A MEASURE-VALUED HJB PERSPECTIVE ON BAYESIAN OPTIMAL ADAPTIVE CONTROL

ALEXANDER M. G. COX, SIGRID KÄLLBLAD, AND CHAORUI WANG

Abstract. We consider a Bayesian adaptive optimal stochastic control problem where a hidden static signal has a *non-separable* influence on the drift of a noisy observation. Being allowed to control the specific form of this dependence, we aim at optimising a cost functional depending on the posterior distribution of the hidden signal. Expressing the dynamics of this posterior distribution in the observation filtration, we embed our problem into a genuinely *infinite-dimensional* stochastic control problem featuring *measure-valued martingales*. We address this problem by use of viscosity theory and approximation arguments. Specifically, we show equivalence to a corresponding weak formulation, characterise the optimal value of the problem in terms of a unique continuous viscosity solution of the associated HJB equation, and construct a piecewise constant and arbitrarily-close-to-optimal control to our main problem of study.

1. Introduction

We consider a Bayesian adaptive optimal stochastic control problem in continuous time on an infinite horizon. There is an underlying static signal which cannot be observed directly but only via a noisy observation. More precisely, the drift of this observation process has a non-linear dependence on the signal and we are allowed to control the specific form of this dependence. We take a Bayesian view on the estimation problem and assume that the prior is known and subsequently update our beliefs. The aim is to optimise an objective depending on both the control and the posterior distribution of the hidden signal. The typical interpretation is that we are searching for the hidden 'signal' and based on what we've learned so far we may choose different 'search actions' which might be more or less effective but at the same time more or less costly.

In mathematical terms, the observation process is given by

$$\mathrm{d}Y_t^u = h(u_t, X)\mathrm{d}t + \mathrm{d}W_t, \quad Y_0^u = 0,$$

where X is an unobservable random variable, u is the chosen control, and the function h specifies the spatial dependence on the signal corresponding to different actions. For any given control, the observation process generates a filtration, say $\mathcal{Y}^u$, which is explicitly control dependent. The posterior distribution, also referred to as the filter, is then given

Date: January 29, 2025.

Alexander M. G. Cox, Department of Mathematical Sciences, University of Bath, Bath, U. K..
e-mail: a.m.g.cox@bath.ac.uk .

Sigrid Källblad, Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden. e-mail: sigrid.kallblad@math.kth.se. The author gratefully acknowledges financial support from the Swedish Research Council (VR) under grant 2020-03449.

Chaorui Wang, Department of Mathematical Sciences, University of Bath, Bath, U. K..
e-mail: cw2581@bath.ac.uk. The author is supported by a scholarship from the EPSRC Centre for Doctoral Training in Statistical Applied Mathematics at Bath (SAMBa) under the project EP/S022945/1 and the Chinese Council Scholarship.

by

$$\pi_t^u = \mathrm{Law}(X|\mathcal{Y}_t^u).$$

Given a discount rate β and a running cost function, say k, depending on the control and the current state of the posterior distribution, the aim is to optimise an objective of the form

$$\mathbb{E}\int_0^\infty e^{-\beta t}k(\pi_t^u, u_t)\mathrm{d}t;$$

the optimisation is over a class of controls u which are adapted to their own generated filtration $\mathcal{Y}^u$, that is, controls that rely only on the information on X generated from observing Y^u.

We address this problem by use of dynamic programming arguments and stochastic control methods. To this end, we first express the dynamics of the posterior distribution in the observation filtration; the filter can then be viewed as a controlled measure-valued martingale whose dynamics is driven by a Brownian motion. We then introduce two thereto related formulations: A weak problem formulation and an approximate formulation where one restricts to piecewise constant controls. A priori, it is not clear that either of these formulations coincide. However, our main result establishes that the value of our original problem, the value of the weak formulation, and the limiting value of the approximate formulation as the time grid is refined, are equal in value. *A posteriori*, it is thus clear that the additional information available within the weak formulation does not alter the value of the problem. The proof bears resemblance to the approximation approach introduced in [4] as well as the stochastic Perron method developed in [5]; specifically, relying on the viscosity theory for MVMs developed in [17], it proceeds by establishing that the limiting and weak value functions are sub and super solutions, respectively, of an associated HJB equation to which appropriate comparison results applies. In particular, we obtain that the value of the problem can be characterised in terms of the unique continuous viscosity solution of this HJB equation. Finally, having constructed an optimiser for a piecewise constant problem with a fine enough time grid, returning to our original problem and relying on uniqueness in law for the controlled filtering equation, we construct a piecewise constant and arbitrarily-close-to-optimal strategy for our main problem of interest.

Our main problem of study is notably given in a strong formulation where the observation filtration has an explicit dependence on the control. At the same time, admissible controls are constrained to depend on the information provided by the observation process only. When aiming at picking a 'good control' there is thus a trade-off present since such controls must balance improving the estimation and minimizing the cost functional. The effect of considering strong formulations within stochastic control problems with unobservable components was recently reviewed in [15]. In particular, they illuminated the link between controlling the information flow and the notions of *dual effect* (referring to the interplay between the control's effect on the state and its influence on the estimation of unobservable components) and the much-studied trade-off between *exploration and exploitation* (referring to the knowledge of the unobservable component contra cost optimisation); we refer to that article for further details and references on the topic. Within the literature on optimal control under partial information, controls that are restricted to depend on the observation process only are often referred to as 'strict-sense' controls in contrast to so-called 'wide-sense' controls which are allowed to depend on a somewhat larger filtration. The latter were introduced in [23] and used in e.g. [21] and [7] where it was shown, among other things, that the value of their problem remained the same regardless of whether strict- or wide-sense controls were used; our main result is in line with this. In this literature, it is also common to consider a fixed filtered probability space on which admissible controls are defined and then let the controlled dynamics enter the problem via a change of measure. Crucially, the relevant filtration then becomes independent of the chosen control (see e.g. [8] for the case of strict-sense controls and [31] for the case of wide-sense controls). It is also common to optimise over tuples of this form. Our main result verifies that for the problem at hand, such variants of weak formulations would agree in value.

Our work is notably inspired by the contribution of [15]. Following their approach, we introduce our strong problem formulation by defining a class of pre-admissible controls and restricting in turn to the subset of controls which rely on the information provided by the observations only. Our usage of stochastic Perron and Barles-Souganides type arguments for establishing equivalence between our main problem of interest and the weak and approximate formulations thereof, is also closely inspired by their arguments. The key contribution of our work is that we allow for a more general dependence on the signal in the drift coefficient of the observation process; while we restrict to an action space enforcing a certain polynomial structure, we crucially allow for infinite polynomials. While this might seem an innocent extension at first, it has far-reaching consequences: In contrast to the problem studied in [15], our problem can *not* be reduced to a finite-dimensional one. This calls for a genuinely infinite-dimensional analysis working directly with the probability-measure–valued filters. We carry out such a study by relying on the viscosity theory developed for stochastic control problems featuring MVMs in [17]. In particular, we herein develop approximation arguments analogous to the results of Barles-Souganides [4] for this infinite-dimensional set-up; we find those developments to be of independent interest and hope that they will be useful also in other contexts. Apart from extending the existing theory on Bayesian optimal control problems, we thus also provide a concrete example of how the viscosity theory of [17] can be applied; in particular, for the first time, we establish continuity of the value function for this genuinely infinite-dimensional set-up. Our analysis also includes an original result on the stability of controlled MVMs in terms of their initial condition, a result which we also find to be of independent interest.

We now discuss the related literature. Within a more applied context, Bayesian search problems have been studied in e.g. [1, 18, 25, 42]. This literature aside, there are essentially three streams of research to which our problem relates:

First, our problem bears resemblance to the problem of stochastic control under partial information; the difference being that our signal is kept constant while we control the observation rather than the signal itself. For this class of problems, the existence of wide-sense optimal controls were studied in [22], [23] and [21]; see also [13] and [27]. The predominant approach in this literature has been to address the original problem via a reformulation of the objective enabling viewing it as a control problem where the conditional distribution of the hidden signal plays the role of the controlled state variable. The approach taken herein is of course closely related to this stream of research and we will review the relevant literature in more detail in Section 2.

Second, motivated by aims similar to ours, closely related problems have been considered within an optimal stopping framework. For example, in [19], the authors let an unobservable state influence the drift of the state process and their aim is to estimate this drift as accurately but also as quickly as possible. Considering a Bayesian framework they formulate it as an optimal stopping problem which they subsequently address. An extended version of such a stopping problem was also studied in [20]. We refer to those articles for further references.

Third, our formulation notably includes the problem of minimising, over admissible controls, an objective of the form[1]

$$\mathbb{E}\int_0^\infty e^{-\beta t}\tilde{k}(X, u_t)\mathrm{d}t, \tag{1}$$

where $\tilde{k}$ is some given cost function; this can be seen by defining $k(\mu, v) = \mu(\tilde{k}(\cdot, v))$, for any probability measure μ and action v. Closely related problems have been studied previously in the literature:

- (*i*) In [7], the authors consider a problem where the observation process follows a dynamics corresponding to letting $h(v, x) = vx$, for real-valued controls, and where the aim is to optimise an objective similar to (1) but with a running cost depending on the observation process itself. Specifically, they consider a cost of the form $\tilde{k}(Y_t^u)$ for $\tilde{k}(y) = y^2$, $y \in \mathbb{R}$. In [31], their results were extended to the case of $\tilde{k}$ being an arbitrary even, convex function of exponential growth. The corresponding finite-horizon problem was studied in [32]. Those articles obtained explicit results but under rather restrictive assumptions on the prior.
- (*ii*) More recently, the above-mentioned article [15], considered an observation process whose dynamics is again similar to ours but with $h(u_t, X)$ replaced by $b(t, Y_t^u, u_t)X$, for some function b, and with a similar controlled volatility coefficient appearing in front of the noise term; they also allowed for a multi-dimensional signal. They considered the problem of optimising a finite-horizon objective with a running cost of the form $\tilde{k}(t, Y_t^u, X, u_t)$ and a terminal cost of analogous form.

While those papers motivated our work, we note that our set-up is different: while we crucially allow for a *non-linear* dependence on the signal in the drift of the observation process as well as a *non-linear* dependence on the posterior distribution in the cost functional, neither the drift nor the cost has any explicit dependence on the observation process in our set-up. In order to render the presentation focused on the difficulties arising from the non-linear dependence on the signal, we choose to develop the theory for the case without such an explicit dependence on the observation process. We believe, however, that our results could be extended to cover also this situation but leave it for future research to thoroughly investigate this.

The remainder of the article is organized as follows: In Section 2 we introduce our main problem of interest, express the dynamics of the involved processes in the observation filtration, and discuss the related literature; in Section 3 we introduce our weak problem formulation as well as the piecewise constant approximate formulation and provide our main result; Section 4 presents our stability results and Section 5 is devoted to the proof of our main result.

Notation. Throughout, μ_0 will be a given fixed probability measure on $\mathbb{R}$. We denote by $\mathcal{P}$ the space of probability measures on $\mathbb{R}$ whose support is contained in the support of μ_0; we equip this space with the topology of weak convergence rendering it a Polish space. Since we will assume that μ_0 has compact support, we can work with e.g. the Wasserstein-1 metric, which we denote by $\mathcal{W}$. That μ_0 has compact support also implies

[1] Notably there is still a trade-off present when choosing the control since it affects both the cost and the estimation of the unknown signal; as before, this is a consequence of the fact that we consider a strong formulation where the choice of the control influences the available information through the filtration.

that the space $\mathcal{P}$ is compact. We denote by $\mathcal{P}^s$ the (closed) subset of probability measures supported in one single point. We write C_b for the bounded continuous functions on $\mathbb{R}$.

2. Problem formulation

2.1. **Controls and objective.** Consider a given filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ satisfying the usual conditions and supporting a Brownian motion W and an independent $\mathcal{F}_0$-measurable random variable X with law μ_0; the latter will play the role of our unknown signal. We suppose throughout that μ_0 has compact support; that is, there exists some $R > 0$ such that

$$\operatorname{supp}(\mu_0) \subset (-R, R).$$

In order to define our *set of actions*, we let $K > 0$ be arbitrary but fixed and introduce a closed set $\mathcal{U}$ such that

$$\mathcal{U} \subseteq \left\{ v = (v_1, v_2, \dots) : v_i \in \mathbb{R},\, i \in \mathbb{N}, \text{ and } \sum_{i=1}^{\infty} (R^i v_i)^2 \leq K \right\};$$

we here consider the subspace topology inherited from the product topology and note that $\mathcal{U}$ thus defined is a Polish space. We also define

$$h(v, x) = \sum_{i=1}^{\infty} v_i x^i, \quad v \in \mathcal{U},\ x \in (-R, R); \tag{2}$$

we note that this function is well-defined and jointly measurable, that the definition fixes $h(v, 0) \equiv 0$, and that $x \mapsto h(v, x)$ is continuous uniformly in v on $\operatorname{supp}(\mu_0)$ (c.f. Lemma 4.1 below).

Our controls will be progressively measurable processes taking their values in this action space. More pertinently, we first define a set of pre-admissible controls by

$$\mathcal{A}^{pre} := \{u : \Omega \times [0, \infty) \to \mathcal{U} \colon u \text{ is } \mathbb{F}\text{-progressively measurable}\}.$$

For $u \in \mathcal{A}^{pre}$, the observation process $Y^u = (Y_t^u)_{t \geq 0}$ is then given by

$$\mathrm{d}Y_t^u = h(u_t, X)\mathrm{d}t + \mathrm{d}W_t, \quad Y_0^u = 0. \tag{3}$$

For $u \in \mathcal{A}^{pre}$, we let $\mathcal{Y}^u = (\mathcal{Y}_t^u)_{t \geq 0}$ be the (control-dependent) observation filtration; that is, the filtration generated by the process Y^u and completed by the $\mathbb{P}$-null sets. The set of *admissible controls* is finally defined as

$$\mathcal{A} := \{u \in \mathcal{A}^{pre} : u \text{ is } \mathcal{Y}^u\text{-progressively measurable}\};$$

for further discussion on how this type of closed-loop controls relate to alternative definitions, we refer to [15, Remark 2.4].

We consider a discount rate $\beta > 0$ and a measurable *cost function* $k : \mathcal{P} \times \mathcal{U} \to \mathbb{R}$ to be given; throughout, we impose the following *standing assumption*:

Assumption 2.1. *The cost function k is bounded and $k(\cdot, v)$ is continuous on $\mathcal{P}$ uniformly in $v \in \mathcal{U}$.*

For $u \in \mathcal{A}$, we define the posterior distribution $\pi^u = (\pi_t^u)_{t \geq 0}$ by

$$\pi_t^u = \operatorname{Law}(X | \mathcal{Y}_t^u); \tag{4}$$

note that π^u takes values in $\mathcal{P}$ and that $\pi_0^u = \mu_0$. With each $u \in \mathcal{A}$, we then associate the following *cost functional*:

$$J(u) = \mathbb{E} \int_0^{\infty} e^{-\beta t} k(\pi_t^u, u_t)\mathrm{d}t. \tag{5}$$

Our main problem of interest is the one of minimising $J(u)$ over all $u \in \mathcal{A}$.

2.2. **MVMs and SDEs.** For $u \in \mathcal{A}$, π^u is a measure-valued martingale (MVM) in the sense of [16, Definition 2.1]. More pertinently, taking the signal to be constant in the classical (controlled) Kushner-Stratonovich equation (see e.g. [41, Proposition 7.2.8]), we have that π^u is the solution to [2]

$$\mathrm{d}\pi^u_t(\varphi) = \big(\pi^u_t\big(\varphi(\cdot)h(u_t,\cdot)\big) - \pi^u_t\big(\varphi\big)\pi^u_t\big(h(u_t,\cdot)\big)\big)\big(\mathrm{d}Y^u_t - \pi^u_t(h(u_t,\cdot))\mathrm{d}t\big), \quad \varphi \in C_b, \tag{6}$$

equipped with the initial condition $\pi^u_0 = \mu_0$. Following the filtering literature, for $u \in \mathcal{A}$, we define the innovations process by

$$\mathrm{d}I^u_t = (h(u_t, X) - \pi^u_t(h(u_t,\cdot)))\,\mathrm{d}t + \mathrm{d}W_t, \quad I^u_0 = 0;$$

it is a $(\mathcal{Y}^u, \mathbb{P})$-Brownian motion. Note that (3) and (6) can then be re-written as

$$\mathrm{d}Y^u_t = \pi^u_t(h(u_t,\cdot))\mathrm{d}t + \mathrm{d}I^u_t, \quad Y^u_0 = 0. \tag{7}$$

and

$$\mathrm{d}\pi^u_t(\varphi) = \big(\pi^u_t\big(\varphi(\cdot)h(u_t,\cdot)\big) - \pi^u_t\big(\varphi\big)\pi^u_t\big(h(u_t,\cdot)\big)\big)\mathrm{d}I^u_t, \ \varphi \in C_b, \ \ \pi^u_0 = \mu_0; \tag{8}$$

the process π^u can thus be viewed as a controlled MVM of the specific form considered in [17] (c.f. equation (2.3) therein).

The links to MVMs and the above equations will become crucial for our upcoming analysis. Indeed, the benefit of linking our problem to equation (8) is that it enables relating the problem to a more standard (albeit still infinite-dimensional) stochastic control problem featuring a controlled state-process being driven by a Brownian motion; the idea being that if one manages to construct a solution to that auxiliary problem, one should be able to construct a solution to the original problem too. To see this, suppose for illustrative purposes that we've been given some probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ supporting a Brownian motion B, and have found a control in feedback form provided by a function $\hat{u} : \mathcal{P} \to \mathcal{U}$ for which the SDE

$$\mathrm{d}\tilde{\pi}_t(\varphi) = \left(\tilde{\pi}_t\left(\varphi(\cdot)h\left(\hat{u}(\tilde{\pi}_t),\cdot\right)\right) - \tilde{\pi}_t\left(\varphi\right)\tilde{\pi}_t\left(h\left(\hat{u}(\tilde{\pi}_t),\cdot\right)\right)\right)\mathrm{d}B_t, \quad \tilde{\pi}_0 = \mu_0, \tag{9}$$

admits a pathwise unique solution $\tilde{\pi}$. Then one would hope to be able to construct[3] a control $u^* \in \mathcal{A}$ for which the objective attains the same value as the objective evaluated at the tuple $(\tilde{\pi}, \hat{u}(\tilde{\pi}_\cdot))$. Provided (i) that $\hat{u}$ is optimal for the problem of optimising the objective over solutions to (9) for some suitable class of controls, and (ii) that the value of this problem coincides with the value of our original problem of interest, it would follow that u^* were optimal for that original problem. In the upcoming analysis, the role of $\hat{u}$ will be played by piecewise constant controls in feedback form on a finite time grid, and we will establish that the value of the respective problem formulations

[2]A-priori, setting $\mathcal{L} \equiv 0$ in [41, Proposition 7.2.8], we obtain (8) for $\varphi \in C^2_b$ with all of its derivatives bounded. However, for any $\varphi \in C_b$, there exists a sequence $\{\varphi_n\}$ belonging to this class and converging pointwise to φ. Since, by Lemma 4.1 below, $\{h(v,x) : v \in \mathcal{U}, x \in \mathrm{supp}(\mu_0)\}$ is a bounded subset of $\mathbb{R}$,

$$\mathbb{E}\int_0^t \left(\int_{\mathbb{R}} (\varphi - \varphi_n)(x)\left(h(u_s,x) - \pi^u_s\left(h(u_s,\cdot)\right)\right)\pi^u_s(\mathrm{d}x)\right)^2 \mathrm{d}s \xrightarrow[n\to\infty]{} 0,$$

which implies that (8) holds for each $\varphi \in C_b$.

[3]Indeed, returning to equations (3) and (6) with u_t replaced by $\hat{u}(\pi^u_t)$ in all instances, given that $\hat{u}$ is nice enough, these SDEs admit a pathwise unique solution, say (Y, π), which is adapted to the filtration generated by Y. Defining $u^* \in \mathcal{A}^{pre}$ by $u^*_t = \hat{u}(\pi_t)$, (3) gives that $Y = Y^{u^*}$ and it follows that the pair (Y, π) is adapted to $\mathcal{Y}^{u^*}$; hence $u^* \in \mathcal{A}$. Since the solution to (6) is pathwise unique, we also have that $\pi = \pi^{u^*}$ and therefore $u^*_t = \hat{u}(\pi^{u^*}_t)$. In consequence, (Y^{u^*}, π^{u^*}) must solve the system of SDEs (7) and (8) with u_t replaced by $\hat{u}(\pi^{u^*}_t)$ and I^u by I^{u^*} which is a $(\mathcal{Y}^{u^*}, \mathbb{P})$-BM. Since pathwise uniqueness implies uniqueness in law for the SDE (9), or equivalently for (8), evaluating the objective at (π^{u^*}, u^*) and $(\tilde{\pi}, \hat{u}(\tilde{\pi}_\cdot))$ must yield the same value, which verifies the claim.

coincide by linking them to the unique continuous viscosity solution of an associated HJB equation.

2.3. **Related equations in the literature.** Within the literature on *stochastic control under partial information* the above approach of linking the problem of original interest to a more standard control problem (involving measure-valued processes) is often referred to as the *separation principle* (although there is some ambiguity in how this terminology is used). We here recall the various related equations which have emerged within this context and review the related literature.

To this end, recall that one is here typically interested in a problem which is similar to ours but where the hidden signal follows itself an Itô dynamics; specifically, one is allowed to control that signal process, say X_t, $t \geq 0$, while the observation is given by (3) for some fixed (not controlled[4]) function $h(x)$, $x \in \mathbb{R}$, and X is replaced by X_t. The aim is to optimise an objective of the form (1) — again with X replaced by the controlled process X_t. The nomenclature 'partially observable' refers to the fact that the controls must be adapted to the filtration generated by the observation process; controls given in feedback form in terms of the signal are for example not admissible. As discussed in the introduction, one often considers wide-sense controls that are adapted to a somewhat larger filtration; it is also common to consider various weak formulations where the explicit control-dependence of the filtration is circumvented.

Kushner-Stratonovich equations. In its original form, the separation principle amounts to rewrite the objective onto the form (5), where now

$$\pi_t^u = \text{Law}(X_t | \mathcal{Y}_t^u);$$

this filter then solves the so-called *Kushner-Stratonovich equation* which is similar to (8) but with some additional terms appearing due to the dynamics of X itself — the control then appears through those additional terms while typically $h(v,x) \equiv h(x)$. At the cost of the involved processes taking values in the space of probability measures, the resulting problem of optimising (5) over solutions to the KS equation is conceptually simpler since it is a fully observable problem; for example, controls given in feedback form in terms of the filter are admissible. Provided an optimiser can be found, one would aim at closing the loop by constructing a solution to the original problem along the lines outlined above.

While offering a conceptually simple and elegant formulation, this problem turns out to be notoriously difficult to solve in practice, and explicit results are few to date. Particular attention has however been paid to the case when X_0 is Gaussian and the various coefficients appearing in the dynamics of the signal and the observation process are linear. The resulting filter, referred to as the *Kalman-Bucy filter*, is then Gaussian too with its variance process solving a deterministic Ricatti equation and its mean process an SDE. The separated problem thus becomes a finite-dimensional problem which can be addressed using standard methods. For various results in this direction we refer to [24], [8, Chapter 7] and [41, Section 7.5]; see in particular [41, Proposition 7.3.9] for a result analogous to the above description of how to construct an optimiser to the original problem based on an optimiser to the separated one.

Zakai equations. A more common approach has been to employ the separation principle in conjunction with the so-called *Zakai equation*. The approach is based on introducing a measure $\mathbb{Q}_t^u \sim \mathbb{P}$ under which Y^u is a Brownian motion on $[0,t]$ (independent of X_0

[4]Notably, [41] and [2] also allow for controlled observations.

and the Brownian motion driving the signal). By Girsanov, this is achieved by defining $\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}_t^u = \Lambda_t^u$ with $\mathrm{d}\Lambda_t^u = h(X_t, u_t)\Lambda_t^u \mathrm{d}Y_t^u$. Defining the measure-valued process

$$\sigma_t^u(\varphi) := \mathbb{E}^{\mathbb{Q}_t^u}[\varphi(X_t)\Lambda_t^u|\mathcal{Y}_t^u], \quad \varphi \in C_b, \tag{10}$$

making a change of measure, and, in turn, applying the tower property, one may then rewrite the objective as follows

$$\mathbb{E}^{\mathbb{Q}_T^u}\left[\int_0^\infty e^{-\beta t}\sigma_t^u\left(\tilde{k}(\cdot, u_t)\right)\mathrm{d}t\right]; \tag{11}$$

meanwhile, making use of the properties of $\mathbb{Q}_t^u$, one obtains that the process σ^u solves the so-called Zakai equation. For our set-up with a constant signal and controlled observations, this equation would take the form

$$\mathrm{d}\sigma_t^u(\varphi) = \sigma_t^u\left(h(u_t, \cdot)\varphi\right)\mathrm{d}Y_t^u. \tag{12}$$

For the case with a signal following a controlled Itô dynamics there are however additional terms present; the control notably enters through those terms and typically $h(v, x) \equiv h(x)$. The process σ^u is referred to as the unnormalised process: by Bayes' rule, $\pi_t^u(\varphi) = \sigma_t^u(\varphi)/\sigma_t^u(1)$, $\varphi \in C_b$.

The main point is that one may again view the problem of optimising (11) over controlled solutions to the Zakai equation (c.f. (12)) as a fully observed stochastic control problem (with a measure-valued controlled process). To study this separated problem has been the predominant approach in the literature. Most frequently, it is assumed that the unnormalised conditional distribution admits a density: the Zakai equation then reduces to the so-called *Duncan-Mortensen-Zakai equation* for the density process itself. Within such a set-up, there are numerous contributions to date: In [8, Chapter 8] the problem is dealt with by use of the stochastic maximum principle as well as via the DPP combined with semi group arguments. In [30] (c.f. [28, 29]) the associated HJB-equation is studied and it is proven that the value function is indeed a viscosity solution (albeit in a rather weak sense which does not enable getting comparison). In [21], the question of existence of optimal controls is discussed (within the class of randomised controls). In [36], it is assumed that the density process belongs to L^2; it is then established that the value function is indeed the unique viscosity solution to the corresponding HJB equation (see also [35, 37]). In [26], the weaker requirement that the density process belongs to some weighted L^2-space is imposed (still a Hilbert space); also here, it is then proven that the value function is the unique viscosity solution to the HJB equation. Finally, motivated by similar questions, [38] established an Itô formula for solutions to the KS and Zakai equations without imposing any assumptions on the existence of densities; in [39], viscosity solutions were discussed for some related equations.

Yet another version of the separation principle. The articles [2] and [3] considered yet a different implementation of the separation principle. By first making a change of measure and using again the tower property, they noticed that the objective can also be rewritten as

$$\mathbb{E}^{\mathbb{Q}_T^u}\left[\int_0^\infty \int e^{-\beta t}\tilde{k}(x, u_t)z\ \mathrm{d}\pi_t^u(x, z)\mathrm{d}t\right], \tag{13}$$

where π_t^u is now defined as the joint distribution of (X_t, Λ_t^u) under $\mathbb{Q}_t^u$ given $\mathcal{Y}_t^u$. Recalling the dynamics for Λ_t and rewriting the dynamics for X in terms of the observation process, one might then write down the equation for π^u.

Indeed, effectively one is then back in a set-up similar to the one considered when deriving the KS equation with the difference that the only thing that is being observed is some correlated noise (note that in this case the unnormalised and normalised conditional probabilities coincide); in effect, one has thus obtained a simpler dynamics at the cost of introducing an additional factor. For our case with a constant signal and controlled observations, the equation would take the form

$$\mathrm{d}\pi_t^u(\varphi) = \frac{1}{2}\int (h(u_t,x)z)^2 \frac{\partial^2\varphi}{\partial z^2}(x,z)\mathrm{d}\pi_t^u(x,z)\mathrm{d}t + \int h(u_t,x)z\frac{\partial\varphi}{\partial z}(x,z)\mathrm{d}\pi_t^u(x,z)\mathrm{d}Y_t^u;$$

for the general case there are however additional terms present. The main point is that one may now consider the separated problem where the objective is given by (13) and where the controlled probability-measure valued processes π^u follows a certain dynamics.

In [3], this separated problem is studied in a Markovian set-up; the article notably imposes no assumptions on the existence of a density. They consider the corresponding problem using randomised controls and first establish that the randomised problem and the original problem coincide. For the randomised problem they then consider the HJB equation and prove that the value function is the unique viscosity solution. To do so, they make use of the Lions derivative; that is, they work in a lifted space and rely on the Itô formula in terms of the Lions derivative as formulated in e.g. [12]. In [2], the authors consider the analogous (separated) problem with non-Markovian dynamics and address the problem using BSDE methods.

Notably, [2] also allows for controlled observations. We leave for future research to investigate whether their approach may provide further insights compared to the results obtained herein.

A finite-dimensional reduction. In [15], the authors consider a set-up more closely related to ours where the signal is constant and the effect of the control enters through the observation. Specifically, their observation process follows a dynamics similar to (3) but with $h(u_t, X)$ replaced by $b(t, Y_t^u, u_t)X$ for some function b; they also let the signal be multi-dimensional and allow for a controlled volatility coefficient but we omit those features here. For this set-up, σ^u is given by (10) with $X_t \equiv X$ and $\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}_t^u = \Lambda_t^u = \mathcal{E}(X\Upsilon^u)_t$, where

$$\mathrm{d}\Upsilon_t^u = b(t, Y_t^u, u_t)\mathrm{d}Y_t^u, \quad \Upsilon_0^u = 0. \tag{14}$$

Hence, defining, $F[\varphi](v,\gamma) = \int \varphi(x)e^{xv-\frac{1}{2}x^2\gamma}\mu_0(\mathrm{d}x)$, it follows that

$$\sigma_t^u(\varphi) = \mathbb{E}^{\mathbb{Q}_t^u}\left[\varphi(X)e^{X\Upsilon_t^u-\frac{1}{2}X^2\langle\Upsilon^u\rangle_t}|\mathcal{Y}_t^u\right] = F[\varphi]\left(\Upsilon_t^u, \langle\Upsilon^u\rangle_t\right).$$

In particular, applying Itô's formula and making use of (14) and the fact that $\frac{1}{2}F[\varphi]_{vv} + F[\varphi]_\gamma = 0$ and $F[\varphi]_v = F[\varphi\cdot\mathrm{id}]$, it is straightforward to verify that σ^u thus defined satisfies

$$\mathrm{d}\sigma_t^u(\varphi) = F[\varphi]_v\left(\Upsilon_t^u, \langle\Upsilon^u\rangle_t\right)\mathrm{d}\Upsilon_t^u = \sigma_t^u(\varphi\cdot\mathrm{id})b(t, Y_t^u, u_t)\mathrm{d}Y_t^u, \tag{15}$$

which is what the Zakai equation reduces to in this set-up (c.f. (12)).

The filter is thus fully characterised by the current state of the processes Υ^u and $\langle\Upsilon^u\rangle$; in [15], this is exploited in order to address the problem of optimising, over $u \in \mathcal{A}$, the finite-horizon analogue of an objective of the form (1) — with the running cost $\tilde{k}(X, u_t)$ replaced by $\tilde{k}(t, Y_t^u, X, u_t)$ — as a finite-dimensional problem. More pertinently, note that for $u \in \mathcal{A}$, the innovations process, which is a $(\mathcal{Y}^u, \mathbb{P})$-BM, is given by

$$\mathrm{d}I_t^u = (X - \pi_t^u(\mathrm{id}))\, b(t, Y_t^u, u_t)\mathrm{d}t + \mathrm{d}W_t, \quad I_0^u = 0;$$

in consequence,

$$\mathrm{d}Y_t^u = \frac{F[\mathrm{id}]}{F[1]}\left(\Upsilon_t^u, \langle \Upsilon^u\rangle_t\right) b(t, Y_t^u, u_t)\mathrm{d}t + \mathrm{d}I_t^u, \quad Y_0^u = 0. \tag{16}$$

Combining (14) with (16), they observe that the dynamics of $(Y^u, \Upsilon^u, \langle\Upsilon^u\rangle)$ can be described by a system of SDEs driven by a $\mathbb{P}$-BM. In parallel, they express the objective as an expected value under $\mathbb{P}$ of a function depending on precisely $(Y^u, \Upsilon^u, \langle\Upsilon^u\rangle)$; this can be achieved by rewriting the objective as an expected value of a functional depending on the observation and the filter (in analogy with the KS approach) and subsequently express the filter in terms of $(\Upsilon^u, \langle\Upsilon^u\rangle)$. This leaves them with a separated problem in a more standard form, the study of which forms the basis for their approach.

As mentioned in [15, Remark 2.7], a similar approach may be employed when restricting to controls which are polynomials of some fixed degree; we crucially go beyond such an assumption in our work. It is worth noticing that when turning an optimiser of their separated problem into an optimiser for the original problem, they focus on piecewise constant controls; we will here follow a similar approach. In order to further relate their set-up to ours, making use of (15), (16) and Itô's formula, we obtain that

$$\mathrm{d}\pi_t^u(\varphi) = \left(\pi_t^u(\varphi\,\mathrm{id}) - \pi_t^u(\varphi)\pi_t^u(\mathrm{id})\right) b(t, Y_t^u, u_t)\mathrm{d}I_t^u.$$

For $b(t, Y_t^u, u_t) = b(u_t)$, this is precisely what (8) reduces to when $h(u_t, X) = b(u_t)X$.

3. Weak and approximate formulation and main result

In this section we provide alternative problem formulations and give our main result, which on the one hand states that all those formulations coincide, and on the other hand characterises an arbitrarily-close-to-optimal strategy for our main problem of interest.

3.1. Weak problem formulation.

We first introduce our weak formulation where there notably are no restrictions on the information used for the strategies:

Definition 3.1 (Weak controls)**.** A tuple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W, \xi, u)$ is a *weak admissible control* if it satisfies:

(*i*) $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space satisfying the usual conditions;
(*ii*) W is an $(\mathbb{F}, \mathbb{P})$-Brownian motion;
(*iii*) u is an $\mathcal{U}$-valued $\mathbb{F}$-progressively measurable process;
(*iv*) ξ is a continuous $\mathcal{P}$-valued MVM such that

$$\mathrm{d}\xi_s(\varphi) = \Big(\xi_s\big(\varphi(\cdot)h(u_s,\cdot)\big) - \xi_s(\varphi)\xi_s\big(h(u_s,\cdot)\big)\Big)\mathrm{d}W_s, \quad \varphi \in C_b. \tag{17}$$

To avoid overly cumbersome notation, we often call (ξ, u) an admissible control without explicitly mentioning the other objects of the tuple. For $\mu \in \mathcal{P}$, we denote by $\mathcal{A}^{weak}(\mu)$ the set of all weak admissible controls (ξ, u) which satisfy $\xi_0 = \mu$; we note that this set is non-empty. In turn, we define

$$V^{weak}(\mu) = \inf_{\mathcal{A}^{weak}(\mu)} \mathbb{E}\int_0^\infty e^{-\beta t} k(\xi_t, u_t)\mathrm{d}t. \tag{18}$$

Our main result will establish that the value of this weak formulation agrees with the value of our original problem of interest.

3.2. **Approximate problem formulation.** We here introduce an approximation of the original problem of interest by restricting to piecewise constant controls.

To this end, here, let $\Omega = C([0,\infty),\mathbb{R})$, denote by W the canonical process, and let $\mathbb{P}$ be the Wiener measure under which W is a Brownian motion. Let $\mathbb{F}^W$ denote the raw filtration generated by the canonical process and augmented by the null sets (but not completed). For $n \in \mathbb{N}$, define the dyadic step size $\delta_n = 2^{-n}$ and consider the associated discrete time grid $\mathbb{T}^n = \{k\delta_n : k \in \mathbb{N}\}$. For $n \in \mathbb{N}$, let

$$\mathcal{A}^n = \Big\{u : [0,\infty)\times\Omega \to \mathcal{U} : u_t = \sum_{r\in\mathbb{T}^n} u_r \mathbf{1}_{[r,r+\delta_n)}(t),\ \ t\in[0,\infty),\ u_r \text{ is } \mathcal{F}^W_r\text{-meas.}\Big\}.$$

For $\mu \in \mathcal{P}$ and $u \in \mathcal{A}^n$, $n \in \mathbb{N}$, let $\xi^{u;\mu}$ be the strong solution of

$$\mathrm{d}\xi_t(\varphi) = \left(\xi_t\left(\varphi h(u_t,\cdot)\right) - \xi_t\left(\varphi\right)\xi_t\left(h(u_t,\cdot)\right)\right)\mathrm{d}W_t,\ \varphi\in C_b,\ \ \xi_0 = \mu; \tag{19}$$

since μ has compact support and $h(v,\cdot)$ is continuous for $v \in \mathcal{U}$ (see Lemma 4.1 below), existence and pathwise uniqueness of solutions to (19) is guaranteed by [34, Theorem 2.1]. For $\mu \in \mathcal{P}$ and $u \in \mathcal{A}^n$, $n \in \mathbb{N}$, we define

$$\hat{J}(u;\mu) := \mathbb{E}\int_0^\infty e^{-\beta t} k\left(\xi^{u;\mu}_t, u_t\right)\mathrm{d}t = \mathbb{E}\sum_{r\in\mathbb{T}^n}\int_r^{r+\delta_n} e^{-\beta t} k\left(\xi^{u;\mu}_t, u_r\right)\mathrm{d}t.$$

For $n \in \mathbb{N}$ and $\mu \in \mathcal{P}$, the approximate problem formulation is then given by

$$V^n(\mu) = \inf_{u\in\mathcal{A}^n} \hat{J}(u;\mu).$$

Our main result will establish that these approximate problems converge, as the partition grows finer, to the value of our original problem of study. In order to formalise this, we define the limiting function, for $\mu \in \mathcal{P}$, by

$$V^+(\mu) = \lim_{n\to\infty} V^n(\mu);$$

it is well defined since V^n is bounded and monotone in n.

3.3. **Associated HJB equation.** In order to formulate the HJB equation, we define, for $\mu \in \mathcal{P}$, $r \in \mathbb{R}$ and $\varphi \in C(\mathbb{R}^2)$,

$$H(\mu,r,\varphi) = \beta r + \sup_{v\in\mathcal{U}}\left\{-k(\mu,v) - \frac{1}{2}\int_{\mathbb{R}\times\mathbb{R}}\varphi(x,z)\sigma(v,\mu;\mathrm{d}x)\sigma(v,\mu;\mathrm{d}z)\right\},$$

where, for $\mu \in \mathcal{P}$ and $v \in \mathcal{U}$,

$$\sigma(v,\mu;\mathrm{d}x) = (h(v,x) - \mu(h(v,\cdot)))\mu(\mathrm{d}x).$$

The HJB-equation associated with our problem is then given by (c.f. the dynamics in (17) and objective in (18)):

$$H\left(\mu, u(\mu), \frac{\partial^2 u}{\partial\mu^2}(\cdot,\cdot;\mu)\right) = 0, \quad \mu\in\mathcal{P}; \tag{20}$$

for functions $u : \mathcal{P} \to \mathbb{R}$ we here make use of the notion of derivative as it was defined in [17, Section 4], which is essentially the same notion as the one referred to as the *linear functional derivative* in [11, Section 5.4].

Remark 3.2. When $\mu = \delta_x \in \mathcal{P}^s$, (20) simplifies to

$$u(\delta_x) = \frac{1}{\beta}\inf_{v\in\mathcal{U}} k(\delta_x, v).$$

This can be interpreted as a kind of boundary condition: Since an MVM starting at a Dirac measure δ_x stays there for all times, by definition,

$$V^{weak}(\delta_x) = V^+(\delta_x) = \frac{1}{\beta} \inf_{v \in \mathcal{U}} k(\delta_x, v).$$

3.4. **Our main result.** We are now ready to provide our main result; its proof is reported in Section 5.3. When talking about solutions of equation (20) we make use of the notion of viscosity solution as it was introduced in [17]; for the reader's convenience, this definition is recalled in Definition 5.1 below. We also recall the standing Assumption 2.1.

Theorem 3.3. *It holds that*

$$V^{weak}(\mu_0) = \inf_{u \in \mathcal{A}} J(u) = V^+(\mu_0).$$

Moreover, on $\mathcal{P}$,

$$V^{weak} = V^+,$$

and this function is the unique continuous viscosity solution of (20).

Finally, for any $\varepsilon > 0$, we can find $n \in \mathbb{N}$ and a function $\hat{u}^n : \mathcal{P} \to \mathcal{U}$, such that $u^* \in \mathcal{A}$ given in feedback form by

$$u^*_t = \sum_{r \in \mathbb{T}^n} \hat{u}^n\big(\pi^{u^*}_r\big) \mathbf{1}_{[r, r+\delta_n)}(t), \quad t \in [0, \infty), \tag{21}$$

satisfies

$$J(u^*) \leq \inf_{u \in \mathcal{A}} J(u) + \varepsilon.$$

Remark 3.4. The functions V^+ and V^{weak} effectively correspond to a piecewise constant approximation and a weak formulation of the *separated problem* associated with our main problem of study. Thanks to the above result, it is however possible to draw conclusions also about alternative formulations of the original problem itself: (i) Restricting in our main problem of study to controls which are piecewise constant (over arbitrary partitions) would not alter the value of the problem. (ii) Since the specifics of the underlying probability space do not have an impact on the value of the problem, considering a weak formulation where one also optimises over the underlying probability tuples, would still yield the same value. (iii) In the literature on stochastic control under partial information, it is common to consider so-called wide-sense controls which are adapted to a filtration which is larger than the observation filtration but still small enough for the filter to satisfy the associated filtering equation; allowing for controls of this type (adopting the definitions to the present context) would also not alter the value of the problem.

4. Stability properties of controlled MVMs

In this section we establish stability properties of our controlled MVMs with respect to the initial condition; these properties will become crucial in the upcoming analysis. We note that the proofs of these results do not make any use of the upcoming viscosity theory. Specifically, the proof of the first result relies on stability properties for classical SDEs, and the proof of the second result makes use of ideas similar to those used in the proof of [34, Theorem 2.1]; c.f. also the proof of [22, Theorem 3].

Before proceeding to the stability results, we establish a basic property of our action space:

Lemma 4.1. *Let $\kappa < R$. Let $h(v,x)$ be given by (2) for $v \in \mathcal{U}$ and $x \in [-\kappa,\kappa]$. Then $h(v,\cdot)$ is continuous on $[-\kappa,\kappa]$ uniformly in $v \in \mathcal{U}$. In particular, $\{h(v,x) : v \in \mathcal{U}, x \in [-\kappa,\kappa]\}$ is a bounded subset of $\mathbb{R}$.*

Proof. For all $v \in \mathcal{U}$, $x \in [-\kappa,\kappa]$ and $n,m \in \mathbb{N}$,

$$\sum_{i=n}^{m} |v_i x^i| \le \sqrt{K}\sqrt{\sum_{i=n}^{m}\left(\frac{\kappa}{R}\right)^{2i}}. \tag{22}$$

Since $\sum_{i=1}^{\infty}(\kappa/R)^{2i}$ converges, we obtain that $\sum_{i=1}^{\infty} v_i x^i$ is absolutely convergent, for all $x \in [-\kappa,\kappa]$ and $v \in \mathcal{U}$. Next, relying once again on (22), we have for every $\varepsilon > 0$, that there exists some $n \in \mathbb{N}$, such that for all $x,y \in [-\kappa,\kappa]$ and $v \in \mathcal{U}$,

$$|h(v,y) - h(v,x)| \le \sum_{i=1}^{n} |v_i||x^i - y^i| + \varepsilon \le \sqrt{K}\sqrt{\sum_{i=1}^{n}\left(\frac{x^i - y^i}{R^i}\right)^2} + \varepsilon.$$

For each $i = 1,\dots,n$, there exists some $\delta_i > 0$, such that if $|x-y| \le \delta_i$, then $\frac{|x^i - y^i|}{R^i} \le \frac{\varepsilon}{\sqrt{Kn}}$. Provided that $|x-y| \le \min_{i=1,\dots,n}\delta_i$, it then holds that $|h(v,y) - h(v,x)| \le 2\varepsilon$, for all $v \in \mathcal{U}$, which completes the proof. □

Proposition 4.2. *Let $N \in \mathbb{N}$ and consider some fixed N points $\{x_1,\dots,x_N\} \subset \operatorname{supp}(\mu_0)$; let $\mathcal{P}_N = \{\mu \in \mathcal{P} : \operatorname{supp}(\mu) \subseteq \{x_1,\dots,x_N\}\}$.*

(i) On any filtered probability space supporting a Brownian motion W and a progressively measurable $\mathcal{U}$-valued process u, equation (17) equipped with an initial condition $\mu \in \mathcal{P}_N$ admits a pathwise unique solution; it is $\mathcal{P}_N$-valued a.s.

(ii) For any $t > 0$, there exists a constant c depending on N and t, such that for any filtered probability space supporting a Brownian motion W and a progressively measurable $\mathcal{U}$-valued processes u, and for any initial conditions $\mu,\nu \in \mathcal{P}_N$,

$$\frac{\mathbb{E}\|((\xi_t^{u;\mu} - \xi_t^{u;\nu})(\{x_1\}),\dots,(\xi_t^{u;\mu} - \xi_t^{u;\nu})(\{x_N\}))\|}{\|((\mu-\nu)(\{x_1\}),\dots,(\mu-\nu)(\{x_N\}))\|} \le c,$$

where $\xi^{u;\mu}$ and $\xi^{u;\nu}$ denote the respective solutions to equation (17) and $\|\cdot\|$ is the Euclidean norm.

Proof. (i). Suppose $\xi^{u;\mu}$ is a solution to (17) with respect to an initial condition $\mu \in \mathcal{P}_N$. Then $\xi_t^{u;\mu}$ remains supported on $\{x_1,\dots,x_N\}$ a.s. for all $t \ge 0$, and $\xi^{u;\mu}(\{x_i\})$, $i = 1,\dots,N$, are non-negative martingales which satisfy

$$\mathrm{d}\xi_t^{u;\mu}(\{x_i\}) = \xi_t^{u;\mu}(\{x_i\})\left(h(u_t,x_i) - \sum_{n=1}^{N}\xi_t^{u;\mu}(\{x_n\})h(u_t,x_n)\right)\mathrm{d}W_t.$$

Let Δ^{N-1} be the N-dimensional standard simplex and define the function $\tilde{\sigma} : \mathcal{U} \times \Delta^{N-1} \to \mathbb{R}^N$ by

$$\tilde{\sigma}(v,\theta) = \left(\theta_1\left(h(v,x_1) - \sum_{n=1}^{N}\theta_n h(v,x_n)\right),\dots,\theta_N\left(h(v,x_N) - \sum_{n=1}^{N}\theta_n h(v,x_n)\right)\right),$$

If we denote $\tilde{\xi}_t^{u;\mu} = (\xi_t^{u;\mu}(\{x_i\}),\dots,\xi_t^{u;\mu}(\{x_N\}))$, then $\tilde{\xi}^{u;\mu}$ is a Δ^{N-1}-valued process which satisfies the N-dimensional SDE

$$\mathrm{d}\tilde{\xi}_t^{u;\mu} = \tilde{\sigma}(u_t,\tilde{\xi}_t^{u;\mu})\mathrm{d}W_t. \tag{23}$$

Conversely, suppose $\tilde{\xi}^{u;\mu}$ is a solution to (23) with initial condition $\tilde{\xi}_0^{u;\mu} \in \Delta^{N-1}$. We define the measure-valued process $\hat{\xi}^{u;\mu}$ by $\hat{\xi}_t^{u;\mu}(f) := \sum_{i=1}^N (\tilde{\xi}_t^{u;\mu})_i f(x_i)$, $f \in C_b$, $t \geq 0$. Then,

$$\begin{aligned}
\mathrm{d}\hat{\xi}_t^{u;\mu}(f) &= \left(\sum_{i=1}^N (\tilde{\xi}_t^{u;\mu})_i f(x_i) h(u_t, x_i) - \sum_{i=1}^N (\tilde{\xi}_t^{u;\mu})_i f(x_i) \sum_{n=1}^N (\tilde{\xi}_t^{u;\mu})_n h(u_t, x_n) \right) \mathrm{d}W_t \\
&= \left(\hat{\xi}_t^{u;\mu} (f h(u_t, \cdot)) - \hat{\xi}_t^{u;\mu} (f) \, \hat{\xi}_t^{u;\mu} (h(u_t, \cdot)) \right) \mathrm{d}W_t,
\end{aligned}$$

which means that the process $\hat{\xi}^{u;\mu}$ is a solution to (17). Since $\tilde{\xi}_0^{u;\mu} \in \Delta^{N-1}$, $\hat{\xi}_0^{u;\mu} \in \mathcal{P}_N$; since every measure-valued process satisfying (17) has preserved mass, $\hat{\xi}^{u;\mu}$ is almost surely $\mathcal{P}_N$-valued. Therefore, the finite-dimensional SDE (23) provides an equivalent representation of (17) when the initial condition has finite support. It is thus sufficient to show that $\tilde{\sigma}(v, \theta)$ is Lipschitz continuous in θ uniformly for all $v \in \mathcal{U}$.

Let θ and η be two distinct points in Δ^{N-1} and $\|h\| = \sup\{h(v, x) : v \in \mathcal{U}, x \in \mathrm{supp}(\mu_0)\}$, which is well-defined by Lemma 4.1. For each $k = 1, \ldots, N$, we have

$$\begin{aligned}
|\tilde{\sigma}_k(v, \theta) - \tilde{\sigma}_k(v, \eta)| &= \left| (\theta_k - \eta_k) h(v, x_k) + \eta_k \sum_{n=1}^N \eta_n h(v, x_n) - \theta_k \sum_{n=1}^N \theta_n h(v, x_n) \right| \\
&\leq |\theta_k - \eta_k| \|h\| + \sum_{n=1}^N |\eta_k \eta_n - \theta_k \eta_n + \theta_k \eta_n - \theta_k \theta_n| \|h\| \\
&\leq 2|\theta_k - \eta_k| \|h\| + \theta_k \sum_{n=1}^N |\theta_n - \eta_n| \|h\|.
\end{aligned}$$

Therefore, applying the triangle inequality and the inequality $(a_1 + \cdots + a_N)^2 \leq N(a_1^2 + \cdots + a_N^2)$, for all $a_i \in \mathbb{R}$, we have

$$\|\tilde{\sigma}(v, \theta) - \tilde{\sigma}(v, \eta)\| \leq 2\|h\| \|\theta - \eta\| + \|h\| \sqrt{N} \|\theta - \eta\| \leq C \|\theta - \eta\|,$$

where C depends on N and $\|h\|$.

(ii). Denote $\tilde{\mu} = (\mu(\{x_1\}), \ldots, \mu(\{x_N\}))$ and $\tilde{\nu} = (\nu(\{x_1\}), \ldots, \nu(\{x_N\}))$. Let $\tilde{\xi}^{u;\mu}$ and $\tilde{\xi}^{u;\nu}$ be solutions to (23) with initial conditions $\tilde{\mu}$ and $\tilde{\nu}$ respectively. Since the function $\tilde{\sigma}(v, \cdot)$ is uniformly Lipschitz for $v \in \mathcal{U}$, by [33, Corollary 2.5.5],

$$\left(\mathbb{E} \left[\left\| \tilde{\xi}_t^{u;\mu} - \tilde{\xi}_t^{u;\nu} \right\| \right] \right)^2 \leq 4 \|\tilde{\mu} - \tilde{\nu}\|^2 + L(C) \|\tilde{\mu} - \tilde{\nu}\|^2 \int_0^t e^{(4C^2+1)(t-s)} \mathrm{d}s,$$

where $L(C)$ is a constant dependent on the uniform Lipschitz constant C of $\tilde{\sigma}$. Hence,

$$\mathbb{E} \left[\left\| \tilde{\xi}_t^{u;\mu} - \tilde{\xi}_t^{u;\nu} \right\| \right] \leq c \|\tilde{\mu} - \tilde{\nu}\|,$$

where c depends on N, t and $\|h\|$. □

Theorem 4.3. *Let $\mu \in \mathcal{P}$ and let μ_k be a sequence in $\mathcal{P}$ such that $\mathcal{W}(\mu_k, \mu) \to 0$, as $k \to \infty$. Consider a filtered probability space supporting a Brownian motion W and a sequence of progressively measurable $\mathcal{U}$-valued processes u_k, $k \in \mathbb{N}$, such that equation* (17) *admits pathwise unique solutions for $u = u_k$ and the initial conditions μ_k and μ, respectively; denote the solutions by $\xi^{u_k;\mu_k}$ and $\xi^{u_k;\mu}$, $k \in \mathbb{N}$. Then, for any $t > 0$,*

$$\mathbb{E} \, |\xi_t^{u_k;\mu_k}(f) - \xi_t^{u_k;\mu}(f)| \xrightarrow[k \to \infty]{} 0, \quad \text{for all } f \in C_b.$$

In particular, there exists a subsequence, which we still index by k, along which

$$\mathcal{W}\,(\xi_t^{u_k;\mu_k}, \xi_t^{u_k;\mu}) \xrightarrow[k\to\infty]{} 0, \quad a.s.$$

Proof. Let u be a $\mathcal{U}$-valued progressively measurable process such that (17) admits pathwise unique solutions $\xi^{u;\mu_k}$ and $\xi^{u;\mu}$ with initial conditions μ_k and μ, $k \in \mathbb{N}$. For ease of notation, we define

$$\delta_t^k(f) = \mathbb{E}\left[|(\xi_t^{u;\mu_k} - \xi_t^{u;\mu})(f)|^2\right], \quad f \in C_b.$$

Step 1: We first show that $\delta_t^k(f) \to 0$ for all $f \in C_b$ when $k \to \infty$, which implies the L_1-convergence. Denote by $\|f\|$ the sup-norm of $f \in C_b$ on $\mathrm{supp}(\mu_0)$, and $\|h\| = \sup\{h(v,x) : v \in \mathcal{U}, x \in \mathrm{supp}(\mu_0)\}$, which is well-defined by Lemma 4.1. Note that for any $k \in \mathbb{N}$,

$$\delta_t^k(f) \le 4\|f\|^2. \tag{24}$$

By (17), and by adding and subtracting the hybrid term $\xi_s^{u;\mu_k}(f)\xi_s^{u;\mu}(h(u_s,\cdot))$, $\delta_t^k(f)$ can be rewritten as

$$\begin{aligned}\delta_t^k(f) = \mathbb{E}\Bigg[\Bigg|(\mu_k - \mu)(f) + \int_0^t \Big\{ &(\xi_s^{u;\mu_k} - \xi_s^{u;\mu})\,(fh(u_s,\cdot)) \\ &+ \xi_s^{u;\mu_k}(f)\,(\xi_s^{u;\mu}\,(h\,(u_s,\cdot)) - \xi_s^{u;\mu_k}\,(h\,(u_s,\cdot))) \\ &+ \xi_s^{u;\mu}(h(u_s,\cdot))\,(\xi_s^{u;\mu}(f) - \xi_s^{u;\mu_k}(f))\Big\}\mathrm{d}W_s\Bigg|^2\Bigg].\end{aligned}$$

Applying the Itô isometry and the inequality $(a_1 + \cdots + a_n)^2 \le n(a_1^2 + \cdots + a_n^2)$, for all $a_i \in \mathbb{R}$, with $n = 2$ and $n = 3$, we get

$$\begin{aligned}\delta_t^k(f) \le{}& 2|(\mu_k - \mu)(f)|^2 + 6\int_0^t \mathbb{E}\left[|(\xi_s^{u;\mu_k} - \xi_s^{u;\mu})\,(fh(u_s,\cdot))|^2\right]\mathrm{d}s \\ &+ 6\|f\|^2\int_0^t \mathbb{E}\left[|(\xi_s^{u;\mu_k} - \xi_s^{u;\mu})\,(h(u_s,\cdot))|^2\right]\mathrm{d}s + 6\|h\|^2\int_0^t \delta_s^k(f)\mathrm{d}s.\end{aligned} \tag{25}$$

Note that for all $f \in C_b$,

$$\begin{aligned}\mathbb{E}\left[|(\xi_s^{u;\mu_k} - \xi_s^{u;\mu})\,(fh(u_s,\cdot))|^2\right] &= \mathbb{E}\left[\left|(\xi_s^{u;\mu_k} - \xi_s^{u;\mu})\left(f\sum_{i=1}^\infty (u_s)_i \mathrm{id}^i\right)\right|^2\right] \\ &= \mathbb{E}\left[\left|\sum_{i=1}^\infty (u_s)_i R^i\,(\xi_s^{u;\mu_k} - \xi_s^{u;\mu})\left(f\frac{\mathrm{id}^i}{R^i}\right)\right|^2\right] \\ &\le \mathbb{E}\left[\left(\sum_{i=1}^\infty (u_s)_i^2 R^{2i}\right)\left(\sum_{i=1}^\infty \left((\xi_s^{u;\mu_k} - \xi_s^{u;\mu})\left(f\frac{\mathrm{id}^i}{R^i}\right)\right)^2\right)\right] \\ &\le K\sum_{i=1}^\infty \delta_s^k\left(f\frac{\mathrm{id}^i}{R^i}\right).\end{aligned}$$

Substituting the above estimate in (25), we obtain the estimate

$$\begin{aligned}\delta_t^k(f) \le{}& 2|(\mu_k - \mu)(f)|^2 \\ &+ 6\int_0^t \left(K\sum_{i=1}^\infty \delta_s^k\left(f\frac{\mathrm{id}^i}{R^i}\right) + K\|f\|^2\sum_{i=1}^\infty \delta_s^k\left(\frac{\mathrm{id}^i}{R^i}\right) + \|h\|^2\delta_s^k(f)\right)\mathrm{d}s.\end{aligned}$$

To further ease the notation, we denote

$$D_k(f) = 2|(\mu_k - \mu)(f)|^2, \quad \psi = \frac{\mathrm{id}}{R}, \quad L = \sum_{i=1}^{\infty} \left\|\psi^i\right\|^2, \quad M = 6\max\left(K, \|h\|^2\right),$$

and get a neater estimate

$$\delta_t^k(f) \leq D_k(f) + M\int_0^t \left(\sum_{i=1}^{\infty} \delta_s^k\left(f\psi^i\right) + \|f\|^2 \sum_{i=1}^{\infty} \delta_s^k\left(\psi^i\right) + \delta_s^k(f)\right) \mathrm{d}s. \tag{26}$$

We first apply the bound (24) to (26) and get a new estimate

$$\begin{aligned}\delta_t^k(f) &\leq D_k(f) + M\int_0^t \left(\sum_{i=1}^{\infty} 4\left\|f\psi^i\right\|^2 + \|f\|^2 \sum_{i=1}^{\infty} 4\left\|\psi^i\right\|^2 + 4\|f\|^2\right)\mathrm{d}s \\ &\leq D_k(f) + 4M(2L+1)\|f\|^2 t. \end{aligned} \tag{27}$$

We can apply (27) to $f\psi^i$, ψ^i, $i \in \mathbb{N}$, and f in (26), and again obtain a second new estimate

$$\begin{aligned}\delta_t^k(f) \leq D_k(f) + M&\left(\sum_{i=1}^{\infty} D_k(f\psi^i) + \|f\|^2 \sum_{i=1}^{\infty} D_k(\psi^i) + D_k(f)\right) t \\ &+ 4M^2(2L+1)^2 \frac{t^2}{2!}\|f\|^2. \end{aligned} \tag{28}$$

Again, we can apply (28) to $f\psi^i$, ψ^i, $i \in \mathbb{N}$, and f in (26), and obtain a third new estimate

$$\begin{aligned}\delta_t^k(f) \leq D_k(f) + M&\left(\sum_{i=1}^{\infty} D_k(f\psi^i) + \|f\|^2 \sum_{i=1}^{\infty} D_k(\psi^i) + D_k(f)\right) t \\ + M^2\Bigg(&\sum_{i_1,i_2=1}^{\infty} D_k(f\psi^{i_1+i_2}) + L\|f\|^2 \sum_{i=1}^{\infty} D_k(\psi^i) + \sum_{i=1}^{\infty} D_k(f\psi^i) \\ &+ \|f\|^2 \sum_{i_1,i_2=1}^{\infty} D_k(\psi^{i_1+i_2}) + L\|f\|^2 \sum_{i=1}^{\infty} D_k(\psi^i) + \|f\|^2 \sum_{i=1}^{\infty} D_k(\psi^i) \\ &+ \sum_{i=1}^{\infty} D_k(f\psi^i) + \|f\|^2 \sum_{i=1}^{\infty} D_k(\psi^i) + D_k(f)\Bigg)\frac{t^2}{2!} \\ + 4M^3&(2L+1)^3 \frac{t^3}{3!}\|f\|^2. \end{aligned}$$

Thus, we can do this iteratively to (26). If we regard (27) as the first iteration, then for the m^{th} iteration, $m \geqslant 2$, $m \in \mathbb{N}$, we obtain the estimate

$$\delta_t^k(f) \leq D_k(f) + M\Delta_1^k t + \cdots + M^{m-1}\Delta_{m-1}^k \frac{t^{m-1}}{(m-1)!} + 4M^m(2L+1)^m \frac{t^m}{m!}\|f\|^2, \tag{29}$$

where each $\Delta_j^k, j = 1, \ldots, m-1$, consists of of 3^j terms and is a linear combination of components of the forms

$$\sum_{i_1,\ldots,i_j=1}^{\infty} D_k(f\psi^{i_1+\cdots+i_j}), \ldots, D_k(f), \sum_{i_1,\ldots,i_j=1}^{\infty} D_k(\psi^{i_1+\cdots+i_j}), \ldots, \sum_{i=1}^{\infty} D_k(\psi^i), \tag{30}$$

with coefficients depending on L and $\|f\|$. Note that for any $j = 1, \dots, m-1$,

$$\sum_{i_1,\dots,i_j=1}^{\infty} D_k(f\psi^{i_1+\dots+i_j}) \le \sum_{i_1,\dots,i_j=1}^{\infty} 8\|f\psi^{i_1+\dots+i_j}\|^2 \le 8\|f\|^2 L^j,$$

which means the series is convergent. Thus, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\sum_{i_1,\dots,i_j=1}^{\infty} D_k(f\psi^{i_1+\dots+i_j}) < \sum_{i_1,\dots,i_j=1}^{N} D_k(f\psi^{i_1+\dots+i_j}) + \varepsilon.$$

Since $f\psi^i$ are Lipschitz for all $i \in \mathbb{N}$ and ε is arbitrary, this means that every component in (30) converges to 0 as $k \to \infty$; since each Δ_j^k has finitely many components with finite coefficients, it follows that all the Δ_j^k, $j = 1, \dots, m-1$, converge to 0 when $k \to \infty$. Therefore, from (29) we get

$$\lim_{k\to\infty} \delta_t^k(f) \le 4M^m(2L+1)^m \frac{t^m}{m!}\|f\|^2.$$

Sending $m \to \infty$, we obtain the desired result. Note, in particular, that the estimates (24) and (26) do not depend on u, so all the consequent iteration estimates do not depend on u. Hence, the convergence is uniform in u.

Step 2: Since we have the L_1-convergence, for every $f \in C_b$, there is a subsequence, which we still index by k, such that

$$|\xi_t^{u_k;\mu_k}(f) - \xi_t^{u_k;\mu}(f)| \xrightarrow[k\to\infty]{} 0 \quad \text{a.s.}$$

We pick a countable subset $\{f_i\}_{i=1}^{\infty} \subset C_b$, where $f_1 \equiv 1$ and $\{f_i\}_{i=2}^{\infty}$ coincides with the polynomials with rational coefficients on $\operatorname{supp}(\mu_0)$. By a diagonalisation argument, there exists a subsequence, which we still index by k, such that

$$|\xi_t^{u_k;\mu_k}(f_i) - \xi_t^{u_k;\mu}(f_i)| \xrightarrow[k\to\infty]{} 0 \quad \text{a.s.,} \quad \text{for all } f_i \in \{f_i\}_{i=1}^{\infty}.$$

Since there are countably many f_i, there exists a probability one set, say A, such that for all $\omega \in A$,

$$|\xi_t^{u_k;\mu_k}(\omega)(f_i) - \xi_t^{u_k;\mu}(\omega)(f_i)| \xrightarrow[k\to\infty]{} 0, \quad \text{for all } f_i \in \{f_i\}_{i=1}^{\infty}.$$

Note that $\{f_i\}_{i=1}^{\infty}$ strongly separates points in $\operatorname{supp}(\mu_0)$ and is closed under multiplication. Therefore, by [9, Theorem 6(a)], along this subsequence we have $\xi^{u_k;\mu_k} \Rightarrow \xi^{u_k;\mu}$ a.s. Since $\operatorname{supp}(\mu_0)$ is compact, this gives us the desired result. □

5. Viscosity theory and proof of main result

In order to establish our main result, we will study in detail the limit of the auxiliary problem. By definition, on $\mathcal{P}$,

$$V^{weak} \le V^+. \tag{31}$$

Our strategy is to show that V^+ is a viscosity subsolution of an HJB equation, a property that ultimately relies on the fact that the DPP holds for the auxiliary problem. From [17], we already know that V^{weak} is a viscosity solution of the same equation, and that this equation does satisfy a comparison principle. Having established that the involved functions are sufficiently smooth, we may thus deduce that (31) holds with equality. Constructing an optimal control for an approximate problem with sufficiently fine grid and closing the loop along the lines outlined in Section 2.2, we may then complete the proof of our main result. The remainder of this section is devoted to carrying out this scheme.

We note that a similar approach was used for the closely related finite-dimensional problem studied in [15]. More pertinently, the idea of justifying an approximation by relying on comparison for the associated HJB equation effectively goes back to [4], meanwhile, the approach of sandwiching the main problem of interest between functions which more easily can be verified to satisfy the HJB equation, resembles the stochastic Perron method developed in [5]. The results in this section develop arguments of this type for the infinite-dimensional problem at hand.

We first recall the definition of viscosity solutions which we make use of in this paper. To this end, let $C^2(\mathcal{P})$ denote the class of functions which are twice continuously differentiable in the sense of [17, Definition 4.7]. Further, recall that continuous MVMs have decreasing support in the sense that, with probability one, $\operatorname{supp}(\xi_t) \subseteq \operatorname{supp}(\xi_s)$, for $s \le t$; see [17, Remark 2.3.(ii)]. Motivated by this fact, consider the partial order $\preceq$ defined on $\mathcal{P}$ by

$$\mu \preceq \nu \quad \Longleftrightarrow \quad \operatorname{supp}(\mu) \subseteq \operatorname{supp}(\nu);$$

MVMs are then decreasing with respect to this order. Utilising this property, the notion of viscosity solutions is then defined as follows:

Definition 5.1 (Definition 6.4 in [17])**.** A function $u : \mathcal{P} \to \mathbb{R}$ is a *viscosity subsolution* (resp. *supersolution*) of (20) if

$$\liminf_{\mu\to\bar\mu,\mu\preceq\bar\mu} H\left(\mu, \varphi(\mu), \frac{\partial^2\varphi}{\partial\mu^2}(\cdot,\cdot;\mu)\right) \le 0 \quad \left(\text{resp. } \limsup_{\mu\to\bar\mu,\mu\preceq\bar\mu} H\left(\mu, \varphi(\mu), \frac{\partial^2\varphi}{\partial\mu^2}(\cdot,\cdot;\mu)\right) \ge 0\right)$$

holds for all $\bar\mu \in \mathcal{P}$ and $\varphi \in C^2(\mathcal{P})$ such that

$$\varphi(\bar\mu) = \limsup_{\mu\to\bar\mu,\ \mu\preceq\bar\mu} u(\mu) \quad \left(\text{resp. } \varphi(\bar\mu) = \liminf_{\mu\to\bar\mu,\ \mu\preceq\bar\mu} u(\mu)\right),$$

and $\varphi(\mu) \ge u(\mu)$ (resp. $\varphi(\mu) \le u(\mu)$) for all $\mu \preceq \bar\mu$.

It is a viscosity solution if it is both a sub- and supersolution.

5.1. **DPP for piecewise constant controls.** The main result of this section is the dynamic programming principle for the auxiliary problem. We start by establishing a continuity property of the auxiliary value function.

Proposition 5.2. *Let $n \in \mathbb{N}$. The function $\hat J(u;\cdot)$ is continuous on $\mathcal{P}$ uniformly in $u \in \mathcal{A}^n$. In consequence, the function V^n is continuous on $\mathcal{P}$.*

Proof. The second part follows immediately from the first since, for any $\mu, \nu \in \mathcal{P}$,

$$|V^n(\mu) - V^n(\nu)| \le \sup_{u\in\mathcal{A}^n} \left|\hat J(u;\mu) - \hat J(u;\nu)\right|.$$

To argue the first part, let $\mu \in \mathcal{P}$ and consider a sequence μ_k in $\mathcal{P}$ such that $\mu_k \to \mu$, and an arbitrary sequence u_k in $\mathcal{A}^n$. For any $\varepsilon > 0$, there exists some $T > 0$ such that, for all $k \in \mathbb{N}$,

$$\begin{aligned} \left|\hat J(u_k;\mu_k) - \hat J(u_k;\mu)\right| &\le \mathbb{E}\int_0^T \left|k\big(\xi_t^{u_k;\mu_k}, (u_k)_t\big) - k\big(\xi_t^{u_k;\mu}, (u_k)_t\big)\right| \mathrm{d}t + \varepsilon \\ &\le T w\left(\mathbb{E}\,\mathcal{W}\left(\xi_T^{u_k;\mu_k}, \xi_T^{u_k;\mu}\right)\right) + \varepsilon, \end{aligned} \tag{32}$$

where we used that, by the standing Assumption 2.1, $\mu \mapsto k(\mu, v)$ is continuous, uniformly in $v \in \mathcal{U}$, and thus admits a concave and continuous modulus of continuity w on the compact set $\mathcal{P}$, and the fact that $\mathcal{W}(\xi^{u_k;\mu_k}, \xi^{u_k;\mu})$ is a submartingale for $k \in \mathbb{N}$ (c.f. [6, Proposition 4.1.(i)]).

Now, for any subsequence of (μ_k, u_k), according to Theorem 4.3, there exists a further subsequence, which we still index by k, along which $\mathcal{W}(\xi_T^{u_k;\mu_k}, \xi_T^{u_k;\mu})$ converges to zero a.s., and along which the first term on the right-hand side of (32) thus converges to zero; since any subsequence of the original sequence admits a further subsequence with this property, the convergence must hold also along the original sequence. Since $\varepsilon > 0$ was arbitrarily chosen, this completes the proof. □

Next, we argue that a certain 'pseudo Markov property' holds within our context. To this end, let $n \in \mathbb{N}$ and recall the definition of δ_n from Section 3.2. Given a control $u \in \mathcal{A}^n$, we define a new control $u^{\delta_n,\omega} \in \mathcal{A}^n$, for $\omega \in \Omega$, by

$$u_t^{\delta_n,\omega}(\tilde{\omega}) = u_{\delta_n+t}(\omega \otimes_{\delta_n} \tilde{\omega}), \quad (t, \tilde{\omega}) \in [0,\infty) \times \Omega,$$

where

$$(\omega \otimes_{\delta_n} \tilde{\omega})(t) = \omega(t)\mathbf{1}_{[0,\delta_n)}(t) + (\omega(\delta_n) + \tilde{\omega}(t-\delta_n) - \tilde{\omega}(0))\mathbf{1}_{[\delta_n,\infty)}(t), \quad t \geq 0.$$

The following result can then be deduced following the arguments in [14] and making use of the pathwise uniqueness of the flow solving (19) (the arguments simplify considerably since we here consider piecewise constant controls):

Lemma 5.3. *Let $n \in \mathbb{N}$ and $u \in \mathcal{A}^n$. Then, for every $\mu \in \mathcal{P}$ and $\mathbb{P}$-a.a. ω,*

$$\mathbb{E}\left[\int_{\delta_n}^{\infty} e^{-\beta(t-\delta_n)} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t \big| \mathcal{F}_{\delta_n}\right](\omega) = \int_\Omega \int_0^\infty e^{-\beta t} k\left(\xi_t^{u^{\delta_n,\omega}, \xi_{\delta_n}^{u;\mu}(\omega)}(\tilde{\omega}), u_t^{\delta_n,\omega}(\tilde{\omega})\right) \mathrm{d}t \, \mathrm{d}\mathbb{P}(\tilde{\omega}).$$

Thanks to the above continuity and pseudo Markov properties, we may establish the DPP following the arguments developed in [10]. In turn, by use of the DPP, we may prove existence of an ε-optimal strategy for the auxiliary problem. We present those results next; for completeness, we provide the full proofs in the present set-up in Appendix A.

For $v \in \mathcal{U}$, we write $\xi^{v;\mu}$ for the solution to (19) using the constant control $u_s \equiv v$.

Theorem 5.4 (DPP)**.** *Let $n \in \mathbb{N}$. For every $\mu \in \mathcal{P}$,*

$$V^n(\mu) = \inf_{v \in \mathcal{U}} \mathbb{E}\left[\int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{v;\mu}, v\right) \mathrm{d}t + e^{-\beta\delta_n} V^n\left(\xi_{\delta_n}^{v;\mu}\right)\right].$$

Corollary 5.5. *For any $n \in \mathbb{N}$ and $\varepsilon > 0$, there exists a measurable function $\hat{u}^n : \mathcal{P} \to \mathcal{U}$, such that for $\mu \in \mathcal{P}$,*

$$\hat{J}(u^{*,n}; \mu) \leq V^n(\mu) + \varepsilon,$$

where $u^{*,n} \in \mathcal{A}^n$ is given by

$$u_t^{*,n} = \sum_{r \in \mathbb{T}^n} \hat{u}^n\left(\xi_r^{u^{*,n},\mu}\right) \mathbf{1}_{[r,r+\delta_n)}(t), \quad t \in [0,\infty). \tag{33}$$

5.2. **Subsolution property of limiting problem.** In this section we establish the subsolution property of V^+. To ease the notation, we introduce the effective state space for an MVM starting at a measure $\bar{\mu} \in \mathcal{P}$; it is given by the set

$$D_{\bar{\mu}} = \{\mu \in \mathcal{P} \colon \mu \preceq \bar{\mu}\}.$$

Making use of the fact that, for any $u \in C^2(\mathcal{P})$, the left-hand side of (20) is lower semicontinuous in μ, and the fact that V^+ is upper semicontinuous on $\mathcal{P}$ (being the decreasing limit of continuous functions by Proposition 5.2), we note that V^+ is a subsolution of (20), if and only if, for any $\bar{\mu} \in \mathcal{P}$ and $\varphi \in C^2(\mathcal{P})$, one has the implication

$$\varphi(\bar{\mu}) = V^+(\bar{\mu}) \text{ and } \varphi > V^+ \text{ on } D_{\bar{\mu}} \setminus \{\bar{\mu}\} \implies H\left(\bar{\mu}, \varphi(\bar{\mu}), \frac{\partial^2 \varphi}{\partial \mu^2}(\cdot, \cdot; \bar{\mu})\right) \leq 0. \tag{34}$$

Indeed, the 'only if' is immediate. Meanwhile, the 'if' follows by use of the same arguments as employed in the proof of [17, Lemma 6.6].

Lemma 5.6. *Let $\bar{\mu} \in \mathcal{P}$ and $\varphi \in C^2(\mathcal{P})$ such that $\varphi(\bar{\mu}) = V^+(\bar{\mu})$ and $V^+ - \varphi < 0$ on $D_{\bar{\mu}} \setminus \{\bar{\mu}\}$. Then, there exist sequences $m_n \to \infty$ and $\mu_n \to \bar{\mu}$, $\mu_n \in D_{\bar{\mu}}$, such that μ_n attains the maximum of $V^{m_n} - \varphi$ on $D_{\bar{\mu}}$ and $V^{m_n}(\mu_n) \to V^+(\bar{\mu})$. In particular, $V^{m_n} \leq \varphi + \kappa_n$, where $\kappa_n := (V^{m_n} - \varphi)(\mu_n) \downarrow 0$.*

Proof. For $n \in \mathbb{N}$, let $N_n \geq n$ and $\varepsilon_n \leq 1/n$ such that, for all $m \geq N_n$, $\mu, \nu \in D_{\bar{\mu}} \cap B_{\varepsilon_n}(\bar{\mu})$,

$$V^m(\mu) \leq V^+(\bar{\mu}) + 1/n \quad \text{and} \quad |\varphi(\nu) - \varphi(\mu)| \leq 1/n;$$

existence is guaranteed by Proposition 5.2. By the upper semicontinuity of V^+ (c.f. Proposition 5.2), there exists $\eta_n > 0$ such that $V^+ - \varphi \leq -2\eta_n$ on $D_{\bar{\mu}} \setminus B_{\varepsilon_n}(\bar{\mu})$. By possibly making N_n even larger, we may ensure that, for all $m \geq N_n$, $\mu \in D_{\bar{\mu}} \setminus B_{\varepsilon_n}(\bar{\mu})$,

$$(V^m - \varphi)(\mu) < -\eta_n. \tag{35}$$

Indeed, for any $\nu \in D_{\bar{\mu}} \setminus B_{\varepsilon_n}(\bar{\mu})$, there exist N^ν and ε^ν such that (35) holds for all $m \geq N^\nu$, $\mu \in B_{\varepsilon^\nu}(\nu)$; a finite covering argument yields the claim.

Now, for $n \in \mathbb{N}$, pick $m_n \geq N_n$ and let

$$\mu_n := \operatorname*{arg\,max}_{\mu \in D_{\bar{\mu}}} (V^{m_n} - \varphi)(\mu).$$

Since $(V^{m_n} - \varphi)(\bar{\mu}) \geq 0$, it follows that $\mu_n \in B_{\varepsilon_n}(\bar{\mu})$. We note that

$$V^{m_n}(\mu_n) - V^+(\bar{\mu}) \geq \varphi(\mu_n) - \varphi(\bar{\mu}) \geq -1/n,$$

which completes the proof. □

Theorem 5.4 together with Lemma 5.6 immediately gives the following version of the DPP; we here focus on the inequality which will be needed for proving the subsolution property below and note that this inequality relies on the 'deep inequality' of the DPP:

Corollary 5.7. *Let $\bar{\mu} \in \mathcal{P}$ and $\varphi \in C^2(\mathcal{P})$ such that $\varphi(\bar{\mu}) = V^+(\bar{\mu})$ and $V^+ - \varphi < 0$ on $D_{\bar{\mu}} \setminus \{\bar{\mu}\}$. Then there exist sequences $m_n \to \infty$, $\mu_n \to \bar{\mu}$, $\mu_n \in D_{\bar{\mu}}$, and $\kappa_n \downarrow 0$, such that*

$$\varphi(\mu_n) \leq \inf_{v \in \mathcal{U}} \mathbb{E}\left[\int_0^{\delta_{m_n}} e^{-\beta s} k\left(\xi_s^{v;\mu_n}, v\right) \mathrm{d}s + e^{-\beta \delta_{m_n}} \varphi\left(\xi_{\delta_{m_n}}^{v;\mu_n}\right)\right] - \left(1 - e^{-\beta \delta_{m_n}}\right) \kappa_n.$$

Next we establish a certain 'consistency property':

Proposition 5.8. *Let $\bar{\mu} \in \mathcal{P}$. For any $\varphi \in C^2(\mathcal{P})$,*

$$\liminf_{\substack{h \downarrow 0 \\ \mu \to \bar{\mu}}} \frac{1}{h}\left\{\varphi(\mu) - \inf_{v \in \mathcal{U}} \mathbb{E}\left[\int_0^h e^{-\beta s} k\left(\xi_s^{v;\mu}, v\right) \mathrm{d}s + e^{-\beta h} \varphi\left(\xi_h^{v;\mu}\right)\right]\right\}$$
$$\geq H\left(\bar{\mu}, \varphi(\bar{\mu}), \frac{\partial^2 \varphi}{\partial \mu^2}(\cdot, \cdot, \bar{\mu})\right).$$

Proof. Let $\mu \in \mathcal{P}$, $v \in \mathcal{U}$ and $\varphi \in C^2(\mathcal{P})$. By Lemma 4.1, $h(v,\cdot)$ is bounded on $\operatorname{supp}(\mu_0)$. Hence, for any $f \in C_b$,

$$\int_0^t \left(\int_{\mathbb{R}} f(x)\, |h(v,x) - \xi_s^{v;\mu}(h(v,\cdot))|\, \xi_s^{v;\mu}(\mathrm{d}x) \right)^2 \mathrm{d}s < \infty, \quad a.s. \tag{36}$$

We may thus apply the Itô formula [17, Theorem 5.1] to obtain

$$\begin{aligned} e^{-\beta t}\varphi\left(\xi_t^{v;\mu}\right) - \varphi(\mu) &= \int_0^t e^{-\beta s}\left(-\beta\varphi(\xi_s^{v;\mu}) + L\varphi(\xi_s^{v;\mu}, v)\right)\mathrm{d}s \\ &\quad + \int_0^t e^{-\beta s} \int_{\mathbb{R}} \frac{\partial \varphi}{\partial \mu}(x; \xi_s^{v;\mu})\sigma_s(\mathrm{d}x)\mathrm{d}W_s, \end{aligned} \tag{37}$$

where $\sigma(\mu, v; \mathrm{d}x) = (h(v,x) - \mu(h(v,\cdot)))\mu(\mathrm{d}x)$ and

$$L\varphi(\mu, v) = \frac{1}{2} \int_{\mathbb{R}\times\mathbb{R}} \frac{\partial^2 \varphi}{\partial \mu^2}(x, y; \mu)\sigma(\mu, v; \mathrm{d}x)\sigma(\mu, v; \mathrm{d}y).$$

Making once again use of (36), we see that the last integral in (37) is in fact a martingale; we thus obtain

$$e^{-\beta t}\mathbb{E}\left[\varphi\left(\xi_t^{v;\mu}\right)\right] - \varphi(\mu) = \mathbb{E}\int_0^t e^{-\beta s}\Big(-\beta\varphi(\xi_s^{v;\mu}) + L\varphi(\xi_s^{v;\mu}, v)\Big)\mathrm{d}s, \quad t \geq 0. \tag{38}$$

Next, consider sequences $h_k \downarrow 0$ and $\mu_k \to \mu$; without loss of generality, suppose that $h_k \leq 1$, $k \in \mathbb{N}$. For $v \in \mathcal{U}$, by Lemma 4.1, $h(v,\cdot)$ is bounded and continuous on $\operatorname{supp}(\mu_0)$ and therefore $L\varphi(\cdot, v)$ is continuous and thus admits a concave and continuous modulus of continuity, say w, on the compact set $\mathcal{P}$. Moreover, $\mathcal{W}(\xi^{v;\mu_k}, \xi^{v;\mu})$ is a submartingale for $k \in \mathbb{N}$ (c.f. [6, Proposition 4.1.(i)]). Hence,

$$\frac{1}{h_k}\mathbb{E}\left[\int_0^{h_k} \Big| L\varphi(\xi_s^{v;\mu_k}, v) - L\varphi(\xi_s^{v;\mu}, v)\Big|\mathrm{d}s\right] \leq w\left(\mathbb{E}\,\mathcal{W}\left(\xi_1^{v;\mu_k}, \xi_1^{v;\mu}\right)\right) \xrightarrow[n\to\infty]{} 0, \tag{39}$$

where the convergence follows from Theorem 4.3. Indeed, according to this result, for any subsequence of μ_k, there exists a further subsequence, which we still index by k, along which $\mathcal{W}(\xi_1^{v;\mu_k}, \xi_1^{v;\mu}) \to 0$ a.s., and along which the right-hand side of (39) thus converges to zero; since any subsequence of the original sequence admits a further subsequence with this property, the convergence must hold also along the original sequence. The same argument applies when replacing $L\varphi(\cdot, v)$ by $\beta\varphi(\cdot)$ and $k(\cdot, v)$, making use of the standing Assumption 2.1.

Finally, for $\mu \in \mathcal{P}$, define

$$\mathcal{T}_h\varphi(\mu) := \inf_{v\in\mathcal{U}} \mathbb{E}\left[\int_0^h e^{-\beta s} k\left(\xi_s^{v;\mu}, v\right)\mathrm{d}s + e^{-\beta h}\varphi\left(\xi_h^{v;\mu}\right)\right].$$

Fix now $\bar{\mu} \in \mathcal{P}$ and let $h_k \downarrow 0$ and $\mu_k \to \bar{\mu}$ such that

$$\liminf_{\substack{h\downarrow 0\\ \mu\to\bar{\mu}}} \frac{1}{h}\big(\varphi(\mu) - \mathcal{T}_h\varphi(\mu)\big) = \lim_{k\to\infty} \frac{1}{h_k}\big(\varphi(\mu_k) - \mathcal{T}_{h_k}\varphi(\mu_k)\big).$$

Let now also $v \in \mathcal{U}$ be fixed. Making use of (38) and (39), we then obtain

$$\begin{aligned}
&\lim_{k\to\infty} \frac{1}{h_k}\big(\varphi(\mu_k) - \mathcal{T}_{h_k}\varphi(\mu_k)\big)\\
&\qquad \geq \limsup_{k\to\infty} \frac{1}{h_k}\left\{\varphi(\mu_k) - \mathbb{E}\Big[\int_0^{h_k} e^{-\beta s} k\left(\xi_s^{v;\mu_k}, v\right) \mathrm{d}s + e^{-\beta h_k}\varphi\left(\xi_{h_k}^{v;\mu_k}\right)\Big]\right\}\\
&\qquad = \limsup_{k\to\infty} \frac{1}{h_k}\mathbb{E}\Big[\int_0^{h_k} e^{-\beta s}\Big(\beta\varphi(\xi_s^{v;\mu_k}) - k\left(\xi_s^{v;\mu_k}, v\right) - L\varphi(\xi_s^{v;\mu_k}, v)\Big)\mathrm{d}s\Big]\\
&\qquad \geq \liminf_{k\to\infty} \frac{1}{h_k}\mathbb{E}\Big[\int_0^{h_k} e^{-\beta s}\Big(\beta\varphi(\xi_s^{v;\mu}) - k\left(\xi_s^{v;\mu}, v\right) - L\varphi(\xi_s^{v;\mu}, v)\Big)\mathrm{d}s\Big]\\
&\qquad = \beta\varphi(\mu) - k\left(\mu, v\right) - L\varphi(\mu, v),
\end{aligned}$$

where we used dominated convergence in the last step. Since $v \in \mathcal{U}$ was arbitrarily chosen, we may now take the supremum over $v \in \mathcal{U}$ to conclude. □

We are now ready to establish that the limit of the auxiliary problem is a subsolution of the HJB equation.

Theorem 5.9. *The function V^+ is a viscosity subsolution of* (20).

Proof. For $\bar{\mu} \in \mathcal{P}^s$, the subsolution property reduces to $\beta V^+(\bar{\mu}) \leq \inf_{v\in\mathcal{U}} k(\bar{\mu}, v)$, which for Dirac measures holds by the definition of V^+; c.f. Remark 3.2. Let $\bar{\mu} \in \mathcal{P} \setminus \mathcal{P}^s$ and consider $\varphi \in C^2(\mathcal{P})$ such that $\varphi(\bar{\mu}) = V^+(\bar{\mu})$ and $V^+ - \varphi < 0$ on $D_{\bar{\mu}} \setminus \{\bar{\mu}\}$. By Corollary 5.7, we can then find sequences $m_n \to \infty$ and $\mu_n \to \bar{\mu}$, such that

$$\begin{aligned}
0 &\geq \limsup_{n\to\infty} \frac{1}{\delta_{m_n}}\left\{\varphi(\mu_n) - \inf_{v\in\mathcal{U}}\mathbb{E}\left[\int_0^{\delta_{m_n}} e^{-\beta s} k\left(\xi_s^{v;\mu_n}, v\right) \mathrm{d}s + e^{-\beta\delta_{m_n}}\varphi\left(\xi_{\delta_{m_n}}^{v;\mu_n}\right)\right]\right\}\\
&\geq \liminf_{\substack{h\downarrow 0\\ \mu\to\bar{\mu}}} \frac{1}{h}\left\{\varphi(\mu) - \inf_{v\in\mathcal{U}}\mathbb{E}\left[\int_0^{h} e^{-\beta s} k\left(\xi_s^{v;\mu}, v\right) \mathrm{d}s + e^{-\beta h}\varphi\left(\xi_h^{v;\mu}\right)\right]\right\}\\
&\geq H\left(\bar{\mu}, \varphi(\bar{\mu}), \frac{\partial^2\varphi}{\partial\mu^2}(\cdot,\cdot,\bar{\mu})\right),
\end{aligned}$$

where the last inequality follows from Proposition 5.8. According to (34), this completes the proof. □

5.3. **Proof of main result.** We first recall two results from [17] which ensure that we do indeed have comparison for the HJB equation and that the weak value function is indeed a viscosity solution of this equation. To this end, given N points $x_1, \dots, x_N \in \operatorname{supp}(\mu_0)$, note that any given function $u : \mathcal{P} \to \mathbb{R}$ induces a function $\tilde{u} : \Delta^{N-1} \to \mathbb{R}$ defined by

$$\tilde{u}(p_1, \dots, p_N) = u(p_1\delta_{x_1} + \cdots + p_n\delta_{x_n}). \tag{40}$$

Theorem 5.10 (Theorem 9.1 in [17])**.** *Let u and v be viscosity sub- and supersolutions, respectively, of* (20)*. Suppose the following hold:*

(i) for any N points $x_1, \dots, x_N \in \operatorname{supp}(\mu_0)$, the functions $\tilde{u}$ and $\tilde{v}$ obtained from u and v via (40) *are, respectively, upper and lower semicontinuous on Δ^{N-1};*
(ii) for any $\mu \in \mathcal{P}$, there exists a sequence of finitely supported $\mu_n \in \mathcal{P}$ such that

$$u(\mu) \leq \liminf_{n\to\infty} u(\mu_n) \quad \text{and} \quad v(\mu) \geq \limsup_{n\to\infty} v(\mu_n).$$

Then, $u \leq v$. In particular, (20) *admits a unique continuous viscosity solution.*

Proof. By Lemma 4.1, the set $\{h(v,x) : v \in \mathcal{U}, x \in \mathrm{supp}(\mu_0)\}$ is a bounded subset of $\mathbb{R}$. Moreover, by the standing Assumption 2.1, $\mu \mapsto k(\mu, v)$ is continuous on $\mathcal{P}$ uniformly in $v \in \mathcal{U}$. Hence, Assumptions (i) and (ii) of [17, Theorem 9.1] are satisfied. Note first that [17, Lemma 9.2 and 9.3] still hold when relaxing their continuity assumptions on $\tilde{u}, \tilde{v}$ to only imposing our (i). Moreover, by our (ii),

$$(u-v)(\mu) \leq \liminf_{n\to\infty}(u-v)(\mu_n).$$

The result can therefore be deduced along the same lines as [17, Theorem 9.1] (making obvious modifications to account for the fact that our $\mathcal{P}$ is a subset of the space considered therein). □

Theorem 5.11 (Theorem 6.2 in [17]). *The function V^{weak} is a viscosity solution of* (20).

Proof. By Lemma 4.1, $\{h(v,x) : v \in \mathcal{U}, x \in \mathrm{supp}(\mu_0)\}$ is a bounded subset of $\mathbb{R}$, and $h(v,\cdot)$ is continuous on $\mathrm{supp}(\mu_0)$. By the standing Assumption 2.1, we also have that $\mu \mapsto k(\mu,v)$ is continuous for every $v \in \mathcal{U}$. Hence, Assumptions (i)–(iii) of [17, Theorem 6.2] are satisfied, and the result follows. □

We recall that V^+ is upper semicontinuous on $\mathcal{P}$; we now argue that it is even continuous. Indeed, note that

$$\left|V^+(\mu) - V^+(\nu)\right| \leq \sup_{n\in\mathbb{N}} |V^n(\mu) - V^n(\nu)| \leq \sup_{n\in\mathbb{N}} \sup_{u\in\mathcal{A}^n} \left|\hat{J}(u;\mu) - \hat{J}(u;\nu)\right|. \tag{41}$$

It thus suffices to argue that the function $\hat{J}(u;\cdot)$ is continuous on $\mathcal{P}$ uniformly in $u \in \cup_{n\in\mathbb{N}}\mathcal{A}^n$; this follows however by use of the same arguments as used to prove Proposition 5.2. We next establish continuity properties of the weak value function:

Lemma 5.12. *Assumptions (i) and (ii) of Theorem 5.10 hold for $v = V^{weak}$.*

Proof. (i) For a finitely supported $\mu \in \mathcal{P}$, according to Proposition 4.2 (i), given a tuple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W, u)$ such that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space, W is an $(\mathbb{F}, \mathbb{P})$-Brownian motion, and u is an $\mathcal{U}$-valued and $\mathbb{F}$-progressively measurable process, there exists a pathwise unique solution to (17) equipped with the initial condition μ. Hence,

$$V^{weak}(\mu) = \inf_{(\Omega,\mathcal{F},\mathbb{F},\mathbb{P},W,u)} \mathbb{E}\int_0^\infty e^{-\beta t} k(\xi_t^{u;\mu}, u_t)\mathrm{d}t,$$

where the infimum is taken over all such tuples and the process $\xi^{u;\mu}$ denotes that associated solution to (17). In consequence, for any two finitely supported measures $\mu, \nu \in \mathcal{P}$,

$$\left|V^{weak}(\mu) - V^{weak}(\nu)\right| \leq \sup_{(\Omega,\mathcal{F},\mathbb{F},\mathbb{P},W,u)} \mathbb{E}\int_0^\infty e^{-\beta t} \left|k(\xi_t^{u;\mu}, u_t) - k(\xi_t^{u;\nu}, u_t)\right| \mathrm{d}t.$$

Moreover, for any $\varepsilon > 0$, there exists some $T > 0$ such that, for any finitely supported measures $\mu, \nu \in \mathcal{P}$, and any tuple $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W, u)$,

$$\begin{aligned}\mathbb{E}\int_0^\infty e^{-\beta t} \left|k(\xi_t^{u;\mu}, u_t) - k(\xi_t^{u;\nu}, u_t)\right| \mathrm{d}t &\leq \mathbb{E}\int_0^T \left|k(\xi_t^{u;\mu}, u_t) - k(\xi_t^{u;\nu}, u_t)\right| \mathrm{d}t + \varepsilon \\ &\leq Tw\left(\mathbb{E}\,\mathcal{W}\left(\xi_T^{u;\mu}, \xi_T^{u;\nu}\right)\right) + \varepsilon,\end{aligned}$$

where we used that, by the standing Assumption 2.1, $\mu \mapsto k(\mu,v)$ is continuous, uniformly in $v \in \mathcal{U}$, and thus admits a concave and continuous modulus of continuity w on the compact set $\mathcal{P}$, and the fact that $\mathcal{W}(\xi^{u;\mu}, \xi^{u;\nu})$ is a submartingale (c.f. [6, Proposition 4.1.(i)]).

Restricting now to $\mu, \nu \in \mathcal{P}(\{x_1, \ldots, x_N\})$ and applying Proposition 4.2 (ii), we have that

$$\mathbb{E}\, \mathcal{W}\left(\xi_T^{u;\mu}, \xi_T^{u;\nu}\right) \leq \max_{i \neq j} |x_i - x_j| \sqrt{N} \mathbb{E} \left\| ((\xi_T^{u;\mu} - \xi_T^{u;\nu})(\{x_1\}), \ldots, (\xi_T^{u;\mu} - \xi_T^{u;\nu})(\{x_N\})) \right\|$$
$$\leq \max_{i \neq j} |x_i - x_j| \sqrt{N} c \left\| ((\mu - \nu)(\{x_1\}), \ldots, (\mu - \nu)(\{x_1\})) \right\|,$$

where $\|\cdot\|$ denotes the Euclidean norm and c is a constant depending on T, N and $\sup\{h(v,x) : v \in \mathcal{U}, x \in \mathrm{supp}(\mu_0)\}$. Since, with the identification between $\alpha \in \Delta^{N-1}$ and $\mu \in \mathcal{P}(\{x_1, \ldots, x_N\})$ given by $\mu = \alpha_1 \delta_{x_1} + \cdots + \alpha_N \delta_{x_N}$, the Euclidean and the Wasserstein distances are topologically equivalent, and since the obtained bound is independent of the chosen tuple, this yields continuity of V^{weak} on $\mathcal{P}(\{x_1, \ldots, x_N\})$.

(ii) Let $\mu \in \mathcal{P}$ and let $\mu_k \in \mathcal{P}$ be a sequence of finitely supported measures such that $\mathcal{W}(\mu_k, \mu) \to 0$. For $\varepsilon > 0$, let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W, \xi, u) \in \mathcal{A}^{weak}(\mu)$ such that

$$\mathbb{E} \int_0^\infty e^{-\beta t} k(\xi_t, u_t) \mathrm{d}t \leq V^{weak}(\mu) + \varepsilon/2.$$

In turn, let $\xi^{u;\mu_k}$ be the solution to (17) constructed on the same probability space and with respect to the same control u but for the initial condition μ_k; it exists and is pathwise unique by Proposition 4.2 (i). Then, there exists some $T > 0$, such that

$$\limsup_{k \to \infty} V^{weak}(\mu_k) - V^{weak}(\mu) \leq \limsup_{k \to \infty} \mathbb{E} \int_0^T e^{-\beta t} \left(k(\xi_t^{u;\mu_k}, u_t) - k(\xi_t, u_t) \right) \mathrm{d}t + \varepsilon.$$

Making use of Theorem 4.3 and the standing Assumption 2.1, and arguing along the same lines as in the proof of Proposition 5.2, we obtain that the limsup on the right-hand side must equal zero. Since ε was arbitrarily chosen, this completes the proof. □

We are now ready to conclude the proof of our main result.

Proof of Theorem 3.3. Given $\varepsilon > 0$, let $n \in \mathbb{N}$ such that $V^n(\mu_0) \leq V^+(\mu_0) + \varepsilon$ and let $\hat{u}^n$ be the function provided by Corollary 5.5 for which $\hat{J}(u^{*,n}; \mu_0) \leq V^n(\mu_0) + \varepsilon$ when $u^{*,n} \in \mathcal{A}^n$ is defined by (33) (in terms of $\hat{u}^n$). Define, in turn, u^* by (21) (in terms of $\hat{u}^n$). Specifically, this feedback control is defined as the solution to the system (3) and (4); since the feedback function in this case is piecewise constant, it is well defined and thus belongs to $\mathcal{A}$.

We recall that $\pi_0^{u^*} = \mu_0$ and that π^{u^*} satisfies (8). Making again use of the fact that the control is piecewise constant, according to [34, Theorem 2.1], we have pathwise uniqueness for the SDE (8) with u given by (21) (viewing (8) as an equation for π^{u^*} with I^{u^*} a fixed $(\mathcal{Y}^{u^*}, \mathbb{P})$-Brownian motion), or, equivalently, for the SDE (19) with u given by (33). We therefore also have uniqueness in law (c.f. [40, proof of Theorem V.3]) and it follows that $J(u^*) = \hat{J}(u^{*,n}; \mu_0)$. In consequence,

$$J(u^*) \leq V^+(\mu_0) + 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrarily chosen, we obtain in particular that

$$V^{weak}(\mu_0) \leq \inf_{u \in \mathcal{A}} J(u) \leq V^+(\mu_0). \tag{42}$$

Next, by Theorem 5.9 and Theorem 5.11, $u = V^+$ and $v = V^{weak}$ are sub- and supersolutions, respectively, of (20). Moreover, by Lemma 5.12 and the preceding discussion (c.f. (41)), the continuity assumptions of Theorem 5.10 hold for $u = V^+$ and $v = V^{weak}$. Application of this comparison result then yields $V^+ \leq V^{weak}$ on $\mathcal{P}$, which implies that both (31) and (42) hold as equalities. Applying Theorem 5.10 once again, making once again use of the continuity of V^+ (c.f. (41)), yields the uniqueness and continuity claims. □

Appendix A. Proof of the DPP

Proof of Theorem 5.4. '$\geq$': For any $\mu \in \mathcal{P}$ and $u \in \mathcal{A}^n$, by use of Lemma 5.3 we obtain

$$\begin{aligned}
\hat{J}(u;\mu) &= \mathbb{E}\left[\int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t + e^{-\beta\delta_n}\mathbb{E}\left[\int_{\delta_n}^{\infty} e^{-\beta(t-\delta_n)} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t \big| \mathcal{F}_{\delta_n}\right]\right] \\
&= \int_\Omega \Bigg\{ \int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}(\omega), u_t(\omega)\right) \mathrm{d}t \\
&\quad + e^{-\beta\delta_n} \int_\Omega \int_0^\infty e^{-\beta t} k\Big(\xi_t^{u^{\delta_n,\omega}, \xi_{\delta_n}^{u;\mu}(\omega)}(\tilde{\omega}), u_t^{\delta_n,\omega}(\tilde{\omega})\Big) \mathrm{d}t \, \mathrm{d}\mathbb{P}(\tilde{\omega}) \Bigg\} \mathrm{d}\mathbb{P}(\omega) \qquad (43) \\
&\geq \int_\Omega \Bigg\{ \int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}(\omega), u_t(\omega)\right) \mathrm{d}t + e^{-\beta\delta_n} V^n\left(\xi_{\delta_n}^{u;\mu}(\omega)\right) \Bigg\} \mathrm{d}\mathbb{P}(\omega) \\
&= \mathbb{E}\left[\int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t + e^{-\beta\delta_n} V^n\left(\xi_{\delta_n}^{u;\mu}\right)\right];
\end{aligned}$$

taking the infimum over $u \in \mathcal{A}^n$ on both sides yields the desired inequality.

'$\leq$': Let $\varepsilon > 0$ and let $\{u^\mu\}_{\mu\in\mathcal{P}} \subset \mathcal{A}^n$ be a family such that

$$\hat{J}(u^\mu;\mu) \leq V^n(\mu) + \varepsilon/3, \quad \text{for every } \mu \in \mathcal{P}.$$

From Proposition 5.2 we have uniform continuity of V^n and $\hat{J}(u;\cdot)$, $u \in \mathcal{A}^n$; hence, there exists $r > 0$ such that

$$V^n(\mu) - V^n(\nu) \leq \varepsilon/3 \quad \text{and} \quad \hat{J}(u^\mu;\mu) - \hat{J}(u^\mu;\nu) \geq -\varepsilon/3, \quad \text{for all } \nu \in B_r(\mu).$$

Clearly $\{B_r(\mu) : \mu \in \mathcal{P}\}$ forms an open cover of $\mathcal{P}$ and thus there exists a finite sequence $\mu_1, \dots \mu_N \in \mathcal{P}$ such that $\mathcal{P} \subset \cup_{i=1}^N B_r(\mu_i)$. Defining

$$A_1 := B_r(\mu_1), \qquad A_i := B_r(\mu_i) \setminus \{A_1 \cup \cdots \cup A_{i-1}\}, \quad i = 2, \dots N,$$

we have that $\mathcal{P} \subset \cup_{i=1}^N A_i$ and $A_i \cap A_j = \emptyset$, $i \neq j$. Defining $u^i = u^{\mu_i}$, $i = 1, \dots, N$, it then holds that

$$\hat{J}(u^i;\mu) \leq \hat{J}(u^i;\mu_i) + \varepsilon/3 \leq V^n(\mu_i) + 2\varepsilon/3 \leq V^n(\mu) + \varepsilon, \quad \text{for all } \mu \in A_i. \qquad (44)$$

Next, let $\mu \in \mathcal{P}$ and let $u \in \mathcal{A}^n$ be an arbitrary control. In turn, define a control $\hat{u} \in \mathcal{A}^n$ by

$$\hat{u}_t(\omega) = \begin{cases} u_t(\omega) & t \in [0,\delta_n) \\ \sum_{i=1}^N \mathbf{1}_{A_i}\left(\xi_{\delta_n}^{u;\mu}(\omega)\right) u_{t-\delta_n}^i\left(\theta_{\delta_n}\omega\right) & t \in [\delta_n, \infty) \end{cases} \qquad (t,\omega) \in [0,\infty) \times \Omega,$$

where

$$\theta_{\delta_n}\omega(t) = \omega(\delta_n + t) - \omega(\delta_n).$$

Making once again use of Lemma 5.3, and the fact that for each $i = 1, \dots, N$ and $\omega \in \Omega$ such that $\xi_{\delta_n}^{u;\mu}(\omega) \in A_i$, it holds that

$$\hat{u}_t^{\delta_n,\omega}(\tilde{\omega}) = u_t^i\left(\tilde{\omega}\right), \quad (t,\tilde{\omega}) \in [0,\infty) \times \Omega,$$

we obtain for $\mathbb{P}$-a.a. ω,

$$\begin{aligned}
\mathbb{E}\Big[\int_{\delta_n}^{\infty} & e^{-\beta(t-\delta_n)} k\left(\xi_t^{\hat{u};\mu}, \hat{u}_t\right) \mathrm{d}t \big| \mathcal{F}_{\delta_n}\Big](\omega) \\
&= \int_\Omega \int_0^\infty e^{-\beta t} k\bigg(\xi_t^{\hat{u}^{\delta_n,\omega}, \xi_{\delta_n}^{\hat{u};\mu}(\omega)}(\tilde{\omega}), \hat{u}_t^{\delta_n,\omega}(\tilde{\omega})\bigg) \mathrm{d}t\ \mathrm{d}\mathbb{P}(\tilde{\omega}) \\
&= \sum_{i=1}^N \mathbf{1}_{A_i}\left(\xi_{\delta_n}^{u;\mu}(\omega)\right)\left\{\int_\Omega \int_0^\infty e^{-\beta t} k\bigg(\xi_t^{u^i, \xi_{\delta_n}^{u;\mu}(\omega)}(\tilde{\omega}), u_t^i(\tilde{\omega})\bigg) \mathrm{d}t\ \mathrm{d}\mathbb{P}(\tilde{\omega})\right\} \\
&\le \sum_{i=1}^N \mathbf{1}_{A_i}\left(\xi_{\delta_n}^{u;\mu}(\omega)\right)\left\{V^n\left(\xi_{\delta_n}^{u;\mu}(\omega)\right) + \varepsilon\right\} = V^n\left(\xi_{\delta_n}^{u;\mu}\right)(\omega) + \varepsilon.
\end{aligned}$$

In consequence,

$$\begin{aligned}
V^n(\mu) &\le \hat{J}(\hat{u};\mu) \\
&= \mathbb{E}\left[\int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{\hat{u};\mu}, \hat{u}_t\right) \mathrm{d}t + e^{-\beta\delta_n}\mathbb{E}\left[\int_{\delta_n}^\infty e^{-\beta(t-\delta_n)} k\left(\xi_t^{\hat{u};\mu}, \hat{u}_t\right) \mathrm{d}t \big| \mathcal{F}_{\delta_n}\right]\right] \\
&\le \mathbb{E}\left[\int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t + e^{-\beta\delta_n} V^n\left(\xi_{\delta_n}^{u;\mu}\right)\right] + e^{-\beta\delta_n}\varepsilon.
\end{aligned}$$

Taking the infimum over $u \in \mathcal{A}^n$ and using that $\varepsilon > 0$ was arbitrary, yields the claim. □

Proof of Corollary 5.5. Let $\varepsilon > 0$ and let $\{A_i\}_{i=1}^N \subset \mathcal{P}$ and $\{u^i\}_{i=1}^N \subset \mathcal{A}^n$ be defined as in the proof of Theorem 5.4. In turn, define $\hat{u}^n : \mathcal{P} \to \mathcal{U}$ by

$$\hat{u}^n(\mu) := \sum_{i=1}^N \mathbf{1}_{A_i}(\mu)\, u_0^i, \quad \mu \in \mathcal{P}.$$

Note that, thanks to (44) and (43), for all $\mu \in \mathcal{P}$,

$$V^n(\mu) \ge \mathbb{E}\left[\int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{\hat{u}^n(\mu);\mu}, \hat{u}^n(\mu)\right) \mathrm{d}t + e^{-\beta\delta_n} V^n\left(\xi_{\delta_n}^{\hat{u}^n(\mu);\mu}\right)\right] - \varepsilon. \tag{45}$$

Let $\mu \in \mathcal{P}$ and define $u \equiv u^{*,n} \in \mathcal{A}^n$ by (33) with respect to $\hat{u}^n$; it remains to argue that u is 'ε-optimal'. By definition, (45) holds with $\hat{u}^n(\mu)$ replaced by u. In turn, making once again use of property (45) together with the fact that $\hat{u}^n(\xi_{\delta_n}^{u;\mu}(\omega)) = u_0^{\delta_n,\omega}$, for each $\omega \in \Omega$, and Lemma 5.3, we obtain

$$\begin{aligned}
V^n\left(\xi_{\delta_n}^{u;\mu}\right)(\omega) \ge &\int_\Omega \bigg\{\int_0^{\delta_n} e^{-\beta t} k\bigg(\xi_t^{\hat{u}^n\left(\xi_{\delta_n}^{u;\mu}(\omega)\right);\xi_{\delta_n}^{u;\mu}(\omega)}(\tilde{\omega}), \hat{u}^n\left(\xi_{\delta_n}^{u;\mu}(\omega)\right)\bigg)\mathrm{d}t \\
&\qquad + e^{-\beta\delta_n} V^n\bigg(\xi_{\delta_n}^{\hat{u}^n\left(\xi_{\delta_n}^{u;\mu}(\omega)\right);\xi_{\delta_n}^{u;\mu}(\omega)}(\tilde{\omega})\bigg)\bigg\}\mathrm{d}\mathbb{P}(\tilde{\omega}) - \varepsilon \\
= &\int_\Omega \bigg\{\int_0^{\delta_n} e^{-\beta t} k\bigg(\xi_t^{u^{\delta_n,\omega};\xi_{\delta_n}^{u;\mu}(\omega)}(\tilde{\omega}), u_t^{\delta_n,\omega}(\tilde{\omega})\bigg)\mathrm{d}t \\
&\qquad + e^{-\beta\delta_n} V^n\bigg(\xi_{\delta_n}^{u^{\delta_n,\omega};\xi_{\delta_n}^{u;\mu}(\omega)}(\tilde{\omega})\bigg)\bigg\}\mathrm{d}\mathbb{P}(\tilde{\omega}) - \varepsilon \\
= &\,\mathbb{E}\bigg[\int_{\delta_n}^{2\delta_n} e^{-\beta(t-\delta_n)} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t + e^{-\beta\delta_n} V^n\left(\xi_{2\delta_n}^{u;\mu}\right) \big| \mathcal{F}_{\delta_n}\bigg](\omega) - \varepsilon.
\end{aligned}$$

In consequence,

$$\mathbb{E}\left[\int_0^{\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t + e^{-\beta\delta_n} V^n\left(\xi_{\delta_n}^{u;\mu}\right)\right]$$
$$\geq \mathbb{E}\left[\int_0^{2\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}, u_t\right) \mathrm{d}t + e^{-\beta 2\delta_n} V^n\left(\xi_{2\delta_n}^{u;\mu}\right)\right] - \varepsilon e^{-\beta\delta_n}.$$

Repeated use of the same kind of arguments yields, for $l \in \mathbb{N}$,

$$V^n(\mu) \geq \mathbb{E}\left[\int_0^{l\delta_n} e^{-\beta t} k\left(\xi_t^{u;\mu}, u\right) \mathrm{d}t + e^{-\beta l\delta_n} V^n\left(\xi_{l\delta_n}^{u;\mu}\right)\right] - \varepsilon\big(1 + e^{-\beta\delta_n} + \cdots + e^{-\beta(l-1)\delta_n}\big).$$

Sending l to infinity, making use of the fact that V^n is bounded and that the geometric series converges, we may conclude. $\square$

References

[1] David Assaf and Shmuel Zamir. Optimal sequential search: a Bayesian approach. *Ann. Statist.*, 13(3):1213–1221, 1985.

[2] Elena Bandini, Andrea Cosso, Marco Fuhrman, and Huyên Pham. Backward SDEs for optimal control of partially observed path-dependent stochastic systems: a control randomization approach. *Ann. Appl. Probab.*, 28(3):1634–1678, 2018.

[3] Elena Bandini, Andrea Cosso, Marco Fuhrman, and Huyên Pham. Randomized filtering and Bellman equation in Wasserstein space for partial observation control problem. *Stochastic Process. Appl.*, 129(2):674–711, 2019.

[4] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[5] Erhan Bayraktar and Mihai Sîrbu. Stochastic Perron's method for Hamilton-Jacobi-Bellman equations. *SIAM J. Control Optim.*, 51(6):4274–4294, 2013.

[6] Mathias Beiglböck, Alexander M. G. Cox, Martin Huesmann, and Sigrid Källblad. Measure-valued martingales and optimality of bass-type solutions to the skorokhod embedding problem. *arXiv preprint arXiv:1708.07071*, 2017.

[7] Václav E. Beneš, Ioannis Karatzas, and Raymond W. Rishel. The separation principle for a Bayesian adaptive control problem with no strict-sense optimal law. In *Applied stochastic analysis (London, 1989)*, volume 5 of *Stochastics Monogr.*, pages 121–156. Gordon and Breach, New York, 1991.

[8] Alain Bensoussan. *Stochastic control of partially observable systems.* Cambridge University Press, Cambridge, 1992.

[9] Douglas Blount and Michael A. Kouritzin. On convergence determining and separating classes of functions. *Stochastic Process. Appl.*, 120(10):1898–1907, 2010.

[10] Bruno Bouchard and Nizar Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM J. Control Optim.*, 49(3):948–962, 2011.

[11] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. I*, volume 83 of *Probability Theory and Stochastic Modelling.* Springer, Cham, 2018. Mean field FBSDEs, control, and games.

[12] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. II*, volume 84 of *Probability Theory and Stochastic Modelling.* Springer, Cham, 2018. Mean field games with common noise and master equations.

[13] Norbert Christopeit. Existence of optimal stochastic controls under partial observation. *Z. Wahrsch. Verw. Gebiete*, 51(2):201–213, 1980.

[14] Julien Claisse, Denis Talay, and Xiaolu Tan. A pseudo-Markov property for controlled diffusion processes. *SIAM J. Control Optim.*, 54(2):1017–1029, 2016.

[15] Samuel N Cohen, Christoph Knochenhauer, and Alexander Merkel. Optimal adaptive control with separable drift uncertainty. *arXiv preprint arXiv:2309.07091*, 2023.

[16] Alexander M. G. Cox and Sigrid Källblad. Model-independent bounds for Asian options: a dynamic programming approach. *SIAM J. Control Optim.*, 55(6):3409–3436, 2017.

[17] Alexander M. G. Cox, Sigrid Källblad, Martin Larsson, and Sara Svaluto-Ferro. Controlled measure-valued martingales: a viscosity solution approach. *Ann. Appl. Probab.*, 34(2):1987–2035, 2024.

[18] Huanyu Ding and David A Castañón. Optimal solutions for classes of adaptive search problems. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 5749–5754. IEEE, 2015.

[19] Erik Ekström, Ioannis Karatzas, and Juozas Vaicenavicius. Bayesian sequential least-squares estimation for the drift of a Wiener process. *Stochastic Process. Appl.*, 145:335–352, 2022.
[20] Erik Ekström and Yuqiong Wang. Stopping problems with an unknown state. *J. Appl. Probab.*, 61(2):515–528, 2024.
[21] Nicole El Karoui, Du' Hùù Nguyen, and Monique Jeanblanc-Picqué. Existence of an optimal Markovian filter for the control under partial observations. *SIAM J. Control Optim.*, 26(5):1025–1061, 1988.
[22] Wendell H. Fleming. Measure-valued processes in the control of partially-observable stochastic systems. *Appl. Math. Optim.*, 6(3):271–285, 1980.
[23] Wendell H. Fleming and Étienne Pardoux. Optimal control for partially observed diffusions. *SIAM J. Control Optim.*, 20(2):261–285, 1982.
[24] Wendell H. Fleming and Raymond W. Rishel. *Deterministic and stochastic optimal control.* Applications of Mathematics, No. 1. Springer-Verlag, Berlin-New York, 1975.
[25] V Fox, Jeffrey Hightower, Lin Liao, Dirk Schulz, and Gaetano Borriello. Bayesian filtering for location estimation. *IEEE pervasive computing*, 2(3):24–33, 2003.
[26] Fausto Gozzi and Andrzej Święch. Hamilton-Jacobi-Bellman equations for the optimal control of the Duncan-Mortensen-Zakai equation. *J. Funct. Anal.*, 172(2):466–510, 2000.
[27] U. G. Haussmann. On the existence of optimal controls for partially observed diffusions. *SIAM J. Control Optim.*, 20(3):385–407, 1982.
[28] Omar Hijab. Partially observed control of Markov processes. I. *Stochastics Stochastics Rep.*, 28(2):123–144, 1989.
[29] Omar Hijab. Partially observed control of Markov processes. II. *Stochastics Stochastics Rep.*, 28(3):247–262, 1989.
[30] Omar Hijab. Partially observed control of Markov processes. III. *Ann. Probab.*, 18(3):1099–1125, 1990.
[31] Ioannis Karatzas and Daniel L. Ocone. The resolvent of a degenerate diffusion on the plane, with application to partially observed stochastic control. *Ann. Appl. Probab.*, 2(3):629–668, 1992.
[32] Ioannis Karatzas and Daniel L. Ocone. The finite-horizon version for a partially-observed stochastic control problem of Beneš and Rishel. *Stochastic Anal. Appl.*, 11(5):569–605, 1993.
[33] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.
[34] Hiroshi Kunita. Asymptotic behavior of the nonlinear filtering errors of Markov processes. *J. Multivariate Anal.*, 1:365–393, 1971.
[35] P.-L. Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. I. The case of bounded stochastic evolutions. *Acta Math.*, 161(3-4):243–278, 1988.
[36] P.-L. Lions. Viscosity solutions of fully nonlinear second order equations and optimal stochastic control in infinite dimensions. II. Optimal control of Zakai's equation. In *Stochastic partial differential equations and applications, II (Trento, 1988)*, volume 1390 of *Lecture Notes in Math.*, pages 147–170. Springer, Berlin, 1989.
[37] P.-L. Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of viscosity solutions for general second-order equations. *J. Funct. Anal.*, 86(1):1–18, 1989.
[38] Mattia Martini. Kolmogorov equations on spaces of measures associated to nonlinear filtering processes. *arXiv preprint arXiv:2107.11865*, 2021.
[39] Mattia Martini. Kolmogorov equations on the space of probability measures associated to the nonlinear filtering equation: the viscosity approach. *arXiv preprint arXiv:2202.11072*, 2022.
[40] J. Szpirglas. Sur l'équivalence d'équations différentielles stochastiques à valeurs mesures intervenant dans le filtrage markovien non linéaire. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 14(1):33–59, 1978.
[41] Ramon Van Handel. Stochastic calculus, filtering, and stochastic control. *Course notes., URL http://www. princeton. edu/rvan/acm217/ACM217.pdf*, 2007.
[42] Liang Yu, Qiang Han, Xianguo Tuo, and Wanchun Tian. A survey of probabilistic search based on bayesian framework. In *2019 4th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pages 930–9305. IEEE, 2019.