

# VISCOSITY SOLUTIONS IN NON-COMMUTATIVE VARIABLES

WILFRID GANGBO, DAVID JEKEL, KYEONGSIK NAM, AARON Z. PALMER

ABSTRACT. Motivated by parallels between mean field games and random matrix theory, we develop stochastic optimal control problems and viscosity solutions to Hamilton-Jacobi equations in the setting of non-commutative variables. Rather than real vectors, the inputs to the equation are tuples of self-adjoint operators from a tracial von Neumann algebra. The individual noise from mean field games is replaced by a free semi-circular Brownian motion, which describes the large- $n$  limit of Brownian motion on the space of self-adjoint matrices. We introduce a classical common noise from mean field games into the non-commutative setting as well, allowing the problems to combine both classical and non-commutative randomness.

## 1. INTRODUCTION

**1.1. Context and motivation.** We aim to develop free non-commutative analogs of mean fields games, stochastic control, and Hamilton–Jacobi–Bellman equations. Random matrix theory has a long history since its appearance in 1928, where empirical covariance matrices of measured data naturally form a random matrix ensemble and the eigenvalues play a crucial role in principal component analysis [84]. Wigner studied the spectral distributions of random matrices with independent entries (and especially the *Gaussian unitary ensemble* or GUE) motivated by the observed similarity of random spectral distributions to energy levels of atomic nuclei [83]. Mean field games, introduced by Caines–Malhamé–Huang [54] and Lasry–Lions [63], describe competitive interactions between a large number of agents through the continuum approximation using their bulk density; see also Gomes–Saúde [46] for a survey of mean field game models and Carmona–Delarue [15] for a more exhaustive coverage. The study of mean field games deploys and advances methods from statistical physics, and is motivated by applications to modeling social behavior, economics, and finance. An influential mathematical development of Cardaliaguet–Delarue–Lasry–Lions [12] was an infinite-dimensional PDE approach to the convergence of the  $n$ -player games to the mean field game limit (see also [40, 42]). The infinite-dimensional mean field game PDE has also been developed using analytic solutions by Gangbo–Mészáros–Mou–Zhang [41] and Bansil–Mészáros–Mou [4] (see also [56]).

Connections between mean field games and random matrix theory have been apparent for some time. For instance, recent studies in game theory were motivated by the empirical observations that the distinctive behavioral patterns of bus drivers in some cities relate to previously observed experiments in quantum physics. The dynamics of these bus patterns correspond to paths of eigenvalues for  $n \times n$  self-adjoint random matrices. In the same vein, the Dyson game [14] described a mean field games approach matrix Brownian motion where the positions of  $n$  players correspond to the eigenvalues of the matrix. Hence, the empirical distribution of the player positions represents the empirical spectral distribution of the matrix, i.e. the measure with mass  $1/n$  at each of the eigenvalues. The Nash–optimal trajectories of the game are given by Dyson Brownian motions, first introduced by Freeman Dyson [34]. The studies in [14] and [34] connect to a rich area of dynamical systems such as the theory of Calogero–Moser–Sutherland models which involve completely integrable Hamiltonians; see Menon [66]

for further discussion. The Dyson Brownian motion also has a natural Riemannian interpretation using motion by mean curvature which was discovered by Huang, Inauen, and Menon [53]. The connections with mean field games and more generally PDE theory led to a new approach for analyzing the spectrum of single-matrix models developed by Bertucci–Debah–Lasry–Lions [6].

We want to explore the connection between mean field games and random matrix theory in the setting of multimatrix models with several self-adjoint matrices that do not necessarily commute with each other. Unlike a single self-adjoint matrix, the empirical distribution of  $d$ -tuple of self-adjoint matrices cannot simply be described by the positions of  $n$  eigenvalues or any classical measure on  $\mathbb{R}^d$ ; it is fundamentally non-commutative, so that one is forced to deal directly with the non-commutative moments  $\text{tr}(X_{i_1} \dots X_{i_k})$  or to formulate some analog of smooth test functions for non-commuting variables. The large- $n$  limiting behavior of such multimatrix models is the domain of free probability theory, developed in large part by Voiculescu, where the role of probabilistic independence is played by free independence modeled on the behavior of free products of groups [73, 79]. The large- $n$  limits are modeled by tuples  $(X_1, \dots, X_d)$  of non-commuting random variables, which are operators in a von Neumann algebra, i.e. a non-commutative measure space.

The difference between the classical and non-commutative settings can be described in terms of the choice of observables and the symmetry group acting on the system as follows. In the classical case, the observables correspond to commuting (diagonal) matrices with the symmetry of the permutation group representing interchangeability. In the non-commutative case, the observables are self-adjoint matrices or operators (as in quantum mechanics), and the symmetry group is the unitary group representing invariance under a change of coordinate basis. In the von Neumann algebraic framework, the symmetries can be understood as trace-preserving automorphisms and, more generally, inclusions of von Neumann algebras. In our framework, this arises from the assumption that problem data is given by ‘tracial  $W^*$ -functions,’ along with structural assumptions about the model noise, which we discuss later.

Despite the inherent non-commutative nature of multimatrix models that makes them less amenable to classical techniques, there is ample motivation to relate multimatrix models with mean field games, especially the role of non-commutative stochastic analysis and stochastic optimization problems in free probability. Biane–Capitaine–Guionnet [9] used a stochastic optimization problem to study large deviations theory for matrix Brownian motion, and hence also free entropy. This analysis was carried further by Dabrowski in the study of entropy for free Gibbs laws [29], and the related Hamilton–Jacobi–Bellman equation appeared in [57]. These works occur in the larger context of free stochastic analysis [9, 47, 30, 61], free entropy and information theory [75, 76, 78, 80, 51, 50], and free optimal transport [10, 48, 59, 38, 58] (and the bibliography listed here is far from exhaustive). We also briefly note the emerging research in matrix-valued optimal transport [17, 16] as well as entropic optimal transport for density operators in the quantum setting, e.g. [13, 31, 37]; although this is distinct from the free probability setting studied here, the similarity of approaches in [33, 38] suggests potential connections.

This work will formulate mean-field-like stochastic control problems in a free probability setting motivated by the large- $n$  behavior of random matrix theory, and study the resulting Hamilton–Jacobi equations in tracial von Neumann algebras; these can be understood as a non-local PDE in the single matrix case and the multimatrix case they are evolution equations in Hilbert space. In a subsequent paper, we plan to show that under appropriate convexity assumptions our free model accurately describes the large- $n$  limiting behavior of the associated random matrix models.

Our model will incorporate, for the first time in the free probability setting, both *common noise* and *free individual noise*. To motivate this feature, we consider the analogous models of mean field games that incorporate common noise and classical individual noise. The classical individual noise is modeled by independent Brownian motions, which lead to a second order Laplacian term in the evolution of the player densities. The individual noise captures variation and uncertainty in the individual player dynamics. The common noise is modeled by a single Brownian motion that is shared in the dynamics of all the players, which results in a stochastic term in the partial differential equation for the player densities. The common noise captures environmental uncertainty and exogenous variations in the system. Both these sources of noise are important for financial applications as well as population modeling. The individual noise is also motivated from physics by the need to capture unstructured interactions between particles in the nucleus of an atom. Whether sources of non-commutative randomness that arise in games or other systems would also be captured by similar freely independent noise is an intriguing aspect of the problem for further study into applications. For background on the role of individual and common noise in mean field games, we refer the reader to Carmona-Delarue [15]; Chapter 1.4 for economic applications and Chapter 2.1 for a typical mathematical set-up.

The study of Hamilton–Jacobi equations on infinite dimensional spaces started several decades ago, for instance in Hilbert space settings [22, 25, 26]. This was followed by studies on the Wasserstein space, whose differential structure has been proven to coincide with that of a quotient Hilbert space [45]. The study in the current manuscript is the non commutative analog of the recent developments which appeared in [1, 5, 18, 39, 43, 44], dealing with either no noise, the common noise or the individual noise. In this study, we trade the Wasserstein space with the set of non commutative laws.

**1.2. Control problems in the multimatrix and von Neumann algebraic settings.** For concreteness, we first describe the stochastic control problems we are interested in the setting of  $n \times n$  matrix tuples before moving on to the general von Neumann algebraic setting. The state space will be  $M_n(\mathbb{C})_{\text{sa}}^d$ , the set of  $d$ -tuples of self-adjoint  $n \times n$  complex matrices. In the mean field games analogy, due to the non-commutative nature of the problem, one cannot isolate individual players (just as a quantum graph does not have a distinct set of vertices), but conceptually we may think of the multimatrix as describing “ $M_n(\mathbb{C})$ -many points” in  $\mathbb{R}^d$ . We are interested in the large  $n$  behavior of stochastic control problems on  $M_n(\mathbb{C})_{\text{sa}}^d$ , i.e. we take  $n \rightarrow \infty$  with  $d$  fixed.

The free individual noise in the multimatrix setting comes from a Brownian motion  $\widehat{W}_t^n$  on the space of  $d$ -tuples of self-adjoint matrices. For each  $j = 1, \dots, d$ , the  $\widehat{W}_t^{n,j}$  is a random Gaussian self-adjoint matrix whose entries are independent up to Hermitian symmetry. The distribution of  $t^{-1/2}\widehat{W}_t^{n,j}$  is the *Gaussian unitary ensemble*, so called because its distribution is invariant under unitary conjugation. Hence, we refer to  $\widehat{W}_t^{n,j}$  and  $\widehat{W}_t^n$  as a *GUE( $n$ ) Brownian motion*. The large- $n$  limiting behavior of this process is described by a free Brownian motion (see e.g. [7, 9, 8]). The free Brownian motion is a process  $S_t = (S_t^1, \dots, S_t^d)$  of self-adjoint operators in a von Neumann algebra  $\mathcal{A}$  (a certain type of algebra of bounded operators on a Hilbert space) with a tracial state  $\tau : \mathcal{A} \rightarrow \mathbb{C}$  representing the expectation; these operators are not random in the classical sense but come from a non-commutative analogue of probability spaces. The spectral distribution of each  $t^{-1/2}S_t^j$  is given by Wigner’s semicircular measure  $(1/2\pi)\mathbf{1}_{[-2,2]}(x)\sqrt{4-x^2}dx$ , which is the almost sure limit of the empirical spectral distribution of the GUE matrix (see e.g. [2, §2]). The different coordinates and the different increments of the free Brownian motion are *freely independent*, a non-commutative form of independence

named for its relationship with free products of groups and operator algebras (see e.g. [82], [2, §5], and §2.2). Voiculescu's theory of asymptotic freeness shows that the joint moments  $\text{tr}_N(W_{t_1}^{n,j_1} \dots W_{t_k}^{n,j_k})$  converge almost surely to the deterministic limit  $\tau(S_{t_1}^{j_1} \dots S_{t_k}^{j_k})$  [73, 79]. The deterministic limiting behavior arises from high-dimensional concentration of measure (see [79, 49, 65]).

Meanwhile, the common noise in the multimatrix setting is still a single *classical* Brownian motion that is shared along the diagonal of each of the matrices. While the individual noise is described by a deterministic limit using free Brownian motion, the classical randomness of common noise persists in the large- $n$  limit. Hence, the equation that we formulate in the free probability setting incorporates both non-commutative randomness in the individual noise and classical randomness in the common noise; this combination itself is a new technical challenge that has not been studied much before.

The multimatrix stochastic control problems are formulated as follows. We fix two nonnegative parameters  $\beta_C$  and  $\beta_F$ , which we refer to respectively as the *common noise* and the *free individual noise* parameters. Fix a probability space  $(\Omega, \mathbb{P})$  and a standard one-dimensional Brownian motion  $(W_t^0)_{t \in [0, T]}$  on  $\Omega$ , which we embed in the matrix algebra by multiplying by the identity matrix  $\mathbf{1}_{M_n(\mathbb{C})}$ . We let  $(\widehat{W}_t^{n,j})_{t \in [0, T]}^{j=1, \dots, d}$  denote  $d$  many i.i.d. GUE( $n$ ) Brownian motions on  $\Omega$ , independent of  $(W_t^0)_{t \in [0, T]}$ . In order to feature the simplest example of the problems of interest, let us start with a time-dependent random control  $(\alpha_t^n)_{t \in [0, T]}$  taking values in  $M_n(\mathbb{C})_{\text{sa}}^d$ . Under appropriate conditions, there exists a unique random path  $t \in (t_0, T) \mapsto X_t^{n,j}[\alpha^n]$ , a solution to the stochastic differential equation

$$dX_t^{n,j}[\alpha^n] = \alpha_t^{n,j} dt + \beta_C \mathbf{1}_{M_n(\mathbb{C})} dW_t^0 + \beta_F d\widehat{W}_t^{n,j}, \quad j \in \{1, \dots, d\}, \quad (1.1)$$

where at the initial time  $(X_{t_0}^{n,j}[\alpha])_{j=1, \dots, d}$  is a prescribed deterministic  $d$ -tuple  $x_0^n \in M_n(\mathbb{C})_{\text{sa}}^d$ . Given a Lagrangian  $L_{M_n(\mathbb{C})}$  representing a running cost and a function  $g_{M_n(\mathbb{C})}$  representing a terminal value function, we search for an optimal control, in an admissible class  $\widehat{\mathcal{A}}_{M_n(\mathbb{C})}^{t_0, T}$ , for the problem

$$\widehat{V}_{M_n(\mathbb{C})}(t_0, x_0^n) = \inf_{\alpha \in \widehat{\mathcal{A}}_{M_n(\mathbb{C})}^{t_0, T}} \left\{ \mathbb{E} \left[ \int_{t_0}^T L_{M_n(\mathbb{C})}(X_t[\alpha], \alpha_t) dt + g_{M_n(\mathbb{C})}(X_T[\alpha]) \right] : X_{t_0}[\alpha] = x_0^n \right\}. \quad (1.2)$$

Our goal is to study the large- $n$  limit of the value function analogously to the mean-field limit where the number of players  $n$  tends to infinity, but in the multimatrix setting, we have to first discuss what we even mean by convergence. In the mean field setting, convergence is understood as convergence of the empirical measures averaged over the players to a limit measure. Similarly, for a single matrix, one studies convergence of the empirical spectral measure. But as mentioned before, the joint distributions of non-commuting random variables are not described by classical measures. Rather, the non-commuting random variables are elements  $(X_1, \dots, X_d)$  in a von Neumann algebra  $\mathcal{A}$  (a certain type of algebra of bounded operators on a Hilbert space) with a tracial state  $\tau : \mathcal{A} \rightarrow \mathbb{C}$  representing the expectation. The probability distribution of these operators is its *non-commutative law* described as the linear functional  $\lambda_X$  on non-commutative polynomials  $\mathbb{C}\langle x_1, \dots, x_d \rangle$  sending each test polynomial  $p$  to its expectation  $\tau(p(X_1, \dots, X_d))$  when evaluated on our given tuple of random variables. The space of non-commutative laws of  $d$ -tuples  $(X_1, \dots, X_d)$  with  $\sup_j \|X_j\|_\infty \leq R$  has a weak- $*$  topology that makes it into a compact space. Analogous to classical probability, one can think of the space of non-commutative laws as a quotient space consisting of equivalence classes of  $d$ -tuples of random variables, where for von Neumann algebras  $\mathcal{A}$  and  $\mathcal{B}$  along with tracial

states  $\tau_{\mathcal{A}}$  and  $\tau_{\mathcal{B}}$ , we say that  $X \in L^2(\mathcal{A})_{\text{sa}}^d$  and  $Y \in L^2(\mathcal{B})_{\text{sa}}^d$  with  $\|X\|_{\infty}, \|Y\|_{\infty} \leq R$  satisfy  $X \sim Y$  if  $\tau_{\mathcal{A}}(p(X)) = \tau_{\mathcal{B}}(p(Y))$  for all  $p \in \mathbb{C}\langle x_1, \dots, x_d \rangle$ . If  $X, Y \in M_n(\mathbb{C})_{\text{sa}}^d$ , then  $X \sim Y$  if and only if there is a unitary  $U$  such that  $UX_jU^* = Y_j$  for all  $j$  (although this is not exactly true for tuples in general von Neumann algebras).

Thus the question of convergence for the multimatrix control problems can be formulated as follows: Suppose that  $x_0^n$  is a  $d$ -tuple of self-adjoint  $n \times n$  matrices and the operator norms of  $x_0^n$  are bounded as  $n \rightarrow \infty$ . Suppose that the non-commutative laws of  $x_0^n$  converge in the weak-\* topology. Then do the value functions  $\hat{V}_{M_n(\mathbb{C})}(t_0, x_0^n)$  converge as  $n \rightarrow \infty$ , and if so, how do we describe the limit? This paper will formulate a candidate for the limit in the setting of free probability and tracial von Neumann algebras. We will set up the stochastic control problems, prove a dynamic programming principle, and define viscosity solutions. Because these developments alone are sufficiently involved, we will address convergence of the matrix models in the large- $n$  limit in a follow-up paper.

The infinite-dimensional analogue of the stochastic control problem (1.1) is as follows. Let  $\mathcal{A} = (A, \tau)$  be a tracial von Neumann algebra, i.e. a von Neumann algebra  $A$  with a chosen (faithful normal) tracial state  $\tau$ . Let  $L^2(\mathcal{A})$  be the completion of  $A$  with respect to  $\|x\|_{L^2(\mathcal{A})} = \tau(x^*x)^{1/2}$ . Assume that  $\mathcal{A}$  is large enough to contain a family of freely independent semi-circular Brownian motions  $S_t^j$ , for  $j = 1, \dots, d$ , that are freely independent from our chosen initial condition  $x_0 \in L^2(\mathcal{A})_{\text{sa}}^d$ . Moreover, the control  $(\alpha_t)_{t \in [0, T]}$  will be a process taking values in  $L^2(\mathcal{A})_{\text{sa}}^d$ , the subset of  $d$ -tuples of self-adjoint operators in  $L^2(\mathcal{A})$ , satisfying appropriate independence conditions (namely, the increments of the free Brownian motion  $S_t$  for  $t \geq t_0$  are freely independent of the values of  $\alpha_t$  for  $t \leq t_0$ , and the increments of the common Brownian motion  $W_t^0$  for  $t \geq t_0$  are probabilistically independent of  $\alpha_t$  for  $t \leq t_0$ ). Then let  $X_t = (X_t^j)_{j=1, \dots, d}$  be the solution to the stochastic differential equation

$$dX_t^j = \alpha_t^j dt + \beta_C \mathbf{1}_{\mathcal{A}} dW_t^0 + \beta_F dS_t^j, \quad j \in \{1, \dots, d\}, \quad (1.3)$$

with initial condition  $X_{t_0}[\alpha] = x_0 \in L^2(\mathcal{A})_{\text{sa}}^d$ . We emphasize that in contrast to the expression multiplied by  $\beta_F$  in (1.1) being stochastic, the analogous expression in (1.3) with the free Brownian motion is not random in the classical sense.

In the special case  $d = 1$ , we may understand solutions to (1.3) in terms of their law  $\mu_t \in \mathcal{P}_2(\mathbb{R})$ , where  $\mathcal{P}_2(\mathbb{R})$  denotes the set of probability measures having finite second moments. Equation (1.3) expresses the following stochastic non-linear non-local partial differential equation, supposing  $\alpha_t = a_t(X_t)$ ,

$$d\mu_t = -\partial_x(a_t \mu_t)dt - \partial_x(\beta_C \mu_t)dW_t^0 + \frac{\beta_C^2}{2}\partial_x^2 \mu_t dt + \frac{\beta_F^2}{2}\partial_x(h[\mu_t] \mu_t)dt,$$

where  $h[\mu](y) := \text{p.v.} \int_{\mathbb{R}} \frac{\mu(dx)}{y-x}$  denotes the Hilbert transform (see §B.4).

In light of the existing theory of viscosity solutions on Hilbert spaces, given a family  $(L_{\mathcal{A}})_{\mathcal{A}}$  of Lagrangians, a family  $(g_{\mathcal{A}})_{\mathcal{A}}$  of value functions, an admissible set of control policies  $\mathbb{A}_{\mathcal{A}}^{t_0, T}$  and  $x_0 \in L^2(\mathcal{A})_{\text{sa}}^d$ , it is tempting to set

$$\tilde{V}_{\mathcal{A}}(t_0, x_0) := \inf_{\alpha \in \mathbb{A}_{\mathcal{A}}^{t_0, T}} \left\{ \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{A}}(X_t[\alpha], \alpha_t) dt + g_{\mathcal{A}}(X_T[\alpha]) \right], X_{t_0}[\alpha] = x_0 \right\}. \quad (1.4)$$

The problem with this naïve formulation is that  $\tilde{V}_{\mathcal{A}}$  lacks the consistency properties necessary to make it well-defined on the quotient spaces of interest; it is not invariant under embedding the non-commutative probability space  $\mathcal{A}$  into a larger one  $\mathcal{B}$  even if we assume such invariance

for the functions  $L$  and  $g$ . Specifically, assume that the families of Lagrangians  $(L_{\mathcal{A}})_{\mathcal{A}}$  and value functions  $(g_{\mathcal{A}})_{\mathcal{A}}$  satisfy the consistency property that

$$g_{\mathcal{B}}(\iota x_0) = g_{\mathcal{A}}(x_0), \quad \text{for all } x_0 \in L^2(\mathcal{A})_{\text{sa}}^d$$

for any  $W^*$ -algebra  $\mathcal{B}$  and a  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ . In general,

$$\tilde{V}_{\mathcal{B}}(t_0, \iota(x_0)) \neq \tilde{V}_{\mathcal{A}}(t_0, x_0).$$

This behavior should be contrasted with the classical setting, in which any probability space that supports a Brownian motion independent of the initial condition would yield the same answer for stochastic optimization problems. The issue of different possible behaviors in different larger tracial von Neumann algebras already arose in [38] in the study of Monge–Kantorovich duality for the non-commutative  $L^2$ -Wasserstein metric of Biane and Voiculescu [10]. To obtain invariance for our value function, we will take the infimum over all possible embeddings into larger tracial von Neumann algebras, and hence the relevant functions to study are

$$\bar{V}_{\mathcal{A}}(t_0, x_0) := \inf_{(\mathcal{B}, \iota)} \left\{ \tilde{V}_{\mathcal{B}}(t_0, \iota x_0) : \iota : \mathcal{A} \rightarrow \mathcal{B} \text{ is a tracial } W^* \text{-embedding} \right\}. \quad (1.5)$$

In order to show that this function has the desired invariance property, we use a novel joint embedding lemma that allows us to embed two given von Neumann algebras with non-commutative filtrations into a larger one while identifying the two copies of the free Brownian motion and initial condition (see §3.4).

We aim to show that the value function  $\bar{V}$  is a viscosity solution for a certain Hamilton–Jacobi equation, which also requires formulating the definition of viscosity solution for the non-commutative setting. Abstractly, the Hamilton–Jacobi equation appears on the space of non-commutative laws as

$$\begin{aligned} -\partial_t V(t, \lambda) + H(\lambda, -\partial V(t, \lambda)) - \frac{\beta_C^2}{2} \Delta V(t, \lambda) - \frac{\beta_F^2}{2} \Theta V(t, \lambda) &= 0, \\ V(T, \lambda) &= g(\lambda). \end{aligned}$$

In this equation,  $H = (H_{\mathcal{A}})_{\mathcal{A}}$  is a tracial  $W^*$ -function of the state and generalized momentum, precisely defined in (3.14) later. The operator  $\Delta$  is a common noise Laplacian, which corresponds to a second-order differential operator, for example, defined on the functions of the Hilbert spaces  $C^2(L^2(\mathcal{A})_{\text{sa}}^d)$ . Finally, the operator  $\Theta$  is a free individual noise Laplacian, which is defined by means of introducing a freely independent free Brownian motion, and hence also requires enlarging the tracial von Neumann algebra and relies on invariance under embeddings.

We close with some brief comments on convergence of the multimatrix value functions to the free limit, which will be studied in a follow-up paper. First, it is essential to assume stronger continuity properties for the functions  $L$  and  $g$  than simply their being Lipschitz tracial  $W^*$ -functions. The reason is that, when  $d > 1$ , the non-commutative Wasserstein distance studied by Biane and Voiculescu [10] gives a much stronger topology than the weak- $*$  topology on the space of non-commutative laws of variables bounded by  $R$ , and in fact it is impossible for the non-commutative laws of multimatrices to converge in Wasserstein distance unless the limiting von Neumann algebra is amenable as shown in [38, §5.4–5.5]. Hence, for convergence we would assume weak- $*$  continuity for  $L$  and  $g$  in addition to the hypotheses in this paper. Secondly, we must also deal with the fact that enlarging the von Neumann algebra could alter the value of the infimum, hence the large- $n$  limit of the infima in the matrix models can in general be greater than the infimum given in  $\bar{V}$ . We will remove this issue by making a stronger convexity assumption on  $L$  and  $g$  called  $E$ -convexity [38]; this is analogous to the way that some convergence results in mean field games require displacement convexity for Hamilton–Jacobi equation [40] or Lasry–Lions monotonicity condition [12].

**1.3. Overview of results and organization.** We approach the problem of understanding Hamilton-Jacobi equations on non-commutative spaces by first defining optimal control problems in von Neumann algebras. We start in §2 with the necessary background on tracial  $W^*$  (von Neumann) algebras, amalgamated free products, and non-commutative laws (which play the role of probability distributions in this theory). In §3, we describe the setup of these stochastic optimization problems. After introducing motivating examples in §3.1, we spell out the general assumptions in §3.2, and then develop the properties of the Hamiltonian in §3.3 and the value function in §3.4.

Our framework relies on the notion of a *tracial  $W^*$ -function*, that is, a real-valued function that is defined in a consistent way on all tracial  $W^*$ -algebras, so that the output only depends on the non-commutative law of the input, similar to how probabilistic phenomena are independent of the choice of probability space. We also use the related notion of tracial  $W^*$ -vector fields. Like tracial  $W^*$ -function, these are defined in terms of consistency with respect to tracial  $W^*$ -embeddings of a tracial  $W^*$ -algebra into a larger one. We show that although the naïve value function  $\tilde{V}_A$  is not necessarily a tracial  $W^*$ -function,  $\bar{V}_A$  will be under our assumptions. We assume Lipschitz conditions for the Lagrangian and terminal cost functions, and establish that the resulting control-theoretic Hamiltonian and value function are tracial  $W^*$ -functions and satisfy a similar Lipschitz estimate. While our assumptions are far from the greatest possible generality, we are guided by a handful of diverse examples, including an Eikonal equation, Linear-Quadratic-Gaussian framework, and a controlled von Neumann equation.

In addition, we study  *$E$ -convexity*, a convexity condition from [38] that also takes into account tracial  $W^*$ -embeddings  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  by demanding that if  $E : \mathcal{B} \rightarrow \mathcal{A}$  is the conditional expectation adjoint to  $\iota$ , then  $f^{\mathcal{A}} \circ E \leq f^{\mathcal{B}}$ . We show that  $E$ -convexity for the terminal and running cost functions implies that  $\bar{V}_A = \tilde{V}_A$ , thus in this case enlarging the von Neumann algebra does not produce any smaller infimum (see Lemma 3.10, and we also obtain  $E$ -convexity for the value function (see Lemma 3.11).

After the development of free stochastic control problems in §3, we move on to the concept of viscosity solutions in §4. We consider two notions of viscosity solution. The first notion in §4.1 is defined on the space of non-commutative laws and uses tracial  $W^*$ -functions as test functions. We show that the value function is a sub- and super-solution in our definition by showing it satisfies the sub- and super-dynamic programming principle in Propositions 4.7 and 4.8. In our framework, the subsolution property can already be shown for the functions  $\tilde{V}_A$  that work with stochastic processes in a fixed tracial  $W^*$ -algebra  $\mathcal{A}$ , but the supersolution property requires consider all possible embeddings into a larger tracial  $W^*$ -algebra. This slight asymmetry is a new subtlety that arises in the non-commutative setting due to different possible behaviors of embeddings into larger tracial  $W^*$ -algebras.

The second notion of viscosity solution that we consider in §4.2 is based on the standard notion of viscosity solution in Hilbert spaces applied to the non-commutative  $L^2$  space  $L^2(\mathcal{A})_{sa}^d$  associated to a tracial  $W^*$ -algebra  $\mathcal{A}$ . In this setting, we do not handle the individual noise, only the common noise. However, in this setting, we can show a comparison principle for viscosity solutions that implies uniqueness (Theorem 4.13) based on the comparison principle in the existing Hilbert-space theory.

Finally, Section 5 discusses our main examples in further depth, including an Eikonal equation, Linear-Quadratic-Gaussian framework, and a controlled von Neumann equation.

The appendices provide necessary technical background and detail. In §A, we give background on vector-valued and stochastic differential equations. In §B, we explain how to compute the free probabilistic Laplacian used in our differential equations on a certain class of test functions. In particular, this section provides a general class of cost functions satisfying

our assumptions. In §C, we give background on amalgamated free products and show that two different non-commutative filtrations and Brownian motions can be jointly embedded into another non-commutative filtration such that the given Brownian motions are identified. This is the technical machinery needed to establish the consistency of the value function  $\bar{V}$  that makes it a well-defined tracial  $W^*$ -function. These computations may also be of interest for free probability in general.

**1.4. Acknowledgements.** W.G. was supported by NSF grant DMS-2154578 and Air Force grant FA9550-18-1-0502. D.J. was partially supported by the National Science Foundation (US) grant DMS-2002826, the National Sciences and Engineering Research Council (Canada) grant RGPIN-2017-05650, and the Independent Research Fund of Denmark grant 1026-00371B. K.N. was supported by the National Research Foundation of Korea (RS-2019-NR040050). A.Z.P. also acknowledges the support of Air Force grant FA9550-18-1-0502.

We thank Dimitri Shlyakhtenko for numerous conversations that motivated this work.

## 2. PRELIMINARIES

**2.1. Von Neumann Algebras.** We first formulate the framework of non-commutative probability spaces, following similar conventions as in [38] (note that further bibliography on von Neumann algebras is given there as well). Recall  $A$  is a  $W^*$ -algebra (von Neumann algebra) if  $A$  is a unital  $C^*$ -algebra together with an operator norm  $\|\cdot\|_\infty$  such that  $A$  as a Banach space is the dual of some Banach space  $A_*$ . If  $\tau \in A^*$  is a faithful normal trace in the dual Banach space, then we call  $\mathcal{A} = (A, \tau)$  a *tracial  $W^*$ -algebra* or a *non-commutative probability space*. The intuition behind non-commutative probability spaces is that the elements of  $A$  are non-commutative random variables, and the trace  $\tau$  is analogous to the expectation.

The GNS construction [69] produces a Hilbert space  $L^2(\mathcal{A})$  as follows: A pre-inner product can be defined on  $\mathcal{A}$  by

$$\langle X, Y \rangle_{L^2(\mathcal{A})} := \tau(X^* Y), \text{ for } X, Y \in L^2(\mathcal{A}).$$

Faithfulness of  $\tau$  implies that this is non-degenerate. Thus,  $\mathcal{A}$  be completed to a Hilbert space  $L^2(\mathcal{A})$ ; we continue to use the same notation  $\langle \cdot, \cdot \rangle_{L^2(\mathcal{A})}$  for the inner product on the completion. Moreover, for  $d$ -tuples  $X \in L^2(\mathcal{A})^d$ , we will indicate the component of the  $d$ -tuple with superscripts, as  $X = (X^1, \dots, X^d)$  and  $Y = (Y^1, \dots, Y^d)$ , and denote the inner product on  $L^2(\mathcal{A})^d$  with the subscript  $L^2(\mathcal{A})$  as

$$\langle X, Y \rangle_{L^2(\mathcal{A})} := \sum_{j=1}^d \tau(X^{j*} Y^j).$$

In either case, we define the Hilbert space norm by

$$\|X\|_{L^2(\mathcal{A})} := \sqrt{\langle X, X \rangle_{L^2(\mathcal{A})}}.$$

We let  $L^\infty(\mathcal{A}) \subset L^2(\mathcal{A})$  be the collection of elements in  $L^2(\mathcal{A})$  that are bounded in the operator norm, and we may consider  $A \subset L^\infty(\mathcal{A})$ . The operator norm also naturally extends to  $d$ -tuples by

$$\|X\|_\infty := \max_{j \in \{1, \dots, d\}} \|X^j\|_\infty, \text{ for } X = (X^1, \dots, X^d) \in L^\infty(\mathcal{A})^d.$$

Let  $1_{\mathcal{A}}$  be the identity element in  $L^2(\mathcal{A})$ . We denote by  $L^2(\mathcal{A})_{sa}^d$  the collection of self-adjoint elements in  $L^2(\mathcal{A})^d$ . We will also consider  $1_{\mathcal{A}} \in L^2(\mathcal{A})_{sa}^d$  to be the  $d$ -tuple with the identity in



each component, and  $\mathbf{e}_{\mathcal{A}}^j \in L^2(\mathcal{A})_{sa}^d$  to have the identity in the  $j$ th component and 0 in other components, for  $j \in \{1, \dots, d\}$ .

Next, we introduce an important notion of tracial  $W^*$ -embedding. We say that  $\iota : \mathcal{A} = (A, \tau) \rightarrow \mathcal{B} = (B, \rho)$  is a *tracial  $W^*$ -embedding* if it is unital  $*$ -homomorphism (it respects addition, multiplication, and adjoints) that is also trace-preserving, meaning that

$$\rho(\iota X) = \tau(X) \quad \text{for any } X \in A.$$

For such a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , there is an induced linear isometry  $\iota : L^2(\mathcal{A})^d \rightarrow L^2(\mathcal{B})^d$  where we use the same notation. For a concise notation, throughout the paper, when we write  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , we always assume that  $\iota$  is a tracial  $W^*$ -embedding. Its adjoint, or the conditional expectation, is denoted by  $E : \mathcal{B} \rightarrow \mathcal{A}$ . For more details on this setup, see [38].

Here are some examples:

- The space of  $n \times n$  complex matrices,  $M_n(\mathbb{C})$ , where the normalized trace (resp. inner product) is given by

$$\text{tr}_n(A) := \frac{1}{n} \text{tr}(A), \quad \langle A, B \rangle_{\text{tr}_n} := \text{tr}_n(A^* B).$$

In this case, we may identify  $M_n(\mathbb{C}) \cong L^2(\mathcal{A}) \cong L^\infty(\mathcal{A})$  for  $\mathcal{A} = (M_n(\mathbb{C}), \text{tr}_n)$ .

We denote by  $\mathcal{U}_n$  the set of unitary matrices  $U \in M_n(\mathbb{C})$ . For any unitary matrix  $U$ , it defines a isometry by a conjugation  $X \mapsto U^* X U$ . We say that two  $d$ -tuples  $X$  and  $Y$  of self-adjoint elements in  $M_n(\mathbb{C})_{sa}^d$  are equivalent if there exists  $U \in \mathcal{U}_n$  such that  $Y^j = U^* X^j U$  for all  $j = 1, \dots, d$ .

- We recover a classical probabilistic setting by fixing a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , where  $\Omega$  is a separable metric space and  $\mathcal{F}$  is the associated Borel  $\sigma$ -algebra. We consider the  $C^*$ -algebra  $A = C(\Omega; \mathbb{C})$ , the collection of continuous functions from  $\Omega$  to  $\mathbb{C}$ . Algebra products are given by pointwise products, and the trace corresponds to the expectation,

$$\tau(X) := \mathbb{E}[X].$$

The GNS construction produces the Hilbert space of equivalence classes of  $\mathbb{P}$ -square-integrable functions, for  $\mathcal{A} = (A, \tau)$ :

$$L^2(\mathcal{A}) \cong L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{C}).$$

The self-adjoint elements are simply the real-valued random variables.

Given a tracial von Neumann algebra  $\mathcal{A} = (A, \tau)$  and  $S \subseteq A$ , we denote  $W^*(S)$  the von Neumann subalgebra generated by  $S$ , i.e., the intersection of all tracial von Neumann algebras containing  $S$ , equipped with the trace  $\tau|_{W^*(S)}$ . Note that the inclusion map  $W^*(S) \rightarrow \mathcal{A}$  is a tracial  $W^*$ -embedding. For tracial von Neumann subalgebras  $(\mathcal{B}_j)_{j \in J}$  in  $\mathcal{A}$ , we write

$$\bigvee_{j \in J} \mathcal{B}_j = W^*\left(\bigcup_{j \in J} \mathcal{B}_j\right)$$

for their *join*, that is, the von Neumann subalgebra that they generate, equipped with the restriction of the trace  $\tau$ .

**2.2. Free independence and free products.** In non-commutative probability, there is an analog of independence known as *free independence* that relates closely to free products of tracial von Neumann algebras (and of groups).

If  $\mathcal{A} = (A, \tau)$  is a tracial  $W^*$ -algebra and  $\{\mathcal{A}_j = (A_j, \tau) : j \in J\}$  are tracial  $W^*$ -subalgebras of  $\mathcal{A}$  with an index set  $J$ , we say that  $\{\mathcal{A}_j : j \in J\}$  are *freely independent* if for all positive

integers  $n$  and  $j : \{1, \dots, n\} \rightarrow J$  such that  $j(k) \neq j(k+1)$  for  $k = 1, \dots, n-1$ ,

$$\tau \left( \prod_{k=1}^n (a_k - \tau(a_k)) \right) = 0, \quad \text{for all } (a_1, \dots, a_n) \in A_{j(1)} \times \dots \times A_{j(n)},$$

where the terms in the product are understood to be multiplied in order from left to right; see [74, 82]. More generally, if  $\{\mathcal{A}_j : j \in J\}$  is a collection of tracial  $W^*$ -algebras in  $\mathcal{A}$  containing a common subalgebra  $\mathcal{B}$  and if  $E_{\mathcal{B}}$  denote the trace-preserving conditional expectation  $\mathcal{A} \rightarrow \mathcal{B}$ , then we say that  $\{\mathcal{A}_j : j \in J\}$  are *freely independent with amalgamation over  $\mathcal{B}$*  or *freely independent over  $\mathcal{B}$*  if for all positive integers  $n$  and  $j : \{1, \dots, n\} \rightarrow J$  such that  $j(k) \neq j(k+1)$  for  $k = 1, \dots, n-1$ ,

$$E_{\mathcal{B}} \left( \prod_{k=1}^n (a_k - E_{\mathcal{B}}(a_k)) \right) = 0, \quad \text{for all } (a_1, \dots, a_n) \in A_{j(1)} \times \dots \times A_{j(n)}.$$

In the case where  $\mathcal{B} = \mathbb{C}$ , this reduces to plain free independence as defined above. We remark that, given the inclusions  $\mathcal{B} \rightarrow \mathcal{A}_j$  and  $\tau|_{\mathcal{A}_j}$ , free independence with amalgamation over  $\mathcal{B}$  uniquely determines the trace of any product  $a_1 \dots a_k$  where  $a_i$  is from  $\mathcal{A}_{i_j}$  [77, Proposition 1.3].

As in classical probability, we often need to construct independent joins of non-commutative probability spaces. Given tracial von Neumann algebras  $\mathcal{A}_j$  for  $j \in J$  containing a common subalgebra  $\mathcal{B}$ , there exists a tracial von Neumann algebra  $*_{\mathcal{B}}(\mathcal{A}_j)_{j \in J}$  containing  $\mathcal{B}$  with trace-preserving unital  $*$ -homomorphisms  $\iota_j : \mathcal{A}_j \rightarrow *_{\mathcal{B}}(\mathcal{A}_j)_{j \in J}$  such that  $\iota_j|_{\mathcal{B}} = \text{id}$  and  $(\iota_j(\mathcal{A}_j))_{j \in J}$  are freely independent over  $\mathcal{B}$  and  $*_{\mathcal{B}}(\mathcal{A}_j)_{j \in J}$  is generated by  $(\iota_j(\mathcal{A}_j))_{j \in J}$ . This algebra  $*_{\mathcal{B}}(\mathcal{A}_j)_{j \in J}$  is unique up to a canonical isomorphism (one respecting the inclusions  $\iota_j$ ), and is called the *free product of  $\mathcal{A}_j$  (with amalgamation) over  $\mathcal{B}$* . In the case  $\mathcal{B} = \mathbb{C}$ , it is called simply the free product of the  $\mathcal{A}_j$ 's and denoted  $*(\mathcal{A}_j)_{j \in J}$ . Note also that in the case of two algebras, we will sometimes use the notation  $\mathcal{A}_1 *_B \mathcal{A}_2$ . For more background on free independence and free products with amalgamation, [74, §5], [82, §3.8], [68, p. 384-385], [77, §1], [72, §III], [11, §4.7]. Because we sometimes deal with multiple different inclusions of the same two algebras, let us explicitly state what we mean by the existence and uniqueness of the free product.

**Lemma 2.1.** *Let  $J$  be any index set. Let  $(\mathcal{A}_j)_{j \in J}$  and  $\mathcal{B}$  be tracial  $W^*$ -algebras, and let  $\varphi_j : \mathcal{B} \rightarrow \mathcal{A}_j$  be a tracial  $W^*$ -embedding. Then there exists a tracial  $W^*$ -algebra  $\mathcal{C}$  and tracial  $W^*$ -embeddings  $\iota_j : \mathcal{A}_j \rightarrow \mathcal{C}$  such that*

- (1)  $\varphi = \iota_j \circ \varphi_j : \mathcal{B} \rightarrow \mathcal{C}$  is independent of  $j$ .
- (2) The images  $(\iota_j(\mathcal{A}_j))_{j \in J}$  are freely independent with amalgamation over  $\varphi(\mathcal{B})$ .
- (3)  $\mathcal{C}$  is generated by  $(\iota_j(\mathcal{A}_j))_{j \in J}$ .

Moreover, if  $\tilde{\mathcal{C}}$  and  $\tilde{\iota}_j$  are another tracial  $W^*$ -algebra and tracial  $W^*$ -embeddings satisfying these properties, then there exists a unique isomorphism  $\Phi : \mathcal{C} \rightarrow \tilde{\mathcal{C}}$  such that  $\Phi \circ \iota_j = \tilde{\iota}_j$  for  $j \in J$ .

In particular, this implies the following observation, which we will use throughout to switch perspectives between free independence and free products.

**Corollary 2.2.** *Let  $J$  be any index set. Let  $\mathcal{A}$  be a tracial  $W^*$ -algebra and  $(\mathcal{A}_j)_{j \in J}$  be a collection of subalgebras containing a common subalgebra  $\mathcal{B}$ . Let  $*_{\mathcal{B}}(\mathcal{A}_j)_{j \in J}$  be the free product with amalgamation over  $\mathcal{B}$  and let  $\iota_j$  be the canonical inclusion of  $\mathcal{A}_j$  into this free product. Then the following are equivalent:*

- (1)  $(\mathcal{A}_j)_{j \in J}$  are freely independent in  $\mathcal{A}$  with amalgamation over  $\mathcal{B}$ .

(2) *There exists a tracial  $W^*$ -isomorphism  $\phi : *_\mathcal{B}(\mathcal{A}_j)_{j \in J} \rightarrow \bigvee_{j \in J} \mathcal{A}_j \subseteq \mathcal{A}$  such that the following diagram commutes for each  $j \in J$ :*

$$\begin{array}{ccc} \mathcal{A}_j & & \\ \downarrow \iota_j & \searrow & \\ *_\mathcal{B}(\mathcal{A}_j)_{j \in J} & \xrightarrow{\phi} & \bigvee_{j \in J} \mathcal{A}_j, \end{array}$$

where the diagonal map is the inclusion.

*Proof.* (1)  $\implies$  (2) because of the uniqueness of the free product up to canonical isomorphism (Lemma 2.1).

For (2)  $\implies$  (1), first note that conditional expectations commute with isomorphisms, that is, if  $\mathcal{P} \subseteq \mathcal{M}_1$  is an inclusion of tracial  $W^*$ -algebras and  $\phi : \mathcal{M}_1 \rightarrow \mathcal{M}_2$  is an isomorphism, then  $\phi \circ E_{\mathcal{P}} = E_{\phi(\mathcal{P})} \circ \phi$ . In particular,  $\phi$  preserves the conditional expectation onto  $\mathcal{B}$ . Since the vanishing moment conditions for free independence of  $(\iota_j(\mathcal{A}_j))_{j \in J}$  over  $\mathcal{B}$  hold in  $*_{\mathcal{B}}(\mathcal{A}_j)_{j \in J}$ , it follows that they hold for  $(\mathcal{A}_j)_{j \in J}$  in  $\bigvee_{j \in J} \mathcal{A}_j \subseteq \mathcal{A}$  as well.  $\square$

The free product construction allows us, for instance, to build a larger non-commutative probability space containing any two given non-commutative probability spaces, and which agree on a common subspace. More precisely, if  $\mathcal{A}_1$  and  $\mathcal{A}_2$  are non-commutative probability spaces containing  $\mathcal{B}$ , then there is some non-commutative probability space  $\mathcal{A}$  containing  $\mathcal{B}$ , which also contains both  $\mathcal{A}_1$  and  $\mathcal{A}_2$  (namely,  $\mathcal{A} = \mathcal{A}_1 *_\mathcal{B} \mathcal{A}_2$ ). Further properties of amalgamated free products that we will use in our arguments are given in Appendix C.

**2.3. Non-commutative Laws and the weak\* topology.** In this section, we describe the analog of laws or probability distributions for non-commutative random variables from a tracial von Neumann algebra. Following the terminology in [38], we denote by  $\mathbb{W}$  a set of representatives of the isomorphism classes of tracial  $W^*$ -algebras with separable predual, so that each tracial von Neumann algebras with separable predual is isomorphic to a unique element of  $\mathbb{W}$ .

We denote by  $\text{NCP}_d := \mathbb{C}\langle x_1, \dots, x_d \rangle$ , the universal unital algebra generated by variables  $x_1, \dots, x_d$ . Note that  $\text{NCP}_d$  can be equipped with a unique  $*$ -operation (i.e., an antilinear involution satisfying  $(xy)^* = y^*x^*$ ) such that  $x_j = x_j^*$ , which makes it into a  $*$ -algebra. We then define  $\Sigma_{d,R}$  to be the linear functionals  $\lambda : \text{NCP}_d \rightarrow \mathbb{C}$  that satisfy

$$\lambda(1) = 1, \quad \lambda(pp^*) \geq 0, \quad \lambda(pq) = \lambda(qp) \quad \forall p, q \in \text{NCP}_d$$

which are  $R$ -exponentially bounded, meaning that for every  $k \in \mathbb{N}$  and any monomial  $\phi$  of degree  $k$ ,

$$|\lambda(\phi)| \leq R^k.$$

For  $\mathcal{A} \in \mathbb{W}$ , there is a natural map  $\lambda : \{X \in L^\infty(\mathcal{A})_{sa} : \|X\|_\infty \leq R\} \rightarrow \Sigma_{d,R}$  given by

$$\lambda_X(p) := \tau(p(X)) \quad \text{for } p \in \text{NCP}_d.$$

It is shown in [38] that given  $\lambda \in \Sigma_{d,R}$ , one can construct  $\mathcal{A} \in \mathbb{W}$  and find  $X \in L^\infty(\mathcal{A})_{sa}^d$  such that  $\lambda_X = \lambda$ . The Wasserstein metric is defined for  $\lambda_1, \lambda_2 \in \Sigma_{d,R}$  as

$$d_W^2(\lambda_1, \lambda_2) := \inf \left\{ \|X_1 - X_2\|_{L^2(\mathcal{A})}^2 : \mathcal{A} \in \mathbb{W}, X_1, X_2 \in L^\infty(\mathcal{A})_{sa}^d, \lambda_{X_1} = \lambda_1, \lambda_{X_2} = \lambda_2 \right\}.$$

We let  $\Sigma_d^\infty$  denote the union of  $\Sigma_{d,R}$  over all  $R > 0$  with the natural equivalence.

Since we want to work with random variables in the  $L^2$  space of the von Neumann algebra, we define  $\Sigma_d^2$  to be the closure of  $\Sigma_d^\infty$  in the Wasserstein metric. We will show that laws in

$\Sigma_d^2$  can be represented by elements of  $L^2(\mathcal{A})_{sa}^d$  for some  $\mathcal{A} \in \mathbb{W}$ , by proving an equivalent construction of  $\Sigma_d^2$ .

Alternatively,  $\tilde{\Sigma}_d^2$  is defined as the equivalence classes of operators in  $\sqcup_{\mathcal{A} \in \mathbb{W}} L^2(\mathcal{A})_{sa}^d$ , with the relation that  $X \in L^2(\mathcal{A})_{sa}^d$  is equivalent to  $Y \in L^2(\mathcal{B})_{sa}^d$  if there exists a coupling that consists of  $\mathcal{C} \in \mathbb{W}$  and tracial  $W^*$ -embeddings  $\iota_1 : \mathcal{A} \rightarrow \mathcal{C}$  and  $\iota_2 : \mathcal{B} \rightarrow \mathcal{C}$  such that  $\iota_1(X) = \iota_2(Y)$ . Given  $X \in L^2(\mathcal{A})_{sa}^d$ , we denote its law by  $\tilde{\lambda}_X \in \tilde{\Sigma}_d^2$ . This space is naturally equipped with a quotient metric, which is equivalent to the non-commutative Wasserstein metric defined as

$$\tilde{d}_W^2(\lambda, \mu) := \inf_{\mathcal{A} \in \mathbb{W}} \left\{ \|X - Y\|_{L^2(\mathcal{A})}^2 : X, Y \in L^2(\mathcal{A})_{sa}^d \text{ such that } \tilde{\lambda}_X = \lambda \text{ and } \tilde{\lambda}_Y = \mu \right\}. \quad (2.1)$$

**Lemma 2.3.** *The definition of  $\tilde{\Sigma}_d^2$  as a quotient space with respect to the couplings of von Neumann algebra is equivalent to  $\Sigma_d^2$  as the closure of  $\Sigma_d^\infty$  under the Wasserstein metric.*

*Proof.* Given  $\tilde{\lambda} \in \tilde{\Sigma}_d^2$ , we may find  $\mathcal{A} \in \mathbb{W}$  and  $X \in L^2(\mathcal{A})_{sa}^d$  such that  $\tilde{\lambda}_X = \tilde{\lambda}$ . By the density of  $L^\infty(\mathcal{A})_{sa}^d$  in  $L^2(\mathcal{A})_{sa}^d$ , we take a sequence  $\{X_i\}_{i \geq 1}$  in  $L^\infty(\mathcal{A})_{sa}^d$  that converges to  $X$  in  $L^2(\mathcal{A})_{sa}^d$ . Clearly,  $\tilde{\lambda}_{X_i}$  converges in the Wasserstein metric, showing that  $\tilde{\Sigma}_d^2 \subset \Sigma_d^2$ .

For the other direction, we consider a sequence with  $\lambda_i \in \Sigma_d^\infty$  and  $d_W^2(\lambda_i, \lambda_{i+1}) \leq 2^{-i}$  for  $i \in \{1, 2, \dots\}$ . We construct by induction a sequence of von Neumann algebra,  $\{\mathcal{A}_i\}_{i \geq 1}$ , representing the non-commutative law  $\lambda_i$  by  $X_i \in L^\infty(\mathcal{A}_i)_{sa}^d$  with tracial  $W^*$ -embeddings  $\iota_i : \mathcal{A}_i \rightarrow \mathcal{A}_{i+1}$  so that

$$\|\iota_i(X_i) - X_{i+1}\|_{L^2(\mathcal{A}_{i+1})}^2 = d_W^2(\lambda_i, \lambda_{i+1}).$$

From the definition of the Wasserstein metric, there is a von Neumann algebra  $\mathcal{B}$  and  $\tilde{X}, \tilde{Y} \in L^\infty(\mathcal{B})_{sa}^d$  such that  $\lambda_{\tilde{X}} = \lambda_i$ ,  $\lambda_{\tilde{Y}} = \lambda_{i+1}$ , and

$$\tilde{d}_W^2(\lambda_i, \lambda_{i+1}) = \|\tilde{X} - \tilde{Y}\|_{L^2(\mathcal{B})}^2.$$

We may define a tracial  $W^*$ -embedding  $\tilde{\iota} : W^*(\tilde{X}) \rightarrow \mathcal{A}_i$  by identifying  $\iota(\tilde{X}) = X$ . We then define  $\mathcal{A}_{i+1}$  as the free product of  $\mathcal{A}_i$  and  $\mathcal{B}$  with amalgamation over  $W^*(\tilde{X})$ . We let  $\iota_i$  be the tracial  $W^*$ -embedding of  $\mathcal{A}_i$  into  $\mathcal{A}_{i+1}$ , which satisfies the inductive hypothesis.

We next define  $\mathcal{A}_\infty$  as the inductive limit of the  $\mathcal{A}_i$ , and one can consider the sequence  $\{\tilde{X}_i\}_{i \geq 1}$  taking values in  $L^\infty(\mathcal{A}_\infty)_{sa}^d$ . It follows that

$$\|\tilde{X}_i - \tilde{X}_{i+1}\|_{L^2(\mathcal{A}_\infty)}^2 = \|\iota(X_i) - X_{i+1}\|_{L^2(\mathcal{A}_{i+1})}^2 \leq 2^{-i}.$$

Therefore, we have a Cauchy sequence and a limit point  $X \in L^2(\mathcal{A}_\infty)_{sa}^d$ , which defines  $\tilde{\lambda}_X \in \tilde{\Sigma}_d^2$  and shows  $\Sigma_d^2 \subset \tilde{\Sigma}_d^2$ .  $\square$

Here is a list of some properties for non-commutative laws:

- It was shown in [38, §5.5] that  $\Sigma_{d,R}$  is not separable with the Wasserstein metric (and therefore neither is  $\Sigma_d^2$ ).
- The spaces  $\Sigma_{d,R}$  (which can be viewed as subsets of  $\Sigma_d^2$  with the operator norm bounded by  $R$ ) are compact in the weak\* topology of convergence for the evaluation of every polynomial in  $\text{NCP}_d$ .
- The continuous functions  $C(\Sigma_d^2)$  can be identified with tracial  $W^*$ -functions  $(f_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  such that  $f_{\mathcal{A}} \in C(L^2(\mathcal{A})_{sa}^d)$  for all  $\mathcal{A} \in \mathbb{W}$ , where we recall the following definition of tracial function from [38, Definition 3.5].

**Definition 2.4** (Definition 3.5 in [38]). A tracial  $W^*$ -function on  $\Sigma_d^2$  with values in  $(-\infty, +\infty]$  is a collection of functions  $f_{\mathcal{A}} : L^2(\mathcal{A})_{sa}^d \rightarrow (-\infty, +\infty]$  for  $\mathcal{A} \in \mathbb{W}$ , such that whenever  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  is a tracial  $W^*$ -embedding,  $f_{\mathcal{A}} = f_{\mathcal{B}} \circ \iota$  (here  $\iota$  is extended to a map  $L^2(\mathcal{A})_{sa}^d \rightarrow L^2(\mathcal{B})_{sa}^d$ ).

### 3. CONTROL PROBLEM WITH FREE INDIVIDUAL NOISE AND COMMON NOISE

**3.1. Example Problems.** Before introducing the general problems we will consider, we introduce a handful of the motivating problems. These examples will be further worked out in Section 5.

**3.1.1. Quadratic cost.** The simplest type of drift is when the control space corresponds directly with the drift of the process, there is a single common noise, and free individual noise for each element. The dynamics may be expressed in  $L^2(\mathcal{A})_{sa}^d$ , for a von Neuman algebra  $\mathcal{A} \in \mathbb{W}$ , as

$$dX_t^j = \alpha_t^j dt + \beta_C \mathbf{1}_{\mathcal{A}} dW_t^0 + \beta_F dS_t^j, \quad j \in \{1, \dots, d\}, \quad (3.1)$$

where  $(W_t^0)_{t \in [0, T]}$  is a standard Brownian motion that affects all elements proportional to the identity element, and  $(S_t^j)_{t \in [0, T]}^{j \in \{1, \dots, d\}}$  are freely independent semi-circular processes. These processes represent a source of non-commutative noise, for example, arising from the Gaussian unitary ensemble of random matrices.

In Section 5, we solve exactly the problem of minimizing a quadratic terminal cost involving the second moments along with first moments squared,

$$g_{\mathcal{A}}(X) = \sum_{i=1}^d \sum_{j=1}^d g_{ij}^0 \tau(X^i X^j) + \sum_{i=1}^d \sum_{j=1}^d g_{ij}^1 \tau(X^i) \tau(X^j).$$

That is, given  $t_0 \in [0, T]$  and a non-commutative law  $\lambda_0$ , setting  $\mathbb{A}_{\mathcal{A}}^{t_0, T}$  to be the set of admissible control policies for  $\mathcal{A} \in \mathbb{W}$  (defined in Section 3.2.2 later precisely), we solve the variational problem

$$\bar{V}(t_0, \lambda_0) := \inf_{\mathcal{A} \in \mathbb{W}} \inf_{\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}}^{t_0, T}} \left\{ \mathbb{E} \left[ \int_{t_0}^T \frac{1}{2} \|\alpha_t\|_{L^2(\mathcal{A})}^2 dt + g_{\mathcal{A}}(X_T) \right] : \lambda_{X_{t_0}} = \lambda_0, X \text{ solves (3.1) on } [t_0, T] \right\}.$$

One can generalize this example in several ways. For instance, one can add a running cost  $\phi_{\mathcal{A}}(X_t)$  to the integral in the cost for some tracial  $W^*$ -function  $(\phi_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$ , and  $(g_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  can be any tracial  $W^*$ -function. In this case, the corresponding Hamiltonian can be expressed as a tracial  $W^*$ -function on  $\Sigma_{2d}^2$ ,  $(H_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$ , by

$$H_{\mathcal{A}}(X, P) = \frac{1}{2} \|P\|_{L^2(\mathcal{A})}^2 - \phi_{\mathcal{A}}(X).$$

The value function  $\bar{V}$  defined above solves the Hamilton-Jacobi equation

$$-\partial_t \bar{V}(t, \lambda) + \frac{1}{2} \|\partial \bar{V}(t, \lambda)\|_{\lambda}^2 - \frac{\beta_C^2}{2} \Delta \bar{V}(t, \lambda) - \frac{\beta_F^2}{2} \Theta \bar{V}(t, \lambda) = \phi(\lambda),$$

$$\bar{V}(T, \lambda) = g(\lambda),$$

in a manner to be described later.

This example arises in the theory of large deviation for random matrices, which we will explore in more detail in the subsequent paper.

**3.1.2. Eikonal controls.** For a related example, we suppose that the controls are constrained so that  $\|\alpha\|_{L^2(\mathcal{A})} \leq 1$ . This results in a similar equation with the 1-homogeneous Hamiltonian  $H_{\mathcal{A}}(P) = \|P\|_{L^2(\mathcal{A})}$  rather than the quadratic one. Alternatively, instead of restricting the  $L^2$  norm of the control, it is possible to restrict the operator norm  $\|\alpha\|_{\infty} \leq 1$ , resulting in an  $L^1$  norm, i.e.,  $H_{\mathcal{A}}(P) = \tau(|P|)$  where  $|P| = P_+ + P_-$  given the decomposition into positive and negative definite parts  $P = P_+ - P_-$ . These examples have nonsmooth solutions, as illustrated in Section 5.

3.1.3. *Controlled von Neumann equation.* We can consider the dynamics driven by the commutator of the state  $X_t$  and the control  $\alpha_t$ :

$$dX_t^j = i[X_t^j, \alpha_t^j] dt, \quad j \in \{1, \dots, d\}. \quad (3.2)$$

This equation is motivated by the von Neumann equation in quantum mechanics where a system evolves through unitary conjugation. Indeed, a straightforward computation shows that

$$X_t^j = U_t^j X_0^j (U_t^j)^*$$

where  $U_t^j$  is the unitary given by

$$dU_t^j = i\alpha_t^j U_t^j dt.$$

In Section 5.3, we show that when the value function

$$\bar{V}(t_0, \lambda_0) := \inf_{\mathcal{A} \in \mathbb{W}} \inf_{\tilde{\alpha} \in \mathcal{A}^{t_0, T}} \left\{ \int_{t_0}^T \frac{1}{2} \|\alpha_t\|_{L^2(\mathcal{A})}^2 dt + g_{\mathcal{A}}(X_T) : \lambda_{X_{t_0}} = \lambda_0, (X_t)_t \text{ solves (3.2) on } [t_0, T] \right\}$$

admits a minimizer, we have a *non-commutative Hopf-Lax formula*

$$\begin{aligned} & \bar{V}(t_0, \lambda_0) \\ &= \inf_{\mathcal{A} \in \mathbb{W}} \inf_{\alpha \in L^2(\mathcal{A})_{sa}^d} \left\{ \frac{(T - t_0)}{2} \|\alpha\|_{L^2(\mathcal{A})}^2 + g_{\mathcal{A}}(e^{-i\alpha(T-t_0)} X e^{i\alpha(T-t_0)}) : X \in L^2(\mathcal{A})_{sa}^d, \lambda_X = \lambda_0 \right\}. \end{aligned}$$

Another motivation for this setup is Voiculescu's liberation theory [80], a version of information theory where additive perturbations are replaced by perturbations through unitary conjugation, so for instance the analogue of semi-circular Brownian motion to an initial condition would be conjugating the initial condition by a tuple of free unitary Brownian motions as in [7]. Similarly, the non-commutative Hopf-Lax formula above is the liberation analogue of the standard Hopf-Lax formula based on additive perturbations. More generally, one could consider both a control term and stochastic term in the von Neumann equation, setting

$$dU_t^j = i(dS_t^j + \alpha_t^j dt) U_t^j$$

and

$$dX_t^j = i[dS_t^j + \alpha_t^j dt, X_t^j]$$

for a free Brownian motion  $S_t^j$ , although we will focus in this paper on the deterministic case.

This example does not quite satisfy our general assumptions in §3.2 since we want the Lagrangian to extend to  $X$  and  $\alpha$  in  $L^2$  and be a globally Lipschitz function plus a quadratic function of  $\alpha$ . However, it is an important example of a drift which is an  $E$ -linear tracial vector-field as defined in §3.2.3(b) below. In other words, whenever there is a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ ,

$$[\iota X, \iota \alpha] = \iota [X, \alpha] \text{ and } E[\iota X, \beta] = [X, E\beta] \quad \text{for any } X, \alpha \in L^2(\mathcal{A})_{sa}^d \text{ and } \beta \in L^2(\mathcal{B})_{sa}^d,$$

where  $E : \mathcal{B} \rightarrow \mathcal{A}$  denotes the conditional expectation adjoint to  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ .

**3.2. Setup and Assumptions for the General Problem.** Given a tracial  $W^*$ -algebra  $\mathcal{A}$ , the “state space” in mean field games terminology corresponds will be the space of the self-adjoint  $d$ -tuples in the space  $L^2(\mathcal{A})_{sa}^d$ . Throughout this section, we assume that  $T > 0$  is a given terminal time. We now describe the terminology and setup for the Brownian motions, control policies, drift function, and cost functions for the general non-commutative stochastic optimization problem (1.5). Throughout the paper, **Assumption A** will refer to the conditions on the control policies, drift, and cost functions in §3.2.2, §3.2.3, and §3.2.4. Another more restrictive set of hypotheses called **Assumption B** will be given in §3.2.5.

3.2.1. *Classical and free Brownian motions.* We assume to be given a non-atomic complete filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$  which supports the Brownian motion  $(W_t^0)_{t \in [0, T]} \in C([0, T]; L^2(\Omega, \mathcal{F}, \mathbb{P}))$ :

- (a)  $W_0^0 = 0$ .
- (b) For  $0 \leq s \leq t \leq T$ ,  $W_t^0 - W_s^0$  is normally distributed with mean 0 and variance  $t - s$ .
- (c) For any sequence of times  $0 = t_0 \leq t_1 \leq t_2 \leq \dots \leq t_{k-1} \leq t_k = T$ , the collection of increments  $W_{t_{j+1}}^0 - W_{t_j}^0$  for  $j \in \{0, 1, \dots, k-1\}$  are (mutually) independent.

We assume that  $\mathcal{F} = \sigma(W_s^0 : 0 \leq s \leq T)$  and  $\mathcal{F}_t = \sigma(W_s^0 : 0 \leq s \leq t)$  for  $t \in [0, T]$ .

Throughout the paper, we write products between the Brownian motions and the von Neumann algebra, corresponding to the canonical embedding of scalars by  $\mathbf{1}_{\mathcal{A}} W_t^0 \in L^\infty(\mathcal{A})_{sa}$  or  $\mathbf{1}_{\mathcal{A}} W_t^0 \in L^\infty(\mathcal{A})_{sa}^d$ . We emphasize that the Wiener probability space may be fixed throughout, whereas it is not possible to fix the non-commutative probability space.

One can similarly define a free semi-circular process. Analogous to the notion of filtration of  $\sigma$ -algebra, we call an increasing collection of sub-von Neumann algebra a free filtration. Let  $0 \leq t_0 \leq t_1 \leq T$ . For a given  $\mathcal{A} \in \mathbb{W}$  and a free filtration  $(\mathcal{A}_t)_{t \in [t_0, t_1]}$ , we say that a  $d$ -dimensional process  $(S_t)_{t \in [t_0, t_1]} \in C([t_0, t_1]; L^\infty(\mathcal{A})_{sa}^d)$  is a *free semi-circular process (or free Brownian motion) compatible with the free filtration  $(\mathcal{A}_t)_{t \in [t_0, t_1]}$*  on the interval  $[t_0, t_1]$ , if it satisfies the following properties:

- (a)  $S_{t_0} = 0$ .
- (b) For  $t_0 \leq s \leq t \leq t_1$  and  $l \in \{1, \dots, d\}$ , the increment  $S_t^l - S_s^l$  is semi-circularly distributed with mean 0 and variance  $t - s$ , and the components  $\{S_t^l - S_s^l\}_{l=1}^d$  are freely independent.
- (c)  $S_t \in L^\infty(\mathcal{A}_t)_{sa}^d$  for all  $t \in [t_0, t_1]$ .
- (d) For  $t_0 \leq s \leq t \leq t_1$ ,  $S_t - S_s$  is freely independent of  $\mathcal{A}_s$ .

Note that (c) and (d) together imply that for  $t_0 = s_0 \leq s_1 \leq \dots \leq s_k = t_1$ , the collection of increments  $S_{s_{j+1}}^l - S_{s_j}^l$  for  $j \in \{0, 1, \dots, k-1\}$  and  $l \in \{1, \dots, d\}$  are freely independent (but the converse is not true since  $(\mathcal{A}_t)_t$  could be larger than algebra generated by the initial conditions and semi-circulars).

We list now the general assumptions on the problem that we will refer to as Assumption **A**.

3.2.2. *Sets of the controls and admissible control policies.* We assume that controls in  $\mathcal{A}$  belong to some subset  $\mathbb{A}_{\mathcal{A}}$  of  $L^2(\mathcal{A})_{sa}^d$  which satisfies

- (a)  $\mathbb{A}_{\mathcal{A}} \subset L^2(\mathcal{A})_{sa}^d$  is closed and convex.
- (b)  $0 \in \mathbb{A}_{\mathcal{A}}$ .
- (c) For any  $\mathcal{B} \in \mathbb{W}$  and a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  (with its adjoint  $E$ ), we have

$$\iota \mathbb{A}_{\mathcal{A}} \subset \mathbb{A}_{\mathcal{B}} \quad \text{and} \quad E \mathbb{A}_{\mathcal{B}} \subset \mathbb{A}_{\mathcal{A}}.$$

Given  $\mathcal{A} \in \mathbb{W}$ ,  $[t_0, t_1] \subset [0, T]$ , and  $x_0 \in L^2(\mathcal{A})_{sa}^d$ , we let  $\mathbb{A}_{\mathcal{A}, x_0}^{t_0, t_1}$  be the collection of *admissible* control policies

$$\tilde{\alpha} = \left( (\alpha_t)_{t \in [t_0, t_1]}, (\mathcal{A}_t)_{t \in [t_0, t_1]}, (S_t)_{t \in [t_0, t_1]} \right)$$

that satisfy the following properties:

- (a)  $(\mathcal{A}_t)_{t \in [t_0, t_1]}$  is a free filtration, i.e., an increasing sequence of tracial  $W^*$  subalgebras of  $\mathcal{A}$ , such that  $x_0 \in L^2(\mathcal{A}_{t_0})_{sa}^d$ .
- (b)  $(S_t)_{t \in [t_0, t_1]}$  is a  $d$ -dimensional free semi-circular process compatible with  $(\mathcal{A}_t)_{t \in [t_0, t_1]}$ .

- (c) For every  $s \in [t_0, t_1]$ , we have  $(\alpha_t)_{t \in [t_0, s]} \in L^2([t_0, s] \times (\Omega, \mathcal{F}_s, \mathbb{P}); \mathbb{A}_{\mathcal{A}_s})$ , i.e.,  $(\alpha_t)_{t \in [t_0, t_1]}$  is progressively measurable and freely progressive.

Note that progressive measurability is in general a stronger condition than adaptedness. However if each path is continuous, then adaptedness does imply progressive measurability.

*Remark 3.1.* We may have  $\mathbb{A}_{\mathcal{A}, x_0}^{t_0, t_1} = \emptyset$  even when  $\mathbb{A}_{\mathcal{A}} \neq \emptyset$ , since not every tracial  $W^*$ -algebra  $\mathcal{A}$  contains a free semi-circular process.

**3.2.3. Drift function.** We consider drift functions  $b_{\mathcal{A}} : L^2(\mathcal{A})_{sa}^d \times \mathbb{A}_{\mathcal{A}} \rightarrow L^2(\mathcal{A})_{sa}^d$  that associate a ‘tangent vector’ for every state and control. We make the following assumptions.

- (a)  $(b_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  defines a tracial vector-field in the sense that for any  $\mathcal{A}, \mathcal{B} \in \mathbb{W}$  with a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ ,  $X \in L^2(\mathcal{A})_{sa}^d$  and  $\alpha \in \mathbb{A}_{\mathcal{A}} \subset L^2(\mathcal{A})_{sa}^d$ ,

$$\iota b_{\mathcal{A}}(X, \alpha) = b_{\mathcal{B}}(\iota X, \iota \alpha). \quad (3.3)$$

- (b)  $(b_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is *E-affine* in the sense that for any  $\mathcal{A} \in \mathbb{W}$  and  $X \in L^2(\mathcal{A})_{sa}^d$ ,  $\alpha \mapsto b_{\mathcal{A}}(X, \alpha)$  is affine, also for any tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  with its adjoint  $E : L^2(\mathcal{B})_{sa}^d \rightarrow L^2(\mathcal{A})_{sa}^d$ ,

$$b_{\mathcal{A}}(X, E \alpha) = E b_{\mathcal{B}}(\iota X, \alpha) \text{ for all } \alpha \in \mathbb{A}_{\mathcal{B}}. \quad (3.4)$$

- (c)  $(b_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is uniformly continuous on bounded sets, i.e. for any  $M > 0$  there is a modulus of continuity  $\omega_M$  such that for all  $\mathcal{A} \in \mathbb{W}$ ,  $X, Y \in L^2(\mathcal{A})$  and  $\alpha, \beta \in \mathbb{A}_{\mathcal{A}}$  with  $\max\{\|X\|_{L^2(\mathcal{A})}, \|Y\|_{L^2(\mathcal{A})}, \|\alpha\|_{L^2(\mathcal{A})}, \|\beta\|_{L^2(\mathcal{A})}\} \leq M$ ,

$$\|b_{\mathcal{A}}(X, \alpha) - b_{\mathcal{A}}(Y, \beta)\|_{L^2(\mathcal{A})}^2 \leq \omega_M(\|X - Y\|_{L^2(\mathcal{A})}^2 + \|\alpha - \beta\|_{L^2(\mathcal{A})}^2).$$

- (d) There is a constant  $\bar{C} > 0$  such that for all  $\mathcal{A} \in \mathbb{W}$ ,  $X_1, X_2 \in L^2(\mathcal{A})$  and  $\alpha \in \mathbb{A}_{\mathcal{A}}$ ,

$$\|b_{\mathcal{A}}(X_1, \alpha) - b_{\mathcal{A}}(X_2, \alpha)\|_{L^2(\mathcal{A})} \leq \bar{C} \|X_1 - X_2\|_{L^2(\mathcal{A})}. \quad (3.5)$$

*Remark 3.2.* These assumptions on the drift, especially (d), imply that for any  $M > 0$ , there exists a constant  $v_M > 0$  such that for all  $X \in L^2(\mathcal{A})_{sa}^d$  with  $\|X\|_{L^2(\mathcal{A})} \leq M$  and  $\alpha, \alpha' \in \mathbb{A}_{\mathcal{A}}$ ,

$$\|b_{\mathcal{A}}(X, \alpha) - b_{\mathcal{A}}(X, \alpha')\|_{L^2(\mathcal{A})} \leq v_M \|\alpha - \alpha'\|_{L^2(\mathcal{A})} \text{ and } \|b_{\mathcal{A}}(X, \alpha)\|_{L^2(\mathcal{A})} \leq v_M (1 + \|\alpha\|_{L^2(\mathcal{A})}). \quad (3.6)$$

**3.2.4. Cost functions.** We consider a running cost  $L_{\mathcal{A}} : L^2(\mathcal{A})_{sa}^d \times \mathbb{A}_{\mathcal{A}} \rightarrow \mathbb{R}$  and a terminal cost  $g_{\mathcal{A}} : L^2(\mathcal{A})_{sa}^d \rightarrow \mathbb{R}$  that satisfy the following assumptions.

- (a) Both  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  and  $(g_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  are tracial  $W^*$ -functions. Equivalently,  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  may be defined on the space of joint non-commutative laws in  $\Sigma_{2d}^2$  and  $(g_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a function on the space of non-commutative laws in  $\Sigma_d^2$ .
- (b)  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is *E-convex* in the control variable, meaning that for any  $\mathcal{A} \in \mathbb{W}$  and  $X \in L^2(\mathcal{A})_{sa}^d$ ,  $\alpha \mapsto L_{\mathcal{A}}(X, \alpha)$  is convex, and for any tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  with its adjoint  $E : L^2(\mathcal{B})_{sa}^d \rightarrow L^2(\mathcal{A})_{sa}^d$ ,

$$L_{\mathcal{A}}(X, E \alpha) \leq L_{\mathcal{B}}(\iota X, \alpha) \text{ for all } \alpha \in \mathbb{A}_{\mathcal{B}}. \quad (3.7)$$

- (c) Similar to the drift, we assume that  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is uniformly continuous on bounded sets.
- (d) There exists a constant  $C_1 > 0$  such that for all  $\mathcal{A} \in \mathbb{W}$ ,  $X \in L^2(\mathcal{A})_{sa}^d$  and  $\alpha \in \mathbb{A}_{\mathcal{A}}$ ,

$$\begin{aligned} -C_1 + \frac{1}{C_1} \|\alpha\|_{L^2(\mathcal{A})}^2 &\leq L_{\mathcal{A}}(X, \alpha) \leq C_1 (1 + \|X\|_{L^2(\mathcal{A})} + \|\alpha\|_{L^2(\mathcal{A})}^2), \\ -C_1 &\leq g_{\mathcal{A}}(X) \leq C_1 (1 + \|X\|_{L^2(\mathcal{A})}). \end{aligned} \quad (3.8)$$



Also  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  and  $(g_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  are Lipschitz with respect to  $X$ : There exists a constant  $C_2 > 0$  such that for all  $\mathcal{A} \in \mathbb{W}$ ,  $X_1, X_2 \in L^2(\mathcal{A})_{sa}^d$  and  $\alpha \in \mathbb{A}_{\mathcal{A}}$ ,

$$\begin{aligned} |L_{\mathcal{A}}(X_1, \alpha) - L_{\mathcal{A}}(X_2, \alpha)| &\leq C_2 \|X_1 - X_2\|_{L^2(\mathcal{A})}, \\ |g_{\mathcal{A}}(X_1) - g_{\mathcal{A}}(X_2)| &\leq C_2 \|X_1 - X_2\|_{L^2(\mathcal{A})}. \end{aligned} \quad (3.9)$$

**3.2.5. Additional assumptions (Assumption B).** As mentioned above the assumptions in Sections 3.2.2, 3.2.3 and 3.2.4 are referred to **Assumption A**. Sometimes to simplify the analysis, we also assume the following conditions, which we refer to as **Assumption B**:

- (a) The control set is  $\mathbb{A}_{\mathcal{A}} = L^2(\mathcal{A})_{sa}^d$  and the drift function has the form

$$b_{\mathcal{A}}(X, \alpha) = \alpha.$$

- (b) The Lagrangian  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is jointly  $E$ -convex in  $(X, \alpha)$  and the terminal cost  $(g_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is  $E$ -convex. This means that for any  $\mathcal{A} \in \mathbb{W}$ , the maps  $(X, \alpha) \mapsto L_{\mathcal{A}}(X, \alpha)$  and  $X \mapsto g_{\mathcal{A}}(X)$  are convex, and for any tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  with its adjoint  $E : L^2(\mathcal{B})_{sa}^d \rightarrow L^2(\mathcal{A})_{sa}^d$ ,

$$\begin{aligned} L_{\mathcal{A}}(E X, E \alpha) &\leq L_{\mathcal{B}}(X, \alpha) \text{ for all } X \in L^2(\mathcal{B})_{sa}^d \text{ and } \alpha \in \mathbb{A}_{\mathcal{B}}, \\ g_{\mathcal{A}}(E X) &\leq g_{\mathcal{B}}(X) \text{ for all } X \in L^2(\mathcal{B})_{sa}^d. \end{aligned}$$

**3.2.6. Stochastic differential equations on the von Neumann algebra.** Now we describe how the process  $X_t$  in the variational problem (1.5) is constructed from the control as a solution to a free stochastic differential equation (SDE). We specify diffusion coefficients  $\beta_C \geq 0$  and  $\beta_F \geq 0$ . Let  $\mathcal{A} \in \mathbb{W}$  and  $[t_0, t_1] \subset [0, T]$ . For  $x_0 \in L^2(\mathcal{A})_{sa}^d$  and  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, t_1}$ , we consider the SDE on  $L^2(\mathcal{A})_{sa}^d$ :

$$\begin{cases} dX_t = b_{\mathcal{A}}(X_t, \alpha_t) dt + \beta_C \mathbf{1}_{\mathcal{A}} dW_t^0 + \beta_F dS_t, \\ X_{t_0} = x_0. \end{cases} \quad (3.10)$$

We consider strong solutions such that  $t \mapsto X_t$  is continuous in the  $L^2$  norm, which are adapted to the common noise and freely adapted; this means more explicitly that  $\omega \mapsto X_t(\omega)$  is  $\mathcal{F}_t$  measurable and  $X_t \in L^2(\mathcal{A}_t)_{sa}^d$   $\mathbb{P}$ -a.s. for each  $t \in [t_0, t_1]$ , and satisfies the integral equation, expressed here component-wise,

$$X_t^j = x_0^j + \int_{t_0}^t b_{\mathcal{A}}^j(X_s, \alpha_s) ds + \beta_C \mathbf{1}_{\mathcal{A}} (W_s^0 - W_{t_0}^0) + \beta_F (S_s^j - S_{t_0}^j), \quad j = 1, \dots, d.$$

Here  $(S_t)_{t \in [t_0, T]}$  denotes the free Brownian motion on  $[t_0, T]$ . Note that  $S_{t_0} = 0$ .

We denote by  $X_t[t_0, x_0, \tilde{\alpha}]$  the solution of (3.10) on  $[t_0, t_1]$ , or  $X_t[\tilde{\alpha}]$  when  $(t_0, x_0)$  are clear from the context. The cost associated to a control policy  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, T}$  is defined as

$$\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{A}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{A}}(X_T[\tilde{\alpha}]) \right].$$

*Remark 3.3.* Observe that under the assumptions above, if  $\mathcal{A}, \mathcal{B} \in \mathbb{W}$  and  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  is a tracial  $W^*$ -embedding, then  $Y := \iota X$  satisfies the analogue of (3.10) where  $(X, b_{\mathcal{A}}, \alpha, S)$  is replaced by  $(Y, b_{\mathcal{B}}, \iota \alpha, \iota S)$ , where the original filtration  $\mathcal{A}_t$  (embedded in  $\mathcal{B}$ ) is used as the filtration in  $\mathcal{B}$  as well.

**3.2.7. The value function.** As mentioned in the introduction, we want to allow the ambient algebra  $\mathcal{B}$  to vary in the optimization problems. We will therefore define several versions of the value function. The first  $\tilde{V}_{\mathcal{A}}$  only looks at filtrations in a fixed algebra, the second  $\bar{V}_{\mathcal{A}}$  allows arbitrary extensions of the given algebra  $\mathcal{A}$ , and the third  $\bar{V}$  is defined on the space of non-commutative laws.

For  $\mathcal{A} \in \mathbb{W}$ , we consider the value function on  $L^2(\mathcal{A})_{sa}^d$ , defined as follows: For  $t_0 \in [0, T]$  and  $x_0 \in L^2(\mathcal{A})_{sa}^d$ ,

$$\tilde{V}_{\mathcal{A}}(t_0, x_0) := \inf_{\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, T}} \left\{ \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{A}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{A}}(X_T[\tilde{\alpha}]) \right] : X_{t_0}[\tilde{\alpha}] = x_0 \right\}, \quad (3.11)$$

where  $X_t[\tilde{\alpha}]$  denotes a solution to (3.10) and  $\mathbb{E}$  denotes expectation with respect to the common noise. Note that if  $\mathcal{A}$  does not support a free Brownian motion  $(S_t)_{t \in [t_0, T]}$  freely independent of  $x_0$ , i.e. if  $\mathbb{A}_{\mathcal{A}, x_0}^{t_0, T}$  is an empty set, then this definition will result in  $+\infty$ .

The eventual value function is defined as follows: For  $\mathcal{A} \in \mathbb{W}$ ,

$$\bar{V}_{\mathcal{A}}(t_0, x_0) := \inf_{\iota: \mathcal{A} \rightarrow \mathcal{B}} \tilde{V}_{\mathcal{B}}(t_0, \iota x_0), \quad (3.12)$$

where the infimum is performed over the set of  $(\mathcal{B}, \iota)$  such that  $\mathcal{B} \in \mathbb{W}$  and  $\iota: \mathcal{A} \rightarrow \mathcal{B}$  is a tracial  $W^*$ -embedding. The function  $\bar{V}_{\mathcal{A}}$  will be our main object of study, while  $\tilde{V}_{\mathcal{A}}$  is largely an intermediate step in making the definition and lacks many desirable properties such as continuous dependence on the initial condition (since for some choices of  $\mathcal{A}$ , a small perturbation of  $x_0$  can preclude the existence of a freely independent Brownian motion).

We will also view this value function as a function on the space of non-commutative laws, i.e. the non-commutative Wasserstein space. For  $t_0 \in [0, T]$  and  $\lambda \in \Sigma_d^2$ , take any  $\mathcal{A} \in \mathbb{W}$  and  $x_0 \in \mathcal{A}$  such that  $\lambda_{x_0} = \lambda$ , and then define the value function

$$\bar{V}(t_0, \lambda) := \bar{V}_{\mathcal{A}}(t_0, x_0). \quad (3.13)$$

We will show that this is well-defined, i.e. independent of the particular  $\mathcal{A}$  and  $x_0$  used to represent  $\lambda$ , in Lemma 3.9 below.

**3.2.8. The Hamiltonian.** The Hamilton-Jacobi equation satisfied by the value function includes the *Hamiltonian*  $H$  which is the Fenchel-Legendre dual of the Lagrangian. In the non-commutative setting, the Hamiltonian is defined as follows: For a tracial function  $L_{\mathcal{A}}: L^2(\mathcal{A})_{sa}^d \times \mathbb{A}_{\mathcal{A}} \rightarrow \mathbb{R}$ , let

$$H_{\mathcal{A}}(X, P) := \sup_{\iota: \mathcal{A} \rightarrow \mathcal{B} \in \mathbb{W}} \sup_{\alpha \in \mathbb{A}_{\mathcal{B}}} \left\{ \langle b_{\mathcal{B}}(\iota X, \alpha), \iota P \rangle_{L^2(\mathcal{B})} - L_{\mathcal{B}}(\iota X, \alpha) \right\}, \quad X, P \in L^2(\mathcal{A})_{sa}^d. \quad (3.14)$$

We will show that the Hamiltonian is a tracial  $W^*$ -function in Lemma 3.4 below.

**3.2.9. Generalizations.** Before going on to prove properties of the value function that follow from Assumption A, we remark that the setup could be significantly generalized, although this would make the analysis correspondingly more complicated, and hence we leave it for future work.

First, we could include non-constant coefficients in the diffusion terms. For the free semi-circular process, this might correspond to tracial  $W^*$ -tensor field  $\eta_{\mathcal{A}}: L^2(\mathcal{A})_d^{sa} \rightarrow L^2(\mathcal{A})_d^{sa} \otimes L^2(\mathcal{A})_d^{sa}$ . We refer to [9, 71, 3] for a theory of free stochastic differential equations. Most of the theory we develop will go through, but we will remark at certain points where complications arise.

Many important examples of controlled partial differential equations feature drifts that are unbounded. Similar, unbounded operators commonly arise in quantum mechanics. In these cases, the drift may be defined on a dense subset and continuous with respect to a stronger topology while we impose that  $X \rightarrow \bar{C} X - b_{\mathcal{A}}(X, \alpha)$  is monotone, i.e.:

$$\langle b_{\mathcal{A}}(X_1, \alpha) - b_{\mathcal{A}}(X_2, \alpha), X_1 - X_2 \rangle_{L^2(\mathcal{A})} \leq \bar{C} \|X_1 - X_2\|_{L^2(\mathcal{A})}^2. \quad (3.15)$$

Significant effort was made in the theory of infinite dimensional viscosity solutions to handle such cases. Due to the complications when handling such unbounded terms, we do not include them in our current analysis. We note that the example of the controlled von Neumann equation  $b_{\mathcal{A}}(X, \alpha) = i[X, \alpha]$  does not satisfy the uniform continuity requirements but does satisfy (3.15) with  $\bar{C} = 0$ .

**3.3. Properties of the Hamiltonian.** In this section, we establish basic properties of the Hamiltonian and show that is a tracial  $W^*$ -function. We first show that under only the assumption that  $(b_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$  vector-field and  $(\mathcal{L}_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$ -function, the Hamiltonian is a tracial  $W^*$ -function. The additional assumptions will allow the Hamiltonian to be realized in a single von Neumann algebra, and we will adopt Assumption **A** for the remainder after this lemma, although some analogous results may hold under weaker assumptions.

**Lemma 3.4.** *If  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$ -function and  $(b_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$  vector-field, then  $(H_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$ -function.*

*Proof.* Let  $\mathcal{A}, \mathcal{B} \in \mathbb{W}$  and  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  be a tracial  $W^*$ -embedding. Note that if  $\mathcal{C} \in \mathbb{W}$  and  $\kappa : \mathcal{B} \rightarrow \mathcal{C}$  is a tracial  $W^*$ -embedding then  $\kappa \circ \iota : \mathcal{A} \rightarrow \mathcal{C}$  is a tracial  $W^*$ -embedding. Thus for any  $X, P \in L^2(\mathcal{A})_{sa}^d$  and  $\alpha \in \mathbb{A}_{\mathcal{C}}$ , we have

$$H_{\mathcal{A}}(X, P) \geq \left\langle b_{\mathcal{C}}(\kappa \circ \iota X, \alpha), \kappa \circ \iota P \right\rangle_{L^2(\mathcal{C})} - L_{\mathcal{C}}(\kappa \circ \iota X, \alpha).$$

Maximizing over the set of  $(\mathcal{C}, \kappa)$ , we obtain that  $H_{\mathcal{A}}(X, P) \geq H_{\mathcal{B}}(\iota X, \iota P)$ .

To show the converse inequality, suppose that  $\mathcal{B}_2 \in \mathbb{W}$  and  $\iota_2 : \mathcal{A} \rightarrow \mathcal{B}_2$  is a tracial  $W^*$ -embedding. Let  $\mathcal{C} \in \mathbb{W}$  be the free product of  $\mathcal{B}$  and  $\mathcal{B}_2$  with amalgamation over  $\mathcal{A}$ . By Lemma 2.1, there exist tracial  $W^*$ -embeddings  $\phi_1 : \mathcal{B} \rightarrow \mathcal{C}$  and  $\phi_2 : \mathcal{B}_2 \rightarrow \mathcal{C}$  such that  $\phi_1 \circ \iota = \phi_2 \circ \iota_2$ . Note that for any  $\alpha \in \mathbb{A}_{\mathcal{B}_2}$ ,

$$\begin{aligned} \left\langle b_{\mathcal{B}_2}(\iota_2 X, \alpha), \iota_2 P \right\rangle_{\mathcal{B}_2} - L_{\mathcal{B}_2}(\iota_2 X, \alpha) &= \left\langle b_{\mathcal{C}}(\phi_2 \circ \iota_2 X, \phi_2 \alpha), \phi_2 \circ \iota_2 P \right\rangle_{L^2(\mathcal{C})} - L_{\mathcal{C}}(\phi_2 \circ \iota_2 X, \phi_2 \alpha) \\ &= \left\langle b_{\mathcal{C}}(\phi_1 \circ \iota X, \phi_2 \alpha), \phi_1 \circ \iota P \right\rangle_{L^2(\mathcal{C})} - L_{\mathcal{C}}(\phi_1 \circ \iota X, \phi_2 \alpha) \\ &\leq H_{\mathcal{B}}(\iota X, \iota P). \end{aligned}$$

Maximizing over the set of  $(\mathcal{B}_2, \iota_2)$ , we conclude that  $H_{\mathcal{A}}(X, P) \leq H_{\mathcal{B}}(\iota X, \iota P)$ .  $\square$

In the following lemma, we show that the Hamiltonian inherits both structure and estimates from our assumptions. The estimate we obtain in (3.17) below will be indispensable for the theory of viscosity solutions developed in [55, 64].

**Lemma 3.5.** *Under Assumption **A**, the following hold.*

(i) *The Hamiltonian satisfies*

$$H_{\mathcal{A}}(X, P) = \sup_{\alpha \in \mathbb{A}_{\mathcal{A}}} \{ \langle b_{\mathcal{A}}(X, \alpha), P \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(X, \alpha) \} \quad \text{for } X, P \in L^2(\mathcal{A})_{sa}^d. \quad (3.16)$$

(ii)  *$(H_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$ -function (in both variables) which is  $E$ -convex in the second variable.*

(iii)  $H_{\mathcal{A}}$  is uniformly continuous on bounded sets. Also for any  $X, Y \in L^2(\mathcal{A})_{sa}^d$  and  $\gamma > 1$ ,

$$H_{\mathcal{A}}(X, \gamma(X - Y)) - H_{\mathcal{A}}(Y, \gamma(X - Y)) \geq - \left( \frac{2\bar{C} + C_2^2}{2} \gamma \|X - Y\|_{L^2(\mathcal{A})}^2 + \frac{1}{2\gamma} \right). \quad (3.17)$$

*Proof.* (i) We first note that  $\geq$  in (3.16) follows immediately from taking  $\mathcal{B} = \mathcal{A}$  in (3.14). For the other direction, by assumptions (3.4) and (3.7), for any  $X \in L^2(\mathcal{A})_{sa}^d$ ,  $\mathcal{B} \in \mathbb{W}$ , and  $\alpha \in \mathbb{A}_{\mathcal{B}}$ , we have

$$\langle b_{\mathcal{A}}(X, E\alpha), P \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(X, E\alpha) \geq \langle b_{\mathcal{B}}(\iota X, \alpha), \iota P \rangle_{L^2(\mathcal{B})} - L_{\mathcal{B}}(\iota X, \alpha).$$

Thus for every  $\mathcal{B} \in \mathbb{W}$  and  $\alpha \in \mathbb{A}_{\mathcal{B}}$ , we may bound the argument of the supremum in (3.14) using  $E\alpha \in \mathbb{A}_{\mathcal{A}}$ , proving  $\leq$  in (3.16).

(ii) By Lemma 3.4,  $(H_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial function. To prove the  $E$ -convexity of  $(H_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$ , consider  $\mathcal{A}, \mathcal{B} \in \mathbb{W}$  with a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ . Then for any  $X \in L^2(\mathcal{A})_{sa}^d$ ,  $\alpha \in \mathbb{A}_{\mathcal{A}}$ , and  $P \in L^2(\mathcal{B})_{sa}^d$ ,

$$\langle b_{\mathcal{A}}(X, \alpha), EP \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(X, \alpha) = \langle b_{\mathcal{B}}(\iota X, \iota \alpha), P \rangle_{L^2(\mathcal{B})} - L_{\mathcal{B}}(\iota X, \iota \alpha) \leq H_{\mathcal{B}}(\iota X, P).$$

Maximizing over  $\alpha \in \mathbb{A}_{\mathcal{A}}$ , we use (3.16) to conclude that  $H_{\mathcal{A}}(X, EP) \leq H_{\mathcal{B}}(\iota X, P)$ .

(iii) Uniform continuity on bounded sets follows from the uniform continuity on bounded sets of  $b_{\mathcal{A}}$  and  $L_{\mathcal{A}}$  along with the coercive lower bound on  $L_{\mathcal{A}}$  in (3.8), which allows us to restrict to a bounded set for  $\alpha \in \mathbb{A}_{\mathcal{A}}$ . To provide more details, fix some  $M > 0$  and  $X, P \in L^2(\mathcal{A})_{sa}^d$  with  $\max\{\|X\|_{L^2(\mathcal{A})}, \|P\|_{L^2(\mathcal{A})}\} \leq M$ . We first note that the Hamiltonian is bounded below by, using that  $0 \in \mathbb{A}_{\mathcal{A}}$ ,

$$H_{\mathcal{A}}(X, P) \geq \langle b_{\mathcal{A}}(X, 0), P \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(X, 0) \geq -M \|b_{\mathcal{A}}(X, 0)\|_{L^2(\mathcal{A})} - C_1(1 + M),$$

and by (3.6)

$$\|b_{\mathcal{A}}(X, 0)\|_{L^2(\mathcal{A})} \leq v_M.$$

Hence there is a constant  $C_M > 0$  depending only on  $M$  such that

$$H_{\mathcal{A}}(X, P) \geq -C_M. \quad (3.18)$$

In light of (3.16), there exists  $\bar{\alpha} \in \mathbb{A}_{\mathcal{A}}$  such that

$$H_{\mathcal{A}}(X, P) \leq 1 + \langle b_{\mathcal{A}}(X, \bar{\alpha}), P \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(X, \bar{\alpha}).$$

Then by conditions (3.6) and (3.8),

$$H_{\mathcal{A}}(X, P) + \frac{1}{C_1} \|\bar{\alpha}\|_{L^2(\mathcal{A})}^2 \leq 1 + v_M(1 + \|\bar{\alpha}\|_{L^2(\mathcal{A})})M + C_1. \quad (3.19)$$

By (3.18) and (3.19), for some constant  $M'$  that depends only on  $M$ ,

$$\|\bar{\alpha}\|_{L^2(\mathcal{A})} \leq M'. \quad (3.20)$$

We can go back to (3.19) to conclude that  $H_{\mathcal{A}}(X, P) \leq C_M$  (by increasing the value of  $C_M$  if necessary).

Now given  $X_1, P_1, X_2, P_2 \in L^2(\mathcal{A})_{sa}^d$  with  $\max\{\|X_1\|_{L^2(\mathcal{A})}, \|X_2\|_{L^2(\mathcal{A})}, \|P_1\|_{L^2(\mathcal{A})}, \|P_2\|_{L^2(\mathcal{A})}\} \leq M$  and  $\epsilon > 0$ , one can find  $\alpha_1 \in \mathbb{A}_{\mathcal{A}}$  with  $\|\alpha_1\|_{L^2(\mathcal{A})} \leq M'$  (see (3.20)) such that

$$\begin{aligned} H_{\mathcal{A}}(X_1, P_1) - H_{\mathcal{A}}(X_2, P_2) &\leq \epsilon + \langle b_{\mathcal{A}}(X_1, \alpha_1), P_1 \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(X_1, \alpha_1) - \langle b_{\mathcal{A}}(X_2, \alpha_1), P_2 \rangle_{L^2(\mathcal{A})} + L_{\mathcal{A}}(X_2, \alpha_1) \\ &\leq \epsilon + \|b_{\mathcal{A}}(X_1, \alpha_1) - b_{\mathcal{A}}(X_2, \alpha_1)\|_{L^2(\mathcal{A})} \|P_1\|_{L^2(\mathcal{A})} \\ &\quad + \|b_{\mathcal{A}}(X_2, \alpha_1)\|_{L^2(\mathcal{A})} \|P_1 - P_2\|_{L^2(\mathcal{A})} + |L_{\mathcal{A}}(X_1, \alpha_1) - L_{\mathcal{A}}(X_2, \alpha_1)| \\ &\leq \epsilon + (1 + M) C_2 \|X_1 - X_2\|_{L^2(\mathcal{A})} + v_M(1 + M') \|P_1 - P_2\|_{L^2(\mathcal{A})}. \end{aligned}$$

Taking  $\epsilon \rightarrow 0$ , we establish a uniform continuity of  $H_{\mathcal{A}}$  on bounded sets.

To prove (3.17), for given  $\gamma > 1$ ,  $\epsilon > 0$  and  $X, Y \in L^2(\mathcal{A})_{sa}^d$ , let  $\alpha_2 \in \mathbb{A}_{\mathcal{A}}$  be such that

$$H_{\mathcal{A}}(Y, \gamma(X - Y)) \leq \epsilon + \langle b_{\mathcal{A}}(Y, \alpha_2), \gamma(X - Y) \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(Y, \alpha_2).$$

Then

$$\begin{aligned} & H_{\mathcal{A}}(X, \gamma(X - Y)) - H_{\mathcal{A}}(Y, \gamma(X - Y)) \\ & \geq \langle b_{\mathcal{A}}(X, \alpha_2), \gamma(X - Y) \rangle_{L^2(\mathcal{A})} - \langle b_{\mathcal{A}}(Y, \alpha_2), \gamma(X - Y) \rangle_{L^2(\mathcal{A})} - L_{\mathcal{A}}(X, \alpha_2) + L_{\mathcal{A}}(Y, \alpha_2) - \epsilon \\ & \geq -\bar{C}\gamma \|X - Y\|_{L^2(\mathcal{A})}^2 - C_2 \|X - Y\|_{L^2(\mathcal{A})} - \epsilon \geq -\left(\frac{2\bar{C} + C_2^2}{2}\gamma \|X - Y\|_{L^2(\mathcal{A})}^2 + \frac{1}{2\gamma}\right) - \epsilon, \end{aligned}$$

using the Lipschitz bound conditions (3.5) and (3.9). Taking  $\epsilon \rightarrow 0$  completes the proof.  $\square$

We now introduce a theorem on well-posedness for solutions to the SDE. This is essentially [85, Theorem 6.16] in a Hilbert space setting. Much more general results can be readily found, for example, in [36], but the generality complicates things, so we present a simpler setting.

We could use a general Itô formula, such as proven later in Lemma 4.5, but for the theorem we need only to consider the simpler case of the test functions made up of the norm squared, which we can handle explicitly.

**Theorem 3.6.** *Suppose that Assumption **A** holds. Given  $\mathcal{A} \in \mathbb{W}$ ,  $x_0 \in L^2(\mathcal{A}_{t_0})_{sa}^d$ , and  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, t_1}$ , there exists a unique solution  $(X_t)_{t \in [t_0, t_1]} \in C([t_0, t_1]; L^2(\Omega, \mathcal{F}, \mathbb{P}; L^2(\mathcal{A})_{sa}^d))$  to (3.10) with  $X_{t_0} = x_0$  which is adapted in the sense that for every  $s \in [t_0, t_1]$ ,  $X_s \in L^2(\Omega, \mathcal{F}_s, \mathbb{P}; L^2(\mathcal{A}_s)_{sa}^d)$ . Furthermore, letting*

$$M := \|x_0\|_{L^2(\mathcal{A})}^2, \quad N := \mathbb{E}\left[\int_{t_0}^{t_1} \|\alpha_t\|_{L^2(\mathcal{A})}^2 dt\right],$$

there exists  $\tilde{C}_{M,N} > 0$ , depending only on  $M$  and  $N$ , such that for  $t_0 \leq t \leq t_1$ ,

$$\mathbb{E}\left[\|X_t - x_0\|_{L^2(\mathcal{A})}^2\right] \leq \tilde{C}_{M,N}(t - t_0). \quad (3.21)$$

Also, there exists  $\tilde{C} > 0$  depending only on  $\bar{C}$  (see (3.5)) such that if  $\hat{X}$  is a solution with another random initial condition  $\hat{x}_0 \in L^2(\Omega, \mathcal{F}_{t_0}, \mathbb{P}; L^2(\mathcal{A}_{t_0})_{sa}^d)$  and the same control, then

$$\mathbb{E}\left[\|X_t - \hat{X}_t\|_{L^2(\mathcal{A})}^2\right] \leq \tilde{C} \mathbb{E}[\|x_0 - \hat{x}_0\|_{L^2(\mathcal{A})}^2]. \quad (3.22)$$

*Proof.* Using a standard Picard iteration method, we can find a unique solution for small times that is adapted to the filtration; see Appendix A.2. Let  $\delta > 0$  be arbitrary and recall the constant  $v_M$  from (3.6). Then, using Assumption **A**,

$$\begin{aligned} & \frac{d}{dt} \mathbb{E}\left[\|X_t - x_0\|_{L^2(\mathcal{A})}^2\right] \\ & = 2\mathbb{E}\left[\langle X_t - x_0, b_{\mathcal{A}}(X_t, \alpha_t) \rangle_{L^2(\mathcal{A})}\right] + d(\beta_C^2 + \beta_F^2) \\ & = 2\mathbb{E}\left[\langle X_t - x_0, b_{\mathcal{A}}(X_t, \alpha_t) - b_{\mathcal{A}}(x_0, \alpha_t) \rangle_{L^2(\mathcal{A})} + \langle X_t - x_0, b_{\mathcal{A}}(x_0, \alpha_t) \rangle_{L^2(\mathcal{A})}\right] + d(\beta_C^2 + \beta_F^2) \\ & \leq 2\mathbb{E}\left[\bar{C}\|X_t - x_0\|_{L^2(\mathcal{A})}^2\right] + \|X_t - x_0\|_{L^2(\mathcal{A})} v_M (1 + \|\alpha_t\|_{L^2(\mathcal{A})}) + d(\beta_C^2 + \beta_F^2) \\ & \leq (2\bar{C} + 1 + \delta) \mathbb{E}[\|X_t - x_0\|_{L^2(\mathcal{A})}^2] + \frac{v_M^2}{\delta} \mathbb{E}[\|\alpha_t\|_{L^2(\mathcal{A})}^2] + v_M^2 + d(\beta_C^2 + \beta_F^2), \end{aligned}$$

where the first identity follows from Appendix A.2. By Grönwall's inequality,

$$\mathbb{E} \left[ \|X_t - x_0\|_{L^2(\mathcal{A})}^2 \right] \leq y_t,$$

where  $y_t$  is defined by, setting  $r := 2\bar{C} + 1$  and  $D_M := v_M^2 + d(\beta_C^2 + \beta_F^2)$ ,

$$\begin{aligned} y_t &:= e^{(r+\delta)(t-t_0)} \left( \int_{t_0}^t e^{-(r+\delta)(s-t_0)} \left( \frac{v_M^2}{\delta} \mathbb{E} [\|\alpha_s\|_{L^2(\mathcal{A})_{sa}}^2] + D_M \right) ds \right) \\ &\leq e^{(r+\delta)(t-t_0)} \left( \frac{v_M^2}{\delta} N + (t-t_0) D_M \right). \end{aligned}$$

Choosing  $\delta := \frac{1}{t-t_0}$ , we arrive at (3.21) for some constant  $\tilde{C}_{M,N}$ .

To show (3.22), note that

$$\begin{aligned} \frac{d}{dt} \mathbb{E} [\|X_t - \hat{X}_t\|_{L^2(\mathcal{A})}^2] &= 2\mathbb{E} [\langle X_t - \hat{X}_t, b_{\mathcal{A}}(X_t, \alpha_t) - b_{\mathcal{A}}(\hat{X}_t, \alpha_t) \rangle_{L^2(\mathcal{A})}] \\ &\leq 2\mathbb{E} [\|X_t - \hat{X}_t\|_{L^2(\mathcal{A})} \|b(X_t, \alpha_t) - b(\hat{X}_t, \alpha_t)\|_{L^2(\mathcal{A})}] \\ &\leq 2\bar{C} \mathbb{E} [\|X_t - \hat{X}_t\|_{L^2(\mathcal{A})}^2]. \end{aligned}$$

Then (3.22) follows from the Grönwall's inequality.  $\square$

**3.4. Properties of the value function.** In this section, we prove well-definedness and continuity properties for the value function, including showing that  $\bar{V}$  is well-defined on the space of non-commutative laws.

We begin by proving an auxiliary lemma about amalgamating several different non-commutative filtrations and associated Brownian motions. Of course, the analogs of these lemmas in classical probability would be proved using conditional distributions, but for the non-commutative setting we will use amalgamated free products. In the following lemma, we describe how to glue together different choices of filtration and free Brownian motion on an interval  $[t_1, T]$  together with a fixed choice of filtration and free Brownian motion on  $[t_0, t_1]$ , where  $0 \leq t_0 \leq t_1 \leq T$ . While the general case is used in Proposition 4.8, the applications in this section are mostly in the case  $t_0 = t_1$ . These lemmas require significant technical effort to prove, so in order to maintain the flow of this section, we include detailed arguments in the appendix (§C.2).

**Lemma 3.7** (Lemma C.5). *Let  $0 \leq t_0 \leq t_1 \leq T$  and  $K$  be any index set. Let  $\mathcal{A}$  be a tracial von Neumann algebra equipped with a filtration  $(\mathcal{A}_t)_{t \in [t_0, t_1]}$  and a compatible  $d$ -variable free Brownian motion  $(S_t^0)_{t \in [t_0, t_1]}$ . Let  $(\mathcal{B}^k)_{k \in K}$  be a family of tracial von Neumann algebras and let  $\iota_k : \mathcal{A} \rightarrow \mathcal{B}^k$  be a tracial  $W^*$ -embedding. Suppose each  $\mathcal{B}^k$  has a filtration  $(\mathcal{B}_t^k)_{t \in [t_1, T]}$  and a compatible  $d$ -variable free Brownian motion  $(S_t^k)_{t \in [t_1, T]}$ . Assume that  $\iota_k(\mathcal{A}_{t_1}) \subseteq \mathcal{B}_{t_1}^k$ .*

*Then there exists an algebra  $\mathcal{B}$ , a tracial  $W^*$ -embeddings  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  and  $\tilde{\iota}_k : \mathcal{B}^k \rightarrow \mathcal{B}$ , a filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$ , and a compatible  $d$ -variable free Brownian motion  $(S_t)_{t \in [t_0, T]}$  such that the following hold:*

- (1) *We have  $\tilde{\iota}_k \circ \iota_k|_{\mathcal{A}_{t_0}} = \iota|_{\mathcal{A}_{t_0}}$  for each  $k \in K$ .*
- (2) *We have  $\iota(S_t^0) = S_t$  for  $t \in [t_0, t_1]$ .*
- (3) *We have  $\tilde{\iota}_k(S_t^k) = S_t - S_{t_1}$  for  $t \in [t_1, T]$ .*
- (4) *We have  $\iota(\mathcal{A}_t) \subseteq \mathcal{B}_t$  for  $t \in [t_0, t_1]$ .*
- (5) *We have  $\tilde{\iota}_k(\mathcal{B}_t^k) \subseteq \mathcal{B}_t$  for  $t \in [t_1, T]$ .*

*Furthermore, for any given  $k_0 \in K$ , the embedding  $\iota$  can be taken to be  $\iota = \tilde{\iota}_{k_0} \circ \iota_{k_0}$ .*

The following is an important application of the previous lemma, which will allow us to choose filtrations in  $\mathcal{B}$  such that  $\mathcal{B}_{t_0} \supseteq \iota(\mathcal{A})$ . In turn, the freedom in choosing the initial subalgebra of filtration provided by Lemma 3.8 will help us to apply Lemma 3.7.

**Lemma 3.8** (Lemma C.6). *Suppose that Assumption A holds. Let  $\mathcal{A}$  be a tracial  $W^*$ -algebra, let  $x_0 \in L^2(\mathcal{A})_{sa}^d$ , and let  $t_0 \in [0, T]$ . Then for every  $\epsilon > 0$ , there exists a tracial  $W^*$ -algebra  $\mathcal{B}$ , a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , and a control policy  $\alpha \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, T}$  associated with a filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$  and compatible  $d$ -variable free Brownian motion  $(S_t)_{t \in [t_0, T]}$ , such that*

- (1)  $\iota(\mathcal{A}) \subseteq \mathcal{B}_{t_0}$
- (2)  $\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{B}}(X_T[\alpha]) \right] \leq \bar{V}_{\mathcal{A}}(t_0, x_0) + \epsilon.$

With Lemmas 3.7 and 3.8, we can now prove various properties of the value function that we want, starting with consistency condition of being a tracial  $W^*$ -function.

**Lemma 3.9.** *Suppose that Assumption A holds. Then for all  $t_0 \in [0, T]$ ,  $(\bar{V}_{\mathcal{A}}(t_0, \cdot))_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$ -function.*

*Proof.* Let  $\mathcal{A}, \mathcal{B} \in \mathbb{W}$  and  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  be a tracial  $W^*$ -embedding. Note that if  $\mathcal{C} \in \mathbb{W}$  and  $\kappa : \mathcal{B} \rightarrow \mathcal{C}$  is a tracial  $W^*$ -embedding then  $\kappa \circ \iota : \mathcal{A} \rightarrow \mathcal{C}$  is also a tracial  $W^*$ -embedding and thus,  $\bar{V}_{\mathcal{A}}(t_0, x_0) \leq \tilde{V}_{\mathcal{C}}(t_0, \kappa \circ \iota x_0)$  for any  $x_0 \in L^2(\mathcal{A})_{sa}^d$ . Minimizing over all  $\mathcal{C} \in \mathbb{W}$  and  $\kappa : \mathcal{B} \rightarrow \mathcal{C}$ , we conclude that  $\bar{V}_{\mathcal{A}}(t_0, x_0) \leq \bar{V}_{\mathcal{B}}(t_0, \iota x_0)$ .

We now show the reverse direction. Given  $\epsilon > 0$ , there exists  $\mathcal{B}^1 \in \mathbb{W}$  and a tracial  $W^*$ -embedding  $\iota_1 : \mathcal{A} \rightarrow \mathcal{B}^1$  such that

$$\tilde{V}_{\mathcal{B}^1}(t_0, \iota_1 x_0) < \bar{V}_{\mathcal{A}}(t_0, x_0) + \epsilon. \quad (3.23)$$

Let  $\tilde{\alpha}^1 = ((\alpha_t^1)_t, (\mathcal{B}_t^1)_t, (S_t^1)_t) \in \mathbb{A}_{\mathcal{B}^1, \iota_1 x_0}^{t_0, T}$  be a control policy such that

$$\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}^1}(X_t[\tilde{\alpha}^1], \alpha_t^1) dt + g_{\mathcal{B}^1}(X_T[\tilde{\alpha}^1]) \right] < \tilde{V}_{\mathcal{B}^1}(t_0, \iota_1 x_0) + \epsilon, \quad (3.24)$$

where  $X_t[\tilde{\alpha}^1]$  denotes a solution to (3.10) (with  $\mathcal{A}$  replaced by  $\mathcal{B}^1$ ) on  $L^2(\mathcal{B}^1)_{sa}^d$  with the initial condition  $X_{t_0}[\tilde{\alpha}^1] = \iota_1 x_0$ .

Using Lemma 2.1, one can take a tracial  $W^*$ -algebra  $\mathcal{C}$  and tracial  $W^*$ -embeddings  $\phi : \mathcal{B} \rightarrow \mathcal{C}$  and  $\phi_1 : \mathcal{B}^1 \rightarrow \mathcal{C}$  such that  $\phi \circ \iota = \phi_1 \circ \iota_1$ . In light of Remark 3.3, setting

$$X_t := \phi_1(X_t[\tilde{\alpha}^1]) \in C([t_0, T]; L^2(\Omega, \mathcal{F}, \mathbb{P}; L^2(\mathcal{C})_{sa}^d)),$$

$$\tilde{\alpha} := ((\phi_1 \alpha_t^1)_t, (\phi_1 \mathcal{B}_t^1)_t, (\phi_1 S_t^1)_t) \in \mathbb{A}_{\mathcal{C}, \phi_1(\iota_1 x_0)}^{t_0, T}$$

(i.e. in particular  $\alpha_t := \phi_1 \alpha_t^1$ ) and

$$S_t := \phi_1 S_t^1 \quad \text{for } t \in [t_0, T],$$

we have

$$\begin{cases} dX_t = b_{\mathcal{C}}(X_t, \alpha_t) dt + \beta_{\mathcal{C}} \mathbb{1}_{\mathcal{C}} dW_t^0 + \beta_F dS_t, \\ X_{t_0} = \phi_1(\iota_1 x_0). \end{cases}$$

Thus interchangeably using the notations  $X_t$  and  $X_t[\tilde{\alpha}]$ ,

$$\begin{aligned} \tilde{V}_{\mathcal{C}}(t_0, \phi_1(\iota_1 x_0)) &\leq \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{C}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{C}}(X_T[\tilde{\alpha}]) \right] \\ &= \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}^1}(X_t[\tilde{\alpha}^1], \alpha_t^1) dt + g_{\mathcal{B}^1}(X_T[\tilde{\alpha}^1]) \right]. \end{aligned}$$

Therefore using this along with the fact that  $\phi \circ \iota = \phi_1 \circ \iota_1$ , we conclude that

$$\tilde{V}_{\mathcal{C}}(t_0, \phi(\iota x_0)) \stackrel{(3.24)}{\leq} \tilde{V}_{\mathcal{B}^1}(t_0, \iota_1 x_0) + \epsilon \stackrel{(3.23)}{\leq} \bar{V}_{\mathcal{A}}(t_0, x_0) + 2\epsilon.$$

Since  $\phi : \mathcal{B} \rightarrow \mathcal{C}$  is a tracial  $W^*$ -embedding,  $\bar{V}_{\mathcal{B}}(t_0, \iota x_0) \leq 2\epsilon + \bar{V}_{\mathcal{A}}(t_0, x_0)$ . By the arbitrariness of  $\epsilon > 0$ , we infer  $\bar{V}_{\mathcal{B}}(t_0, \iota x_0) \leq \bar{V}_{\mathcal{A}}(t_0, x_0)$ .  $\square$

In the next lemma, we verify that two versions of values functions (3.11) and (3.12) are same, under the additional Assumption **B**, provided that  $\mathcal{A}$  admits a free Brownian motion freely independent of the initial condition. To clarify the terminology in the statement below, for a function  $f : X \rightarrow \mathbb{R}$  and  $X_0 \subseteq X$ , we say that the infimum  $\inf_{x \in X} f(x)$  is *witnessed by*  $X_0$  if  $\inf_{x \in X_0} f(x) = \inf_{x \in X} f(x)$ , but note that this does not require that the infimum is *achieved* in  $X_0$  or even in  $X$ .

**Lemma 3.10.** *Suppose that Assumptions **A** and **B** hold.*

1. *Let  $\mathcal{A} \in \mathbb{W}$ , and let  $x_0 \in L^2(\mathcal{A})_{\text{sa}}^d$  and  $t_0 \in [0, T]$ . Suppose that  $\mathcal{A}$  admits a  $d$ -variable free Brownian motion  $(S_t^0)_{t \in [t_0, T]}$  freely independent of  $W^*(x_0)$ . For  $t \in [t_0, T]$ , let*

$$\mathcal{A}_t^0 := W^*(x_0, (S_s^0)_{s \in [t_0, t]}).$$

*Then we have*

$$\bar{V}_{\mathcal{A}}(t_0, x_0) = \tilde{V}_{\mathcal{A}}(t_0, x_0),$$

*and the infimum in (3.11) is witnessed by control policies  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, T}$  that use the given filtration  $(\mathcal{A}_t^0)_{t \in [t_0, T]}$  and free Brownian motion  $(S_t^0)_{t \in [t_0, T]}$ .*

2. *Let  $\mathcal{A} \in \mathbb{W}$ . Let  $\mathcal{C}$  be a tracial von Neumann algebra generated by a  $d$ -variable free Brownian motion  $(S_t^1)_{t \in [t_0, T]}$ , and define  $\mathcal{C}_{t_0, t} := W^*(S_s^1 : s \in [t_0, t])$  for  $t \in [t_0, T]$ . Let  $\iota_1 : \mathcal{A} \rightarrow \mathcal{A} * \mathcal{C}$  and  $\iota_2 : \mathcal{C} \rightarrow \mathcal{A} * \mathcal{C}$  be the inclusions associated to the free product. Then for any  $x_0 \in L^2(\mathcal{A})_{\text{sa}}^d$ ,*

$$\bar{V}_{\mathcal{A}}(t_0, x_0) = \tilde{V}_{\mathcal{A} * \mathcal{C}}(t_0, \iota_1(x_0)),$$

*and the infimum is witnessed by control policies that use the fixed filtration  $(\iota_1(\mathcal{A}) \vee \iota_2(\mathcal{C}_{t_0, t}))_{t \in [t_0, T]}$  and free Brownian motion  $(\iota_2(S_t^1))_{t \in [t_0, T]}$ .*

**Proof. Step 1: Conditioning control policy.** Let  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  be a tracial  $W^*$ -embedding and let  $\tilde{\alpha} = ((\alpha_t)_{t \in [t_0, T]}, (\mathcal{B}_t)_{t \in [t_0, T]}, (S_t)_{t \in [t_0, T]})$  be a control policy in  $\mathcal{B}$ . Let

$$\mathcal{B}_t^0 := W^*(\iota(x_0), (S_s)_{s \in [t_0, t]});$$

it is straightforward to check that the Brownian motion  $(S_t)_{t \in [t_0, T]}$  is compatible with the filtration  $(\mathcal{B}_t^0)_{t \in [t_0, T]}$ . For  $t \in [t_0, T]$ , let  $E_{\mathcal{B}_t^0} : \mathcal{B} \rightarrow \mathcal{B}_t^0$  be the trace-preserving conditional expectation, and define

$$\alpha_t^0 := E_{\mathcal{B}_t^0}[\alpha_t].$$

We claim that  $\alpha_t^0$  is measurable as a function on  $[t_0, T]$  times the underlying probability space. Since the Hilbert space  $L^2(\mathcal{B})$  is assumed to be separable, it suffices to check weak measurability, so fix some  $y \in L^2(\mathcal{B})$  and by linearity we can assume without loss of generality that  $y$  is self-adjoint. Then

$$\langle \alpha_t^0, y \rangle_{L^2(\mathcal{B})} = \langle E_{\mathcal{B}_t^0}[\alpha_t], y \rangle_{L^2(\mathcal{B})} = \langle \alpha_t, E_{\mathcal{B}_t^0}[y] \rangle_{L^2(\mathcal{B})}.$$

Since  $t \mapsto \alpha_t$  is measurable, it suffices to check that  $t \mapsto E_{\mathcal{B}_t^0}[y]$  is weakly measurable. Let  $z \in L^2(\mathcal{B})_{\text{sa}}$ . Then,

$$\langle E_{\mathcal{B}_t^0}[y], z \rangle_{L^2(\mathcal{B})} = \langle E_{\mathcal{B}_t^0}[y], E_{\mathcal{B}_t^0}[z] \rangle_{L^2(\mathcal{B})} = \frac{1}{4} \left( \|E_{\mathcal{B}_t^0}[x + y]\|_{L^2(\mathcal{B})}^2 - \|E_{\mathcal{B}_t^0}[x - y]\|_{L^2(\mathcal{B})}^2 \right),$$



and  $\|E_{\mathcal{B}_t^0}[x+y]\|_{L^2(\mathcal{B})}^2$  along with  $\|E_{\mathcal{B}_t^0}[x-y]\|_{L^2(\mathcal{B})}^2$  are increasing functions of  $t$ , hence measurable. Hence,  $t \mapsto \alpha_t^0$  is measurable as desired, and by construction it is adapted to  $(\mathcal{B}_t^0)_t$ .

**Step 2: Conditioning trajectory.** Let  $\tilde{\alpha}^0 = ((\alpha_t^0)_{t \in [t_0, T]}, (\mathcal{B}_t^0)_{t \in [t_0, T]}, (S_t)_{t \in [t_0, T]})$  be the control policy in  $\mathcal{B}$ . We claim that

$$X_t[\tilde{\alpha}^0] = E_{\mathcal{B}_t^0}(X_t[\tilde{\alpha}]),$$

which in particular implies that the initial condition is  $X_{t_0}[\tilde{\alpha}^0] = X_{t_0}[\tilde{\alpha}] = \iota(x_0)$ . Note that

$$E_{\mathcal{B}_t^0}[X_t[\tilde{\alpha}]] = x_0 + \int_{t_0}^t E_{\mathcal{B}_s^0}[\alpha_s] ds + \beta_C \mathbb{1}_{\mathcal{A}} (W_t^0 - W_{t_0}^0) + \beta_F (S_t - S_{t_0}) = X_t[\tilde{\alpha}],$$

while

$$X_t[\tilde{\alpha}^0] = x_0 + \int_{t_0}^t E_{\mathcal{B}_s^0}[\alpha_s] ds + \beta_C \mathbb{1}_{\mathcal{A}} (W_t^0 - W_{t_0}^0) + \beta_F (S_t - S_{t_0}) = X_t[\tilde{\alpha}].$$

Hence, it suffices to show that for  $s \leq t$ , we have

$$E_{\mathcal{B}_t^0}[\alpha_s] = E_{\mathcal{B}_s^0}[\alpha_s].$$

For this, we use a certain fact about conditional expectations inside free products (Lemma C.3). Let  $\mathcal{D} := W^*(S_{s'} - S_s : s' \in [s, t])$ , and recall that  $\mathcal{D}$  is freely independent of  $\mathcal{B}_s$  since the free Brownian motion  $(S_t)_{t \in [t_0, T]}$  is compatible with the filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$ . Thus,  $\mathcal{B}_s \vee \mathcal{D} \cong \mathcal{B}_s * \mathcal{D}$  (Corollary 2.2); moreover, the subalgebra  $\mathcal{B}_t^0 = \mathcal{B}_s^0 \vee \mathcal{D}$  in  $\mathcal{B}$  corresponds to the subalgebra  $\mathcal{B}_s^0 * \mathcal{D}$  in  $\mathcal{B}_s * \mathcal{D}$ . Let  $\bar{\iota} : \mathcal{B}_s \rightarrow \mathcal{B}_s * \mathcal{D}$  be the canonical inclusion into the free product; then Lemma C.3 shows that

$$E_{\mathcal{B}_s^0 * \mathcal{D}} \circ \bar{\iota}(\alpha_s) = \bar{\iota} \circ E_{\mathcal{B}_s^0}[\alpha_s] \text{ in } \mathcal{B}_s * \mathcal{D},$$

hence

$$E_{\mathcal{B}_t^0}[\alpha_s] = E_{\mathcal{B}_s^0}[\alpha_s] \text{ in } \mathcal{B},$$

which proves our claim.

**Step 3: Comparing the value functions.** For the control policy  $\tilde{\alpha}^0$  defined above, by  $E$ -convexity of  $(L_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  and  $(g_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  (see Assumption B),

$$\begin{aligned} \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}^0], \alpha_t^0) dt + g_{\mathcal{B}}(X_T[\tilde{\alpha}]) \right] &= \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(E_{\mathcal{B}_t^0}[X_t[\tilde{\alpha}^0]], E_{\mathcal{B}_t^0}[\alpha_t^0]) dt + g_{\mathcal{B}}(E_{\mathcal{B}_T^0}[X_T[\tilde{\alpha}^0]]) \right] \\ &\leq \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{B}}(X_T[\tilde{\alpha}]) \right]. \end{aligned}$$

**Step 4: Transformation from  $\mathcal{B}_t^0$  to  $\mathcal{A}_t^0$ .** Recall that

$$\mathcal{B}_T^0 = W^*(\iota(x_0), (S_t)_{t \in [t_0, T]}) \subseteq \mathcal{B}, \quad \mathcal{A}_T^0 = W^*(x_0, (S_t^0)_{t \in [t_0, T]}) \subseteq \mathcal{A}.$$

Because the tracial von Neumann algebra generated by a free Brownian motion is unique up to a canonical isomorphism, and the free product of two given tracial von Neumann algebras is unique up to a canonical isomorphism (Lemma 2.1), similar to the beginning of the proof of Lemma 3.7, we conclude there is a unique tracial  $W^*$ -isomorphism  $\phi : \mathcal{B}_T^0 \rightarrow \mathcal{A}_T^0$  such that  $\phi(\iota(x_0)) = x_0$  and  $\phi(S_t) = S_t^0$  for  $t \in [t_0, T]$ . In particular, we also have  $\phi(\mathcal{B}_t^0) = \mathcal{A}_t^0$  for  $t \in [t_0, T]$ . Let  $\alpha_t' := \phi(\alpha_t^0)$  for  $t \in [t_0, T]$ , which is now a process in  $\mathcal{A}_T^0 \subseteq \mathcal{A}$ ; and consider the control policy  $\tilde{\alpha}' = ((\alpha_t')_{t \in [t_0, T]}, (\mathcal{A}_t^0)_{t \in [t_0, T]}, (S_t^0)_{t \in [t_0, T]})$  in  $\mathcal{A}$ . It is immediate that  $X_t[\tilde{\alpha}'] = \phi(X_t[\tilde{\alpha}^0])$ . Since  $L$  and  $g$  are tracial  $W^*$ -functions, the value achieved by the control policy  $\tilde{\alpha}'$  is the same as the value achieved by  $\tilde{\alpha}^0$ . Therefore, we have shown that the infimum

for  $\bar{V}_{\mathcal{B}}(t_0, \iota(x_0))$  is witnessed by control policies in  $\mathcal{A}$  that use the given free Brownian motion  $(S_t^0)_{t \in [t_0, T]}$  and filtration  $(\mathcal{A}_t^0)_{t \in [t_0, T]}$ , and so in particular  $\bar{V}_{\mathcal{A}}(t_0, x_0) = \bar{V}_{\mathcal{A}}(t_0, x_0)$ .

**Step 5: Proof of second claim.** Finally, we prove the second part of Lemma. Since the free Brownian motion  $(\iota_2(S_t^1))_{t \in [t_0, T]}$  is freely independent of the initial condition  $\iota_1(x_0)$ , the preceding argument applies and shows that the infimum in the definition of  $\bar{V}_{\mathcal{A} \ast \mathcal{C}}(t_0, \iota_1(x_0))$  is witnessed by control policies using the given free Brownian motion  $(\iota_2(S_t^1))_{t \in [t_0, T]}$  and the filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$  given by  $\mathcal{B}_t = W^*(\iota_1(x_0), (\iota_2(S_s^1))_{s \in [t_0, t]})$ . Of course,  $\mathcal{B}_t \subseteq \iota_1(\mathcal{A}) \vee \iota_2(\mathcal{C}_{t_0, t})$ , and the free Brownian motion  $(\iota_2(S_t^1))_{t \in [t_0, T]}$  is still compatible with this larger filtration. Hence, any control policy using this free Brownian motion and the smaller filtration is also a valid control policy with respect to the larger filtration.  $\square$

**Lemma 3.11.** *Suppose Assumptions A and B hold. Then for any  $t_0 \in [0, T]$ ,  $(\bar{V}_{\mathcal{A}}(t_0, \cdot))_{\mathcal{A} \in \mathbb{W}}$  is  $E$ -convex.*

*Proof. Step 1: Convexity of  $x \mapsto \bar{V}_{\mathcal{A}}(t_0, x)$ .* Let  $\mathcal{A} \in \mathbb{W}$  be any tracial  $W^*$ -algebra and  $t_0 \in [0, T]$ . Fix  $x_0, x_1 \in L^2(\mathcal{A})_{sa}^d$  and  $\epsilon > 0$ . By Lemma 3.8, for  $k = 0, 1$ , there exist  $\mathcal{B}^k \in \mathbb{W}$ , a tracial  $W^*$ -embedding  $\iota_k : \mathcal{A} \rightarrow \mathcal{B}^k$ , and a control policy

$$\tilde{\alpha}^k := \left( (\alpha_t^k)_{t \in [t_0, T]}, (\mathcal{B}_t^k)_{t \in [t_0, T]}, (S_t^k)_{t \in [t_0, T]} \right) \in \mathbb{A}_{\mathcal{B}^k, \iota_k(x_k)}^{t_0, T} \quad (3.25)$$

such that  $\iota_k(\mathcal{A}) \subseteq \mathcal{B}_{t_0}^k$ ,  $X_{t_0}[\tilde{\alpha}^k] = \iota_k(x_k)$  and

$$\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}^k}(X_t[\tilde{\alpha}^k], \alpha_t^k) dt + g_{\mathcal{B}^k}(X_T[\tilde{\alpha}^k]) \right] \leq \bar{V}_{\mathcal{A}}(t_0, x_k) + \epsilon. \quad (3.26)$$

Now we apply Lemma 3.7 with  $t_0 = t_1$  and  $\mathcal{A}_{t_0} = \mathcal{A}_{t_1} = \mathcal{A}$  (since the time interval has length zero, the free Brownian motion on  $[t_0, t_1]$  reduces to 0). Recalling  $\iota_k(\mathcal{A}) \subset \mathcal{B}_{t_0}^k$ , by Lemma 3.7, there exists a tracial  $W^*$ -algebra  $\mathcal{B}$ , tracial  $W^*$ -embeddings  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ ,  $\tilde{\iota}_k : \mathcal{B}^k \rightarrow \mathcal{B}$ , and a filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$  with a  $d$ -variable free Brownian motion  $(S_t)_{t \in [t_0, T]}$  such that

- (1)  $\tilde{\iota}_k(S_t^k) = S_t$  for each  $t \in [t_0, T]$ .
- (2)  $\tilde{\iota}_k(\mathcal{B}_t^k) \subset \mathcal{B}_t$  for each  $t \in [t_0, T]$ .
- (3)  $(S_t)_{t \in [t_0, T]}$  is a Brownian motion compatible with the filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$ .
- (4)  $\tilde{\iota}_k \circ \iota_k|_{\mathcal{A}} = \iota|_{\mathcal{A}}$ .

By (3.25),

$$\tilde{\iota}_k \tilde{\alpha}^k = \left( (\tilde{\iota}_k \alpha_t^k)_{t \in [t_0, T]}, (\tilde{\iota}_k \mathcal{B}_t^k)_{t \in [t_0, T]}, (\tilde{\iota}_k S_t^k)_{t \in [t_0, T]} \right) \in \mathbb{A}_{\tilde{\iota}_k \mathcal{B}^k, \tilde{\iota}_k \circ \iota_k(x_k)}^{t_0, T}.$$

Since  $(\tilde{\iota}_k \alpha_t^k)_{t \in [t_0, s]} \in L^2([t_0, s] \times (\Omega, (\tilde{\iota}_k \mathcal{B}_t^k)_t, \mathbb{P}))$ , (2) implies that  $(\tilde{\iota}_k \alpha_t^k)_{t \in [t_0, s]} \in L^2([t_0, s] \times (\Omega, (\mathcal{B}_t)_t, \mathbb{P}))$ . This, together with (2) and (3), implies that

$$\beta^k := \left( (\tilde{\iota}_k \alpha_t^k)_{t \in [t_0, T]}, (\mathcal{B}_t)_{t \in [t_0, T]}, (S_t)_{t \in [t_0, T]} \right) \in \mathbb{A}_{\mathcal{B}, \tilde{\iota}_k \circ \iota_k(x_k)}^{t_0, T}.$$

Observe that

$$X_t[\beta^k] = X_t[\tilde{\iota}_k(\tilde{\alpha}^k)] = \tilde{\iota}_k(X_t[\tilde{\alpha}^k]), \quad X_{t_0}[\beta^k] = \tilde{\iota}_k \circ \iota_k(x_k).$$

Since  $L$  and  $g$  are tracial functions, we have

$$\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\beta^k], \tilde{\iota}_k \alpha_t^k) dt + g_{\mathcal{B}}(X_T[\beta^k]) \right] = \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}^k}(X_t[\tilde{\alpha}^k], \alpha_t^k) dt + g_{\mathcal{B}^k}(X_T[\tilde{\alpha}^k]) \right]$$

$$\stackrel{(3.26)}{\leq} \bar{V}_{\mathcal{A}}(t_0, x_k) + \epsilon. \quad (3.27)$$

For  $s \in (0, 1)$ , we set

$$\tilde{\beta}^s := \left( (\beta_t^s)_{t \in [t_0, T]}, (\mathcal{B}_t)_{t \in [t_0, T]}, (S_t)_{t \in [t_0, T]} \right), \quad \beta_t^s := (1-s)\tilde{\iota}_0\alpha_t^0 + s\tilde{\iota}_1\alpha_t^1,$$

and

$$z_s := (1-s)\tilde{\iota}_0 \circ \iota_0(x_0) + s\tilde{\iota}_1 \circ \iota_1(x_1).$$

Since  $\tilde{\beta}^s \in \mathbb{A}_{\mathcal{B}, z_s}^{t_0, T}$ , we infer that

$$\bar{V}_{\mathcal{B}}(t_0, z_s) \leq \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\tilde{\beta}^s], \beta_t^s) dt + g_{\mathcal{B}}(X_T[\tilde{\beta}^s]) \right].$$

We use the fact that

$$X_t[\tilde{\beta}^s] = (1-s)X_t[\beta^0] + sX_t[\beta^1], \quad \beta_t^s = (1-s)\tilde{\iota}_0\alpha_t^0 + s\tilde{\iota}_1\alpha_t^1, \quad X_{t_0}[\tilde{\beta}^s] = z_s$$

along with the convexity property of  $L_{\mathcal{B}}$  and  $g_{\mathcal{B}}$  to conclude that

$$\begin{aligned} \bar{V}_{\mathcal{B}}(t_0, z_s) &\leq (1-s)\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\beta^0], \tilde{\iota}_0\alpha_t^0) dt + g_{\mathcal{B}}(X_T[\beta^0]) \right] \\ &\quad + s\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\beta^1], \tilde{\iota}_1\alpha_t^1) dt + g_{\mathcal{B}}(X_T[\beta^1]) \right]. \end{aligned}$$

Combining this with (3.27), we obtain

$$\bar{V}_{\mathcal{B}}(t_0, z_s) \leq (1-s)\bar{V}_{\mathcal{A}}(t_0, x_0) + s\bar{V}_{\mathcal{A}}(t_0, x_1) + \epsilon. \quad (3.28)$$

Since  $\tilde{\iota}_k \circ \iota_k(x_k) = \iota(x_k)$ ,

$$z_s = (1-s)\iota(x_0) + s\iota(x_1) = \iota((1-s)x_0 + sx_1).$$

Since  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  is a tracial  $W^*$ -embedding and  $\bar{V}$  is a tracial  $W^*$ -function (see Lemma 3.9), (3.28) implies that

$$\bar{V}_{\mathcal{A}}(t_0, (1-s)x_0 + sx_1) \leq (1-s)\bar{V}_{\mathcal{A}}(t_0, x_0) + s\bar{V}_{\mathcal{A}}(t_0, x_1) + \epsilon.$$

Since  $\epsilon > 0$  was arbitrary, we have shown convexity of  $\bar{V}_{\mathcal{A}}(t_0, \cdot)$ .

**Step 2:  $E$ -convexity of  $(\bar{V}_{\mathcal{A}}(t_0, \cdot))_{\mathcal{A} \in \mathbb{W}}$ .** Since we already showed that  $\bar{V}$  is a tracial  $W^*$ -function (Lemma 3.9) and we showed convexity of  $\bar{V}_{\mathcal{A}}$  in Step 1, it only remains to show that for any tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  and its adjoint  $E : \mathcal{B} \rightarrow \mathcal{A}$ , we have

$$\bar{V}_{\mathcal{A}}(t_0, Ey_0) \leq \bar{V}_{\mathcal{B}}(t_0, y_0), \quad \forall y_0 \in L^2(\mathcal{B})_{sa}^d. \quad (3.29)$$

As in Lemma 3.10, let  $\mathcal{C}$  be the tracial  $W^*$ -algebra generated by a free Brownian motion  $(S_t)_{t \in [t_0, T]}$ . Let  $\bar{\iota}_1 : \mathcal{B} \rightarrow \mathcal{B} * \mathcal{C}$  and  $\bar{\iota}_2 : \mathcal{C} \rightarrow \mathcal{B} * \mathcal{C}$  be the inclusions from the free product construction. By Lemma 3.10, for each  $\epsilon > 0$ , there exists a control policy of the form

$$\tilde{\alpha} = ((\alpha_t)_{t \in [t_0, T]}, (\bar{\iota}_1(\mathcal{A}) \vee \bar{\iota}_2(\mathcal{C}_{t_0, t}))_{t \in [t_0, T]}, (\bar{\iota}_2(S_t))_{t \in [t_0, T]}),$$

where  $\mathcal{C}_{t_0, t} := W^*(S_s : s \in [t_0, t])$ , such that

$$\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B} * \mathcal{C}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{B} * \mathcal{C}}(X_T[\tilde{\alpha}]) \right] \leq \bar{V}_{\mathcal{B}}(t_0, y_0) + \epsilon.$$

Now let  $\iota' : \mathcal{A} * \mathcal{C} \rightarrow \mathcal{B} * \mathcal{C}$  be the tracial  $W^*$ -embedding induced from the embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  (Lemma C.2), and let  $E' : \mathcal{B} * \mathcal{C} \rightarrow \mathcal{A} * \mathcal{C}$  be the corresponding conditional expectation. We want to define a control policy  $\tilde{\beta}$  in  $\mathcal{A} * \mathcal{C}$  by

$$\beta_t = E' \alpha_t, \quad t \in [t_0, T],$$

where the filtration is  $(\bar{\iota}_1 \circ \iota(\mathcal{A}) \vee \bar{\iota}_2(\mathcal{C}_{t_0, t}))_{t \in [t_0, T]}$  and the Brownian motion is the same  $(\bar{\iota}_2(S_t))_{t \in [t_0, T]}$ . Measurability of  $\beta_t$  follows from the fact that  $E'$  is a contraction from  $L^2(\mathcal{B} * \mathcal{C})$  to  $L^2(\mathcal{A} * \mathcal{C})$ , hence continuous.

We next need to check that  $\beta_t \in \bar{\iota}_1 \circ \iota(\mathcal{A}) \vee \bar{\iota}_2(\mathcal{C}_{t_0, t})$ . Here note by associativity (Lemma C.1), we have a canonical isomorphism  $\mathcal{B} * \mathcal{C} \cong \mathcal{B} * \mathcal{C}_{t_0, t} * \mathcal{C}_{t, T}$ , and this also restricts to an isomorphism  $\mathcal{A} * \mathcal{C} \cong \mathcal{A} * \mathcal{C}_{t_0, t} * \mathcal{C}_{t, T}$ . Let  $\iota'_t : \mathcal{A} * \mathcal{C}_{t_0, t} \rightarrow \mathcal{B} * \mathcal{C}_{t_0, t}$  be the inclusion and let  $E'_t : \mathcal{B} * \mathcal{C}_{t_0, t} \rightarrow \mathcal{A} * \mathcal{C}_{t_0, t}$  be the corresponding conditional expectation. Let  $\phi_t : \mathcal{B} * \mathcal{C}_{t_0, t} \rightarrow \mathcal{B} * \mathcal{C}$  be the canonical inclusion. By Lemma C.3, we have

$$E' \circ \phi_t = \phi_t \circ E'_t \text{ on } \mathcal{B} * \mathcal{C}_{t_0, t}.$$

Recall  $\alpha_t \in \bar{\iota}_1(\mathcal{A}) \vee \bar{\iota}_2(\mathcal{C}_{t_0, t}) \subseteq \mathcal{B} * \mathcal{C}$ , that is,  $\alpha_t$  is in the image of  $\phi_t$ , and therefore the above identity implies that  $E' \alpha_t$  is in the image of  $\phi_t \circ E'_t$ . This means that  $\beta_t$  is in the image of  $\mathcal{A} * \mathcal{C}_{t_0, t}$  or  $\beta_t \in \bar{\iota}_1 \circ \iota(\mathcal{A}) \vee \bar{\iota}_2(\mathcal{C}_{t_0, t})$ , as desired. Hence,  $\tilde{\beta}$  is a valid control policy using the asserted filtration.

Finally, noting that  $X_t[\tilde{\beta}] = E' X_t[\tilde{\alpha}]$ , by  $E$ -convexity of  $L$  and  $g$  (Assumption B), we have

$$\begin{aligned} \bar{V}_{\mathcal{A}}(t_0, Ey_0) &\leq \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{A} * \mathcal{C}}(X_t[\tilde{\beta}], \beta_t) dt + g_{\mathcal{A} * \mathcal{C}}(X_T[\tilde{\beta}]) \right] \\ &\leq \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B} * \mathcal{C}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{B} * \mathcal{C}}(X_T[\tilde{\alpha}]) \right] \\ &\leq \bar{V}_{\mathcal{B}}(t_0, y_0) + \epsilon. \end{aligned}$$

Since  $\epsilon > 0$  was arbitrary, we have  $\bar{V}_{\mathcal{A}}(t_0, Ey_0) \leq \bar{V}_{\mathcal{B}}(t_0, y_0)$  as desired.  $\square$

**Proposition 3.12.** *Suppose that Assumption A holds. For  $\mathcal{A} \in \mathbb{W}$ ,  $(t, x) \mapsto \bar{V}_{\mathcal{A}}(t, x)$  is continuous and bounded such that for  $(t_0, x_0) \in [0, T] \times L^2(\mathcal{A})_{sa}^d$ ,*

$$-C_1(1+T) \leq \bar{V}_{\mathcal{A}}(t_0, x_0) \leq C(\|x_0\|_{L^2(\mathcal{A})} + 1). \quad (3.30)$$

*Also, it is Lipschitz in space and Hölder in time:*

$$|\bar{V}_{\mathcal{A}}(t_1, x_1) - \bar{V}_{\mathcal{A}}(t_2, x_2)| \leq C \|x_1 - x_2\|_{L^2(\mathcal{A})} + C_{M_2} \sqrt{|t_1 - t_2|} \quad (3.31)$$

for  $(t_1, t_2, x_1, x_2) \in [0, T] \times [0, T] \times L^2(\mathcal{A})_{sa}^d \times L^2(\mathcal{A})_{sa}^d$ , where  $M_2 := \max\{\|x_1\|_{L^2(\mathcal{A})}, \|x_2\|_{L^2(\mathcal{A})}\}$ .

*Equivalently, for the value function  $\bar{V}$  defined on the space of laws in (3.13), we have*

$$-C_1(1+T) \leq \bar{V}(t_0, \lambda_0) \leq C_{M_3} \quad \text{for } (t_0, \lambda_0) \in [0, T] \times \Sigma_d^2, \quad (3.32)$$

where  $M_3$  is the second moment of  $\lambda_0$ , and

$$|\bar{V}(t_1, \lambda_1) - \bar{V}(t_2, \lambda_2)| \leq C d_W(\lambda_1, \lambda_2) + C_{M_4} \sqrt{|t_1 - t_2|} \quad (3.33)$$

for  $(t_0, \lambda_1, \lambda_2) \in [0, T] \times \Sigma_d^2 \times \Sigma_d^2$ , where  $M_4$  denotes the maximum of the second moments of  $\lambda_1, \lambda_2$ .

*Proof.* We first prove a version of (3.30) for the function  $\tilde{V}_{\mathcal{A}}(t_0, x_0)$ . The lower bound follows immediately from the lower bounds in (3.8) with any control policy, and the upper bound in

(3.30) follows from considering the trivial control  $\alpha_t \equiv 0$ . Indeed, the corresponding trajectory  $(X_t)_t$  becomes

$$X_t = x_0 + \beta_C \mathbb{1}_{\mathcal{A}}(W_t^0 - W_{t_0}^0) + \beta_F(S_t - S_{t_0}).$$

Thus, for any  $0 \leq t \leq T$ ,

$$\mathbb{E}[\|X_t\|_{L^2(\mathcal{A})}] \leq \|x_0\|_{L^2(\mathcal{A})} + (\beta_C + \beta_F)\sqrt{t}.$$

By the upper bound condition on  $\mathcal{L}_{\mathcal{A}}$  and  $g_{\mathcal{A}}$  in (3.8), we obtain the upper bound in (3.30) for  $\tilde{V}_{\mathcal{A}}(t_0, x_0)$ . Hence, the estimate (3.30) follows by taking the infimum over all tracial  $W^*$ -embeddings  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , noting that  $\|x_0\|_{L^2(\mathcal{A})} = \|\iota x_0\|_{L^2(\mathcal{B})}$ .

Now, we prove Lipschitz continuity in the spatial variable. Let  $x_1 \in L^2(\mathcal{A})_{sa}^d$ ,  $t_0 \in [0, T]$  and  $\epsilon > 0$ . By Lemma 3.8, there exists a tracial  $W^*$ -algebra  $\mathcal{B}$ , a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , and a control policy  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{B}, \iota x_1}^{t_0, T}$  associated with a filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$  and a compatible free Brownian motion  $(S_t)_{t \in [t_0, T]}$ , such that  $\iota(\mathcal{A}) \subseteq \mathcal{B}_{t_0}$  and

$$\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{B}}(X_T[\alpha]) \right] \leq \bar{V}_{\mathcal{A}}(t_0, x_1) + \epsilon, \quad (3.34)$$

where  $(X_t[\tilde{\alpha}])_{t \in [t_0, T]}$  satisfies the initial condition  $\hat{X}_{t_0}[\tilde{\alpha}] = \iota x_1$ . We remark, which will be useful later, that from (3.30) and the lower bound (3.8), such control policy  $\tilde{\alpha}$  satisfies a priori bound

$$\mathbb{E} \left[ \int_{t_0}^T \|\alpha_t\|_{L^2(\mathcal{B})}^2 dt \right] \leq C(\|x_1\|_{L^2(\mathcal{A})} + T + 1). \quad (3.35)$$

Note that  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{B}, \iota x_2}^{t_0, T}$ , since  $\iota x_2 \in \iota(\mathcal{A}) \subseteq \mathcal{B}_{t_0}$ . Let  $(\hat{X}_t[\tilde{\alpha}])_{t \in [t_0, T]}$  be a solution to (3.10) in  $L^2(\mathcal{B})_{sa}^d$  with the same filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$  and free Brownian motion  $(S_t)_{t \in [t_0, T]}$ , satisfying the initial condition  $\hat{X}_{t_0}[\tilde{\alpha}] = \iota x_2$ . Then using (3.34),

$$\begin{aligned} & \bar{V}_{\mathcal{A}}(t_0, x_2) - \bar{V}_{\mathcal{A}}(t_0, x_1) \\ & \leq \epsilon + \mathbb{E} \left[ \int_{t_0}^T \left( L_{\mathcal{B}}(\hat{X}_t[\tilde{\alpha}], \alpha_t) - L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) \right) dt + g_{\mathcal{B}}(\hat{X}_T[\tilde{\alpha}]) - g_{\mathcal{B}}(X_T[\tilde{\alpha}]) \right] \\ & \stackrel{(3.9)}{\leq} \epsilon + \mathbb{E} \left[ \int_{t_0}^T C_2 \|\hat{X}_t[\tilde{\alpha}] - X_t[\tilde{\alpha}]\|_{L^2(\mathcal{B})} dt + C_2 \|\hat{X}_T[\tilde{\alpha}] - X_T[\tilde{\alpha}]\|_{L^2(\mathcal{B})} \right] \\ & \stackrel{(3.22)}{\leq} \epsilon + C_2(1 + T) \sqrt{\tilde{C}} \|\iota x_2 - \iota x_1\|_{L^2(\mathcal{B})} = \epsilon + C_2(1 + T) \sqrt{\tilde{C}} \|x_2 - x_1\|_{L^2(\mathcal{A})}. \end{aligned}$$

As  $\epsilon > 0$  is arbitrary, we deduce a Lipschitz bound (3.31) in the spatial variable.

We now show Hölder continuity in time. Take  $t_1, t_2 \in [0, T]$ ,  $x_0 \in L^2(\mathcal{A})_{sa}^d$  and  $\epsilon > 0$ . We may assume  $t_1 < t_2$ . By definition of  $\bar{V}_{\mathcal{A}}$  in (3.12), we take a tracial  $W^*$ -algebra  $\mathcal{B}$ , a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , and a control policy  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_1, T}$  associated with free Brownian motion  $(S_t)_{t \in [t_1, T]}$  and a filtration  $(\mathcal{B}_t)_{t \in [t_1, T]}$  such that

$$\mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_t[\alpha], \alpha_t) dt + g_{\mathcal{B}}(X_T[\alpha]) \right] \leq \bar{V}_{\mathcal{A}}(t_1, x_0) + \epsilon. \quad (3.36)$$

Set  $N := \mathbb{E} \left[ \int_{t_1}^T \|\alpha_t\|_{L^2(\mathcal{B})}^2 dt \right]$ , which is bounded independent of  $t_1$  and  $t_2$ , due to (3.35). We define  $\tilde{\alpha}'$  to be a restriction of  $\tilde{\alpha}$  to  $[t_2, T]$ , associated with free Brownian motion  $(S_t - S_{t_2})_{t \in [t_2, T]}$  and a filtration  $(\mathcal{B}_t)_{t \in [t_2, T]}$ . Note that  $\tilde{\alpha}' \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_2, T}$ , since  $\iota x_0 \in \mathcal{B}_{t_1} \subseteq \mathcal{B}_{t_2}$ . Let  $(\hat{X}_t[\tilde{\alpha}'])_{t \in [t_2, T]}$

be the solution to (3.10) in  $L^2(\mathcal{B})_{sa}^d$ , satisfying the initial condition  $\hat{X}_{t_2}[\tilde{\alpha}'] = \iota x_0$ . By (3.21) in Theorem 3.6, setting  $M := \|\iota x_0\|_{L^2(\mathcal{B})} = \|x_0\|_{L^2(\mathcal{A})}$ ,

$$\mathbb{E}\left[\|\hat{X}_{t_2}[\tilde{\alpha}'] - X_{t_2}[\tilde{\alpha}]\|_{L^2(\mathcal{B})}^2\right] = \mathbb{E}\left[\|X_{t_2}[\tilde{\alpha}] - \iota x_0\|_{L^2(\mathcal{B})}^2\right] \leq \tilde{C}_{M,N}(t_2 - t_1).$$

Hence for  $t \in [t_2, T]$ , it follows from (3.22) that

$$\mathbb{E}\left[\|\hat{X}_t[\tilde{\alpha}'] - X_t[\tilde{\alpha}]\|_{L^2(\mathcal{B})}^2\right] \leq \tilde{C} \mathbb{E}\left[\|\hat{X}_{t_2}[\tilde{\alpha}'] - X_{t_2}[\tilde{\alpha}]\|_{L^2(\mathcal{B})}^2\right] \leq \tilde{C}\tilde{C}_{M,N}(t_2 - t_1).$$

Thus, using this along with the assumptions on  $L_{\mathcal{B}}$  and  $g_{\mathcal{B}}$ ,

$$\begin{aligned} & \bar{V}_{\mathcal{A}}(t_2, x_0) - \bar{V}_{\mathcal{A}}(t_1, x_0) \\ & \stackrel{(3.36)}{\leq} \epsilon + \mathbb{E}\left[-\int_{t_1}^{t_2} L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) dt + \int_{t_2}^T \left(L_{\mathcal{B}}(\hat{X}_t[\tilde{\alpha}'], \alpha'_t) - L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t)\right) dt\right. \\ & \quad \left.+ g_{\mathcal{B}}(\hat{X}_T[\tilde{\alpha}']) - g_{\mathcal{B}}(X_T[\tilde{\alpha}])\right] \\ & \leq \epsilon + 2C_1(t_2 - t_1) + C_2(1 + T)\sqrt{\tilde{C}\tilde{C}_{M,N}}\sqrt{t_2 - t_1}. \end{aligned}$$

As  $\epsilon > 0$  is arbitrary, we deduce a Lipschitz bound (3.31) in time variable.

Finally, (3.32) and (3.33) follow immediately from the same bounds by selecting representatives of  $\lambda_1, \lambda_2$  in a von Neumann algebra.  $\square$

#### 4. VISCOSITY SOLUTIONS

In this section, we develop a new theory of viscosity solutions on the space of non-commutative laws. In particular, we show that the value function for the stochastic optimal control problem described in the previous section is a viscosity solution. Although we do not establish a comparison principle on the space of non-commutative laws, we can do this in the case where there is no common noise by relating our problem to already developed theory of viscosity solutions on Hilbert space.

The notion of viscosity solution was invented to understand non-smooth solutions of first and second-order elliptic and parabolic PDE. The basic idea is that whenever the solution  $u$  has a Taylor approximation from above or from below (or equivalently can be touched from above or below by a smooth test function), then substituting the first and second order terms from the Taylor expansion in place of the gradient and Hessian in the differential equation will produce an inequality in one direction. The theory of viscosity solutions was first developed in [35, 20, 19] and then has been extended to infinite-dimensional Hilbert space [24, 27, 55]. The first author and Tudorascu adapted the theory of viscosity solutions to Hamilton-Jacobi equation on the Wasserstein space [45]. The notion of viscosity solution was applied in the random matrix setting in [57], which is one motivation for the present work.

We consider the following equation on the space of non-commutative laws  $\Sigma_d^2$ :

$$\begin{aligned} -\partial_t V(t, \lambda) + H(\lambda, -\partial V(t, \lambda)) - \frac{\beta_C^2}{2} \Delta V(t, \lambda) - \frac{\beta_F^2}{2} \Theta V(t, \lambda) &= 0, \\ V(T, \lambda) &= g(\lambda). \end{aligned} \tag{4.1}$$

In Definition 4.2, we define viscosity solutions for this equation through the related collection of equations for each von Neumann algebra  $\mathcal{A} \in \mathbb{W}$ ,

$$-\partial_t V_{\mathcal{A}}(t, X) + H_{\mathcal{A}}(X, -\nabla V_{\mathcal{A}}(t, X)) - \frac{\beta_C^2}{2} \Delta_{\mathcal{A}} V(t, X) - \frac{\beta_F^2}{2} \Theta_{\mathcal{A}} V(t, X) = 0, \tag{4.2}$$

$$V_{\mathcal{A}}(T, X) = g_{\mathcal{A}}(X),$$

where the common noise Laplacian  $\Delta_{\mathcal{A}}$  and the free individual noise Laplacian  $\Theta_{\mathcal{A}}$  will be defined below. In fact, we will take a supremum of the left-hand side over the von Neumann algebras, leading to some asymmetry between the arguments for subsolutions and supersolutions. Note that this is a parabolic equation expressed backwards in time, with the terminal condition at time  $T$ .

Given a function  $V : [0, T] \times \Sigma_d^2 \rightarrow \mathbb{R}$ , for each  $\mathcal{A} \in \mathbb{W}$  we define

$$V_{\mathcal{A}}(t, X) := V(t, \lambda_X), \quad X \in L^2(\mathcal{A})_{sa}^d \text{ and } t \in [0, T]. \quad (4.3)$$

This is consistent with our notation for the value function  $(\bar{V}_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  and  $\bar{V}$  in (3.13), due to Lemma 3.9.

For every  $\mathcal{A} \in \mathbb{W}$ , viscosity subsolutions and supersolutions to (4.2) can be defined in a standard way on the Hilbert space  $L^2(\mathcal{A})_{sa}^d$ . However, as will be seen later, this is the case only if we assume that free individual noise is *not* present, since the operator  $\Theta$  is not defined on every  $L^2(\mathcal{A})_{sa}^d$ . The standard class of test functions on the Hilbert space  $L^2(\mathcal{A})_{sa}^d$  is given by

$$\mathcal{X}_{\mathcal{A}} := \left\{ \phi \in C^1([0, T] \times L^2(\mathcal{A})_{sa}^d) : \nabla^2 \phi \in C([0, T] \times L^2(\mathcal{A})_{sa}^d; \text{BL}(L^2(\mathcal{A})_{sa}^d)) \right\},$$

where BL denotes the space of bounded linear operators. Examples of  $\Phi \in \mathcal{X}_{\mathcal{A}}$  include,  $\Phi(x) = \|x - x_0\|_{L^2(\mathcal{A})}^2$  for  $x_0 \in L^2(\mathcal{A})_{sa}^d$  fixed, and  $\Phi(x) = g(P_N x)$  where  $P_N$  is a projection onto a  $N$ -dimensional subspace and  $g$  is a smooth function.

It is not so straightforward to choose a set of smooth test functions on  $\Sigma_d^2$  that are tracial  $W^*$ -functions. For first order equations, one could get around this by expressing the notion of viscosity solutions using subdifferentials on the Wasserstein space, which is done in the commutative setting in [45], but this does not immediately help with second-order equations. Instead, in our setting of non-commutative laws, following the idea of an  $L$ -derivative [12], we define the collection  $\mathcal{X}_{\Sigma}$  of admissible test functions  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  on  $[0, T] \times \Sigma_d^2$  that satisfy the following properties:

- (a) For any  $\mathcal{A} \in \mathbb{W}$ ,  $U_{\mathcal{A}} : [0, T] \times L^2(\mathcal{A})_{sa}^d \rightarrow \mathbb{R}$  and for each  $t \in [0, T]$ ,  $(U_{\mathcal{A}}(t, \cdot))_{\mathcal{A} \in \mathbb{W}}$  is a tracial  $W^*$ -function.
- (b) For any  $\mathcal{A} \in \mathbb{W}$ ,  $U_{\mathcal{A}} \in C^{1,1}([0, T] \times L^2(\mathcal{A})_{sa}^d)$  (i.e. first derivatives, both in time and spatial variables, are Lipschitz).
- (c) For any  $\mathcal{A} \in \mathbb{W}$ ,  $X, A \in L^2(\mathcal{A})_{sa}^d$ ,  $B \in L^\infty(\mathcal{A})_{sa}^d$  and  $t \in [0, T]$ , the partial second derivatives  $\text{Hess } U_{\mathcal{A}}(t, X)[A, B]$  exists and  $t \mapsto \text{Hess } U_{\mathcal{A}}(t, X)[A, B]$  is continuous.
- (d) There is a constant  $K > 0$  such that, with  $\mathcal{A}, X, A, B, t$  as in (c) and  $Y \in L^2(\mathcal{A})_{sa}^d$ ,

$$|\text{Hess } U_{\mathcal{A}}(t, X)[A, B] - \text{Hess } U_{\mathcal{A}}(t, Y)[A, B]| \leq K \|X - Y\|_{L^2(\mathcal{A})} \|A\|_{L^2(\mathcal{A})} \|B\|_{L^\infty(\mathcal{A})}. \quad (4.4)$$

On the space  $\mathcal{X}_{\Sigma}$ , we define the infinite-dimensional operators that will be part of the Hamilton-Jacobi-Bellman equation. For  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$  or  $U_{\mathcal{A}} \in \mathcal{X}_{\mathcal{A}}$ , the *common noise Laplacian* is defined to be

$$\Delta_{\mathcal{A}} U(t, X) := \text{Hess } U_{\mathcal{A}}(t, X)[\mathbb{1}_{\mathcal{A}}, \mathbb{1}_{\mathcal{A}}], \quad X \in L^2(\mathcal{A})_{sa}^d,$$

where we recall that  $\mathbb{1}_{\mathcal{A}} \in L^2(\mathcal{A})_{sa}^d$  has the algebra unit in every component.

When defining the free individual noise Laplacian, it is not sufficient to consider a function  $U_{\mathcal{A}}$  defined on a single von Neumann algebra  $\mathcal{A}$ , but instead we should use a tracial  $W^*$ -function  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$ . This is because not all von Neumann algebras support freely independent semicircle laws. For  $\mathcal{A} \in \mathbb{W}$ , we consider a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  where  $\mathcal{B}$  contains

a  $d$ -dimensional semicircle element  $S = (S^1, \dots, S^d)$  that is freely independent of  $\iota(\mathcal{A})$ . Then, for  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$ , the *free individual noise Laplacian* is defined to be

$$\Theta_{\mathcal{A}} U(t, X) := \sum_{l=1}^d \text{Hess } U_{\mathcal{B}}(t, \iota X) [S^l \mathbf{e}_{\mathcal{A}}^l, S^l \mathbf{e}_{\mathcal{A}}^l], \quad X \in L^2(\mathcal{A})_{sa}^d, \quad (4.5)$$

where  $\mathbf{e}_{\mathcal{A}}^l \in L^2(\mathcal{A})_{sa}^d$  has the algebra unit in the  $l$ -th component and zero in the other components.

This definition is motivated by the classical fact that the Laplacian of a function on  $\mathbb{R}^d$  can be expressed as

$$\Delta u(x) = \mathbb{E}[\langle \text{Hess} u(x) Z, Z \rangle],$$

where  $Z$  is a standard Gaussian random vector in  $\mathbb{R}^d$ . A similar approach is used in [59, §4.3] and [61], while many previous works gave a more explicit definition of the Laplacian for non-commutative polynomials, power series, and the like, in terms of non-commutative derivative operations and traces; see Appendix B for more detail. We point out that such explicit computations are not necessarily possible for tracial  $W^*$ -functions in general since a function might be smooth with respect to the Wasserstein distance but *not* continuous with respect to (weak-\*) convergence in non-commutative law, and hence unable to be approximated by trace polynomials. Hence, it is necessary for us to define the Laplacian directly in terms of free semi-circulars.

*Remark 4.1.*

- i. Note that for  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$ , (4.5) does *not* depend on the choice of von Neumann algebra  $\mathcal{B}$  or the  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ . Indeed, the common noise Laplacian and free individual noise Laplacian are tracial  $W^*$ -functions.
- ii. With the tuple from the definition (4.5),  $\iota X + \sqrt{t-s} \beta_F S$  has the same non-commutative law as  $\iota X + \beta_F (S_t - S_s)$  where  $S_t - S_s$  is freely independent from  $\iota X$ . This holds because  $S_t - S_s$  has the same law as  $\sqrt{t-s} S$ , and both are freely independent of  $\iota X$ . This fact will be used in the proof of the mixed-Itô formula, Lemma 4.5.

**4.1. Intrinsic Viscosity Solution.** For a metric space  $\mathcal{Y}$ , let  $\text{USC}(\mathcal{Y})$  denote the upper-semicontinuous and bounded above functions, and  $\text{LSC}(\mathcal{Y})$  denote the lower-semicontinuous and bounded below functions. We say that a function  $\Phi \in \text{LSC}(\mathcal{Y})$  *touches* a function  $U \in \text{USC}(\mathcal{Y})$  *from above (below)* at  $y \in \mathcal{Y}$  if  $\Phi(y) = U(y)$  and  $\Phi(y') \geq U(y')$  (resp.  $\Phi(y') \leq U(y')$ ) for all  $y' \in \mathcal{Y}$ .

**Definition 4.2.** Suppose that  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \text{USC}([0, T] \times \Sigma_d^2)$  (resp.  $\text{LSC}([0, T] \times \Sigma_d^2)$ ). We say that  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a *free viscosity sub(super)solution* of (4.1) if

- (1)  $U_{\mathcal{A}}(T, x) \leq (\geq) g_{\mathcal{A}}(x)$  for any  $\mathcal{A} \in \mathbb{W}$  and  $x \in L^2(\mathcal{A})_{sa}^d$ .
- (2) Whenever  $(\Phi_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$  touches  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  from above (below) at  $(t_0, \lambda_0) \in [0, T] \times \Sigma_d^2$ ,

$$\sup_{\mathcal{A} \in \mathbb{W}, x_0 \in L^2(\mathcal{A})_{sa}^d, \lambda_{x_0} = \lambda_0} \left\{ -\partial_t \Phi_{\mathcal{A}}(t_0, x_0) + H_{\mathcal{A}}(x_0, -\nabla \Phi_{\mathcal{A}}(t_0, x_0)) - \frac{\beta_C^2}{2} \Delta_{\mathcal{A}} \Phi(t_0, x_0) - \frac{\beta_F^2}{2} \Theta_{\mathcal{A}} \Phi(t_0, x_0) \right\} \leq (\geq) 0.$$

We say that a continuous function  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  on  $[0, T] \times \Sigma_d^2$  is a *free viscosity solution* if it is a free viscosity subsolution *and* supersolution.

We first establish an Itô formula, which will then be used to establish a dynamic programming principle. The dynamic programming principle is then used to prove that the value function, defined in (3.13), is a free viscosity subsolution.



4.1.1. *Mixed-Itô formula.* We essentially need two versions of Taylor's theorem. The following lemma provides the first-order Taylor's theorem.

**Lemma 4.3.** *For any  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$ ,  $X, Y \in L^2(\mathcal{A})_{sa}^d$  and  $t \in [0, T]$ ,*

$$|U_{\mathcal{A}}(t, Y) - U_{\mathcal{A}}(t, X) - \langle \nabla U_{\mathcal{A}}(t, X), Y - X \rangle_{L^2(\mathcal{A})}| \leq \|U_{\mathcal{A}}\|_{C^{1,1}} \|Y - X\|_{L^2(\mathcal{A})}^2.$$

*Proof.* This follows directly from  $U_{\mathcal{A}} \in C^{1,1}([0, T] \times L^2(\mathcal{A})_{sa}^d)$ .  $\square$

The next lemma provides the second order Taylor's theorem.

**Lemma 4.4.** *For any  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$ ,  $t \in [0, T]$  and  $X, Y \in L^2(\mathcal{A})_{sa}^d$  with  $X - Y \in L^{\infty}(\mathcal{A})_{sa}^d$ ,*

$$\begin{aligned} & |U_{\mathcal{A}}(t, Y) - U_{\mathcal{A}}(t, X) - \langle \nabla U_{\mathcal{A}}(t, X), Y - X \rangle_{L^2(\mathcal{A})} - \frac{1}{2} \text{Hess } U_{\mathcal{A}}(t, X)[Y - X, Y - X]| \\ & \leq \frac{K}{2} \|Y - X\|_{L^2(\mathcal{A})} \|Y - X\|_{L^2(\mathcal{A})} \|Y - X\|_{L^{\infty}(\mathcal{A})}, \end{aligned}$$

where  $K > 0$  is a constant from (4.4).

*Proof.* Fix  $X \in L^2(\mathcal{A})_{sa}^d$ ,  $t \in [0, T]$  and let

$$\phi(H) := U_{\mathcal{A}}(t, X + H) - U_{\mathcal{A}}(X) - \langle \nabla U_{\mathcal{A}}(t, X), H \rangle_{L^2(\mathcal{A})} - \frac{1}{2} \text{Hess } U_{\mathcal{A}}(t, X)[H, H].$$

We have that, for  $H \in L^2(\mathcal{A})_{sa}^d$  and  $B \in L^{\infty}(\mathcal{A})_{sa}^d$ ,

$$\begin{aligned} \langle \nabla \phi(H), B \rangle_{L^2(\mathcal{A})} &= \langle \nabla U_{\mathcal{A}}(t, X + H) - \nabla U_{\mathcal{A}}(t, X), B \rangle_{L^2(\mathcal{A})} - \text{Hess } U_{\mathcal{A}}(t, X)[H, B] \\ &= \int_0^1 \left( \text{Hess } U_{\mathcal{A}}(t, X + sH)[H, B] - \text{Hess } U_{\mathcal{A}}(t, X)[H, B] \right) ds. \end{aligned}$$

By the condition (4.4),

$$|\langle \nabla \phi(H), B \rangle_{L^2(\mathcal{A})}| \leq K \|H\|_{L^2(\mathcal{A})} \|H\|_{L^2(\mathcal{A})} \|B\|_{L^{\infty}(\mathcal{A})}.$$

It follows that

$$\begin{aligned} \phi(H) - \phi(0) &= \int_0^1 \langle \nabla \phi(tH), H \rangle_{L^2(\mathcal{A})} dt \\ &\leq \frac{K}{2} \|H\|_{L^2(\mathcal{A})} \|H\|_{L^2(\mathcal{A})} \|H\|_{L^{\infty}(\mathcal{A})}, \end{aligned}$$

yielding the second order Taylor's theorem.  $\square$

We are now ready to state a free version of the mixed-Itô formula.

**Lemma 4.5** (Mixed Itô formula). *Suppose that Assumption **A** holds. For  $\mathcal{A} \in \mathbb{W}$ ,  $t_0 \in [0, T]$ ,  $x_0 \in L^2(\mathcal{A})_{sa}^d$ , and  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, T}$ , let  $(X_t)_{t \in [t_0, T]}$  be the corresponding solution to (3.10). Then for any  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$  and  $t \in [t_0, T]$ ,*

$$\begin{aligned} \mathbb{E}[U_{\mathcal{A}}(t, X_t)] &= U_{\mathcal{A}}(t_0, x_0) + \mathbb{E} \left[ \int_{t_0}^t \left( \partial_t U_{\mathcal{A}}(s, X_s) + \langle \nabla U_{\mathcal{A}}(s, X_s), b_{\mathcal{A}}(X_s, \alpha_s) \rangle_{L^2(\mathcal{A})} \right. \right. \\ &\quad \left. \left. + \frac{\beta_C^2}{2} \Delta_{\mathcal{A}} U(s, X_s) + \frac{\beta_F^2}{2} \Theta_{\mathcal{A}} U(s, X_s) \right) ds \right]. \end{aligned} \quad (4.6)$$

*Proof.* Fix  $U = (U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$ ,  $t \in [t_0, T]$  and  $\delta > 0$ . Then we take a partition  $t_0 < t_1 < \dots < t_N = t$  with  $t_{i+1} - t_i < \delta$  for  $i \in \{0, \dots, N-1\}$ . We define

$$Y_{t_i}^1 := X_{t_i} + \beta_C \mathbb{1}_{\mathcal{A}}(W_{t_{i+1}}^0 - W_{t_i}^0),$$

$$Y_{t_i}^2 := Y_{t_i}^1 + \beta_F(S_{t_{i+1}} - S_{t_i}),$$

and for  $r \in [t_i, t_{i+1})$ ,

$$Y_r^3 = Y_{t_i}^2 + \int_{t_i}^r b_{\mathcal{A}}(X_s, \alpha_s) ds.$$

Then,

$$\begin{aligned} \mathbb{E} \left[ \|Y_r^3 - X_r\|_{L^2(\mathcal{A})}^2 \right] &= \mathbb{E} \left[ \|\beta_C \mathbf{1}_{\mathcal{A}}(W_{t_{i+1}}^0 - W_r^0) + \beta_F(S_{t_{i+1}} - S_r)\|_{L^2(\mathcal{A})}^2 \right] \\ &\leq d\delta(\beta_F^2 + \beta_C^2). \end{aligned} \quad (4.7)$$

We write

$$\begin{aligned} U_{\mathcal{A}}(t_{i+1}, X_{t_{i+1}}) - U_{\mathcal{A}}(t_i, X_{t_i}) &= \left( U_{\mathcal{A}}(t_{i+1}, X_{t_{i+1}}) - U_{\mathcal{A}}(t_i, Y_{t_i}^2) \right) \\ &\quad + \left( U_{\mathcal{A}}(t_i, Y_{t_i}^2) - U_{\mathcal{A}}(t_i, Y_{t_i}^1) \right) + \left( U_{\mathcal{A}}(t_i, Y_{t_i}^1) - U_{\mathcal{A}}(t_i, X_{t_i}) \right). \end{aligned} \quad (4.8)$$

To control the first term of RHS above, by the fundamental theorem of Calculus and the chain rule,

$$U_{\mathcal{A}}(t_{i+1}, X_{t_{i+1}}) - U_{\mathcal{A}}(t_i, Y_{t_i}^2) = \int_{t_i}^{t_{i+1}} \left( \partial_t U_{\mathcal{A}}(s, Y_s^3) + \langle \nabla U_{\mathcal{A}}(s, Y_s^3), b_{\mathcal{A}}(X_s, \alpha_s) \rangle_{L^2(\mathcal{A})} \right) ds.$$

Note that we have, for any  $\epsilon > 0$ ,

$$\begin{aligned} &\mathbb{E} \left[ \left| \langle \nabla U_{\mathcal{A}}(s, Y_s^3), b_{\mathcal{A}}(X_s, \alpha_s) \rangle_{L^2(\mathcal{A})} - \langle \nabla U_{\mathcal{A}}(s, X_s), b_{\mathcal{A}}(X_s, \alpha_s) \rangle_{L^2(\mathcal{A})} \right| \right] \\ &\leq \|U_{\mathcal{A}}\|_{C^{1,1}} \sqrt{\mathbb{E}[\|Y_s^3 - X_s\|_{L^2(\mathcal{A})}^2]} \sqrt{\mathbb{E}[\|b_{\mathcal{A}}(X_s, \alpha_s)\|_{L^2(\mathcal{A})}^2]} \\ &\stackrel{(4.7)}{\leq} \|U_{\mathcal{A}}\|_{C^{1,1}} \sqrt{d\delta(\beta_F^2 + \beta_C^2)} \left( \frac{1}{2\epsilon} + \frac{\epsilon}{2} \mathbb{E}[\|b_{\mathcal{A}}(X_s, \alpha_s)\|_{L^2(\mathcal{A})}^2] \right). \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} \left[ \left| \partial_t U_{\mathcal{A}}(s, Y_s^3) - \partial_t U_{\mathcal{A}}(s, X_s) \right| \right] &\leq \|U_{\mathcal{A}}\|_{C^{1,1}} \sqrt{\mathbb{E}[\|Y_s^3 - X_s\|_{L^2(\mathcal{A})}^2]} \\ &\stackrel{(4.7)}{\leq} \|U_{\mathcal{A}}\|_{C^{1,1}} \sqrt{d\delta(\beta_F^2 + \beta_C^2)}. \end{aligned}$$

Hence for any  $i \in \{0, \dots, N-1\}$

$$\begin{aligned} &\left| \mathbb{E} \left[ U_{\mathcal{A}}(t_{i+1}, X_{t_{i+1}}) - U_{\mathcal{A}}(t_i, Y_{t_i}^2) - \int_{t_i}^{t_{i+1}} \left( \partial_t U_{\mathcal{A}}(s, X_s) + \langle \nabla U_{\mathcal{A}}(s, X_s), b_{\mathcal{A}}(X_s, \alpha_s) \rangle_{L^2(\mathcal{A})} \right) ds \right] \right| \\ &\leq \|U_{\mathcal{A}}\|_{C^{1,1}} \sqrt{d\delta(\beta_F^2 + \beta_C^2)} \left( (t_{i+1} - t_i) \left( 1 + \frac{1}{2\epsilon} \right) + \frac{\epsilon}{2} \int_{t_i}^{t_{i+1}} \mathbb{E}[\|b_{\mathcal{A}}(X_s, \alpha_s)\|_{L^2(\mathcal{A})}^2] ds \right). \end{aligned}$$

To control the last term of RHS in (4.8), by the second order Taylor's theorem (Lemma 4.4),

$$\begin{aligned} &\left| \mathbb{E} \left[ U_{\mathcal{A}}(t_i, Y_{t_i}^1) - U_{\mathcal{A}}(t_i, X_{t_i}) - \langle \nabla U_{\mathcal{A}}(t_i, X_{t_i}), Y_{t_i}^1 - X_{t_i} \rangle_{L^2(\mathcal{A})} \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \text{Hess } U_{\mathcal{A}}(t_i, X_{t_i}) [Y_{t_i}^1 - X_{t_i}, Y_{t_i}^1 - X_{t_i}] \right] \right| \\ &\leq \mathbb{E} \left[ \frac{K}{2} \|Y_{t_i}^1 - X_{t_i}\|_{L^2(\mathcal{A})} \|Y_{t_i}^1 - X_{t_i}\|_{L^2(\mathcal{A})} \|Y_{t_i}^1 - X_{t_i}\|_{L^\infty(\mathcal{A})} \right] \\ &\leq K (d\beta_C^2)^{\frac{3}{2}} \sqrt{\delta} (t_{i+1} - t_i). \end{aligned}$$

By independence of  $W_{t_{i+1}}^0 - W_{t_i}^0$  and  $\mathcal{F}_{t_i}$ ,

$$\mathbb{E}[\langle \nabla U_{\mathcal{A}}(t_i, X_{t_i}), Y_{t_i}^1 - X_{t_i} \rangle_{L^2(\mathcal{A})}] = 0,$$

and

$$\mathbb{E}\left[\frac{1}{2}\text{Hess } U_{\mathcal{A}}(t_i, X_{t_i})[Y_{t_i}^1 - X_{t_i}, Y_{t_i}^1 - X_{t_i}]\right] = (t_{i+1} - t_i)\mathbb{E}\left[\frac{\beta_C^2}{2}\Delta_{\mathcal{A}}U(t_i, X_{t_i})\right].$$

In addition, using (4.4) and (3.21) from Theorem 3.6 with the same definition of  $M$  and  $N$ , for  $s \in [t_i, t_{i+1}]$ ,

$$\left|\mathbb{E}\left[\frac{\beta_C^2}{2}\Delta_{\mathcal{A}}U(t_i, X_{t_i})\right] - \mathbb{E}\left[\frac{\beta_C^2}{2}\Delta_{\mathcal{A}}U(s, X_s)\right]\right| \leq \frac{K\beta_C^2}{2}\sqrt{\tilde{C}_{M,N}}\delta.$$

Thus,

$$\begin{aligned} & \left|\mathbb{E}\left[U_{\mathcal{A}}(t_i, Y_{t_i}^1) - U_{\mathcal{A}}(t_i, X_{t_i}) - \int_{t_i}^{t_{i+1}} \frac{\beta_C^2}{2}\Delta_{\mathcal{A}}U(s, X_s)ds\right]\right| \\ & \leq (K(d\beta_C^2)^{\frac{3}{2}} + \frac{K\beta_C^2}{2}\sqrt{\tilde{C}_{M,N}})\sqrt{\delta}(t_{i+1} - t_i). \end{aligned}$$

To control the middle term of RHS in (4.8), let  $\mathcal{B} \in \mathbb{W}$  and  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  be a tracial  $W^*$ -embedding such that  $\mathcal{B}$  contains a  $d$ -variable semi-circular element  $S = (S^1, \dots, S^d)$  freely independent of  $\iota(\mathcal{A})$ . The tuples  $Y_{t_i}^2$  and  $\iota Y_{t_i}^1 + \sqrt{t_{i+1} - t_i}\beta_F S$  have the same non-commutative law, so, by the tracial property of  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$ ,

$$U_{\mathcal{A}}(t_i, Y_{t_i}^2) - U_{\mathcal{A}}(t_i, Y_{t_i}^1) = U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1 + \sqrt{t_{i+1} - t_i}\beta_F S) - U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1).$$

By free independence of  $S$  and  $\iota Y_{t_i}^1$ ,

$$\langle \nabla U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1), \beta_F S \rangle_{L^2(\mathcal{B})} = 0.$$

Thus, using the second order Taylor's theorem along with the fact  $\|S\|_{L^\infty(\mathcal{A})} = 2\|S\|_{L^2(\mathcal{A})} = 2$ ,

$$\begin{aligned} & \left|U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1 + \sqrt{t_{i+1} - t_i}\beta_F S) - U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1) - \frac{t_{i+1} - t_i}{2}\text{Hess } U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1)[\beta_F S, \beta_F S]\right| \\ & \leq K(d\beta_F^2)^{\frac{3}{2}}\sqrt{\delta}(t_{i+1} - t_i). \end{aligned}$$

Note that, for  $j_1 \neq j_2$  in  $\{1, \dots, d\}$ ,

$$\text{Hess } U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1)[\beta_F S^{j_1} \mathbf{e}_{\mathcal{A}}^{j_1}, \beta_F S^{j_2} \mathbf{e}_{\mathcal{A}}^{j_2}] = 0,$$

by free independence of  $S^{j_1}, S^{j_2}$  and  $\iota Y_{t_i}^1$  along with the tracial property of  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$ . Thus, again using (4.4) and (3.21) from Theorem 3.6,

$$\begin{aligned} & \mathbb{E}\left[\frac{t_{i+1} - t_i}{2}\text{Hess } U_{\mathcal{B}}(t_i, \iota Y_{t_i}^1)[\beta_F S, \beta_F S] - \int_{t_i}^{t_{i+1}} \frac{\beta_F^2}{2}\Theta_{\mathcal{A}}U(s, X_s)ds\right] \\ & \leq \frac{K\beta_F^2}{2}\sqrt{\tilde{C}_{M,N}}\sqrt{\delta}(t_{i+1} - t_i). \end{aligned}$$

Putting the above estimates together and taking the summation over  $i = 0, \dots, N-1$ , the difference of LHS and RHS in (4.6) is bounded by

$$C\sqrt{\delta}\left(\left(1 + \frac{1}{\epsilon}\right)(t - t_0) + \epsilon \int_{t_0}^t \mathbb{E}[\|b_{\mathcal{A}}(X_s, \alpha_s)\|_{L^2(\mathcal{A})}^2]ds\right),$$

where  $C$  is a constant that does not depend on  $\delta$  or  $\epsilon$ . Using bounds from Theorem 3.6 and the bounds on  $b_{\mathcal{A}}$ ,

$$\mathbb{E}[\|b_{\mathcal{A}}(X_s, \alpha_s)\|_{L^2(\mathcal{A})}^2] \leq C,$$

where  $C$  only depends on  $M = \|x_0\|_{L^2(\mathcal{A})}$  and  $N = \mathbb{E} \left[ \int_{t_0}^T \|\alpha_s\|_{L^2(\mathcal{A})}^2 ds \right]$ . We can then set  $\epsilon := \delta^{\frac{1}{4}}$  and sending  $\delta \rightarrow^+ 0$  proves the theorem.  $\square$

**4.1.2. Dynamic Programming Principle.** We now show that the solution to (3.10) satisfies a Markov property analogous to the classical result in [85, Lemma 3.2].

**Lemma 4.6.** *Consider  $\mathcal{A} \in \mathbb{W}$ ,  $x_0 \in L^2(\mathcal{A})_{sa}^d$  and  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, T}$ . Then for any  $0 \leq t_0 \leq t_1 \leq T$ , there exist a collection of control policies  $\tilde{\alpha}^{\bar{\omega}} \in \mathbb{A}_{\mathcal{A}, X_{t_1}[\tilde{\alpha}](\bar{\omega})}^{t_1, T}$  for  $\bar{\omega} \in \Omega$  such that*

(1) For  $\mathbb{P}$ -a.e.  $\bar{\omega} \in \Omega$ ,

$$\begin{aligned} & \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{A}}(X_T[\tilde{\alpha}]) \mid \mathcal{F}_{t_1} \right] (\bar{\omega}) \\ &= \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}], \alpha_t^{\bar{\omega}}) dt + g_{\mathcal{A}}(X_T[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}]) \right]; \end{aligned} \quad (4.9)$$

(2) For every  $\bar{\omega} \in \Omega$ ,  $(\alpha_t^{\bar{\omega}})_{t \in [t_1, T]}$  is independent of  $\mathcal{F}_{t_1}$ .

*Proof.* For  $s, t \in [0, T]$  we let  $s \wedge t := \min\{s, t\}$ . Define  $\mathcal{B}_t = \mathcal{B}(\{u(s \wedge t)_{0 \leq s \leq T} : u \in C([0, T])\})$  to be the Borel  $\sigma$ -algebra of paths stopped at time  $t$  (and continued by a constant). We then let  $\mathcal{B}_{t+} = \cup_{s > t} \mathcal{B}_s$  be the right limit. By Theorem 2.10 in [85, Chapter 2], there exists  $\eta : [0, T] \times C([0, T]) \rightarrow L^2(\mathcal{A})_{sa}^d$ , which is progressively measurable with respect to  $(\mathcal{B}_{t+})_{0 \leq t \leq T}$ , such that for  $t \in [t_0, T]$

$$\alpha_t(\omega) = \eta(t, (W_{s \wedge t}^0(\omega))_{0 \leq s \leq T}).$$

For each  $\omega, \bar{\omega} \in \Omega$  and  $t \in [t_1, T]$ , we now define

$$\alpha_t^{\bar{\omega}}(\omega) := \eta(t, (W_{s \wedge t_1}^0(\bar{\omega}) + W_{s \wedge t}^0(\omega) - W_{s \wedge t_1}^0(\omega))_{0 \leq s \leq T}).$$

Using this, we define the control policy  $\tilde{\alpha}^{\bar{\omega}} \in \mathbb{A}_{\mathcal{A}, X_{t_1}[\tilde{\alpha}](\bar{\omega})}^{t_1, T}$  with the same free semicircular process (we take the increment  $S'_t = S_t - S_{t_1}$  for  $t' \in [t_1, T]$ ) and free filtration as in  $\tilde{\alpha}$ . This is a valid control policy because it inherits the progressive measurability from  $\eta$  by composition with a continuous process. In particular, for  $\mathbb{P}$ -a.e.  $\bar{\omega} \in \Omega$ ,  $(\alpha_t^{\bar{\omega}})_{t_1 \leq t \leq T}$  is progressively measurable with respect to the filtration  $(\mathcal{F}_t)_{t \in [t_1, T]}$ . Since  $\alpha_t^{\bar{\omega}}$  is a function of the increments  $W_s^0 - W_{t_1}^0$  for  $s \in [t_1, t]$ , which are independent from  $\mathcal{F}_{t_1}$ , we deduce that  $(\alpha_t^{\bar{\omega}})_{t \in [t_1, T]}$  is independent of  $\mathcal{F}_{t_1}$ .

To show (4.9), we argue that the free stochastic differential equation for  $t \in [t_1, T]$  may be viewed on the probability space  $\hat{\Omega} := \{u \in C([0, T]) : u(s) = 0, s \in [0, t_1]\}$ . We define  $\phi : \Omega \rightarrow \hat{\Omega}$  by  $\phi(\omega) = (W_s^0(\omega) - W_{s \wedge t_1}^0(\omega))_{0 \leq s \leq T}$ . We let  $\hat{\mathbb{P}} := \phi_{\#} \mathbb{P}$  and consider the filtration  $(\hat{\mathcal{F}}_t)_{t_1 \leq t \leq T}$  generated by the canonical stochastic process  $\hat{W}_t^0 := u(t)$ . We claim this makes  $(\hat{W}_t^0)_{t_1 \leq t \leq T}$  a standard Brownian motion on  $(\hat{\Omega}, (\hat{\mathcal{F}}_t)_{t_1 \leq t \leq T}, \hat{\mathbb{P}})$ . Let  $f \in C(\mathbb{R}^{k+n})$  and consider real numbers  $0 \leq b_k < \dots < b_0 < t_1 \leq a_1 < a_2 < \dots < a_n$ . Defining  $\hat{f} : \hat{\Omega} \rightarrow \mathbb{R}$  by

$$\hat{f}(\hat{\omega}) = f(\hat{\omega}(b_k), \dots, \hat{\omega}(b_1), \hat{\omega}(a_1), \dots, \hat{\omega}(a_n)), \quad \text{for all } \hat{\omega} \in \hat{\Omega}, \quad (4.10)$$

we have that

$$\hat{f} \circ \phi(\omega) = f(0, \dots, 0, W_{a_1}^0(\omega) - W_{t_1}^0(\omega), \dots, W_{a_n}^0(\omega) - W_{t_1}^0(\omega)). \quad (4.11)$$

Exploiting (4.11), we compute  $\int_{\Omega} \hat{f} \circ \phi(\omega) \mathbb{P}(d\omega)$  which by definition is nothing but  $\int_{\hat{\Omega}} \hat{f}(\hat{\omega}) \hat{\mathbb{P}}(d\hat{\omega})$ . We use the arbitrariness of  $\hat{f}$  to deduce that  $\hat{\mathbb{P}}$  is the restriction of Wiener measure to  $\hat{\Omega}$ .

We denote by  $\mathbb{P}^{\bar{\omega}}$  (resp.  $\mathbb{E}^{\bar{\omega}}$ ) the conditional probability measure (resp. conditional expectation) given  $\mathcal{F}_{t_1}$  at  $\bar{\omega}$ , i.e., for any random variable  $Y \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ ,

$$\mathbb{E}^{\bar{\omega}}[Y] = \int_{\Omega} Y(\omega) \mathbb{P}^{\bar{\omega}}(d\omega) = \mathbb{E}[Y \mid \mathcal{F}_{t_1}](\bar{\omega}).$$

We continue to assume that  $\hat{f}$  is as in (4.10). Since the increment  $(W_s^0 - W_{t_1}^0)_{s \in [t_1, T]}$  is independent of  $\mathcal{F}_{t_1}$ , we exploit (4.11) to check that

$$\int_{\hat{\Omega}} \hat{f}(\hat{\omega}) (\phi_{\#} \mathbb{P}^{\bar{\omega}})(d\hat{\omega}) = \int_{\hat{\Omega}} \hat{f}(\hat{\omega}) \hat{\mathbb{P}}(d\hat{\omega}). \quad (4.12)$$

By the arbitrariness of  $\hat{f}$ , we infer  $\phi_{\#} \mathbb{P}^{\bar{\omega}} = \hat{\mathbb{P}}$ . For  $t \in [t_1, T]$ , we let  $\mathcal{F}_t^{\bar{\omega}}$  be the completion of the  $\sigma$ -algebra generated by  $\{W_s^0 - W_{t_1}^0 : s \in [t_1, t]\}$  with respect to the measure  $\mathbb{P}^{\bar{\omega}}$ . Since  $\mathbb{P}^{\bar{\omega}} \ll \mathbb{P}$  for  $\mathbb{P}$ -a.s.  $\bar{\omega}$ , this completion results in a larger filtration (despite having restricted the time interval) where  $\mathcal{F}_t \subset \mathcal{F}_t^{\bar{\omega}}$  for each  $t \in [t_1, T]$ .

For each  $\bar{\omega} \in \Omega$  and  $t \in [t_1, T]$ , we define a control policy on  $\hat{\Omega}$  by

$$\hat{\alpha}_t^{\bar{\omega}}(u) := \eta(t, (W_{s \wedge t_1}^0(\bar{\omega}) + u_{s \wedge t})_{0 \leq s \leq T}), \quad u \in \hat{\Omega}.$$

We have that

$$\hat{\alpha}_t^{\bar{\omega}}(\phi(\omega)) = \alpha_t^{\bar{\omega}}(\omega) \text{ for all } \omega \in \Omega \text{ and } t \in [t_1, T] \quad (4.13)$$

and

$$\hat{\alpha}_t^{\bar{\omega}}(\phi(\omega)) = \alpha_t(\omega) = \alpha_t^{\bar{\omega}}(\omega) \text{ for } \mathbb{P}^{\bar{\omega}} \text{ almost surely } \omega, \text{ and for all } t \in [t_1, T]. \quad (4.14)$$

The second inequality follows as  $W_{s \wedge t_1}^0(\omega) = W_{s \wedge t_1}^0(\bar{\omega})$  for all  $s \in [0, t_1]$  for  $\mathbb{P}^{\bar{\omega}}$ -almost every  $\omega$ . Furthermore,

$$\alpha_{\tau}(\omega) = \alpha_{\tau}(\bar{\omega}) \quad \text{for } \mathbb{P}^{\bar{\omega}} \text{ almost surely } \omega, \text{ and for all } \tau \in [0, t_1], \quad (4.15)$$

which implies

$$X_{t_1}[\tilde{\alpha}](\omega) = X_{t_1}[\tilde{\alpha}](\bar{\omega}) \quad \text{for } \mathbb{P}^{\bar{\omega}} \text{ almost surely } \omega \quad (4.16)$$

We now consider the control policy  $\hat{\alpha}^{\bar{\omega}} \in \hat{\mathbb{A}}_{\mathcal{A}, X_{t_1}[\tilde{\alpha}](\bar{\omega})}^{t_1, T}$ , using the same free filtration as for  $\tilde{\alpha}$  and the free Brownian motion  $S'_t := S_t - S_{t_1}$  for  $t \in [t_1, T]$ . We let  $(\hat{X}_t[\hat{\alpha}^{\bar{\omega}}])_{t_1 \leq t \leq T}$  denote the solution to the following SDE on the probability space  $(\hat{\Omega}, (\hat{\mathcal{F}}_t)_{t_1 \leq t \leq T}, \hat{\mathbb{P}})$  with Brownian motion  $(\widehat{W}_t^0)_{t_1 \leq t \leq T}$  and the initial condition  $X_{t_1}[\tilde{\alpha}](\bar{\omega})$ :

$$d\hat{X}_t = b_{\mathcal{A}}(\hat{X}_t, \hat{\alpha}_t) dt + \beta_C \mathbb{1}_{\mathcal{A}} d\widehat{W}_t^0 + \beta_F dS'_t. \quad (4.17)$$

Theorem 3.6 applies with this probability space, showing well-posedness. Using (4.13), we see that  $(\hat{X}_t[\hat{\alpha}^{\bar{\omega}}] \circ \phi)_{t_1 \leq t \leq T}$  is a solution of (3.10) on the original probability space  $\Omega$  with the control policy  $\tilde{\alpha}^{\bar{\omega}}$ . It follows that for all  $t \in [t_1, T]$  and  $\mathbb{P}$  almost surely

$$\hat{X}_t[\hat{\alpha}^{\bar{\omega}}] \circ \phi = X_t[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}] \quad (4.18)$$

and

$$\begin{aligned} & \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}}(\hat{X}_t[\hat{\alpha}^{\bar{\omega}}], \hat{\alpha}_t^{\bar{\omega}}) dt + g_{\mathcal{A}}(\hat{X}_T[\hat{\alpha}^{\bar{\omega}}]) \right] \\ &= \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}], \alpha_t^{\bar{\omega}}) dt + g_{\mathcal{A}}(X_T[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}]) \right]. \end{aligned} \quad (4.19)$$

Note that  $(W_t^0 - W_{t_1}^0)_{t_1 \leq t \leq T}$  is a Brownian motion with filtration  $(\mathcal{F}_t^{\bar{\omega}})_{t_1 \leq t \leq T}$  on  $(\Omega, \mathcal{F}_T^{\bar{\omega}}, \mathbb{P}^{\bar{\omega}})$ . Let  $X_t^{\bar{\omega}}[\tilde{\alpha}]$  denote the solution to (3.10) on the probability space  $(\Omega, \mathcal{F}_T^{\bar{\omega}}, (\mathcal{F}_t^{\bar{\omega}})_{t_1 \leq t \leq T}, \mathbb{P}^{\bar{\omega}})$  with the initial condition  $X_{t_1}[\tilde{\alpha}](\bar{\omega})$ :

$$\begin{aligned} X_t^{\bar{\omega}}[\tilde{\alpha}](\omega) &= X_{t_1}[\tilde{\alpha}](\bar{\omega}) + \int_{t_1}^t b_{\mathcal{A}}(X_s^{\bar{\omega}}[\tilde{\alpha}](\omega), \alpha_s(\omega)) ds \\ &\quad + \beta_C \mathbb{1}_{\mathcal{A}}(W_t^0(\omega) - W_{t_1}^0(\omega)) + \beta_F (S_t - S_{t_1}). \end{aligned} \quad (4.20)$$

This holds  $\mathbb{P}^{\bar{\omega}}$  almost everywhere, and  $(X_t^{\bar{\omega}}[\tilde{\alpha}])_{t_1 \leq t \leq T}$  is  $(\mathcal{F}_t^{\bar{\omega}})_{t_1 \leq t \leq T}$  progressively measurable. Since  $\mathbb{P}^{\bar{\omega}} \ll \mathbb{P}$  and  $\mathcal{F}_t \subset \mathcal{F}_t^{\bar{\omega}}$ , for  $\mathbb{P}$  a.s.  $\bar{\omega}$  we also have that  $X_t[\tilde{\alpha}]$  satisfies (4.20)  $\mathbb{P}^{\bar{\omega}}$  a.s. and is  $(\mathcal{F}_t^{\bar{\omega}})_{t_1 \leq t \leq T}$  progressively measurable. Thus, using (4.16) and the uniqueness of solutions we deduce that for  $\mathbb{P}$  a.s.  $\bar{\omega}$ , we have  $X_t^{\bar{\omega}}[\tilde{\alpha}](\omega) = X_t[\tilde{\alpha}](\omega)$  for  $\mathbb{P}^{\bar{\omega}}$  almost every  $\omega$ . By the definition of the conditional probability

$$\mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{A}}(X_T[\tilde{\alpha}]) \middle| \mathcal{F}_{t_1} \right] (\bar{\omega}) = \mathbb{E}^{\bar{\omega}} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t^{\bar{\omega}}[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{A}}(X_T^{\bar{\omega}}[\tilde{\alpha}]) \right].$$

Finally, we have that  $\mathbb{P}^{\bar{\omega}}$  almost surely

$$\hat{X}_t[\hat{\alpha}^{\bar{\omega}}] \circ \phi = X_t^{\bar{\omega}}[\tilde{\alpha}],$$

and thus recalling  $\phi_{\#} \mathbb{P}^{\bar{\omega}} = \hat{\mathbb{P}}$ ,

$$\begin{aligned} &\mathbb{E}^{\bar{\omega}} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t^{\bar{\omega}}[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{A}}(X_T^{\bar{\omega}}[\tilde{\alpha}]) \right] \\ &= \hat{\mathbb{E}} \left[ \int_{t_1}^T L_{\mathcal{A}}(\hat{X}_t[\hat{\alpha}^{\bar{\omega}}], \hat{\alpha}_t^{\bar{\omega}}) dt + g_{\mathcal{A}}(\hat{X}_T[\hat{\alpha}^{\bar{\omega}}]) \right] \\ &\stackrel{(4.19)}{=} \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}], \alpha_t^{\bar{\omega}}) dt + g_{\mathcal{A}}(X_T[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}]) \right]. \end{aligned}$$

Therefore we conclude the proof.  $\square$

**Proposition 4.7.** *Suppose that Assumption A holds. The value function  $\bar{V}$ , defined in (3.13), satisfies the sub-dynamic programming principle: For any  $\mathcal{A} \in \mathbb{W}$ ,  $x_0 \in L^2(\mathcal{A})_{sa}^d$  and  $0 \leq t_0 \leq t_1 \leq T$ ,*

$$\bar{V}_{\mathcal{A}}(t_0, x_0) \leq \inf_{\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, t_1}} \left\{ \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{A}}(X_s[\tilde{\alpha}], \alpha_s) ds + \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}]) \right] : X_{t_0}[\tilde{\alpha}] = x_0 \right\}. \quad (4.21)$$

Furthermore,  $\bar{V}$  is a free viscosity subsolution in the sense of Definition 4.2.

*Proof.* First note that if  $\mathbb{A}_{\mathcal{A}, x_0}^{t_0, t_1}$  is empty, i.e.,  $\mathcal{A}$  does not support a free Brownian motion independent of  $x_0$ , then RHS of (4.21) is  $+\infty$  and the inequality holds. Hence let us assume that  $\mathcal{A}$  supports a free Brownian motion on  $[t_0, t_1]$  independent of  $x_0$ .

Given an arbitrary control policy  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, t_1}$ , we will construct a larger space  $\mathcal{B} \in \mathbb{W}$  with a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  and control policy  $\tilde{\alpha}' \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, T}$  which is  $\epsilon$ -optimal on  $[t_1, T]$ , i.e.,

$$\mathbb{E} \left[ \bar{V}_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}']) \right] \geq -\epsilon + \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}'], \alpha'_t) dt + g_{\mathcal{B}}(X_T[\tilde{\alpha}']) \right]. \quad (4.22)$$

Furthermore, the control policy will be consistent on  $[t_0, t_1]$  in the sense that  $\alpha'_t = \iota \alpha_t$  and  $X_t[\tilde{\alpha}'] = \iota X_t[\tilde{\alpha}]$  on  $[t_0, t_1]$ . Indeed, given this consistency, we have

$$\begin{aligned} \bar{V}_{\mathcal{A}}(t_0, x_0) &\leq \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}'], \alpha'_t) dt + g_{\mathcal{B}}(X_T[\tilde{\alpha}']) \right] \\ &= \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_t[\tilde{\alpha}'], \alpha'_t) dt \right] + \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}'], \alpha'_t) dt + g_{\mathcal{B}}(X_T[\tilde{\alpha}']) \right] \\ &\stackrel{(4.22)}{\leq} \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{A}}(X_t[\tilde{\alpha}], \alpha_t) dt \right] + \mathbb{E} \left[ \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}]) \right] + \epsilon, \end{aligned}$$

where the last inequality follows from the tracial property of the value function  $(\bar{V}_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$ . Taking  $\epsilon \rightarrow 0^+$  proves the sub-dynamic programming principle.

In order to construct a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  along with a consistent policy  $\tilde{\alpha}' \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, T}$  satisfying (4.22), we reduce to taking a finite collection of points in the von Neumann algebra that approximate  $X_{t_1}[\tilde{\alpha}]$  and apply Lemma 3.7 to amalgamate spaces with nearly optimal control policies for each point. Let  $(\mathcal{A}_t)_{t \in [t_0, T]}$  be the free filtration and  $(S_t^0)_{t \in [t_0, T]}$  be the compatible free Brownian motion of  $\tilde{\alpha}$ .

We recall that by adaptedness,  $X_{t_1}[\tilde{\alpha}] \in L^2(\mathcal{A}_{t_1})_{sa}^d$ . By separability of the space  $L^2(\mathcal{A}_{t_1})_{sa}^d$ , there is a countable dense collection  $\{z_k\}_{k=1}^\infty$  in  $L^2(\mathcal{A}_{t_1})_{sa}^d$ . For a positive integer  $N$ , let  $K^{N,r} := \cup_{k=1}^N B_r(z_k)$  be the union of balls of radius  $r > 0$  around the first  $N$  points. Then, for any  $r > 0$ ,

$$\lim_{N \rightarrow \infty} \mathbb{P}(X_{t_1}[\tilde{\alpha}] \notin K^{N,r}) = \mathbb{P}(X_{t_1}[\tilde{\alpha}] \notin \cup_{k=1}^\infty B_r(z_k)) = 0. \quad (4.23)$$

We claim that for any  $r > 0$  and  $\epsilon > 0$ , there exists sufficiently large  $N$  such that

$$\mathbb{E} \left[ \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}]) \right] \geq -\epsilon + \mathbb{E} \left[ \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K^{N,r}} \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}]) \right]. \quad (4.24)$$

Indeed, by Hölder inequality,

$$\begin{aligned} \mathbb{E} \left[ \left| \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \notin K^{N,r}} \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}]) \right| \right] &\leq \sqrt{\mathbb{P}(X_{t_1}[\tilde{\alpha}] \notin K^{N,r})} \sqrt{\mathbb{E} |\bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}])|^2} \\ &\stackrel{(3.30)}{\leq} C \sqrt{\mathbb{P}(X_{t_1}[\tilde{\alpha}] \notin K^{N,r})} \sqrt{1 + \mathbb{E} \|X_{t_1}[\tilde{\alpha}]\|_{L^2(\mathcal{A})}^2}. \end{aligned}$$

By (3.21) in Theorem 3.6,

$$\mathbb{E} [\|X_{t_1}[\tilde{\alpha}]\|_{L^2(\mathcal{A})}^2] \leq 2\mathbb{E} [\|X_{t_1}[\tilde{\alpha}] - x_0\|_{L^2(\mathcal{A})}^2] + 2\|x_0\|_{L^2(\mathcal{A})}^2 \leq \tilde{C}_{M,N} (t_1 - t_0) + 2\|x_0\|_{L^2(\mathcal{A})}^2.$$

This together with (4.23) verifies (4.24).

Take a partition  $K^{N,r} = \sqcup_{k=1}^N K_k$  such that  $K_k \subset B_r(z_k)$ . We choose

$$r = \frac{\epsilon}{4C_2(1+T)\sqrt{\tilde{C}}}, \quad (4.25)$$

using the Lipschitz coefficient of Proposition 3.12 so that  $\bar{V}(t_1, \cdot)$  does not oscillate by more than  $\epsilon$  in a ball of radius  $r$ . Setting  $c_k := \mathbb{P}(X_{t_1}[\tilde{\alpha}] \in K_k)$ ,

$$\mathbb{E} \left[ \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K^{N,r}} \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}]) \right] = \sum_{k=1}^N \mathbb{E} \left[ \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K_k} \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}]) \right]$$

$$\begin{aligned}
&\geq \sum_{k=1}^N \mathbb{E} \left[ \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K_k} (\bar{V}_{\mathcal{A}}(t_1, z_k) - \epsilon) \right] \\
&\geq -\epsilon + \sum_{k=1}^N c_k \bar{V}_{\mathcal{A}}(t_1, z_k).
\end{aligned} \tag{4.26}$$

Thus combining this with (4.24), we obtain

$$\sum_{k=1}^N c_k \bar{V}_{\mathcal{A}}(t_1, z_k) \leq \mathbb{E} [\bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}])] + 2\epsilon. \tag{4.27}$$

We now apply Lemma 3.8, for each  $k \in \{1, \dots, N\}$ , to infer the existence of a tracial  $W^*$ -algebra  $\mathcal{B}^k$ , a tracial  $W^*$ -embedding  $\iota_k : \mathcal{A} \rightarrow \mathcal{B}^k$ , a control policy  $\tilde{\alpha}^k \in \mathbb{A}_{\mathcal{B}^k, \iota_k z_k}^{t_1, T}$  with filtration  $(\mathcal{B}_t^k)_{t \in [t_1, T]}$  and compatible  $d$ -variable free Brownian motion  $(S_t^k)_{t \in [t_1, T]}$ , such that  $\iota_k(\mathcal{A}) \subset \mathcal{B}_{t_1}^k$  and

$$\bar{V}_{\mathcal{A}}(t_1, z_k) \geq -\epsilon + \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}^k}(X_t[t_1, \iota_k z_k, \tilde{\alpha}^k], \alpha_t^k) dt + g_{\mathcal{B}^k}(X_T[t_1, \iota_k z_k, \tilde{\alpha}^k]) \right]. \tag{4.28}$$

Furthermore, we may assume that  $(\alpha_t^k)_{t \in [t_1, T]}$  is independent of  $\mathcal{F}_{t_1}$ . To accomplish this, we first take  $\bar{\omega} \in \Omega$  such that

$$\begin{aligned}
&\mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}^k}(X_t[t_1, \iota_k z_k, \tilde{\alpha}^k], \alpha_t^k) dt + g_{\mathcal{B}^k}(X_T[t_1, \iota_k z_k, \tilde{\alpha}^k]) \right] \\
&\geq \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}^k}(X_t[t_1, \iota_k z_k, \tilde{\alpha}^k], \alpha_t^k) dt + g_{\mathcal{B}^k}(X_T[t_1, \iota_k z_k, \tilde{\alpha}^k]) \mid \mathcal{F}_{t_1} \right](\bar{\omega}).
\end{aligned}$$

By Lemma 4.6 with  $t_0 = t_1$ , redefining  $\alpha_t^k := (\alpha_t^k)^{\bar{\omega}}$ , we arrive at a process, satisfying (4.28), that is  $\mathcal{F}_{t_1}$ -independent for all  $t \in [t_1, T]$ , as required.

By construction we have  $\iota_k(\mathcal{A}_{t_1}) \subset \mathcal{B}_{t_1}^k$ , so we may apply the amalgamation Lemma 3.7 to find a larger algebra  $\mathcal{B}$  with tracial  $W^*$  embeddings  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  and  $\tilde{\iota}_k : \mathcal{B}^k \rightarrow \mathcal{B}$ , a filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$ , and a compatible free Brownian motion  $(S_t)_{t \in [t_0, T]}$  such that  $\iota(S_t^0 - S_{t_0}^0) = S_t - S_{t_0}$  for  $t \in [t_0, t_1]$  and  $\tilde{\iota}_k(S_t^k - S_{t_1}^k) = S_t - S_{t_1}$  for each  $k$  and  $t \in [t_1, T]$ . We define a control policy  $\tilde{\alpha}'_t \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, T}$  using the filtration  $(\mathcal{B}_t)_t$ , free Brownian motion  $(S_t)_t$  and the control  $(\alpha'_t)_t$  defined as follows:  $\alpha'_t := \iota \alpha_t$  for  $t \in [t_0, t_1)$ , and on  $[t_1, T]$

$$\alpha'_t := \begin{cases} \tilde{\iota}_k \alpha_t^k & \text{on the event } X_{t_1}[\tilde{\alpha}] \in K_k \text{ for some } k = 1, \dots, N, \\ 0 & \text{otherwise.} \end{cases} \tag{4.29}$$

Note that the properties of Lemma 3.7 imply that  $(\alpha'_t)_{t \in [t_0, T]}$  is adapted to  $(\mathcal{B}_t)_{t \in [t_0, T]}$ . Also,  $(\alpha'_t)_{t \in [t_0, T]}$  is adapted to  $(\mathcal{F}_t)_{t \in [t_0, T]}$ .

For  $t \in [t_0, t_1]$ , both  $\iota X_t[t_0, x_0, \tilde{\alpha}]$  and  $X_t[t_0, \iota x_0, \tilde{\alpha}']$  solve the same SDE and thus agree by Theorem 3.6. Hence they are consistent, i.e.  $X_t[\tilde{\alpha}'] = \iota X_t[\tilde{\alpha}]$  (we omit the initial time and position for the sake of readability). Similarly, for  $t \in [t_1, T]$ , on the event  $X_{t_1}[\tilde{\alpha}] \in K_k$ ,

$$\begin{aligned}
d\tilde{\iota}_k X_t[t_1, \iota_k X_{t_1}[\tilde{\alpha}], \tilde{\alpha}^k] &= \tilde{\iota}_k b_{\mathcal{B}^k}(X_t[t_1, \iota_k X_{t_1}[\tilde{\alpha}], \tilde{\alpha}^k], \alpha_t^k) dt + \beta_C \tilde{\iota}_k \mathbb{1}_{\mathcal{B}^k} dW_t^0 + \beta_F \tilde{\iota}_k dS_t^k \\
&= b_{\mathcal{B}}(\tilde{\iota}_k X_t[t_1, \iota_k X_{t_1}[\tilde{\alpha}], \tilde{\alpha}^k], \alpha'_t) dt + \beta_C \mathbb{1}_{\mathcal{B}} dW_t^0 + \beta_F dS_t,
\end{aligned}$$

recalling that the free SDE formulation is shorthand for an integral form where  $\int_{t_1}^t \beta_F \tilde{\iota}_k dS_t^k = \beta_F \tilde{\iota}_k (S_t^k - S_{t_1}^k)$  for  $t \in [t_1, T]$ . Thus it follows that  $X_t[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}'] = \tilde{\iota}_k X_t[t_1, \iota_k X_{t_1}[\tilde{\alpha}], \tilde{\alpha}^k]$ , since they solve the same SDE and thus agree by Theorem 3.6.



Then we bound RHS of (4.22) as follows: Noting that  $X_t[\tilde{\alpha}'] = X_t[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}']$ ,

$$\begin{aligned}
 & \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_t[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}'], \alpha'_t) dt + g_{\mathcal{B}}(X_T[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}']) \right] \\
 &= \mathbb{E} \left[ \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_t[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}'], \alpha'_t) dt + g_{\mathcal{B}}(X_T[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}']) \mid \mathcal{F}_{t_1} \right] \right] \\
 &\leq \sum_{k=1}^N \mathbb{E} \left[ \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_t[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}'], \alpha'_t) dt + g_{\mathcal{B}}(X_T[t_1, \iota X_{t_1}[\tilde{\alpha}], \tilde{\alpha}']) \mid \mathcal{F}_{t_1} \right] \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K_k} \right] \\
 &\quad + C \mathbb{E} \left[ 1 + \|X_{t_1}[\tilde{\alpha}]\|_{L^2(\mathcal{A})}^2 + \int_{t_1}^T \|\alpha'_t\|_{L^2(\mathcal{A})}^2 dt \right] \mathbb{P}(X_{t_1}[\tilde{\alpha}] \notin K^{N,r}), \tag{4.30}
 \end{aligned}$$

where we used the upper bound condition (3.8) and (3.21) in Theorem 3.6 to obtain the last term. Observe that by the Lipschitz condition (3.9) along with (3.22), for any  $\mathcal{A}' \in \mathbb{W}$ ,  $y_1, y_2 \in L^2(\mathcal{A}')_{sa}^d$  and a control  $\beta = (\beta_t)_t$ ,

$$\begin{aligned}
 & \left| \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}'}(X_t[t_1, y_1, \beta], \beta_t) dt + g_{\mathcal{A}'}(X_T[t_1, y_1, \beta]) \right] \right. \\
 & \quad \left. - \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}'}(X_t[t_1, y_2, \beta], \beta_t) dt + g_{\mathcal{A}'}(X_T[t_1, y_2, \beta]) \right] \right| \leq C \|y_1 - y_2\|_{L^2(\mathcal{A})}. \tag{4.31}
 \end{aligned}$$

Now recall the definition of  $(\alpha'_t)_t$  in (4.29) and  $\|X_{t_1}[\tilde{\alpha}] - z_k\|_{L^2(\mathcal{A})} \leq r$  on the event  $X_{t_1}[\tilde{\alpha}] \in K_k$ . The independence of  $(\alpha'_t)_{t \in [t_1, T]}$  and  $\mathcal{F}_{t_1}$  also implies independence of  $(X_t[t_1, \iota_k z_k, \tilde{\alpha}^k])_{t \in [t_1, T]}$  and  $\mathcal{F}_{t_1}$ . We get that the first term in (4.30) is bounded by

$$\begin{aligned}
 & \sum_{k=1}^N \mathbb{E} \left[ \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}^k}(X_t[t_1, \iota_k X_{t_1}[\tilde{\alpha}], \tilde{\alpha}^k], \alpha'_t) dt + g_{\mathcal{B}^k}(X_T[t_1, \iota_k X_{t_1}[\tilde{\alpha}], \tilde{\alpha}^k]) \mid \mathcal{F}_{t_1} \right] \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K_k} \right] \\
 &\stackrel{(4.31)}{\leq} \sum_{k=1}^N \mathbb{E} \left[ \left( \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}^k}(X_t[t_1, \iota_k z_k, \tilde{\alpha}^k], \alpha'_t) dt + g_{\mathcal{B}^k}(X_T[t_1, \iota_k z_k, \tilde{\alpha}^k]) \right] + Cr \right) \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K_k} \right] \\
 &\stackrel{(4.28)}{\leq} \sum_{k=1}^N \mathbb{E} \left[ (\bar{V}_{\mathcal{A}}(t_1, z_k) + \epsilon) \mathbb{1}_{X_{t_1}[\tilde{\alpha}] \in K_k} \right] + Cr \\
 &\leq \sum_{k=1}^N c_k \bar{V}_{\mathcal{A}}(t_1, z_k) + \epsilon + Cr \\
 &\stackrel{(4.27)}{\leq} \mathbb{E}[\bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}])] + 3\epsilon + Cr \stackrel{(4.25)}{=} \mathbb{E}[\bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}])] + C'\epsilon.
 \end{aligned}$$

Next, the second term in (4.30) converges to 0, due to (4.23) and the boundedness of  $\mathbb{E}[\|X_{t_1}[\tilde{\alpha}]\|_{L^2(\mathcal{A})}^2]$  (see (3.21) in Theorem 3.6). Thus noting  $\bar{V}_{\mathcal{B}}(t_1, \iota X_{t_1}[\tilde{\alpha}]) = \bar{V}_{\mathcal{A}}(t_1, X_{t_1}[\tilde{\alpha}])$ , we obtain (4.22) and thus establish the sub-dynamic programming principle.

We now prove the viscosity subsolution property. Suppose that  $(\Phi_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$  touches  $(\bar{V}_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  from above at  $(t_0, \lambda_0) \in [0, T] \times \Sigma_d^2$ . Let  $\mathcal{A} \in \mathbb{W}$  and  $x_0 \in L^2(\mathcal{A})_{sa}^d$  such that  $\lambda_0 = \lambda_{x_0}$ . Take a larger von Neumann algebra  $\mathcal{B}$ , along with a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , which contains a free semi-circular process  $(S_t)_t$  freely independent of  $\iota(\mathcal{A})$ . Then for any  $\bar{\alpha} \in \mathbb{A}_{\mathcal{A}}$ , consider  $\tilde{\alpha} = ((\alpha_t)_t, (\mathcal{A}_t)_t, (S_t)_t) \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, t_1}$  such that  $\alpha_t = \iota \bar{\alpha}$  for  $t \in [t_0, T]$ , and then let  $X_t[\tilde{\alpha}]$  be

a solution to (3.10) on  $L^2(\mathcal{B})_{sa}^d$  with an initial condition  $X_{t_0}[\tilde{\alpha}] = \iota x_0$ . As the non-commutative law of  $\iota x_0$ , as an element in  $\mathcal{B}$ , is  $\lambda_0$ , for any  $0 \leq t_0 < t_1 \leq T$ ,

$$\begin{aligned} \Phi_{\mathcal{B}}(t_0, \iota x_0) &= \bar{V}_{\mathcal{B}}(t_0, \iota x_0) \stackrel{(4.21)}{\leq} \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds + \bar{V}_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}]) \right] \\ &\leq \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds + \Phi_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}]) \right]. \end{aligned}$$

Using Lemma 4.5, abbreviating  $X_t[\tilde{\alpha}]$  to  $X_t$ ,

$$\begin{aligned} 0 \leq \mathbb{E} \left[ \int_{t_0}^{t_1} \left( L_{\mathcal{B}}(X_s, \iota \bar{\alpha}) + \partial_t \Phi_{\mathcal{B}}(s, X_s) \right. \right. \\ \left. \left. + \langle \nabla \Phi_{\mathcal{B}}(s, X_s), b_{\mathcal{B}}(X_s, \iota \bar{\alpha}) \rangle_{L^2(\mathcal{B})} + \frac{\beta_C^2}{2} \Delta_{\mathcal{B}} \Phi(s, X_s) + \frac{\beta_F^2}{2} \Theta_{\mathcal{B}} \Phi(s, X_s) \right) ds \right]. \end{aligned}$$

Dividing by  $t_1 - t_0$  and then taking the limit  $t_1 \rightarrow^+ t_0$ , we find that

$$\begin{aligned} -\partial_t \Phi_{\mathcal{B}}(t_0, \iota x_0) - \langle \nabla \Phi_{\mathcal{B}}(t_0, \iota x_0), b_{\mathcal{B}}(\iota x_0, \iota \bar{\alpha}) \rangle_{L^2(\mathcal{B})} - L_{\mathcal{B}}(\iota x_0, \iota \bar{\alpha}) \\ - \frac{\beta_C^2}{2} \Delta_{\mathcal{B}} \Phi(t_0, \iota x_0) - \frac{\beta_F^2}{2} \Theta_{\mathcal{B}} \Phi(t_0, \iota x_0) \leq 0. \end{aligned} \quad (4.32)$$

Note that as  $\bar{\alpha} \in L^2(\mathcal{A})_{sa}^d$ ,

$$\langle \nabla \Phi_{\mathcal{B}}(\iota x_0, \iota \bar{\alpha}), b_{\mathcal{B}}(\iota x_0, \iota \bar{\alpha}) \rangle_{L^2(\mathcal{B})} = \langle \nabla \Phi_{\mathcal{A}}(t_0, x_0), b_{\mathcal{A}}(x_0, \bar{\alpha}) \rangle_{L^2(\mathcal{A})}.$$

Hence since all the functions in (4.32) are tracial  $W^*$ -functions, taking the supremum over all  $\bar{\alpha} \in \mathbb{A}_{\mathcal{A}}$ , by (3.16) in Lemma 3.5,

$$-\partial_t \Phi_{\mathcal{A}}(t_0, x_0) + H_{\mathcal{A}}(x_0, -\nabla \Phi_{\mathcal{A}}(t_0, x_0)) - \frac{\beta_C^2}{2} \Delta_{\mathcal{A}} \Phi(t_0, x_0) - \frac{\beta_F^2}{2} \Theta_{\mathcal{A}} \Phi(t_0, x_0) \leq 0.$$

As this holds for arbitrary  $\mathcal{A} \in \mathbb{W}$  which contains a non-commutative law  $\lambda_0$ , one can take a supremum over all such  $\mathcal{A} \in \mathbb{W}$  and thus establish the free viscosity subsolution property of  $\bar{V}$ .  $\square$

We contrast this with the proof that the value function on a given von Neumann algebra satisfies the superdynamic programming principle and is a supersolution.

**Proposition 4.8.** *Suppose that Assumption A holds. The value function  $\bar{V}$ , defined in (3.13), satisfies the super-dynamic programming principle: For any  $\mathcal{A} \in \mathbb{W}$ ,  $x_0 \in L^2(\mathcal{A})_{sa}^d$  and  $0 \leq t_0 \leq t_1 \leq T$ ,*

$$\bar{V}_{\mathcal{A}}(t_0, x_0) \geq \inf_{\iota: \mathcal{A} \rightarrow \mathcal{B}} \inf_{\tilde{\alpha} \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, t_1}} \left\{ \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds + \bar{V}_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}]) \right] : X_{t_0}[\tilde{\alpha}] = \iota x_0 \right\}.$$

Furthermore,  $\bar{V}$  is a free viscosity supersolution in the sense of Definition 4.2.

*Proof.* For any  $t_0 \in [0, T]$ ,  $x_0 \in L^2(\mathcal{A})_{sa}^d$  and  $\epsilon > 0$ , there exist  $\mathcal{B} \in \mathbb{W}$  along with a tracial  $W^*$ -embedding  $\iota: \mathcal{A} \rightarrow \mathcal{B}$  and a control policy  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, T}$  such that for the corresponding solution  $(X_t[\tilde{\alpha}])_t$  to (3.10) on  $L^2(\mathcal{B})_{sa}^d$  with  $X_{t_0}[\tilde{\alpha}] = \iota x_0$ ,

$$\begin{aligned} \bar{V}_{\mathcal{A}}(t_0, x_0) &\geq -\epsilon + \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds + g_{\mathcal{B}}(X_T[\tilde{\alpha}]) \right] \\ &= -\epsilon + \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds + \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds + g_{\mathcal{B}}(X_T[\tilde{\alpha}]) \mid \mathcal{F}_{t_1} \right] \right] \end{aligned}$$

$$\begin{aligned}
 &= -\epsilon + \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds \right. \\
 &\quad \left. + \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_s[t_1, X_{t_1}[\tilde{\alpha}], \tilde{\alpha}], \alpha_s) ds + g_{\mathcal{B}}(X_T[t_1, X_{t_1}[\tilde{\alpha}], \tilde{\alpha}]) \mid \mathcal{F}_{t_1} \right] \right]
 \end{aligned}$$

By Lemma 4.6, for  $\mathbb{P}$ -a.s.  $\bar{\omega}$ , there exists a control policy  $\tilde{\alpha}^{\bar{\omega}}$  such that

$$\begin{aligned}
 &\mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{B}}(X_s[t_1, X_{t_1}[\tilde{\alpha}], \tilde{\alpha}], \alpha_s) ds + g_{\mathcal{B}}(X_T[t_1, X_{t_1}[\tilde{\alpha}], \tilde{\alpha}]) \mid \mathcal{F}_{t_1} \right] (\bar{\omega}) \\
 &= \mathbb{E} \left[ \int_{t_1}^T L_{\mathcal{A}}(X_t[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}], \alpha_t^{\bar{\omega}}) dt + g_{\mathcal{A}}(X_T[t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega}), \tilde{\alpha}^{\bar{\omega}}]) \right] \\
 &\geq \bar{V}_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}](\bar{\omega})).
 \end{aligned}$$

Combining the above inequalities, we obtain

$$\bar{V}_{\mathcal{A}}(t_0, x_0) \geq -\epsilon + \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_s[\tilde{\alpha}], \alpha_s) ds + \bar{V}_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}]) \right].$$

As  $\epsilon > 0$  is arbitrary, we establish the super-dynamic programming principle.

Next, in order to show that  $\bar{V}$  is a free viscosity supersolution, we assume that  $(\Phi_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}} \in \mathcal{X}_{\Sigma}$  touches  $(\bar{V}_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  from below at  $(t_0, \lambda_0) \in [0, T) \times \Sigma_d^2$ . Let  $\mathcal{A} \in \mathbb{W}$  and  $x_0 \in L^2(\mathcal{A})_{sa}^d$  such that  $\lambda_0 = \lambda_{x_0}$ . For  $t_1 \in (t_0, T]$  and  $\epsilon > 0$ , we choose  $\mathcal{B} \in \mathbb{W}$  along with a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  and  $\tilde{\alpha} \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, t_1}$  such that

$$\begin{aligned}
 \epsilon(t_1 - t_0) &\geq \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) ds - \bar{V}_{\mathcal{A}}(t_0, x_0) + \bar{V}_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}]) \right] \\
 &\geq \mathbb{E} \left[ \int_{t_0}^{t_1} L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) ds - \Phi_{\mathcal{B}}(t_0, \iota x_0) + \Phi_{\mathcal{B}}(t_1, X_{t_1}[\tilde{\alpha}]) \right].
 \end{aligned}$$

Abbreviating  $X_t[\tilde{\alpha}]$  to  $X_t$ , by Lemma 4.5,

$$\begin{aligned}
 \epsilon &\geq \frac{1}{t_1 - t_0} \mathbb{E} \left[ \int_{t_0}^{t_1} \left( L_{\mathcal{B}}(X_t, \alpha_t) + \partial_t \Phi_{\mathcal{B}}(t, X_t) + \langle \nabla \Phi_{\mathcal{B}}(t, X_t), b_{\mathcal{B}}(X_t, \alpha_t) \rangle_{L^2(\mathcal{B})} \right. \right. \\
 &\quad \left. \left. + \frac{\beta_C^2}{2} \Delta_{\mathcal{B}} \Phi(t, X_t) + \frac{\beta_F^2}{2} \Theta_{\mathcal{B}} \Phi(t, X_t) \right) dt \right] \\
 &\geq \frac{1}{t_1 - t_0} \mathbb{E} \left[ \int_{t_0}^{t_1} \left( \partial_t \Phi_{\mathcal{B}}(t, X_t) - H_{\mathcal{B}}(X_t, -\nabla \Phi_{\mathcal{B}}(t, X_t)) + \frac{\beta_C^2}{2} \Delta_{\mathcal{B}} \Phi(t, X_t) + \frac{\beta_F^2}{2} \Theta_{\mathcal{B}} \Phi(t, X_t) \right) dt \right].
 \end{aligned}$$

Taking  $\epsilon \rightarrow^+ 0$  and then  $t_1 \rightarrow^+ t_0$ , setting  $y_0 := \iota x_0$ ,

$$-\partial_t \Phi_{\mathcal{B}}(t_0, y_0) + H_{\mathcal{B}}(y_0, -\nabla \Phi_{\mathcal{B}}(t_0, y_0)) - \frac{\beta_C^2}{2} \Delta_{\mathcal{B}} \Phi(t_0, y_0) - \frac{\beta_F^2}{2} \Theta_{\mathcal{B}} \Phi(t_0, y_0) \geq 0.$$

This shows that  $\bar{V}$  is a free viscosity supersolution.  $\square$

By Propositions 4.7 and 4.8, we establish the following theorem.

**Theorem 4.9.** *Suppose that Assumption A holds. Then the value function  $\bar{V}$ , defined in (3.13), is a free viscosity solution.*

**4.2. Hilbert Space Viscosity Solution.** An alternative definition is to make use of the Hilbert spaces in the following way. As will be explained below, this definition only applies when there is *no* free individual noise, and thus throughout this section we assume that  $\beta_F = 0$ . Recall that for a function  $V : [0, T] \times \Sigma_d^2 \rightarrow \mathbb{R}$ , for each  $\mathcal{A} \in \mathbb{W}$ ,  $V_{\mathcal{A}}$  is defined as in (4.3).

**Definition 4.10.** We say that a  $V \in \text{USC}([0, T] \times \Sigma_d^2)$  is a *Hilbert space viscosity subsolution* to (4.1) if for every Von Neumann algebra  $\mathcal{A} \in \mathbb{W}$ ,  $V_{\mathcal{A}}$  is a viscosity subsolution (in a Hilbert sense) to (4.2) on  $L^2(\mathcal{A})_{sa}^d$ .

Recall that the free individual noise Laplacian (4.5) is defined for a tracial  $W^*$ -function, *not* for a function on a fixed von Neumann algebra. Thus the above definition makes only sense in the absence of the free individual noise, i.e.  $\beta_F = 0$ .

**Definition 4.11.** We say that a  $V \in \text{LSC}([0, T] \times \Sigma_d^2)$  is a *Hilbert space viscosity supersolution* to (4.1) if for every  $(t, \lambda) \in [0, T] \times \Sigma_d^2$ , there exist a Von Neumann algebra  $\mathcal{A} \in \mathbb{W}$  and  $x \in L^2(\mathcal{A})_{sa}^d$  with  $\lambda = \lambda_x$ , such that there is a viscosity supersolution  $U_{\mathcal{A}}$  (in a Hilbert sense) to (4.2) on  $L^2(\mathcal{A})_{sa}^d$  with  $U_{\mathcal{A}}(t, x) = V(t, \lambda)$ .

We say  $V \in C([0, T] \times \Sigma_d^2)$  is a *Hilbert space viscosity solution* if it is both a Hilbert space viscosity subsolution and supersolution. We will omit some details from the proof that the value function is a Hilbert space viscosity solution, as they can be found in the literature on stochastic control in Hilbert spaces and are mostly repetitive with our results for the intrinsic viscosity solution.

**Proposition 4.12.** *Suppose that Assumption A holds. If there is no free noise (i.e.  $\beta_F = 0$ ), then the value function  $\bar{V}$ , defined in (3.13), is a Hilbert space viscosity solution.*

*Proof.* We first establish that  $\bar{V}$  is a Hilbert space viscosity subsolution. This follows from the subdynamic programming principle (Proposition 4.7), along with an Itô formula for the test functions  $\Phi \in \chi_{\mathcal{A}}$ . These results are standard and also very similar to the arguments above.

To establish that  $\bar{V}$  is a Hilbert space viscosity supersolution, consider that, by a standard Hilbert space theory, for any  $\mathcal{A} \in \mathbb{W}$ , the value function  $\tilde{V}_{\mathcal{A}}$ , defined in (3.11), is a viscosity solution on  $L^2(\mathcal{A})_{sa}^d$ . We fix  $(t_0, \lambda_0) \in [0, T] \times \Sigma_d^2$  and consider a minimizing sequence of Von Neumann algebras,  $\mathcal{A}_i \in \mathbb{W}$ , states  $x_i \in L^2(\mathcal{A}_i)_{sa}^d$  such that  $\lambda_{x_i} = \lambda_0$ , and controls,  $\tilde{\alpha}^i \in \mathbb{A}_{\mathcal{A}, x_i}^{t_0, T}$  for the variational problem (3.12). By considering the free product of these Von Neumann algebras,  $\mathcal{A} = *(\mathcal{A}_i)_{i=1}^{\infty}$ , and then taking  $x_0 \in \mathcal{A}$  such that  $\lambda_{x_0} = \lambda_0$ , we can restrict the control policies in  $\mathcal{A}$  to obtain

$$\bar{V}(t_0, \lambda_0) = \tilde{V}_{\mathcal{A}}(t_0, x_0) = \inf_{\tilde{\alpha} \in \mathbb{A}_{\mathcal{A}, x_0}^{t_0, T}} \left\{ \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{A}}(X_s[\tilde{\alpha}], \alpha_s) ds + g_{\mathcal{A}}(X_T[\tilde{\alpha}]) \right] \right\}.$$

Since  $\tilde{V}_{\mathcal{A}}$  is a viscosity supersolution of (4.2) on  $L^2(\mathcal{A})_{sa}^d$  and touches  $\bar{V}_{\mathcal{A}}$  from below at  $(t, \lambda_0)$ , it follows that  $\bar{V}_{\mathcal{A}}$  is supersolution. This holds for all  $(t_0, \lambda_0)$  so  $\bar{V}$  is a Hilbert space viscosity supersolution of (4.1).  $\square$

Using the comparison principle for the viscosity solutions to Hamilton-Jacobi equations on the (infinite-dimensional) Hilbert spaces (see [21, 23, 64, 27, 28, 55, 36]), one can deduce the comparison principle for the Hilbert space viscosity subsolution and supersolution.

**Theorem 4.13.** *Suppose that Assumption A holds. If  $V^1$  is a Hilbert space viscosity subsolution to (4.1) in the sense of Definition 4.10 and  $V^2$  is a Hilbert space viscosity supersolution to (4.1) in the sense of Definition 4.11, then  $V^1(t, \lambda) \leq V^2(t, \lambda)$  for all  $(t, \lambda) \in [0, T] \times \Sigma_d^2$ .*

In particular, the value function  $\bar{V}$ , defined in (3.13), is a unique Hilbert space viscosity solution to (4.1).

*Proof.* We first prove that for any fixed  $\mathcal{A} \in \mathbb{W}$ , viscosity solutions on  $L^2(\mathcal{A})_{sa}^d$  satisfy the comparison principle. We do this by checking the assumptions F6-F9 from [55, Section 6]. The result from [55] allows for more general term and singular drifts and infinite-dimensional diffusions, which complicate the matter. The equation is expressed using a perturbation equation with a stronger norm, which we may take to be  $h(x) = \frac{1}{2}\|x\|_{L^2(\mathcal{A})}^2$ . The perturbed equation is

$$F_{\mathcal{A},\delta}^{\pm}(x, u(t, x), \nabla u(t, x), \text{Hess } u(t, x)) = H_{\mathcal{A}}(x, -\nabla u(t, x) \mp \delta x) - \frac{\beta_C^2}{2} \Delta_{\mathcal{A}} u(t, x) \mp \frac{\beta_C^2 d \delta}{2}.$$

Unlike most cases considered in [55], this function is continuous at  $\delta = 0$ , and we can readily check all the conditions for  $\delta = 0$ .

- Assumption F6 states monotonicity of dependence of  $F_{\mathcal{A},\delta}^{\pm}$  on  $u$ . This assumption always holds as we have no dependence of  $F_{\mathcal{A},\delta}^{\pm}$  on  $u$  itself.
- Assumption F7 combines both a structural assumption on the diffusion and a quantitative estimate of the Hamiltonian. Since these terms are separate in our equation we may address them separately. For the diffusion, fix  $\theta > 1$ , and for  $A, B \in \text{BL}(L^2(\mathcal{A})_{sa}^d)$  such that for all  $a, b \in L^2(\mathcal{A})_{sa}^d$  and some  $\gamma > 1$

$$\langle a, Aa \rangle_{L^2(\mathcal{A})} + \langle b, Bb \rangle_{L^2(\mathcal{A})} \leq \theta \gamma \|a - b\|_{L^2(\mathcal{A})}^2.$$

To satisfy the hypothesis we must show the monotonicity of the equation (up to an error, which is not needed). This follows from

$$-\frac{\beta_C^2}{2} \langle \mathbb{1}_{\mathcal{A}}, A \mathbb{1}_{\mathcal{A}} \rangle_{L^2(\mathcal{A})} - \frac{\beta_C^2}{2} \langle \mathbb{1}_{\mathcal{A}}, B \mathbb{1}_{\mathcal{A}} \rangle_{L^2(\mathcal{A})} \geq -\frac{\beta_C^2 \theta \gamma}{2} \|\mathbb{1}_{\mathcal{A}} - \mathbb{1}_{\mathcal{A}}\|_{L^2(\mathcal{A})}^2 = 0.$$

For the Hamiltonian, for  $\gamma > 1$  and  $X, Y \in L^2(\mathcal{A})_{sa}^d$ , we must show that

$$H_{\mathcal{A}}(X, -\gamma(X - Y)) - H_{\mathcal{A}}(Y, -\gamma(X - Y)) \geq -\omega_1 \left( \gamma \|X - Y\|_{L^2(\mathcal{A})}^2 + \frac{1}{\gamma} \right)$$

for a function  $\omega_1$  on  $[0, \infty)$  with  $\lim_{y \rightarrow +0} \omega_1(y) = 0$ . This has been shown in Lemma 3.5 with

$$\omega_1(y) = \max \left\{ \bar{C} + C_2^2, \frac{1}{2} \right\} y.$$

- Hypothesis F8 states uniform continuity of the Hamiltonian and the diffusion coefficient on bounded sets. The uniform continuity of the Hamiltonian is shown in Lemma 3.5, and our diffusion coefficient is constant.
- This assumption F9 states convergence with respect to finite-dimensional approximations of the Hessian term. For the common noise, it is entirely in the one-dimensional space spanned by  $\mathbb{1}_{\mathcal{A}}$ , so the one-dimensional approximation is exact, satisfying F9.

Having met the hypotheses, we have the comparison principle for each fixed  $\mathcal{A} \in \mathbb{W}$ .

Now let us prove that  $V^1 \leq V^2$ . By Definition 4.11, for each  $(t, \lambda) \in [0, T] \times \Sigma_d^2$ , there exist a Von Neumann algebra  $\mathcal{A} \in \mathbb{W}$ ,  $x \in L^2(\mathcal{A})_{sa}^d$  with  $\lambda_x = \lambda$ , and a supersolution  $U_{\mathcal{A}}^2$  of (4.2) on  $L^2(\mathcal{A})_{sa}^d$  such that  $U_{\mathcal{A}}^2(t, x) = V^2(t, \lambda)$ . Also by Definition 4.10, for such  $\mathcal{A} \in \mathbb{W}$ , the function  $V_{\mathcal{A}}^1$ , defined as in (4.3) through a function  $V^1$ , is a viscosity subsolution of (4.2). Therefore, by the comparison principle on  $L^2(\mathcal{A})_{sa}^d$ , we have  $V_{\mathcal{A}}^1(t, x) \leq U_{\mathcal{A}}^2(t, x)$  and thus

$$V^1(t, \lambda) = V_{\mathcal{A}}^1(t, x) \leq U_{\mathcal{A}}^2(t, x) = V^2(t, \lambda).$$

The comparison principle and Proposition 4.12 immediately imply the existence and uniqueness of Hilbert space viscosity solutions.  $\square$

## 5. EXAMPLES

**5.1. Linear-quadratic setting.** We first consider the setting of Assumption **B** when  $L_{\mathcal{A}}$  and  $g_{\mathcal{A}}$  are both quadratic. These examples admit exact solutions, although they show less non-commutativity due to the cyclic property of the trace.

We consider the case where

$$\mathbb{A}_{\mathcal{A}} = L^2(\mathcal{A})_{sa}^d, \quad b_{\mathcal{A}}(X, \alpha) = \alpha, \quad L_{\mathcal{A}}(X, \alpha) = \frac{1}{2} \|\alpha\|_{L^2(\mathcal{A})}^2.$$

For coefficients  $g_{ij}^0, g_{ij}^1 \in \mathbb{R}$  such that  $g_{ij}^0 = g_{ji}^0$  and  $g_{ij}^1 = g_{ji}^1$ , let

$$g_{\mathcal{A}}(X) := \sum_{i=1}^d \sum_{j=1}^d g_{ij}^0 \tau(X^i X^j) + \sum_{i=1}^d \sum_{j=1}^d g_{ij}^1 \tau(X^i) \tau(X^j),$$

showcasing the two types of quadratic terms that may arise. More general linear-quadratic terms can be handled similarly.

The Hamilton-Jacobi equation has the form

$$-\partial_t U_{\mathcal{A}}(t, X) + \frac{1}{2} \|\nabla U_{\mathcal{A}}(t, X)\|_{L^2(\mathcal{A})}^2 - \frac{\beta_C^2}{2} \Delta_{\mathcal{A}} U(t, X) - \frac{\beta_F^2}{2} \Theta_{\mathcal{A}} U(t, X) = 0, \quad (5.1)$$

with the terminal condition

$$U_{\mathcal{A}}(T, X) = g_{\mathcal{A}}(X).$$

We begin with a candidate solution of the form

$$U_{\mathcal{A}}(t, X) = e(t) + \sum_{i=1}^d \sum_{j=1}^d a_{ij}^0(t) \tau(X^i X^j) + \sum_{i=1}^d \sum_{j=1}^d a_{ij}^1(t) \tau(X^i) \tau(X^j)$$

such that  $a_{ij}^0, a_{ij}^1 \in \mathbb{R}$  and  $a_{ij}^0(t) = a_{ji}^0(t)$  and  $a_{ij}^1(t) = a_{ji}^1(t)$ . Then we have

$$\nabla^j U_{\mathcal{A}}(t, X) = 2 \sum_{i=1}^d (a_{ij}^0(t)) X^i + 2 \sum_{i=1}^d (a_{ij}^1(t)) \tau(X^i) \mathbf{1}_{\mathcal{A}}.$$

This yields

$$\frac{1}{2} \|\nabla U_{\mathcal{A}}(t, X)\|_{L^2(\mathcal{A})}^2 = 2 \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \left( (a_{ik}^0(t) a_{kj}^0(t) \tau(X^i X^j) + (a_{ik}^0(t) + a_{ik}^1(t)) a_{kj}^1(t) \tau(X^i) \tau(X^j) \right).$$

The Hessian of  $U_{\mathcal{A}}$  is given by

$$\text{Hess } U_{\mathcal{A}}(t, X)[A, B] = 2 \sum_{i=1}^d \sum_{j=1}^d \left( a_{ij}^0(t) \tau(A^i B^j) + a_{ij}^1(t) \tau(A^i) \tau(B^j) \right),$$

implying that the common noise Laplacian is given by

$$\Delta_{\mathcal{A}} U(t, X) = 2 \sum_{i=1}^d \sum_{j=1}^d (a_{ij}^0(t) + a_{ij}^1(t)).$$

The free individual noise Laplacian is

$$\Theta_{\mathcal{A}} U(t, X) = 2 \sum_{i=1}^d a_{ii}^0(t).$$

Computing the coefficient of  $\tau(X^i X^j)$  term in (5.1),

$$-(a_{ij}^0)'(t) + 2 \sum_{k=1}^d a_{ik}^0(t) a_{kj}^0(t) = 0,$$

with  $a_{ij}^0(T) = g_{ij}^0$ . Computing the coefficient of  $\tau(X^i) \tau(X^j)$  term in (5.1),

$$-(a_{ij}^1)'(t) + 2 \sum_{k=1}^d (a_{ik}^0(t) + a_{ik}^1(t)) a_{kj}^1(t) = 0,$$

with  $a_{ij}^1(T) = g_{ij}^1$ . The constant term in (5.1) reads as

$$-e'(t) - \beta_C^2 \sum_{i=1}^d \sum_{j=1}^d (a_{ij}^0(t) + a_{ij}^1(t)) - \beta_F^2 \sum_{i=1}^d a_{ii}^0(t) = 0,$$

and  $e(T) = 0$ .

We can also easily express this solution on the space of non-commutative laws. Recall that we consider  $\lambda$  to be a map from the space of polynomials of  $d$  complex variables to  $\mathbb{C}$ . For  $i = 1, \dots, d$ , define  $q^i$  to be the canonical polynomial  $q^i(X) := X^i$ .

The value function of non-commutative laws is then given by

$$V(t, \lambda) := e(t) + \sum_{i=1}^d \sum_{j=1}^d \left( a_{ij}^0(t) \lambda(q^i q^j) + a_{ij}^1(t) \lambda(q^i) \lambda(q^j) \right).$$

**5.2. Eikonal example.** Next, we consider the free Eikonal equation to illustrate a nonsmooth solution. We set

$$\mathbb{A}_{\mathcal{A}} = \{ \alpha \in L^2(\mathcal{A})_{sa}^d : \|\alpha\|_{L^2(\mathcal{A})} \leq 1 \}, \quad b_{\mathcal{A}}(X, \alpha) = \alpha, \quad L_{\mathcal{A}}(X, \alpha) = 1, \quad \beta_C = \beta_F = 0.$$

We consider the terminal cost of

$$g_{\mathcal{A}}(X) = d_W(\bar{\lambda}, \lambda_X),$$

where  $\bar{\lambda}$  is a given non-commutative law and  $d_W$  denotes the non-commutative  $L^2$ -Wasserstein distance (recall (2.1) for the definition). This example satisfies Assumption **A**. However, the terminal condition is not  $E$ -convex, violating Assumption **B**.

We claim that the following function  $\bar{V}$  is a free viscosity solution:

$$\bar{V}(t, \lambda) := \max\{T - t, d_W(\bar{\lambda}, \lambda)\}. \quad (5.2)$$

In order to deduce that  $\bar{V}$  is a free viscosity subsolution, we may observe that the maximum of two free viscosity subsolutions is a free viscosity subsolution and check the terms individually. For  $\mathcal{A} \in \mathbb{W}$  and  $X \in \mathcal{A}$ , set

$$U_{\mathcal{A}}^1(t, X) := T - t, \quad U_{\mathcal{A}}^2(t, X) := d_W(\bar{\lambda}, \lambda_X), \quad U_{\mathcal{A}} := \max\{U_{\mathcal{A}}^1, U_{\mathcal{A}}^2\}.$$

Firstly, for any  $\mathcal{A} \in \mathbb{W}$ ,  $U_{\mathcal{A}}^1$  satisfies  $U_{\mathcal{A}}^1(T, X) = 0 \leq g_{\mathcal{A}}(X)$  and

$$-\partial_t U_{\mathcal{A}}^1(t, X) + \|\nabla U_{\mathcal{A}}^1(t, X)\|_{L^2(\mathcal{A})} - 1 = 0.$$

Since  $U_{\mathcal{A}}^1$  is a smooth function, we verify that  $(U_{\mathcal{A}}^1)_{\mathcal{A} \in \mathbb{W}}$  is a viscosity solution.

Now we show that  $(U_{\mathcal{A}}^2)_{\mathcal{A} \in \mathbb{W}}$  is a free viscosity subsolution. Since  $X \mapsto d_W(\bar{\lambda}, \lambda_X)$  is 1-Lipschitz, if  $(\phi_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial function which touches  $(U_{\mathcal{A}}^2)_{\mathcal{A} \in \mathbb{W}}$  from above at  $(t_0, \lambda_0)$ , then for all  $\mathcal{A} \in \mathbb{W}$ ,  $t \geq 0$  and  $X, X_0 \in L^2(\mathcal{A})_{sa}^d$  with  $\lambda_0 = \lambda_{X_0}$ ,

$$\phi_{\mathcal{A}}(t_0, X_0) - \phi_{\mathcal{A}}(t, X) \leq d_W(\bar{\lambda}, \lambda_{X_0}) - d_W(\bar{\lambda}, \lambda_X) \leq \|X - X_0\|_{L^2(\mathcal{A})}.$$

This implies that  $\partial_t \phi_{\mathcal{A}}(t_0, X_0) = 0$  and  $\|\nabla \phi_{\mathcal{A}}(t_0, X_0)\|_{L^2(\mathcal{A})} \leq 1$ , verifying that  $(U_{\mathcal{A}}^2)_{\mathcal{A} \in \mathbb{W}}$  is a viscosity subsolution.

To show that  $\bar{V}$  is a viscosity supersolution, let  $\lambda_1$  be an arbitrary non-commutative law in  $\Sigma_d^2$ . Suppose that  $\phi = (\phi_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  is a tracial function which touches  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  from below at  $(t_1, \lambda_1)$  with  $t_1 \in [0, T)$ . Let  $\mathcal{A}$  be a von Neumann algebra such that there exist  $X_1, \bar{X} \in L^2(\mathcal{A})_{sa}^d$  satisfying  $\lambda_{\bar{X}} = \bar{\lambda}$ ,  $\lambda_{X_1} = \lambda_1$  and  $d_W(\bar{\lambda}, \lambda_1) = \|\bar{X} - X_1\|_{L^2(\mathcal{A})}$ .

In the case  $T - t_1 \geq d_W(\bar{\lambda}, \lambda_1)$ , we have  $\phi_{\mathcal{A}}(t_1, X_1) = T - t_1$  and  $\phi_{\mathcal{A}}(t, X_1) \leq T - t$ . Thus  $\partial_t \phi_{\mathcal{A}}(t_1, X_1) = -1$ , implying

$$-\partial_t \phi_{\mathcal{A}}(t_1, X_1) + \|\nabla \phi_{\mathcal{A}}(t_1, X_1)\|_{L^2(\mathcal{A})} - 1 \geq 0. \quad (5.3)$$

In the other case  $T - t_1 < d_W(\bar{\lambda}, \lambda_1)$ , we have  $U_{\mathcal{A}} = U_{\mathcal{A}}^2$  near  $(t_1, X_1)$ . Thus  $\phi_{\mathcal{A}}(t_1, X_1) = d_W(\bar{\lambda}, \lambda_1)$  and  $\phi_{\mathcal{A}}(t, X_1) \leq d_W(\bar{\lambda}, \lambda_1)$  for any  $t$  near  $t_1$ , implying that  $\partial_t \phi_{\mathcal{A}}(t_1, X_1) = 0$ . Also note that  $\lambda_1 \neq \bar{\lambda}$  and so,  $\|\bar{X} - \cdot\|_{L^2(\mathcal{A})}$  is differentiable at  $X_1$ . Since  $\phi_{\mathcal{A}}(t_1, \cdot) - \|\bar{X} - \cdot\|_{L^2(\mathcal{A})}$  has the maximum at  $X_1$ , its derivative exists at  $X_1$  and vanishes there. As the  $L^2(\mathcal{A})$ -norm of a derivative of  $\|\bar{X} - \cdot\|_{L^2(\mathcal{A})}$  is equal to 1 at any differentiable point, we have  $\|\nabla \phi_{\mathcal{A}}(t_1, X_1)\|_{L^2(\mathcal{A})} = 1$ , implying (5.3). Therefore we conclude that  $\bar{V}$  is a viscosity supersolution.

To illustrate this example further, suppose that  $d = 1$  and  $\bar{\lambda}$  has a spectral measure  $\mu \in \mathcal{P}_2(\mathbb{R})$ , where  $\mathcal{P}_2(\mathbb{R})$  denotes the set of probability measures on  $\mathbb{R}$  having a finite second moment. Then, on  $\mathcal{A} = \mathbb{C}$ , for any real number  $x$  (i.e. a self-adjoint element in  $\mathbb{C}$ ), we have

$$\bar{V}_{\mathbb{C}}(t, x) = \max \left\{ T - t, \inf_{f \in L^{\infty}[0,1]} \int_0^1 |x - f(\omega)| d\omega \right\} = \max \left\{ T - t, \int_{\mathbb{R}} |x - y| d\mu(y) \right\}.$$

We get

$$\partial_x \int_{\mathbb{R}} |x - y| d\mu(y) = -\mu((x, \infty)) + \mu((-\infty, x)).$$

Clearly, we have  $-\partial_t \bar{V}_{\mathbb{C}}(t, x) + |\partial_x \bar{V}_{\mathbb{C}}(t, x)| \leq 1$ , but the supersolution property does not hold everywhere so long as  $\mu$  is not supported on a point. This particular example shows that, although  $\bar{V}$  defined in (5.2) is always a *free* viscosity solution, for a *fixed* von Neumann algebra  $\mathcal{A}$ , the function  $\bar{V}_{\mathcal{A}}$  constructed from  $\bar{V}$  via (4.3) may not be a viscosity supersolution on  $L^2(\mathcal{A})_{sa}$  in a usual sense.

**5.3. Controlled von Neumann equation.** We consider the dynamics corresponding to the evolution of  $d$ -quantum system, where the state plays the role of a density matrix (i.e. the one satisfying  $\tau(X) = 1$  along with  $X \geq 0$ ) and the control plays the role of the Hamiltonian. This is given by the drift, with the commutator is applied element-wise,

$$b_{\mathcal{A}}(X, \alpha) = i[X, \alpha].$$

Assume that there is neither common noise nor free individual noise, and consider the quadratic cost

$$L_{\mathcal{A}}(X, \alpha) = \frac{1}{2} \|\alpha\|_{L^2(\mathcal{A})}^2$$

along with a general terminal cost  $g_{\mathcal{A}}(X)$ .

The drift  $b_{\mathcal{A}}$  defined above is a tracial vector-field and satisfies  $E$ -linearity in the control and (3.15) for  $\bar{C} = 0$ , but does not satisfy the uniform continuity or Lipschitz continuity conditions in Assumption **A**. Instead, it would make sense to restrict a state space to  $L^{\infty}(\mathcal{A})_{sa}^d$ , which is preserved under the dynamics.



A seemingly remarkable property is that the optimal control  $(\alpha_t)_t$  turns out to be constant along trajectories. The Hamiltonian is given by

$$\begin{aligned} H_{\mathcal{A}}(X, P) &= \sup_{\alpha \in L^2(\mathcal{A})_{sa}^d} \left\{ \langle i[X, \alpha], P \rangle - \frac{1}{2} \|\alpha\|_{L^2(\mathcal{A})}^2 \right\} \\ &= \sup_{\alpha \in L^2(\mathcal{A})_{sa}^d} \left\{ \langle i[P, X], \alpha \rangle - \frac{1}{2} \|\alpha\|_{L^2(\mathcal{A})}^2 \right\} \\ &= \frac{1}{2} \| [P, X] \|_{L^2(\mathcal{A})}^2. \end{aligned}$$

Thus, the minimization can be done of the terminal cost in the form of a non-commutative Hopf-Lax formula:

$$\tilde{V}_{\mathcal{A}}(t, X) = \inf_{\alpha \in L^2(\mathcal{A})_{sa}^d} \left\{ \frac{(T-t)}{2} \|\alpha\|_{L^2(\mathcal{A})}^2 + g_{\mathcal{A}}(e^{-i\alpha(T-t)} X e^{i\alpha(T-t)}) \right\}. \quad (5.4)$$

The following theorem partly verifies that the optimal control is constant in time if the minimizer exists.

**Theorem 5.1.** *Suppose that  $\mathcal{A} \in \mathbb{W}$ ,  $\mathbb{A}_{\mathcal{A}} = L^2(\mathcal{A})_{sa}^d$  and  $b_{\mathcal{A}}(X, \alpha) = i[X, \alpha]$  with neither common noise nor free individual noise. Let  $(\alpha_t)_{t \in [0, T]} \in \mathbb{A}_{\mathcal{A}}$  be a minimizer of  $\tilde{V}_{\mathcal{A}}$  in (3.11) with the quadratic Lagrangian  $L_{\mathcal{A}}(X, \alpha) = \frac{1}{2} \|\alpha\|_{L^2(\mathcal{A})}^2$ , where  $(X_t)_t \in L^\infty([0, T]; L^\infty(\mathcal{A})_{sa}^d) \cap H^1([0, T]; L^2(\mathcal{A})_{sa}^d)$  is determined by the dynamics (3.10). Then  $(\alpha_t)_{t \in [0, T]}$  is constant in time.*

*Proof.* We consider variations corresponding to a curve  $\xi \in C^1([0, T]; L^\infty(\mathcal{A})_{sa}^d)$  with  $\xi_0 = \xi_T = 0$ . We will show that for any  $s \in \mathbb{R}$ , setting

$$\beta_{t,s} := \int_0^1 e^{-irs\xi_t} \xi'_t e^{irs\xi_t} dr, \quad t \in [0, T],$$

the trajectory

$$t \mapsto \hat{X}_{t,s} := e^{-is\xi_t} X_t e^{is\xi_t}.$$

is a solution to

$$\frac{d}{dt} \hat{X}_{t,s} = -i[s\beta_{t,s} + e^{-is\xi_t} \alpha_t e^{is\xi_t}, \hat{X}_{t,s}]. \quad (5.5)$$

In other words, we have variations of the control (parametrized by  $s \in \mathbb{R}$ ) of the form

$$t \mapsto \hat{\alpha}_{t,s} := s\beta_{t,s} + e^{-is\xi_t} \alpha_t e^{is\xi_t}$$

that result in solutions with  $\hat{X}_{0,s} = X_0$  and  $\hat{X}_{T,s} = X_T$  for any  $s \in \mathbb{R}$ .

To verify that  $(\hat{X}_{t,s})_t$  is a solution to (5.5), first we use Duhamel's formula to compute

$$\begin{aligned} \frac{d}{dt} e^{-is\xi_t} &= \int_0^1 e^{-irs\xi_t} (-is\xi'_t) e^{-i(1-r)s\xi_t} dr \\ &= -is\beta_{t,s} e^{-is\xi_t} \end{aligned}$$

and

$$\begin{aligned} \frac{d}{dt} e^{is\xi_t} &= \int_0^1 e^{irs\xi_t} (is\xi'_t) e^{i(1-r)s\xi_t} dr \\ &= \int_0^1 e^{i(1-u)s\xi_t} (is\xi'_t) e^{iu s\xi_t} du \\ &= is e^{is\xi_t} \beta_{t,s}. \end{aligned}$$

We then have that

$$\frac{d}{dt} \widehat{X}_{t,s} = -i s \beta_{t,s} e^{-is\xi_t} X_t e^{is\xi_t} + e^{-is\xi_t} (-i[\alpha_t, X_t]) e^{is\xi_t} + i s e^{-is\xi_t} X_t e^{is\xi_t} \beta_{t,s},$$

which agrees with (5.5) seeing as

$$e^{-is\xi_t} (-i[\alpha_t, X_t]) e^{is\xi_t} = -i[e^{-is\xi_t} \alpha_t e^{is\xi_t}, e^{-is\xi_t} X_t e^{is\xi_t}].$$

Now the variation of the cost at  $s = 0$  is given by

$$\begin{aligned} & \left. \frac{d}{ds} \right|_{s=0} \int_0^T \frac{1}{2} \|s \beta_{t,s} + e^{-is\xi_t} \alpha_t e^{-is\xi_t}\|_{L^2(\mathcal{A})}^2 dt \\ &= \left. \frac{d}{ds} \right|_{s=0} \int_0^T \left( \frac{s^2}{2} \|\beta_{t,s}\|_{L^2(\mathcal{A})}^2 + s \langle \beta_{t,s}, e^{-is\xi_t} \alpha_t e^{-is\xi_t} \rangle_{L^2(\mathcal{A})} + \frac{1}{2} \|\alpha_t\|_{L^2(\mathcal{A})}^2 \right) dt \\ &= \int_0^T \langle \xi'_t, \alpha_t \rangle_{L^2(\mathcal{A})} dt. \end{aligned}$$

By the optimality of  $(\alpha_t)_{t \in [0, T]}$ , the above quantity is zero for all  $\xi \in C^1([0, T]; L^2(\mathcal{A})_{sa}^d)$  with  $\xi_0 = \xi_T = 0$ , which implies that  $(\alpha_t)_{t \in [0, T]}$  is constant in time.  $\square$

## APPENDIX A. DIFFERENTIAL EQUATIONS

**A.1. Differentiation of vector-valued functions.** Throughout this section, we interchangeably use the notations  $\|\cdot\|_2$  and  $\|\cdot\|_{L^2(\mathcal{A})}$  for a tracial von Neumann algebra  $\mathcal{A}$ .

**Lemma A.1** (Facts about vector-valued  $W^{1,2}$  space). *Let  $\mathcal{A}$  be a tracial von Neumann algebra. For any function  $A : [0, T] \rightarrow L^2(\mathcal{A})$ , define*

$$m(A) := \sup_{0=t_0 < t_1 < \dots < t_n=T} \sum_{j=1}^n \|A_{t_j} - A_{t_{j-1}}\|_2^2.$$

*If  $\alpha : [0, T] \rightarrow L^2(\mathcal{A})$  is Bochner measurable with  $\int_0^T \|\alpha_t\|_2^2 dt < \infty$ , then  $A_t := \int_0^t \alpha_s ds$  is well-defined and continuous (in  $t$ ) which satisfies  $m(A) = \int_0^T \|\alpha_t\|_2^2 dt$ . Conversely, if  $t \mapsto A_t$  is continuous with  $m(A) < \infty$  and  $A_0 = 0$ , then there exist a measurable  $\alpha : [0, T] \rightarrow L^2(\mathcal{A})$  such that  $A_t = \int_0^t \alpha_s ds$  for  $t \in [0, T]$  and  $\int_0^T \|\alpha_t\|_2^2 dt = m(A)$ .*

*Proof.* First, let  $\alpha : [0, T] \rightarrow L^2(\mathcal{A})$  be given with  $\int_0^T \|\alpha_t\|_2^2 dt < \infty$ . Let  $0 \leq a < b \leq T$  and  $f \in L^2(\mathcal{A})$ . Then

$$\begin{aligned} \int_a^b |\langle \alpha_t, f \rangle_2| dt &\leq \int_a^b \|\alpha_t\|_2 \|f\|_2 dt \leq \left( \int_a^b \|\alpha_t\|_2^2 dt \right)^{1/2} \left( \int_a^b \|f\|_2^2 dt \right)^{1/2} \\ &= (b-a)^{1/2} \left( \int_a^b \|\alpha_t\|_2^2 dt \right)^{1/2} \|f\|_2. \end{aligned}$$

Thus,  $f \mapsto \int_a^b \langle \alpha_t, f \rangle_2 dt$  defines a bounded linear functional on  $L^2(\mathcal{A})$  and hence there is a unique vector, which we denote by  $\int_a^b \alpha_t dt$ , such that

$$\left\langle \int_a^b \alpha_t dt, f \right\rangle_2 = \int_a^b \langle \alpha_t, f \rangle_2 dt \quad \text{for any } f \in L^2(\mathcal{A}).$$

Furthermore,

$$\left\| \int_a^b \alpha_t dt \right\|_2^2 = \sup_{\|f\|_2 \leq 1} \left| \int_a^b \langle \alpha_t, f \rangle_2 dt \right|^2 \leq (b-a) \int_a^b \|\alpha_t\|_2^2 dt.$$

Let  $A_t := \int_0^t \alpha_s ds$ . Fix a partition  $0 = t_0 < t_1 < \dots < t_n = T$ , and note that

$$\sum_{j=1}^n \frac{1}{t_j - t_{j-1}} \|A_{t_j} - A_{t_{j-1}}\|_2^2 \leq \sum_{j=1}^n \frac{1}{t_j - t_{j-1}} \left\| \int_{t_{j-1}}^{t_j} \alpha_t dt \right\|_2^2 \leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|\alpha_t\|_2^2 dt.$$

and so  $m(A) \leq \int_0^T \|\alpha_t\|_2^2 dt$ . To show the opposite inequality, since  $\alpha_t$  is in the Bochner  $L^2$  space, it can be approximated by linear combinations of indicator functions times fixed vectors, and we can also arrange approximations by indicator functions of intervals. Hence, given  $\epsilon > 0$ , there exists some partition  $0 = t_0 < t_1 < \dots < t_n = T$  and vectors  $f_1, \dots, f_n \in L^2(\mathcal{A})$  such that

$$\left( \int_0^T \|\alpha_t - \beta_t\|_2^2 dt \right)^{1/2} < \epsilon, \text{ where } \beta_t := \sum_{j=1}^n \mathbf{1}_{(t_{j-1}, t_j]}(t) f_j.$$

Let  $B_t := \int_0^t \beta_s ds$ . Note that

$$m(B) \geq \sum_{j=1}^n \frac{1}{t_j - t_{j-1}} \left\| \int_{t_{j-1}}^{t_j} \beta_t dt \right\|_2^2 = \sum_{j=1}^n (t_j - t_{j-1}) \|f_j\|^2 = \int_0^T \|\beta_t\|_2^2 dt.$$

It is not hard to check that  $m(A)^{1/2}$  is a seminorm (allowing the value  $+\infty$ ) on functions  $[0, T] \rightarrow L^2(\mathcal{A})$ . Hence,

$$\begin{aligned} m(A)^{1/2} &\geq m(B)^{1/2} - m(A - B)^{1/2} \geq \|\beta\|_{L^2([0, T], L^2(\mathcal{A}))} - \|\alpha - \beta\|_{L^2([0, T], L^2(\mathcal{A}))} \\ &\geq \|\alpha\|_{L^2([0, T], L^2(\mathcal{A}))} - 2\|\alpha - \beta\|_{L^2([0, T], L^2(\mathcal{A}))} \\ &\geq \|\alpha\|_{L^2([0, T], L^2(\mathcal{A}))} - 2\epsilon. \end{aligned}$$

Since  $\epsilon > 0$  was arbitrary, we have  $m(A) = \int_0^T \|\alpha_t\|_2^2 dt$  as desired.

Next, suppose we are given  $A$  with  $m(A) < \infty$ . We will construct  $\alpha$  again using the self-duality of Hilbert spaces. Let  $H_0 \subseteq L^2([0, T], L^2(\mathcal{A}))$  (Bochner  $L^2$  space) be the linear span of functions of the form  $\mathbf{1}_{(a, b]}(t)f$  for  $f \in L^2(\mathcal{A})$ . We define a linear map  $\ell : H_0 \rightarrow \mathbb{C}$  by

$$\ell : \sum_{j=1}^n \mathbf{1}_{(s_j, t_j]} f_j \mapsto \sum_{j=1}^n \langle A_{t_j} - A_{s_j}, f_j \rangle_2;$$

It is straightforward to check that this is well-defined, i.e. independent of the decomposition of the given function on  $H_0$ . To check that the map  $\ell$  is bounded, we consider a given function  $f$  in  $H_0$ ; let  $\{t_0, \dots, t_n\}$  be a partition that contains all the endpoints of intervals in the given decomposition of  $f$ , and express  $f$  as  $\sum_{j=1}^n \mathbf{1}_{(t_{j-1}, t_j]} f_j$ . Then

$$\begin{aligned} |\ell(f)| &\leq \sum_{j=1}^n |\langle A_{t_j} - A_{t_{j-1}}, f_j \rangle_2| \\ &\leq \sum_{j=1}^n \frac{1}{(t_j - t_{j-1})^{1/2}} \|A_{t_j} - A_{t_{j-1}}\|_2 (t_j - t_{j-1})^{1/2} \|f_j\|_2 \\ &\leq \left( \sum_{j=1}^n \frac{1}{t_j - t_{j-1}} \|A_{t_j} - A_{t_{j-1}}\|_2^2 \right)^{1/2} \left( \sum_{j=1}^n (t_j - t_{j-1}) \|f_j\|_2^2 \right)^{1/2} \\ &\leq m(A)^{1/2} \|f\|_{L^2([0, T], L^2(\mathcal{A}))}. \end{aligned}$$

Since  $\ell$  is bounded, it extends from  $H_0$  to all of  $L^2([0, T], L^2(\mathcal{A}))$ , and there exists  $\alpha \in L^2([0, T], L^2(\mathcal{A}))$  with  $\ell(f) = \langle f, \alpha \rangle_{L^2([0, T], L^2(\mathcal{A}))}$ , and  $\|\alpha\|_{L^2([0, T], L^2(\mathcal{A}))} \leq m(A)^{1/2}$ . Then to check that  $A_t = \int_0^t \alpha_s ds$ , note that for  $f \in L^2(\mathcal{A})$ , we have

$$\left\langle \int_0^t \alpha_s ds, f \right\rangle_{L^2(\mathcal{A})} = \langle \alpha, \mathbf{1}_{[0, t]} f \rangle_{L^2([0, T], L^2(\mathcal{A}))} = \ell(f) = \langle A_t - A_0, f \rangle_{L^2(\mathcal{A})}.$$

□

**Lemma A.2.** *The following statements hold.*

1. *Let  $\alpha : [0, T] \rightarrow L^2(\mathcal{A})$  be Bochner measurable and define  $A_t := \int_0^t \alpha_s ds$  for  $t \in [0, T]$ . Then  $t \mapsto \|\alpha_t\|_{L^\infty(\mathcal{A})}$  is measurable and*

$$\|\alpha\|_{L^\infty([0, T], L^\infty(\mathcal{A}))} = \sup_{0 \leq s < t \leq T} \frac{\|A_t - A_s\|_{L^\infty(\mathcal{A})}}{t - s}.$$

(Here both sides may be infinite.)

2. *For any  $A : [0, T] \rightarrow L^2(\mathcal{A})$ ,*

$$m(A)^{1/2} \leq T^{1/2} \sup_{0 \leq s < t \leq T} \frac{\|A_t - A_s\|_{L^\infty(\mathcal{A})}}{t - s},$$

*so in particular if  $A$  is Lipschitz in  $\|\cdot\|_{L^\infty(\mathcal{A})}$ , then there exists a compatible  $\alpha : [0, T] \rightarrow L^2(\mathcal{A})$  (i.e.  $A_t := \int_0^t \alpha_s ds$ ).*

*Proof.* Note that

$$\|x\|_{L^\infty(\mathcal{A})} = \sup_{a, b \in \mathcal{A}, \|a\|_2, \|b\|_2 \leq 1} |\tau_{\mathcal{A}}(axb)|.$$

This shows that  $x \mapsto \|x\|_{L^\infty(\mathcal{A})}$  is a lower semi-continuous function on  $L^2(\mathcal{A})$ , both in the norm topology and the weak-\* topology. In particular, it is a Borel-measurable function, and hence  $t \mapsto \|\alpha_t\|_{L^\infty(\mathcal{A})}$  is measurable.

Now note that for  $s < t$ , we have

$$\|A_t - A_s\|_{L^\infty(\mathcal{A})} = \left\| \int_s^t \alpha_u du \right\|_{L^\infty(\mathcal{A})} \leq \int_s^t \|\alpha_u\|_{L^\infty(\mathcal{A})} du \leq (t - s) \|\alpha\|_{L^\infty([0, T], L^\infty(\mathcal{A}))}.$$

For the opposite direction, first note that the image of  $\alpha$  is contained in  $L^2(\mathcal{A}_0)$  for some  $L^2$ -separable von Neumann subalgebra of  $\mathcal{A}$ . Furthermore, choose some countable set  $C \subseteq \mathcal{A}_0$  that is dense in the unit ball of  $L^2(\mathcal{A})$ , so that

$$\|\alpha_t\|_{L^\infty(\mathcal{A})} = \sup_{a, b \in C} |\tau_{\mathcal{A}}(a\alpha_t b)|$$

Note that for  $a, b \in C$ ,

$$\int_s^t \tau_{\mathcal{A}}(a\alpha_u b) du = \int_s^t \langle \alpha_u, (ba)^* \rangle du = \tau_{\mathcal{A}}(a(A_t - A_s)b).$$

Write  $L$  for the Lipschitz norm of  $t \mapsto A_t$  with respect to  $\|\cdot\|_{L^\infty(\mathcal{A})}$ . By the Lebesgue differentiation theorem, we have for a.e.  $t$  that

$$|\tau_{\mathcal{A}}(a\alpha_t b)| = \lim_{\epsilon \rightarrow 0} |\tau_{\mathcal{A}}(a(A_{t+\epsilon} - A_t)b)| \leq L \|a\|_{L^2(\mathcal{A})} \|b\|_{L^2(\mathcal{A})} \leq L.$$

Since  $C$  is countable, this holds for all  $a, b \in C$  simultaneously, for a.e.  $t \in [0, T]$ . Hence, the  $L^\infty[0, T]$ -norm of  $\|\alpha_t\|_{L^\infty(\mathcal{A})}$  is bounded by  $L$ . □

**A.2. Stochastic integrals and differential equations.** First, we should define what the stochastic integral means. For our purposes, it should be sufficient to assume  $\mathbf{A}$ , and we really don't have to worry about any free stochastic integrals.

To complete the proof of Theorem 3.6, we should show the existence of an adapted solution using a Picard iteration scheme. A simple version of the proof uses the weighted norm on  $C([t_0, t_1]; L^2(\Omega, \mathcal{F}, \mathbb{P}; L^2(\mathcal{A})_{sa}^d))$  given by, for some constant  $\gamma > 0$  chosen later,

$$\|X\|_{C_\gamma}^2 := \sup_{t \in [t_0, t_1]} e^{-2\gamma(t-t_0)} \mathbb{E}[\|X_t\|_{L^2(\mathcal{A})}^2].$$

For  $X \in C([t_0, t_1]; L^2(\Omega, \mathcal{F}, \mathbb{P}; L^2(\mathcal{A})_{sa}^d))$ , define  $\Phi(X) = (\Phi_t(X))_{t \in [t_0, t_1]}$  as

$$\Phi_t(X) := \int_{t_0}^t b_{\mathcal{A}}(X_s, \alpha_s) ds + \beta_C \mathbb{1}_{\mathcal{A}} W_t^0 + \beta_F S_t.$$

Then for  $X, Y \in C([t_0, t_1]; L^2(\Omega, \mathcal{F}, \mathbb{P}; L^2(\mathcal{A})_{sa}^d))$ ,

$$\Phi_t(X) - \Phi_t(Y) = \int_{t_0}^t (b_{\mathcal{A}}(X_s, \alpha_s) - b_{\mathcal{A}}(Y_s, \alpha_s)) ds.$$

From this we obtain

$$\begin{aligned} & \frac{d}{dt} e^{-2\gamma(t-t_0)} \mathbb{E}[\|\Phi_t(X) - \Phi_t(Y)\|_{L^2(\mathcal{A})}^2] \\ &= e^{-2\gamma(t-t_0)} \mathbb{E}\left[2\langle b_{\mathcal{A}}(X_t, \alpha_t) - b_{\mathcal{A}}(Y_t, \alpha_t), \Phi_t(X) - \Phi_t(Y) \rangle_{L^2(\mathcal{A})} - 2\gamma \|\Phi_t(X) - \Phi_t(Y)\|_{L^2(\mathcal{A})}^2\right] \\ &\leq e^{-2\gamma(t-t_0)} \mathbb{E}\left[\frac{\bar{C}^2}{\gamma} \|X_t - Y_t\|_{L^2(\mathcal{A})}^2 - \gamma \|\Phi_t(X) - \Phi_t(Y)\|_{L^2(\mathcal{A})}^2\right]. \end{aligned}$$

We end up with

$$\|\Phi(X) - \Phi(Y)\|_{C_\gamma}^2 \leq \frac{\bar{C}^2}{\gamma^2} \|X - Y\|_{C_\gamma}^2,$$

making  $\Phi$  a contraction map w.r.t.  $C_\gamma$ -norm when  $\gamma > \bar{C}$ . At each time  $t$ , the fixed-point iteration converges in  $L^2(\mathcal{A}_t)$ , making the solution freely adapted.

We may also directly calculate that for almost every  $t$ ,

$$\begin{aligned} \frac{d}{dt} \mathbb{E}[\|X_t - x_0\|_{L^2(\mathcal{A})}^2] &= \lim_{h \rightarrow +0} h^{-1} \mathbb{E}[\|X_{t+h} - x_0\|_{L^2(\mathcal{A})}^2 - \|X_t - x_0\|_{L^2(\mathcal{A})}^2] \\ &= \lim_{h \rightarrow +0} h^{-1} \mathbb{E}[\|X_{t+h} - X_t\|_{L^2(\mathcal{A})}^2 + 2\langle X_{t+h} - X_t, X_t - x_0 \rangle_{L^2(\mathcal{A})}] \\ &= \lim_{h \rightarrow +0} \mathbb{E}[h^{-1} \beta_C^2 d(W_{t+h}^0 - W_t^0)^2 + \beta_F^2 \|S_{t+h} - S_t\|_{L^2(\mathcal{A})}^2 \\ &\quad + 2\langle h^{-1} \int_t^{t+h} b_{\mathcal{A}}(X_s, \alpha_s) ds, X_t - x_0 \rangle_{L^2(\mathcal{A})}] \\ &= d(\beta_C^2 + \beta_F^2) + \mathbb{E}[\langle b_{\mathcal{A}}(X_t, \alpha_t), X_t - x_0 \rangle_{L^2(\mathcal{A})}]. \end{aligned}$$

## APPENDIX B. COMPUTATION OF FREE LAPLACIAN

Here we describe the free Laplacian for special test functions (such as polynomials) in terms of Voiculescu's non-commutative derivatives. These computations are well-established in free probability. See e.g. [32, §3], [62, §14.1], [59, §4.3].

We recall that  $\text{NCP}_d$  denotes the  $*$ -algebra of non-commutative polynomials in  $d$  self-adjoint variables. Because Voiculescu's free difference quotient maps  $\text{NCP}_d$  into the tensor product space  $\text{NCP}_d \otimes \text{NCP}_d$ , we recall the following definitions.

**B.1. Tensor products.** For any  $*$ -algebras  $A$  and  $B$ , the algebraic tensor product  $A \otimes B$  is also a  $*$ -algebra with the multiplication and  $*$ -operation satisfying

$$(X \otimes Y)(W \otimes Z) = (XW) \otimes (ZY)$$

and

$$(X \otimes Y)^* = X^* \otimes Y^*.$$

Moreover, in the case of  $A \otimes A$ , for  $X, Y, Z \in A$ , we write

$$(X \otimes Y) \# Z = XZY,$$

and extend this function of  $X \otimes Y$  linearly to the tensor product  $A \otimes A$ .

Given a tracial von Neumann algebra  $\mathcal{A} = (A, \tau)$ , the algebraic tensor product  $A \otimes A$  can be equipped with the trace  $\tau \otimes \tau$  and then completed to a tracial von Neumann algebra. The inner product is then given by

$$\langle X \otimes Y, Z \otimes W \rangle = \tau(X^*Z)\tau(Y^*W).$$

The von Neumann algebraic tensor product with completion is also denoted by, for example,  $A \otimes A$ .

**B.2. The free difference quotient.** Voiculescu's  $j$ th free difference quotient is the map  $\partial_{x_j} : \text{NCP}_d \rightarrow \text{NCP}_d \otimes \text{NCP}_d$  given, for a monomial, by

$$\partial_{x_j} x_{i_1} x_{i_2} \dots x_{i_m} := \sum_{i_k=j} (x_{i_1} \dots x_{i_{k-1}}) \otimes (x_{i_{k+1}} \dots x_{i_m}).$$

In the cases where  $i_k = j$  for  $k = 1$ , we take the left part of the tensor to be the identity, and when  $k = m$ , we take the right part to be the identity. This operation is extended linearly to  $\text{NCP}_d$ . Note that if  $\mathcal{A}$  is a tracial  $W^*$ -algebra and  $x \in \mathcal{A}_{\text{sa}}^d$ , then  $\partial_{x_j} p$  can be evaluated at  $x$  to produce an element in  $A \otimes A$ .

As motivation for this definition, we recall the following fact; see e.g. [62, Lemma 14.1.3], [59, Lemma 3.17].

**Lemma B.1.** *Let  $p \in \text{NCP}_d$ . Fix a tracial von Neumann algebra  $\mathcal{A} = (A, \tau)$ . Then  $p$  defines a map  $A_{\text{sa}}^d \rightarrow A$  which is Fréchet differentiable and satisfies*

$$\left. \frac{d}{dt} \right|_{t=0} p(x + ty) = \sum_{j=1}^d \partial_{x_j} p(x) \# y_j.$$

The non free difference quotients also satisfy a chain rule; recall the multiplication in the tensor product space defined earlier.

**Lemma B.2** (Chain rule). *Let  $f \in \text{NCP}_d$  and let  $g = (g_1, \dots, g_d) \in \text{NCP}_{d'}^d$ . Then for  $i \in \{1, \dots, d'\}$ ,*

$$\partial_{x_i}(f \circ g)(x) = \sum_{j=1}^d \partial_{x_j} f(g(x)) \partial_{x_i} g_j(x).$$

*Remark B.3.* In the case of a single variable, the non-commutative derivative  $\partial_x$  relates closely to difference quotients. Indeed,  $\partial_x$  maps  $\mathbb{C}[x]$  into  $\mathbb{C}[x] \otimes \mathbb{C}[x]$ , and there is a canonical algebra isomorphism  $\Phi : \mathbb{C}[x] \otimes \mathbb{C}[x] \rightarrow \mathbb{C}[x, y]$  such that  $\Phi(x \otimes 1) = x$  and  $\Phi(1 \otimes x) = y$ . Note

$$\partial_x[x^k] = x^{k-1} \otimes 1 + x^{k-2} \otimes x + \dots + 1 \otimes x^{k-1},$$

and thus

$$\Phi[\partial_x[x^k]] = x^{k-1} + x^{k-2}y + \dots + y^{k-1} = \frac{x^k - y^k}{x - y} \in \mathbb{C}[x, y].$$

Hence, by linearity, for all  $p \in \mathbb{C}[x]$ ,

$$\Phi[\partial_x p(x)] = \frac{p(x) - p(y)}{x - y} \in \mathbb{C}[x, y].$$

In a classical probabilistic setting, the space  $\mathbb{C}[x, y]$  will often be equipped with a product measure representing identical copies of independent random variables. Suppose that  $\tau$  is the trace of von Neumann algebra and  $X \in L^2(\mathcal{A})_{sa}$  has spectral measure  $\mu \in \mathcal{P}_2(\mathbb{R})$ , i.e., for  $p \in \mathbb{C}[x]$ ,

$$\tau(p(X)) = \int_{\mathbb{R}} p(x) \mu(dx).$$

This identification extends to the tensor product using the isomorphism  $\Phi$ . For  $A \in C[x] \otimes C[x]$  we have

$$(\tau \otimes \tau)(A(X)) = \int_{\mathbb{R}} \int_{\mathbb{R}} \Phi[A](x, y) \mu(dx) \mu(dy).$$

An example that has driven a lot of research in free probability is the case of resolvents (see [81]). Although resolvents are not polynomials, they can be expressed through power series, and so the definition of the non-commutative can be extended to them without much difficulty.

For resolvents, we may compute, for example,

$$(1 - \lambda x)^{-1} - (1 - \lambda y)^{-1} = \lambda(1 - \lambda x)^{-1}(x - y)(1 - \lambda y)^{-1}$$

and it follows that

$$\partial_x(1 - \lambda x)^{-1} = \lambda(1 - \lambda x)^{-1} \otimes (1 - \lambda x)^{-1}.$$

For another example we make use of later, consider  $\psi(x) = \arctan(x)$ . We have for a single-variable that the derivative can be computed using the standard derivative of  $\arctan$ ,

$$\begin{aligned} \frac{d}{dt} \psi(tX) &= X(1 + t^2 X^2)^{-1} \\ &= \frac{1}{2} X[(1 + itX)^{-1} + (1 - itX)^{-1}] \\ &= \frac{i}{2t} [(1 + itX)^{-1} - (1 - itX)^{-1}]. \end{aligned}$$

We find that we can write the  $\arctan$  as an integral of resolvents

$$\arctan(X) = \int_0^1 \frac{i}{2t} [(1 + itX)^{-1} - (1 - itX)^{-1}] dt.$$

The non-commutative derivative is then given by

$$\partial_x \arctan(X) = \int_0^1 \frac{1}{2} [(1 + itX)^{-1} \otimes (1 + itX)^{-1} + (1 - itX)^{-1} \otimes (1 - itX)^{-1}] dt.$$

**B.3. The cyclic derivative.** The *cyclic derivative*  $\mathcal{D}_{x_j}^\circ : \text{NCP}_d \rightarrow \text{NCP}_d$  is defined for non-commutative monomials by

$$\mathcal{D}_{x_j}^\circ x_{i_1} x_{i_2} \dots x_{i_m} := \sum_{i_k=j} (x_{i_{k+1}} \dots x_{i_m} x_{i_1} \dots x_{i_{k-1}}),$$

and extended linearly to non-commutative polynomials. Equivalently,  $\mathcal{D}_{x_j}^\circ p$  is obtained from  $\partial_{x_j} p$  by applying the map  $X \otimes Y \mapsto YX$ . From this we see that  $\tau(y \mathcal{D}_{x_j}^\circ p(x)) = \tau(\partial_{x_j} p(x) \# y)$  whenever  $x$  is a self-adjoint tuple in a tracial von Neumann algebra and  $y$  is an element of it.

We remark that if  $p$  is self-adjoint, then  $\partial_{x_j} p$  is also self-adjoint. We denote by  $\mathcal{D}^\circ p$  the row vector with entries  $\mathcal{D}_{x_j}^\circ p$ .

The cyclic derivative arises naturally from computing the gradient of the trace of a non-commutative polynomial as follows.

**Lemma B.4.** *Let  $p \in \text{NCP}_d$ . Fix a tracial von Neumann algebra  $\mathcal{A} = (A, \tau)$ . Then  $\tau \circ p$  defines a map  $A_{\text{sa}}^d \rightarrow \mathbb{C}$  which is Fréchet differentiable and satisfies*

$$\nabla[\tau(p(x))] = \mathcal{D}^\circ p(x).$$

*Proof.* Differentiability follows immediately from Lemma B.1 since  $\tau$  is linear. Moreover, for  $y \in A_{\text{sa}}^d$ , we have

$$\langle y, \nabla p(x) \rangle_{L^2(\mathcal{A})} = \frac{d}{dt} \Big|_{t=0} \tau[p(x+ty)] = \sum_{j=1}^d \tau(\partial_{x_j} p(x) \# y_j) = \sum_{j=1}^d \tau(y_j \mathcal{D}_{x_j}^\circ p(x)) = \langle y, \mathcal{D}^\circ p(x) \rangle_{L^2(\mathcal{A})}.$$

Since  $y$  is arbitrary,  $\nabla[\tau(p(x))] = \mathcal{D}^\circ p(x)$ .  $\square$

**B.4. Cylindrical test functions.** A useful class of test functions where one can compute the free Laplacian are cylindrical functions of the form

$$U_{\mathcal{A}}(X) = g\left(\tau((\phi_1 \circ \psi)(X)), \dots, \tau((\phi_m \circ \psi)(X))\right)$$

for  $g \in C^{2,1}(\mathbb{R}^m; \mathbb{R})$ , self-adjoint  $\phi_1, \dots, \phi_m \in \text{NCP}_d$ , and  $\psi(x) = \arctan(x)$  applied component-wise. Clearly,  $(U_{\mathcal{A}})_{\mathcal{A} \in \mathbb{W}}$  are tracial  $W^*$ -functions. In fact, they fall into the category of non-commutative smooth functions studied in [59], as one can see from §3.4 and §4.2 of that paper. The free Laplacian computed here for cylindrical functions will thus be a special case of the free Laplacian in [59, §4.3], but here we want to give a shorter and more self-contained development for this concrete family.

We use the notation  $g_o$  to denote the partial derivative with respect to the  $o$ th component of  $g$  and  $g_{oq}$  to denote the second partial derivative with respect to the  $o$  and  $q$ th components. We abbreviate by

$$g_o = g_o\left(\tau((\phi_1 \circ \psi)(X)), \dots, \tau((\phi_m \circ \psi)(X))\right)$$

and

$$g_{oq} = g_{oq}\left(\tau((\phi_1 \circ \psi)(X)), \dots, \tau((\phi_m \circ \psi)(X))\right).$$

Then by the chain rule we have

$$\partial_{x_j}(\phi_m \circ \psi)(X) = (\partial_{x_j} \phi_m \circ \psi)(X) \partial_x \psi(X_j).$$

Similarly, we may define the cyclic derivative  $\mathcal{D}_{x_j}^\circ(\phi_m \circ \psi)(X)$ .

By Lemma B.4 and the chain rule, we have that, for  $j \in \{1, \dots, d\}$ ,

$$(\nabla U_{\mathcal{A}})^j(X) = \sum_{o=1}^m g_o \mathcal{D}_{x_j}^\circ(\phi_o \circ \psi)(X).$$

We define the non-commutative derivative, for  $i, j \in \{1, \dots, d\}$ , naturally to be

$$\partial_{x_i}(\nabla U_{\mathcal{A}})^j(X) = \sum_{o=1}^m g_o \partial_{x_i} \mathcal{D}_{x_j}^\circ(\phi_o \circ \psi)(X) \in L^2(\mathcal{A} \otimes \mathcal{A}, \tau \otimes \tau).$$

**Proposition B.5.** *For a cylindrical function  $U$  as given above, we have  $U \in \mathcal{X}_\Sigma$  and*

$$\Theta_{\mathcal{A}} U(X) = \sum_{i=1}^d (\tau \otimes \tau) \left( \partial_{x_i}(\nabla U_{\mathcal{A}})^i(X) \right).$$



*Proof.* There are two tensors arising in the calculation of the Hessian of  $U_{\mathcal{A}}$  that act in different ways. The first is the non-commutative derivative,  $\partial_{x_i}(\nabla U_{\mathcal{A}})^j(X)$  defined above. We also collect terms involving the second derivatives of  $g$  by

$$(\nabla^2 U_{\mathcal{A}})^{ij}(X) = \sum_{o=1}^m \sum_{q=1}^m g_{oq} \mathcal{D}_{x_i}^o(\phi_o \circ \psi)(X) \otimes \mathcal{D}_{x_j}^o(\phi_q \circ \psi)(X).$$

The Hessian of  $U$  can be expressed as

$$\begin{aligned} & \text{Hess } U_{\mathcal{A}}(X)[A, B] \\ &= \sum_{i=1}^d \sum_{j=1}^d \left( \langle (\nabla^2 U_{\mathcal{A}})^{ij}(X), A^i \otimes B^j \rangle_{L^2(\mathcal{A} \otimes \mathcal{A})} + \langle \partial_{x_i}(\nabla U_{\mathcal{A}})^j(X) \# A^i, B^j \rangle_{L^2(\mathcal{A})} \right). \end{aligned} \quad (\text{B.1})$$

We check condition (d) of the definition  $\mathcal{X}_{\Sigma}$  as the other conditions are simpler. We fix  $\mathcal{A} \in \mathbb{W}$ ,  $X, Y, A \in L^2(\mathcal{A})_{sa}^d$  and  $B \in L^\infty(\mathcal{A})_{sa}^d$ . Furthermore, we consider the case that  $U_{\mathcal{A}}(X) = \tau((\phi \circ \psi)(x))$  for a monomial  $\phi(x) = \prod_{k=1}^m x_{i_k} \in \text{NCP}_d$ , which showcases the key calculation. We have

$$\begin{aligned} & \langle (\nabla U_{\mathcal{A}})^j(X), A \rangle_{L^2(\mathcal{A})} \\ &= \frac{1}{2} \int_0^1 \sum_{i_k=j} \sum_{\rho=\pm 1} \tau(\psi(X_{i_1}) \dots \psi(X_{i_{k-1}})(1 + \rho i t X_j)^{-1} A_j (1 + \rho i t X_j)^{-1} \psi(X_{i_{k+1}}) \dots \psi(X_{i_m})) dt. \end{aligned}$$

We have when computing the Hessian finitely many terms that include either, for  $j, j' \in \{1, \dots, d\}$  and  $\rho, \rho' \in \{\pm 1\}$  and  $i_k = j$  and  $i_{k'} = 1$ ,

$$\begin{aligned} & \tau(\psi(X_{i_1}) \dots \psi(X_{i_{k-1}})(1 + \rho i t X_j)^{-1} A_j (1 + \rho i t X_j)^{-1} \psi(X_{i_{k+1}}) \\ & \dots \psi(X_{i_{k'-1}})(1 + \rho' i t X_{j'})^{-1} B_{j'} (1 + \rho' i t X_{j'})^{-1} \psi(X_{i_{k'+1}}) \dots \psi(X_{i_m})), \end{aligned}$$

or, when  $j = j'$  and  $k = k'$ ,

$$- \rho i t \tau(\psi(X_{i_1}) \dots \psi(X_{i_{k-1}})(1 + \rho i t X_j)^{-1} A_j (1 + \rho i t X_j)^{-1} B_j (1 + \rho i t X_j)^{-1} \psi(X_{i_{k+1}}) \dots \psi(X_{i_m})).$$

When we consider the terms appearing in  $|\text{Hess } U_{\mathcal{A}}(X)[A, B] - \text{Hess } U_{\mathcal{A}}(Y)[A, B]|$  we may telescope and rearrange the terms into either the form, where  $M_1, M_2, M_3$  have bounded  $L^\infty$  norm (consider  $\|\arctan(X_j)\|_\infty \leq \frac{\pi}{2}$ ),

$$\tau([\psi(X_i) - \psi(Y_i)] M_1 A_j M_2 B_{j'} M_3) \leq \left(\frac{\pi}{2}\right)^{m-2} \|X_i - Y_i\|_{L^2(\mathcal{A})} \|A_j\|_{L^2(\mathcal{A})} \|B_{j'}\|_{L^\infty(\mathcal{A})}$$

or

$$\tau([(1 + \rho i t X_i)^{-1} - (1 + \rho i t Y_i)^{-1}] A_j M_2 B_{j'} M_3) \leq \left(\frac{\pi}{2}\right)^{m-1} \|X_i - Y_i\|_{L^2(\mathcal{A})} \|A_j\|_{L^2(\mathcal{A})} \|B_{j'}\|_{L^\infty(\mathcal{A})}.$$

It follows that condition (d) in the definition of  $\mathcal{X}_{\Sigma}$  is satisfied with the constant  $K$  depending only on the degree of the monomial because of the number of terms that arise in this calculation. For more general cylindrical functions the result holds with  $K$  also depending on Lipschitz estimates for  $g$ .

We now compute the free Laplacian for a general cylindrical function. Assume that  $\iota : \mathcal{A} = (A, \tau) \rightarrow \mathcal{B} = (B, \rho)$  is a tracial  $W^*$ -embedding and  $\mathcal{B}$  contains a semi-circular  $S = (S^1, \dots, S^d) \in \mathcal{B}$  freely independent of  $\iota(A)$ . For the terms with the second-order derivatives of  $g$ , we see that, for example, for  $l = 1, \dots, d$ ,

$$\langle (\nabla^2 U_{\mathcal{B}})^{ll}(X), S^l \otimes S^l \rangle_{L^2(\mathcal{B} \otimes \mathcal{B})}$$

$$= \sum_{o=1}^m \sum_{q=1}^m g_{oq} \langle \mathcal{D}_{x_l}^o(\phi_o \circ \psi)(X), S^l \rangle_{L^2(\mathcal{B})} \langle \mathcal{D}_{x_l}^o(\phi_q \circ \psi)(X), S^l \rangle_{L^2(\mathcal{B})} = 0,$$

because  $S^l$  is freely independent of  $\mathcal{D}_{x_l}^o(\phi_o \circ \psi)(X)$  and  $S^l$  has a zero trace.

For the term with the non-commutative second derivative, for each  $i$  and  $X$ , we can decompose the non-commutative derivative into finitely many simple tensors (see the calculations above) like

$$\partial_{x_i}(\nabla U_{\mathcal{A}})^i(X) = \sum_{j=1}^M a^j \otimes b^j.$$

We then have that, using the free Wick law (which follows here directly from free independence) to evaluate,

$$\begin{aligned} \text{Hess } U_{\mathcal{B}}(X)[\mathbf{e}^i S, \mathbf{e}^i S] &= \langle \partial_{x_i}(\nabla U_{\mathcal{B}})^i(X) \# S^i, S^i \rangle_{L^2(\mathcal{B})} \\ &= \sum_{j=1}^M \langle a^j S^i b^j, S^i \rangle_{L^2(\mathcal{B})} \\ &= \sum_{j=1}^M \rho(S^i a^j S^i, b^j) \\ &= \sum_{j=1}^M \tau(a^j) \tau(b^j) \\ &= \sum_{j=1}^M (\tau \otimes \tau)(a^j \otimes b^j) \\ &= (\tau \otimes \tau)(\partial_{x_i}(\nabla U_{\mathcal{A}})^i(X)). \end{aligned}$$

□

We can also consider the special case of  $d = 1$  and relate it to the Wasserstein derivative of  $U(\mu)$  where  $\mu$  denotes the probability measure on  $\mathbb{R}$  and the Wasserstein derivative is denoted by  $\partial U(\mu)(x)$ . We will only describe this here as a formal computation. We see that if  $\mathcal{A}$  is a tracial  $W^*$ -algebra and  $X \in L^2(\mathcal{A})_{sa}$  with law  $\mu$  then, since  $\partial U(\mu)$  is a Borel function, the spectral theory of bounded normal operators allows us to define the operator  $\partial U(\mu)(X)$ . When  $U$  is differentiable at  $\mu$  and  $\mathcal{U}_{\mathcal{A}}$  is differentiable at  $X$ , i.e., the cylindrical test functions above, one checks that the identity

$$\partial U(\mu)(X) = \nabla U_{\mathcal{A}}(X)$$

holds (note the resemblance with the identity in [45, Corollary 3.22], which instead involves  $\partial U(\mu) \circ X$ ).

Using the canonical isomorphism  $\Phi : \mathbb{C}[x] \otimes \mathbb{C}[x] \rightarrow \mathbb{C}[x, y]$  above we have

$$\Phi[\partial_x \partial U(\mu)(x)] = \frac{\partial U(\mu)(x) - \partial U(\mu)(y)}{x - y}.$$

The free Laplacian can then be expressed using the identification of the trace tensor product with the double integral as

$$\Theta_{\mathcal{A}} U(X) = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\partial U(\mu)(x) - \partial U(\mu)(y)}{x - y} \mu(dx) \mu(dy)$$

$$= 2 \int_{\mathbb{R}} \left( P.V. \int_{\mathbb{R}} \frac{\mu(dy)}{x-y} \right) \partial U(\mu)(x) \mu(dx).$$

The final form involves the Hilbert transform of the density, provided it exists. For background and similar computations, see [75, §5], [78, Proposition 3.5].

## APPENDIX C. AMALGAMATED FREE PRODUCTS

**C.1. Properties of free products of von Neumann algebras.** In this section, we record several useful properties of amalgamated free products and free independence. The first property is a form of associativity for free independence, which is well-known (see e.g. [2, Exercise 5.3.8], [67, Exercise 5.25], [60, Example 5.22]). The proof is similar to, but much easier, than the proof of Lemma C.4 below, so we omit it.

**Lemma C.1** (Associativity of free independence). *Let  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  be von Neumann subalgebras of a tracial von Neumann algebra  $\mathcal{M}$ . Let  $\mathcal{A} \vee \mathcal{B}$  be the von Neumann subalgebra generated by  $\mathcal{A}$  and  $\mathcal{B}$ . Then the following are equivalent:*

- (1)  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  are freely independent.
- (2)  $\mathcal{A}$  and  $\mathcal{B}$  are freely independent, and  $\mathcal{A} \vee \mathcal{B}$  and  $\mathcal{C}$  are freely independent.

The next property relates free independence with restriction to subalgebras.

**Lemma C.2** (Free independence passes to subalgebras). *Let  $J$  be any index set.*

- (1) *Suppose that  $\mathcal{A}$  is a non-commutative probability space containing  $\mathcal{C}$ . Let  $\mathcal{C} \subseteq \mathcal{B}_j \subseteq \mathcal{A}_j \subseteq \mathcal{A}$ . If the  $\mathcal{A}_j$ 's are freely independent over  $\mathcal{B}$ , then the  $\mathcal{B}_j$ 's are freely independent over  $\mathcal{C}$ .*
- (2) *Similarly, suppose non-commutative probability spaces  $\mathcal{C} \subseteq \mathcal{B}_j \subseteq \mathcal{A}_j$  are given. Then there is a unique tracial  $W^*$ -embedding  $*_{\mathcal{C}}(\mathcal{B}_j)_{j \in J} \rightarrow *_{\mathcal{C}}(\mathcal{A}_j)_{j \in J}$  that restricts to the given inclusions  $\mathcal{B}_j \rightarrow \mathcal{A}_j$  for each  $j$ .*

*Proof.* (1) For the  $\mathcal{B}_j$ 's to be freely independent over  $\mathcal{C}$  means that for all positive integers  $n$  and  $j : \{1, \dots, n\} \rightarrow J$  such that  $j(k) \neq j(k+1)$  for  $k = 1, \dots, n-1$ ,

$$E_{\mathcal{C}} \left( \prod_{k=1}^n (b_k - E_{\mathcal{C}}(b_k)) \right) = 0, \quad \text{for all } (b_1, \dots, b_n) \in \mathcal{B}_{j(1)} \times \dots \times \mathcal{B}_{j(n)}.$$

Since  $\mathcal{B}_{j(1)} \times \dots \times \mathcal{B}_{j(n)} \subseteq \mathcal{A}_{j(1)} \times \dots \times \mathcal{A}_{j(n)}$ , the free independence of the  $\mathcal{A}_j$ 's over  $\mathcal{C}$  implies free independence of the  $\mathcal{B}_j$ 's over  $\mathcal{C}$ .

(2) This follows from claim (1), since we can apply Corollary 2.2 to the subalgebra of  $*_{\mathcal{C}}\mathcal{A}_j$  generated by the images of  $\mathcal{B}_j$ 's.  $\square$

Next, we consider the behavior of certain conditional expectations in free products. The following fact is well-known in the study of von Neumann algebras.

**Lemma C.3.** *Consider an amalgamated free product  $\mathcal{A} *_C \mathcal{B}$  of tracial von Neumann algebras, and let  $\mathcal{C} \subseteq \mathcal{A}_0 \subseteq \mathcal{A}$ . Let  $\iota : \mathcal{A} \rightarrow \mathcal{A} *_C \mathcal{B}$  be the canonical inclusion. Using the previous lemma, view  $\mathcal{A}_0 *_C \mathcal{B}$  as a subset of  $\mathcal{A} *_C \mathcal{B}$ . Then for  $x \in \mathcal{A}$ , we have*

$$E_{\mathcal{A}_0 *_C \mathcal{B}} \circ \iota(x) = \iota \circ E_{\mathcal{A}_0}(x).$$

*Proof.* We regard  $\mathcal{A}$  as a subset of  $\mathcal{A} *_C \mathcal{B}$  in this proof, and hence suppress  $\iota$  in the notation. Let  $x' = x - E_{\mathcal{A}_0}(x)$ . It suffices to show that  $x'$  is orthogonal to  $\mathcal{A}_0 *_C \mathcal{B}$  inside  $\mathcal{A} *_C \mathcal{B}$ . We recall (see [68, p. 684-685]) that an  $L^2$ -dense subset of  $\mathcal{A}_0 *_C \mathcal{B}$  is spanned by alternating products of elements  $y_i$  from  $\mathcal{A}$  and  $z_i$  from  $\mathcal{B}$  respectively that are orthogonal to  $\mathcal{C}$ ; for instance, an

alternating product that starts with an element from  $\mathcal{A}$  and ends with an element from  $\mathcal{B}$  would be  $y_1 z_1 \dots y_k z_k$ . Hence, it suffices to show that  $x'$  is orthogonal to these elements. Note

$$x'(y_1 z_1 \dots y_k z_k) = (x' y_1) z_1 \dots y_k z_k$$

is another alternating product of elements from  $\mathcal{A}$  and  $\mathcal{B}$  that are orthogonal to  $\mathcal{C}$ ; indeed,

$$E_{\mathcal{C}}[x' y_1] = E_{\mathcal{C}} \circ E_{\mathcal{A}_0}[x' y_1] = E_{\mathcal{C}}[E_{\mathcal{A}_0}[x'] y_1] = 0,$$

and by assumption  $y_j \in \mathcal{A}_0 \subseteq \mathcal{A}$  and  $z_j \in \mathcal{B}$ . Therefore, by free independence of  $\mathcal{A}$  and  $\mathcal{B}$  over  $\mathcal{C}$ , we have  $E_{\mathcal{C}}[(x' y_1) z_1 \dots y_k z_k] = 0$  as desired. The other cases where the alternating string starts with an element of  $\mathcal{B}$  or ends with an element of  $\mathcal{A}$  are similar.  $\square$

The last and most substantial result is a free product manipulation for which several cases and related facts exist in the literature [52, Proposition 4.1]. We remark that the free independence manipulation of [70, Lemma 2.6] which changes the base subalgebra is somewhat related but does not imply the statement here.

**Lemma C.4.** *Let  $(\mathcal{A}_j)_{j \in J}$  be a family of tracial  $W^*$ -algebra containing a common subalgebra  $\mathcal{B}$ , let  $\mathcal{D} \subseteq \mathcal{B}$ , and let  $\mathcal{C}$  be another tracial  $W^*$ -algebra containing  $\mathcal{D}$ . For each  $i \in J$ , let*

$$\varphi_i : \mathcal{B} *_D \mathcal{C} \rightarrow \mathcal{A}_i *_D \mathcal{C}$$

*be the embedding given by Lemma C.2 applied to the inclusion  $\mathcal{B} \rightarrow \mathcal{A}_i$  and the identity  $\mathcal{C} \rightarrow \mathcal{C}$ . Let  $*_{\mathcal{B} *_D \mathcal{C}} (\mathcal{A}_j *_D \mathcal{C})_{j \in J}$  be the amalgamated free product over  $\mathcal{B} *_D \mathcal{C}$ , where  $\mathcal{B} *_D \mathcal{C}$  is viewed as a subalgebra of  $\mathcal{A}_i *_D \mathcal{C}$  via  $\varphi_i$ . Then we have a unique tracial  $W^*$ -isomorphism*

$$*_{\mathcal{B} *_D \mathcal{C}} \left( \mathcal{A}_j *_D \mathcal{C} \right)_{j \in J} \xrightarrow[\cong]{\Phi} \left( *_{\mathcal{B}} (\mathcal{A}_j)_{j \in J} \right) *_D \mathcal{C}$$

*which behaves in the natural way on each copy of  $\mathcal{A}_j$  and each copy of  $\mathcal{C}$ . More precisely, there is a unique isomorphism such that the following diagrams commute for each  $i \in J$ :*

$$\begin{array}{ccc} & \mathcal{A}_i & \\ \swarrow \rho_i & & \searrow \lambda_i \\ \mathcal{A}_i *_D \mathcal{C} & & *_{\mathcal{B}} (\mathcal{A}_j)_{j \in J} \\ \downarrow \tilde{\lambda}_i & & \downarrow \rho \\ *_{\mathcal{B} *_D \mathcal{C}} (\mathcal{A}_j *_D \mathcal{C})_{j \in J} & \xrightarrow{\Phi} & (*_{\mathcal{B}} (\mathcal{A}_j)_{j \in J}) *_D \mathcal{C}, \end{array}$$

*where  $\rho_i$ ,  $\lambda_i$ ,  $\tilde{\lambda}_i$ , and  $\rho$  are the canonical inclusions associated to the free product constructions, and similarly,*

$$\begin{array}{ccc} \mathcal{A}_i *_D \mathcal{C} & \xleftarrow{\sigma_i} & \mathcal{C} \\ \downarrow \tilde{\lambda}_i & & \downarrow \sigma \\ *_{\mathcal{B} *_D \mathcal{C}} (\mathcal{A}_j *_D \mathcal{C})_{j \in J} & \xrightarrow{\Phi} & (*_{\mathcal{B}} (\mathcal{A}_j)_{j \in J}) *_D \mathcal{C}, \end{array}$$

*where  $\sigma_i$  and  $\sigma$  are the canonical inclusions of  $\mathcal{C}$  associated to the free product constructions.*

*Proof.* Let  $\lambda_i : \mathcal{A}_i \rightarrow *_{j \in J} \mathcal{A}_j$  be the embedding given by the free product construction. By Lemma C.2 (2),  $\lambda_i$  and the identity map on  $\mathcal{B}$  induce an embedding

$$\pi_i : \mathcal{A}_i *_D \mathcal{C} \rightarrow (*_{\mathcal{B}} (\mathcal{A}_j)_{j \in J}) *_D \mathcal{C}.$$

Our goal is to show that the images  $\pi_j(\mathcal{A}_j *_D \mathcal{C})$  are freely independent with amalgamation over  $\mathcal{B} *_D \mathcal{C}$  in  $(*_{\mathcal{B}} (\mathcal{A}_j)_{j \in J}) *_D \mathcal{C}$ , and that they generate it. Indeed, if we can show this, then

the uniqueness of the free product with amalgamation over  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$  up to canonical isomorphism (Lemma 2.1) will imply the existence and uniqueness of the isomorphism  $\Phi$  with the desired properties. Moreover, the fact that  $(*_B(\mathcal{A}_j)_{j \in J}) *_{\mathcal{D}} \mathcal{C}$  is generated by the images of  $\pi_i$  is immediate because it is generated by the images of  $\mathcal{A}_j$  and  $\mathcal{C}$ . Thus, the main challenge is showing free independence over  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$ .

In the entire rest of the argument, we work in  $(*_B(\mathcal{A}_j)_{j \in J}) *_{\mathcal{D}} \mathcal{C}$  as the ambient space and view  $\mathcal{A}_i *_{\mathcal{D}} \mathcal{C}$  as a subset of it via the map  $\pi_i$ , and we therefore suppress  $\pi_i$  in our notation. To show free independence of  $\mathcal{A}_j *_{\mathcal{D}} \mathcal{C}$  over  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$ , it suffices to prove free independence over  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$  for weak-operator dense  $*$ -subalgebras of  $\mathcal{A}_j * \mathcal{C}$ ; this is well known and follows from a straightforward limiting argument as in [74, Proposition 2.5.7]. Therefore, given  $j : \{1, \dots, N\} \rightarrow J$  and  $X_i$  in some dense subalgebra of  $\mathcal{A}_{j(i)} *_{\mathcal{D}} \mathcal{C}$  with  $E_{\mathcal{B} *_{\mathcal{D}} \mathcal{C}}[X_i] = 0$ , we seek to show that  $E_{\mathcal{B} *_{\mathcal{D}} \mathcal{C}}[X_1 \dots X_N] = 0$ . We rely on the decomposition of elements in the free product  $\mathcal{A}_j *_{\mathcal{D}} \mathcal{C}$  as linear combinations of strings in  $\mathcal{A}_j$  and  $\mathcal{C}$ . Consider the dense subalgebra of  $\mathcal{A}_j *_{\mathcal{D}} \mathcal{C}$  generated by  $\mathcal{A}_j$  and  $\mathcal{C}$ ; these elements can be expressed as linear combinations of products of elements from  $x_i$  from  $\mathcal{A}_j$  and  $y_i$  from  $\mathcal{C}$  respectively that are orthogonal to  $\mathcal{D}$  (for instance, an alternating product that starts with an element from  $\mathcal{A}_j$  and ends with an element from  $\mathcal{C}$  would be  $x_1 y_1 \dots x_k y_k$ ). Thus our dense subalgebra can be expressed as

$$\mathcal{C} \oplus \bigoplus_{n=0}^{\infty} \mathcal{C} \cdot (\mathcal{A}_j^{\circ} \cdot \mathcal{C}^{\circ})^n \cdot \mathcal{A}_j^{\circ} \cdot \mathcal{C},$$

where  $\mathcal{A}_j^{\circ} = \{a \in \mathcal{A}_j : E_{\mathcal{D}}[a] = 0\}$  and  $\mathcal{C}^{\circ} = \{c \in \mathcal{C} : E_{\mathcal{D}}[c] = 0\}$ . These summands are, in fact, orthogonal to each other [68, p. 684-685]. Moreover, on each one, the inner product is given by

$$\tau((y_0 x_1 y_1 \dots x_k y_k)^* y'_0 x'_1 y'_1 \dots x'_k y'_k) = \tau(y_k^* E_{\mathcal{D}}[\dots x_1^* E_{\mathcal{D}}[y_0^* y_0] x'_1 \dots] y'_k).$$

In other words, we have a bimodule decomposition,

$$\mathcal{C} \oplus \bigoplus_{n=0}^{\infty} \mathcal{C} \otimes_{\mathcal{D}} (\mathcal{A}_j^{\circ} \otimes_{\mathcal{D}} \mathcal{C}^{\circ})^{\otimes n} \otimes_{\mathcal{D}} \mathcal{A}_j^{\circ} \otimes_{\mathcal{D}} \mathcal{C}.$$

This follows for instance from the way that the amalgamated free product is constructed (see [11, §4.7] and use for the  $\mathcal{H}_i$ 's the  $\mathcal{A}_j$ - $\mathcal{D}$  correspondence generated by  $\mathcal{A}_j$  and  $\mathcal{C}$ - $\mathcal{D}$ -correspondence generated by  $\mathcal{C}$ ). In summary, we have a dense subalgebra

$$\mathcal{A}_j *_{\mathcal{D}} \mathcal{C} \supseteq \mathcal{C} \oplus \bigoplus_{n=0}^{\infty} \mathcal{C} \otimes_{\mathcal{D}} (\mathcal{A}_j^{\circ} \otimes_{\mathcal{D}} \mathcal{C}^{\circ})^{\otimes n} \otimes_{\mathcal{D}} \mathcal{A}_j^{\circ} \otimes_{\mathcal{D}} \mathcal{C}.$$

representing that each element of  $\mathcal{A}_j *_{\mathcal{D}} \mathcal{C}$  is expressed as a sum of products  $c_1 a_0 c_2 \dots a_{n-1} c_n$  where  $a_j \in \mathcal{A}_j$  with  $E_{\mathcal{D}}[a_j] = 0$  and  $c_j \in \mathcal{C}$  with  $E_{\mathcal{D}}[c_j] = 0$  except for the first and last  $c_j$ 's, and the products for different lengths of strings are orthogonal to each other. Now we further decompose

$$\begin{aligned} \mathcal{A}_j^{\circ} &= \mathcal{B}^{\circ} \oplus \mathcal{A}_j^{\bullet}, & \text{where} \\ \mathcal{B}^{\circ} &= \{b \in \mathcal{B} : E_{\mathcal{D}}[b] = 0\}, \\ \mathcal{A}_j^{\bullet} &= \{a \in \mathcal{A}_j : E_{\mathcal{B}}[a] = 0\}. \end{aligned}$$

Hence, we obtain a dense subalgebra

$$\mathcal{A}_j *_{\mathcal{D}} \mathcal{C} \supseteq \mathcal{C} \oplus \bigoplus_{n=0}^{\infty} \bigoplus_{H_{\ell} \in \{\mathcal{B}^{\circ}, \mathcal{A}_j^{\bullet}\}} \mathcal{C} \otimes_{\mathcal{D}} H_1 \otimes_{\mathcal{D}} \mathcal{C}^{\circ} \otimes_{\mathcal{D}} \dots \otimes_{\mathcal{D}} H_n \otimes_{\mathcal{D}} \mathcal{C},$$

where the summands are orthogonal. When we restrict to the terms where all the  $H_j$ 's are  $\mathcal{B}^\circ$ , that produces a dense subalgebra of  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$ . Thus, since  $X_i$  is orthogonal to  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$ , we see that  $X_i$  can be expressed as a sum of terms in  $\mathcal{C} \otimes_{\mathcal{D}} H_1 \otimes_{\mathcal{D}} \mathcal{C}^\circ \otimes_{\mathcal{D}} \cdots \otimes_{\mathcal{D}} H_n \otimes_{\mathcal{D}} \mathcal{C}$  where at least one of the  $H_\ell$ 's is  $\mathcal{A}_j^\bullet$ . By taking linear combinations, we can assume without loss of generality that  $X_i$  is an element of one of the spaces  $L^2(\mathcal{C}) \otimes_{\mathcal{D}} H_1 \otimes_{\mathcal{D}} L^2(\mathcal{C})^\circ \otimes_{\mathcal{D}} \cdots \otimes_{\mathcal{D}} H_n \otimes_{\mathcal{D}} L^2(\mathcal{C})$ , so that

$$X_i = y_{i,0} x_{i,1} y_{i,1} \cdots x_{i,k_i} y_{i,n_i},$$

where  $y_{i,\ell} \in L^2(\mathcal{C})$  for all  $\ell$ ,  $y_{i,\ell} \in L^2(\mathcal{C})^\circ$  for  $0 < \ell < n_i$ , and  $x_{i,\ell} \in L^2(\mathcal{B})^\circ$  or  $x_{i,\ell} \in L^2(\mathcal{A}_{j(i)})^\bullet$ .

Now we proceed to show that  $X_1 \cdots X_n$  is orthogonal to  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$  using induction on the total length  $n_1 + \cdots + n_i$ . Since the orthogonal complement of  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$  is a bimodule over  $\mathcal{C}$ , we can factor out the first term  $y_{1,0}$  and thus assume without loss of generality that  $y_{1,0} = 1$ . Similarly, for  $i > 1$ , the term  $y_{i,0}$  can be combined with  $y_{i-1,n_{i-1}}$ , so without loss of generality  $y_{i,0} = 1$ . Next, each  $y_{i,n_i}$  can be expressed as  $E_{\mathcal{D}}[y_{i,n_i}] + (y_{i,n_i} - E_{\mathcal{D}}[y_{i,n_i}])$  and hence by linearity we can assume without loss of generality that each  $y_{i,n_i}$  is either in  $\mathcal{D}$  or in  $L^2(\mathcal{C})^\circ$ .

We want to apply the free product structure of  $\mathcal{A} *_{\mathcal{D}} \mathcal{C}$  where  $\mathcal{A} = *_B \mathcal{A}_j$ . Note as above that there is a dense subalgebra

$$\begin{aligned} L^2(\mathcal{A} *_{\mathcal{D}} \mathcal{C}) &\subseteq \mathcal{C} \oplus \bigoplus_{n=0}^{\infty} \mathcal{C} \otimes_{\mathcal{D}} (\mathcal{A}^\circ \otimes_{\mathcal{D}} \mathcal{C}^\circ)^{\otimes_{\mathcal{D}} n} \otimes_{\mathcal{D}} \mathcal{A}^\circ \otimes_{\mathcal{D}} \mathcal{C} \\ &\cong \mathcal{C} \oplus \bigoplus_{n=0}^{\infty} \bigoplus_{H_\ell \in \{\mathcal{B}^\circ, \mathcal{A}^\bullet\}} \mathcal{C} \otimes_{\mathcal{D}} H_1 \otimes_{\mathcal{D}} \mathcal{C}^\circ \otimes_{\mathcal{D}} \cdots \otimes_{\mathcal{D}} H_n \otimes_{\mathcal{D}} \mathcal{C}. \end{aligned}$$

Meanwhile,

$$X_1 \cdots X_n = (x_{1,1} y_{1,1} \cdots x_{1,n_1} y_{1,n_1}) \cdots (x_{N,1} y_{N,1} \cdots x_{N,n_N} y_{N,n_N})$$

It is important to keep track of what happens when we multiply  $x_{i,n_i}$ ,  $y_{i,n_i}$ , and  $x_{i+1,1}$  since  $y_{i,n_i}$  is not necessarily in  $\mathcal{C}^\circ$ . We distinguish for each  $i < N$  several cases:

- **Case 1:**  $y_{i,n_i} \in \mathcal{D}$ , and  $x_{i,n_i}$  and  $x_{i+1,1}$  are in  $\mathcal{B}^\circ$ .
- **Case 2:**  $y_{i,n_i} \in \mathcal{D}$ , and  $x_{i,n_i} \in \mathcal{A}_{j(i)}^\bullet$  and  $x_{i+1,1} \in \mathcal{B}^\circ$ .
- **Case 3:**  $y_{i,n_i} \in \mathcal{D}$ , and  $x_{i,n_i} \in \mathcal{B}^\circ$  and  $x_{i+1,1} \in \mathcal{A}_{j(i+1)}^\bullet$ .
- **Case 4:**  $y_{i,n_i} \in \mathcal{D}$ , and  $x_{i,n_i} \in \mathcal{A}_{j(i)}^\bullet$  and  $x_{i+1,1} \in \mathcal{A}_{j(i+1)}^\bullet$ .
- **Case 5:**  $y_{i,n_i} \in \mathcal{C}^\circ$ .

We will first show that if there is any term in Cases 1-3, we can reduce the claim to a shorter string and apply the induction hypothesis, and moreover that if all the terms fall under Cases 4 and 5, we obtain the desired orthogonality from the free product structure of  $\mathcal{A} *_{\mathcal{D}} \mathcal{C}$ .

**Case 1:** Note that  $x_{i,n_i} y_{i,n_i} x_{i+1,1} \in \mathcal{B}$  and thus can be written as  $d + b$  where  $d \in \mathcal{D}$  and  $b \in \mathcal{B}^\circ$ . Thus,

$$\begin{aligned} X_i X_{i+1} &= [x_{i,1} y_{i,1} \cdots x_{i,n_i-1} (y_{i,n_i-1} d y_{i+1,1})] [x_{i+1,2} \cdots x_{i+1,n_{i+1}} y_{i,n_{i+1}}] \\ &\quad + [x_{i,1} y_{i,1} \cdots x_{i,n_i-1} y_{i,n_i-1}] [b y_{i+1,1} \cdots x_{i+1,n_{i+1}} y_{i,n_{i+1}}], \end{aligned}$$

and hence  $X_1 \cdots X_n$  is a linear combination of two terms of the same form but with a lower total degree. Hence, by induction hypothesis, it is orthogonal to  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$  as desired.

**Case 2:** Since  $\mathcal{A}_{j(i)}^\bullet$  is a bimodule over  $\mathcal{B}$ , we have that  $x_{i,n_i} y_{i,n_i} x_{i+1,1} \in L^2(\mathcal{A}_{j(i)})^\bullet$ . Hence,

$$X_i X_{i+1} = [x_{i,1} y_{i,1} \cdots x_{i,n_i-1} y_{i-1,n_{i-1}} (x_{i,n_i} y_{i,n_i} x_{i+1,1}) y_{i+1,1}] [x_{i+1,2} y_{i+1,2} \cdots x_{i+1,n_{i+1}} y_{i,n_{i+1}}],$$

and so  $X_1 \cdots X_n$  has been expressed as a term of the same form but with a lower degree, and we may apply the induction hypothesis.

**Case 3:** The argument is symmetrical to Case 2.

At this point, we can assume that every index  $i$  is in Case 4 or Case 5. Besides, since we want to show that the element is orthogonal to  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$ , we can factor out any terms at the front that come from  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$ , and thus assume that  $x_{1,1} \in \mathcal{A}_{i(1)}^\bullet$ . We claim that  $X_1 \dots X_N$  can be expressed as

$$\begin{aligned} & a_{1,1} d_{1,1} \dots a_{1,k_1-1} d_{1,k_1-1} a_{1,k_1} b_{1,1} c_{1,1} \dots b_{1,\ell_1} c_{1,\ell_1} \\ & \dots \\ & a_{K,1} d_{K,1} \dots a_{K,k_K} d_{1,k_K} c_{1,K} b_{1,K} \dots c_{1,\ell_K} b_{1,\ell_K} c_{1,\ell_K+1}, \end{aligned}$$

where

- each  $a_{i,1} d_{i,1} \dots a_{i,k_i} d_{i,k_i}$  is an alternating product with  $a_{i,j}$  in one of the  $\mathcal{A}_t^\bullet$ 's and  $d_{i,j} \in \mathcal{D}$ ,
- each  $c_{i,j} \in \mathcal{C}^\circ$ ,
- each  $b_{i,j} \in \mathcal{B}^\circ$ .

This factorization arises as follows. The  $d_{i,j}$  terms come from the  $y_{i,n_i}$  terms that are in  $L^2(\mathcal{D})$  under Case 4. The  $c_{i,j}$  terms come from all the rest of the  $y_{i,j}$  terms, i.e. those which are in  $\mathcal{C}^\circ$ , and so in particular the junctures between  $X_i$  and  $X_{i+1}$  in Case 5 will produce  $c_{i,j}$  terms. Thus, each  $a_{i,1} d_{i,1} \dots a_{i,k_i} d_{i,k_i}$  corresponds to consecutive occurrences of Case 5; here intermediate terms  $d_{i,j} a_{i,j} d_{i,j+1}$  only occur when  $a_{i,j}$  comes from some  $X_{i'}$  of the form  $x_{i',1} y_{i',1}$ . Now we notice that  $a_{1,1} d_{1,1} \dots a_{1,k_1-1} d_{1,k_1-1} a_{1,k_1} \in \mathcal{A}^\bullet$  since the  $\mathcal{A}_j$ 's are freely independent over  $\mathcal{B}$ . Therefore,  $X_1 \dots X_N$  is an element in

$$\mathcal{C} \otimes_{\mathcal{D}} H_1 \otimes_{\mathcal{D}} \mathcal{C}^\circ \otimes_{\mathcal{D}} \dots \otimes_{\mathcal{D}} H_n \otimes_{\mathcal{D}} \mathcal{C},$$

where  $H_\ell \in \{\mathcal{A}^\bullet, \mathcal{B}^\circ\}$ , and there is at least one occurrence of  $\mathcal{A}^\bullet$ . Therefore, it follows from the direct sum decomposition that  $X_1 \dots X_N$  is orthogonal to  $\mathcal{B} *_{\mathcal{D}} \mathcal{C}$  as desired.  $\square$

## C.2. Amalgamation of filtrations and free Brownian motions.

**Lemma C.5** (Lemma 3.7). *Let  $0 \leq t_0 \leq t_1 \leq T$  and  $K$  be any index set. Let  $\mathcal{A}$  be a tracial von Neumann algebra equipped with a filtration  $(\mathcal{A}_t)_{t \in [t_0, t_1]}$  and a compatible  $d$ -variable free Brownian motion  $(S_t^0)_{t \in [t_0, t_1]}$ . Let  $(\mathcal{B}^k)_{k \in K}$  be a family of tracial von Neumann algebras and let  $\iota_k : \mathcal{A} \rightarrow \mathcal{B}^k$  be a tracial  $W^*$ -embedding. Suppose each  $\mathcal{B}^k$  has a filtration  $(\mathcal{B}_t^k)_{t \in [t_1, T]}$  and a compatible  $d$ -variable free Brownian motion  $(S_t^k)_{t \in [t_1, T]}$ . Assume that  $\iota_k(\mathcal{A}_{t_1}) \subseteq \mathcal{B}_{t_1}^k$ .*

*Then there exists an algebra  $\mathcal{B}$ , a tracial  $W^*$ -embeddings  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  and  $\tilde{\iota}_k : \mathcal{B}^k \rightarrow \mathcal{B}$ , a filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$ , and a compatible  $d$ -variable free Brownian motion  $(S_t)_{t \in [t_0, T]}$  such that the following hold:*

- (1) *We have  $\tilde{\iota}_k \circ \iota_k|_{\mathcal{A}_{t_0}} = \iota|_{\mathcal{A}_{t_0}}$  for each  $k \in K$ .*
- (2) *We have  $\iota(S_t^0) = S_t$  for  $t \in [t_0, t_1]$ .*
- (3) *We have  $\tilde{\iota}_k(S_t^k) = S_t - S_{t_1}$  for  $t \in [t_1, T]$ .*
- (4) *We have  $\iota(\mathcal{A}_t) \subseteq \mathcal{B}_t$  for  $t \in [t_0, t_1]$ .*
- (5) *We have  $\tilde{\iota}_k(\mathcal{B}_t^k) \subseteq \mathcal{B}_t$  for  $t \in [t_1, T]$ .*

*Furthermore, for any given  $k_0 \in K$ , the embedding  $\iota$  can be taken to be  $\iota = \tilde{\iota}_{k_0} \circ \iota_{k_0}$ .*

*Proof.* We want to construct  $\mathcal{B}$  as an amalgamated free product of the  $\mathcal{B}^k$ 's over the subalgebra generated by  $\mathcal{A}_{t_1}$  and a free Brownian motion.

First, let  $\mathcal{Z}$  be the von Neumann algebra generated by a free Brownian motion  $(Z_t)_{t \in [t_1, T]}$  with initial condition  $Z_{t_1} = 0$ . The existence of such an algebra follows from Voiculescu's free Gaussian construction. It is also well known in free probability that  $(\mathcal{Z}, (Z_t)_{t \in [t_1, T]})$  is unique

up to isomorphism, meaning that if  $(\tilde{\mathcal{Z}}, (\tilde{Z}_t)_{t \in [t_1, T]})$  is another tracial von Neumann algebra generated by a free Brownian motion, then there is a unique trace-preserving  $*$ -isomorphism  $\phi : \mathcal{Z} \rightarrow \tilde{\mathcal{Z}}$  such that  $\phi(Z_t) = \tilde{Z}_t$  for  $t \in [t_1, T]$ . We explain the proof briefly here for completeness. Consider non-commutative polynomials in variables indexed by  $t \in [t_1, T]$  and  $j = 1, \dots, d$ ; this is a well-defined  $*$ -algebra where each polynomial only depends on finitely many variables, even though there are uncountably many variables overall. We have  $\tau(p(Z_t : t \in [t_1, T])) = \tau(p(\tilde{Z}_t : t \in [t_1, T]))$  for every  $p$  since the joint moments of the  $S_t$ 's are uniquely determined by it being a Brownian motion. It follows that  $\|p(Z_t : t \in [t_1, T])\|_{L^2} = \|p(\tilde{Z}_t : t \in [t_1, T])\|_{L^2}$ , hence there is an isometric map  $L^2(W^*(Z_t : t \in [t_1, T])) \rightarrow L^2(W^*(\tilde{Z}_t : t \in [t_1, T]))$ . Since the operator norm can be evaluated by testing the inner products against polynomials, we also obtain  $\|p(Z_t^k : t \in [t_1, T])\| = \|p(\tilde{Z}_t^k : t \in [t_1, T])\|$ , and so this restricts to a isomorphism of tracial  $W^*$ -algebras  $W^*(Z_t : t \in [t_1, T]) \rightarrow W^*(\tilde{Z}_t : t \in [t_1, T])$ .

Let  $\tilde{\mathcal{A}}$  be the free product  $\mathcal{A}_{t_1} * \mathcal{Z}$ , and denote the associated embeddings

$$\lambda_1 : \mathcal{A}_{t_1} \rightarrow \tilde{\mathcal{A}}, \quad \lambda_2 : \mathcal{Z} \rightarrow \tilde{\mathcal{A}}.$$

We claim that there is a unique tracial  $W^*$ -embedding

$$\varphi_k : \tilde{\mathcal{A}} \rightarrow \mathcal{B}^k$$

such that

$$\varphi_k \circ \lambda_1 = \iota_k|_{\mathcal{A}_{t_1}}, \quad \varphi_k \circ \lambda_2(Z_t) = S_t^k \text{ for } t \in [t_1, T]. \quad (\text{C.1})$$

First, by uniqueness of the tracial  $W^*$ -algebra generated by a free Brownian motion proved above, we obtain an isomorphism  $\mathcal{Z} \rightarrow W^*(S_t^k : t \in [t_1, T])$  that sends  $Z_t$  to  $S_t^k$ . Then because  $\iota_k \circ \lambda_1(\mathcal{A}_{t_1})$  is contained in  $\mathcal{B}_{t_1}^k$  and the Brownian motion  $(S_t^k)_{t \in [t_1, T]}$  is compatible with the filtration  $(\mathcal{B}_t^k)_{t \in [t_1, T]}$ , we have that  $\iota_k \circ \lambda_1(\mathcal{A}_{t_1})$  is freely independent of  $W^*(S_t^k : t \in [t_1, T])$ . Since the free product is unique up to canonical isomorphism, we have a unique isomorphism  $\varphi_k$  from  $\tilde{\mathcal{A}} = \mathcal{A}_{t_1} * \mathcal{Z}$  to the subalgebra of  $\mathcal{B}^k$  generated by  $\iota_k \circ \lambda_2(\mathcal{A}_{t_1})$  and  $W^*(S_t^k : t \in [t_1, T])$  satisfying the desired conditions (C.1).

Now let  $\mathcal{B}$  be the amalgamated free product

$$\mathcal{B} = \ast_{\tilde{\mathcal{A}}} (\mathcal{B}^k)_{k \in K}$$

with respect to the inclusions  $\varphi_k : \tilde{\mathcal{A}} \rightarrow \mathcal{B}^k$  described above, and let  $\tilde{\iota}_k : \mathcal{B}^k \rightarrow \mathcal{B}$  be the canonical inclusion of  $\mathcal{B}^k$  into the amalgamated free product, so that  $\psi := \tilde{\iota}_k \circ \varphi_k : \tilde{\mathcal{A}} \rightarrow \mathcal{B}$  is independent of  $k$ . To obtain the embedding from  $\mathcal{A}$  into  $\mathcal{B}$ , fix  $k_0 \in K$  arbitrary and let  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  be given by  $\iota = \tilde{\iota}_{k_0} \circ \iota_{k_0}$ . (This embedding will not be unique since our requirements in the theorem statement only determine  $\iota|_{\mathcal{A}_{t_0}}$ .) The filtration  $\mathcal{B}_t$  is defined as follows:

$$\mathcal{B}_t = \psi \circ \lambda_1(\mathcal{A}_t) \text{ for } t \in [t_0, t_1]$$

and

$$\mathcal{B}_t = W^*(\tilde{\iota}_k(\mathcal{B}_t^k) : k \in K) \text{ for } t \in [t_1, T].$$

This is a valid filtration since we assumed that  $\iota_k(\mathcal{A}_{t_1}) \subseteq \mathcal{B}_{t_1}^k$ . The condition that  $\iota(\mathcal{A}_t) \subseteq \mathcal{B}_t$  for  $t \in [t_0, t_1]$  is immediate since  $\iota|_{\mathcal{A}_{t_1}} = \psi|_{\mathcal{A}_{t_1}}$ . Similarly, for  $t \in [t_1, T]$ , we have  $\tilde{\iota}_k(\mathcal{B}_t^k) \subseteq \mathcal{B}_t$  by construction.

Define a  $d$ -variable free Brownian motion  $(S_t)_{t \in [t_0, T]}$  as follows:

- Set  $S_t = \psi(S_t^0)$  for  $t \in [t_0, t_1]$ , which is well-defined since  $S_t \in \tilde{\mathcal{A}}_{t_1}$ .
- Set  $S_t = S_{t_0} + \psi(Z_t)$  for  $t \in [t_1, T]$ . Note that  $S_t = S_{t_0} + \tilde{\iota}_k(S_t^k)$  as well since  $\psi = \tilde{\iota}_k \circ \varphi_k$ .



To check that this is actually a free Brownian motion, note that  $(S_t : t \in [t_0, t_1])$  and  $(S_t - S_{t_1} : t \in [t_1, T])$  are free Brownian motions, and by construction,  $(S_t - S_{t_1} : t \in [t_1, T]) = (\psi(Z_t) : t \in [t_1, T])$  is freely independent of  $\psi(\mathcal{A}_{t_0})$  which contains  $(S_t : t \in [t_0, t_1])$ . From here it is easy to check the free independence of increments in  $[t_0, T]$  by splitting into cases based on the intervals  $[t_0, t_1]$  and  $[t_1, T]$ .

Finally, we need to check that the free Brownian motion  $(S_t)_{t \in [t_0, T]}$  is compatible with the filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$ . To see adaptedness, note that by construction, we have

$$S_t = \psi(S_t^0) \in \psi(\mathcal{A}_t) = \mathcal{B}_t \text{ for } t \in [t_0, t_1]$$

and

$$S_t = \tilde{\iota}_k(S_t^k) + \psi(S_{t_1}^0) \in \tilde{\iota}_k(\mathcal{B}_t^k) + \psi(\mathcal{A}_{t_1}) = \tilde{\iota}_k(\mathcal{B}_t^k) \text{ for } t \in [t_1, T].$$

Then we have to check that for  $t \in [t_0, T]$ , we have free independence of  $\mathcal{B}_t$  and  $W^*(S_s - S_t : s \in [t, T])$ .

First consider the case where  $t \in [t_0, t_1]$ . Then  $\mathcal{B}_t = \psi(\mathcal{A}_t)$  and  $S_s - S_t$  for  $s \geq t$  are both in the image of  $\psi$ . Specifically, we have  $S_s = \psi(\tilde{S}_s^0)$ , where

$$\tilde{S}_t^0 = \begin{cases} S_t^0, & t \in [t_0, t_1] \\ S_{t_1}^0 + Z_t, & t \in [t_1, T]. \end{cases}$$

Since  $\psi$  is trace-preserving, it suffices to show that  $\mathcal{A}_t$  and  $W^*(\tilde{S}_s^0 - \tilde{S}_t^0)$  are freely independent in  $\tilde{\mathcal{A}}$ . For each of notation, for  $t_1 \leq a \leq b \leq T$ , let

$$\mathcal{C}_{a,b} = W^*(\tilde{S}_s^0 - \tilde{S}_a^0 : s \in [a, b]).$$

Since  $S_t^0$  is compatible with the filtration on  $\mathcal{A}$ , we know that  $\mathcal{A}_t$  and  $\mathcal{C}_{t,t_1}$  are freely independent. Furthermore, by the construction of  $\tilde{\mathcal{A}}$  as a free product, we see that  $\mathcal{C}_{t_1,T}$  is freely independent of  $\mathcal{A}_{t_1}$ , and in particular freely independent of the  $W^*$ -algebra generated by  $\mathcal{A}_t$  and  $\mathcal{C}_{t,t_1}$ . By the associativity property of free products (Lemma C.1), it follows that  $\mathcal{A}_t$  and  $\mathcal{C}_{t,t_1}$  and  $\mathcal{C}_{t_1,T}$  are freely independent. Then by associativity again,  $\mathcal{A}_t$  is freely independent of the  $W^*$ -algebra generated by  $\mathcal{C}_{t,t_1}$  and  $\mathcal{C}_{t_1,T}$ , which is  $\mathcal{C}_{t,T}$ . This is the claim we wanted to prove.

Next, consider the case where  $t \in [t_1, T]$ , which is the more difficult case. Using the notation above, we need to show that  $\psi(\mathcal{C}_{t,T})$  is freely independent of  $\mathcal{B}_t$ . Let  $\tilde{\iota}_k(\mathcal{B}_t^k) \vee \psi(\mathcal{C}_{t,T})$  be the von Neumann subalgebra of  $\mathcal{B}$  generated by  $\tilde{\iota}_k(\mathcal{B}_t^k)$  and  $\psi(\mathcal{C}_{t,T})$ , which is canonically isomorphic to the free product  $\tilde{\iota}_k(\mathcal{B}_t^k) * \psi(\mathcal{C}_{t,T})$ . By Lemma C.2, since

$$\psi \circ \lambda_1(\tilde{\mathcal{A}}) \subseteq \tilde{\iota}_k(\mathcal{B}_t^k) \vee \psi(\mathcal{C}_{t,T}) \subseteq \tilde{\iota}_k(\mathcal{B}_k),$$

we see that the algebras  $(\tilde{\iota}_k(\mathcal{B}_t^k) \vee \psi(\mathcal{C}_{t,T}))_{k \in K}$  are freely independent with amalgamation over  $\psi \circ \lambda_1(\tilde{\mathcal{A}})$ . Hence, by Corollary 2.2,

$$\mathcal{B}_t \vee \psi(\mathcal{C}_{t,T}) = \bigvee_{k \in K} (\tilde{\iota}_k(\mathcal{B}_{t,k}) \vee \psi(\mathcal{C}_{t,T}))$$

is canonically isomorphic to the free product with amalgamation over  $\psi \circ \lambda_1(\tilde{\mathcal{A}})$

$$*_\mathcal{A}(\mathcal{B}_{t,k} \vee \varphi_k(\mathcal{C}_{t,T}))_{k \in K}.$$

Furthermore, by the choice of filtration on  $\mathcal{B}_k$ , we have that  $\mathcal{B}_t^k$  is freely independent of  $W^*(S_s^k - S_t^k : s \in [t, T]) = \varphi_k(\mathcal{C}_{t,T})$ . Hence, we obtain an isomorphism

$$\sigma : \mathcal{B} \supseteq \mathcal{B}_t \vee \psi(\mathcal{C}_{t,T}) \rightarrow *_\mathcal{A}(\mathcal{B}_{t,k} * \mathcal{C}_{t,T})_{k \in K}$$

such that, writing  $\sigma_k$  for the canonical embedding of  $\mathcal{B}_{t,k} * \mathcal{C}_{t,T}$  into the free product, we have

$$\sigma \circ \tilde{\iota}_k|_{\mathcal{B}_t^k} = \sigma_k|_{\mathcal{B}_t^k}$$

and

$$\sigma \circ \psi|_{\mathcal{C}_{t,T}} = \sigma_k \circ \varphi_k|_{\mathcal{C}_{t,T}}.$$

Finally, recall by construction that

$$\tilde{\mathcal{A}} = \mathcal{A}_{t_1} * \mathcal{C}_{t_1,T} \cong \mathcal{A}_{t_1} * \mathcal{C}_{t_1,t} * \mathcal{C}_{t,T} = \mathcal{D} * \mathcal{C}_{t,T},$$

where we set  $\mathcal{D} = \mathcal{A}_{t_1} * \mathcal{C}_{t_1,t}$ . Hence, we have

$$\mathcal{B}_t \vee \psi(\mathcal{C}_{t,T}) \cong_{\mathcal{D} * \mathcal{C}_{t,T}}^* (\mathcal{B}_t^k * \mathcal{C}_{t,T})_{k \in K},$$

where the isomorphism respects the inclusions of various subalgebras as described above. We claim that there is a unique isomorphism

$$\Phi : {}_{\mathcal{D} * \mathcal{C}_{t,T}}^* (\mathcal{B}_t^k * \mathcal{C}_{t,T})_{k \in K} \rightarrow \left( {}_{\mathcal{D}}^* (\mathcal{B}_t^k)_{k \in K} \right) * \mathcal{C}_{t,T},$$

which respects the inclusions of various algebras in the way one would expect, namely, the following diagrams commute for each  $i \in K$ :

$$\begin{array}{ccc} & \mathcal{B}_t^i & \\ \swarrow \rho_i & & \searrow \lambda'_i \\ \mathcal{B}_t^i * \mathcal{C}_{t,T} & & {}_{\mathcal{D}}^* (\mathcal{B}_t^k)_{k \in K} \\ \downarrow \tilde{\lambda}'_i & & \downarrow \rho \\ {}_{\mathcal{D} * \mathcal{C}_{t,T}}^* (\mathcal{B}_t^k * \mathcal{C}_{t,T})_{k \in K} & \xrightarrow{\Phi} & ({}_{\mathcal{D}}^* (\mathcal{B}_t^k)_{k \in K}) * \mathcal{C}_{t,T}, \end{array}$$

and

$$\begin{array}{ccc} \mathcal{B}_t^i * \mathcal{C}_{t,T} & \xleftarrow{\sigma'_i} & \mathcal{C}_{t,T} \\ \downarrow \tilde{\lambda}'_i & & \downarrow \sigma' \\ {}_{\mathcal{D} * \mathcal{C}_{t,T}}^* (\mathcal{B}_t^k * \mathcal{C}_{t,T})_{j \in J} & \xrightarrow{\Phi} & ({}_{\mathcal{D}}^* (\mathcal{B}_t^k)_{k \in K}) * \mathcal{C}_{t,T}, \end{array}$$

where the maps  $\rho_i$ ,  $\lambda'_i$ ,  $\tilde{\lambda}'_i$ ,  $\rho$ ,  $\sigma'_i$ ,  $\sigma'$  in these diagrams are the canonical inclusions associated to the respective free product constructions (see Lemma 2.1). We give a proof of this claim in Lemma C.4 in the appendix. Overall, we have the following chain of isomorphisms:

$$\mathcal{B} \supseteq \mathcal{B}_t \vee \psi(\mathcal{C}_{t,T}) \xrightarrow{\sigma} {}_{\tilde{\mathcal{A}}}^* (\mathcal{B}_{t,k} * \mathcal{C}_{t,T})_{k \in K} \rightarrow {}_{\mathcal{D} * \mathcal{C}_{t,T}}^* (\mathcal{B}_t^k * \mathcal{C}_{t,T})_{k \in K} \xrightarrow{\Phi} \left( {}_{\mathcal{D}}^* (\mathcal{B}_t^k)_{k \in K} \right) * \mathcal{C}_{t,T},$$

which all respect the inclusions of  $\mathcal{B}_t^k$  and  $\mathcal{C}_{t,T}$  in the way we expect. In the right-most expression  $({}_{\mathcal{D}}^* (\mathcal{B}_t^k)_{k \in K}) * \mathcal{C}_{t,T}$ , the subalgebra generated by the images of  $(\mathcal{B}_t^k)_{k \in K}$  is freely independent of  $\mathcal{C}_{t,T}$ ; hence, mapping these subalgebras back into  $\mathcal{B}$  by the chain of isomorphisms, we see that  $\bigvee_{k \in K} \tilde{\iota}_k(\mathcal{B}_{t,k})$  is freely independent of  $\psi(\mathcal{C}_{t,T})$  in  $\mathcal{B}$ , which is what we wanted to prove.  $\square$

We continue with the proof of Lemma 3.8 which we restate here.

**Lemma C.6** (Lemma 3.8). *Suppose that Assumption A holds. Let  $\mathcal{A}$  be a tracial  $W^*$ -algebra, let  $x_0 \in L^2(\mathcal{A})_{\text{sa}}^d$ , and let  $t_0 \in [0, T]$ . Then for every  $\epsilon > 0$ , there exists a tracial  $W^*$ -algebra  $\mathcal{B}$ , a tracial  $W^*$ -embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$ , and a control policy  $\alpha \in \mathbb{A}_{\mathcal{B}, \iota x_0}^{t_0, T}$  associated with a filtration  $(\mathcal{B}_t)_{t \in [t_0, T]}$  and compatible  $d$ -variable free Brownian motion  $(S_t)_{t \in [t_0, T]}$ , such that*

- (1)  $\iota(\mathcal{A}) \subseteq \mathcal{B}_{t_0}$
- (2)  $\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\tilde{\alpha}], \alpha_t) dt + g_{\mathcal{B}}(X_T[\alpha]) \right] \leq \bar{V}_{\mathcal{A}}(t_0, x_0) + \epsilon.$

*Proof.* First, by definition of  $\bar{V}_{\mathcal{A}}(t_0, x_0)$ , there exists some tracial  $W^*$ -embedding  $\iota_0 : \mathcal{A} \rightarrow \mathcal{B}^0$  and some control policy  $\tilde{\alpha}^0$  in  $\mathbb{A}_{\mathcal{B}^0, \iota_0(x_0)}^{t_0, T}$  such that

$$\mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}^0}(X_t[\tilde{\alpha}^0], \alpha_t^0) dt + g_{\mathcal{B}^0}(X_T[\alpha^0]) \right] \leq \bar{V}_{\mathcal{A}}(t_0, x_0) + \epsilon.$$

Let  $(\mathcal{B}_t^0)_{t \in [t_0, T]}$  be the corresponding filtration and  $(S_t^0)_{t \in [t_0, T]}$  the corresponding free Brownian motion.

Next, we define a tracial von Neumann algebra  $\mathcal{B}^1$  as follows. By Voiculescu's free Gaussian functor construction, there exists a tracial von Neumann algebra  $\mathcal{C}$  generated by a  $d$ -variable free Brownian motion  $(Z_t)_{t \in [t_0, T]}$  with initial condition  $Z_{t_0} = 0$ . For  $t_0 \leq a \leq b \leq T$ , let  $\mathcal{C}_{a,b}$  be the von Neumann subalgebra of  $\mathcal{C}$  generated by  $(Z_s - Z_a)_{s \in [a, b]}$ . Let  $\mathcal{B}^1$  be the free product  $\mathcal{A} * \mathcal{C}$ . Let

$$\iota_1 : \mathcal{A} \rightarrow \mathcal{B}^1, \quad \psi : \mathcal{C} \rightarrow \mathcal{B}^1$$

be the canonical inclusions from the free product construction. Define  $S_t^1 = \psi(Z_t)$  for  $t \in [t_0, T]$ . Define the filtration

$$\mathcal{B}_t^1 = \iota_1(\mathcal{A}) \vee \psi(\mathcal{C}_{t_0, t}), \quad t \in [t_0, T].$$

We claim that  $(S_t^1)_{t \in [t_0, T]}$  is compatible with the filtration  $(\mathcal{B}_t^1)_{t \in [t_0, T]}$ . To prove this, first note that it is adapted because  $S_t^1 = \psi(Z_t) \in \psi(\mathcal{C}_{t_0, t}) \subseteq \mathcal{B}_t^1$ . Next, we show that for each  $t \in [t_0, T]$ , the algebra  $W^*(S_s^1 - S_t^1 : s \in [t, T])$  is freely independent of  $\mathcal{B}_t^1$ . Note that  $W^*(S_s^1 - S_t^1 : s \in [t, T]) = \psi(\mathcal{C}_{t, T})$ . Recall that  $\mathcal{C}_{t_0, t}$  and  $\mathcal{C}_{t, T}$  are freely independent. Hence,  $\psi(\mathcal{C}_{t_0, t})$  and  $\psi(\mathcal{C}_{t, T})$  are freely independent of each other, and  $\iota_1(\mathcal{A})$  is freely independent of  $\psi(\mathcal{C}_{t_0, t}) \vee \psi(\mathcal{C}_{t, T})$ . Therefore, by associativity of free products (Lemma C.1),  $\iota_1(\mathcal{A})$  and  $\psi(\mathcal{C}_{t_0, t})$  and  $\psi(\mathcal{C}_{t, T})$  are freely independent. Using associativity again, it follows that  $\psi(\mathcal{C}_{t, T})$  is freely independent of  $\iota_1(\mathcal{A}) \vee \psi(\mathcal{C}_{t_0, t}) = \mathcal{B}_t^1$ , which is what we wanted to prove.

Now we apply Lemma 3.7 using the index set  $K = \{0, 1\}$  and using the algebras  $\mathcal{B}^0$  and  $\mathcal{B}^1$  with the associated filtrations and Brownian motions above. Set  $t_1 = t_0$ , and for the filtration in  $\mathcal{A}$  on the degenerate time interval  $[t_0, t_0]$ , which reduces to just a single subalgebra, use  $\mathcal{A}_{t_0} = W^*(x_0)$ . Then  $\iota_0(\mathcal{A}_{t_0}) \subseteq \mathcal{B}_{t_0}^0$  because the definition of admissible control policies requires that  $x_0 \in L^2(\mathcal{B}_{t_0}^0)_{\text{sa}}^d$ . Also,  $\iota_1(\mathcal{A}_{t_0}) \subseteq \iota_1(\mathcal{A}) = \mathcal{B}_{t_0}^1$  by construction. Let  $\mathcal{B}$ ,  $\tilde{\iota}_0$ ,  $\tilde{\iota}_1$ ,  $(\mathcal{B}_t)_{t \in [t_0, T]}$ , and  $(S_t)_{t \in [t_0, T]}$  be as in the previous lemma. By that lemma, we may take the embedding  $\iota : \mathcal{A} \rightarrow \mathcal{B}$  to be given by  $\iota = \tilde{\iota}_1 \circ \iota_1$ . Thus, by (5) of the previous lemma,

$$\iota(\mathcal{A}) = \tilde{\iota}_1 \circ \iota_1(\mathcal{A}) = \tilde{\iota}_1(\mathcal{B}_{t_0}^1) \subseteq \mathcal{B}_{t_0},$$

which verifies the first condition we wanted to prove. Let  $\alpha := \tilde{\iota}_0 \circ \alpha^0$ . By construction,

$$\alpha_t = \tilde{\iota}_0(\alpha_t^0) \in \tilde{\iota}_0(\mathcal{B}_t^0) \subseteq \mathcal{B}_t,$$

so that  $((\alpha_t)_{t \in [t_0, T]}, (\mathcal{B}_t)_{t \in [t_0, T]}, (S_t)_{t \in [t_0, T]})$  is an admissible control policy in  $\mathcal{B}$ . Note that  $X_t[\alpha] = \tilde{\iota}_0(X_t[\alpha^0])$ . Since  $L$  and  $g$  are tracial  $W^*$ -functions, we have

$$\begin{aligned} \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}}(X_t[\alpha], \alpha_t) dt + g_{\mathcal{B}}(X_T[\alpha]) \right] &= \mathbb{E} \left[ \int_{t_0}^T L_{\mathcal{B}^0}(X_t[\alpha^0], \alpha_t^0) dt + g_{\mathcal{B}^0}(X_T[\alpha^0]) \right] \\ &\leq \bar{V}_{\mathcal{A}}(t_0, x_0) + \epsilon, \end{aligned}$$

which verifies the second condition that we wanted to prove.  $\square$

## REFERENCES

- [1] Luigi Ambrosio and Jin Feng. On a class of first-order hamilton-jacobi equations in metric spaces. *J. Diff. Eq.*, 256(7):2194–2245, 2014.
- [2] Greg W Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*. Number 118 in Cambridge series in advanced mathematics. Cambridge university press, 2010.
- [3] Michael Anshelevich. Itô formula for free stochastic integrals. *Journal of Functional Analysis*, 188(1):292–315, 2002.
- [4] Mohit Bansil, Alpár R Mészáros, and Chenchen Mou. Global well-posedness of displacement monotone degenerate mean field games master equations. *arXiv:2308.16167v2*, 2024.
- [5] Erhan Bayraktar, Ibrahim Ekren, and Xin Zhang. Comparison of viscosity solutions for a class of second order pdes on the wasserstein space. *ArXiv:2309.05040*, 2023.
- [6] Charles Bertucci, Mérouane Debbah, Jean-Michel Lasry, and Pierre-Louis Lions. A spectral dominance approach to large random matrices. *Journal de Mathématiques Pures et Appliquées*, 164:27–56, 2022.
- [7] Philippe Biane. *Free Brownian Motion, Free Stochastic Calculus and Random Matrices*, volume 12 of *Fields Institute Communications*, pages 1–19. American Mathematical Society, Providence, RI, 1997.
- [8] Philippe Biane, Mireille Capitaine, and Alice Guionnet. Large deviation bounds for matrix brownian motion. *Inventiones mathematicae*, 152:433–459, 2003.
- [9] Philippe Biane and Roland Speicher. Stochastic calculus with respect to free brownian motion and analysis on wigner space. *Probability theory and related fields*, 112:373–409, 1998.
- [10] Philippe Biane and Dan Voiculescu. A free probability analogue of the wasserstein metric on the trace-state space. *Geometric & Functional Analysis GAFA*, 11(6):1125–1138, 2001.
- [11] Nathaniel P. Brown and Narutaka Ozawa. *C\*-algebras and finite-dimensional approximations*, volume 88 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2008.
- [12] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The master equation and the convergence problem in mean field games:(ams-201)*. Princeton University Press, 2019.
- [13] Eric A. Carlen and Jan Maas. Gradient flow and entropy inequalities for quantum markov semigroups with detailed balance. *Journal of Functional Analysis*, 273(5):1810–1869, 2017.
- [14] René Carmona, Mark Cerenzia, and Aaron Zeff Palmer. The dyson and coulomb games. In *Annales Henri Poincaré*, volume 21, pages 2897–2949. Springer, 2020.
- [15] René Carmona, François Delarue, et al. *Probabilistic theory of mean field games with applications I*. Springer, 2018.
- [16] Yongxin Chen, Wilfrid Gangbo, Tryphon T Georgiou, and Allen Tannenbaum. On the matrix monge-kantorovich problem. *European Journal of Applied Mathematics*, 31(4):574–600, 2020.
- [17] Yongxin Chen, Tryphon T Georgiou, and Allen Tannenbaum. Matrix optimal mass transport: a quantum mechanical approach. *IEEE Transactions on Automatic Control*, 63(8):2612–2619, 2017.
- [18] Andrea Cosso, Fausto Gozzi, Idris Kharroubi, Huyên Pham, and Mauro Rosestolato. Master bellman equation in the wasserstein space: Uniqueness of viscosity solutions. *Transactions of the American Mathematical Society*, 377(01):31–83, 2024.
- [19] Michael G Crandall, Lawrence C Evans, and P-L Lions. Some properties of viscosity solutions of hamilton-jacobi equations. *Transactions of the American Mathematical Society*, 282(2):487–502, 1984.
- [20] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations. *Transactions of the American mathematical society*, 277(1):1–42, 1983.
- [21] Michael G Crandall and Pierre-Louis Lions. Hamilton-jacobi equations in infinite dimensions i. uniqueness of viscosity solutions. *Journal of functional analysis*, 62(3):379–396, 1985.
- [22] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations in infinite dimensions. i. uniqueness of viscosity solutions. *J. Funct. Anal.*, 62, 1985.
- [23] Michael G Crandall and Pierre-Louis Lions. Hamilton-jacobi equations in infinite dimensions. ii. existence of viscosity solutions. *Journal of Functional Analysis*, 65(3):368–405, 1986.
- [24] Michael G Crandall and Pierre-Louis Lions. Hamilton-jacobi equations in infinite dimensions, iii. *Journal of functional analysis*, 68(2):214–247, 1986.
- [25] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations in infinite dimensions. ii. existence of viscosity solutions. *J. Funct. Anal.*, 65, 1986.
- [26] Michael G Crandall and Pierre-Louis Lions. Viscosity solutions of hamilton-jacobi equations in infinite dimensions. iii. existence of viscosity solutions. *J. Funct. Anal.*, 65, 1986.
- [27] Michael G Crandall and PL Lions. Viscosity solutions of hamilton-jacobi equations in infinite dimensions. iv. hamiltonians with unbounded linear terms. *Journal of Functional Analysis*, 90(2):237–283, 1990.

- [28] Michael G Crandall and PL Lions. Viscosity solutions of hamilton-jacobi equations in infinite dimensions. v. unbounded linear terms and b-continuous solutions. *Journal of functional analysis*, 97(2):417–465, 1991.
- [29] Yoann Dabrowski. A Laplace principle for Hermitian Brownian motion and free entropy I: the convex functional case. arXiv:1604.06420, 2017.
- [30] Yoann Dabrowski. A non-commutative path space approach to stationary free stochastic differential equations. arxiv:1006.4351, 2010.
- [31] Giacomo De Palma and Dario Trevisan. Quantum optimal transport with quantum channels. *Ann. Henri Poincaré*, 22:3199–3234, 2021.
- [32] Bruce K. Driver, Brian C. Hall, and Todd Kemp. The large- $n$  limit of the Segal-Bargmann transform on  $u_n$ . *Journal of Functional Analysis*, 265(11):2585 – 2644, 2013.
- [33] Rocco Duvenhage. Quadratic Wasserstein metrics for von Neumann algebras via transport plans. *Journal of Operator Theory*, 88(2):289–308, 2022.
- [34] Freeman J Dyson. A brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6):1191–1198, 1962.
- [35] Lawrence C Evans. On solving certain nonlinear partial differential equations by accretive operator methods. *Israel Journal of Mathematics*, 36:225–247, 1980.
- [36] Giorgio Fabbri, Fausto Gozzi, and Andrzej Swiech. Stochastic optimal control in infinite dimension. *Probability and Stochastic Modelling. Springer*, 2017.
- [37] Dario Feliciangeli, Augusto Gerolin, and Lorenzo Portinale. A non-commutative entropic optimal transport approach to quantum composite systems at positive temperature. *Journal of Functional Analysis*, page 109963, 2023.
- [38] Wilfrid Gangbo, David Jekel, Kyeongsik Nam, and Dimitri Shlyakhtenko. Duality for optimal couplings in free probability. *Communications in Mathematical Physics*, 396(3):903–981, 2022.
- [39] Wilfrid Gangbo, Sergio Mayorga, and Andrzej Swiech. Finite dimensional approximations of hamilton-jacobi-bellman equations in spaces of probability measures. *SIAM J. Math. Anal.*, 53(2):1320–1356, 2021.
- [40] Wilfrid Gangbo and Alpár R. Mészáros. Global well-posedness of master equations for deterministic displacement convex potential mean field games. *Communications on Pure and Applied Mathematics*, 75(12):2685–2801, 2022.
- [41] Wilfrid Gangbo, Alpár R Mészáros, Chenchen Mou, and Jianfeng Zhang. Mean field games master equations with nonseparable hamiltonians and displacement monotonicity. *The Annals of Probability*, 50(6):2178–2217, 2022.
- [42] Wilfrid Gangbo and Andrzej Swiech. Existence of a solution to an equation arising from mean field games. *J. Differential Equations*, 259(11):6573–6643, 2015.
- [43] Wilfrid Gangbo and Andrzej Swiech. Metric viscosity solution to hamilton-jacobi equations depending on local slopes. *Calculus of Variations and Partial Differential Equations*, 54(1):1183 – 1218, 2015.
- [44] Wilfrid Gangbo, Nguyen Truyen, and Adrian Tudorascu. Hamilton-jacobi equations in the wasserstein space. *Meth. Appl. Anal.*, 15(2):155–184, 2008.
- [45] Wilfrid Gangbo and Adrian Tudorascu. On differentiability in the wasserstein space and well-posedness for hamilton-jacobi equations. *Journal de Mathématiques Pures et Appliquées*, 125:119–174, 2019.
- [46] Diogo A. Gomes and João Saúde. Mean field games models—a brief survey. *Dyn Games Appl*, 4:110–154, 2014.
- [47] Alice Guionnet and Dimitri Shlyakhtenko. Free diffusions and matrix models with strictly convex interaction. *Geometric and Functional Analysis*, 18(6):1875–1916, 2009.
- [48] Alice Guionnet and Dimitri Shlyakhtenko. Free monotone transport. *Inventiones Mathematicae*, 197(3):613–661, 09 2014.
- [49] Alice Guionnet and Ofer Zeitouni. Concentration of the Spectral Measure for Large Matrices. *Electronic Communications in Probability*, 5:119–136, 2000.
- [50] F. Hiai and Y. Ueda. Free transportation cost inequalities for noncommutative multi-variables. *Infinite Dimensional Analysis, Quantum Probability, and Related Topics*, 9:391–412, 2006.
- [51] Fumio Hiai, Dénes Petz, and Yoshimichi Ueda. Free transportation cost inequalities via random matrix approximation. *Probab. Theory Related Fields*, 130(2):199–221, 2004.
- [52] Cyril Houdayer. On some free products of von neumann algebras which are free araki-woods factors. *Int. Math. Res. Not. IMRN*, 2007(23):21, 2007.
- [53] Ching-Peng Huang, Dominik Inauen, and Govind Menon. Motion by mean curvature and dyson brownian motion. *Electron. Commun. Probab.*, 28(34):1–10, 2023.

- [54] Minyi Huang, Roland P Malhamé, and Peter E Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in information and systems*, 2006.
- [55] Hitoshi Ishii. Viscosity solutions of nonlinear second-order partial differential equations in hilbert spaces. *Communications in partial differential equations*, 18(3-4):601–650, 1993.
- [56] Espen Robstad Jakobsen, Artur Rutkowski, and Rocco Duvenhage. The master equation for mean field game systems with fractional and nonlocal diffusions. *arXiv:2305.18867*, 88(2):289–308, 2022.
- [57] David Jekel. An elementary approach to free entropy theory for convex potentials. *Analysis & PDE*, 13(8):2289–2374, 2020.
- [58] David Jekel. Optimal transport for types and convex analysis for definable predicates in tracial  $w^*$ -algebras. *Journal of Functional Analysis*, 287(9):110583, 2024.
- [59] David Jekel, Wuchen Li, and Dimitri Shlyakhtenko. Tracial smooth functions of non-commuting variables and the free wasserstein manifold. *Dissertationes Mathematicae*, 580:1–150, 2022.
- [60] David Jekel and Weihua Liu. An operad of non-commutative independences defined by trees. *Dissertationes Mathematicae*, 553:1–100, 2020.
- [61] David A Jekel, Todd A Kemp, and Evangelos A Nikitopoulos. A martingale approach to noncommutative stochastic calculus. *arXiv preprint arXiv:2308.09856*, 2023.
- [62] David Andrew Jekel. *Evolution equations in non-commutative probability*. University of California, Los Angeles, 2020.
- [63] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.
- [64] Pierre-Louis Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. part i: The case of bounded stochastic evolutions. 1988.
- [65] Elizabeth Meckes and Mark Meckes. Spectral measures of powers of random matrices. *Electron. Commun. Probab.*, 18:1–13, 2013.
- [66] Govind Menon. The complex burgers’ equation, the hciz integral and the calogero-moser system. *preprint*, 2017.
- [67] Alexandru Nica and Roland Speicher. *Lectures on the Combinatorics of Free Probability*. London Mathematical Society Lecture Note Series. Cambridge University Press, 2006.
- [68] Sorin Popa. Markov traces on universal jones algebras and subfactors of finite index. *Invent Math*, 111:375–405, 1993.
- [69] Irving E Segal. Irreducible representations of operator algebras. 1947.
- [70] Dimitri Shlyakhtenko. Free entropy with respect to a completely positive map. *American Journal of Mathematics*, 122(1):45–81, 02 2000.
- [71] Roland Speicher. A new example of ‘independence’ and ‘white noise’. *Probability theory and related fields*, 84(2):141–159, 1990.
- [72] Roland Speicher. Combinatorial theory of the free product with amalgamation and operator-valued free probability theory. *Mem. Amer. Math. Soc.*, 132(627), 1998.
- [73] Dan Voiculescu. Limit laws for random matrices and free products. *Inventiones mathematicae*, 104(1):201–220, 1991.
- [74] Dan-Virgil Voiculescu. Symmetries of some reduced free product  $C^*$ -algebras. In Huzihiro Araki, Calvin C. Moore, Șerban-Valentin Stratila, and Dan-Virgil Voiculescu, editors, *Operator Algebras and their Connections with Topology and Ergodic Theory*, pages 556–588. Springer Berlin Heidelberg, Berlin, Heidelberg, 1985.
- [75] Dan-Virgil Voiculescu. The analogues of entropy and Fisher’s information in free probability, I. *Communications in Mathematical Physics*, 155(1):71–92, 1993.
- [76] Dan-Virgil Voiculescu. The analogues of entropy and of Fisher’s information in free probability, II. *Inventiones Mathematicae*, 118:411–440, 1994.
- [77] Dan-Virgil Voiculescu. Operations on certain non-commutative operator-valued random variables. In *Recent advances in operator algebras*, number 232 in Astérisque, pages 243–275. Societe mathematique de France, 1995.
- [78] Dan-Virgil Voiculescu. The analogues of entropy and of Fisher’s information in free probability V. *Inventiones Mathematicae*, 132:189–227, 1998.
- [79] Dan-Virgil Voiculescu. A strengthened asymptotic freeness result for random matrices with applications to free entropy. *Internat. Math. Res. Not. IMRN*, (1):41–63, 1998.
- [80] Dan-Virgil Voiculescu. The analogues of entropy and of fisher’s information measure in free probability theory, VI: Liberation and mutual free information. *Advances in Mathematics*, 146:101–166, 1999.

- [81] Dan-Virgil Voiculescu. The coalgebra of the difference quotient and free probability. *International Mathematics Research Notices*, 2000(2):79–106, 2000.
- [82] Dan-Virgil Voiculescu, Kenneth J. Dykema, and Alexandru Nica. *Free Random Variables*, volume 1 of *CRM Monograph Series*. American Mathematical Society, Providence, 1992.
- [83] Eugene P Wigner. Random matrices in physics. *SIAM review*, 9(1):1–23, 1967.
- [84] John Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.
- [85] Jiongmin Yong and Xun Yu Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Science & Business Media, 1999.