Covariance control problems over martingales with fixed terminal distribution arising from game theory

Fabien Gensbittel

COVARIANCE CONTROL PROBLEMS OVER MARTINGALES WITH FIXED TERMINAL DISTRIBUTION ARISING FROM GAME THEORY.

FABIEN GENSBITTEL

Abstract. We study several aspects of covariance control problems over martingale processes in $\mathbb{R}^d$ with constraints on the terminal distribution, arising from the theory of repeated games with incomplete information. We show that these control problems are the limits of discrete-time stochastic optimization problems called problems of maximal variation of martingales meaning that sequences of optimizers for the problems of length n, seen as piecewise constant processes on the uniform partition of $[0,1]$, define relatively compact sequences having all their limit points in the set of optimizers of the control problem. Optimal solutions of this limit problem are then characterized using convex duality techniques and the dual problem is shown to be an unconstrained stochastic control problem characterized by a second order nonlinear PDE of HJB type. We deduce from this dual relationship that solutions of the control problem are the images by the spatial gradient of the solution of the HJB equation of the solutions of the dual stochastic control problem using tools from optimal transport theory.

1. Introduction

We study in this work several aspects of constrained covariance control problems of the form

$$W_{ac}(\mu) \triangleq \sup_{X \in \mathcal{M}_{ac}(\preceq_\mu)} \mathbb{E}[\int_0^1 r(\frac{d}{ds}\langle X\rangle_s)ds], \tag{1}$$

where $\mathcal{M}_{ac}(\preceq_\mu)$ is the set of distributions of martingales $(X_t)_{t\in[0,1]}$ with continuous trajectories, having a quadratic variation process $(\langle X\rangle_t)_{t\in[0,1]}$ which is absolutely continuous with respect to the Lebesgue's measure, and such that the law of X_1 is dominated by μ in the sense of convex ordering[1]. Our aim is to characterize the solutions of this problem and to relate them with the limits of the maximizers of discrete-time functionals Ψ_n defined below arising from the study of repeated games with incomplete information. The functionals Ψ_n have been introduced in De Meyer [9] in order to solve the problem of optimal revelation over time for an informed agent in financial exchange games (see also Gensbittel [13] for the multi-dimensional extension). The maximizers of these discrete-time optimization problems are equilibrium price processes in these games. Our main convergence results (Theorems 1.3 and 1.4 below) identify the continuous-time limits of these price processes as solutions of (1). Moreover, our motivation for studying both the continuous-time and the discrete-time problems in the same work is motivated by the fact that the control problem cannot be directly interpreted as a continuous-time game of the same type as the games introduced in [9]. Indeed, convergence involves a Central Limit Theorem, and thus a loss of information on the data of the discrete-time problem. Let us mention however that control problems similar to (1) but depending on the position of the martingale and not on its infinitesimal covariance appear in the study of differential games with incomplete information (see Cardaliaguet-Rainer [7]).

In a first part, the value function of the control problem (1) will be shown to be the limit of the discrete-time optimization problems constructed with Ψ_n and called problems of maximal variation of martingales. These problems generalize the problem of maximal L^1-variation introduced in Mertens-Zamir [18] to more general functions than the L^1-norm. In a second part, we analyze the convex dual problem of (1) which is shown to be an unconstrained stochastic control problem (actually a simple case of the G-expectation introduced by Peng [20]) characterized by a nonlinear second-order HJB equation. We finally prove that the primal solutions of (1) are the images by the gradient of the solution of the HJB equation of the dual solutions, using tools from Optimal Transport theory.

The problem of maximal variation. Given some real-valued function V defined on the set of probabilities over $\mathbb{R}^d$, let us introduce a functional called the V-variation, defined over the set $\mathcal{M}_n(\mu)$ of $\mathbb{R}^d$-valued martingales of length n whose terminal distribution is dominated by μ in the sense of convex ordering[1]. The V-variation of length n of the martingale $(L_k)_{k=1,..,n}$ is defined as

$$\Psi_n[V]((L_k)_{k=1,..,n}) \triangleq \mathbb{E}[\sum_{k=1}^n V([\![L_k - L_{k-1} \mid (L_i, i \leq k-1)]\!])],$$

Key words and phrases. Stochastic Control, Repeated Games, HJB equations, Duality.
This paper is a revised version of the second chapter of the authors's Ph.d. thesis [12].

[1] ν is dominated by μ (denoted $\nu \preceq \mu$) if $\int f d\nu \leq \int f d\mu$ for all closed convex functions f (see Definition 5.6)

where $[\![L_k - L_{k-1} \mid (L_i, i \leq k-1)]\!]$ denotes the conditional law[2] of $L_k - L_{k-1}$ given $(L_i, i \leq k-1)$ with the convention $L_0 = \mathbb{E}[L_1]$. The normalized value function of the above problem is denoted

$$V_n(\mu) \triangleq \frac{1}{\sqrt{n}} \sup_{(L_k)_{k=1,..,n} \in \mathcal{M}_n(\mu)} \Psi_n[V]((L_k)_{k=1,..,n}).$$

The asymptotic behavior of such discrete-time functionals has been recently studied in De Meyer [9] for the case $d = 1$. The main result in [9] is two-fold. At first, a characterization of the limit $V_\infty = \lim_n V_n$ as a maximal covariance function is given but without the corresponding continuous-time control formulation introduced in the present work. Then, it is shown that any sequence of asymptotically optimal martingales for V_n, considered as piecewise constant continuous-time processes, converges in law to a specific continuous-time martingale called Continuous Martingale of Maximal Variation (CMMV) when n goes to ∞. The most surprising aspect of this result is that the law of the limit process CMMV does not depend on V, and neither does V_∞ up to a multiplicative constant. We will show that in the general case, discrete-time maximizers still converge to the set of solutions of (1), which is not necessarily reduced to a point and characterized by a dual HJB equation. The invariance property in higher dimension is expressed through the integral cost r of the limiting control problem which is the upper envelope of V with respect to equivalence classes of laws having the same covariance matrices.

Assumptions on V. We introduce five assumptions denoted A1-A5 on the function V. A1-A4 are the natural generalizations of the assumptions given in [9], while A5 is specific to the multi-dimensional case. Let Δ^2 denote the set of probabilities with finite second-order moments over $\mathbb{R}^d$ and Δ_0^2 the subset of centered probabilities. Let L^2 denote a space of $\mathbb{R}^d$-valued square-integrable random variables defined on some atomless probability space. We assume that the function $V : \Delta_0^2 \to \mathbb{R}$ has the following properties:

(A1) $V \geq 0$ and has no degenerate directions: $\forall x \in \mathbb{R}^d, \exists \mu \in \Delta_0^2$ such that $\mu(\mathbb{R}x) = 1$ and $V(\mu) > 0$.
(A2) V is γ-Lipschitz for the Wasserstein distance[3] of order p for some $p \in [1, 2)$.
(A3) V is positively 1-homogenous: for all centered random variable $X \in L^2$ and $\lambda > 0$,

$$V([\![\lambda X]\!]) = \lambda V([\![X]\!]), \text{ where } [\![X]\!] \text{ denotes the law of } X.$$

(A4) V is concave on Δ_0^2 (seen as a convex subset of the space of Radon measures on $\mathbb{R}^d$).

The last assumption requires the introduction of the auxiliary functions r and R. The function r is an upper envelope that depends only on the covariance matrices of the probabilities in Δ_0^2 (denoted $cov(\mu)$). Precisely, r and R are defined by

$$\forall P \in S_+^d,\ r(P) \triangleq \sup_{\nu \in \Delta_0^2 : cov(\nu) = P} V(\nu)\,; \qquad \forall \mu \in \Delta_0^2,\ R(\mu) \triangleq r(cov(\mu)), \tag{2}$$

where S_+^d denotes the set of non-negative symmetric matrices of size d. Note also that R defines naturally a function on L^2 by $Y \to R([\![Y]\!]) = r(cov(Y))$. Our last assumption is

(A5) R is quasiconvex on L^2 i.e. $\forall \alpha \in \mathbb{R}$, $\{Y \in L^2 \mid R([\![Y]\!]) \leq \alpha\}$ is convex in L^2.

Remark 1.1. *Note that the function R is concave on Δ_0^2 and convex on L^2 (from A1,A3 and A5), hence for different linear structures. If $d = 1$, it is easy to check that A5 is always true and that $r = \sqrt{.}$ up to a multiplicative constant (see Proposition 4.3).*

A simple example fulfilling A1-A5 is given by the L^p-norm $\mu \to \|\mu\|_p \triangleq (\int |x|^p d\mu(x))^{\frac{1}{p}}$ for some $p \in [1, 2)$. A larger class of functions is obtained by considering the upper envelopes of maximal covariance functions (see section 5.1)

$$\mu \to \sup_{\nu \in I} C(\mu, \nu), \tag{3}$$

where $I \subset \Delta_0^2$ is convex, has uniformly bounded moments of order q for some $q > 2$, and contains some ν such that $cov(\nu)$ in non-degenerate. The function C is defined by

$$C(\mu, \nu) \triangleq \sup_{[\![X]\!] = \mu\,,\, [\![Y]\!] = \nu} \mathbb{E}[\langle X, Y \rangle], \tag{4}$$

where $\langle ., . \rangle$ denotes the scalar product in $\mathbb{R}^d$, and the maximum is over all the joint distributions of pairs (X, Y) fulfilling the marginal constraints $[\![X]\!] = \mu$ and $[\![Y]\!] = \nu$.

[2]Recall that $[\![L_k - L_{k-1} \mid (L_i, i \leq k-1)]\!]$ defines a $\sigma(L_i, i \leq k-1)$-measurable random variable with values in the set of probabilities over $\mathbb{R}^d$ (see e.g. Proposition 7.26 in [3]).

[3]We will assume without loss of generality in the proofs that $1 < p < 2$.

Main results. In order to state the first result, we need the following definition.

Definition 1.2. *Given the function r defined above, the subsets F, G and Γ of S_+^d are defined by*

$$F \triangleq \{P \in S_+^d \,:\, r(P) \leq 1\},\;\; G \triangleq \{P \in S_+^d \,:\, \sup_{M \in \mathbb{M}_d \,:\, MM^T \in F} Tr(\sqrt{P}M) \leq 1\} \;\; and \;\; \Gamma \triangleq co(G),$$

where $co(.)$ denotes the convex hull, $\mathbb{M}_d$ the set of $d \times d$ matrices and $\sqrt{P}$ the non-negative square root of P.

G is actually the"polar" set of F induced by the linear structure of L^2 (see Lemma 2.3). Our first main Theorem shows that $\lim_n V_n$ depends only on V through Γ, hence through the auxiliary function r.

Theorem 1.3. *Under assumptions A1-A5, the limit V_∞ of the sequence V_n exists and is given by*

$$\forall \mu \in \Delta^2,\;\; \lim_{n \to \infty} V_n(\mu) = V_\infty(\mu) \triangleq \max_{[\![(Z_t)_{t\in[0,1]}]\!] \in Q_\Gamma \,,\, [\![L]\!] = \mu} \mathbb{E}[\langle L, Z_1 \rangle]$$

where Q_Γ is the compact convex set of laws of martingales $(Z_t)_{t\in[0,1]}$ with continuous trajectories whose quadratic covariation process $\langle Z \rangle$ is such that with probability 1

$$Z_0 = 0 \quad and \quad \forall\, 0 \leq s < t \leq 1, \qquad (t-s)^{-1}(\langle Z \rangle_t - \langle Z \rangle_s) \in \Gamma. \tag{5}$$

The proof of this Theorem has two distinct parts. The first one shows that the function V_∞ is an upper bound for $limsup_n V_n$, and relies on Limit Theorems for martingales (see Proposition 3.4). The second part shows that V_∞ is a lower bound for $liminf_n V_n$. This lower bound property relies on the reformulation of the problem V_∞ as the covariance control problem (1) (Lemma 3.6), which allows us to prove in Proposition 3.5 that for an ε-optimal $X \in \mathcal{M}_{ac}(\preceq_\mu)$, there exists a sequence of discretizations $X^n = (X_k^n)_{k=1,..,n}$ of X that are asymptotically ε-optimal for V_n (i.e. such that $liminf_n \Psi_n[V](X^n) \geq W_{ac}(X) - \varepsilon$). We emphasize that our approximation procedure is not only the usual time-discretization, since we have to introduce a second level of discretization based on the Central Limit Theorem for the Wasserstein distance.

The second part of this work is devoted to characterize the maximizers of (1) and to relate them to the limits of optimizers of $\Psi_n[V]$. Precisely, given a discrete-time process $(L_1, .., L_n)$, the continuous-time version of this process is defined by

$$X_t^n \triangleq L_{\lfloor nt \rfloor} \quad \text{for } t \in [0,1],$$

where $\lfloor a \rfloor$ denotes the greatest integer less or equal to a. We aim to characterize the limits in law of the continuous-time versions of asymptotically optimal sequences in $\mathcal{M}_n(\mu)$ for the problem $V_n(\mu)$. At first, we introduce the following reformulation of V_∞

$$V_\infty(\mu) = W(\mu) \triangleq \max_{X \in \mathcal{M}(\preceq_\mu)} H(X),$$

where $\mathcal{M}(\preceq_\mu)$ is the set of distributions of martingales $(X_t)_{t\in[0,1]}$ with càdlàg trajectories whose final distribution is dominated by μ. The functional H is defined in section 4.1 and extends the integral functional given in (1) to the set $\mathcal{M}(\preceq_\mu)$. This second formulation is introduced in order to obtain compactness, and to show that the set of maximizers of W contains the set of accumulation points of the maximizers of the discrete-time problems.

Theorem 1.4. *Let (L^n) be an asymptotically maximizing sequence of $V_n(\mu)$ in $\mathcal{M}_n(\mu)$. Then the continuous-time versions of these martingales define a weakly relatively compact sequence of laws for the Meyer-Zheng topology (see* [19]*) and any limit point belongs to*

$$\mathcal{P}_\infty(\mu) \triangleq \operatorname*{argmax}_{X \in \mathcal{M}(\preceq_\mu)} H(X).$$

We deduce directly from this result the former results obtained in [9] for the particular case $d = 1$ (see Proposition 4.3). In order to study the general case, we introduce the convex dual problem of V_∞ defined on the set of proper closed convex functions $Conv(\mathbb{R}^d)$ by

$$V_\infty^*(\phi) \triangleq \sup_{[\![(Z_t)_{t\in[0,1]}]\!] \in Q_\Gamma} \mathbb{E}[\phi^*(Z_1)],$$

where ϕ^* denotes the Fenchel transform of ϕ. A dual equality is proved in Proposition 4.4 using results appearing in the theory of Optimal Transport. This dual problem is then shown to be a PDE problem of HJB type appearing in Stochastic Control theory (Proposition 4.6). This dual formulation is used to derive a characterization of the elements of $\mathcal{P}_\infty$. Let us mention here the following result which is a Corollary of the main Verification Theorem 4.13.

Theorem 1.5. *Let $u(t,x)$ be the unique viscosity solution of the following HJB equation*

$$\begin{cases} -\frac{\partial}{\partial t}u - \frac{1}{2}\sup_{P\in\Gamma} Tr(P\nabla^2 u) &= 0 \quad in \quad [0,1) \times \mathbb{R}^d \\ u(1,x) &= f(x) \quad in \quad \mathbb{R}^d \end{cases} \tag{6}$$

where f is a C^1 Lipschitz-convex function on $\mathbb{R}^d$. Assume that u is a classical $C^{1,2}$ solution. Let Z be a martingale whose law $\mathbb{P}$ is in Q_Γ and such that

$$\frac{d}{dt}\langle Z\rangle_t \in \underset{P\in\Gamma}{argmax}\ Tr(P\nabla^2 u(t,Z_t))\ dt\otimes d\mathbb{P} \text{ almost surely.} \tag{7}$$

Then, if $\mu \triangleq [\![\nabla f(Z_1)]\!]$, the set $\mathcal{P}_\infty(\mu)$ is exactly the set of laws of the martingales

$$(X_t)_{t\in[0,1]} = (\nabla u(t,\widetilde{Z}_t))_{t\in[0,1]},$$

where the law of $\widetilde{Z}$ runs through all the laws in Q_Γ verifying (7) and $[\![\nabla f(\widetilde{Z}_1)]\!] = \mu$.

The paper is organized as follows. Section 2 presents the main properties of the maximal variation problem. Section 3 is devoted to the proof of Theorem 1.3 and section 4 to the characterization of the solutions of the control problem. The last section is an appendix collecting some classical results reproduced for convenience of the reader and/or because precise references are difficult to find, it also contains some technical proofs which can be omitted at a first reading.

Glossary of notations. In order to lighten several proofs and statements that will appear throughout this work, some non-standard notations will be used, essentially in Probability Theory. A glossary of the most frequent ones is provided below for the convenience of the reader. Notations which are related to Stochastic Processes Theory will be given at the beginning of section 3

- $\Delta(E)$ denotes the set of probabilities defined on the borel σ-field of a topological space E, endowed with the usual weak* topology. $\Delta^2(\mathbb{R}^d)$ (Δ^2 for short) is the subset of probabilities with finite second order moments. Δ_0^2 denotes the subset of laws with zero mean.
- Let E, E' be two separable metric spaces and A,A' two subsets of $\Delta(E)$ and $\Delta(E')$. $\mathcal{P}(A,A')$ denotes the set of probabilities over $E\times E'$ whose marginal distributions over E and E' belongs respectively to A and A'.If $A=\{\mu\}$, we simply write $\mathcal{P}(\mu,A')$.
- Let E be a Polish space and X be an E-valued random variable defined on $(\Omega,\mathcal{F},\mathbb{P})$.
  - $[\![X]\!]$ denotes the law of X.
  - Given a sub-σ-field $\mathcal{G}$, $[\![X \mid \mathcal{G}]\!]$ denotes a version of the conditional law of X given $\mathcal{G}$, hence a $\mathcal{G}$ measurable random variable with values in $\Delta(E)$.
- d_{W_q} denotes the Wasserstein distance of order q (see section 5).

2. Properties of the discrete-time problem.

In this section, we study the auxiliary functions R and r. Next, using their properties, we provide an upper bound for the V-variation which will be a key argument for the main convergence result in section 3.

2.1. Properties of the auxiliary functions R and r.

The next Lemma is based on [8].

Lemma 2.1. *For all $P,Q\in S_+^d$ and $\mu\in\Delta^2$ such that $cov(\mu)=P$, we have*

$$\sup_{\nu\in\Delta_0^2: cov(\nu)\le Q} C(\mu,\nu) = \sup_{\nu\in\Delta_0^2: cov(\nu)=Q} C(\mu,\nu) = Tr\left((\sqrt{P}Q\sqrt{P})^{\frac{1}{2}}\right) = \sup_{D\in\mathbb{M}_d\,:\,DD^T=Q} Tr(N^T D)$$

where the last equality holds for any N such that $NN^T=P$ (in particular for $N=\sqrt{P}$).

Proof. If $X\sim\mu$ and $Y\sim\nu$ are given random variables such that $Q-cov(\nu)\ge 0$, we can construct a variable Z independent of (X,Y) such that $\mathbb{E}[Z]=0$ and $cov(Z)=Q-cov(\nu)$. It follows that $\mathbb{E}[\langle X,Y+Z\rangle]=\mathbb{E}[\langle X,Y\rangle]$ and $cov(Y+Z)=Q$, which proves the first equality. The second equality follows from Theorem 2.1 in [8], where a characterization is given which implies moreover that the supremum is reached. For the third equality, given a variable X of law μ, define $U=N^{-1}(X-\mathbb{E}[X])$, where $(.)^{-1}$ denotes the Moore-Penrose pseudo-inverse. Since $cov(U)\le I_d$, we can construct a random variable V with values in $Ker(N)$, independent of X and such that $cov(U+V)=I_d$. It follows that $X=N(U+V)$ and with $Y=D(U+V)$, we have $\mathbb{E}[\langle X,Y\rangle]=Tr(DN^T)$. This implies the result since the supremum is reached with $D=MN$ for $M=\sqrt{P}^{-1}(\sqrt{P}Q\sqrt{P})^{1/2}\sqrt{P}^{-1}$. □

In the following, L^2 denotes the space $L^2([0,1],dx;\mathbb{R}^d)$ and L_0^2 its subspace of centered random variables.

Definition 2.2. *The polar set C° of $C\subset L_0^2$ is defined by $C^\circ \triangleq \{X\in L_0^2 : \sup_{Y\in C}\mathbb{E}[\langle X,Y\rangle]\le 1\}$.*

The following Lemma lists the main properties of r.

Lemma 2.3. *The function r is non-negative, concave, non-decreasing, continuous on S_+^d and*

$$r(P) = \max_{\mu\in\Delta_0^2\,:\,cov(\mu)\le P} V(\mu), \tag{8}$$

$$\forall\lambda>0,\qquad r(\lambda P)=\sqrt{\lambda}r(P), \tag{9}$$

$$\forall M\in\mathbb{M}_d,\quad r(MM^T)=\max_{N\in\mathbb{M}_d\,:\,NN^T\in G} Tr(MN). \tag{10}$$

Moreover, G is a compact neighborhood of 0 *in* S_+^d *and* $M \to r(MM^T)$ *is Lipschitz.*

Proof. Note at first that the d_{W_p}-closure of $\{\nu \in \Delta_0^2 \,:\, cov(\nu) = P\}$ is $\{\nu \in \Delta_0^2 \,:\, cov(\nu) \leq P\}$ (see Lemma 5.13), so that (8) follows from A2. Since cov is linear and V is 1-homogenous, non-negative and concave, the non-negativeness, concavity and (9) are obvious. Note that the subset $\{\mu \in \Delta_0^2 : cov(\mu) \leq P\}$ is d_{W_p}-compact since moments of order $2 > p$ are uniformly bounded. The continuity of r follows therefore from Berge's Maximum Theorem (see [2] p116) since the set-valued mapping $P \to \{\mu \in \Delta_0^2 \,:\, cov(\mu) \leq P\}$ is both upper and lower semi-continuous when Δ_0^2 is endowed with the metric d_{W_p}. Using then that r is continuous, (9) and A1, F is a compact neighborhood of 0 in S_+^d. R being sublinear in L_0^2, it is the support function of the polar set of

$$\widehat{F} \triangleq \{X \in L_0^2 \,|\, cov(X) \in F\} = \{X \in L_0^2 \,|\, R([\![X]\!]) \leq 1\}.$$

Let us prove that $\widehat{F}^\circ = \widehat{G} \triangleq \{X \in L_0^2 \,:\, cov(X) \in G\}$. Since r is nondecreasing,

$$\widehat{F} = \cup_{Q \in F} \{Y \in L_0^2 \,:\, cov(Y) \leq Q\}.$$

Next, we claim that if $X \in L_0^2$ is μ-distributed, then

$$\sup_{Y \in L_0^2 \,:\, cov(Y) \leq Q} \mathbb{E}[\langle X, Y \rangle] = \sup_{\nu \in \Delta_0^2 : cov(\nu) \leq Q} C(\mu, \nu). \tag{11}$$

The left-hand side is obviously lower or equal than the right-hand side. To prove the converse, given $X \in L^2$ and $\pi \in \mathcal{P}(\mu, \nu)$ such that $cov(\nu) \leq Q$, we can construct a pair (X, Y) of law π on an enlarged probability space, and replace Y by $\phi(X) \triangleq \mathbb{E}[Y \mid X] \in L_0^2$. We check easily that $\mathbb{E}[\langle X, Y \rangle] = \mathbb{E}[\langle X, \phi(X) \rangle]$ and that $cov(\phi(X)) \leq cov(Y) \leq Q$. It follows from Lemma 2.1 that with $cov(X) = P$

$$\sup_{Y \in \widehat{F}} \mathbb{E}[\langle X, Y \rangle] = \sup_{Q \in F} \sup_{Y \in L_0^2 \,:\, cov(Y) \leq Q} \mathbb{E}[\langle X, Y \rangle] = \sup_{Q \in F} \sup_{M \,:\, MM^T = Q} Tr(M\sqrt{P}),$$

which proves that $\widehat{F}^\circ = \widehat{G}$. R is therefore the support function of $\widehat{G}$ and (10) follows from the definition of G. Using lemma 2.1, we have

$$G = \{P \in S_+^d \mid \sup_{Q \in F} Tr\left((\sqrt{P} Q \sqrt{P})^{\frac{1}{2}}\right) \leq 1\}.$$

This equality implies that G is itself a compact neighborhood of 0. Indeed, F being compact the above supremum defines a continuous function. This function is positive for $P \neq 0$ since $Q = \lambda P \in F$ for sufficiently small $\lambda > 0$ and this proves that G is a neighborhood of 0. Compactness of G follows then from (9) and directly implies the announced Lipschitz property for $M \to r(MM^T)$. □

Our main result in this section is the following upper bound for V, which is a modification of R that takes into account the Lipschitz assumption A2.

Proposition 2.4.

$$\forall \mu \in \Delta_0^2,\ V(\mu) \leq R'(\mu) \triangleq \sup_{\nu \in T} C(\mu, \nu),$$

where $T \triangleq \{\nu \in \Delta_0^2 \,:\, cov(\nu) \in \Gamma\,,\ \|\nu\|_{p'} \leq 2\gamma\}$ *and* p' *is the conjugate exponent of* p.

Proof. For $q \geq 1$ and $m \geq 0$, define $B_m^q \triangleq \{\mu \in \Delta_0^2 \,:\, \|\mu\|_q \leq m\}$ and $\widetilde{B}_m^q \triangleq \{X \in L_0^2 \,:\, [\![X]\!] \in B_m^q\}$. Recall also the definitions of $\widehat{F}, \widehat{G}$ given in the proof of the previous Lemma. We claim that

$$co(\widehat{F} \cup \widetilde{B}_{1/\gamma}^p) \text{ is a closed convex set and is included in } \{X \in L_0^2 \mid V([\![X]\!]) \leq 1\}. \tag{12}$$

Using A5 and Lemma 2.3, $\widehat{F}$ is weakly compact and convex. Since $\widetilde{B}_{1/\gamma}^p$ is weakly closed and convex, the convex envelope is weakly closed, hence closed in L_0^2 for the norm topology. Let $X \in \widehat{F}$, $Y \in \widetilde{B}_{1/\gamma}^p$ and $\lambda \in [0, 1]$. Using the Lipschitz property of V, we deduce

$$V([\![\lambda X + (1-\lambda)Y]\!]) \leq V([\![\lambda X]\!]) + \gamma \|(1-\lambda)Y\|_{L^p} \leq \lambda r(cov(X)) + (1-\lambda) \leq 1,$$

which proves (12). Define $\widehat{\Gamma} \triangleq \{X \in L_0^2 \,:\, cov(X) \in \Gamma\}$. From the definition of T,

$$\forall X \in L_0^2,\ R'([\![X]\!]) = sup\{\mathbb{E}[\langle X, Y \rangle] \mid Y \in \widehat{\Gamma} \cap \widetilde{B}_{2\gamma}^{p'}\}. \tag{13}$$

The proof of (13) proceeds as for (11) since T is stable by conditional expectations. Note that by definition $\widehat{F}^\circ = \widehat{G} \subset \widehat{\Gamma}$ and that $(\widetilde{B}_{1/\gamma}^p)^\circ \subset \widetilde{B}_{2\gamma}^{p'}$, which follows from the classical $L^p/L^{p'}$ duality (the coefficient 2 appears since the case of equality in Hölder's inequality is not necessarily attained for centered random variables). We deduce that $(\widehat{F} \cup \widetilde{B}_{1/\gamma}^p)^\circ \subset (\widehat{\Gamma} \cap \widetilde{B}_{2\gamma}^{p'})$ and using properties of support functions, the inclusion

$$(\widehat{\Gamma} \cap \widetilde{B}_{2\gamma}^{p'})^\circ = \{X \in L_0^2 \mid R'(X) \leq 1\} \subset co(\widehat{F} \cup \widetilde{B}_{1/\gamma}^p) = (\widehat{F} \cup \widetilde{B}_{1/\gamma}^p)^{\circ\circ}.$$

Finally, we deduce from the preceding inclusions that

$$\{X \in L_0^2 \mid R'([\![X]\!]) \leq 1\} \subset \{X \in L_0^2 \mid V([\![X]\!]) \leq 1\} \tag{14}$$

which concludes the proof since these functions are positively homogenous (A3). □

2.2. **Properties of the V-variation.** We provide here an upper bound based on the inequality proved in Proposition 2.4. At first, as it will be convenient to consider martingales defined with respect to a larger filtration than the filtration generated by the process itself, let us now introduce an equivalent formulation of the V-variation.

Definition 2.5. *$\mathfrak{M}_n(\mu)$ is the collection of martingales $(L_k, \mathcal{F}_k)_{k=1,..,n}$ defined of some filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_k)_{k=1,..,n}, \mathbb{P})$, of length n and whose final distribution is dominated by μ ($[\![L_n]\!] \preceq \mu$). By convention, we set $\mathcal{F}_0 \triangleq \{\Omega, \emptyset\}$.*

With a slight abuse of notations, we extend the definition of the V-variation to martingales in $\mathfrak{M}_n(\mu)$ by

$$\Psi_n[V]((L_k, \mathcal{F}_k)_{k=1,..,n}) = \mathbb{E}[\sum_{k=1}^{n} V([\![L_k - L_{k-1} \mid \mathcal{F}_{k-1}]\!])]. \tag{15}$$

Lemma 2.6.

$$V_n(\mu) = \frac{1}{\sqrt{n}} \sup_{((L_k,\mathcal{F}_k)_{k=1,..,n}) \in \mathfrak{M}_n(\mu)} \Psi_n[V]((L_k, \mathcal{F}_k)_{k=1,..,n}). \tag{16}$$

Proof. Given a distribution $[\![(L_1, .., L_n)]\!] \in \mathcal{M}_n(\mu)$, then the two notions of V-variation agree if we define $(\mathcal{F}_k^L)_{k=1,..,n}$ as the natural filtration of $(L_1, .., L_n)$, i.e.

$$\Psi_n[V]((L_k)_{k=1,..,n}) = \Psi_n[V]((L_k, \mathcal{F}_k^L)_{k=1,..,n}), \text{ with } \mathcal{F}_k^L = \sigma(L_1, .., L_k) \text{ for } k = 1, .., n \text{ and } \mathcal{F}_0 = \{(\mathbb{R}^d)^n, \emptyset\}.$$

This proves that V_n is not greater than the right-hand side of (16). To prove the reverse inequality, let $((L_k, \mathcal{F}_k)_{k=1,..,n}) \in \mathfrak{M}_n(\mu)$. Since V is concave and d_{W_p}-Lipschitz, it follows from Jensen's inequality (Lemma 5.10 in the appendix) that for all $k = 1, .., n$

$$V([\![L_k - L_{k-1} \mid \mathcal{F}_{k-1}]\!]) \leq V([\![L_k - L_{k-1} \mid \mathcal{F}_{k-1}^L]\!]).$$

The proof follows then by summation over k. □

Notation 2.7. *In order to shorten notations, the function V is extended to Δ^2 by the relation*

$$\forall X \in L^2, \quad V([\![X]\!]) \triangleq V([\![X - \mathbb{E}[X]]\!]). \tag{17}$$

The same convention will be used in the next sections with the functions R and R'. Using the above convention, it follows from the martingale property that

$$\Psi_n[V]((L_k, \mathcal{F}_k)_{k=1,..,n}) = \mathbb{E}[\sum_{k=1}^{n} V([\![L_k \mid \mathcal{F}_{k-1}]\!])].$$

This relation also holds for the R'-variation, which will be denoted $\Psi_n[R']$.

Using the preceding results, we obtain the following upper bound for V_n.

Lemma 2.8.

$$\sqrt{n} V_n(\mu) \leq \sup_{(L_k,\mathcal{F}_k)_{k=1,..,n} \in \mathfrak{M}_n(\mu)} \Psi_n[R']((L_k, \mathcal{F}_k)_{k=1,..,n}).$$

Proof. Since we proved in Proposition 2.4 that $V \leq R'$, we have

$$\Psi_n[V]((L_k, \mathcal{F}_k)_{k=1,..,n}) \leq \Psi_n[R']((L_k, \mathcal{F}_k)_{k=1,..,n}),$$

for any martingale. The conclusion follows by taking the supremum over $\mathfrak{M}_n(\mu)$. □

Let us now reformulate this upper bound in a more tractable way.

Definition 2.9. *Define T^n as the set of distributions $\nu \in \Delta((\mathbb{R}^d)^n)$ of sequences $(S_1, ..S_n)$ such that*

$$\forall k = 1, .., n, \qquad [\![S_k \mid S_1, ..., S_{k-1}]\!] \in T, \ \nu - \ a.s.$$

Lemma 2.10.

$$\sup_{(L_k,\mathcal{F}_k)_{k=1,..,n} \in \mathfrak{M}_n(\mu)} \Psi_n[R']((L_k, \mathcal{F}_k)_{k=1,..,n}) = \max_{[\![(S_k)_{k=1,..,n}]\!] \in T^n, [\![L]\!] \preceq \mu} \mathbb{E}[\langle L, \sum_{k=1}^{n} S_k \rangle].$$

Proof. At first, note that T^n is convex and weakly compact since $\nu \in T^n$ is equivalent to

$$\mathbb{E}_\nu[f(S_1, .., S_{k-1}) S_k] = 0, \quad \mathbb{E}_\nu[f(S_1, .., S_{k-1}) |S_k|^{p'}] \leq (2\gamma)^{p'} \mathbb{E}_\nu[f(S_1, .., S_{k-1})],$$
$$\forall P \in S_+^d, \ \mathbb{E}_\nu[f(S_1, .., S_{k-1}) Tr(P S_k S_k^T)] \leq \mathbb{E}_\nu[f(S_1, .., S_{k-1})] sup\{Tr(PQ) \ : \ Q \in \Gamma\},$$

for all $k = 1, .., n$, where f runs through all nonnegative continuous functions bounded by 1. Indeed, using monotone or dominated convergence, these equalities extend to indicator functions, and the equivalence follows

easily. Since all these constraints are affine and continuous, it defines a closed convex set, and relative compactness follows from the uniform bound on the moments of order 2. Existence of a maximum follows therefore from the Lemmas 5.1 and 5.3. For any law of martingale $[\![(L_k)_{k=1,..,n}]\!] \in \mathcal{M}_n(\mu)$, denoting $(\mathcal{F}_k^L)_{k=1,..,n}$ the natural filtration of $(L_k)_{k=1,..,n}$, we will prove

$$\Psi_n[R']((L_k,\mathcal{F}_k^L)_{k=1,..,n}) \leq \max_{S\in T^n, [\![L]\!]\preceq\mu} \mathbb{E}[\langle L, \sum_{k=1}^n S_k\rangle]. \tag{18}$$

Recall that for $\kappa \in \Delta^2$, with R' extended on Δ^2 by the relation (17)

$$R'(\kappa) = \sup_{\nu\in T} C(\kappa,\nu) = \max_{\pi\in\mathcal{P}(\kappa,T)} \int \langle x,y\rangle d\pi(x,y).$$

The set-valued map $\kappa \to \mathcal{P}(\kappa,T)$ is compact valued and upper semi-continuous, and the map $\pi \to \int\langle x,y\rangle d\pi(x,y)$ is continuous on $\mathcal{P}(B_r,T)$ for any $r \geq 0$ where $B_r \triangleq \{\kappa \in \Delta^2 : \|\kappa\|_2 \leq r\}$. Therefore, using a Measurable Selection Theorem (see Proposition 7.33 in [3]), the set-valued map $\kappa \to argmax\{\int\langle x,y\rangle d\pi(x,y) \mid \pi \in \mathcal{P}(\kappa,T)\}$ admits a measurable selection $f(\kappa)$ on B_r for any $r>0$ and thus on Δ^2. Since the martingale has finite second order moments, the conditional second order moments are almost surely finite and there exists a family of versions of the conditional laws $[\![L_k \mid L_1,..,L_{k-1}]\!]$ with values in Δ^2. Up to enlarging the probability space, we assume the existence of a sequence $(U_i)_{i=1,..,n}$ of independent uniform random variables independent of $(L_1,..,L_n)$. Then we can construct[4] a sequence of random variables $(S_1,..,S_n)$ as a measurable function of $(L_k,U_k)_{k=1,..,n}$ such that the conditional laws are optimal, i.e.

$$\forall k=1,..,n, \qquad [\![(L_k,S_k) \mid L_1,..,L_{k-1}]\!] = f([\![L_k \mid L_1,..,L_{k-1}]\!]) \ a.s.$$

By construction, and using the martingale property

$$\mathbb{E}[\langle L_n,S_k\rangle \mid L_1,..,L_{k-1}] = \mathbb{E}[\langle L_k,S_k\rangle \mid L_1,..,L_{k-1}] = R'([\![L_k \mid L_1,..,L_{k-1}]\!]).$$

We deduce by summation that $\Psi_n[R']((L_k,\mathcal{F}_k^L)_{k=1,..,n}) = \mathbb{E}[\langle L_n, \sum_{k=1}^n S_k\rangle]$ and inequality 18 follows. The converse inequality is straightforward. Given a pair $(L,(S_k)_{k=1,..,n})$, define a martingale by projecting (using conditional expectations) L on the natural filtration of $(S_k)_{k=1,..,n}$ and the proof follows from the definition of R'. □

3. Convergence to the control problem.

This section is devoted to the proof of Theorem 1.3, which is divided in the two Propositions 3.4 and 3.5.

3.1. **Convergence of the upper bound.** We study here the asymptotic behavior of the upper bound introduced in the preceding section, starting from the formulation obtained in Lemma 2.10. The main result is Proposition 3.4 below and is based on classical Limit Theorems for martingales. For this reason, let us recall some standard notations from the theory of stochastic processes.

Notation 3.1.

- $\mathbb{D}([0,1],\mathbb{R}^d)$ *: set of càdlàg functions endowed with the Skorokhod topology.*
- $C([0,1],\mathbb{R}^d)$ *: set of continuous functions identified with a subset of* $\mathbb{D}([0,1],\mathbb{R}^d)$.
- $\mathcal{M}$ *(resp.* $\mathcal{M}^c$*) denotes the subset of* $\Delta(\mathbb{D}([0,1],\mathbb{R}^d))$ *(resp.* $\Delta(C([0,1],\mathbb{R}^d))$*) of martingale distributions.*
- *For a martingale* $(Z_t)_{t\in[0,1]}$ *we denote* $(\mathcal{F}_t^Z)_{t\in[0,1]}$ *the right-continuous filtration it generates defined by* $\mathcal{F}_t^Z \triangleq \cap_{s>t}\sigma(Z_u, u\leq s)$ *and by* $\langle Z\rangle$ *its predictable quadratic covariation process.*

Notation 3.2.

- Q_Γ *is the subset of probabilities* $\mathbb{P}$ *in* $\mathcal{M}^c$ *such that, with* $\mathbb{P}$*-probability 1,*

$$Z_0 = 0 \quad and \quad \forall\, 0\leq s<t\leq 1, \qquad (t-s)^{-1}(\langle Z\rangle_t - \langle Z\rangle_s) \in \Gamma, \tag{19}$$

  where Z *denotes the canonical coordinate process on* $C([0,1],\mathbb{R}^d)$.

- $Q_\Gamma(t)$ *denotes the set of laws of variables* Z_t *when the law of the process* Z *runs through* Q_Γ.
- $\pi_t(Q_\Gamma)$ *denotes the set of laws of processes* $(Z_s)_{s\leq t}$ *when the law of the process* Z *runs through* Q_Γ.

In the sequel, $\mathbb{M}_d$ and S_+^d are endowed with the Euclidean norm $|M| = Tr(MM^T)^{1/2}$.

Lemma 3.3. Q_Γ *is closed, convex and tight (hence compact) and is a face of the convex set* $\mathcal{M}^c$.

Proof. Fix $\mathbb{P}\in Q_\Gamma$, then $\sum_{i=1}^d(\langle Z^i\rangle_t - \langle Z^i\rangle_s) \leq \sqrt{d}C_\Gamma(t-s)$, with $C_\Gamma = sup\{|M| \mid M\in\Gamma\}$. Hence using the Propositions VI.3.35 and VI.4.13 in [15], Q_Γ is tight and using Proposition VI.6.29 in [15], for any sequence $\mathbb{P}_n \in Q_\Gamma$ converging to some limit $\mathbb{P}$, we have that the sequence of distributions of $(Z^n,\langle Z^n\rangle)$ under $\mathbb{P}_n$ converges to the law of $(Z,\langle Z\rangle)$ under $\mathbb{P}$ in $\Delta(C([0,1],\mathbb{R}^d\times S_+^d))$. As a consequence, the sequence of laws of $\langle Z^n\rangle$ converges to the law of $\langle Z\rangle$ so that $\mathbb{P}$ fulfills property (19) and thus belongs to Q_Γ (since the set of continuous functions

[4] see Theorem 5.8 and the following discussion in the appendix.

verifying (19) is closed). To prove convexity, if $\mathbb{P} = \lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2$ with $\mathbb{P}_1, \mathbb{P}_2 \in Q_\Gamma$ and $\lambda \in (0,1)$, then for $i = 1, 2$, it follows from the characterization of the quadratic covariation that

$$\forall \varepsilon > 0, \qquad \mathbb{P}_i(d_\Gamma(\frac{1}{t-s}T^n_{s,t}(Z)) \geq \varepsilon) \underset{n\to\infty}{\longrightarrow} 0, \tag{20}$$

where $T^n_{s,t}(Z) = \sum_{k=0}^{\infty}(Z_{(s+\frac{k+1}{n})\wedge t} - Z_{(s+\frac{k}{n})\wedge t})(Z_{(s+\frac{k+1}{n})\wedge t} - Z_{(s+\frac{k}{n})\wedge t})^T$ and $d_\Gamma(x)$ is the usual distance between x and and the compact set Γ. Therefore the same property holds for $\mathbb{P}$ and this implies (19) (the property holds with probability 1 for s, t rational, and therefore for all s, t by continuity), which in turn implies $\mathbb{P} \in Q_\Gamma$. Finally, if $\mathbb{P} = \lambda\mathbb{P}_1 + (1-\lambda)\mathbb{P}_2$ with $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}^c$, $\lambda \in (0,1)$ and $\mathbb{P} \in Q_\Gamma$ then property (20) holds for $\mathbb{P}$. This property holds then also for $\mathbb{P}_1$ and $\mathbb{P}_2$, and this implies $\mathbb{P}_1, \mathbb{P}_2 \in Q_\Gamma$. □

The following result is the upper bound part of Theorem 1.3.

Proposition 3.4.

$$\underset{n\to\infty}{limsup}\, V_n(\mu) \leq \underset{[\![(Z_t)_{t\in[0,1]}]\!]\in Q_\Gamma, [\![L]\!]\preceq\mu}{sup} \mathbb{E}[\langle L, Z_1\rangle]$$

Proof. Using Lemma 2.10, we have to prove that

$$\underset{n\to\infty}{limsup} \underset{[\![(S_k)_{k=1,..,n}]\!]\in T^n, [\![L]\!]\preceq\mu}{sup} \mathbb{E}[\langle L, \frac{1}{\sqrt{n}}\sum_{k=1}^{n} S_k\rangle] \leq \underset{[\![(Z_t)_{t\in[0,1]}]\!]\in Q_\Gamma, [\![L]\!]\preceq\mu}{sup} \mathbb{E}[\langle L, Z_1\rangle].$$

Let $(L^n, (S^n_k)_{k=1,..,n})$ be a maximizing sequence. Let us define $\mathbb{P}_n$ as the set of distributions of the continuous-time processes $Z^n_t = n^{-\frac{1}{2}}\sum_{k=1}^{\lfloor nt\rfloor} S^n_k$. The sequence $\mathbb{P}_n$ is tight since Z^n_t are martingales with respect to the right-continuous filtration $\mathcal{F}^n_t = \sigma(S^n_k,\ k \leq \lfloor nt\rfloor)$ and their predictable quadratic covariation is C-tight. To prove the last point, note that $\langle Z^n\rangle_t$ is piecewise constant on the intervals $[\frac{k}{n}, \frac{k+1}{n})$ and that

$$n(\langle Z^n\rangle_{\frac{k+1}{n}} - \langle Z^n\rangle_{\frac{k}{n}}) = \mathbb{E}[S^n_{k+1}(S^n_{k+1})^T \mid S^n_1, ..S^n_k] \in \Gamma. \tag{21}$$

Since Γ is bounded by the constant C_Γ, the trace of this matrix-valued process is strongly majorized by the process $t \to \frac{\sqrt{d}C_\Gamma\lfloor nt\rfloor}{n}$ so that the associated sequence of laws is C-tight (see Proposition VI.3.35 in [15]). To prove that the sequence $\mathbb{P}_n$ is itself C-tight, it's sufficient according to Lemma VI.3.26 in [15] to prove that

$$\forall \varepsilon > 0,\ \mathbb{P}_n(\underset{t\in[0,1]}{sup} \mid \Delta Z^n_t \mid > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0,$$

where $\Delta Z^n_t = Z^n_t - Z^n_{t-}$ is the jump of Z^n at time t. We have:

$$\mathbb{P}_n(\underset{t\in[0,1]}{sup} \mid \Delta Z^n_t \mid > \varepsilon) \leq \sum_{k=0}^{n-1}\mathbb{P}_n(\mid S^n_{k+1} - S^n_k \mid > \varepsilon\sqrt{n}) \leq \sum_{k=0}^{n-1}\frac{\mathbb{E}^{\mathbb{P}_n}[\mid S^n_{k+1} - S^n_k \mid^{p'}]}{(\varepsilon\sqrt{n})^{p'}} \leq n\frac{(2\gamma)^{p'}}{(\varepsilon\sqrt{n})^{p'}} \underset{n\to\infty}{\longrightarrow} 0.$$

Suppose now that some subsequence still denoted $\mathbb{P}_n$ converges to $\mathbb{P}$. Then the sequence of laws $\mathbb{Q}_n \in \mathbb{D}([0,1], \mathbb{R}^d \times S^d_+)$ of $(Z^n, \langle Z^n\rangle)$ is also C-tight (corollary VI.3.33 in [15]) and converges to some law $\mathbb{Q}$ (up to the extraction of some subsequence) of a process (Z, A) such that Z has law $\mathbb{P}$. Now the sequences of processes Z^n and $Z^n(Z^n)^T - \langle Z^n\rangle$ are martingales with respect to $\mathcal{F}^n$ and uniformly integrable since respectively bounded in L^2 and $L^{p'/2}$. Applying Proposition IX.1.12 in [15] to each coordinate of these processes, we conclude that Z and $ZZ^T - A$ are martingales relative to the filtration $\mathcal{F}$ generated by (Z, A). The process A is $\mathcal{F}$-predictable since it is $\mathcal{F}$-adapted and has continuous trajectories. Therefore, $\mathbb{P}(\forall t \in [0,1],\ \langle Z\rangle_t = A_t) = 1$ and this implies that for all $0 \leq s < t \leq 1$ and $\varepsilon > 0$,

$$\mathbb{P}(d_\Gamma(\frac{1}{t-s}(\langle Z\rangle_t - \langle Z\rangle_s)) > \varepsilon) \leq \liminf_n \mathbb{P}^n(d_\Gamma(\frac{1}{t-s}(\langle Z^n\rangle_t - \langle Z^n\rangle_s)) > \varepsilon). \tag{22}$$

Using then (21):

$$\mathbb{P}^n(\frac{1}{t-s}(\langle Z^n\rangle_t - \langle Z^n\rangle_s) \in \frac{\lfloor nt\rfloor - \lfloor ns\rfloor}{n(t-s)}\Gamma) = 1 \Rightarrow \mathbb{P}^n(d_\Gamma(\frac{1}{t-s}(\langle Z^n\rangle_t - \langle Z^n\rangle_s)) > |1 - \frac{\lfloor nt\rfloor - \lfloor ns\rfloor}{n(t-s)}|C_\Gamma) = 0.$$

This last equality implies that the right-hand side of (22) is equal to zero for all ε, which in turn implies (19) and we deduce finally that $\mathbb{P} \in Q_\Gamma$. The conclusion follows now easily, any sequence of maximizing joint distributions (L^n, Z^n_1) is tight in $\Delta(\mathbb{R}^d \times \mathbb{R}^d)$ and from the preceding discussion it converges to the law of (L, Z_1) fulfilling the constraints $[\![L]\!] \preceq \mu$ and $[\![Z_1]\!] \in Q_\Gamma(1)$ by construction. Since Z^n_1 has bounded second order moments and L^n has uniformly integrable second order moments (its law is dominated by μ), we have from Lemma 5.3 that

$$\mathbb{E}[\langle L^n, Z^n_1\rangle] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[\langle L, Z_1\rangle].$$

□

3.2. **The control problem and the discretization procedure.** The main result of this section is the lower bound part of Theorem 1.3 given below.

Proposition 3.5.

$$\liminf_{n\to\infty} V_n(\mu) \geq V_\infty(\mu).$$

This Proposition will be proved using the first reformulation W_{ac} of V_∞ announced in the introduction. The key argument of the proof is a two-scales discretization of the control problem based on a Central Limit Theorem. Let us at first prove that both problems are equal.

Lemma 3.6.

$$V_\infty(\mu) = W_{ac}(\mu) \triangleq \sup_{X\in\mathcal{M}_{ac}(\preceq_\mu)} \mathbb{E}[\int_0^1 r(\frac{d}{ds}\langle X\rangle_s)ds],$$

where $\mathcal{M}_{ac}(\preceq_\mu) \subset \mathcal{M}^c$ is the subset of distributions of martingales $(X_t)_{t\in[0,1]}$ whose final distribution is dominated by μ, and such that with probability 1, the quadratic variation process $(\langle X\rangle_t)_{t\in[0,1]}$ is absolutely continuous with respect to Lebesgue's measure. Moreover, the supremum in W_{ac} can be restricted to martingales with respect to a fixed d-dimensional Brownian filtration.

Proof. We assume without loss of generality that $\mu \in \Delta_0^2$. We prove at first that $W_{ac} \leq V_\infty$. Let X be a martingale whose law is in $\mathcal{M}_{ac}(\preceq_\mu)$. Then there exists on an extension[5] denoted $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$ of our filtered probability space a d-dimensional Brownian motion W and an $\mathcal{F}$-progressively measurable process $q_s \in \mathbb{M}_d$ such that $X_t = \int_0^t q_s dW_s$ (see e.g. [16] Theorem 3.4.2). Moreover, we have $\langle X\rangle_t = \int_0^t q_s q_s^T ds$. Define the progressively measurable process $\sigma_s = \phi(q_s)$ where ϕ is some measurable selection of the set-valued map $M \in \mathbb{M}_d \to argmax\{Tr(MN) \mid N \in \mathbb{M}_d : NN^T \in G\}$. The law of the process $(\int_0^t \sigma_s dW_s)_{t\in[0,1]}$ is by construction in Q_Γ and we have

$$V_\infty(\mu) \geq \mathbb{E}[\langle X_1, \int_0^1 \sigma_s dW_s\rangle] = \mathbb{E}[\int_0^1 Tr(q_s\sigma_s)ds] = \mathbb{E}[\int_0^1 r(\tfrac{d}{ds}\langle X\rangle_s)ds], \tag{23}$$

where the last equality follows from Lemma 2.3. Let us prove the reverse inequality $V_\infty \leq W_{ac}$. Consider the canonical space $C([0,1],\mathbb{R}^d)$ endowed with the standard d dimensional Wiener measure $\mathbb{P}_0$. Let $(B_t)_{t\in[0,1]}$ denote the canonical process, $\mathcal{F}^B$ its natural filtration and $\mathcal{H}_G$ be the set of $\mathbb{M}_d$-valued $\mathcal{F}^B$-progressively measurable processes ρ such that $\rho\rho^T \in G$. Define $\widetilde{Q}_G(1)$ as the set of laws of variables $\int_0^1 \rho_s dB_s$ with $\rho \in \mathcal{H}_G$. Then, using Caratheodory's Theorem, $\widetilde{Q}_G(1)$ is dense in $Q_\Gamma(1)$ (see Lemma 5.12). Using Lemma 5.3, it follows that $V_\infty(\mu) = sup\{C(\mu,\nu) \mid \nu \in \widetilde{Q}_G(1)\}$. From this equality, for all $\varepsilon > 0$, there exists an ε-optimal pair $(L, (Z_t)_{t\in[0,1]})$ defined on the same probability space as B such that $Z_t = \int_0^t \sigma_s dB_s$ for some $\mathcal{F}^B$ progressive process σ such that $\sigma_s\sigma_s^T \in G$, $[\![L]\!] \preceq \mu$, and $\mathbb{E}[\langle L, Z_1\rangle] \geq V_\infty(\mu) - \varepsilon$. We can assume that L is $\mathcal{F}_1^B$-measurable up to replace L by its conditional expectation given $\mathcal{F}_1^B$. Using the predictable representation property of the Brownian filtration, there exist an $\mathcal{F}^B$ progressive process λ_s such that $L = \int_0^1 \lambda_s dB_s$. We deduce that

$$V_\infty(\mu) - \varepsilon \leq \mathbb{E}[\langle L, Z_1\rangle] = \mathbb{E}[\int_0^1 Tr(\lambda_s\sigma_s^T)ds] \leq \mathbb{E}[\int_0^1 r(\lambda_s\lambda_s^T)ds] \leq W_{ac}(\mu),$$

which completes the proof of the second inequality and of the last assertion concerning the Brownian filtration. □

Some Technical Results. The proof of the Proposition 3.5 is based on the following three technical Lemmas 3.7, 3.8 and 3.9, whose proofs are standard and therefore postponed to section 5.2. The first Lemma is the usual Central Limit Theorem for the Wasserstein distance. Let $RC^1(q,C) \triangleq \{\mu \in \Delta_0^2 : cov(\mu) = I_d,\ \|\mu\|_q \leq C\}$. Define then $RC^n(q,C)$ as the set of rescaled convolutions of these distributions, precisely all distributions of the variables $(n^{-\frac{1}{2}}\sum_{k=1}^n S_i)$, where $(S_i)_{i=1,..,n}$ is an i.i.d. sequence of law $\mu \in RC^1(q,C)$. We will also use the notation $\mu^{\otimes n}$ for the law (in $\Delta((\mathbb{R}^d)^n)$) of $(S_i)_{i=1,..,n}$.

Lemma 3.7. *Using the previous notations and with $\mathcal{N}(0,I_d)$ being the standard centered gaussian distribution in $\mathbb{R}^d$, we have for all $q > 2$:*

$$\lim_{n\to\infty} \sup_{\nu\in RC^n(q,C)} d_{W_2}(\nu, \mathcal{N}(0,I_d)) = 0.$$

Moreover, for any fixed q, C, there exists a measurable selection $\mu \in RC^1(q,C) \to \pi(\mu) \in \mathcal{P}(\mu^{\otimes n}, \mathcal{N}(0,I_d))$ such that:

$$\mathbb{E}_{\pi(\mu)}[\|n^{-\frac{1}{2}}\textstyle\sum_{k=1}^n S_i - N\|^2] \leq \displaystyle\sup_{\nu\in RC^n(q,C)} d_{W_2}(\nu, \mathcal{N}(0,I_d)) \text{ with } [\![(S_i)_{i=1,..,n}, N]\!] = \pi(\mu).$$

Due to the Lipschitz property of V with respect to the Wasserstein distance of order p, we have the following approximation results

[5] All the extensions we consider in this work are always the canonical Wiener extensions as defined in e.g. [14].

Lemma 3.8. *For all $q > 2$ we have $\lim_{C\to\infty} z(q,C) = 0$ with*

$$z(q,C) = \sup_{M\in\mathbb{M}_d\,:\,|M|=1} \left(r(MM^T) - \sup_{\nu\in RC^1(q,C)} V(M\sharp\nu) \right),$$

where $M\sharp\nu$ denotes the image probability of ν induced by the linear map $x \to Mx$. Moreover, there exists a measurable selection $M \in \mathbb{M}_d \to \chi(M) \in RC^1(q,C)$ such that

$$r(MM^T) - V(M\sharp\chi(M)) \le |M| z(q,C).$$

Lemma 3.9. *Let $(X_k, Y_k)_{k=1,..,n}$ be two $\mathbb{R}^d$-valued martingales defined on the same probability space with respect to the same filtration $(\mathcal{F}_k)_{k=1,..,n}$. Then*

$$|\Psi_n[V]((X_k,\mathcal{F}_k)_{k=1,..,n}) - \Psi_n[V]((Y_k,\mathcal{F}_k)_{k=1,..,n})| \le \sqrt{n}\gamma \|X_n - Y_n\|_{L^2}.$$

Let us turn to the proof of the main Proposition.

Proof of Proposition 3.5. Let B be a d-dimensional Brownian motion and $\mathcal{F}^B$ its natural filtration. We assume without loss of generality that $\mu \in \Delta_0^2$. According to Lemma 3.6, for all $\varepsilon > 0$, there exists an $\mathcal{F}^B$ martingale $(L_t = \int_0^t \lambda_s dB_s)_{t\in[0,1]}$ such that $[\![L_1]\!] \preceq \mu$ and

$$\mathbb{E}[\textstyle\int_0^1 r(\lambda_s\lambda_s^T)ds] \ge V_\infty(\mu) - \varepsilon.$$

Applying Lemma 5.11 to λ, there exists a sequence of simple processes λ^n, constant on the intervals $[\frac{k}{n}, \frac{k+1}{n})$ such that $\mathbb{E}[\int_0^1 |\lambda_s - \lambda_s^n|^2\, ds] \underset{n\to\infty}{\longrightarrow} 0$. Let us denote $\lambda_s^n = \sum_{k=1}^n u_k^n \mathbb{1}_{[\frac{k-1}{n},\frac{k}{n})}(s)$ and note that u_1^n is deterministic. Using the regularity for r given in Lemma 2.3, we deduce

$$\tfrac{1}{n}\mathbb{E}[\textstyle\sum_{k=1}^n r(u_k^n(u_k^n)^T)] = \mathbb{E}[\int_0^1 r(\lambda_s^n(\lambda_s^n)^T)] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[\int_0^1 r(\lambda_s\lambda_s^T)ds]. \tag{24}$$

The idea of this proof is to construct a discrete-time approximation of the martingale $(L_t)_{t\in[0,1]}$ using two steps of discretization. The first step is the usual time-discretization on the intervals $[\frac{k}{n}, \frac{k+1}{n})$ and the second acts on the integrator B. Each increment $\Delta_k^n B \triangleq B_{k/n} - B_{(k-1)/n}$ will be replaced by a sufficiently long normalized sum of i.i.d. random variables whose laws will be chosen in order for the V-variation to be close to the R-variation.

Up to enlarging the probability space, we assume that there is a sequence $(U_i)_{i\in\mathbb{N}^*}$ of uniform random variables independent of B. Let us fix $C > 0$ and $q > 2$. According to Lemma 3.7, given a sequence ε_n converging to zero, there exists an increasing sequence N_n of integers such that

$$\forall m \ge N_n, \quad \sup_{\mu\in RC^m(q,C)} d_{W_2}(\mu, \mathcal{N}(0,I_d)) \le \varepsilon_n.$$

For a vector $(N(k,n))_{k=1,..,n}$ of integers such that $N(k,n) \ge N_n$, define the partial sums $D(k,n) = \sum_{i=1}^k N(i,n)$ and $D(0,n) = 0$. Using the notations of Lemma 3.8, define the sequence $(\nu_k^n)_{k=1,..,n}$ of $\mathbb{R}^d$-valued transition probabilities by $\nu_k^n = \chi(u_k^n)$, having the property that for any variable Y such that $[\![Y \mid u_k^n]\!] = \nu_k^n$

$$r(u_k^n(u_k^n)^T) - V([\![u_k^n Y \mid u_k^n]\!]) \le |u_k^n| z(q,C), \tag{25}$$

where $z(q,C)$ is defined in Lemma 3.8. This allows us to construct by induction (on $k = 1,..,n$) a family of random variables $(S_i)_{i=1,..,D(n,n)}$ and a filtration $(\mathcal{H}_i)_{i=1,..,D(n,n)}$ (both depending on n and of the chosen sequence $N(k,n)$) as follows. Consider the measurable selection given by Lemma 3.7,

$$\pi(\nu_k^n) \in \mathcal{P}((\nu_k^n)^{\otimes N(k,n)}, \mathcal{N}(0,I_d)).$$

Define $\mathcal{H}_0 = \sigma(\emptyset)$. At step k, using the variable U_k as a generator, construct the sequence $(S_i)_{i=D(k-1,n)+1,..,D(k,n)}$ such that the conditional law of $((S_i)_{i=D(k-1,n)+1,..,D(k,n)}, \sqrt{n}\Delta_k^n B)$ given $\mathcal{H}_{D(k-1,n)}$ is $\pi(\nu_k^n)$[6] and the filtration

$$\mathcal{H}_i = \sigma((u_{k+1}^n, \Delta_k^n B), k \le k^*(i)\,;\, S_j, j \le i) \text{ for } i = D(k-1,n)+1,..,D(k,n),$$

where $k^*(i)$ is defined by the relation $D(k^*(i),n) \le i < D(k^*(i)+1,n)$. It follows from the Lemmas 3.7 and 3.8 that

$$\mathbb{E}\left[\left| n^{\frac12}\Delta_k^n B - \sum_{i=D(k-1,n)+1}^{D(k,n)} \frac{S_i^n}{\sqrt{N(k,n)}} \right|^2 |\mathcal{H}_{D(k-1,n)} \right] \le \varepsilon_n^2, \;\; [\![S_i \mid \mathcal{H}_{i-1}]\!] = \nu_{k^*(i)}^n \in RC^1(q,C). \tag{26}$$

Consider then the martingale $(M_i = \mathbb{E}[L \mid \mathcal{H}_i], i = 0,..,D(n,n))$ and its approximation

$$\widetilde{M}_i = \sum_{k=1}^{k^*(i)} \sum_{j=D(k-1,n)+1}^{D(k,n)} \frac{u_k^n S_j}{\sqrt{nN(k,n)}} + \sum_{j=D(k^*(i),n)+1}^{i} \frac{u_{k^*(i)+1}^n S_j}{\sqrt{nN(k^*(i),n)}},$$

[6] See the discussion following Theorem 5.8 in the appendix.

which is also an $\mathcal{H}$-martingale. Using Lemma 3.9, we have

$$\left|\Psi_{D(n,n)}[V]((M_i,\mathcal{H}_i)_{i=1,..,D(n,n)}) - \Psi_{D(n,n)}[V]((\widetilde{M}_i,\mathcal{H}_i)_{i=1,..,D(n,n)})\right| \leq \gamma\sqrt{D(n,n)}\left\|L-\widetilde{M}_{D(n,n)}\right\|_{L^2}, \tag{27}$$

where we replaced $M_{D(n,n)}$ by L using the martingale property and Jensen's inequality. Moreover:

$$\begin{aligned}\left\|L-\widetilde{M}_{D(n,n)}\right\|_{L^2} &\leq \mathbb{E}[\int_0^1 |\lambda_s-\lambda_s^n|^2 ds]^{\frac{1}{2}} + \mathbb{E}\left[\sum_{k=1}^n \mathbb{E}[|u_k^n(\Delta_k^n B - \sum_{i=D(k-1,n)+1}^{D(k,n)} \frac{S_i^n}{\sqrt{nN(k+1,n)}})|^2 \mid \mathcal{F}^B_{k-1/n}]\right]^{\frac{1}{2}} \\ &\leq \mathbb{E}[\int_0^1 |\lambda_s-\lambda_s^n|^2 ds] + \alpha\varepsilon_n\mathbb{E}[\int |\lambda_s^n|^2 ds]^{\frac{1}{2}} \underset{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

using (26) and where α is a constant such that $|Px| \leq \alpha|P||x|$ for all $P \in \mathbb{M}_d$ and $x \in \mathbb{R}^d$. Using (27), these inequalities reduce our problem to the study of the V-variation of $\widetilde{M}$.

$$\begin{aligned}\Psi_{D(n,n)}[V]((\widetilde{M}_i,\mathcal{H}_i)_{i=1,..,D(n,n)}) &= \mathbb{E}\left[\sum_{k=1}^n \sum_{i=D(k-1,n)+1}^{D(k,n)} V([\![\widetilde{M}_i \mid \mathcal{H}_{i-1}]\!])\right] \\ &\geq \mathbb{E}\left[\sum_{k=1}^n \sum_{i=D(k-1,n)+1}^{D(k,n)} \frac{r(u_k^n(u_k^n)^T) - |u_k^n|z(q,C)}{\sqrt{nN(k,n)}}\right] = \mathbb{E}\left[\sum_{k=1}^n \frac{1}{n}\sqrt{nN(k,n)}(r(u_k^n(u_k^n)^T) - |u_k^n|z(q,C))\right] \\ &\geq \sqrt{D(n,n)}\left(\mathbb{E}[\int_0^1 r(\lambda_s^n(\lambda_s^n)^T)ds] - z(q,C)\mathbb{E}[\int_0^1 |\lambda_s^n|ds]\right. \\ &\quad \left. - \max_{k=1,..,n}\left|\frac{\sqrt{nN(k,n)}}{\sqrt{D(n,n)}} - 1\right|\left|\mathbb{E}[\int_0^1 r(\lambda_s^n(\lambda_s^n)^T)ds] - z(q,C)\mathbb{E}[\int_0^1 |\lambda_s^n|ds]\right|\right),\end{aligned}$$

where the first inequality follows from (25). Using the former results, for any sequence of vectors $N(k,n)$ indexed by n such that

$$\max_{k=1,..,n}\left|\frac{\sqrt{nN(k,n)}}{\sqrt{D(n,n)}} - 1\right| \underset{n\to\infty}{\longrightarrow} 0, \tag{28}$$

we have

$$\liminf_{n\to\infty} \frac{\Psi_{D(n,n)}[V]((\widetilde{M}_i,\mathcal{H}_i)_{i=1,..,D(n,n)})}{\sqrt{D(n,n)}} \geq V_\infty(\mu) - z(q,C)\mathbb{E}[\int_0^1 |\lambda_s|ds] - \varepsilon.$$

The condition (28) is not restrictive since for fixed n, any vector of integers $N(k,n) \in \{m; m+1\}$ for $m \geq N_n$ is such that

$$\max_{k=1,..,n}\left|\frac{\sqrt{nN(k,n)}}{\sqrt{D(n,n)}} - 1\right| \leq \frac{1}{N_n},$$

and then any value above nN_n is admissible for $D(n,n)$. It implies that

$$\liminf_{n\to\infty} V_n(\mu) \geq V_\infty(\mu) - z(q,C)\mathbb{E}[\int_0^1 |\lambda_s|ds] - \varepsilon.$$

The result follows by sending C to $+\infty$ and ε to 0. □

4. Characterization of the solutions.

In this section we prove at first that the solutions of the discrete-time problem converge to the solution of the control problem (Theorem 1.4). This result is based on a reformulation of the control problem which is proved to admit maximizers. Using this reformulation, we solve directly the problem in the unidimensional case, obtaining thus the main Theorem which was proved in [9]. However, the method we use does not extend to higher dimension. In order to solve the general case, we study the dual problem and prove that it is related to a PDE problem using tools from Optimal Transport theory. The main result is then the Verification Theorem 4.13, which relates primal and dual solutions through the gradient of the solution of the dual PDE.

4.1. **A first reduction.**

Before going to the proof of Theorem 1.4, let us recall some properties of the Meyer-Zheng topology (denoted M-Z hereafter, see [19]) on the space of martingale distributions. This topology is the weak convergence on $\Delta(\mathbb{D}([0,1],\mathbb{R}^d))$ when the set $\mathbb{D}([0,1],\mathbb{R}^d)$ of càdlàg functions is endowed with the topology of convergence in measure with respect to Lebesgue's measure (denoted λ) together with the convergence of the value at time 1 : a sequence y_n converges to y in $\mathbb{D}([0,1],\mathbb{R}^d)$ if

$$\forall\varepsilon>0,\ \lambda(\{|y_n(x)-y(x)| \geq \varepsilon\}) \underset{n\to\infty}{\longrightarrow} 0, \qquad \text{and} \quad y_n(1) \underset{n\to\infty}{\longrightarrow} y(1).$$

Note that this topology is a product topology, i.e. topologies of $\mathbb{D}([0,1],\mathbb{R}^d)$ and $\mathbb{D}([0,1],\mathbb{R})^d$ coincide.

Lemma 4.1. *The sets of (laws of) martingales uniformly bounded in L^q for some $q \geq 1$ and the set $\mathcal{M}(\preceq_\mu)$ of martingales whose law at time 1 is dominated by μ are compact subsets of $\Delta(\mathbb{D}([0,1],\mathbb{R}^d))$ for the M-Z topology.*

Proof. The topology introduced in [19] was defined on $\mathbb{D}([0,\infty),\mathbb{R}^d)$ and the definition given above is just the induced topology on $\mathbb{D}([0,1],\mathbb{R}^d)$ which is seen as the closed subset of functions that remain constant after time 1. The first result is therefore a Corollary of Theorem 2 in [19]. The second follows from the fact that the projection $(X_t)_{t\in[0,1]} \to X_1$ at time 1 is continuous and that the condition $[\![X_1]\!] \preceq \mu$ is closed. □

Lemma 4.2.

$$V_\infty(\mu) = W(\mu) \triangleq \max_{[\![(X_t)_{t\in[0,1]}]\!]\in\mathcal{M}(\preceq_\mu)} H([\![(X_t)_{t\in[0,1]}]\!]) \text{ with } H(\mathbb{P}) \triangleq \max_{[\![(X_t,Z_t)_{t\in[0,1]}]\!]\in\mathcal{M}(\mathbb{P},Q_\Gamma)} \mathbb{E}[\langle X_1, Z_1\rangle], \tag{29}$$

where $\mathcal{M}(\mathbb{P},Q_\Gamma)$ is the set of martingales laws in $\Delta(\mathbb{D}([0,1],\mathbb{R}^{2d}))$ of processes (X,Z) such that $[\![(X_t)_{t\in[0,1]}]\!] = \mathbb{P}$ and $[\![(Z_t)_{t\in[0,1]}]\!] \in Q_\Gamma$ (using the identification of the continuous functions as a subset of $\mathbb{D}$). The set of maximizers is a non empty (M-Z)-compact convex subset of $\mathcal{M}(\preceq_\mu)$ denoted $\mathcal{P}_\infty(\mu)$.

Proof. From the definition of V_∞, we have

$$V_\infty(\mu) = \max_{[\![(Z_t)_{t\in[0,1]}]\!]\in Q_\Gamma, [\![L]\!]\preceq\mu} \mathbb{E}[\langle L, Z_1\rangle].$$

Therefore,

$$W(\mu) = \max_{[\![(X_t,Z_t)_{t\in[0,1]}]\!]\in\mathcal{M}(\preceq_\mu,Q_\Gamma)} \mathbb{E}[\langle X_1, Z_1\rangle] \leq V_\infty(\mu) \text{ where } \mathcal{M}(\preceq_\mu,Q_\Gamma) \triangleq \cup_{\mathbb{P}\in\mathcal{M}(\preceq_\mu)}\mathcal{M}(\mathbb{P},Q_\Gamma),$$

since the marginal distribution of $[\![X_1,(Z_t)_{t\in[0,1]}]\!]$ fulfills the constraints of the definition. For the converse inequality, just define $X_t \triangleq \mathbb{E}[L \mid Z_s,\ s \leq t]$. The set $\mathcal{M}(\preceq_\mu,Q_\Gamma)$ is (M-Z)-compact convex since it is the intersection of the set of martingale distributions uniformly bounded in L^2 by $(C_\Gamma + \|\mu\|_2)$ and of the set $\mathcal{P}(\mathcal{M}(\preceq_\mu),Q_\Gamma)$. Compactness and convexity of $\mathcal{P}(\mathcal{M}(\preceq_\mu),Q_\Gamma)$ follow from Lemma 5.1. Indeed, the M-Z topology is a product topology and it is weaker than the Skorokhod's topology (so that Q_Γ is M-Z compact). The application $[\![(X_t,Z_t)_{t\in[0,1]}]\!] \longrightarrow \mathbb{E}[\langle X_1,Z_1\rangle]$ is (M-Z)-continuous and affine on $\mathcal{M}(\preceq_\mu,Q_\Gamma)$ since the projection at time 1 is linear and continuous and using Lemma 5.3. We deduce that the set of maximizers is nonempty and compact convex. Its marginal projection $\mathcal{P}_\infty(\mu)$ on the first coordinate of the product $\mathbb{D}([0,1],\mathbb{R}^d)^2$ is then compact convex. □

Proof of Theorem 1.4. Using the proof of Lemma 2.10, given an optimal sequence of martingales $((L^n_k)_{k=1,..,n})_{n\in\mathbb{N}}$, we can construct a sequence $((L^n_k,S^n_k)_{k=1,..,n})_{n\in\mathbb{N}}$ such that $[\![(S^n_k)_{k=1,..,n}]\!] \in T^n$ and

$$\Psi_n[V]((L_k)_{k=1,..,n}) \leq \mathbb{E}[\langle L^n_n, \textstyle\sum_{k=1}^n S^n_k\rangle].$$

Define $Z^n_t \triangleq n^{-\frac{1}{2}}\sum_{k=1}^{\lfloor nt\rfloor} S^n_k$ and $X^n_t \triangleq L^n_{\lfloor nt\rfloor}$. Z^n is martingale from the definition of T^n (see Definition 2.9) and the sequence of joint distributions $[\![(X^n_t,Z^n_t)_{t\in[0,1]}]\!]$ is (M-Z)-relatively compact from Lemma 4.1 since $[\![X^n_1]\!] \preceq \mu$ and $cov(Z^n_1) \in \Gamma$. Any limit distribution is a martingale using that the sets of uniformly L^2 bounded martingale's distributions are closed. The marginal laws of the coordinate processes of any limiting distribution are respectively in the compact sets $\mathcal{M}(\preceq_\mu)$ by Lemma 4.1 and Q_Γ using Proposition 3.4 (convergence to an element of Q_Γ holds for a stronger topology along a subsequence). Moreover, using Lemma 5.3, the application

$$[\![(X^n_t,Z^n_t)_{t\in[0,1]}]\!] \longrightarrow \mathbb{E}[\langle X^n_1, Z^n_1\rangle]$$

is continuous and since by hypothesis $\mathbb{E}[\langle X^n_1,Z^n_1\rangle] \underset{n\to\infty}{\longrightarrow} V_\infty(\mu)$, we deduce that the limiting distribution of $[\![(X^n_t)_{t\in[0,1]}]\!]$ belongs to $\mathcal{P}_\infty(\mu)$. □

4.2. **The unidimensional case.** In the following, F_ν, F^{-1}_ν denote respectively the distribution function of a probability $\nu \in \Delta(\mathbb{R})$ and its right-continuous generalized inverse, and $\mathcal{N}(0,\eta)$ is the centered gaussian distribution with variance $\eta > 0$.

Proposition 4.3. *Assume that $d = 1$ and A1-A4. Then for all $\mu \in \Delta^2$, $\mathcal{P}_\infty(\mu)$ is reduced to a point which is the law of the martingale X defined by*

$$\forall t \in [0,1],\ X_t \triangleq \mathbb{E}[f_\mu(B_1) \mid \mathcal{F}^B_t],$$

with B a standard Brownian motion, $f_\mu \triangleq F^{-1}_\mu \circ F_{\mathcal{N}(0,\rho^2)}$, and $\rho \triangleq r(1)$.

Proof. Using Lemma 2.3 (which does not rely on A5), the set $F = \{r \leq 1\}$ is the interval $[0, 1/\rho^2]$ where $\rho^2 > 0$ using A1. The set $\widehat{F}$ is the closed ball in L^2_0 of radius $1/\rho^2$, and this clearly implies A5. Now $\widehat{G}$ is the ball of radius ρ^2 and therefore Q_Γ is the set of distributions of continuous $\mathbb{R}$-valued martingales such that $\langle Z\rangle$ is ρ^2-lipschitz with respect to the time-variable. From Theorem 1.3, we have

$$V_\infty(\mu) = sup\{C(\mu,\nu) \mid \nu \in Q_\Gamma(1)\}.$$

Let us prove that this problem admits a unique solution which is the joint law of the pair $(f_\mu(B_1), B_1)$. At first, note that the Gaussian law $\mathcal{N}(0,\rho^2)$ is the unique maximal element in $Q_\Gamma(1)$ for the convex order and the unique maximizer of the above problem. To see this let h be a non-linear closed convex function, Z a process in Q_Γ, and q be a process with values in $[0,\rho]$ such that $\langle Z\rangle_t = \int_0^t q_s ds$. Given W a standard Brownian motion independent of Z, define

$$B_t = Z_t + \int_0^t \sqrt{\rho - q_s} dW_s.$$

Then B is a Brownian motion with variance ρ^2 and the conditional law of B_1 given $\mathcal{F}_1^Z$ is a gaussian distribution with variance $\int_0^1(\rho - q_s)ds$ (see Proposition 1.1 in [14]). If $Z \neq B$, then $\mathbb{P}(\int_0^1(\rho - q_s)ds > 0) > 0$, and $\mathbb{E}[h(B_1) \mid \mathcal{F}_1^Z] > h(Z_1)$ on this set so that $\mathbb{E}[h(B_1)] > \mathbb{E}[h(Z_1)]$. Using Theorem 5.4, and assuming that μ is not a Dirac mass (otherwise the result is obvious), then there exists a non-linear convex function h such that

$$C(\mu, [\![Z_1]\!]) \leq \langle h^*, \mu\rangle + \langle h, [\![Z_1]\!]\rangle < \langle h^*, \mu\rangle + \langle h, [\![B_1]\!]\rangle = C(\mu, [\![B_1]\!])$$

Finally, the problem $C(\mu, \mathcal{N}(0,\rho^2))$ is known to admit a unique solution (see e.g. [1] Theorem 6.0.2) which is the law of the pair $(f_\mu(B_1), B_1)$. Uniqueness implies that this relation is met by any maximizer of $\mathbb{E}[\langle X_1, Z_1\rangle]$ in $\mathcal{M}(\mu, Q_\Gamma)$: Any optimal (X,Z) is such that $Z = \sqrt{\rho}B$ with B a standard Brownian motion and $X_1 = f_\mu(B_1)$. To conclude, elements in $\mathcal{M}(\mu, Q_\Gamma)$ being martingales, B is an $\mathcal{F}^{X,B}$-Brownian motion. Therefore,

$$\forall t \in [0,1],\ X_t = \mathbb{E}[X_1 \mid \mathcal{F}_t^{X,B}] = \mathbb{E}[f_\mu(B_1) \mid \mathcal{F}_t^B] = u(t, B_t),$$

where u is the solution of the backward heat equation with terminal condition $u(1,x) = f_\mu(x)$. □

We recover exactly the main Theorem given in [9], where this particular martingale is called continuous martingale of maximal variation (CMMV) with terminal law μ. Moreover, the limiting martingale having continuous trajectories, the convergence given in Theorem 1.4 extends to convergence for the usual Skorokhod's topology (see [21]). This method does not work anymore in higher dimension, since uniqueness of a maximal element for the convex order in $Q_\Gamma(1)$ fails. However, the link with a PDE problem outlined in the preceding proof can be generalized as shown in the following results.

4.3. **The dual problem.** From this point, we will restrict our attention to laws concentrated on some fixed compact convex subset $K \subset \mathbb{R}^d$ in order to shorten the proofs. We mention that all our results can be easily extended on Δ^2 using ad hoc growth or integrability assumptions on the dual variables. Let us prove the following dual representation for V_∞ which is very similar to Kantorovitch Duality Theorem (see Theorem 5.5) in the theory of Optimal Transport. In the following, we use the notation $\langle \phi, \mu\rangle \triangleq \mathbb{E}_\mu[\phi(L)]$ for expectations.

Proposition 4.4.

$$\forall \mu \in \Delta(K),\ V_\infty(\mu) = \inf_{\phi \in C(K)} (\langle \phi, \mu\rangle + V_\infty^*(\phi)) = \min_{\phi \in Conv(K)} (\langle \phi, \mu\rangle + V_\infty^*(\phi)) \tag{30}$$

$$with \quad V_\infty^*(\phi) \triangleq \sup_{\nu \in Q_\Gamma(1)} \langle \phi^*, \nu\rangle,$$

where $C(K)$ denotes the set of continuous functions on K, $Conv(K)$ the proper closed convex functions from $\mathbb{R}^d$ to $\mathbb{R} \cup \{+\infty\}$ such that $Dom(f) \subset K$, and ϕ^* the Fenchel transform of ϕ. Let us also denote

$$\partial V_\infty(\mu) \triangleq \operatorname*{argmin}_{\phi \in Conv(K)} (\langle \phi, \mu\rangle + V_\infty^*(\phi)).$$

Proof. Note at first that we can replace the constraint $[\![L]\!] \preceq \mu$ given in the definition of V_∞ by $[\![L]\!] = \mu$ since the maximal covariance functions $C(.,\nu)$ defined in section 5 are nondecreasing for the convex order (see Lemma 5.7). Applying then Theorem 5.4, we obtain

$$V_\infty(\mu) = \max_{\nu \in Q_\Gamma(1)} \inf_{(\phi - \frac{1}{2}|.|^2, \psi - \frac{1}{2}|.|^2) \in C_b(\mathbb{R}^d)^2;\, \phi + \psi \geqslant \langle .,. \rangle} (\langle \phi, \mu\rangle + \langle \psi, \nu\rangle).$$

$Q_\Gamma(1)$ is a compact subset of Δ^2, and weak convergence coincides in this set with the d_{W_2}-convergence since moments of order $q > 2$ are uniformly bounded. Therefore and since the function ψ in the above expression of $V_\infty(\mu)$ has at most quadratic growth, the application $\nu \to \langle \psi, \nu\rangle$ is affine and weakly continuous on $Q_\Gamma(1)$. On the other hand, the application $(\phi, \psi) \to (\langle \phi, \mu\rangle + \langle \psi, \nu\rangle)$ is affine on the convex set

$$\{(\phi, \psi) \in (\frac{1}{2}|.|^2 + C_b(\mathbb{R}^d)) \times (\frac{1}{2}|.|^2 + C_b(\mathbb{R}^d)) : \phi + \psi \geqslant \langle .,. \rangle\},$$

so that the Minmax Theorem ([22]) implies:

$$V_\infty(\mu) = \inf_{(\phi - \frac{1}{2}|.|^2, \psi - \frac{1}{2}|.|^2) \in C_b(\mathbb{R}^d)^2;\, \phi + \psi \geqslant \langle .,. \rangle} (\langle \phi, \mu\rangle + \max_{\nu \in Q_\Gamma(1)} \langle \psi, \nu\rangle).$$

Any $\phi \in C(K)$ is identified with the function equal to ϕ on K and to $+\infty$ otherwise. Since for any pair (ϕ, ψ) we have $(\phi^*)^* \leq \phi$ and $\phi^* \leq \psi$, we infer:

$$V_\infty(\mu) = \inf_{\phi}(\langle \phi, \mu \rangle + \sup_{\nu \in Q_\Gamma(1)} \langle \phi^*, \nu \rangle), \tag{31}$$

where the infimum is taken over convex functions $\phi \in \frac{1}{2}|.|^2 + C_b(\mathbb{R}^d)$. Finally, equality still holds for $\phi \in C(K)$ using Fenchel's Lemma. Let now (ϕ_n, ψ_n) be a minimizing sequence with $\phi_n \in C(K)$ and $\psi_n = \phi_n^*$. Replacing (ϕ_n, ψ_n) by $(\phi_n - \alpha, \psi_n + \alpha)$ with $\alpha = \min_{x \in K} \phi_n(x)$, we obtain that

$$\forall x \in \mathbb{R}^d,\ \phi_n(x) \geq 0, \quad \psi_n(x) \geq -C_K|x|, \quad \psi_n(0) = 0.$$

The functions ψ_n being uniformly Lipschitz, Ascoli's Theorem implies that the sequence ψ_n is relatively compact in $C(\mathbb{R}^d)$ for the uniform convergence on compact sets. Let ψ denote the limit of some convergent subsequence also denoted ψ_n. Pointwise convergence implies that $\psi(0) = 0$, and we deduce therefore from Fatou's Lemma that

$$\sup_{\nu \in Q_\Gamma(1)} \langle \psi, \nu \rangle \leq \liminf_{n \to \infty} \sup_{\nu \in Q_\Gamma(1)} \langle \psi_n, \nu \rangle.$$

For $\ell \in \mathbb{N}$, let ξ_{B_ℓ} the convex indicator function equal to 0 on B_ℓ and $+\infty$ otherwise. For any function f, we define $f^{*\ell} = (f + \xi_{B_\ell})^*$, so that the sequence $f^{*\ell}$ is nondecreasing and converges pointwise to f^*. Using that Fenchel transform is an isometry for the uniform norm, ℓ being fixed, $\psi_n^{*\ell}$ converges uniformly to $\psi^{*\ell}$ when n goes to $+\infty$. Using these notations

$$\int \psi^{*\ell} d\mu = \lim_{n \to \infty} \int \psi_n^{*\ell} d\mu \leq \liminf_{n \to \infty} \int \psi_n^* d\mu.$$

Monotone convergence implies $\lim_{\ell \to \infty} \int \psi^{*\ell} d\mu = \int \psi^* d\mu$, and therefore $\int \psi^* d\mu \leq \liminf_{n \to \infty} \int \phi_n d\mu$. Finally, the pair (ψ^*, ψ) is optimal and has the required properties. □

The next result is quite similar to the characterization given in Theorem 5.5.

Lemma 4.5. *In the following, $\phi \in Conv(K)$, μ denotes the law of the variable L in $\Delta(K)$, and Z is a process whose law is in Q_Γ, both defined on the same probability space. The two following assertions are equivalent*

i) *$L \in \partial\phi^*(Z_1)$ almost surely, and $\mathbb{E}[\phi^*(Z_1)] = \sup_{\nu \in Q_\Gamma(1)} \langle \phi^*, \nu \rangle$.*

ii) *The joint distribution of (L, Z_1) is optimal for $V_\infty(\mu)$ and $\phi \in \partial V_\infty(\mu)$.*

Proof. It follows directly from the definition of V_∞ and Fenchel's Lemma. Indeed, suppose ii)

$$V_\infty(\mu) = \mathbb{E}[\langle L, Z_1 \rangle] \leq \mathbb{E}[\phi(L) + \phi^*(Z_1)] \leq \langle \phi, \mu \rangle + \sup_{\nu \in Q_\Gamma(1)} \langle \phi^*, \nu \rangle = V_\infty(\mu).$$

Therefore, all the above inequalities are equalities, and $\langle L, Z_1 \rangle = \phi(L) + \phi^*(Z_1)$ with probability 1 which proves the result by Fenchel's Lemma. Conversely, if i) is true, then it follows from (30) that

$$V_\infty(\mu) \geq \mathbb{E}[\langle L, Z_1 \rangle] = \mathbb{E}[\phi(L) + \phi^*(Z_1)] = \langle \phi, \mu \rangle + \sup_{\nu \in Q_\Gamma(1)} \langle \phi^*, \nu \rangle \geq V_\infty(\mu),$$

which completes the proof. □

Let us now characterize V_∞^* as a second-order nonlinear PDE problem (HJB) using classical stochastic control results. We know from Proposition 4.4 and Lemma 4.5 that all the optimizers of V_∞ are linked with the optimizers of the dual problem V_∞^*. Moreover, the set of dual variables ϕ can be restricted to the set of functions such that ϕ^* is a C_K-lipschitz function such that $\phi^*(0) = 0$. Define the associated time-dependent value function

$$u : (0,1] \times \mathbb{R}^d \longrightarrow \mathbb{R} : (t, x) \longrightarrow \sup_{\mathbb{P} \in Q_\Gamma} \mathbb{E}_{\mathbb{P}}[\phi^*(x + X_{1-t})].$$

Then we have, with ∇^2 for the spatial hessian matrix:

Proposition 4.6. *The function u is the unique continuous viscosity solution of*

$$\begin{cases} -\frac{\partial}{\partial t} u - \frac{1}{2} \sup_{P \in \Gamma} Tr(P \nabla^2 u) &= 0 \quad in \quad [0,1) \times \mathbb{R}^d \\ u(1, x) &= \phi^*(x) \quad in \quad \mathbb{R}^d \end{cases} \tag{32}$$

in the class of C_K-Lipschitz functions. Moreover, $V_\infty^*(\phi) = u(0,0)$.

Proof. Using Lemma 5.12, we recover the classical framework of stochastic control, and the result can be found in [11]. Note also that our formulation is a special case of the G-expectation (see e.g. [10]) introduced by Peng. □

4.4. **A verification Theorem.** Based on the dual equality (Proposition 4.4) and on the relationships between optimal variables of the primal and dual problems obtained in Lemma 4.5, we will provide a characterization of the set $\mathcal{P}_\infty$ of optimal solutions for our control problem. Our main result is the Verification Theorem 4.13. The main ingredients of the proof are the dynamic programming equation given in Lemma 4.9 and the increasing property of V_∞ given in Lemma 4.10 which allow to prove that the constraint of domination with respect to the convex order is always active.

In order to study the dynamic properties of the control problem, we introduce below a time-dependent value function. Let us also recall Notation 3.2 for the sets $Q_\Gamma(t)$ and $\pi_t(Q_\Gamma)$.

Definition 4.7.

$$U(t,\mu) \triangleq \sup_{\nu \in Q_\Gamma(t)} C(\mu,\nu).$$

Lemma 4.8. *For all $t \in [0,1]$ and $\mu \in \Delta^2$, we have $U(t,\mu) = \sqrt{t} V_\infty(\mu)$.*

Proof. If $(Z_t)_{t\geq 0}$ is a martingale, then $Y_t = (\alpha^{-\frac{1}{2}} Z_{\alpha t})$ is a martingale such that $\langle Y\rangle_t = \alpha^{-1}\langle Z\rangle_{\alpha t}$. It follows easily that $Q_\Gamma(t) = \sqrt{t}\sharp Q_\Gamma(1)$ (i.e. the image probabilities induced by the map $x \to \sqrt{t}x$) and we conclude the proof using that $\nu \to C(\mu,\nu)$ is positively homogenous in the sense (A3). □

In the following Lemma, we prove a dynamic programming equation associated to the control problem.

Lemma 4.9. *Let $\mu_1 \preceq \mu_2 \in \Delta^2$ and (S_1, S_2) be a martingale such that $S_i \sim \mu_i$ for $i = 1, 2$. Then,*

$$\forall t \in [0,1], \quad V_\infty(\mu_2) \geq U(t,\mu_1) + \mathbb{E}[U(1-t, [\![S_2 \mid S_1]\!])].$$

Moreover, for all (law of) martingale $(X_s)_{s\in[0,1]}$ in $\mathcal{P}_\infty(\mu)$ and $t \in [0,1]$, we have

$$V_\infty(\mu) = U(t, [\![X_t]\!]) + \mathbb{E}[U(1-t, [\![X_1 - X_t \mid X_t]\!])].$$

Proof. Let us prove the first assertion. Note at first that the cases $t = 0$ and $t = 1$ follow respectively from Jensen's inequality (Lemma 5.10 in the appendix) and from the convex nondecreasing property of V_∞ (which is a supremum of nondecreasing functions). Let $t \in (0,1)$ and $(\widehat{S}_1, (W_s)_{s\in[0,t]})$ be optimal for the problem $U(t,\mu_1)$, which means

$$[\![\widehat{S}_1]\!] = \mu_1, \ [\![(W_s)_{s\in[0,t]}]\!] \in \pi_t(Q_\Gamma), \text{ and } U(t,\mu_1) = \mathbb{E}[\langle \widehat{S}_1, W_t\rangle].$$

Let $F(x)$ be a version of the conditional law of S_2 given $S_1 = x$ and let Ψ be a measurable selection on Δ^2 of the set-valued mapping (see Proposition 7.33 in [3])

$$\mu \to \operatorname*{argmax}_{\pi \in \mathcal{P}(\mu, \pi_{1-t}(Q_\Gamma))} \int \langle x, y(1-t)\rangle d\pi(x, y(.)).$$

Construct on an enlarged probability space a variable $(\widehat{S_2}, (Y_u)_{u\in[0,1-t]})$ whose conditional law given $(\widehat{S}_1, (W_s)_{s\in[0,t]})$ is $\Psi(F(S_1))$. The law of the process $\widehat{Z}_s \triangleq W_{s\wedge t} + Y_{(s\vee t)-t}$ is in Q_Γ (using e.g. Lemma 6.1.1 in [23] for the martingale property and the fact that $\pi_{1-t}(Q_\Gamma)$ is a closed convex set) and

$$\begin{aligned} U(t,\mu_1) + \mathbb{E}[U(1-t, [\![S_2 \mid S_1]\!])] &= \mathbb{E}[\langle \widehat{S}_1, W_t\rangle] + \mathbb{E}[\mathbb{E}[\langle \widehat{S_2}, Y_{1-t}\rangle \mid \widehat{S}_1, (W_s)_{s\in[0,t]}]] \\ &= \mathbb{E}[\langle \widehat{S_2}, \widehat{Z_1}\rangle] \leq V_\infty(\mu_2). \end{aligned}$$

Let us prove the second assertion. Using Lemma 4.2, for any law in $\mathcal{P}_\infty(\mu)$, there exists a (law of) martingale $(X_s, Z_s)_{s\in[0,1]}$ maximizing $\mathbb{E}[\langle X_1, Z_1\rangle]$ in $\mathcal{M}(\preceq_\mu, Q_\Gamma)$ such that $(X_s)_{s\in[0,1]}$ follows the chosen law. Note at first that the martingale property implies

$$V_\infty(\mu) = \mathbb{E}[\langle X_1, Z_1\rangle] = \mathbb{E}[\langle X_t, Z_t\rangle] + \mathbb{E}[\langle X_1 - X_t, Z_1 - Z_t\rangle]$$

Assume that $\mathbb{E}[\langle X_t, Z_t\rangle] < U(t, [\![X_t]\!])$. Let $\Phi(x)$ be a version of the conditional law of $(X_1 - X_t, (Z_s - Z_t)_{s\geq t})$ given $X_t = x$. That Φ has its values almost surely in $\mathcal{P}(\Delta^2, \pi_{1-t}(Q_\Gamma))$ follows from the fact that the conditional law of $(Z_s - Z_t)_{s\geq t}$ given X_t is almost surely a law of martingale (see e.g. Theorem 1.2.10 in [23]) as well as the face property given in Lemma 3.3. Let $(S, (W_s)_{s\in[0,t]})$ be a pair such that $[\![S]\!] = [\![X_t]\!]$, $[\![(W_s)_{s\in[0,t]}]\!] \in \pi_t(Q_\Gamma)$ and $\mathbb{E}[\langle S, W_t\rangle] = U(t, [\![X_t]\!])$. Construct on a possibly enlarged probability space a pair $(T, (Y_s)_{s\in[0,1-t]})$ whose conditional law given $(S, (W_{s\in[0,t]}))$ is $\Phi(S)$. It follows that $[\![S+T]\!] = \mu$, $[\![(\widehat{Z}_s)_{s\in[0,1]}]\!] \in Q_\Gamma$ with $\widehat{Z}_s \triangleq W_{s\wedge t} + Y_{(s\vee t)-t}$ and

$$\mathbb{E}[\langle S+T, \widehat{Z}_1\rangle] = \mathbb{E}[\langle S, W_t\rangle] + \mathbb{E}[\langle T, Y_{1-t}\rangle] > V_\infty(\mu)$$

which contradicts the definition of $V_\infty(\mu)$. The second part of the proof is similar to the proof of the first assertion. □

Lemma 4.10. *The function V_∞ is strictly increasing with respect to the convex order.*

Proof. Let $\mu_1 \preceq \mu_2 \in \Delta^2$ and let (S_1, S_2) a martingale such that $S_i \sim \mu_i$ for $i = 1, 2$. Assume that $V_\infty(\mu_1) = V_\infty(\mu_2)$. From Lemma 4.9, we have for all $t \in (0, 1)$

$$V_\infty(\mu_2) \geq U(t, \mu_1) + \mathbb{E}[U(1-t, [\![S_2 \mid S_1]\!])] = \sqrt{t} V_\infty(\mu_1) + \sqrt{1-t}\mathbb{E}[V_\infty([\![S_2 \mid S_1]\!])].$$

This implies

$$\mathbb{E}[V_\infty([\![S_2 \mid S_1]\!])] \leq \frac{1-\sqrt{t}}{\sqrt{1-t}} V_\infty(\mu_2),$$

and we deduce that the first term is equal to zero by sending t to 1. In order to conclude that $\mu_1 = \mu_2$, it remains to prove that $V_\infty(\mu) = 0$ implies that μ is a Dirac mass. Recall that $V_\infty(\mu) = sup\{C(\mu, \nu) \mid \nu \in Q_\Gamma(1)\}$. Using Lemma 5.7, it is then sufficient to prove that $Q_\Gamma(1)$ contains a law which is absolutely continuous with respect to Lebesgue's measure. But Q_Γ contains the laws of Brownian motion processes with constant instantaneous covariance equal to a non-degenerate matrix $P \in \Gamma$ (which exists since G is a neighborhood of 0) and this concludes the proof. □

Notation 4.11. *Given some function $\phi \in Conv(\mathbb{R}^d)$, and some $\mathbb{R}^d$-valued random variables Z, X, then*

$$X = \nabla\phi(Z) \text{ means } \begin{cases} \mathbb{P}(Z \in \{x \in \mathbb{R}^d : \nabla\phi \text{ exists}\}) = 1 \\ \mathbb{P}(X \in \partial\phi(Z)) = 1 \end{cases}.$$

In this case, the random variable $\nabla\phi(Z)$ is well-defined and is almost surely equal to $g(Z)$ for any measurable selection g of the subdifferential of ϕ.

Proposition 4.12. *Let $\mu \in \Delta(K)$ and $\phi \in \partial V_\infty(\mu)$. Then for any optimal joint distribution of (L, Z_1) in the problem $V_\infty(\mu)$, we have*

$$L = \nabla\phi^*(Z_1) \quad \text{almost surely.}$$

Moreover, for any (law of) martingale $(X_t, Z_t)_{t\in[0,1]}$ in $\mathcal{M}(\preceq_\mu, Q_\Gamma)$ maximizing $\mathbb{E}[\langle X_1, Z_1\rangle]$, we have for all $t \in [0, 1]$

$$X_t = \nabla u(t, Z_t) \quad \text{almost surely,}$$

where u is the solution of (32).

Proof. Using Lemma 4.5, for any optimal variables (L, Z_1) we have $L \in \partial\phi^*(Z_1)$ almost surely. It follows that

$$V_\infty(\mu) = \mathbb{E}[\langle L, Z_1\rangle] = \mathbb{E}[\langle g(Z_1), Z_1\rangle] = V_\infty(\widehat{\mu}),$$

where μ is the law of L and $\widehat{\mu}$ the law of $g(Z_1) = \mathbb{E}[L \mid Z_1]$. The last equality follows from Lemma 4.5 since $g(Z_1) \in \partial\phi^*(Z_1)$ using that the subdifferential of ϕ^* has closed convex values. On the other hand, $\widehat{\mu} \preceq \mu$. Using Lemma 4.10, V_∞ is strictly increasing and therefore $\mu = \widehat{\mu}$ which implies $L = g(Z_1)$. To conclude, define the variable Y such that its conditional law given Z_1 is uniform on the set[7] $(g(Z_1) + B(0, \varepsilon)) \cap \partial\phi^*(Z_1)$ and a Dirac mass on $\nabla\phi^*(Z_1)$ when this set is reduced to a single point. $Y \in L^2$ and $\widehat{Y} = \mathbb{E}[Y \mid Z_1] \in L^2$ since $|Y - L| \leq \varepsilon$. Applying again the Lemmas 4.5 and 4.10, we deduce as above that $Y = \widehat{Y}$ almost surely, which implies $g(Z_1) = \nabla\phi^*(Z_1)$. Let us now prove the second assertion. Given a martingale $(X_t, Z_t)_{t\in[0,1]}$ in $\mathcal{M}(\preceq_\mu, Q_\Gamma)$ maximizing $\mathbb{E}[\langle X_1, Z_1\rangle]$, it follows from the preceding result that X_1 has law μ. If it was not true, we could construct a variable L of law μ as a function of X_1 and an independent uniform variable Y such that $X_1 = \mathbb{E}[L \mid \mathcal{F}_1^{X,Z}]$. We would have $\mathbb{E}[\langle L, Z_1\rangle] = \mathbb{E}[\langle X_1, Z_1\rangle] = V_\infty(\mu)$ and therefore $L = \nabla\phi^*(Z_1)$ which is absurd. It follows that $X_1 = \nabla\phi^*(Z_1)$ and for all $t \in [0, 1]$

$$X_t = \mathbb{E}[X_1 \mid \mathcal{F}_t^{X,Z}] = \mathbb{E}[\nabla\phi^*(Z_1) \mid \mathcal{F}_t^{X,Z}].$$

But from the second part of Lemma 4.9, we know that $\mathbb{E}[\langle X_t, Z_t\rangle] = U(t, [\![X_t]\!])$. Using then Proposition 4.4 and the Lemmas 4.5, 4.8 and 4.12, we deduce that $X_t = \nabla\psi_t(Z_t)$ almost surely for some function $\psi_t \in Conv(\mathbb{R}^d)$. It follows that X_t is Z_t-measurable so that

$$X_t = \mathbb{E}[X_1 \mid Z_t] = \mathbb{E}[\nabla\phi^*(Z_1) \mid Z_t].$$

Let v denote a measurable selection of $\partial\phi^*$ (hence bounded). We have by definition

$$\forall z, h \in \mathbb{R}^d,\ \phi^*(z + h) \geq \phi^*(z) + \langle v(z), h\rangle.$$

Replacing z by Z_1 and taking conditional expectations, we obtain

$$\begin{aligned} u(t, Z_t + h) &\geq \mathbb{E}[\phi^*(Z_t + h + (Z_1 - Z_t)) \mid Z_t] \\ &\geq \mathbb{E}[\phi^*(Z_t + (Z_1 - Z_t)) \mid Z_t] + \langle \mathbb{E}[v(Z_t + (Z_1 - Z_t)) \mid Z_t], h\rangle. \end{aligned}$$

The process $(Z_t)_{t\in[0,1]}$ being optimal, the standard dynamic programming principle of stochastic control implies $u(t, Z_t) = \mathbb{E}[\phi^*(Z_t + (Z_1 - Z_t)) \mid Z_t]$. Moreover, since $X_1 = v(Z_1)$, $X_t = \mathbb{E}[v(Z_t + (Z_1 - Z_t)) \mid Z_t]$ and it follows that

$$u(t, Z_t + h) \geq u(t, Z_t) + \langle X_t, h\rangle.$$

[7]The probability whose density is the normalized indicator function of the set with respect to the Lebesgue's measure on the affine subspace it generates.

We conclude that $X_t \in \partial u(t, Z_t)$ since the above inequality holds almost surely for a countable dense subset of h in $\mathbb{R}^d$. The end of the proof is similar to the proof of the first assertion. □

We can now state our main Verification Theorem.

Theorem 4.13. *Under the same hypotheses as Proposition 4.12 and if the solution u is C^1 with respect to the space variable, then $\mathcal{P}_\infty(\mu)$ is the set of all laws of processes*

$$(X_t)_{t\in[0,1]} = (\nabla u(t, Z_t))_{t\in[0,1]},$$

where the law of the process Z runs through the set of maximizers of $V^*_\infty(\phi)$ such that $[\![\nabla\phi^*(Z_1)]\!] = \mu$.

Proof. For any (law of) martingale (X, Z) in $\mathcal{M}(\preceq_\mu, Q_\Gamma)$ maximizing H, we have from Proposition 4.12 with probability 1,

$$\forall t \in [0, 1],\ t \text{ rational}, \quad X_t = \nabla u(t, Z_t).$$

The process in the right-hand side has continuous trajectories and X has càdlàg trajectories so that the equality can be extended to all $t \in [0, 1]$. The results follows then from Proposition 4.12. □

Let us finally prove the result announced in the introduction.

Proof of Theorem 1.5. In view of the previous results, we only need to prove that Z is a maximizer of $V^*_\infty(\phi)$ if and only if property (7) is true. But this follows directly from Ito's formula since u is assumed to be $C^{1,2}$. □

5. Appendix

5.1. **Auxiliary results.** We present in this section results about Optimal Transportation and Wasserstein distances. This material is well-known and can be found in [24] or [1].

Lemma 5.1. *Let E,E' be two separable metric spaces and A,A' two tight (resp. closed, convex) subsets of $\Delta(E)$ and $\Delta(E')$. Then the set $\mathcal{P}(A, A')$ is itself tight (resp. closed, convex).*

The Wasserstein distances. The Wasserstein distance of order p is defined on the set $\Delta^p(\mathbb{R}^d)$ of probabilities with finite moment of order p by

$$d_{W_p}(\mu,\nu) \triangleq \min_{\pi\in\mathcal{P}(\mu,\nu)} \Big(\int |y-x|^p d\pi(x,y)\Big)^{\frac{1}{p}} = min\{\|X-Y\|_{L^p} \mid X\sim\mu\ ,\ Y\sim\nu\}.$$

The metric space $(\Delta^p(\mathbb{R}^d), d_{W_p})$ is Polish. Convergence for d_{W_p} is equivalent to classic weak convergence together with the convergence of the moments of order p. The sets of probabilities with uniformly integrable moments of order p are relatively compact. Moreover, we have the following useful Lemma.

Lemma 5.2. *For any continuous function f and $K > 0$ such that $\mid f(x) \mid \leq K(1+ \mid x \mid^p)$, the application*

$$\Delta^p(\mathbb{R}^d) \to \mathbb{R} : \pi \to \int f(x)d\pi(x) \text{ is } d_{W_p} \text{ continuous.}$$

Lemma 5.3. *(Lemma 5.2.4 in [1]) Let $E = E' = \mathbb{R}^d$ and $\pi_n \in \Delta(E \times E')$ be a weakly converging sequence with limit π such that*

$$\sup_n \int |x|^p + |y|^q d\pi_n(x, y) < \infty \quad \text{for some } \ p, q \in (1, \infty) \text{ such that } \frac{1}{p} + \frac{1}{q} = 1.$$

If the sequence of marginals μ_n on E has uniformly integrable moments or order p (resp. ν_n on E' has uniformly integrable moments of order q) then

$$\int \langle x, y\rangle d\pi_n(x, y) \underset{n\to\infty}{\longrightarrow} \int \langle x, y\rangle d\pi(x, y).$$

Maximal covariance functions. These functions are also optimal transport value functions, related to the square Wasserstein distance. Precisely , the maximal covariance between two probabilities on $\mathbb{R}^d$ is defined by

$$C : \Delta^2 \times \Delta^2 \longrightarrow \mathbb{R} : (\mu, \nu) \longrightarrow \max_{\pi\in P(\mu,\nu)} \int \langle x, y\rangle d\pi(x, y).$$

We have then the straightforward relation

$$\forall \mu, \nu \in \Delta^2, \qquad d^2_{W_2}(\mu, \nu) = \|\mu\|_2^2 + \|\nu\|_2^2 - 2C(\mu, \nu).$$

Theorem 5.4. *For all $\mu, \nu \in \Delta^2$, we have the following equalities*

$$\begin{aligned}\max_{\pi\in\mathcal{P}(\mu,\nu)} \int \langle x, y\rangle d\pi(x, y) &= \inf_{(\phi-\frac{1}{2}|.|^2, \psi-\frac{1}{2}|.|^2)\in C_b(\mathbb{R}^d)^2; \phi+\psi\geqslant\langle .,.\rangle} \Big(\int \phi d\mu + \int \psi d\nu\Big)\\ &= \min_{\phi\in Conv(\mathbb{R}^d)} \int \phi d\mu + \int \phi^* d\nu,\end{aligned}$$

where $\phi + \psi \geqslant \langle ., .\rangle$ means $\phi(x) + \psi(y) \geqslant \langle x, y\rangle$ for all $x, y \in \mathbb{R}^d$ and $C_b(\mathbb{R}^d)$ denotes the set of real-valued bounded continuous functions on $\mathbb{R}^d$.

Let us also mention the following characterization where $\partial\phi$ denotes the subdifferential

Theorem 5.5. *For all $\mu, \nu \in \Delta^2$, we have the following equivalence*

$$\pi^* \in \underset{\pi\in\mathcal{P}(\mu,\nu)}{argmax} \int \langle x, y\rangle d\pi(x,y) \Longleftrightarrow \exists \phi \in Conv(\mathbb{R}^d), \quad y \in \partial\phi(x) \quad \pi^* \text{ -almost surely.}$$

Convex order.

Definition 5.6. *The convex order (also called Choquet or Blackwell order) is defined on Δ^2 by*

$$\mu_1 \preceq \mu_2 \Leftrightarrow \forall f \in Conv(\mathbb{R}^d), \int f d\mu_1 \leq \int f d\mu_2. \tag{33}$$

Moreover, $\mu_1 \preceq \mu_2 \in \Delta^2$ iff there exists a martingale X_1, X_2 such that $[\![X_i]\!] = \mu_i$ for $i = 1, 2$ (see [4]).

Let us now list some useful properties

Lemma 5.7. *The set $\{\nu \in \Delta^2 : \nu \preceq \mu\}$ is d_{W_2}-compact (hence weakly compact). The function $\mu \to C(\mu, \nu)$ is nondecreasing for the convex order, strictly if ν is absolutely continuous with respect to the Lebesgue's measure.*

Proof. For the first assertion, uniform integrability of the second order moment follows from the Jensen inequality and the martingale characterization of the convex order. Closedness follows from the convex representation (33) since the map $\mu \to \int f d\mu$ is lower continuous for any $f \in Conv(\mathbb{R}^d)$. For the second assertion, let ν be absolutely continuous and $\mu_1 \preceq \mu_2 \in \Delta^2$. Using Theorem 5.5, we have

$$C(\mu_i, \nu) = \min_{\phi\in Conv(\mathbb{R}^d)} \int \phi d\mu_i + \int \phi^* d\nu.$$

Let ϕ_2 be optimal in the above minimization problem for μ_2. If ϕ_2 is also optimal for μ_1, then Theorem 5.5 implies $\nabla\phi_2 \sharp \nu = \mu_1 = \mu_2$ since ν is absolutely continuous. Therefore, if $\mu_1 \neq \mu_2$, ϕ_2 is not optimal for μ_1, and we deduce from (33) that

$$C(\mu_1, \nu) < \int \phi d\mu_1 + \int \phi^* d\nu \leq \int \phi d\mu_2 + \int \phi^* d\nu = C(\mu_2, \nu).$$

□

Conditional laws. The following Theorem is well-known and allows to construct variables with prescribed conditional laws.

Lemma 5.8. *(Blackwell-Dubins [5])*
Let E be a Polish space with $\Delta(E)$ the set of Borelian probabilities on E, and $([0,1], \mathcal{B}([0,1]), \lambda)$ the unit interval equipped with Lebesgue's measure. There exists a measurable mapping

$$\Phi : [0,1] \times \Delta(E) \longrightarrow E$$

such that for all $\mu \in \Delta(E)$, the law of $\Phi(U, \mu)$ is μ where U is the canonical element in $[0,1]$.

In the proofs of Lemma 2.10 and Proposition 3.5, we use indirectly this result together with the Disintegration Theorem. Precisely :

Lemma 5.9. *Let X, Y be E-valued random variables defined on the same probability space, U an uniform random variable independent of (X, Y) and f a measurable mapping from E to $\Delta(E^2)$. Let $f_1(x)$ be the marginal law of $f(x)$ on the first coordinate. If $f_1(X)$ is a version of the conditional law of Y given X, then there exists a random variable $Z = \varphi(X, Y, U)$ such that $f(X)$ is a version of the conditional law of (Y, Z) given X.*

Proof. One can define using Theorem 5.8 a variable $(\tilde{Y}, \tilde{Z}) = \Phi(U, f(X))$ having the property that $f_1(X)$ is a version of the conditional law of $\tilde{Y}$ given X. Let $g(X, \tilde{Y})$ be a version of the conditional law of $\tilde{Z}$ given $(X, \tilde{Y})$, it follows easily that $Z = \Phi(U, g(X, Y))$ fulfills the required properties. □

Jensen Inequality. Let $p \in [1, \infty)$ and $\Delta^p(\mathbb{R}^d)$ the set of probabilities with finite moment of order p. The vector space M^p of finite signed borel measures μ on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} |x|^p d|\mu| < \infty$ is endowed with initial topology generated by the set $C_p(\mathbb{R}^d)$ of continuous functions on $\mathbb{R}^d$ with at most polynomial growth of order p. Recall that the induced topology on Δ^p is metrizable by the Wasserstein distance d_{W_p}.

Lemma 5.10. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ a probability space, $\mathcal{G} \subset \mathcal{F}$ two sub σ-algebra of $\mathcal{A}$, and f a concave upper semi-continuous mapping from Δ^p to $\mathbb{R}$ which is bounded by $C(1 + d_{W_p}(\delta_0, .))$. Then, for all $\mathbb{R}^d$-valued random variable X with finite moment of order p*

$$f([\![X]\!]) \geq \mathbb{E}[f([\![X \mid \mathcal{F}]\!])] \text{ and } f([\![X \mid \mathcal{G}]\!]) \geq \mathbb{E}[f([\![X \mid \mathcal{F}]\!]) \mid \mathcal{G}] \text{ almost surely.}$$

Proof. Note that all the expectations in the proof are well-defined using the bound on f and the integrability condition on X. Since X has a finite moment of order p, we can assume that the random variable $[\![X \mid \mathcal{F}]\!]$ is Δ^p-valued. Let Φ denote its distribution (in $\Delta(\Delta^p(\mathbb{R}^d))$). f being concave and upper semi-continuous, it is sufficient to prove that $\mu = [\![X]\!]$ is the barycenter of Φ. But, for all $h \in C_p(\mathbb{R}^d)$, it follows from the properties of the conditional expectation that

$$\int \langle h, \nu \rangle d\Phi(\nu) = \mathbb{E}[\mathbb{E}[h(X) \mid \mathcal{F}]] = \mathbb{E}[h(X)] = \langle h, \mu \rangle,$$

which proves the first result. The second assertion follows by the same method. It is sufficient to prove that $[\![X \mid \mathcal{G}]\!]$ is almost surely the barycenter the $\Delta(\Delta^p(\mathbb{R}^d))$-valued $\mathcal{G}$-measurable random variable

$$\Psi = [\![[\![X \mid \mathcal{F}]\!] \mid \mathcal{G}]\!].$$

Applying the previous argument to a well-chosen countable subset $\mathcal{C}_0$ of $C_p(\mathbb{R}^d)$ and by using the definitions of conditional laws and conditional expectations, we have with probability one

$$\forall h \in \mathcal{C}_0, \quad \int \langle h, \nu \rangle d\Psi(\nu) = \mathbb{E}[\mathbb{E}[h(X) \mid \mathcal{F}]\mathcal{G}] = \mathbb{E}[h(X) \mid \mathcal{G}]. \tag{34}$$

Now $\mathcal{C}_0$ can be taken as the union of $x \to (1+|x|^p)$ and of a countable convergence determining subset of $C_b(\mathbb{R}^d)$ (see e.g. [1] p106-107). The property (34) can therefore be extended to all $h \in C_p(\mathbb{R}^d)$. The proof is complete now. □

The two following Lemmas show an useful approximation result for processes on a fixed sequence of partitions for which we didn't find references.

Lemma 5.11. *Let c be a measurable and adapted $\mathbb{R}^d$-valued process defined on some filtered probability space such that*

$$\mathbb{E}[\int_0^1 |c_s|^2 ds] < \infty.$$

Then

i) $\lim_{h\to 0} \mathbb{E}[\int_0^1 |c_t - c_{t-h}|^2 dt] = 0$ *with the convention $c_t = 0$ for $t < 0$.*

ii) *There exists a sequence $\delta_n \in [0,1]$ and a sequence of simple processes c^n such that*

$$lim_{n\to\infty} \mathbb{E}[\int_0^1 |c_t - c_t^n|^2 dt] = 0 \text{ and } \quad c_t^n = \sum_{k=0}^{n-1} c_{\frac{k-\delta_n}{n}} \mathbb{1}_{[\frac{k}{n}, \frac{k+1}{n})}(t).$$

Proof. We only sketch the proof for $d = 1$, the generalization is straightforward. Note that it is a slight modification of Lemma 4.4(d) p96 in [17] (see also [16] problem 2.5 p134). For i), since $c_t = 0$ for $t < 0$, if $c^N = c\mathbb{1}_{|c|\le N}$ then for all $h \ge 0$

$$\mathbb{E}[\int_0^1 (c_{t-h}^N - c_{t-h})^2 dt] \le \mathbb{E}[\int_0^1 (c_t^N - c_t)^2 dt] \underset{N\to+\infty}{\longrightarrow} 0.$$

This implies i) since the above quantity is bounded by ε for sufficiently large N,

$$\begin{aligned}\mathbb{E}[\int_0^1 (c_t - c_{t-h})^2 dt]^{\frac{1}{2}} &\le \mathbb{E}[\int_0^1 (c_t - c_t^N)^2 dt]^{\frac{1}{2}} + \mathbb{E}[\int_0^1 (c_t^N - c_{t-h}^N)^2 dt]^{\frac{1}{2}} + \mathbb{E}[\int_0^1 (c_{t-h}^N - c_{t-h})^2 dt]^{\frac{1}{2}} \\ &\le 2\varepsilon + \mathbb{E}[\int_0^1 (c_t^N - c_{t-h}^N)^2 dt]^{\frac{1}{2}},\end{aligned}$$

and the last term goes to zero with h (see [16] p134). For the second point, define

$$c_t^{(n,\delta)} = \sum_{k=0}^{n-1} c_{\frac{k-\delta}{n}} \mathbb{1}_{[\frac{k}{n}, \frac{k+1}{n})}(t) \quad \text{and} \quad g_n(\delta) = \mathbb{E}[\int_0^1 (c_t - c_t^{n,\delta})^2 dt].$$

Fubini Theorem implies

$$\int_0^1 g_n(\delta) d\delta = \mathbb{E}[\int_0^1 \int_0^1 (c_t - c_t^{(n,\delta)})^2 d\delta dt] = \mathbb{E}[\sum_{k=0}^{n-1} \int_{k/n}^{(k+1)/n} \int_0^1 (c_t - c_t^{(n,\delta)})^2 d\delta dt].$$

For $t \in [\frac{k}{n}, \frac{k+1}{n})$, we have $c_t^{(n,\delta)} = c_{\frac{k-\delta}{n}}$ and by change of variable with $h = t - \frac{k-\delta}{n}$

$$\int_0^1 (c_t - c_t^{(n,\delta)})^2 d\delta = n \int_{t-k/n}^{t-(k-1)/n} (c_t - c_{t-h})^2 dh \le n \int_0^{2/n} (c_t - c_{t-h})^2 dh.$$

We obtain

$$\int_0^1 g_n(\delta) d\delta \le \mathbb{E}[\sum_{k=0}^{n-1} \int_{k/n}^{(k+1)/n} n \int_0^{2/n} (c_t - c_{t-h})^2 dh dt] \le 2 \max_{0\le h\le 2/n} \mathbb{E}[\int_0^1 (c_t - c_{t-h})^2 dt].$$

Finally $\int_0^1 g_n(\delta)d\delta$ converges to zero and we can choose δ_n such that

$$g_n(\delta_n) \leq \int_0^1 g_n(\delta)d\delta,$$

and this concludes the proof. □

Lemma 5.12. *Consider the canonical space $C([0,1],\mathbb{R}^d)$ endowed with the standard d dimensional Wiener measure $\mathbb{P}_0$. Let $(B_t)_{t\in[0,1]}$ denote the canonical process, $\mathcal{F}^B$ its natural filtration and $\mathcal{H}_G$ be the set of $\mathbb{M}_d$-valued $\mathcal{F}^B$-progressively measurable processes ρ such that $\rho\rho^T \in G$. Define $\widetilde{Q}_G(1)$ as the set of laws of variables $\int_0^1 \rho_s dB_s$ with $\rho \in \mathcal{H}_G$. Then $\widetilde{Q}_G(1)$ is dense in $Q_\Gamma(1)$*

Proof. Using Caratheodory's Theorem together with a measurable selection result, we can parameterize points in Γ as follows

$$\forall Q \in \Gamma, \quad Q = \textstyle\sum_{i=1}^m \lambda_i(Q)P_i(Q), \text{ with } m \triangleq \frac{d(d+1)}{2}+1$$

where the λ_i form a convex combination and $P_i \in G$, all these functions being measurable. Let Z be the canonical process defined on the canonical space endowed with a law in Q_Γ. Then there exists on an extended filtered probability space $(\Omega,\mathcal{A},(\mathcal{F}_t)_{t\in[0,1]},\mathbb{P})$ a d-dimensional Brownian motion W and an $\mathcal{F}$-progressively measurable process q_s such that $Z_t = \int_0^t q_s dW_s$ (see e.g. [16] Theorem 3.4.2) such that with probability 1, $q_s q_s^T \in \Gamma$. Let q^n be a sequence of simple processes (given by Lemma 5.11)

$$q_s^n = \textstyle\sum_{k=1}^n b_k^n \mathbb{1}_{[\frac{(k-1)}{n},\frac{k}{n})}(s),$$

approximating q in the Hilbert space $L^2(\Omega\times[0,1],d\mathbb{P}\otimes dt)$, and such that b_1^n is deterministic, b_k^n is $\mathcal{F}_{(k-1)/n}$-measurable and $b_k^n(b_k^n)^T \in \Gamma$. Since the above L^2-convergence implies convergence in law at time 1, it is sufficient to prove that the law of $\int_0^1 q_s^n dW_s$ is in $\tilde{Q}_G(1)$. Let B be a d-dimensional Brownian motion defined on the canonical space, we have to construct a process $\tau \in \mathcal{H}_G$ such that the stochastic integral $\int_0^1 \tau_s dB_s$ has the same law as $\int_0^1 q_s^n dW_s$. Note that the law of $\int_0^1 q_s^n dW_s$ is determined by the law of the vector $(b_k^n \Delta_k^n W)_{k=1,..,n}$ where $\Delta_k^n W = W_{k/n} - W_{(k-1)/n}$. The conditional law of $n b_k^n \Delta_k^n W$ given $\mathcal{F}_{(k-1)/n}$ is a normal distribution with covariance matrix $c_k^n = b_k^n(b_k^n)^T \in \Gamma$. We will construct by induction the process τ and a sequence $(\widehat{c_k^n})_{k=1,..,n}$ such that $\widehat{c_k^n}$ is $\mathcal{F}^B_{(k-1)/n}$-measurable and $\widehat{c_1^n} = c_1^n$. Assume that the process τ on $[0,(k-1)/n)$ and the variables $(\widehat{c_i^n})_{i=1,..,k}$ are given and such that $(\int_0^{(k-1)/n}\tau_s dB_s, \widehat{c_k^n})$ has the same law as $(\int_0^{(k-1)/n} q_s^n dW_s, c_k^n)$. Define then τ on $[\frac{(k-1)}{n},\frac{k}{n})$ as the piecewise constant process equal to $\sqrt{P_i(\widehat{c_k^n})}$ on the interval

$$\left[\tfrac{1}{n}((k-1)+\textstyle\sum_{q=1}^{i-1}\lambda_q(\widehat{c_k^n}))\ ,\ \tfrac{1}{n}((k-1)+\sum_{q=1}^{i}\lambda_q(\widehat{c_k^n}))\right), \tag{35}$$

for $i = 1,..,m$. By construction, the conditional law of $n\int_{(k-1)/n}^{(k)/n}\tau_s dB_s$ given $\int_0^{(k-1)/n}\tau_s dB_s$ is a normal distribution with covariance matrix $\widehat{c_k^n}$ and using our assumption it implies that $\int_0^{k/n}\tau_s dB_s$ has the same law as $\int_0^{k/n} q_s^n dW_s$. Next we construct a variable $\widehat{c_{k+1}^n}$, $\mathcal{F}^B_{k/n}$ measurable, such that the pair $(\int_0^{k/n}\tau_s dB_s, \widehat{c_{k+1}^n})$ has the same law as $(\int_0^{k/n} q_s^n dW_s, c_{k+1}^n)$. To construct $\widehat{c_{k+1}^n}$ with the prescribed conditional law given $\int_0^{k/n}\tau_s dB_s$, it is sufficient to have a diffuse random variable $\mathcal{F}^B_{k/n}$-measurable and independent from $\int_0^{k/n}\tau_s dB_s$. We can construct such a variable as a stochastic integral $\int_{(k-1)/n}^{k/n}\nu_s dB_s^1$ where B^1 is the first coordinate of B. For example, define ν_s as the piecewise constant process taking alternatively the values 1 and -1 on the partition of $[(k-1)/n, k/n)$ obtained by dividing each element of the partition given in (35) into two intervals of equal size. Usual properties of the stochastic integral against a Brownian motion show that this variable has the required properties, and we conclude the proof by induction. □

The latter Lemma rely on very classical density results and is reproduced here for the sake of completeness.

Lemma 5.13. *Let $M \in S_+^d$, and $1 \leq p < 2 < q < +\infty$. Then the d_{W_p}-closure of the set*

$$\{\nu \in \Delta_0^2(\mathbb{R}^K) \mid cov(\nu) = M\ ,\ \|\nu\|_q < +\infty\} \quad is \quad \{\nu \in \Delta_0^2(\mathbb{R}^K) \mid cov(\nu) \leq M\}.$$

Proof. Let $\nu \in \Delta_0^2(\mathbb{R}^K)$ such that $cov(\nu) \leq M$. Let X be a random variable with distribution ν. Define

$$Y_n = X\mathbb{1}_{\{|X|\leq n\}} - \mathbb{E}[X\mathbb{1}_{\{|X|\leq n\}}].$$

Then we can check that $\|Y_n\|_{L^q} \leq 2n$, $cov(Y_n) \leq M$ and $\|X - Y_n\|_{L^p} \underset{n\to\infty}{\longrightarrow} 0$. Let $P = M - cov(X)$. In order to conclude the proof, it is sufficient to construct a sequence of variables $Z_n \in L_q$ independent of Y_n such that $cov(Z_n) = P$, and $\|Z_n\|_{L^p} \underset{n\to\infty}{\longrightarrow} 0$. We would then have

$$d_{W_p}(\nu, [\![Y_n + Z_n]\!]) \leq \|X - (Y_n + Z_n)\|_{L^p} \underset{n\to\infty}{\longrightarrow} 0,$$

and thus the conclusion. Let us now define such a sequence. Let $(U_1,..,U_d)$ be independent uniform random variables on $[0,1]$, independent of the variable X. For all $k\in\mathbb{N}^*$, define

$$f_k:[0,1]\to\mathbb{R}:x\to\begin{cases}0 & if\quad x\leq 1-1/k\\ k^2x-k^2+k & if\quad x>1-1/k\end{cases}.$$

Define then $Z_n=\sqrt{P}((f_n(U_i)^{1/2}sgn(U_i-\alpha_n))_{i=1,..,d})$ with $\alpha_n\in[0,1]$ chosen such that $\mathbb{E}[Z_n]=0$. This sequence has clearly the required properties. □

5.2. **Technical proofs.**

Proof of Lemma 3.7. Let us define

$$D_n=\sup_{\nu\in RC^n(q,C)}d^2_{W_2}(\nu,\mathcal{N}(0,I_d)).$$

Let $(S_i)_{i=1,..,n}$ be an i.i.d. sequence of law $\mu\in RC^1(q,C)$, and let S_i^j denotes the j-th coordinates. Note at first that it follows from the martingale property that $cov(n^{-\frac{1}{2}}\sum_{k=1}^n S_i)=I_d$. In the next inequality, c_q denotes the universal constant of Burkholder's square function inequality for discrete-time martingales (cf [6]):

$$\mathbb{E}[|\sum_{k=1}^n\frac{S_i^j}{\sqrt{n}}|^q]\leq\frac{c_q^q}{n^{q/2}}\mathbb{E}[\left(\sum_{k=0}^{n-1}(S_{k+1}^j-S_k^j)^2\right)^{q/2}]\leq\frac{n^{q/2-1}c_q^q}{n^{q/2}}\sum_{k=0}^{n-1}\mathbb{E}[|S_{k+1}^j-S_k^j|^q]\leq C^qc_q^q.$$

Therefore, moments of order q are uniformly bounded independently of n. Recall that convergence in law together with uniformly bounded moments of order $q>2$ imply d_{W_2}-convergence. Since any maximizing sequence ν_n for D_n fulfills the classical Lindeberg's condition of the the Central Limit Theorem (Theorem VII.5.2 in [15]) for row-wise independent triangular arrays (again, since laws in $RC^1(q,C)$ have bounded q-th order moments), we deduce that

$$D_n=d_{W_2}(\mu_n,\mathcal{N}(0,I_d))\underset{n\to\infty}{\longrightarrow}0.$$

Moreover, $RC^n(q,C)$ is d_{W_2} compact, and the last assertion follows directly from Proposition 7.33 in [3]. □

Proof of Lemma 3.8. Given $q>2$ and $C>0$ define $\Lambda=\{\nu\in\Delta_0^2:cov(\nu)\leq I_d\}$ and recall that

$$RC^1_{q,C}=\{\nu\in\Delta_0^2:cov(\nu)=I_d,\ \|\nu\|_q\leq C\}.$$

For $M\in\mathbb{M}_d$, we have obviously $M\sharp(\Lambda)=\{\nu\in\Delta_0^2:cov(\mu)\leq MM^T\}$ (recall that $M\sharp$ denotes the image probability by the linear map $x\to Mx$). Moreover, using Lemma 2.4, we have

$$r(MM^T)=\sup_{\mu\in M\sharp(\Lambda)}V(\mu)=\sup_{\nu\in\Lambda}V(M\sharp\nu).$$

Since Λ is W_p-compact, there exists a maximum $\nu^*\in\Lambda$ (depending on M). We deduce that

$$r(MM^T)-\sup_{\nu\in RC^1_{q,C}}V(M\sharp\nu)=V(M\sharp\nu^*)-\sup_{\nu\in RC^1_{q,C}}V(M\sharp\nu)\leq\gamma d_{W_p}(M\sharp\nu^*,M\sharp RC^1_{q,C})\leq\alpha\gamma d_{W_p}(\nu^*,RC^1_{q,C}),$$

where α is a constant such that $|Mx|\leq\alpha|M||x|$ for all $M\in\mathbb{M}_d,x\in\mathbb{R}^d$. We conclude using Lemma 5.13 which implies that

$$\sup_{\nu\in\Lambda}d_{W_p}(\nu,RC^1_{q,C})\underset{C\to+\infty}{\to}0.$$

The measurable selection exists from Proposition 7.33 in [3] using that

$$(M,\nu)\in\mathbb{M}_d\times RC^1_{q,C}\to V(M\sharp\nu)$$

is jointly continuous when $RC^1_{q,C}$ is endowed with the d_{W_2}-topology, and in particular compact. □

Proof of Lemma 3.9. Recall that $p\in[1,2)$. With the notation $\Delta X_{k+1}=X_{k+1}-X_k$, we have

$$\begin{aligned}\left|\Psi_n[V]((X_k,\mathcal{F}_k)_{k=1,..,n})-\Psi_n[V]((Y_k,\mathcal{F}_k)_{k=1,..,n})\right|&=\left|\mathbb{E}[\textstyle\sum_{k=0}^{n-1}V([\![\Delta X_{k+1}\mid\mathcal{F}_k]\!])-V([\![\Delta Y_{k+1}\mid\mathcal{F}_k]\!])\right|\\ &\leq\gamma\mathbb{E}\left[\textstyle\sum_{k=0}^{n-1}\mathbb{E}\left[|\Delta X_{k+1}-\Delta Y_{k+1}|^p\mid\mathcal{F}_k\right]^{\frac{1}{p}}\right]\leq\gamma\mathbb{E}\Big[\textstyle\sum_{k=0}^{n-1}\mathbb{E}\Big[\left|\Delta X_{k+1}-\Delta Y_{k+1}\right|^2\mid\mathcal{F}_k\Big]^{\frac{1}{2}}\Big].\end{aligned}$$

Due to Cauchy-Schwarz and Jensen's inequalities, we have

$$\begin{aligned}\mathbb{E}\Big[\textstyle\sum_{k=0}^{n-1}\mathbb{E}\Big[\left|\Delta X_{k+1}-\Delta Y_{k+1}\right|^2\mid\mathcal{F}_k\Big]^{\frac{1}{2}}\Big]&\leq\sqrt{n}\mathbb{E}[\sqrt{\textstyle\sum_{k=0}^{n-1}\mathbb{E}\big[\left|\Delta X_{k+1}-\Delta Y_{k+1}\right|^2\mid\mathcal{F}_k\big]}]\\ &\leq\sqrt{n}\sqrt{\textstyle\sum_{k=0}^{n-1}\mathbb{E}\big[\left|\Delta X_{k+1}-\Delta Y_{k+1}\right|^2\big]}=\sqrt{n}\sqrt{\mathbb{E}[|X_n-Y_n|^2]}.\end{aligned}$$

□

Acknowledgements : The author would like to thank his supervisor Bernard De Meyer who initiated this work and for many fruitful discussions and improvements. The author gratefully acknowledges the support of the Agence Nationale de la Recherche, under grant ANR JEUDY, ANR-10-BLAN 0112.

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Birkhauser, 2008.

[2] C. Berge. *Topological spaces*. Dover, 1997.

[3] D.P. Bertsekas and S.E. Shreve. *Stochastic optimal control: The discrete time case*. Academic Press, 1978.

[4] D. Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953.

[5] D. Blackwell and L.E. Dubins. An extension of Skorohod's almost sure representation theorem. *Proceedings of the American Mathematical Society*, 89(4):691–692, 1983.

[6] D.L. Burkholder. Distribution function inequalities for martingales. *The Annals of Probability*, 1(1):19–42, 1973.

[7] P. Cardaliaguet and C. Rainer. On a continuous-time game with incomplete information. *Math. Oper. Res.*, 34(4):769–794, 2009.

[8] J.A. Cuesta-Albertos, C. Matrán-Bea, and A. Tuero-Diaz. On lower bounds for the L 2-Wasserstein metric in a Hilbert space. *Journal of Theoretical Probability*, 9(2):263–283, 1996.

[9] B. De Meyer. Price dynamics on a stock market with asymmetric information. *Games Econom. Behav.*, 69(1):42–71, 2010.

[10] L. Denis, M. Hu, and S. Peng. Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. *Potential Analysis*, pages 1–23, 2009.

[11] W.H. Fleming and H.M. Soner. *Controlled Markov processes and viscosity solutions*. Springer, 2006.

[12] F. Gensbittel. Analyse asymptotique de jeux répétés à information incomplète. *Ph.D. thesis, manuscript, University Paris 1.*, 2010.

[13] F. Gensbittel. Extensions of the Cav(u) theorem for repeated games with incomplete information on one side. *preprint*, 2011.

[14] J. Jacod. On continuous conditional Gaussian martingales and stable convergence in law. *Séminaire de Probabilités XXXI*, pages 232–246, 1997.

[15] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[16] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. Springer, 1991.

[17] R.S. Lipster and A.N. Shiryaev. *Statistics of random Processes. I: General theory. Second edition.* Springer-Verlag, 2001.

[18] J.-F. Mertens and S. Zamir. The maximal variation of a bounded martingale. *Israel Journal of Mathematics*, 27:252–276, 1977.

[19] P.A. Meyer and W.A. Zheng. Tightness criteria for laws of semimartingales. *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, 20(4):353–372, 1984.

[20] S. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Processes and their Applications*, 118(12):2223–2253, 2008.

[21] M. Pratelli. An alternative proof of a theorem of Aldous concerning convergence in distribution for martingales. *Séminaire de Probabilités XXXIII*, pages 334–338, 1999.

[22] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958.

[23] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*. Springer, 1979.

[24] C. Villani. *Topics in optimal transportation*. Amer Mathematical Society, 2003.

TSE (GREMAQ, Université Toulouse 1 Capitole)
Bureau MF213, Manufacture Des Tabacs, 21, Allée de Brienne, 31000 Toulouse
E-mail address: fabien.gensbittel@univ-tlse1.fr