\xi}(u, \tilde{Y}_u, \alpha_u)|^2 |z|^2 \nu(dz) du \right|^{q/2} \right] \\
& \leq \tilde{c}_q C_J \mathbb{E}^{\mathbb{P}} \left[\hat{K}^q \eta^{q/2} \left(1 + \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s}| \right)^q \right] \\
& \leq \tilde{c}_q c_q C_J \hat{K}^q \eta^{q/2} \mathbb{E}^{\mathbb{P}} \left[1 + \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q + |\tilde{y}|^q \right].
\end{aligned}$$

Choose the value of the parameter ϵ so that $1 - C_1 \eta^q - C_2 \eta^{q/2} := 1 - (\tilde{c}_q \hat{K}^{2q} h^q + \tilde{c}_q C_J \hat{K}^q \eta^q + \tilde{c}_q c_q \hat{K}^q \eta^{q/2} + \tilde{c}_q c_q C_J \hat{K}^q \eta^{q/2}) > 0$. Therefore, if $\eta < \epsilon$, from inequality (2.23) and (2.25), we have that

$$\begin{aligned} \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q \right] &\leq (C_1 \eta^q + C_2 \eta^{q/2}) \mathbb{E}^{\mathbb{P}} \left[1 + \sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q + |\tilde{y}|^q \right] \\ &\leq (C_1 \eta^q + C_2 \eta^{q/2}) \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_\eta} |\tilde{Y}_{t+s} - \tilde{y}|^q \right] \\ &\quad + (C_1 \epsilon^q + C_2 \epsilon^{q/2}) [1 + |\tilde{y}|^q]. \end{aligned}$$

Let

$$C = \frac{(C_1 \eta^q + C_2 \eta^{q/2}) (1 + |\tilde{y}|^q)}{1 - (C_1 \epsilon^q + C_2 \epsilon^{q/2})}.$$

Now, for $\mathbb{P} \in \mathcal{P}(t, x, G)$ the following inequality holds:

$$(2.27) \quad \mathbb{E}^{\mathbb{P}} \left[\sup_{0 \leq s \leq \tau_n} |\tilde{Y}_{t+s} - \tilde{y}|^q \right] \leq C (\eta^{q/2} + \eta^q). \quad \square$$

We require that the set G satisfies the following structural assumption for the proof of the characterization of the value function as a solution of a nonlinear PDE.

Assumption 2.16. We assume that $G(t, x)$ is a closed set, in the standard topology, for all $(t, x) \in [0, T] \times \mathbb{R}^{n+d}$. And we assume that the Hausdorff distance between the sets $G(t, x)$ and $G(s, y)$ has a Lipschitz-like property. That is, there exists a constant K_H such that $d_{\text{Haus}}(G(t, x), G(s, y)) \leq K_H (|t - s| + |x - y|)$ for all $(t, x), (s, y) \in [0, T] \times \mathbb{R}^{n+d}$.

The assumption above is essential to characterize the value function as a solution of a nonlinear PDE, but not to prove the DPP of the value function. We assume that $G(t, x)$ is a closed set, so the maximal is attained in the set $G(t, x)$. The second statement of the assumption requires the function G to be well-behaved. This condition allows us to quantify the nonlinear operator that depends on the set $G(t, x)$.

Denote by $\varphi \in C_b^{2,3}([0, T] \times \mathbb{R}^{n+d+\bar{d}})$ the set of functions on $[0, T] \times \mathbb{R}^{n+d+\bar{d}}$ that have bounded continuous derivatives up to the second and third order in t and x , respectively. We use the next lemma to characterize the value function as a solution of a nonlinear PDE.

LEMMA 2.17. *Let $\varphi \in C_b^{2,3}([0, T] \times \mathbb{R}^{n+d+\bar{d}})$, and let u be a small enough positive real number. If Assumptions 2.12, 2.13, 2.14, and 2.16 hold, then there exists a probability measure $\mathbb{P} \in \mathcal{P}(t, x)$ such that for all $\alpha \in \mathcal{A}_0$,*

$$\begin{aligned} (2.28) \quad &\mathbb{E}^{\mathbb{P}} \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \tilde{\theta}_{t+\tau_s}, \alpha_{t+\tau_s}) \right. \\ &\quad + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^{\text{T}} \partial_{\tilde{y}\tilde{y}}^2 \varphi(t, \tilde{y}) \right) \\ &\quad \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) h(z)) \nu(dz) \right] \\ &\geq -C_A s^{1/2} + \inf_{a \in A} \sup_{\tilde{\theta} \in G(t, x)} \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}, a) + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^{\text{T}} \partial_{\tilde{y}\tilde{y}}^2 \varphi(t, \tilde{y}) \right) \right. \\ &\quad \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) h(z)) \nu(dz) \right], \end{aligned}$$

where the stopping times $\tau_s := \inf\{m : |\tilde{Y}_{t+m} - \tilde{y}| \geq 1\} \wedge s$ for all $s \in [0, u]$, and C_A is a constant that depends on t and x .

Proof. From Assumption 2.16, there exists an optimal a^* that minimizes the supremum term, and there exists an optimal $\tilde{\theta}^{a^*}$ that maximizes the right-hand side of (2.28). The process $\gamma_{t+s} = \text{proj}_{\tilde{\theta}^{a^*}}(G(t+s, \tilde{X}_{t+s}))$ is \mathcal{F}_{t+s} -adapted, where $\text{proj}_x(A)$ is an element in the set A such that the Euclidean distance between the element and x is smallest. The process γ is right-continuous because the process \tilde{X} is right-continuous and due to the Lipschitz property of the mapping G ; see Assumption 2.16. Therefore, the process γ is progressively measurable. Let \mathbb{P}^γ be a measure such that

$$\beta_{t+s}^{\mathbb{P}^\gamma} = \mathbf{b}(t+s, \tilde{X}_{t+s}, \gamma_{t+s}) + \int_{\mathbb{R}} (\hat{h}(\boldsymbol{\xi}(t+s, \tilde{X}_{t+s})z) - \boldsymbol{\xi}(t+s, \tilde{X}_{t+s})h(z)) \nu(dz).$$

It is easy to see that $\mathbb{P}^\gamma \in \mathcal{P}(t, x)$ because of the way we construct the process γ .

Next, we show that \mathbb{P}^γ satisfies (2.28). First, note that

$$(2.29) \quad |\gamma_{t+s} - \tilde{\theta}^{a^*}| \leq d_{\text{Haus}}(G(t+s, \tilde{X}_{t+s}), G(t, x)) \leq K_H (s + |\tilde{X}_{t+s} - x|).$$

Let $C_1 = \max\{|\partial_x \varphi(t, x)|, |\partial_{xx}^2 \varphi(t, x)|\}$. Consider

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}^\gamma} \left[\left| \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \gamma_{t+\tau_s}, \alpha_{t+\tau_s}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}^{a^*}, \alpha_{t+\tau_s}) \right| \right] \\ & \leq C_1 \mathbb{E}^{\mathbb{P}^\gamma} \left[\left| \bar{\mathbf{b}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \gamma_{t+\tau_s}, \alpha_{t+\tau_s}) - \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}^{a^*}, \alpha_{t+\tau_s}) \right| \right] \\ & \leq C_1 (K_L + K_H) \mathbb{E}^{\mathbb{P}^\gamma} \left[\left| \tilde{Y}_{t+\tau_s} - \tilde{y} \right| + \tau_s \right] \\ & \leq C_2 s^{1/2}, \end{aligned}$$

where the second line follows from the bounded derivative of φ , the third line is a result of (2.29) and Assumptions 2.13, 2.14, and 2.16, and the last inequality follows from Lemma 2.15 with $q = 1$. Moreover, C_1, C_2 are constants that depend on t and \tilde{y} . Next, consider

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}^\gamma} \left[\frac{1}{2} \left| \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^\top \partial_{\tilde{y}\tilde{y}}^2 \varphi \right) \right. \right. \\ & \quad \left. \left. - \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, \alpha_{t+\tau_s})^\top \partial_{\tilde{y}\tilde{y}}^2 \varphi \right) \right| \right] \\ & \leq \mathbb{E}^{\mathbb{P}^\gamma} \left[\frac{1}{2} \left| \partial_{\tilde{y}\tilde{y}}^2 \varphi(t, \tilde{y}) \right| \left| \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^\top \right) \right. \right. \\ (2.30) \quad & \quad \left. \left. - \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, \alpha_{t+\tau_s})^\top \right) \right| \right] \\ & \leq C_1 \mathbb{E}^{\mathbb{P}^\gamma} \left[\left| \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^\top \right) \right. \right. \\ & \quad \left. \left. - \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, \alpha_{t+\tau_s})^\top \right) \right| \right], \end{aligned}$$

where the second line follows from the bounded derivative of the function φ . From Assumption 2.12, we have that $|\bar{\sigma}(t+s, \tilde{Y}_{t+s}, a) - \bar{\sigma}(t, \tilde{y}, a)| \leq K_L(s + |\tilde{Y}_{t+s} - \tilde{y}|)$ for all $a \in A$; therefore, the right-hand side of (2.30) obeys the bound

$$\begin{aligned} & C_1 \mathbb{E}^{\mathbb{P}^\gamma} \left[\left| \text{Tr} \left(\bar{\sigma}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) \bar{\sigma}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^\top \right) \right. \right. \\ & \quad \left. \left. - \text{Tr} \left(\bar{\sigma}(t, \tilde{y}, \alpha_{t+\tau_s}) \bar{\sigma}(t, \tilde{y}, \alpha_{t+\tau_s})^\top \right) \right| \right] \\ & \leq C_1 \max\{2K_L, K_L^2\} \mathbb{E}^{\mathbb{P}^\gamma} \left[|\bar{\sigma}(t, \tilde{y}, \alpha_{t+\tau_s})| \left(\tau_s + |\tilde{Y}_{t+\tau_s} - \tilde{y}| \right) + \left(\tau_s + |\tilde{Y}_{t+\tau_s} - \tilde{y}| \right)^2 \right] \leq C_A s^{1/2}, \end{aligned}$$

where the last inequality follows from Lemma 2.15 with $q = 1$ and C_A is a constant that depends on t and \tilde{y} . For the remaining term on the left-hand side of (2.28), we have that

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}} \left[\int_{\mathbb{R}} \left(\varphi(t, \tilde{y} + \bar{\xi}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\xi}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) h(z) \right) \nu(dz) \right. \\ & \quad \left. \int_{\mathbb{R}} \left(\varphi(t, \tilde{y} + \bar{\xi}(t, \tilde{y}, \alpha_{t+\tau_s}) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\xi}(t, \tilde{y}, \alpha_{t+\tau_s}) h(z) \right) \nu(dz) \right] \\ & \leq C \mathbb{E}^{\mathbb{P}} \left[\int_{\mathbb{R}} \left(\left| z (\bar{\xi}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) - \bar{\xi}(t, \tilde{y}, \alpha_{t+\tau_s})) \right| \right. \right. \\ & \quad \left. \left. + \left| (\bar{\xi}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) - \bar{\xi}(t, \tilde{y}, \alpha_{t+\tau_s})) \right| |h(z)| \right) \nu(dz) \right] \\ & \leq C K_L \mathbb{E}^{\mathbb{P}} \left[\tau_s + |\tilde{Y}_{t+\tau_s} - \tilde{y}| \right] \leq C_J s^{1/2}, \end{aligned}$$

where the first inequality follows from the property of the function φ , the second inequality follows from (2.16), and C_J is a constant that depends on t and \tilde{y} . \square

Next, we show that the value function v in (2.6) is a viscosity solution of a Hamilton–Jacobi–Bellman–Isaacs (HJBI) equation.

THEOREM 2.18 (viscosity solution). *If Assumptions 2.12, 2.13, 2.14, and 2.16 hold and the value function v in (2.6) is continuous, then v is a viscosity solution of the PDE*

$$\begin{aligned} & (2.31) \\ & \partial_t v(t, \tilde{y}) + \inf_{a \in A} \sup_{\tilde{\theta} \in G(t, x)} \left[f(t, \tilde{y}, a) + \partial_{\tilde{y}} v \cdot \bar{b}(t, \tilde{y}, \tilde{\theta}, a) + \frac{1}{2} \text{Tr} \left(\bar{\sigma}(t, \tilde{y}, a) \bar{\sigma}(t, \tilde{y}, a)^\top \partial_{\tilde{y}\tilde{y}}^2 v \right) \right. \\ & \quad \left. + \int_{\mathbb{R}} \left(v(t, \tilde{y} + \bar{\xi}(t, \tilde{y}, a) z) - v(t, \tilde{y}) - \partial_{\tilde{y}} v(t, \tilde{y}) \bar{\xi}(t, \tilde{y}, a) h(z) \right) \nu(dz) \right] = 0, \end{aligned}$$

subject to the terminal condition $v(T, \tilde{y}) = g(\tilde{y})$, and recall that A is a compact set taking values in \mathbb{R}^k .

Proof. Let $(t, \tilde{y}) \in [0, T] \times \mathbb{R}^{n+d+\tilde{d}}$, and let $\varphi \in C_b^{2,3}([0, T] \times \mathbb{R}^{n+d+\tilde{d}})$ be such that $v(t, \tilde{y}) = \varphi(t, \tilde{y})$, and $v \geq \varphi$ on $[0, T] \times \mathbb{R}^{n+d+\tilde{d}}$. Recall the DPP holds; therefore

$$\begin{aligned} & (2.32) \\ & 0 = \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} f(t+s, \tilde{Y}_{t+s}, \alpha_{t+s}) ds + v(t+s, \tilde{Y}_{t+s}) - v(t, \tilde{y}) \right] \\ & \geq \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} f(t+s, \tilde{Y}_{t+s}, \alpha_{t+s}) ds + \varphi(t+s, \tilde{Y}_{t+s}) - \varphi(t, \tilde{y}) \right]. \end{aligned}$$

We recall that the stopping times are $\tau_u = \inf\{s : |\tilde{Y}_{t+s} - \tilde{y}| \geq 1\} \wedge u$. Next, we show that v is a viscosity supersolution of (2.31). Let $\mathbb{P} \in \mathcal{P}(t, x)$ and $\alpha \in \mathcal{A}_0$; by Itô's lemma we have that

(2.33)

$$\begin{aligned} & \varphi(t + \tau_u, \tilde{Y}_{t+\tau_u}) - \varphi(t, \tilde{y}) \\ &= \int_0^{\tau_u} \partial_t \varphi(t + s, \tilde{Y}_{t+s}) ds + \int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t + s, \tilde{Y}_{t+s}) dM_{t+s}^{\mathbb{P}} \\ & \quad + \int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t + s, \tilde{Y}_{t+s}) \bar{\mathbf{b}}(t + s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) ds \\ & \quad + \int_0^{\tau_u} \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi(t + s, \tilde{Y}_{t+s}) \right) ds \\ & \quad + \int_0^{\tau_u} \int_{\mathbb{R}} \left(\varphi(t + s, \tilde{Y}_{t+s} + \bar{\boldsymbol{\xi}}(t + s, \tilde{Y}_{t+s}, \alpha_{t+s}) z) - \varphi(t + s, \tilde{Y}_{t+s}) \right. \\ & \quad \left. - \partial_{\tilde{y}} \varphi(t + s, \tilde{Y}_{t+s}) h(z) \right) \nu(dz) ds. \end{aligned}$$

The expectation of the stochastic integral term is zero because the function $\varphi \in C_b^{2,3}([0, T] \times \mathbb{R}^{n+d+\tilde{d}})$, and the process $M^{\mathbb{P}}$ is a martingale under the probability measure \mathbb{P} . Now, consider the expectation of each term on the right-hand side of (2.33). Begin with the third term:

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t + s, \tilde{Y}_{t+s}) \bar{\mathbf{b}}(t + s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) ds \right] \\ & \geq \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t + s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) \right. \\ & \quad \left. - \left| \partial_{\tilde{y}} \varphi(t + s, \tilde{Y}_{t+s}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \right| \left| \bar{\mathbf{b}}(t + s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) \right| ds \right]. \end{aligned}$$

From Lemma 2.15 and Theorem 4.1 in Fadina et al. (2019), write

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} \left| \partial_{\tilde{y}} \varphi(t + s, \tilde{Y}_{t+s}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \right| \left| \bar{\mathbf{b}}(t + s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) \right| ds \right] \\ & \leq C_1 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right), \end{aligned}$$

where C_1 is a constant that depends only on \tilde{y} . Perform similar calculations for the first and the fourth terms on the right-hand side of (2.33) (i.e., $\partial_t \varphi(t + s, \tilde{Y}_{t+s})$ and $\partial_{\tilde{y}\tilde{y}}^2 \varphi(t + s, \tilde{Y}_{t+s})$) and write

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}} \left[\varphi(t + \tau_u, \tilde{Y}_{t+\tau_u}) - \varphi(t, \tilde{y}) \right] \geq -C_2 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right) \\ & \quad + \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} \left(\partial_t \varphi(t, \tilde{y}) + \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t + s, \tilde{Y}_{t+s}, \tilde{\theta}_{t+s}, \alpha_{t+s}) \right. \right. \right. \\ & \quad \left. \left. + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+s}, \alpha_{t+s}) \bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+s}, \alpha_{t+s})^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi(t, \tilde{y}) \right) \right. \right. \\ & \quad \left. \left. + \int_0^{\tau_u} \int_{\mathbb{R}} \left(\varphi(t + s, \tilde{Y}_{t+s} + \bar{\boldsymbol{\xi}}(t + s, \tilde{Y}_{t+s}, \alpha_{t+s}) z) - \varphi(t + s, \tilde{Y}_{t+s}) \right. \right. \right. \\ & \quad \left. \left. \left. - \partial_{\tilde{y}} \varphi(t + s, \tilde{Y}_{t+s}) h(z) \right) \nu(dz) \right] ds \right] \end{aligned}$$

$$\begin{aligned}
&\geq -C_2 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right) + \mathbb{E}^{\mathbb{P}} \left[\tau_u \partial_t \varphi(t, \tilde{y}) \right] \\
&\quad + \int_0^u \left(\mathbb{E}^{\mathbb{P}} \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \tilde{\theta}_{t+\tau_s}, \alpha_{t+\tau_s}) \right. \right. \\
&\quad \left. \left. + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) \bar{\boldsymbol{\sigma}}(t + \tau_s, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s})^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi(t, \tilde{y}) \right) \right. \right. \\
&\quad \left. \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{Y}_{t+\tau_s}, \alpha_{t+\tau_s}) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) h(z)) \nu(dz) \right] \right) ds,
\end{aligned}$$

where C_2 is a constant that depends only on \tilde{y} . From Lemma 2.17, if u is small enough, there exists a probability measure $\tilde{\mathbb{P}} \in \mathcal{P}(t, x)$ such that for all $\alpha \in \mathcal{A}_0$, the inequality in (2.28) holds. Hence, for all $\alpha \in \mathcal{A}_0$ the following equation holds:

$$\begin{aligned}
(2.34) \quad &\sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\varphi(t + \tau_u, \tilde{Y}_{t+\tau_u}) - \varphi(t, \tilde{y}) \right] \geq -C_2 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right) - C_A u^{3/2} \\
&+ \mathbb{E}^{\tilde{\mathbb{P}}} \left[\tau_u \partial_t \varphi(t, \tilde{y}) \right] + \int_0^u \left(\inf_{a \in A} \sup_{\tilde{\theta} \in G(t, x)} \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}, a) + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi \right) \right. \right. \\
&\left. \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) h(z)) \nu(dz) \right] \right) ds.
\end{aligned}$$

Now, consider the running reward of the performance criterion. Use Lemma 2.15 to write

$$\begin{aligned}
&\sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} \left| f(t + s, \tilde{Y}_{t+s}, \alpha_{t+s}) - f(t, \tilde{y}, \alpha_{t+s}) \right| ds \right] \\
&\leq \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} K \left(s + |\tilde{Y}_{t+s} - \tilde{y}| \right) ds \right] \leq C_3 \left(u^{3/2} + u^2 \right).
\end{aligned}$$

Use the triangle inequality to write

$$\begin{aligned}
(2.35) \quad &\sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} f(t + s, \tilde{Y}_{t+s}, \alpha_{t+s}) ds \right] \\
&\geq \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\tilde{\mathbb{P}}} \left[\int_0^{\tau_u} f(t, \tilde{y}, \alpha_{t+s}) ds \right] - C_3 \left(u^{3/2} + u^2 \right) \\
&\geq \inf_{a \in A} \mathbb{E}^{\mathbb{P}} \left[\tau_u f(t, \tilde{y}, a) \right] - C_3 \left(u^{3/2} + u^2 \right),
\end{aligned}$$

where C_3 is a constant that depends only on x . Therefore, from (2.32), (2.34), and (2.35), we have that

$$\begin{aligned}
(2.36) \quad &0 \geq \inf_{a \in A} u \mathbb{E}^{\tilde{\mathbb{P}}} \left[(\tau_u/u) f(t, \tilde{y}, a) \right] + u \left(\mathbb{E}^{\tilde{\mathbb{P}}} \left[(\tau_u/u) \partial_t \varphi(t, \tilde{y}) \right] \right. \\
&\quad \left. + \inf_{a \in A} \sup_{\tilde{\theta} \in G(t, x)} \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}, a) + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi \right) \right. \right. \\
&\quad \left. \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) h(z)) \nu(dz) \right] \right) \\
&\quad \left. - C_4 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right) \right),
\end{aligned}$$

where $C_4 = C_2 + C_3 + C_A$. As u is small enough and $\mathbb{E}^{\tilde{\mathbb{P}}} [\tau_u/u] \rightarrow 1$, write

$$0 \geq \partial_t \varphi(t, \tilde{y}) + \inf_{a \in A} \sup_{\tilde{\theta} \in G(t, x)} \left[f(t, \tilde{y}, a) + \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}, a) + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi \right) \right. \\ \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) h(z)) \nu(dz) \right].$$

Thus, v is a viscosity supersolution of the PDE.

Next, we show that v is a viscosity subsolution of the PDE (2.31). Let $(t, \tilde{y}) \in [0, T] \times \mathbb{R}^{n+d+\tilde{d}}$, and let $\varphi \in C_b^{2,3}([0, T] \times \mathbb{R}^{n+d+\tilde{d}})$ be such that $v(t, \tilde{y}) = \varphi(t, \tilde{y})$ and $v \leq \varphi$ on $[0, T] \times \mathbb{R}^{n+d+\tilde{d}}$. As v satisfies the DPP, we have

$$0 = \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t, x, G)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} f(t+s, \tilde{Y}_{t+s}, \alpha_{t+s}) ds + v(t+s, \tilde{Y}_{t+s}) - v(t, \tilde{y}) \right] \\ \leq \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}(t, x, G)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} f(t+s, \tilde{Y}_{t+s}, \alpha_{t+s}) ds + \varphi(t+s, \tilde{Y}_{t+s}) - \varphi(t, \tilde{y}) \right].$$

Then, for a constant control process $\alpha = a \in A$, we have that

$$0 \leq \sup_{\mathbb{P} \in \mathcal{P}(t, x, G)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} f(t+s, \tilde{Y}_{t+s}, a) ds + \varphi(t+s, \tilde{Y}_{t+s}) - \varphi(t, \tilde{y}) \right].$$

As above in (2.34), we obtain the following inequality for all $\mathbb{P} \in \mathcal{P}(t, x)$:

$$\mathbb{E}^{\mathbb{P}} [\varphi(t + \tau_u, \tilde{Y}_{t+\tau_u}) - \varphi(t, \tilde{y})] \leq C_1 \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right) \\ + \mathbb{E}^{\mathbb{P}} [\tau_u \partial_t \varphi(t, \tilde{y})] + \int_0^u \left(\sup_{\tilde{\theta} \in G(t, x)} \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}, a) \right. \right. \\ \left. \left. + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi(t, \tilde{y}) \right) \right. \right. \\ \left. \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) h(z)) \nu(dz) \right] \right) ds.$$

Similarly,

$$\sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^{\mathbb{P}} \left[\int_0^{\tau_u} |f(t+s, \tilde{Y}_{t+s}, a) - f(t, \tilde{y}, a)| ds \right] \leq C_2 \left(u^{3/2} + u^2 \right),$$

where C_1 and C_2 are constants that depend only on \tilde{y} . We combine these two inequalities to obtain

$$(2.37) \quad 0 \leq u \mathbb{E}^{\mathbb{P}} [(\tau_u/u)] f(t, \tilde{y}, a) + u \left(\mathbb{E}^{\mathbb{P}} [(\tau_u/u)] \partial_t \varphi(t, \tilde{y}) \right. \\ \left. + \sup_{\tilde{\theta} \in G(t, x)} \left[\partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}, a) + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^{\top} \partial_{\tilde{y}\tilde{y}}^2 \varphi \right) \right. \right. \\ \left. \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) h(z)) \nu(dz) \right] \right) \\ \left. + (C_1 + C_2) \left(u^{3/2} + u^2 + u^{5/2} + u^3 \right) \right).$$

Thus, we have that

$$0 \leq \partial_t \varphi(t, \tilde{y}) + \inf_{a \in A} \sup_{\tilde{\theta} \in G(t, x)} \left[f(t, \tilde{y}, a) + \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\mathbf{b}}(t, \tilde{y}, \tilde{\theta}, a) + \frac{1}{2} \text{Tr} \left(\bar{\boldsymbol{\sigma}}(t, \tilde{y}, a) \bar{\boldsymbol{\sigma}}(t, \tilde{y}, a)^T \partial_{\tilde{y}\tilde{y}}^2 \varphi \right) \right. \\ \left. + \int_{\mathbb{R}} (\varphi(t, \tilde{y} + \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) z) - \varphi(t, \tilde{y}) - \partial_{\tilde{y}} \varphi(t, \tilde{y}) \bar{\boldsymbol{\xi}}(t, \tilde{y}, a) h(z)) \nu(dz) \right],$$

because a is an arbitrary constant control in A . Therefore, v is a viscosity subsolution of the PDE in (2.31). \square

The parameter $\tilde{\theta}$ does not appear in the second order term and in the nonlocal term in (2.31) because the parameter θ only appears in the drift term of the process \tilde{Y} . Next, we show that there exists a unique viscosity solution for the PDE in (2.31). This ensures that we can use a numerical method to approximate the value function v when we cannot find an analytic solution.

PROPOSITION 2.19 (uniqueness of viscosity solution). *If Assumptions 2.12, 2.13, 2.14, and 2.16 hold, then the nonlinear PDE in (2.31) has a unique uniformly continuous viscosity solution with at most linear growth in x .*

Proof. Let v_1 and v_2 be a uniformly continuous viscosity subsolution (resp., supersolution) of (2.31) with $v_1(T, \tilde{y}) \leq v_2(T, \tilde{y})$ for $\tilde{y} \in \mathbb{R}^{n+d+\tilde{d}}$. We show that $v_1(t, \tilde{y}) \leq v_2(t, \tilde{y})$ for all $(t, \tilde{y}) \in [0, T] \times \mathbb{R}^{n+d+\tilde{d}}$. We follow an argument from Pham (1998), and for $\beta, \epsilon, \delta, \lambda > 0$ let us define the function Φ in $[0, T] \times \mathbb{R}^{n+d+\tilde{d}} \times \mathbb{R}^{n+d+\tilde{d}}$

(2.38)

$$\Phi(t, \tilde{y}_1, \tilde{y}_2) = v_1(t, \tilde{y}_1) - v_2(t, \tilde{y}_2) - \frac{\beta}{t} - \frac{1}{2\epsilon} |\tilde{y}_1 - \tilde{y}_2|^2 - \delta e^{\lambda(T-t)} (|\tilde{y}_1|^2 + |\tilde{y}_2|^2).$$

Let the function Φ admit a maximum at $(\bar{t}, \bar{y}_1, \bar{y}_2) \in [0, T] \times \mathbb{R}^{n+d+\tilde{d}} \times \mathbb{R}^{n+d+\tilde{d}}$. As in Theorem 4.1 in Pham (1998), we only consider the case when $\bar{t} < T$. Then, it suffices to show that the inequality

$$(2.39) \quad \left[\sup_{a \in A} \left[\sup_{\tilde{\theta} \in G(\bar{t}, \bar{x}_1)} \bar{\mathbf{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a) \left(\frac{1}{\epsilon} (\bar{y}_1 - \bar{y}_2) + 2\delta e^{\lambda(T-\bar{t})} \bar{y}_1 \right) \right. \right. \\ \left. \left. - \sup_{\tilde{\theta} \in G(\bar{t}, \bar{x}_2)} \bar{\mathbf{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}, a) \left(\frac{1}{\epsilon} (\bar{y}_1 - \bar{y}_2) - 2\delta e^{\lambda(T-\bar{t})} \bar{y}_2 \right) \right] \right] \\ \leq C \left(\frac{|\bar{y}_1 - \bar{y}_2|^2}{\epsilon} + 2\delta e^{\lambda(T-\bar{t})} (1 + |\bar{y}_1|^2 + |\bar{y}_2|^2) \right)$$

holds, where C is a positive constant that does not depend on $\epsilon, \delta, \lambda$, and β because all the conditions and all the remaining terms in equation (4.6) in Pham (1998) are satisfied. Next, proceed as in Theorem 4.1 in Pham (1998). For each $\tilde{\theta} \in G(t, \bar{x}_1)$, there exists $\tilde{\theta}^p \in G(\bar{t}, \bar{x}_2)$ such that $|\tilde{\theta} - \tilde{\theta}^p| \leq |\tilde{y}_1 - \tilde{y}_2|$,

(2.40)

$$\bar{\mathbf{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a) \left(\frac{1}{\epsilon} (\bar{y}_1 - \bar{y}_2) + 2\delta e^{\lambda(T-\bar{t})} \bar{y}_1 \right) - \bar{\mathbf{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}^p, a) \left(\frac{1}{\epsilon} (\bar{y}_1 - \bar{y}_2) - 2\delta e^{\lambda(T-\bar{t})} \bar{y}_2 \right) \\ \leq |\bar{\mathbf{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a) - \bar{\mathbf{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}^p, a)| |(\bar{y}_1 - \bar{y}_2)/\epsilon|$$

$$\begin{aligned}
& + 2\delta e^{\lambda(T-\bar{t})} \left(\left| \bar{\mathbf{b}}(\bar{t}, \bar{y}_1, \tilde{\theta}, a) \right| |\bar{y}_1| + \left| \bar{\mathbf{b}}(\bar{t}, \bar{y}_2, \tilde{\theta}, a) \right| |\bar{y}_2| \right) \\
& \leq K_L \frac{|\bar{y}_1 - \bar{y}_2|^2}{\epsilon} + 2\delta e^{\lambda(T-\bar{t})} \left(|\bar{y}_1| + |\bar{y}_2| + |\bar{y}_1|^2 + |\bar{y}_2|^2 \right) \\
& \leq C \left(\frac{|\bar{y}_1 - \bar{y}_2|^2}{\epsilon} + 2\delta e^{\lambda(T-\bar{t})} \left(1 + |\bar{y}_1|^2 + |\bar{y}_2|^2 \right) \right),
\end{aligned}$$

where the third inequality follows from Assumption 2.13 and the linear growth of the function $\bar{\mathbf{b}}$, and C is a positive constant that does not depend on $\epsilon, \delta, \lambda, \beta$. Therefore, (2.40) implies that (2.39) holds because a is arbitrary. Thus, $v_1(t, \tilde{y}) \leq v_2(t, \tilde{y})$ for all $(t, \tilde{y}) \in [0, T] \times \mathbb{R}^{n+d+\bar{d}}$. \square

3. Example and numerical results. In this section, we analyze a classical problem in finance to illustrate the adaptive robust control framework. We derive the optimal acquisition strategy for an agent who employs market orders to purchase a large block of shares in an order driven market; see Cartea et al. (2015).

3.1. Optimal acquisition. At time $t = 0$, an investor must purchase $Q_0 > 0$ shares by the terminal date $T > 0$. The investor sends buy market orders to the limit order book (LOB) of the equity exchange at the speed ν_t . The controlled inventory target Q_t^ν denotes the remaining shares to be purchased over the remaining trading horizon $[t, T]$ for $t \geq 0$, and the target satisfies the SDE $dQ_t^\nu = -\nu_t dt$.

The investor's orders have temporary price impact; i.e., the orders receive worse prices than the midprice, denoted by S_t , because market orders walk the LOB. The price impact is temporary (i.e., the LOB replenishes after each trade), and we assume that the impact is linear in the speed of trading. For example, over a small time-step Δt the investor purchases $\nu_t \Delta t$ shares, and instead of paying the midprice S_t per share, the investor pays $(S_t + k \nu_t) \nu_t \Delta t$, where $k \geq 0$ is the temporary price impact parameter; see Cartea et al. (2015). As the investor buys shares, the process X^ν keeps track of the cumulative cost, which satisfies the SDE $dX_t^\nu = (S_t + k \nu_t) \nu_t dt$. We assume that the midprice of the stock satisfies the SDE

$$dS_t = \theta^* dt + \sigma dB_t^{\mathbb{P}_{\theta^*}} + J_{1+M_t} dM_t,$$

where M is a Poisson process with intensity λ , $M_0 = 0$, and $\{J_1, J_2, \dots\}$ are identically independent random variables that take the value ϵ with probability 1/2 and the value $-\epsilon$ with probability 1/2, where $\epsilon > 0$. The estimate $\hat{\theta}$ satisfies the SDE

$$d\hat{\theta}_t = \beta_t \left[\frac{\tilde{\theta}_t - \hat{\theta}_t}{\sigma^2} dt + \frac{1}{\sigma} dB_t^{\mathbb{P}_{\hat{\theta}}} + \frac{J}{\sigma^2} dM_t \right],$$

where θ^* is an unknown value and $\beta_t = \sigma/(1+t)$ is the learning rate; see estimator (1.9) in Bhudisaksang and Cartea (2021). This choice of β_t simplifies the calculations in our example.

The investor's performance criterion is

$$(3.1) \quad V^\nu = \mathbb{E}^\mathbb{P} [X_T + Q_T (S_T + \eta Q_T)] ,$$

where $\eta \geq 0$ is a terminal penalty parameter, and the investor's adaptive robust problem is given by

$$(3.2) \quad V(t, x, S, \hat{\theta}, q) := \inf_{\nu \in \mathcal{A}} \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^\mathbb{P} [X_T + Q_T (S_T + \eta Q_T)] ,$$

where $\hat{\theta}$ is the estimate of the unknown drift parameter θ^* . Here, the uncertainty set for the estimator of θ^* is

$$(3.3) \quad G(t, \hat{\theta}) = [\hat{\theta} - c/\sqrt{1+t}, \hat{\theta} + c/\sqrt{1+t}] ,$$

where the uncertainty parameter $c > 0$ is a constant and it is easy to check that the function G satisfies Assumptions 2.14 and 2.16. The set of admissible strategies is

$$\mathcal{A} = \left\{ \nu = (\nu_t)_{\{0 \leq t \leq T\}} \mid \nu \text{ is progressively measurable, } \sup_{\mathbb{P} \in \mathcal{P}(0, S_0, \hat{\theta}_0)} \mathbb{E}^\mathbb{P} \left[\int_0^T |\nu_t|^2 dt \right] < \infty \right\} .$$

Next, we show that Assumption 2.4 holds. Use Gronwall's inequality and the definition of the function G to write

$$(3.4) \quad \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^\mathbb{P} \left[\int_t^T (\hat{\theta}_u^2 + S_u^2) du \right] < \infty$$

for all $t \in [0, T]$ and $x = (S_t, \hat{\theta}_t) \in \mathbb{R}^2$. Let $\nu \in \mathcal{A}$; then

$$(3.5) \quad \begin{aligned} \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^\mathbb{P} [X_T] &= \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^\mathbb{P} \left[x + \int_t^T |S_u + k \nu_u| |\nu_u| du \right] \\ &\leq x + \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \left(\mathbb{E}^\mathbb{P} \left[\int_t^T S_u^2 du \right] \mathbb{E}^\mathbb{P} \left[\int_t^T \nu_u^2 du \right] + \mathbb{E}^\mathbb{P} \left[\int_t^T k \nu_u^2 du \right] \right) \\ &< \infty , \end{aligned}$$

where the first inequality follows from the Cauchy inequality and the second inequality follows from (3.4). Next, apply a similar argument to obtain

$$(3.6) \quad \sup_{\mathbb{P} \in \mathcal{P}(t, x)} \mathbb{E}^\mathbb{P} [|Q_T| |S_T + \eta Q_T|] < \infty .$$

Hence, the functions J and w are finite from (3.5) and (3.6). Now, we show that the function V is finite. Let $(t, x, y) \in [0, T] \times \mathbb{R}^{1+d} \times \mathbb{R}^{\tilde{d}}$, for each $\nu \in \mathcal{A}$, we notice that $V(t, x, y) \leq w(t, x, y, \nu) < \infty$. To show that $V(t, x, y) > -\infty$ we consider

$$\begin{aligned} X_T + Q_T (S_T + \eta Q_T) &\geq x + \int_t^T (S_u + k \nu_u) \nu_u du - \frac{S_T^2}{4\eta} \\ &\geq x - \int_t^T \frac{S_u^2}{4k} du - \frac{S_T^2}{4\eta} . \end{aligned}$$

Therefore,

$$V(t, x, y) \geq \sup_{\mathbb{P} \in \mathcal{P}(t, x, G)} \mathbb{E}^{\mathbb{P}} \left[x - \int_t^T \frac{S_u^2}{4k} du - \frac{S_T^2}{4\eta} \right] > -\infty,$$

where the last inequality follows from (3.4). Then, from Theorem 2.18, the value function V satisfies the HJBI

$$\begin{aligned} \partial_t V + \inf_{\nu} \sup_{\tilde{\theta} \in \tau(t, \hat{\theta})} & \left[(S + k\nu) \nu \partial_x V - \nu \partial_q V + \tilde{\theta} \partial_S V \right. \\ & \left. + \frac{\sigma^2}{2} \partial_{SS} V + \frac{\beta_t (\tilde{\theta} - \hat{\theta})}{\sigma^2} \partial_{\tilde{\theta}} V + \beta_t \partial_{\tilde{\theta} S} V + \frac{\beta_t^2}{\sigma^2} \partial_{\tilde{\theta} \tilde{\theta}} V \right] \\ & + \frac{\lambda}{2} \left[V(t, x, S, \hat{\theta} + \beta_t \epsilon / \sigma^2, q) + V(t, x, S, \hat{\theta} - \beta_t \epsilon / \sigma^2, q) - 2V \right] = 0 \end{aligned}$$

subject to the terminal condition $V(T, x, S, \hat{\theta}, q) = x + S q + \eta q^2$.

The proposition below shows the optimal acquisition speed.

PROPOSITION 3.1. *The optimal speed of trading that solves the agent's acquisition problem (see (3.2)) is*

$$(3.7) \quad \nu_t^* = \frac{q}{T - t + k/\eta} + \frac{\partial_q h(t, \hat{\theta}, q)}{2k},$$

where the function h satisfies

$$\begin{aligned} (3.8) \quad \partial_t h - \frac{1}{4k} (\partial_q h)^2 + \left(\hat{\theta} + \frac{c}{\sqrt{1+t}} \right) q + \frac{c \partial_{\hat{\theta}} h}{\sigma(t+1) \sqrt{1+t}} \\ + \frac{\lambda}{2} \left(h(t, \hat{\theta} + \beta_t \epsilon / \sigma^2, q) + h(t, \hat{\theta} - \beta_t \epsilon / \sigma^2, q) - 2h \right) + \frac{\partial_{\hat{\theta} \hat{\theta}} h}{(t+1)^2} = 0 \\ \text{if } q + \frac{\partial_{\hat{\theta}} h}{\sigma(t+1)} \geq 0; \\ \partial_t h - \frac{1}{4k} (\partial_q h)^2 + \left(\hat{\theta} - \frac{c}{\sqrt{1+t}} \right) q - \frac{c \partial_{\hat{\theta}} h}{\sigma(t+1) \sqrt{1+t}} \\ + \frac{\lambda}{2} \left(h(t, \hat{\theta} + \beta_t \epsilon / \sigma^2, q) + h(t, \hat{\theta} - \beta_t \epsilon / \sigma^2, q) - 2h \right) + \frac{\partial_{\hat{\theta} \hat{\theta}} h}{(t+1)^2} = 0 \quad \text{otherwise,} \end{aligned}$$

with terminal condition $h(T, \hat{\theta}, q) = 0$. Moreover, the optimal speed of trading of the adaptive agent (i.e., $c = 0$) is

$$(3.9) \quad \nu^* = \frac{q}{T - t + k/\eta} + \frac{(T - t)(T - t + 2k/\eta) \hat{\theta}}{4k(T - t + k/\eta)}.$$

Proof. It is straightforward to show that the optimal speed of trading in feedback form is given by $\nu^* = (\partial_q V - S)/2k$. Next, use the ansatz

$$V(t, x, S, \hat{\theta}, q) = x + S q + \tilde{h}(t, q, \hat{\theta})$$

and reduce the HJBI equation to

$$(3.10) \quad \begin{aligned} \partial_t \tilde{h} - \frac{1}{4k} \left(\partial_q \tilde{h} \right)^2 + \frac{\lambda}{2} \left(\tilde{h}(t, \hat{\theta} + \beta_t \epsilon / \sigma^2, q) + \tilde{h}(t, \hat{\theta} - \beta_t \epsilon / \sigma^2, q) - 2\tilde{h} \right) \\ + \sup_{\tilde{\theta} \in \tau(t, \hat{\theta})} \left[\tilde{\theta} q + \frac{(\tilde{\theta} - \hat{\theta})}{\sigma(t+1)} \partial_{\tilde{\theta}} \tilde{h} + \frac{1}{(t+1)^2} \partial_{\tilde{\theta}\tilde{\theta}} \tilde{h} \right] = 0, \end{aligned}$$

where $\tilde{h}(T, \hat{\theta}, q) = \eta q^2$. The supremum attains at either $\tilde{\theta} = \hat{\theta} + c/\sqrt{1+t}$ or $\tilde{\theta} = \hat{\theta} - c/\sqrt{1+t}$. Thus, we write

$$(3.11) \quad \begin{aligned} \partial_t \tilde{h} - \frac{1}{4k} \left(\partial_q \tilde{h} \right)^2 + \left(\hat{\theta} + \frac{c}{\sqrt{1+t}} \right) q + \frac{c \partial_{\tilde{\theta}} \tilde{h}}{\sigma(t+1) \sqrt{1+t}} \\ + \frac{1}{2} \left(\tilde{h}(t, \hat{\theta} + \beta_t \epsilon / \sigma^2, q) + \tilde{h}(t, \hat{\theta} - \beta_t \epsilon / \sigma^2, q) - 2\tilde{h} \right) + \frac{\partial_{\tilde{\theta}\tilde{\theta}} \tilde{h}}{(t+1)^2} = 0 \\ \text{if } q + \frac{\partial_{\tilde{\theta}} \tilde{h}}{\sigma(t+1)} \geq 0; \\ \partial_t \tilde{h} - \frac{1}{4k} \left(\partial_q \tilde{h} \right)^2 + \left(\hat{\theta} - \frac{c}{\sqrt{1+t}} \right) q - \frac{c \partial_{\tilde{\theta}} \tilde{h}}{\sigma(t+1) \sqrt{1+t}} \\ + \frac{1}{2} \left(\tilde{h}(t, \hat{\theta} + \beta_t \epsilon / \sigma^2, q) + \tilde{h}(t, \hat{\theta} - \beta_t \epsilon / \sigma^2, q) - 2\tilde{h} \right) + \frac{\partial_{\tilde{\theta}\tilde{\theta}} \tilde{h}}{(t+1)^2} = 0 \quad \text{otherwise.} \end{aligned}$$

Substitute

$$\tilde{h}(t, q, \hat{\theta}) = \frac{k q^2}{T - t + k/\eta} + h(t, q, \hat{\theta})$$

into the equation above to obtain (3.8), and it is straightforward to write the optimal speed of trading as in (3.9).

When the uncertainty parameter is $c = 0$, the system of PDEs in (3.8) reduces to

$$(3.12) \quad \partial_t h - \frac{1}{4k} (\partial_q h)^2 + \hat{\theta} q + \frac{\partial_{\hat{\theta}\hat{\theta}} h}{(t+1)^2} + \frac{\lambda}{2} \left(h(t, \hat{\theta} + \beta_t \epsilon / \sigma^2, q) + h(t, \hat{\theta} - \beta_t \epsilon / \sigma^2, q) - 2h \right) = 0$$

with $h(T, \hat{\theta}, q) = 0$. Next, substitute the ansatz

$$h(t, \hat{\theta}, q) = q \tilde{h}_1(t, \hat{\theta}) + \tilde{h}_2(t, \hat{\theta})$$

into (3.12) to obtain the coupled PDEs:

$$\begin{aligned} \partial_t \tilde{h}_1 - \frac{\tilde{h}_1}{T - t + k/\eta} + \frac{1}{(t+1)^2} \partial_{\hat{\theta}\hat{\theta}} \tilde{h}_1 + \hat{\theta} \\ + \frac{\lambda}{2} \left(\tilde{h}_1(t, \hat{\theta} + \beta_t \epsilon / \sigma^2) + \tilde{h}_1(t, \hat{\theta} - \beta_t \epsilon / \sigma^2) - 2\tilde{h}_1 \right) = 0, \\ \partial_t \tilde{h}_2 - \frac{\tilde{h}_1^2}{4k} + \frac{1}{(t+1)^2} \partial_{\hat{\theta}\hat{\theta}} \tilde{h}_2 + \frac{\lambda}{2} \left(\tilde{h}_2(t, \hat{\theta} + \beta_t \epsilon / \sigma^2) + \tilde{h}_2(t, \hat{\theta} - \beta_t \epsilon / \sigma^2) - 2\tilde{h}_2 \right) = 0. \end{aligned}$$

Let $\tilde{h}_1(t, \hat{\theta}) = \hat{\theta} f^{\hat{\theta}q}(t) + f^q(t)$ and $\tilde{h}_2(t, \hat{\theta}) = \hat{\theta}^2 f^{\hat{\theta}\hat{\theta}}(t) + \hat{\theta} f^{\hat{\theta}}(t) + f^{\hat{\theta}}(t)$, where all the functions depend only on t , and obtain the ordinary differential equation

$$(T - t + k/\eta) \partial_t f^{\hat{\theta}q} - f^{\hat{\theta}q} = -(T - t + k/\eta).$$

Integrate the equation above from t to T to obtain

$$\begin{aligned} f^{\hat{\theta}, q}(t) &= \frac{1}{T-t+k/\eta} \int_t^T (T-s+k/\eta) ds \\ (3.13) \quad &= \frac{(T-t)k/\eta + \frac{1}{2}(T-t)^2}{T-t+k/\eta}. \end{aligned}$$

Substitute the expression above into the optimal speed of trading to obtain (3.9). \square

Note that when the value of the uncertainty parameter c is zero, the optimal strategy does not depend on the size of the jump ϵ or on the arrival rate parameter λ . When $c > 0$, we cannot find a closed-form solution for the function h , so we employ the Crank–Nicolson finite-difference method to solve the PDE in (3.8) and obtain the optimal speed of trading in (3.7). For the set of parameters we study, our numerical solution shows that the term $\partial_q h(t, \hat{\theta}, q)/2k$ in (3.7) is greater (smaller) than the second term on the right-hand side of the optimal speed with $c = 0$ in (3.9) when $q > 0$ ($q < 0$). Also, we find that all else being equal, as the bands of the set G widen (i.e., the larger is the value of the uncertainty parameter c), the value $|\partial_q h(t, \hat{\theta}, q)|$ increases.

Therefore, as the investor is more uncertain about the estimate of the drift of the asset, the speed of trading is adjusted as follows. When the remaining target to purchase is positive (negative), i.e., $q > 0$ ($q < 0$), the investor speeds up the purchase (sales) of shares as the value of the uncertainty parameter c increases. Note that if $q < 0$, the investor holds more shares than the original target Q_0 , so the investor must sell the excess shares before the end of the trading horizon.

In other words, as the investor perceives more uncertainty about the estimate of the drift parameter, the conservative (i.e., robust) strategy is to accelerate the purchase of the shares when the target is positive and to accelerate the sales of shares when the target is negative.

3.2. Performance of adaptive robust strategies. In this subsection, we compare the performance of the adaptive robust strategies with that of three strategies in which the agent knows the true value of the drift parameter, employs a wrong value of the drift parameter, and employs a robust strategy. In the robust strategy, the agent uses the framework derived above but does not learn the value of the unknown parameter. Instead, the agent assumes that the true parameter θ^* lies in the interval $[\underline{\theta}, \bar{\theta}]$ and solves

$$(3.14) \quad v(t, x, y) = \inf_{\alpha \in \mathcal{A}_0} \sup_{\mathbb{P} \in \mathcal{P}} J(t, x, y, \mathbb{P}, \alpha),$$

where the set \mathcal{P} contains all probability measures $\mathbb{P}_{\hat{\theta}}$ such that $\hat{\theta}_u \in [\underline{\theta}, \bar{\theta}]$ for all $u \in [t, T]$.

3.2.1. Adaptive robust optimal acquisition. The terminal time is $T = 20$ minutes, and other model parameters are

$$\begin{aligned} Q_0 &= 10^5, X_0 = 0, S_0 = 10, \theta^* = 0.09, \hat{\theta}_0 = -0.03, \sigma = 0.2, \eta = 10^{-3}, k = 10^{-4}, \\ c &= 0.02, \epsilon = 0.1, \end{aligned}$$

and for the robust strategy, the set of possible values of the parameter θ^* is $[-0.1, 0.2]$.

When the agent believes that θ is the true drift parameter, the optimal speed to trade is given by

$$(3.15) \quad \nu_t^* = \frac{q}{T-t+k/\eta} + \frac{(T-t)(T-t+2k/\eta)\theta}{4(T-t+k/\eta)k}.$$

We compare the adaptive robust strategy in (3.7) with the following strategies: (i) the agent employs (3.15) with the true drift $\theta = \theta^*$, (ii) the agent employs the robust strategy without learning in (3.14), in which case the speed of trading is in (3.15) with $\theta = \bar{\theta} = 0.2$, and (iii) the agent employs (3.15) with the wrong drift parameter $\theta = -0.03$.

We discretize the time space into 2,000 time-steps and employ 1,000 simulations to analyze the performance of the four strategies. The left-hand panel of Figure 1 shows the mean acquisition cost of the strategies. At the terminal date, the lowest mean cost is that of the strategy with the true drift parameter, followed by the adaptive robust strategy—see third column (for $c = 0.02$) in Table 1 below.

The right-hand panel of Figure 1 shows the standard deviation of the acquisition costs. For most of the trading horizon, the adaptive strategy shows the highest value of the standard deviation of the acquisition costs—the standard deviation peaks half-way through and then declines. This sharp increase, followed by a sharp decrease, in the standard deviation of the acquisition costs results from the strategy learning the correct value of the parameter. At every time-step, the adaptive strategy refines the estimate of the drift parameter; thus the speed of trading readjusts, so the costs are more volatile during the first half of the trading horizon.

Figure 2 shows the mean target inventory for the four strategies. Recall that the true, wrong, and robust strategies are deterministic (see (3.15)), so for each strategy the mean of the remaining target is the same as the remaining target for each simulation. Observe that the robust strategy overshoots the target, that is, the robust

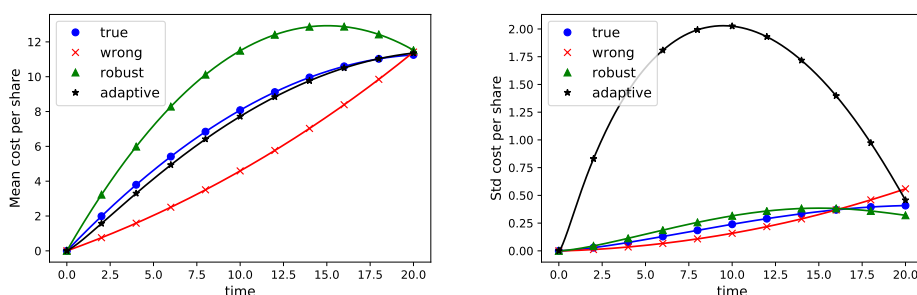


FIG. 1. Mean and standard deviation of the acquisition costs per share.

TABLE 1
Performance of strategies.

	$c = 0.01$		$c = 0.02$		$c = 0.03$		$c = 0.04$	
θ	mean	std	mean	std	mean	std	mean	std
true	11.26	0.41	11.26	0.41	11.26	0.41	11.26	0.41
false	11.41	0.55	11.41	0.55	11.41	0.55	11.41	0.55
robust	11.55	0.32	11.55	0.32	11.55	0.32	11.55	0.32
adaptive	11.38	0.47	11.38	0.46	11.36	0.47	11.37	0.46

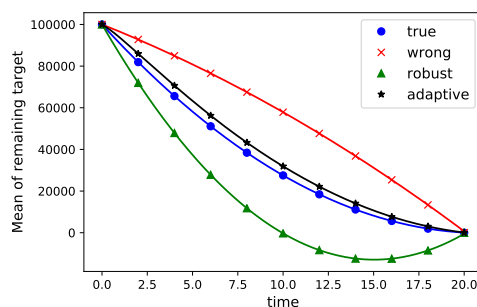


FIG. 2. Target inventory.

TABLE 2
Performance of strategies.

	$T = 10$		$T = 20$		$T = 30$		$T = 40$	
θ	mean	std	mean	std	mean	std	mean	std
true	10.86	0.32	11.25	0.42	11.55	0.45	11.79	0.48
false	10.85	0.36	11.40	0.58	11.99	0.70	12.62	0.84
robust	10.98	0.28	11.54	0.33	12.11	0.40	12.76	0.56
adaptive	10.89	0.36	11.36	0.48	11.77	0.52	12.13	0.60

strategy purchases more than the target and unwinds the extra shares at the end of the trading horizon. It is “optimal” to speculate because the agent expects an increase in the price of the asset—recall that the agent’s estimate of the drift parameter is $\bar{\theta} = 0.2$. Finally, although the mean target is nonnegative for the adaptive robust strategy, we note that in 300 simulations there was at least one time-step when the target became negative (because the agent’s estimate of the drift was high enough to justify speculative purchases).

We repeat the analysis for various values of the uncertainty parameter c in the set G (see (3.3)) and report the findings in Table 1; recall that the case discussed in the figures above is when $c = 0.2$. As the value of the uncertainty parameter c increases, the agent is less certain about the estimate of the drift of the stock price, so the agent speeds up the purchase of the target inventory (see Figure 2 above), and the performance of the strategy worsens.

We repeat the analysis for various values of the terminal date of the trading horizon; see Table 2. Here, the initial acquisition target is $Q_0 = T \times 5 \times 10^3$ to keep the ratio Q_0/T constant.² Our results show that as the terminal date increases, the adaptive robust strategy performs better, relative to the false and robust strategy, because there is more time for the agent to learn the value of the unknown drift parameter. Finally, note that the jump component has a similar effect as that of the volatility because the jump component is symmetric around zero.

4. Conclusions and future research. We proposed a continuous-time version of the adaptive robust methodology in Bielecki, Chen, and Cialenco (2017). In our extension, the underlying process follows a jump-diffusion process with unknown drift, and the agent is continuously estimating the drift while making optimal decisions that

²Note that for Figures 1 and 2, the terminal date is $T = 20$. Therefore, the initial acquisition target shown in those figures is $Q_0 = 20 \times 5 \times 10^3 = 100,000$.

are time-consistent. Our methodology has a balance between making a model robust to misspecification and learning an unknown parameter. Our result is general, and the value function is characterized as the solution of a PDE. As an particular example we considered an optimal execution problem when the agent purchases a large amount of shares and her trades walk the LOB of the exchange. When the agent has enough time to learn the value of the unknown parameter, we showed that the adaptive robust strategy performs better (lower average and lower variance of acquisition costs) than when the agent employs a robust strategy or uses the incorrect parameter estimate.

We propose a number of extensions for future research. Assume that the observed process is noisy, e.g., the stock process has a stochastic drift, which follows an Ornstein–Uhlenbeck process with unknown mean. Another extension is to include a penalty term in the expectation of the objective function instead of considering the unknown parameter inside the confidence region. To this end, one can explore the approach in Bion-Nadal (2009) to impose a penalty for choosing an alternative model in the adaptive robust framework.

Appendix A. Proofs.

A.1. Proof of Lemma 2.5. We follow the proofs of Neufeld and Nutz (2014) and Fadina et al. (2019). From Theorem 2.6 in Neufeld and Nutz (2014), the set $\mathfrak{P}_{sem}^{ac}(\Omega)$ is Borel measurable in $\mathfrak{P}(\Omega)$. Therefore, the set

$$\{(\omega, t, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}_{sem}^{ac}(\Omega) \mid \mathbb{P}(\tilde{X}_t = \omega_t) = 1\}$$

is Borel measurable. Recall that the processes $\gamma^\mathbb{P}$, a , and $F_{\tilde{\omega}, t}^\mathbb{P}$ are the drift characteristic, the volatility characteristic, and the jump characteristic, respectively, of the process X_t under the probability measure \mathbb{P} . From Theorem 2.6 in Neufeld and Nutz (2014),

$$(\tilde{\omega}, s, \mathbb{P}) \mapsto (\tilde{\omega}, s, \gamma_s^\mathbb{P}(\tilde{\omega}), a_s(\tilde{\omega}), F_{\tilde{\omega}, t}^\mathbb{P})$$

is Borel measurable. From Assumption 2.3, the set

$$\{(\tilde{\omega}, s, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}_{sem}^{ac}(\Omega) \mid s > t, (\gamma_s^\mathbb{P}(\tilde{\omega}), a_s(\tilde{\omega}), F_{\tilde{\omega}, t}^\mathbb{P}) \in \mathbf{b}^*(s, \tilde{\omega}_s) \times a_s(\tilde{\omega}_s) \times L_s(\tilde{\omega}_s)\}$$

is Borel measurable. Therefore, the set

$$E := \left\{ (\omega, t, \mathbb{P}, \tilde{\omega}, s) \in \Omega \times [0, T] \times \mathfrak{P}_{sem}^{ac}(\Omega) \times \Omega \times [0, T] \mid s > t, \right. \\ \left. \mathbb{P}(\{\tilde{X}_s = \omega_t, 0 \leq s \leq t\}) = 1, (\gamma_s^\mathbb{P}(\tilde{\omega}), a_s(\tilde{\omega}), F_{\tilde{\omega}, t}^\mathbb{P}) \in \mathbf{b}^*(s, \tilde{\omega}_s) \times a_s(\tilde{\omega}_s) \times L_s(\tilde{\omega}_s) \right\}$$

is Borel measurable. Then, by Fubini's theorem (see Appendix B.1 below), we obtain the Borel measurable mapping

$$(\omega, t, \mathbb{P}, \tilde{\omega}) \mapsto \int_0^T \mathbf{1}_E(\omega, t, \mathbb{P}, \tilde{\omega}, s) \mathbf{1}_{(t, T]}(s) ds.$$

The set

$$E' := \left\{ (\omega, t, \mathbb{P}, \tilde{\omega}) \in \Omega \times [0, T] \times \mathfrak{P}_{sem}^{ac}(\Omega) \times \Omega \mid \int_0^T \mathbf{1}_E(\omega, t, \mathbb{P}, \tilde{\omega}, s) \mathbf{1}_{(t, T]}(s) ds = T - t \right\}$$

is Borel measurable because the inverse image of a Borel measurable mapping is also Borel measurable. Therefore, by a monotone class argument, we have that

$(\omega, t, \mathbb{P}) \mapsto E^{\mathbb{P}}[\mathbf{1}_{E'}(\omega, t, \mathbb{P}, \cdot)]$ is a Borel measurable function. Therefore, the following result completes the proof:

$$\begin{aligned} & \{(\omega, t, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\} \\ &= \left\{ (\omega, t, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}_{sem}^{ac}(\Omega) \mid E^{\mathbb{P}}[\mathbf{1}_{E'}(\omega, t, \mathbb{P}, \cdot)] = 1 \right\}, \end{aligned}$$

where the set on the right-hand side of the equality above is Borel measurable because $(\omega, t, \mathbb{P}) \mapsto E^{\mathbb{P}}[\mathbf{1}_{E'}(\omega, t, \mathbb{P}, \cdot)]$ is a Borel measurable function.

A.2. Proof of Lemma 2.6. From the definition of a probability measure $\mathbb{P}^{\tau, \omega}$, we have that for \mathbb{P} -a.s. $\bar{\omega} \in \Omega$,

$$\begin{aligned} & E^{\mathbb{P}} \left[\int_{\tau}^T f(s, Y_s, \alpha_s) ds + g(Y_T) \mid \mathcal{F}_{\tau} \right] (\bar{\omega}) \\ &= E^{\mathbb{P}^{\tau(\bar{\omega}), \bar{\omega}}} \left[\int_{\tau(\bar{\omega})}^T f \left(s, (Y_s)^{\tau(\bar{\omega}), \bar{\omega}}, \alpha_s^{\tau(\bar{\omega}), \bar{\omega}} \right) ds + g \left(Y_T^{\tau(\bar{\omega}), \bar{\omega}} \right) \right] \\ &= E^{\mathbb{P}^{\tau(\bar{\omega}), \bar{\omega}}} \left[\int_{\tau(\bar{\omega})}^T f \left(s, Y_s, \alpha_s^{\tau(\bar{\omega}), \bar{\omega}} \right) ds + g(Y_T) \right], \end{aligned}$$

where the last equation follows because of the path uniqueness of the process Y .

A.3. Proof of Lemma 2.7. First, due to Assumption 2.4, the functions J , w , and v are finite. Next, we check the regularity of the function J . Use a similar argument to that in Lemma 3.3 in Pham and Wei (2017) to show that for a fixed measure \mathbb{P} the function J is continuous with respect to t, x, y, α . Write the function $J(t, x, y, \mathbb{P}, \alpha)$ as

$$J(t, x, y, \mathbb{P}, \alpha) = \int_{\Omega} \int_t^T f(s, Y_s(\omega), \alpha_s(\omega)) ds + g(Y_T(\omega)) \mathbb{P}(d\omega).$$

For fixed t, x, y, α , the mapping $\omega \mapsto Y_s(\omega)$, $\omega \mapsto \alpha_s(\omega)$ is Borel measurable for all $s \in [t, T]$. Apply Corollary 7.29.1 in Bertsekas and Shreve (1996) to show that the function J is Borel measurable with respect to \mathbb{P} . Therefore, by the Carathéodory theorem (see Appendix B.6) the function J is Borel measurable. Recall that the set $\{(\omega, t, \mathbb{P}) \in \Omega \times [0, T] \times \mathfrak{P}(\Omega) \mid \mathbb{P} \in \mathcal{P}(t, \omega(t))\}$ is Borel measurable. Therefore, all conditions in Lemma A.1 hold, which completes the proof.

A.4. Proof of Lemma A.1.

LEMMA A.1. *Let A, X, Y, Z be metrizable separable spaces. Let J be a Borel measurable function on $X \times Y \times Z$, and let $J(x, y, z)$ be a continuous function in the variable y . Define $w(x, y) := \sup_{z \in B(x)} J(x, y, z)$ and $v^*(x) := \inf_{y \in A} w(x, y)$. Assume that the sets $\{(x, z) \mid z \in B(x)\}$ and $\{(x, y) \mid y \in A\}$ are Borel measurable and w and v^* are finite. Then we have the following results:*

- (a) *The function w is upper semianalytic and $\mathcal{L}_X \otimes B_Y$ -measurable.*
- (b) *The function v^* is universally measurable, and for any $\epsilon > 0$ there exists a universally measurable ϵ -minimax strategy.*

This lemma is a modification of Theorem 1 in Nowak (2010). We replace the σ -compact assumption on the set $B(x)$ with the continuity of the function J on the second variable because the σ -compact assumption is not satisfied in our framework.

(a) From a standard measurable selection, the function w is upper semianalytic, and $w(\cdot, y)$ is \mathcal{L}_X -measurable. Let

$$f_n(x, y, z, b) := J(x, b, z) + n d(y, b), \quad J_n(x, y, z) := \inf_{b \in A} f_n(x, y, z, b),$$

and define the function w_n as a lower envelope for the function w ,

$$w_n(x, y) := \sup_{z \in B(x)} J_n(x, y, z),$$

where d is a metric distance in the space A . Note that the function f_n is continuous in the variables y and b and that the space A is separable. Then, write the function J_n as the infimum of the countable measurable function

$$J_n(x, y, z) = \inf_{k \in \mathbb{N}} f_n(x, y, z, b_k),$$

where $\{b_1, \dots, b_k, \dots\}$ is the countable dense subset in A . Therefore, J_n is Borel measurable and continuous in y . By the DPP, the function w_n is upper semianalytic, and hence $w_n(\cdot, y)$ is \mathcal{L}_X -measurable. Now, use the inequality

$$\begin{aligned} |w_n(x, y) - w_n(x, y')| &\leq \sup_{z \in B(x)} |J_n(x, y, z) - J_n(x, y', z)| \\ &\leq \sup_{z \in B(x)} \sup_{b \in A} |f_n(x, y, z, b) - f_n(x, y', z, b)| \leq n d(y, y') \end{aligned}$$

to check that the function w_n is continuous in y .

Therefore, by the Carathéodory Theorem B.6, the function w_n is $\mathcal{L}_X \otimes B_Y$ -measurable. From the definition of the functions f_n , we have that $J_n \leq J_{n+1}$, $w_n \leq w_{n+1}$, and $w_n \leq w$. Thus, $\lim_{n \rightarrow \infty} w_n(x, y) \leq w(x, y)$. Next, we prove that the reverse inequality holds by showing that

$$\lim_{n \rightarrow \infty} J_n(x, y, z) = J(x, y, z).$$

Assume there exists x, y, z such that $\lim_{n \rightarrow \infty} J_n(x, y, z) < J(x, y, z)$. Therefore, there exists an $\epsilon > 0$ and $N \in \mathbb{N}$ such that $J_n(x, y, z) + \epsilon < J(x, y, z)$ for all $n \geq N$. Hence, there exists a sequence \tilde{b}_n such that

$$J(x, \tilde{b}_n, z) + n d(y, \tilde{b}_n) + \frac{\epsilon}{2} < J(x, y, z)$$

for all $n \geq N$. Because the function $J(x, \cdot, z)$ is continuous, the sequence $\tilde{b}_n \in A$ converges to y , and this leads to a contradiction. Therefore, the limit of function J_n is equal to J , and hence

$$\lim_{n \rightarrow \infty} \sup_{z \in B(x)} J_n(x, y, z) \geq \sup_{z \in B(x)} \lim_{n \rightarrow \infty} J_n(x, y, z) = \sup_{z \in B(x)} J(x, y, z) = w(x, y).$$

Thus, $\lim_{n \rightarrow \infty} w_n(x, y) \geq w(x, y)$, and $\lim_{n \rightarrow \infty} w_n(x, y) = w(x, y)$, so the function w is also $\mathcal{L}_X \otimes B_Y$ -measurable. The rest of the proof follows from Theorem 1 in Nowak (2010).

Appendix B.

THEOREM B.1 (Fubini's theorem). *Let μ and ν be σ -finite outer measures on X and Y , respectively. For any non-negative $\mu \times \nu$ -measurable function f ,*

$$\begin{aligned} x &\mapsto f(x, y) \text{ is } \mu\text{-measurable for } \nu\text{-a.e.}, \\ y &\mapsto \int_X f(x, y) d\mu(x) \text{ is } \mu\text{-measurable}. \end{aligned}$$

DEFINITION B.2 (Hausdorff metric). *Let (M, d) be a metric space, and let X, Y be nonempty subsets of the space M . The Hausdorff metric is given by*

$$d_{\text{haus}}(X, Y) := \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}.$$

DEFINITION B.3 (analytic set and analytically measurable function).

- (i) *Let E be a Borel space; then a subset B is an analytic set in E if there is another Borel space F and a Borel subset $A \subseteq E \times F$ such that $B = \pi_E(A)$. A subset $C \subseteq E$ is coanalytic if its complement C^c is analytic.*
- (ii) *A function $g : E \rightarrow \mathbb{R} = \mathbb{R} \cup \infty$ is upper semianalytic if $\{x \in E : g(x) > c\}$ is analytic for every $c \in \mathbb{R}$.*
- (iii) *Let E be a Borel set and $\mathcal{A}(E)$ denote a σ -field generated by all analytic subsets. A function $f : E \rightarrow F$, where F is a Borel set, is analytically measurable if $f^{-1}(C) \in \mathcal{A}(E)$ for every $C \in \mathcal{B}(F)$.*

DEFINITION B.4. *Let (Ω, \mathcal{F}) be a measurable space; the universal completion of \mathcal{F} is the σ -field defined as the intersection of $\mathcal{F}^{\mathbb{P}}$ for all probability measures $\mathbb{P} \in \mathcal{P}(\Omega)$ on (Ω, \mathcal{F}) , i.e.,*

$$\mathcal{F}^U := \bigcap_{\mathbb{P} \in \mathcal{P}(\Omega)} \mathcal{F}^{\mathbb{P}}.$$

A function f from (Ω, \mathcal{F}) to E is called universally measurable if $f^{-1}(A) \in \mathcal{F}^U$ for each $A \in \mathcal{B}(E)$.

DEFINITION B.5. *Let (S, Σ) be a measurable space, and let X and Y be topological spaces. A function $f : S \times X \rightarrow Y$ is a Carathéodory function if*

1. *for each $x \in X$, the function $f^x = f(\cdot, x) : S \rightarrow Y$ is (Σ, \mathcal{B}_Y) -measurable;*
2. *for each $s \in S$, the function $f_s = f(s, \cdot) : X \rightarrow Y$ is continuous.*

THEOREM B.6. *Let (S, Σ) be a measurable space, X a separable metrizable space, and Y a metrizable space. Then every Carathéodory function $f : S \times X \rightarrow Y$ is jointly measurable.*

REFERENCES

- K. BANNÖR, R. KIESEL, A. NAZAROVA, AND M. SCHERER, *Parametric model risk and power plant valuation*, Energy Econ., 59 (2016), pp. 423–434.
- E. BAYRAKTAR AND S. YAO, *On the robust optimal stopping problem*, SIAM J. Control Optim., 52 (2014), pp. 3135–3175.
- V. BERGEN, M. ESCOBAR, A. RUBTSOV, AND R. ZAGST, *Robust multivariate portfolio choice with stochastic covariance in the presence of ambiguity*, Quant. Finance, 18 (2018), pp. 1265–1294.
- D. BERTSEKAS AND S. SHREVE, *Stochastic Optimal Control: The Discrete-Time Case*, Athena Scientific, Nashua, NH, 1996.
- D. BERTSEKAS AND J. TSITSIKLIS, *Neuro-Dynamic Programming*, Athena Scientific, Nashua, NH, 1996.

- T. BHUDISAKSANG AND Á CARTEA, *Online drift estimation for jump-diffusion processes*, Bernoulli, 27 (2021), pp. 2494–2518.
- S. BIAGINI AND M. Ç. PINAR, *The robust Merton problem of an ambiguity averse investor*, Math. Financ. Econ., 11 (2017), pp. 1–24.
- T. BIELECKI, T. CHEN, AND I. CIALENCO, *Recursive construction of confidence regions*, Electron. J. Stat., 11 (2017), pp. 4674–4700.
- T. R. BIELECKI, T. CHEN, I. CIALENCO, A. COUSIN, AND M. JEANBLANC, *Adaptive robust control under model uncertainty*, SIAM J. Control Optim., 57 (2019), pp. 925–946.
- J. BION-NADAL, *Time consistent dynamic risk processes*, Stochastic Process. Appl., 119 (2009), pp. 633–654.
- Á. CARTEA, R. DONNELLY, AND S. JAIMUNGAL, *Algorithmic trading with model uncertainty*, SIAM J. Financial Math., 8 (2017), pp. 635–671.
- Á. CARTEA AND S. JAIMUNGAL, *Irreversible investments and ambiguity aversion*, Int. J. Theor. Appl. Finance, 20 (2017), 1750044.
- Á. CARTEA, S. JAIMUNGAL, AND J. PENALVA, *Algorithmic and High-Frequency Trading*, Cambridge University Press, Cambridge, UK, 2015.
- Á. CARTEA AND L. SÁNCHEZ-BETANCOURT, *The shadow price of latency: Improving intraday fill ratios in foreign exchange markets*, SIAM J. Financial Math., 12 (2021), pp. 254–294.
- T. CHEN AND M. LUDKOVSKI, *A machine learning approach to adaptive robust utility maximization and hedging*, SIAM J. Financial Math., 12 (2021), pp. 1226–1256.
- S. N. COHEN AND R. J. ELLIOTT, *Stochastic Calculus and Applications*, 2nd ed., Springer, Cham, 2015.
- L. DENIS AND M. KERVAREC, *Optimal investment under model uncertainty in nondominated models*, SIAM J. Control Optim., 51 (2013), pp. 1803–1822.
- I. EKREN, N. TOUZI, AND J. ZHANG, *Optimal stopping under nonlinear expectation*, Stochastic Process. Appl., 124 (2014), pp. 3277–3311.
- L. G. EPSTEIN AND S. JI, *Ambiguous volatility, possibility and utility in continuous time*, J. Math. Econom., 50 (2014), pp. 269–282.
- T. FADINA, A. NEUFELD, AND T. SCHMIDT, *Affine processes under parameter uncertainty*, Probab. Uncertain. Quant. Risk, 4 (2019), 5.
- L. P. HANSEN AND T. J. SARGENT, *Robustness and ambiguity in continuous time*, J. Econom. Theory, 146 (2011), pp. 1195–1223.
- J. HOLLENDER, *Lévy-Type Processes Under Uncertainty and Related Nonlocal Equations*, Technische Universität Dresden, 2016.
- A. ISMAIL AND H. PHAM, *Robust Markowitz mean-variance portfolio selection under ambiguous covariance matrix*, Math. Finance, 29 (2019), pp. 174–207.
- S. JAIMUNGAL AND G. SIGLOCH, *Incorporating risk and ambiguity aversion into a hybrid model of default*, Math. Finance, 22 (2012), pp. 57–81.
- N. E. KAROUI AND X. TAN, *Capacities, Measurable Selection and Dynamic Programming Part II: Application in Stochastic Control Problems*, preprint, arXiv:1310.3364 [math.OC], 2013.
- A. NEUFELD AND M. NUTZ, *Measurability of semimartingale characteristics with respect to the probability law*, Stochastic Process. Appl., 124 (2014), pp. 3819–3845.
- A. NOWAK, *On measurable minimax selectors*, J. Math. Anal. Appl., 366 (2010), pp. 385–388.
- M. NUTZ AND R. VAN HANDEL, *Constructing sublinear expectations on path space*, Stochastic Process. Appl., 123 (2013), pp. 3100–3121.
- H. PHAM, *Optimal stopping of controlled jump diffusion processes: A viscosity solution approach*, J. Math. Syst. Estimation Control, 8 (1998), pp. 1–27.
- H. PHAM, *Continuous-Time Stochastic Control and Optimization with Financial Applications*, Springer Science & Business Media, New York, 2009.
- H. PHAM AND X. WEI, *Dynamic programming for optimal control of stochastic McKean–Vlasov dynamics*, SIAM J. Control Optim., 55 (2017), pp. 1069–1101.
- H. PHAM, X. WEI, AND C. ZHOU, *Portfolio diversification and model uncertainty: A robust dynamic mean-variance approach*, Math. Finance, in press, 2021.
- J. SIRIGNANO AND K. SPILIOPOULOS, *Stochastic gradient descent in continuous time*, SIAM J. Financial Math., 8 (2017), pp. 933–961.
- C. SKIADAS, *Smooth ambiguity aversion toward small risks and continuous-time recursive utility*, J. Political Econ., 121 (2013), pp. 775–792.
- H. M. SONER AND N. TOUZI, *Dynamic programming for stochastic target problems and geometric flows*, J. Eur. Math. Soc. (JEMS), 4 (2002), pp. 201–236.