STOCHASTIC CONTROL WITH SIGNATURES

P. BANK, C. BAYER, P.P. HAGER, S. RIEDEL, AND T. NAUEN

ABSTRACT. This paper proposes to parameterize open loop controls in stochastic optimal control problems via suitable classes of functionals depending on the driver's path signature, a concept adopted from rough path integration theory. We rigorously prove that these controls are dense in the class of progressively measurable controls and use rough path methods to establish suitable conditions for stability of the controlled dynamics and target functional. These results pave the way for Monte Carlo methods to stochastic optimal control for generic target functionals and dynamics. We discuss the rather versatile numerical algorithms for computing approximately optimal controls and verify their accurateness in benchmark problems from Mathematical Finance.

1. INTRODUCTION

Many stochastic optimal control approaches are based on dynamic programming. One starts by considering the dynamics of a controlled system driven by some source of randomness, originating for instance from a Brownian motion. Controls are then specified in open loop form as a function(al) adapted to this driving randomness or constructed in closed form as feedback functions depending on the system itself. The evolution of the controlled system and the control jointly result in some cost or reward whose expectation is ultimately sought to be optimized. A key mathematical tool is the dynamic programming principle which in its infinitesimal form as a Hamilton-Jacobi-Bellman (HJB) equation for the value function often takes the form of a nonlinear (integro-)partial differential equation on state space; see [FS06] for a standard reference of this approach. Typically, these equations can be solved only numerically, and even then it remains a daunting task, particularly due to the well known "curse of dimensionality": If one tries to approximate the value function by computing its value in certain points, one quickly needs an impossibly large number of these. Alternative methods where approximations of the value function are parameterized through, e.g., a deep neural network (DNN) have been successfully applied in recent years also in higher dimensions (cf. the survey [GPW23] and [EHJ17, SS18, HPW20] to name just a few). But even these approaches reach their limits in situations where non-Markovian noise as generated, for instance, by a fractional Brownian motion requires one to store the whole history of the controlled system (or of at least parts of it). Even the derivation of the HJB equation can become challenging as one works on an infinite-dimensional state space; see [FGS17] for a recent monograph on optimal control in such functional analytic settings. In a finite-dimensional Markovian setting, parametrizations of closed loop (feedback) controls have been proposed in [GM05] and successfully extended using DNNs in [HE16] (see also [RW20, HH21] and the hybrid method [HPBL21]). Again the transition to the non-Markovian setting driven, e.g., by a fractional Brownian motion is far from evident as the entire history of the system (and maybe even more) has to be incorporated in the control. Alternatively, bespoke finite-dimensional approximations to the controlled system can be used. But these need to be carefully tailored to the problem under investigation; see, e.g., [BD20] and [BB23a, BHT19].

By contrast the approach proposed in the present paper uses what is called path signatures of the driving noise to consider suitable parameterizations of open loop policies. (A closed loop approach has also recently been proposed in [HFH$^+$23] on which we further comment below.)

We will prove that for many problems this signature based class of policies is sufficiently rich to include some which are arbitrarily close to optimal. Studied first by Chen [Che54], signatures were identified by Lyons [Lyo98] as the building block for his by now well developed theory of rough path integration. At any time the signature consists of all the iterated integrals of the components in the path under consideration against each other until the present moment. This infinite-dimensional tensor can be shown to (essentially) encode the full evolution of the path ([HL10], [BGLY16]). Therefore, truncated versions of the signature can be expected to be an efficient finite-dimensional encoding of path information that can be used by numerical approaches. Even so, the space of progressively measurable policies still corresponds to the typically vast space of measurable, non-anticipative mappings on the path space of the driving noise. It is thus remarkable that even *linear* functionals of the signature can be shown to be dense in the space of continuous functionals over compact path spaces. This was also used by [KLPA20] who gave a Stone-Weierstrass argument which is also at the heart of our proof of pathwise denseness of linear or DNN-parameterized families of signature based policies in the class of all admissible ones (cf. Prop. 4.5). For the workhorse model of stochastic optimal control where ones seeks to minimize costs like

$$\mathbb{E}[L(Y^U, U)] = \mathbb{E}\left[\int_0^T f(t, Y_t^U, U_t)\mathrm{d}t + g(Y_T^U)\right]$$

subject to system dynamics such as

(1.1) $$Y_0^U = y_0, \quad \mathrm{d}Y_t^U = b(Y_t^U, U_t)\mathrm{d}t + \sigma(Y_t^U)\mathrm{d}X_t,$$

a stability result (Theorem 4.1) shows that this kind of denseness is sufficient for our purposes (given some integrability assumptions).

Remarkably, this remains true when, instead of the usual choice of X as a Brownian motion, one passes to a stochastic rough driver $\mathbb{X}$ such as a fractional Brownian motion W^H. As a consequence, even for such non-Markovian, inherently infinite-dimensional drivers the same signature-based policies can be used for computing approximately optimal ones. Indeed, machine learning tools such as stochastic gradient descent become applicable as soon as one can efficiently produce Monte Carlo samples of $L(Y^U, U)$ for a given signature policy U. This optimization procedure is therefore highly versatile and largely model-independent: only its dynamics Y^U need to be generated for an open loop control determined from samples of the extended driver's $(t, \mathbb{X}_t)_{t \geq 0}$ (truncated) signature. The latter can for instance be generated even offline for any driver of interest and then re-used for different dynamics driven by the same $\mathbb{X}$.

We illustrate the effectiveness of this optimization procedure by working out two case studies on linear-quadratic stochastic optimal control problems. In either case a fractional Brownian motion makes these problems challenging to address by dynamic programming methods. In the first one, an analytic formula for the problem's value is nonetheless available and our numerically computed signature policies turn out to approximate it very well. In the second case study the benchmark is provided by a highly original signature-based numerical method due to [KLPA20] which even manages to transform the optimization problem into a deterministic one that just draws on the *expected* signature. Here again our more versatile approach performs well. In either case, the signature turns out be needed only up to level 3 or 4, which corresponds to an 5 respectively 8-dimensional encoding (at least when using log-signatures) of the extended fractional Brownian motion's full path $(s, W_s^H)_{s \leq t}$.

Related literature. Initiated by [DFG17], there is ongoing development in the pathwise control of rough differential equations ([AC20, CHT24]) and the associated *rough* HJB equations.

Although these works are only tangentially related to our topic, they require similar stability results for the rough dynamic system with respect to the control. To impose minimal assumptions on the control and the regularity of the rough signal, we present a novel stability result.

The theoretical analysis of continuous-time non-Markovian control problems already constitutes a vast body of literature, a large part of which deals with dynamic stochastic control problems where the driver is Markov but the coefficients are path dependent (e.g. delayed systems [KS97]). We therefore restrain the discussion to works that have an ambition in implementable numerical methods for problems including a non-Markovian driving signal.

The work [LOSar] proposes a discretization method that applies to the optimal drift control of a (path-dependent) differential equation driven by fractional Brownian motion with constant volatility. Exploiting the underlying Brownian filtration they parameterize open loop controls as a functional of a martingale discretization of Brownian motion, which are then trained based on a dynamic programming principle for the (augmented) discretized structure. Notably the authors are able to prove convergence rates towards the optimal value. In general, as we comment in more detail below, such discretiziation methods suffer from a proportional increase in the state space dimension.

The recent paper [HFH⁺23] presents a novel method based on neural rough differential equations, originally introduced in [KMFL20, MSKF21], which serve as a continuous-time analogue to recurrent neural networks. In this sense, the method can be seen as the continuous-time extension of [HH21]. Similar to our work, [HFH⁺23] employs a Monte Carlo-based method by directly parametrizing the control, though in their case, the parameterization is done for feedback (closed loop) controls. Signatures appear in [HFH⁺23] as a special case when using linear activation functions, a perspective the authors use to argue for the density of their approximations. In this case, the signature arises from the system's state rather than its driver. In contrast, our work focusses on open-loop approximations and on proving in a mathematically rigorous manner that the optimal control value can be approximated. Therefore, our research provides new theoretical insights that can also support similar approximation results for the method proposed in [HFH⁺23].

The paper [KLPA20] (and later also [CPASB20]) considers strategies given as linear functionals of the dynamically updated signature of the system's driver. It focusses on optimization problems where the target functional can be expressed as a linear functional given in terms of a strategy dependent shuffle polynomial acting on the expected signature. The shuffle product used there ensures that the space of linear signature functionals forms an algebra, making the class of such target functionals larger than one might be led to believe from the outset, even though this problems class is far from the generic formulation of a stochastic optimal control problem as considered in the present paper.

In [CM24] signature models are used for stochastic portfolio theory, hence, portfolio optimization, see also [FHW23] for a very similar approach. The recent works [CGMSF23, CGSF23] analyze asset price models (without optimization) given as SDEs with coefficients defined as linear functionals of signatures of underlying stochastic processes. They calibrate such models (jointly) to VIX and SPX options.

As an alternative to signature methods, numerical approximation of stochastic optimal control problems in non-Markovian settings can be used. We will specifically comment on two such approaches recently explored for specific examples.

The first such approach is to discretize the problem in time and to then enhance the state variable by the whole history, see, for instance, [BCJ19] for a specific example. Hence, if the original state processes evolved in $\mathbb{R}^d$ and we are using a time-grid of length N, then the enhanced state process evolves in $\mathbb{R}^{dN}$. Obviously, this potentially magnifies the curse of dimensionality, unless we can rely on powerful techniques for dealing with high-dimensional problems (such as

deep learning as used in [BCJ19] or tensor-trains as used in [BEST23]). In principle, the same issue arises for signature methods (i.e., the dimension of the truncated tensor algebra is much higher than the dimension of the original state space), but signatures are empirically seen to be highly efficient encoding of path properties, implying that the dimension of the state space required to effectively approximate the solution is often much lower. Crucially, the dimension of the truncated signature does *not* depend on the time-discretization. This is especially relevant when the underlying process is rough (e.g., fractional Brownian motion with low Hurst index), and therefore requires fine discretization of the stochastic optimal control problem.

A second workable approximation strategy can also be based on a *Markovian approximation* of the non-Markovian state process, see, for instance, [BD20] for an application to portfolio optimization in rough volatility models, as well as [BB23a, BHT19] for American options. In essence, the idea is to consider the same control problem, but for a surrogate state process, which is a Markov process. For instance, when the state process is given in terms of a stochastic Volterra equation, there is a general approach [BD20, BB23a] for constructing multi-factor Markovian approximations essentially of Ornstein-Uhlenbeck type. This approximation will also highly enlarge the dimension of the state space, but, under suitable conditions, the approximation method can be extremely efficient, see [BB23b] for a concrete example. In any case, these constructions are highly problem specific, unlike the general, model-free architecture proposed in this paper. Even if a good approximation to the value function – now expressed in terms of the multi-factor surrogate process – is available, actually applying the corresponding strategy in the context of the original problem might be difficult, as it would require "inverting" the Markovian approximation in a path-wise manner, which is generally not a well-posed problem.

Outline of the paper. In Section 2 we formalize the standard control problem studied in the paper, spelling out and discussing in particular the technical assumptions on the coefficients in the dynamics of Y^U and in the cost functional $L = L(Y^U, U)$.

In Section 3 we provide an introduction to signatures and provide tools needed from rough path theory. We also describe the natural (deterministic!) rough path setting for the analysis of stochastic optimal control problems, namely *stopped rough paths*; see also [KLPA20]. We also recall necessary prerequisites from the theory of rough differential equations (RDEs), and provide a new existence and uniqueness result for RDEs with a function-valued parameter—think of the control as a function in time.

Section 4 introduces two signature-based classes of controls to approximate general progressively measurable controls U. Specifically, we consider *linear signature policies* in $\mathcal{A}_{\text{lin}}$ which are just linear functionals of the present signature. Alternatively, *deep signature controls* from $\mathcal{A}_{\text{DNN}}$ are obtained by using the path's signature (or, more precisely, its log-signature) as input in a neural network. Proposition 4.5 shows that both linear signature controls and deep signature controls are dense in the set $\mathcal{A}$ of admissible controls. By stability results of the solution map of controlled stochastic rough differential equations, convergence of the controls implies convergence of the controlled process $\mathbb{Y}^U$ (Lemma 4.6), and this further implies that the associated costs also converge. Theorem 4.7 then concludes that the infimum of expected costs is the same over the three policy classes $\mathcal{A}$, $\mathcal{A}_{\text{lin}}$, and $\mathcal{A}_{\text{DNN}}$.

Finally, Section 5 gives details of the numerical method suggested, together with a discussion of numerical properties and possible extensions. We then provide two case studies. In the first, the driving process X is a fractional Brownian motion, which we try to force to stay close to 0 by controlling its drift. In order to obtain a well-posed problem, we penalize by the L^2 norm of the control. This problem's value is available analytically in closed form from [BSV17] and can thus be used as a benchmark. We observe excellent accuracy for both linear and deep signature methods.

In the second case study, we revisit an optimal execution problem already studied in [KLPA20]. Here, the driving process, again chosen to be fractional Brownian motion, corresponds to the fundamental price of a financial asset. We want to liquidate a position in that asset while maximizing the proceeds from selling which are adversely affected by price impact. The bespoke optimization procedure of [KLPA20] provides another valuable benchmark that we are able to match with our methodology.

2. An optimal control problem of a rough differential system

Let us start by describing more precisely the class of optimization problems we consider in this paper. We fix an underlying complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. *Admissible controls* will be $(\mathcal{F}_t)$-progressively measurable processes $U : [0, T] \times \Omega \to \mathcal{U}$ taking values in $\mathcal{U}$, a closed, convex and nonempty subset of $\mathbb{R}^k$. The set of such admissible controls will be denoted by $\mathcal{A}$. Any realization $u = U(\omega) : [0, T] \to \mathcal{U}$ can then be viewed as an element of $L^0(\mathrm{d}t; \mathcal{U})$, the space of measurable functions $u : [0, T] \to \mathcal{U}$ equipped with the topology of convergence in (Lebesgue) measure $\mathrm{Leb}(\mathrm{d}t) = \mathrm{d}t$.

We study the optimal control of the differential dynamics

$$(2.1) \qquad \mathrm{d}Y_t^U = b(Y_t^U, U_t)\mathrm{d}t + \sigma(Y_t^U)\mathrm{d}\mathbb{X}_t, \qquad t \in [0, T], \quad Y_0^U = y_0 \in \mathbb{R}^m,$$

where $b : \mathbb{R}^m \times \mathcal{U} \to \mathbb{R}^m$, $\sigma : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{m \times d}$ satisfy the Lipschitz assumptions (4.2) and (4.3) specified below. The stochastic driver $\mathbb{X}$ is a stochastic rough path with finite p-variation for some $p \in [1, \infty)$. The reader can think of $\mathbb{X}$ as a Brownian motion, but also a fractional Brownian motion can be considered. We refer to Section 3 for the precise details and an introduction to rough paths including the definition of the metric rough path space $\Omega_T^{p,0}(\mathbb{R}^m)$ we will consider for our driver $\mathbb{X}$. By Theorem 4.1 there exists then a unique solution Y^U to (2.1) which canonically extends to a stochastic rough path $\mathbb{Y}^U : \Omega \to \Omega_T^p(\mathbb{R}^m)$ and depends continuously on y_0, $\mathbb{X}$ and U. The optimization objective is to minimize the expected costs

$$J(U) := \mathbb{E}[L(\mathbb{Y}^U, U)]$$

among all admissible controls $U \in \mathcal{A}$, where the *cost functional*

$(2.2) \quad L : \mathbb{R}^m \times \Omega_T^p \times L^0(\mathrm{d}t; \mathcal{U}) \to \mathbb{R}$ is continuous and bounded.

It is then clear that there is a finite *optimal cost*

$$\inf_{U \in \mathcal{A}} J(U).$$

In Section 4.3 below we will characterize subclasses of admissible controls, sufficient for approximating the optimal cost. In fact, assuming additionally that

$(2.3) \quad \mathcal{F}_t$ is the completion of $\sigma(\mathbb{X}_s \mid 0 \leq s \leq t)$ by all $\mathbb{P}$-null sets of $\mathcal{F}$, $0 \leq t \leq T$,

it will turn out that it suffices to restrict to controls that are continuous functions on the pathspace Ω_T^p. This fact allows us to deduce that it also suffices to restrict to controls that are (linear) functionals of the signature of the time-augmented path $\widehat{\mathbb{X}}$ defined below.

In the rest of this section we will discuss the above assumptions in more detail and comment on some possible generalizations.

Form of the dynamics. Drift-controlled stochastic differential equations of the form (1.1) find many applications in various fields. Studying the rough path extension (2.1) allows us to treat non-semimartingale drivers $\mathbb{X}$ in a unified framework, including non-Markovian processes such

as fractional Brownian motion. The price to pay for this generality is that we cannot allow for a controlled diffusion part as in

$$(2.4) \qquad \mathrm{d}Y_t^U = b(Y_t^U, U_t)\mathrm{d}t + \sigma(Y_t^U, U_t)\mathrm{d}\mathbb{X}_t.$$

This is because the stochastic part of these dynamics will only be defined rigorously for a general rough path $\mathbb{X}$ if the control U is *controlled* by $\mathbb{X}$, that is, roughly speaking, it cannot fluctuate more than $\mathbb{X}$ itself; see [FH20] for an introduction to rough differential equations using *Gubinelli derivatives*. A stability result for systems of the form (2.4) similar to Theorem 4.1 has recently been established in [AC20] for the case $p \in [2, 3)$ under the assumption that the controls U have finite $p/2$-variation. In contrast to the present work, the authors in [DFG17, AC20] have considered a pathwise "anticipating" control of the system (2.1). In [DFG17] the restriction to drift-controlled system was imposed for non-degeneracy reasons, while [AC20] included control in the volatility by introducing a penalty term that ensures the required regularity of controls similar to a Tychonov regularization.

We stress that the results from Section 4 are not tied to the stability results for rough differential equations. Instead, we can study the optimal control of stochastic differential equations for which we obtain stability under weaker conditions. In fact, referring to [Pro05] one even has sufficient stability for path-dependent stochastic differential equations of the form

$$(2.5) \qquad \mathrm{d}Y_t = b(U_t, Y|_{[0,t]})\mathrm{d}t + \sigma(U_t, Y|_{[0,t]})\mathrm{d}X_t,$$

when X is a continuous semimartingale, assuming that b and σ are bounded and *functional* Lipschitz uniformly in U. Thus, to adapt our approximation results from Section 4 to cover the optimal control of (2.5) we only need to modify the assumption on the cost functional. More precisely, we need to require the continuous dependence of L on Y, not in the rough path sense, but in the uniform topology on $C([0, T]; \mathbb{R}^m)$ (see also the comments on the form of the cost functional below). This even extends to the case of càdlàg semimartingales, where the canonical lift is given by the iterated Marcus-integral (see [FS17] and [CF19]) and corresponding universal approximation theorems have recently been provided in [CPSFar].

Form of the cost functional. A general form of cost functionals satisfying the continuity assumption in (2.2) is given by

$$(2.6) \qquad L(\mathbf{y}, u) := \int_0^T f(t, y_t, u_t)\mathrm{d}t + g(y_T), \quad \mathbf{y} \in \Omega_T^p, \quad u \in L^0(\mathrm{d}t; \mathcal{U}),$$

where $f : [0, T] \times \mathbb{R}^m \times \mathcal{U} \to \mathbb{R}$, $g : \mathbb{R}^m \to \mathbb{R}$ are continuous and bounded. In view of such "classical" cost functionals the definition of L as a map on the rough path space may seem cumbersome. However, this definition does not lead to additional complications and allows us to consider cost functionals involving rough integrals such as

$$L(\mathbf{y}, u) := g\left(\int_0^T f(y_t)\mathrm{d}\mathbf{y}_t\right), \qquad \mathbf{y} \in \Omega_T^p, \quad u \in L^0(\mathrm{d}t; \mathcal{U}),$$

where $f \in \mathrm{Lip}^\gamma(\mathbb{R}^m \times \mathcal{U}; \mathbb{R})$ for some $\gamma > p$ (see the next section for the definition of this space) and $g : \mathbb{R} \to \mathbb{R}$ is continuous and bounded (see [FV10, Section 10.6] for definition of the rough integral and its continuity properties).

Boundedness of the cost functional. Assumption (2.2) allows us to disentangle the problem of approximating a given control from the convergence of the associated costs. More precisely, we prove the approximation of admissible controls in $\mathcal{A}$ by controls in a subclass $\mathcal{A}' \subset \mathcal{A}$ for almost sure convergence in $L^0(\mathrm{d}t; \mathcal{U})$. The convergence of the associated costs then follows from the stability result in Theorem 4.1 and the dominated convergence theorem. If one

were to consider unbounded cost functionals this disentanglement would in general not be possible. Firstly, the set of admissible controls is intrinsically related to the cost functional by $\mathcal{A} \subset \{U \text{ prog. meas.} \mid \mathbb{E}[|L(\mathbb{Y}^U, \mathbb{X})|] < \infty\}$. Secondly, the mode of convergence for approximating controls needs to be chosen accordingly to deduce the convergence of costs. Hence, the theoretical analysis becomes more problem-specific and possibly requires more elaborate universal approximation results for signature functionals.

To be more specific about the last statement, consider a quadratic control problem with a cost functional of the form

$$L(\mathbf{y}, u) := \int_0^T \left\{ a|y_t|^2 + b|u_t|^2 \right\} \mathrm{d}t, \quad \mathbf{y} \in \Omega_T^p, \quad u \in L^0(\mathrm{d}t; \mathcal{U}),$$

where $a, b > 0$ are constants and $\mathcal{U} = \mathbb{R}^k$. We consider examples of such problems in Section 5.2 and 5.3. In this case, assuming square-integrability of $\|\mathbb{X}\|_{p-\mathrm{var}}$, one needs to prove convergence of the approximating sequence of controls in $L^2(\mathrm{d}t \otimes \mathrm{d}\mathbb{P})$. For the sub-class of signature controls this cannot be easily deduced from the universal approximation in Proposition 3.3, since outside the chosen compact sets there is no general way to bound the costs incurred. Note, however, that recently global universal approximation theorems have become available for normalized signatures in [CO22] and on weighted functions spaces in [CST23]. In particular, [BPS23] provide a corresponding universal approximation for processes in $L^2(\mathrm{d}t \otimes \mathrm{d}\mathbb{P})$, thus allowing to directly adapt our theoretical results to the case of quadratic control problems.

Initial value and history. The assumption that $\mathbb{X}$ takes values in $\Omega_T^{p,0}$ implies a deterministic starting value $\mathbb{X}_0$. A straight forward generalization is an inclusion of a finite history of the path, i.e., by assuming that $\mathbb{X}$ is defined on the interval $[t_0, T]$ for some $t_0 < 0$ and then setting $\mathcal{F}_t$ to the completion of $\sigma(\mathbb{X}_s \mid t_0 \le s \le t)$ for all $t \in [0, T]$. In this case, the approximation results from Section 4 are extended by considering controls that are functionals of the signature started at time t_0.

3. Preliminaries on rough analysis and signatures

We are going to introduce the basic definitions and notation needed for the understanding rough differential equations and signatures. These definitions are standard in the rough path literature, we refer to [LCL07, FH20, FV10] for a more detailed exposition.

3.1. **The tensor algebra.** The sequence of iterated integrals of a smooth path satisfies algebraic relations that are consequence of the linearity of the integral and the integration by parts rule. These algebraic properties are most conveniently revealed when organizing these intergals as *tensor series*. In this section we introduce the basic algebraic concepts that will allow us to define signatures and rough paths.

Let V be a finite-dimensional $\mathbb{R}$-vector space with basis $\{e_1, \ldots, e_d\}$. We define the *extended tensor algebra* by setting

$$T((V)) := \prod_{n=0}^{\infty} V^{\otimes n}$$

where $V^{\otimes n}$ denotes the n-th tensor power of V with the convention $V^{\otimes 0} := \mathbb{R}$, $V^{\otimes 1} := V$. The algebraic structure on $T((V))$ is given by componentwise summation and tensor multiplication, i.e., for two tensor series $\mathbf{a} = (a_n)_{n=0}^{\infty}$ and $\mathbf{b} = (b_n)_{n=0}^{\infty}$ in $T((V))$ the tensor product is defined by

$$\mathbf{a} \otimes \mathbf{b} = \left(\sum_{k=0}^{n} a_k \otimes b_{n-k} \right)_{n=0,1,\ldots}.$$

We denote by $\mathbf{0} := (0, 0, \dots)$ and $\mathbf{1} := (1, 0, \dots)$ the neutral elements of summation and multiplication. The *tensor algebra* $T(V) \subset T((V))$ consists of those tensor series with only finitely many non-zero elements, which defines a subalgebra. There is a natural pairing between $T((V))$ and $T(V)$ given by

$$\langle \cdot, \cdot \rangle \colon T(V) \times T((V)) \to \mathbb{R}, \quad \langle \mathbf{a}, \mathbf{b} \rangle := a^0 b^0 + \sum_{k=1}^{\infty} \langle a^k, b^k \rangle,$$

where the summation is finite by the definition of $T(V)$ and $\langle a^k, b^k \rangle$ denotes the dot product[1] on $V^{\otimes k}$. We will frequently use this pairing by associating an element $\ell \in T(V)$ to the linear functional on the extended tensor algebra $\langle \ell, \cdot \rangle : T((V)) \to \mathbb{R}$.

The *truncated tensor algebra* is defined by

$$T^N(V) := \bigoplus_{n=0}^{N} V^{\otimes n}.$$

We define maps $\pi_n \colon T((V)) \to V^{\otimes n}$ and $\pi_{\leq N} \colon T((V)) \to T^N(V)$ by $\pi_n(\mathbf{a}) = a_n$ and $\pi_{\leq N}(\mathbf{a}) = (a_0, \dots, a_N)$. Note that $T^N(V)$ forms an algebra under the truncated tensor multiplication $\mathbf{a} \otimes_N \mathbf{b} := \pi_{\leq N}(\mathbf{a} \otimes \mathbf{b})$, for $\mathbf{a}, \mathbf{b} \in T^N(V)$. However, we will not distinguish between the multiplication symbols on $T^N(V)$, $T(V)$ and $T((V))$ and use $\otimes$ in all cases and write $\mathbf{0}$ and $\mathbf{1}$ for the neutral elements $\pi_{\leq N}(\mathbf{0})$ and $\pi_{\leq N}(\mathbf{1})$ in the truncated tensor algebra. Furthermore, $T^N(V)$ is a finite dimensional vector space which we equip with the norm

$$|\mathbf{a}| = \sum_{n=0}^{N} |a^n|, \qquad \mathbf{a} \in T^N(V),$$

where $|\cdot| = \sqrt{\langle \cdot, \cdot \rangle}$ denotes the Euclidean norm on $V^{\otimes n} \cong \mathbb{R}^{d^n}$ for any $n \in \mathbb{N}$.

The Lie-algebra generated from $\{\mathbf{e}_1, \dots, \mathbf{e}_d\}$, where $\mathbf{e}_i := (0, e_i, 0, \dots) \in T(V)$, and the commutator bracket

$$[\mathbf{a}, \mathbf{b}] = \mathbf{a} \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{a}, \qquad \mathbf{a}, \mathbf{b} \in T(V),$$

is called the *free Lie-algebra* $\mathfrak{g}(V)$ over V. It is a sub-algebra of $T((V))_0 = \{\mathbf{a} \in T((V)) \mid \pi_0(\mathbf{a}) = 0\}$. Its tensor exponential, i.e., its image under the map

$$(3.1) \qquad \exp_{\otimes} \colon T((V))_0 \to T((V)), \quad \mathbf{a} \mapsto \mathbf{1} + \sum_{n=1}^{\infty} \frac{1}{n!} \mathbf{a}^{\otimes n},$$

is called the *free Lie group* $G(V) := \exp_{\otimes}(\mathfrak{g}(V))$, which is a subgroup of $T((V))_1 = \{\mathbf{a} \in T((V)) \mid \pi_0(\mathbf{a}) = 1\}$. Indeed, $(G(V), \otimes)$ is a group with identity $\mathbf{1}$, where for $\mathbf{g} = \exp_{\otimes}(\mathbf{a}) \in G(V)$ the inverse element is given by $\mathbf{g}^{-1} = \exp_{\otimes}(-\mathbf{a})$. The inverse of the operation (3.1) is given by the tensor logarithm

$$(3.2) \qquad \log_{\otimes} \colon T((V))_1 \to T((V))_0, \quad (\mathbf{1} + \mathbf{a}) \mapsto \mathbf{1} + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \mathbf{a}^{\otimes n}.$$

We also set $\mathfrak{g}^N(V) := \pi_{\leq N}(\mathfrak{g}^N(V))$ and $G^N(V) := \pi_{\leq N}(G(V))$, which are the *free nilpotent Lie algebra* and *group* of order N. We equip these spaces with the subspace topology in $T^N(V)$. The truncated tensor exponential and logarithm

$$\exp_{\otimes}^N \colon T^N(V)_0 \to T^N(V)_1, \qquad \log_{\otimes}^N \colon T^N(V)_1 \to T^N(V)_0,$$

---

[1]This inner product on $V^{\otimes k}$ is defined as the bilinear extension of $\langle e_{i_1} \otimes \cdots \otimes e_{i_k}, e_{j_1} \otimes \cdots \otimes e_{j_k} \rangle :=$ $\langle e_{i_1}, e_{j_1} \rangle \cdots \langle e_{i_k}, e_{j_k} \rangle := \delta_{i_1, j_1} \cdots \delta_{i_k, j_k}$, where δ denotes the Kronecker-Delta.

are defined using the corresponding (finite!) power series in the truncated tensor algebra. Hence, these maps are smooth, and, furthermore, it holds $\log_\otimes^N = (\exp_\otimes^N)^{-1}$ and $G^N(V) = \exp_\otimes^N(\mathfrak{g}^N(V))$.

3.2. **Rough paths and their signatures.** One can start by defining the signature for a piecewise smooth path $x : [0, T] \to V$, i.e., assume that the derivative $\dot{x}$ is integrable and has only finitely many points of discontinuity. We define its signature $\mathrm{Sig}(x)_{s,t}$ over the interval $[s, t] \subset [0, T]$ as the tensor series of iterated integrals. More precisely, $\mathrm{Sig}(x)_{s,t} \in T((V))_1$ and for all $n = 1, 2, \ldots$ we set

$$\pi_n(\mathrm{Sig}(x)_{s,t}) := \int_s^t \int_s^{t_k} \cdots \int_s^{t_2} \dot{x}(t_1) \otimes \cdots \otimes \dot{x}(t_n) \, \mathrm{d}t_1 \cdots \mathrm{d}t_n.$$

One can show that as consequence of the integration by parts rule such signatures are elements of the free Lie group $G(V)$, i.e., $\mathrm{Sig}(x) : \Delta_T \to G(V)$, where $\Delta_T := \{(s, t) \in [0, T]^2 \, : \, 0 \le s \le t \le T\}$. Moreover, we have the Chen relation (see [Che57])

$$\mathrm{Sig}(x)_{s,t} = \mathrm{Sig}(x)_{s,u} \otimes \mathrm{Sig}(x)_{u,t}, \qquad 0 \le s \le u \le t \le T.$$

We set $\mathrm{Sig}(x)_t := \mathrm{Sig}(x)_{0,t}$ and also define the truncated signature by $\mathrm{Sig}(x)^{\le N} := \pi_{\le N}(\mathrm{Sig}(x))$.

The space $\Omega_T^p(V)$ of *geometric p-rough paths* for some $p \ge 1$ is constructed as the closure of the set of smooth paths under a p-variation metric that also measures the first $\lfloor p \rfloor$-levels of the signature, where $\lfloor p \rfloor$ denotes the integer part of p. To make this more precise, we first define the *p-variation*-distance of two group valued paths $\mathbf{x}, \mathbf{y} : [0, T] \to G^N(V)$ by

$$d_{p-\mathrm{var};[s,t]}(\mathbf{x}, \mathbf{y}) := \max_{k=1,\ldots,N} \sup_{\mathcal{D} \subset [s,t]} \left(\sum_{t_i \in \mathcal{D}} \left| \pi_k \Big(\mathbf{x}(t_i)^{-1} \otimes \mathbf{x}(t_{i+1}) - \mathbf{y}(t_i)^{-1} \otimes \mathbf{y}(t_{i+1}) \Big) \right|^{\frac{p}{k}} \right)^{\frac{k}{p}}$$

where the supremum ranges over all partitions $\mathcal{D} = \{t_i\}$ of $[s, t]$. We use the notation $d_{p-\mathrm{var}}(\mathbf{x}, \mathbf{y}) := d_{p-\mathrm{var};[0,T]}(\mathbf{x}, \mathbf{y})$. The space of geometric p-rough paths $\Omega_T^p(V)$ is then defined as the set of continuous paths $\mathbf{x} : [0, T] \to G^{\lfloor p \rfloor}(V)$ that satisfy

(i) $d_{p-\mathrm{var};[s,t]}(\mathbf{1}, \mathbf{x}) < \infty$,

(ii) there exists a sequence of piecewise smooth paths $(x_n)_{n \ge 1}$ with $x_n : [0, T] \to V$ such that

$$\lim_{n \to \infty} d_{p-\mathrm{var}}\big(\mathbf{x}, \mathrm{Sig}(x_n)^{\le \lfloor p \rfloor}\big) = 0.$$

Unless stated otherwise we will abbreviate $\Omega_T^p = \Omega_T^p(V)$. On the subset

$$\Omega_T^{p,0} := \{\mathbf{x} \in \Omega_T^p \mid \mathbf{x}(0) = \mathbf{1}\},$$

$d_{p-\mathrm{var}}$ defines a metric. We can similarly equip Ω_T^p with a metric by additionally comparing starting values. With the thereby induced topology $\Omega_T^{p,0}$ and Ω_T^p are Polish spaces.

A key insight for geometric p-rough paths is that signature at high levels $N > \lfloor p \rfloor$ are uniquely determined by the signature up to level $\lfloor p \rfloor$. Indeed, the Lyons Extension Theorem states that every geometric rough path $\mathbf{x} \in \Omega_T^{p,0}$ has a unique lift $\mathrm{Sig}(\mathbf{x}) \in G(V)$, such that $\pi_{\le \lfloor p \rfloor}(\mathrm{Sig}(\mathbf{x})) = \mathbf{x}$ and also for every $N > \lfloor p \rfloor$ it holds $d_{p-\mathrm{var}}\big(\pi_{\le N}(\mathrm{Sig}(\mathbf{x})), \mathbf{1}\big) < \infty$ and the lifting map

$$\Omega_T^{p,0} \to G^N(V), \quad \mathbf{x} \mapsto \mathrm{Sig}(\mathbf{x})_{0,T}^{\le N}$$

is continuous. Note that for a piecewise smooth path $x : [0, T] \to V$ we have $\mathrm{Sig}(x)^{\le \lfloor p \rfloor} \in \Omega_T^{p,0}$ for any $p \ge 1$. Furthermore, the signature in the rough path sense is consistent with signature in the smooth sense, i.e., $\mathrm{Sig}(\mathrm{Sig}(x)^{\le \lfloor p \rfloor}) = \mathrm{Sig}(x)$.

We say that a continuous path $x : [0, T] \to V$ has a *lift* to a geometric p-rough path, if there exists $\mathbf{x} \in \Omega_T^p$ such that $\pi_1(\mathbf{x}) = x$. Note that this lift is not unique, even though there often is a canonical choice. For instance, this is the case for solutions to rough differential equations (see

the Section 3.5) and for many stochastic process. The canonical lift of a Brownian motion B, for example, is obtained by suitable piecewise linear approximations $(B^n)_{n \geq 1}$. It can be shown that for any $p \in (2,3)$ we have

$$\mathrm{Sig}(B^n)^{\leq 2} \xrightarrow[n \to \infty]{\Omega_T^p} \mathbb{B} := \left(1, B, \left(\int_0^{\cdot} B_s^i \circ \mathrm{d}B_s^j\right)_{i,j=1,\ldots,d}\right) \qquad \text{almost surely,}$$

where "$\circ \mathrm{d}B$" denotes the Stratonovich integration with respect to B. In particular, $\mathbb{B}$ defines a Ω_T^p valued random-variable, which we call the *Stratonovich-lift* of the Brownian motion B.

3.3. **Stopped rough paths.** We can define a stochastic rough path as an Ω_T^p-valued random variable. To introduce the notion of adeptedness, we will need to consider functionals of the restriction of a rough path to subintervals of $[0, T]$. This calls for the definition of an appropriate ambient space Λ_T, where we can compare paths defined on different segments of the time line. For rough paths this was originally proposed in [KLPA20], then elaborated in [BHRS23], and is motivated by the functional Itô calculus, see [Dup19, CF10].

We call $\Lambda_T := \bigcup_{t \in [0,T]} \Omega_t^{p,0}$ the *space of stopped rough paths.* Note that

$$\Lambda_T = \{\mathbf{x}|_{[0,t]} \mid \mathbf{x} \in \Omega_T^{p,0}, \quad t \in [0,T]\}.$$

Following [CF10] and [CM24] we also call a map $f : \Lambda_T^p \to \mathbb{R}$ a *non-anticipative functional.* To indicate from which set $\Omega_t^{p,0}$ and element $\mathbf{x} \in \Lambda_T$ is chosen, we will write it as $\mathbf{x} = \mathbf{x}^t$.

To define a suitable topology on Λ_T, note that we can extend a path segment $\mathbf{x}^t \in \Omega_t^{p,0}$ by its constant extrapolation $\mathbf{x}^t(\cdot \wedge t) \in \Omega_T^{p,0}$. We then equip Λ_T with the metric

$$d(\mathbf{x}^t, \mathbf{y}^s) := d_{p-\mathrm{var}}\left(\mathbf{x}^t(\cdot \wedge t), \mathbf{y}^s(\cdot \wedge s)\right) + |t - s|, \qquad \mathbf{x}^t, \mathbf{y}^s \in \Lambda_T.$$

Clearly, the topology on Λ_T is the initial topology of the two maps $\Lambda_T \to \Omega_T^{p,0}$, $\mathbf{x}^t \mapsto \mathbf{x}^t(\cdot \wedge t)$ and $\Lambda_T \to [0,T]$, $\mathbf{x}^t \mapsto t$. We further recall the following result from [BHRS23, Appendix A].

Proposition 3.1. Λ_T *is a Polish space. Furthermore, the topology on Λ_T is the final topology induced by the map $\varphi : [0,T] \times \Omega_T^{p,0} \to \Lambda_T$, $(t, \mathbf{x}) \mapsto \mathbf{x}|_{[0,t]}$.*

3.4. **Time augmention and universal approximation.** In order to use the signature as a feature on the path space, we extend paths by a running time component. This will guarantee that the signature uniquely characterizes the path on the entire interval. In this section we will explain how this extension works for geometric rough paths. We also explain the universality of linear signature maps; see Proposition 3.3.

Let $\mathbf{x} \in \Omega_T^{p,0}$ and let $(x_n)_{n \geq 1}$ be a sequence of piecewise smooth paths such that $\mathrm{Sig}(x_n)^{\leq \lfloor p \rfloor}$ converges to $\mathbf{x}$ in $\Omega_T^{p,0}$ We extend x_n to a piecewise smooth path $\hat{x}_n := t \mapsto (t, x_n(t))$. It then follows from [FV10, Theorem 9.32] that $(\mathrm{Sig}(\hat{x}_n)^{\leq \lfloor p \rfloor})_{n \geq 1}$ forms a Cauchy sequence in $\Omega_T^{p,0}(\mathbb{R} \times V)$ and we denote by $\hat{\mathbf{x}}$ its limit. Furthermore, we see from the same theorem that there exists a constant $C > 0$ depending on p and T such that

$$d_{p-\mathrm{var}}(\mathbf{x}, \mathbf{y}) \leq d_{p-\mathrm{var}}(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \leq C d_{p-\mathrm{var}}(\mathbf{x}, \mathbf{y}), \qquad \mathbf{x}, \mathbf{y} \in \Omega_T^{p,0}.$$

We define by $\hat{\Omega}_T^{p,0} := \{\hat{\mathbf{x}} \mid \mathbf{x} \in \Omega_T^{p,0}\} \subset \Omega_T^{p,0}(\mathbb{R} \times V)$ the set of *time augmented geometric p-rough paths.* The above estimate implies that the embedding $\mathbf{x} \mapsto \hat{\mathbf{x}}$ is continuous and that $\hat{\Omega}_T^{p,0}$ is also a Polish space. We extend the notions from the previous section to these time augmented paths by passing from the stopped rough path $\mathbf{x}^t \in \Lambda_T$ to its augmented version $\hat{\mathbf{x}}^t \in \hat{\Omega}_T^{p,0}$.

We will often use that signatures of time-augmentations depend continuously on the original rough path:

Lemma 3.2. *For any $N \in \mathbb{N}$, $\Lambda_T \to G^N(V) : \mathbf{x}^t \mapsto \mathrm{Sig}(\hat{\mathbf{x}}^t)_{0,t}^{\leq N}$ is continuous.*

Proof. For $\mathbf{x}^t \in \Lambda_T$ define the extrapolated path $\mathbf{x}^{t,T} := \mathbf{x}^t(t \wedge \cdot) \in \Omega_T^{p,0}$ and its time-augmentation by $\hat{\mathbf{x}}^{t,T} \in \hat{\Omega}_T^{p,0}$. We then have by Chen's identity

$$\begin{aligned}
\mathrm{Sig}(\hat{\mathbf{x}}^t)_{0,t} &= \mathrm{Sig}(\hat{\mathbf{x}}^{t,T})_{0,t} \\
&= \mathrm{Sig}(\hat{\mathbf{x}}^{t,T})_{0,T} \otimes (\mathrm{Sig}(\hat{\mathbf{x}}^{t,T})_{t,T})^{-1} \\
&= \mathrm{Sig}(\hat{\mathbf{x}}^{t,T})_{0,T} \otimes \exp_\otimes(-e_1(T-t))
\end{aligned}$$

The statement then follows from continuity of the maps $\mathbf{x}^t \mapsto (t, \mathbf{x}^{t,T})$, $\mathbf{x} \mapsto \hat{\mathbf{x}}$ and $\mathbf{x} \mapsto \mathrm{Sig}(\mathbf{x})_{0,T}^{\leq N}$ for any $N \in \mathbb{N}$ are continuous. $\qquad\square$

We are now ready to state the approximation property of linear signature functionals within the class of continuous non-anticipative functionals.

Proposition 3.3. *For any continuous function $f : \Lambda_T^p \to \mathbb{R}$ and compact set $K \subset \Omega_T^{p,0}$ there exists a sequence $(\ell_n)_{n \geq 1} \subset T(V)$ such that*

$$\lim_{n \to \infty} \sup_{t \in [0,T]} \sup_{\mathbf{x} \in K} \left(\left| f(\mathbf{x}|_{[0,t]}) - \langle \ell_n, \mathrm{Sig}(\hat{\mathbf{x}}|_{[0,t]}) \rangle \right| \right) = 0.$$

This form of universal approximation property goes back to [KLPA20]. Let us sketch its proof since it highlights key structural properties of signatures which underline their usefulness for encoding path information. From the algebraic properties of the signature, it follows that the linear functionals $\{\mathbf{x}^t \mapsto \langle \ell, \mathrm{Sig}(\hat{\mathbf{x}}^t)_{0,t} \rangle \mid \ell \in T(V)\}$ form an algebra (cf. [LCL07, Theorem 2.15]). By Lemma 3.2 it is a sub-algebra of $C(\Lambda_T; \mathbb{R})$. Since the signature characterizes the path (see [HL10], [BGLY16]), the algebra is point separating (see [KLPA20, Lemma B.3] for detailed proof). Finally, by Proposition 3.1, we have that $\varphi([0,T] \times K) \subset \Lambda_T$ is compact. The statement then follows from the Stone-Weierstrass Theorem.

3.5. **Rough differential equations.** Starting as in Section 3.2, we can give meaning to a rough differential equation

$$\text{(3.3)} \qquad \mathrm{d}y(t) = \sigma(y(t))\mathrm{d}\mathbf{x}(t), \quad t \in [0,T], \qquad y(0) = y_0 \in \mathbb{R}^m.$$

where $\mathbf{x} \in \Omega_T^{p,0}(\mathbb{R}^d)$ and $\sigma : \mathbb{R}^m \to \mathbb{R}^{m \times d}$, by first considering it for piecewise smooth paths. Indeed, let $x : [0,T] \to \mathbb{R}^d$ be piecewise smooth and consider the system of ordinary integral equations

$$\text{(3.4)} \qquad y^k(t) = y_0^k + \sum_{i=1}^d \int_0^t \sigma_i^k(y(s)) \dot{x}^i(s) \mathrm{d}s, \qquad k = 1, \ldots, m, \quad t \in [0,T], \quad y_0 \in \mathbb{R}^m.$$

Given that σ is bounded and Lipschitz continuous, the above equation has a unique solution, which we denote by $\Gamma_\sigma(y_0; x)$. We then say that a continuous path $y : [0,T] \to \mathbb{R}^m$ is a solution to the rough differential equation (3.3) if there exists a sequence of piecewise smooth paths $(x_n)_{n \geq 1}$ such that

(i) $\lim_{n \to \infty} d_{p-\mathrm{var}}(\mathbf{x}, \mathrm{Sig}(x_n)^{\leq N}) = 0$,
(ii) $\Gamma_\sigma(y_0; x_n) \xrightarrow[n \to \infty]{} y$ uniformly on $[0,T]$.

In order to state the main existence and uniqueness result for rough differential equation we introduce a convenient Lipschitz space. Let V and W be two Banach spaces. For $\gamma \in (0,1]$ and a map $f : V \to W$ we define

$$\|f\|_{\mathrm{Lip}^\gamma(V;W)} = \max \left\{ \sup_{u \in V} \|f(u)\|, \ \sup_{u,v \in V} \frac{\|f(v) - f(u)\|}{\|u-v\|^\gamma} \right\}$$

and, for $\gamma > 1$ and a $\lfloor\gamma\rfloor$-times Fréchet-differentiable map $f : V \to W$, we define recursively

$$\|f\|_{\mathrm{Lip}^\gamma(V;W)} = \max\left\{\sup_{u\in V}\|f(u)\|,\ \|f'\|_{\mathrm{Lip}^{\gamma-1}(V,L(V;W))}\right\},$$

where $L(V;W)$ denotes the Banach space of bounded linear functions from V to W (see [FV10, Appendix B] for more detail). In the following we will simply write $\|f\|_{\mathrm{Lip}^\gamma} = \|f\|_{\mathrm{Lip}^\gamma(V;W)}$ as the spaces V and W can be inferred from the context. Furthermore, if we require that $\|f\|_{\mathrm{Lip}^\gamma} < \infty$ then we implicitly require that f is $\lfloor\gamma\rfloor$-times differentiable and we set

$$\mathrm{Lip}^\gamma(V;W) := \{f : V \to W \mid \|f\|_{\mathrm{Lip}^\gamma} < \infty\}$$

Given that $\|\sigma\|_{\mathrm{Lip}^\gamma} < \infty$ for some $\gamma > p$ it holds that equation (3.3) has a unique solution $y \in C([0,T];\mathbb{R}^m)$. Moreover, y has a unique lift to a geometric rough path $\mathbf{y} \in \Omega_T^p(\mathbb{R}^m)$ such that for any sequence of piecewise smooth paths (x_n) satisfying (i) and $y_n := \Gamma_\sigma(y_0, x_n)$ it holds

$$\lim_{n\to\infty} d_{p-\mathrm{var}}(\mathbf{y}, \mathrm{Sig}(y_n)^{\leq\lfloor p\rfloor}) = 0.$$

Furthermore, the so called Itô-Lyons solution map

$$\mathbb{R}^m \times \Omega_T^{p,0}(\mathbb{R}^d) \times \mathrm{Lip}^\gamma(\mathbb{R}^m;\mathbb{R}^{m\times d}) \to \Omega_T^p(\mathbb{R}^m)$$
$$(y_0, \mathbf{x}, \sigma) \mapsto \mathbf{y}$$

is locally Lipschitz continuous.

4. Approximation with Signature Controls

We are now ready to provide the main results of this paper, namely that the general, rough and non-Markovian stochastic control problem introduced in Section 2, can be solved by controls, which are functions of the signature of the driving path. In order to do so, we first need a new, general uniform stability result for rough differential equations whose drift depends on a measurable control, see Section 4.1. We continue by showing that progressively measurable admissible controls can be approximated by continuous controls, i.e., continuous function on the space of stopped rough paths, see Section 4.2. In Section 4.3 we derive an appropriate universal approximation theorem (a.k.a. Stone-Weierstrass theorem), which allows us to conclude that signature controls (i.e., controls defined as linear functionals of the signatures or as neural networks applied to the log-signature) are dense in the set of all admissible controls. Finally, we prove that we can solve the original stochastic optimal control problem by restricting admissible controls to signature controls, in the sense that the infimum of the expected cost is equal, see Section 4.4.

4.1. **Drift-controlled rough differential equations.** In the following, we will consider rough differential equations with a controlled drift term

(4.1) $$dy(t) = b(y(t), u(t))dt + \sigma(y(t))d\mathbf{x}(t), \quad t \in [0,T], \qquad y(0) = y_0 \in \mathbb{R}^m,$$

where $u : [0,T] \to \mathcal{U} \subset \mathbb{R}^k$ is measurable and $b : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^m$. The following theorem establishes the existence, uniqueness and stability of solutions, which will be crucial for proving our main approximation result in Section 4.4. Substituting $\mathbf{x} = \mathbb{X}(\omega)$ and $u = U(\omega)$, the solution of (4.1) corresponds to the solution of the controlled stochastic RDE (2.1) for a given realization of the noise as well as the control.

Theorem 4.1. *Let $\mathbf{x} \in \Omega_T^{p,0}(\mathbb{R}^d)$ for some $p \in [1,\infty)$ and assume that*

(4.2) $$\sup_{a\in\mathcal{U}}\|b(\cdot,a)\|_{\mathrm{Lip}^1} < \infty \qquad and \qquad \|\sigma\|_{\mathrm{Lip}^{\gamma+1}} < \infty \quad for\ some\ \gamma > p.$$

Then equation (4.1) *has a unique solution* y *with corresponding lift* $\mathbf{y} \in \Omega_T^p(\mathbb{R}^m)$ *such that*

$$\mathbb{R}^m \times \Omega_T^{p,0}(\mathbb{R}^d): \quad (y_0, \mathbf{x}) \mapsto \mathbf{y},$$

is locally Lipschitz continuous uniformly in u. *Assuming further that*

$$(4.3) \qquad \sup_{a,a' \in \mathcal{U}} \frac{\|b(\cdot, a) - b(\cdot, a')\|_{\text{Lip}^1}}{\|a - a'\|} < \infty$$

and denoting by $L^0(\mathrm{d}t; \mathcal{U}) = \{u : [0,T] \to \mathcal{U} \text{ measurable}\}$ *equipped with the topology of convergence in Lebesgue-measure, then also the map*

$$\Gamma_{b,\sigma} : \mathbb{R}^m \times \Omega_T^{p,0}(\mathbb{R}^d) \times L^0(\mathrm{d}t, \mathcal{U}) \to \Omega_T^p(\mathbb{R}^m), \quad (y_0, \mathbf{x}, u) \mapsto \mathbf{y},$$

is continuous.

Proof. The proof uses the flow decomposition method. Let $\phi \colon [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$ denote the solution to

$$\phi(t, y_0) = y_0 + \int_0^t \sigma(\phi(s, y_0)) \, \mathrm{d}\mathbf{x}(s).$$

From [FV10, Proposition 11.11], we know that ϕ is a flow of $\mathcal{C}^2$-diffeomorphisms and that the first and second derivate of ϕ and its inverse are bounded by a constant depending only on $p, \gamma, \|\sigma\|_{\text{Lip}^{\gamma+1}}$ and the rough path norm of $\mathbf{x}$. From these properties and assumption (4.2), we can conclude that the ordinary differential equation

$$(4.4) \qquad Z_t^{y_0} = y_0 + \int_0^t (D_z\phi(s,z)|_{z=Z_s^{y_0}})^{-1} b(\phi(s, Z_s^{y_0}), u(s)) \, \mathrm{d}s$$

has a unique solution Z^{y_0} for every initial condition y_0. The solution to (4.1) is then given by

$$y(t) = \phi(t, Z_t^{y_0}).$$

We will now construct the rough path lift of y. By definition, there exists a sequence of smooth paths $(x_n)_{n \geq 1}$ s.t.

$$\lim_{n \to \infty} d_{p-\text{var}}\left(\mathbf{x}, \text{Sig}(x_n)^{\leq \lfloor p \rfloor}\right) = 0.$$

Let y_n denote the solution to (4.1) where we replace the driving rough path by $\text{Sig}(x_n)^{\leq \lfloor p \rfloor}$. Since the paths y_n are smooth, $\text{Sig}(y_n)^{\leq \lfloor p \rfloor}$ is well defined. It is not hard to prove that $(\exp_\otimes^{\lfloor p \rfloor}(y_0) \otimes \text{Sig}(y_n)^{\leq \lfloor p \rfloor})_{n \geq 1}$ constitutes a Cauchy sequence in the space $\Omega_T^p(\mathbb{R}^m)$ and that the limit $\mathbf{y}$ is a rough path lift of y that satisfies the stated properties. In fact, the components of $\mathbf{y}$ can be identified as rough integrals (cf. [FH20, Chapter 4]). To see this, we note first that y is controlled by the rough path $\mathbf{x}$ in the sense of Gubinelli (cf. [FH20, Definition 4.6 and Section 4.5]). The second iterated integral is then given by

$$\int_s^t (y(v) - y(s)) \otimes \mathrm{d}y(v) := \int_s^t (y(v) - y(s)) \otimes b(y(v), u(v)) \, \mathrm{d}v$$

$$+ \int_s^t (y(v) - y(s)) \otimes \sigma(y(v)) \, \mathrm{d}\mathbf{x}(v).$$

Since the integrand in the second integral is controlled by $\mathbf{x}$, the integral indeed exists as a rough integral. Furthermore, the rough integral is again controlled by $\mathbf{x}$, thus we can make sense of the third iterated integral of y similarly. By an induction argument, iterated integrals of y of arbitrary order can be defined and it is not hard to show that they coincide with the components of $\mathbf{y}$.

The rest of the proof follows by standard arguments. $\qquad \square$

Remark 4.2. In [DFG17, Theorem 29], a similar statement is formulated in the case of $p \in [2, 3)$ under the (slightly) weaker assumption $\|\sigma\|_{\mathrm{Lip}^\gamma} < \infty$. The reason why we need one additional degree of smoothness here is that we use the flow decomposition method for which also the second derivative of the flow ϕ plays a role in the estimates to ensure global existence and uniqueness of (4.4). The proof of [DFG17, Theorem 29] avoids the flow decomposition, but relies on a yet to be developed solution theory for rough differential equations on Banach spaces containing a drift parameter that is only Lipschitz continuous. With this it should be possible to get a version of Theorem 4.1 that only assumes $\|\sigma\|_{\mathrm{Lip}^\gamma} < \infty$.

4.2. **Approximation of admissible controls by non-anticipative continuous path functions.** We will start by lifting progressively measurable processes with measurable non-anticipative path functionals. The proof is mainly postponed to the Appendix A. Throughout this section we will assume that Assumption (2.3) holds, i.e., $\mathbb{X}$ takes values in $\Omega_T^{p,0}$ and generates the underlying filtration $(\mathcal{F}_t)$.

Proposition 4.3. *For any $(\mathcal{F}_t)$-progressively measurable process $U \colon [0,T] \times \Omega \to \mathbb{R}^k$ there exists a Borel measurable map $\theta \colon \Lambda_T \to \mathbb{R}^k$ such that for*

$$(4.5) \qquad\qquad \theta(\mathbb{X}(\omega)|_{[0,t]}) = U_t(\omega),$$

for $\mathrm{Leb} \otimes \mathbb{P}$-*a.e.* $(t, \omega) \in [0,T] \times \Omega$. *Conversely, for any Borel measurable $\theta \colon \Lambda_T^p \to \mathbb{R}^k$,*

$$(4.6) \qquad\qquad U_t^\theta := \theta(\mathbb{X}|_{[0,t]}), \qquad t \in [0,T],$$

defines an $(\mathcal{F}_t)$-progressively measurbale process.

Proof. Let U be a $(\mathcal{F}_t)$-progressively measurable process. Then $U : [0,T] \times \Omega \to \mathbb{R}^k$ is measurable with respect to $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$. From Lemma A.1 we have that $\mathcal{F}_T$ is the completion of $\mathbb{X}^{-1}(\mathcal{B}(\Omega_T^{p,0}))$. Hence, there exists a measurable map $\eta : [0,T] \times \Omega_T^{p,0} \to \mathbb{R}^k$ such that $U = \eta(\cdot, \mathbb{X})$ upto indistinguishability.

Now fix an arbitrary $t \in [0,T]$. Since U_t is $\mathcal{F}_t$-measurable it follows again by Lemma A.1 that there exists a map $\eta_t : \Omega_t^{p,0} \to \mathbb{R}^k$ such that almost surely $\eta(t, \mathbb{X}) = U_t = \eta_t(\mathbb{X}|_{[0,t]})$. This readily implies that almost surely $\eta(t, \mathbb{X}) = \eta(t, \mathbb{X}_{\cdot \wedge t})$. Finally we define $\theta : \Lambda_T^p \to \mathbb{R}^k$ by setting $\theta(\mathbb{X}|_{[0,t]}) := \eta(t, \mathbb{X}_{\cdot \wedge t})$ which satisfies (4.5). From the continuity of the extension map $\Lambda_T \to [0,T] \times \Omega_T^{p,0}$, $(t, \mathbf{x}^t) := (t, \mathbf{x}^t(\cdot \wedge t))$ (see Section 3.3) we also have that θ is measurable.

Conversely, if U is a process satisfying (4.6) for some Borel-measurable map $\theta : \Lambda_T \to \mathbb{R}^k$, then for any $t \in [0,T]$ and $B \in \mathcal{B}(\mathbb{R}^k)$ we have

$$\{(s, \omega) \in [0,t] \times \Omega \mid U_s(\omega) \in B\} = \{(s, \omega) \in [0,t] \times \Omega \mid \mathbb{X}(\omega)|_{[0,s]} \in \theta^{-1}(B)\}$$
$$= \{(s, \omega) \in [0,t] \times \Omega \mid \varphi(s, \mathbb{X}(\omega)|_{[0,t]}) \in \theta^{-1}(B)\},$$

where $\varphi : [0,t] \times \Omega_t^{p,0} \to \Lambda_t : (s, \mathbf{x}) \mapsto \mathbf{x}|_{[0,s]}$. By Proposition 3.1 the map φ is continuous. Since $\omega \mapsto \mathbb{X}(\omega)|_{[0,t]}$ is $\mathcal{F}_t$ measurable it then follows that the above set is in $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$. $\square$

Based on the above result, we are next going to show that admissible controls can be approximated with *continuous* non-anticipative path functionals, i.e., with controls in the following subclass

$$\mathcal{A}_c := \{U^\theta \mid \theta : \Lambda_T^p \to \mathbb{R} \text{ continuous.}\}$$

Corollary 4.4. *For any admissible control $U \in \mathcal{A}$ there exists a sequence (U^n) in $\mathcal{A}_c$ such that*

$$(4.7) \qquad\qquad U = \lim_{n \to \infty} U^n \qquad \mathrm{Leb} \otimes \mathbb{P}\text{-}almost\ everywhere.$$

We denote by $P_{\mathcal{U}} : E \to \mathcal{U}$ the projection map onto the closed convex set $\mathcal{U}$. Note that by the Hilbert projection theorem (c.f. [Rud87, Theorem 4.10]) this map is unique and continuous.

Proof. The continuity of $\varphi : [0, T] \times \Omega_T^{p,0} \to \Lambda_T^p :, (t, \mathbf{x}) \mapsto \mathbf{x}|_{[0,t]}$ implies that the push forward $\mu = \varphi(\cdot, \mathbb{X}(\cdot))_*[\mathrm{Leb} \otimes \mathbb{P}]$ is a well-defined probability measure on $(\Lambda_T^p, \mathcal{B}(\Lambda_T^p))$. From [Wiś94, p. 148] it follows that every Borel measurable map $\theta \colon \Lambda_T^p \to \mathbb{R}^k$ can be approximated μ-almost surely by continuous maps.

Now let $U \in \mathcal{A}$. By Proposition 4.3 there exits a Borel measurable $\theta \colon \Lambda_T^p \to \mathbb{R}^k$ such that (4.5) holds for $\mathrm{Leb} \otimes \mathbb{P}$-a.e. (t, ω). From the previous paragraph it follows that there exists a sequence of continuous maps $\theta_n : \Lambda_T^p \to \mathbb{R}^k$ for $n \geq 1$ that converges $(\mathrm{Leb} \otimes \mathbb{P})$-almost everywhere towards θ. This implies that

$$U_t(\omega) = \theta(\mathbb{X}|_{[0,t]}(\omega)) = \lim_{n \to \infty} \theta_n(\mathbb{X}|_{[0,t]}(\omega)),$$

for $\mathrm{Leb} \otimes \mathbb{P}$-a.e. $(t, \omega) \in [0, T] \times \Omega$. Finally, composing with the projection map $U^n := P_{\mathcal{U}} \circ \widetilde{\theta}_n \circ \mathbb{X}$ and recalling that $P_{\mathcal{U}}$ is continuous, the claim follows. $\qquad \square$

4.3. **Linear and deep signature controls.** The universal approximation theorem for signatures allows us to approximate any continuous functional on the path space by linear functionals of the signature. This motivates us to consider the following class of *linear signature functionals*

$$\mathcal{T}_{\mathrm{lin}} := \left\{ \theta : \Lambda^T \to E \;\middle|\; \exists \ell_1, \ldots, \ell_k \in T((\mathbb{R}^{d+1})^*) \text{ s.t. } \theta^i(\hat{\mathbf{x}}^t) = \langle \ell_i, \mathrm{Sig}(\hat{\mathbf{x}}^t)_{0,t} \rangle \;\; \forall \, \mathbf{x}^t \in \Lambda_T^p \right\}.$$

In general such functionals $\theta \in \mathcal{T}_{\mathrm{lin}}$ will not lead to admissible strategies U^θ. Recall that $P_{\mathcal{U}}$ is the projection from E onto the convex set $\mathcal{U}$. We then define the set of *linear signature controls*

$$\mathcal{A}_{\mathrm{lin}} := \left\{ U^\theta \;\middle|\; \exists \theta' \in \mathcal{T}_{\mathrm{lin}} \text{ s.t. } \theta = P_{\mathcal{U}} \circ \theta' \right\}.$$

We will argue below that $\mathcal{T}_{\mathrm{lin}} \subset C(\Lambda_T^p; E)$, thus verifying that indeed $\mathcal{A}_{\mathrm{lin}} \subset \mathcal{A}_c \subset \mathcal{A}$. Before doing so we will, however, introduce a second class of signature strategies, based on non-linear functionals. Motivated by the numerical efficiency for optimal stopping problems in [BHRS23], we define the following class of deep neural functionals of the log signature

$$\mathcal{T}_{\mathrm{DNN}} := \left\{ \theta : \Lambda^T \to E \;\middle|\; \exists N \in \mathbb{N}, \; F \in \mathcal{D}^{\eta_{d+1,N},k} \text{ s.t. } \theta(\hat{\mathbf{x}}^t) = F \circ \log_\otimes(\mathrm{Sig}(\hat{\mathbf{x}}^t)_{0,t}^{\leq N}) \;\; \forall \, \mathbf{x}^t \in \Lambda_T^p \right\},$$

where $\log_\otimes : G^N(\mathbb{R}^{d+1}) \to \mathfrak{g}^N(\mathbb{R}^{d+1}) \cong \mathbb{R}^{\eta_{d+1,N}}$ is the truncated tensor logarithm (see Section 3.1), $\eta_{d+1,N}$ is the dimension of the truncated log signature[2] and $\mathcal{D}^{l,k}$ is the class of deep neural networks mapping from $\mathbb{R}^l$ to $\mathbb{R}^k$ with fixed depth $I \geq 1$, number of neurons $q \geq 1$ and activation function[3] $\varphi : \mathbb{R}^q \to \mathbb{R}^q$. More precisely, $\mathcal{D}^{l,k}$ consists of functions of the form

$$F = A_0 \circ \varphi \circ A_1 \circ \varphi \circ \cdots \circ A_I,$$

where $A_I : \mathbb{R}^l \to \mathbb{R}^q$, $A_0 : \mathbb{R}^q \to \mathbb{R}^k$ and $A_i : \mathbb{R}^q \to \mathbb{R}^q$ $(0 < i < I)$ are linear maps. Similar to the linear case, we define the set of *deep signature controls* by

$$\mathcal{A}_{\mathrm{DNN}} := \left\{ U^\theta \;\middle|\; \exists \theta' \in \mathcal{T}_{\mathrm{DNN}} \text{ s.t. } \theta = P_{\mathcal{U}} \circ \theta' \right\}.$$

The following verifies that the signature controls are dense in the set of admissible strategies.

---

[2]We refer to [BHRS23] Section 7 for more detail on the log-signature and the dimension of the step-N nilpotent free Lie-algebra $\mathfrak{g}^N(\mathbb{R}^{d+1})$. Note also that for simplicity we identify $\mathfrak{g}^N(\mathbb{R}^{d+1})$ with $\mathbb{R}^{\eta_N,k}$ in the definition of $\mathcal{T}_{\mathrm{DNN}}$ without formally introducing an isomorphism.

[3]An activation function is a continuous function that is not a polynomial. For example the ReLu function $\varphi(x) = \max\{0, x\}$ (componentwise).

Proposition 4.5. *It holds that $\mathcal{T}_{\mathrm{lin}}, \mathcal{T}_{\mathrm{DNN}} \subset C(\Lambda^T; \mathbb{R}^k)$. Furthermore, given assumption (2.3), it holds that $\mathcal{A}_{\mathrm{lin}}, \mathcal{A}_{\mathrm{DNN}} \subset \mathcal{A}_c$ and for any admissible control $U \in \mathcal{A}$ there exists a sequence $(U^n)_{n \geq 1} \subset \mathcal{A}_{\mathrm{lin}}$ (respectively in $\mathcal{A}_{\mathrm{DNN}}$) such that*

$$U = \lim_{n \to \infty} U^n, \qquad \mathrm{Leb} \otimes \mathbb{P}\text{-almost everywhere.}$$

Based on Corollary 4.4 the proof is similar to the Proposition 7.4 in [BHRS23].

Proof. Recall from Lemma 3.2 that the map $\mathbf{x}^t \mapsto \mathrm{Sig}(\hat{\mathbf{x}}^t)_{0,T}^{\leq N}$ is continuous for any $N \in \mathbb{N}$. Since also the linear map $\langle \ell, \cdot \rangle : T((\mathbb{R}^{d+1})) \to \mathbb{R}$ is continuous for any $\ell \in T(\mathbb{R}^{d+1})$ we readily conclude that $\mathcal{T}_{\mathrm{lin}} \subset C(\Lambda_T; \mathbb{R}^k)$. Similarly, since any $F \in \mathcal{D}^{\eta_N, k}$ is a continuous function and recalling from Section 3.1 that also the truncated logarithm $\log_{\otimes} : G^N(\mathbb{R}^{d+1}) \to \mathfrak{g}^N(\mathbb{R}^{d+1})$ is continuous it follows similarly that $\mathcal{T}_{\mathrm{DNN}} \subset C(\Lambda_T; \mathbb{R}^k)$. The fact that $\mathcal{A}_{\mathrm{lin}}, \mathcal{A}_{\mathrm{DNN}} \subset \mathcal{A}_c$ then follows immediately by the continuity of the projection $P_{\mathcal{U}}$.

To prove the approximating result it suffices by Corollary 4.4 to show that for any $\theta \in C(\Lambda_T; \mathbb{R}^k)$ there exits a sequence $(\theta_n)_{n \geq 1} \subset \mathcal{T}_{\mathrm{lin}}$ (respectively $\mathcal{T}_{\mathrm{DNN}}$), such that for a.e. $\omega \in \Omega$ it holds

$$(4.8) \qquad \theta(\mathbb{X}|_{[0,t]}(\omega)) = \lim_{n \to \infty} \theta_n(\mathbb{X}|_{[0,t]}(\omega)), \qquad \text{for Leb-a.e. } t \in [0, T].$$

To this end, since $\Omega_T^{p,0}$ is a Polish space, we can choose an increasing sequence of compact sets $K_n \subset \Omega_T^{p,0}$ such that $\lim_{n \to \infty} \mathbb{P}(\mathbb{X} \in K_n) = 1$. By Proposition 3.3 for each $n \in \mathbb{N}$ there exits a sequence $(\ell_{n,j})_{j \geq 1} \subset T(\mathbb{R}^{d+1})^{\otimes k}$ such that for $j_n \in \mathbb{N}$ large enough it holds

$$\sup_{t \in [0,T]} \sup_{\mathbf{x} \in K_n} \max_{i=1,\ldots,k} \left(\left| \theta^i(\mathbf{x}|_{[0,t]}) - \langle \ell_j^i, \mathrm{Sig}(\hat{\mathbf{x}}|_{[0,t]}) \rangle \right| \right) \leq \frac{1}{n}, \qquad j \geq j_n.$$

Defining the sequence $(\theta_n)_{n \geq 1}$ by $\theta_n^i(\mathbf{x}^t) := \langle \ell_{j_n}^i, \mathrm{Sig}(\hat{\mathbf{x}}^t)_{0,t} \rangle$ for all $i = 1, \ldots, k$, we readily see that it satisfies (4.8). To prove the approximation result for $\mathcal{A}_{\mathrm{DNN}}$ it now suffices to show that $\mathcal{T}_{\mathrm{DNN}}$ is suitably dense in $\mathcal{T}_{\mathrm{lin}}$. More precisely, it suffices to show that for any compact set $K \subset \Omega_T^{p,0}$, $N \in \mathbb{N}$ and $\ell \in T^N(\mathbb{R}^{d+1})$ there exists a sequence of functions $(F^n)_{n \geq 1} \in \mathcal{D}^{\eta_{d+1,N}, 1}$ such that

$$\lim_{n \to \infty} \sup_{\mathbf{x} \in K} |\langle \ell, \mathrm{Sig}(\hat{\mathbf{x}})_{0,T} \rangle - F^n(\log_{\otimes} \mathrm{Sig}(\hat{\mathbf{x}})_{0,T})| = 0.$$

Recall from Section 3.1 that the map $\log_{\otimes} : G^N(\mathbb{R}^{d+1}) \to \mathfrak{g}^N(\mathbb{R}^{d+1})$ is continuous and invertable with inverse $\exp_{\otimes}$. Hence, in particular the set $\{\log_{\otimes} \mathrm{Sig}(\hat{\mathbf{x}})_{0,T} \mid \mathbf{x} \in K\} \subset \mathfrak{g}^N(\mathbb{R}^{d+1})$ is compact. Since also the map $\mathfrak{g}^N(\mathbb{R}^{d+1}) \to \mathbb{R} : \mathbf{z} \mapsto \langle \ell, \exp_{\otimes}(\mathbf{z}) \rangle$ is continuous, the statement now readily follows from the universal approximation theorem for nerual networks (see e.g. [LLPS93, Theorem 1]). $\qquad \square$

4.4. **Approximation of the optimal costs.** In this section we present our main theoretical result, which states that the optimal expected costs of the problem introduced in Section 2 can be approximated using the classes of signature controls $\mathcal{A}_{\mathrm{lin}}$ and $\mathcal{A}_{\mathrm{DNN}}$. We begin by proving the continuous dependence of the costs with respect to the underlying control. Given the continuity of the cost functional L, this is a direct consequence of the stability result in Theorem 4.1.

Lemma 4.6. *Assume that (4.2)–(2.2) hold and let $(U^n)_{n \geq 1} \subset \mathcal{A}$ and $U \in \mathcal{A}$ s.t.*

$$U = \lim_{n \to \infty} U^n, \qquad \mathrm{Leb} \otimes \mathbb{P}\text{-almost everywhere.}$$

Denote by $\mathbb{Y}^n = \mathbb{Y}^{U^n}$ respectively $\mathbb{Y} = \mathbb{Y}^U$ the solution to the rough differential equation (2.1) with the control U^n respectively U. Then it holds

$$L(\mathbb{Y}, U) = \lim_{n \to \infty} L(\mathbb{Y}^n, U^n), \qquad \text{almost surely.}$$

Proof. We have in particular that $U^n(\omega)$ converges towards $U(\omega)$ in $L^0(\mathrm{d}t, \mathcal{U})$ for a.e. $\omega \in \Omega$. From Theorem 4.1 it follows that

$$\lim_{n \to \infty} d_{p-\mathrm{var}}(\mathbb{Y}^n(\omega), \mathbb{Y}(\omega)) = 0, \qquad \text{for a.e. } \omega \in \Omega.$$

We readily conclude using the continuity assumptions on L. $\qquad\square$

Combining Corollary 4.4 and Proposition 4.5 with the above lemma, the following result is now a direct consequence of the dominated convergence theorem.

Theorem 4.7. *Given assumptions* (4.2)–(2.3), *we have*

$$\inf_{U \in \mathcal{A}} J(U) \;=\; \inf_{U \in \mathcal{A}_c} J(U) \;=\; \inf_{U \in \mathcal{A}_{\mathrm{lin}}} J(U) \;=\; \inf_{U \in \mathcal{A}_{\mathrm{DNN}}} J(U).$$

5. Numerical Method and Examples

In this section, we will introduce a numerical method that arises from the parametrization of admissible controls in Section 4.2. We will then evaluate its performance through two case studies of non-Markovian control problems.

The full code is available at `https://github.com/hagerpa/sigControl`.

5.1. **Numerical Method.** Fixing a signature truncation level $N \geq 1$, the corresponding sub-classes $\mathcal{A}_{\mathrm{lin}}^N \subset \mathcal{A}_{\mathrm{lin}}$ and $\mathcal{A}_{\mathrm{DNN}}^N \subset \mathcal{A}_{\mathrm{DNN}}$ constitute finite-dimensional parameterizations of admissible controls. Indeed, for $\mathcal{A}_{\mathrm{lin}}^N$ the set of parameters is given by the coefficients in the truncated tensor series $\ell_1, \ldots, \ell_k \in T^N((\mathbb{R}^{d+1})^*)$, i.e., the dimension of the parameter space is given by $k \cdot (1 + (d+1) + \cdots (d+1)^N)$. For $\mathcal{A}_{\mathrm{DNN}}^N$, the parameter space is determined by the architecture of the deep neural networks $\mathcal{D}^{\eta_{d+1}, N, k}$, thus characterized by the number of hidden layers I and the number of neurons per layer q. Let us reveal the parametrization in both cases more explicitly by noting that for any $U \in \mathcal{A}_{\mathrm{lin}}^N$ (resp. $\mathcal{A}_{\mathrm{DNN}}^N$) there exists a vector of parameters η that defines a function $\theta(\cdot\,;\eta) : G^N(\mathbb{R}^{d+1}) \to \mathcal{U}$ such that $U_t = \theta(\mathrm{Sig}(\hat{\mathbb{X}})_{0,t}^{\leq N}; \eta)$. For ease of notation we then also write $U(\eta) \in \mathcal{A}_{\mathrm{lin}}^N$ (resp. $\mathcal{A}_{\mathrm{DNN}}^N$).

The next best objective towards solving the control problem is then to minimize $J(U(\eta))$ over η. This is achieved numerically through a Monte Carlo approximation of the expectation and a time discretization for approximating the signature and solving the rough differential equation. We will first present the general structure of the algorithm and then comment on the details below. For the sake of concreteness, we will assume that the cost functional L is of the integral form[4] (2.6).

(1) Fix a time grid $\Pi = \{t_0, t_1, \ldots, t_n\}$ with $0 = t_0 < t_1 < \cdots < t_n = T$.

(2) Generate M independent realizations $\{(\delta \mathbf{S}_j^{(i)})_{j=1,\ldots,n} \mid i = 1, \ldots, M\}$ of $(\mathrm{Sig}(\hat{\mathbb{X}})_{t_{j-1}, t_j}^{\leq N})_{t_j \in \Pi}$; then set $\mathbf{S}_0^{(i)} = \mathbf{1}$ and iteratively $\mathbf{S}_j^{(i)} = \mathbf{S}_{j-1}^{(i)} \otimes \delta \mathbf{S}_j^{(i)}$.

(3) Evaluate the signature controls $U_j^{(i);\eta} = \theta(\mathbf{S}_j^{(i)}; \eta)$ and store the gradients $\nabla_\eta \theta(\mathbf{S}_j^{(i)}; \eta)$, $j = 1, \ldots, n$.

(4) Calculate or use a suitable numerical scheme to find (approximate) solutions $\{(Y_j^{(i);\eta})_{j=0,\ldots,n} \mid i = 1, \ldots, M\}$ on the grid Π of the rough differential equation (2.4) corresponding to the samples $(\delta \mathbf{S}_j^{(i)})$.

---

[4]In general, the discretization of the cost functional depends on its specific form, while under the given assumption, a straightforward discretization is given by the Riemann sum.

(5) Calculate the approximate expected costs

$$(5.1) \qquad \frac{1}{M} \sum_{i=1}^{M} \left\{ \sum_{t_j \in \Pi} f(t_j, Y_j^{(i);\eta}, U_j^{(i);\eta})(t_j - t_{j-1}) + g(Y_n^{(i);\eta}) \right\}$$

and its gradients with respect to η using the previously stored gradients of the control. Then updated η using a step of stochastic gradient descent or a similar method.

(6) Starting from 2., repeat the above procedure with the new parameters several times or until no further improvement.

(7) For a larger number of samples $M' \gg M$ follow the steps 2. - 5. above once again and use (5.1) as an approximation to the optimal control value.

Calculation of the signature. In most cases of interest, we are confined to approximate realizations of the signature $\mathrm{Sig}(\hat{\mathbb{X}})_{t_{j-1},t_j}^{\leq N}$ in step (2) since direct sampling is not possible. Assume that $\mathbb{X}$ is the limit of the lifted piecewise linear approximation of the process $X = \pi_1(\mathbb{X})$, from which we can generate samples[5]. As discussed in Section 3.2, this holds true, for example, in the case of the Stratonovich-lift of a Brownian motion. We can then approximate realizations of the signature by sampling X on a refined grid of the interval $[t_{j-1}, t_j]$ and calculating the signature of the linearly interpolated path. Here a refinement can be beneficial for the optimization procedure, since further information from the evolution of X between t_{j-1} and t_j can be incorporated without the need to re-evaluate the control. The signature of a piecewise linear path can be calculated exactly, and implementations of the underlying algebraic structure are readily accessible, e.g., in the *iisignature* package [RG20]. This package also provides the necessary functionality to join the signatures over consecutive time intervals. We also note that the log-signature of a piecewise linear path can be calculated directly using the Baker–Campbell–Hausdorff formula, and the iisignature package also provides this functionality.

Numerical schemes. Similarly to the signature, calculating the RDE solutions explicitly is mostly not an option. Instead, one can use various numerical schemes to calculate approximate solutions. Having already calculated approximations of the signature increments $(\delta \mathbf{S}_j^{(i)})$ in the previous step, it is natural to use a higher-order Euler (or Taylor-type) scheme given by

$$(5.2) \qquad Y_j^{(i);\eta} = Y_{j-1}^{(i);\eta} + b(Y_{j-1}^{(i);\eta}, U_{j-1}^{(i);\eta})(t_j - t_{j-1}) + \mathcal{E}_{(\sigma)}(Y_{j-1}^{(i);\eta}, \delta \mathbf{S}_j^{(i)}),$$

for all $j = 1, \ldots, n$, starting with $Y_0^{(i);\eta} = y_0$, where

$$\mathcal{E}_{(\sigma)}(y, \delta \mathbf{S}_j^{(i)}) = \sum_{k=1}^{N} \sum_{i_1,\ldots,i_k=1}^{d} \sigma_{[i_1]}(y) \cdots \sigma_{[i_k]}(y) \mathrm{Id}(y) \langle e_{i_1 \ldots i_k}, \delta \mathbf{S}_j \rangle$$

$$= \sigma(Y_{j-1}^{(i);\eta}) \pi_1(\delta \mathbf{S}_j) + \sum_{i_1,i_2=1}^{d} \sum_{l=1}^{m} \left[\sigma_{i_1}^l \frac{\partial}{\partial y_l} \sigma_{i_2} \right](Y_{j-1}^{(i);\eta}) \langle e_{i_1 i_2}, \delta \mathbf{S}_j \rangle + \ldots$$

with $\sigma_{[i]} := \sum_{l=1}^{m} \sigma_i^l \frac{\partial}{\partial y_l}$ and the identity map $\mathrm{Id} : y \mapsto y$. Note that one recovers the usual Euler scheme for $N = 1$ and that the choice $N = 2$ corresponds to the Milstein scheme. In general, to guarantee convergence, one needs to choose $N \geq \lfloor p \rfloor$. This class of schemes were analyzed in [FV10] first. If no refinement of the intervals $[t_{j-1}, t_j]$ is used when calculating the signature increments, we have $\delta \mathbf{S}_j^{(i)} = \exp_{\otimes}^{N}(X_{t_j}^{(i)} - X_{t_{j-1}}^{(i)})$. Such *simplified* higher-order Euler schemes and their convergence rates were analyzed e.g. in [DNT12, FR14, BFRS16]. In general, higher-order Euler schemes are easy to implement, but they require the calculation of higher order derivatives,

---

[5]If direct sampling from the rough path $\mathbb{X}$ is possible, we use piecewise geodesic interpolation instead. However, it's worth noting that even for a Brownian motion, direct sampling from the rough path is not possible.

too, which can be costly in high dimensions. Moreover, they have a bad performance in case the underlying RDE is stiff, in which case implicit schemes typically perform better. In [RR22], a class of simplified Runge-Kutta schemes for RDEs was introduced and analyzed that contains derivative-free and implementable schemes including implicit ones.

Numerical optimization procedure. The calculation of gradients in step (3) and (5) is automatized when using standard machine learning software libraries such as *PyTorch* [PGM+19]. Such packages also provide state of the art variants of stochastic gradient descent. In the numerical examples below the performance was rather insensitive to the specific choice of optimization method and the choice of hyper-parameters. In this case the "Adam" method [KB14] lead to sufficient accuracy.

Linearization of the control problem. In special cases the algebraic properties of the signature can further be employed to transform the optimization problem into a form that allows to efficiently use non-stochastic solvers. This approach was first suggested in [KLPA20] in the context of an optimal execution problem (see also Section 5.3 below). Without going into details, this approach is generally possible if the the system (2.1) can be solved by integration, i.e., when $b(y, u)$ and $\sigma(y)$ are independent of y, and when b and L are of a polynomial form. Given sufficient integrability of $\mathrm{Sig}(\hat{\mathbb{X}})_{0,T}$, the shuffle property of the signature then allows to rewrite the expected costs associated to a signature control $U^\ell = \langle \ell, \mathrm{Sig}(\hat{\mathbb{X}})_{0,\cdot} \rangle$ for some $\ell \in T^N(\mathbb{R}^{d+1})$ as

$$J(U^\ell) = \langle P(\ell), \mathbb{E}[\mathrm{Sig}(\hat{\mathbb{X}})_{0,T}^{\leq N'}] \rangle,$$

where $P : T^N(\mathbb{R}^{d+1}) \to T^{N'}(\mathbb{R}^{d+1})$ with $N' \geq N$ is polynomial in the coefficients of $\ell \in T^N(\mathbb{R}^{d+1})$. Estimating the truncated expected signature using a Monte-Carlo average an optimal ℓ can then be obtain from a deterministic solver for polynomial optimization. As observed in [KLPA20], this reformulation of the control problem preserves quadratic convex structures, and thus allows to solve problems admitting such a structure as quadratic programs.

5.2. **Case study 1: Optimal tracking of fractional Brownian motion.** As a first benchmark example we consider the problem of optimally tracking a fractional Brownian motion with a process whose speed can be controlled at quadratic costs. Letting the Hurst parameter H vary in $(0, 1]$ allows us to test our numerical method in a range of cases outside the Markov and semimartingale regimes typically considered in the literature.

To set the stage, let ξ be a one-dimensional fractional Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and put $X := \xi$ so that ξ is trivially adapted to the augmented filtration generated by $(\sigma(X_s; s \in [0, t]))_{t \geq 0}$. Notice that this way the controller will *not* have access to the full past before time 0 of the fractional Brownian motion; granting this extended access leads to a different, yet equally relevant control problem, which, however, is computationally more demanding and thus left for future experiments at this point.

Now define the controlled process

$$(5.3) \qquad\qquad Y_t^U = y_0 + \int_0^t U_s \mathrm{d}s - \xi_t,$$

where U is a progressively measurable process that represents the speed and direction of the tracking. Since a one-dimensional fractional Brownian motion allows for a rough path lift

$$\mathbb{X}_t = \exp_\otimes^{\lfloor p \rfloor}(X_t), \quad \text{for some } p \in (1/H, 1 + 1/H),$$

this setting fits into our theoretical framework (2.1) by choosing $\mathcal{U} = \mathbb{R}$, $b(y, u) = u$, $\sigma = -1$. We measure the tracking performance by the cost functional

$$(5.4) \qquad L(\mathbb{Y}^U, U) = \frac{1}{2} \int_0^T \left((Y_t^U)^2 + \kappa (U_t)^2 \right) dt$$

where $\kappa > 0$ is the penalization parameter; we also confine ourselves to the natural class of progressively measurable control processes $U \in L^2(\mathbb{P} \otimes dt)$.

Remark 5.1. This kind of tracking problem finds several applications in mathematical finance. For instance, following [BSV17], we can consider ξ to be the hedging strategy of a contingent claim in an idealized frictionless reference market model driven by X; $y_0 + \int_0^\cdot U_s ds$ could be the evolution of a trader's actual hedging position when she is confronted with market impact costs as captured by $\kappa \int_0^T U_t^2 dt$. Due to these costs, the trader has to allow for a nonzero hedging error Y^U and the resulting risk can be measured by $\mathbb{E}\left[\int_0^T \left(\frac{1}{2} (Y_t^U)^2 \right) dt \right]$. Combining this with expected impact costs leads to (5.4).

Our choice of this particular benchmark problem is motivated by its analytic tractability: [BSV17] describes control policies and the problem value even for general targets ξ in. For our particular tracking problem with fractional Brownian motion, this gives:

Theorem 5.2. *The minimal tracking costs for a fractional Brownian motion ξ with Hurst parameter $H \in (0, 1)$ are*

$$\inf_U \mathbb{E}[L(Y^U, U)] = \frac{1}{2} \sqrt{\kappa} \tanh(\tau^\kappa(0)) (y_0)^2$$

$$(5.5) \qquad + \frac{1}{2} \int_0^T \int_0^t \left(\int_s^T (z_H(t, s) - z_H(u, s)) \frac{\cosh(\tau^\kappa(u))}{\sqrt{\kappa} \sinh(\tau^\kappa(t))} du \right)^2 ds dt$$

$$+ \frac{1}{2} \int_0^T \sqrt{\kappa} \tanh(\tau^\kappa(t)) \frac{\left(\int_t^T z_H(u, t) \cosh(\tau^\kappa(u)) du \right)^2}{\kappa \sinh^2(\tau^\kappa(t))} dt < \infty.$$

where, for $0 \leq s \leq t \leq T$, we let $\tau^\kappa(t) := (T - t)/\sqrt{\kappa}$, $0 \leq t \leq T$, and

$$z_H(t, s) := c_H \left(\frac{t}{s} \right)^{H - \frac{1}{2}} (t - s)^{H - 1/2} - (H - \frac{1}{2}) s^{\frac{1}{2} - H} \int_s^t u^{H - \frac{3}{2}} (u - s)^{H - \frac{1}{2}} du,$$

with

$$c_H := \left(\frac{2H \Gamma(\frac{3}{2} - H)}{\Gamma(H + \frac{1}{2}) \Gamma(2 - 2H)} \right)^{\frac{1}{2}}.$$

Proof. The general form of the optimal tracking strategy is given in Theorem 1 of [BSV17]. In conjunction with this, Theorem 3.4 of [BV18] identifies the minimal tracking error as

$$\inf_U \mathbb{E}[L(Y^U, U)] = \frac{1}{2} \sqrt{\kappa} \tanh(\tau^\kappa(0))(x - \hat{\xi}_0)^2 + \frac{1}{2} \mathbb{E}\left[\int_0^T (\xi_t - \hat{\xi}_t)^2 dt \right]$$

$$+ \frac{1}{2} \mathbb{E}\left[\int_0^T \sqrt{\kappa} \tanh(\tau^\kappa(t)) d\langle \hat{\xi} \rangle_t \right]$$

The tracking process $\hat{\xi}$ introduced there takes here the form

$$\hat{\xi}_t := \int_t^T \mathbb{E}\left[\xi_u | \mathcal{F}_t \right] \frac{\cosh(\tau^\kappa(u))}{\sqrt{\kappa} \sinh(\tau^\kappa(t))} du, \quad 0 \leq t \leq T.$$

For $\mathcal{F}_t$ generated by $\sigma(\xi_s, s \in [0, t])$, the conditional expectation $\mathbb{E}[\xi_u | \mathcal{F}_t]$ is computed in Theorem 4.2 in [NP00]. With z_H as above we obtain

$$\mathbb{E}[\xi_u | \mathcal{F}_t] = \int_0^t z_H(u, s) dB_s$$

where

$$B_t := \frac{2H}{c_H} \int_0^t s^{H-\frac{1}{2}} dM_s^H, \quad t \geq 0$$

is the standard Brownian motion constructed from the Gaussian martingale obtained from the fBM ξ via $M_t^H := \int_0^t w_H(t, s) d\xi_s$, $t \geq 0$, with

$$w_H(t, s) := \frac{s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}}{2H \int_0^1 u^{\frac{3}{2}-H}(1-u)^{\frac{1}{2}-H} du}, \quad s < t.$$

Straightforward computations starting from this show that

$$\mathbb{E}\left[\int_0^T (\xi_t - \hat{\xi}_t)^2 dt\right] = \int_0^T \int_0^t \left(\int_s^T (z_H(t, s) - z_H(u, s)) K(t, u) du\right)^2 ds dt$$

and

$$\langle \hat{\xi} \rangle_t = \int_0^t \frac{\left(\int_s^T z_H(u, s) \cosh(\tau^\kappa(u)) du\right)^2}{\kappa \sinh^2(\tau^\kappa(s))} ds, \quad t \in [0, T].$$

Our formula for the minimal tracking error follows now by plugging the identities from the last two displays into the general cost formula stated above. $\qquad\square$

	H	1/16	1/8	1/4	1/2	3/4	1
th. optimum		0.293	0.264	0.206	0.124	0.071	0.034
	$N = 1$	0.329	0.286	0.223	0.146	0.101	0.073
	$N = 2$	0.315	0.275	0.211	0.127	0.076	0.041
$\mathcal{A}_{\mathrm{lin}}$	$N = 3$	0.310	0.273	0.210	**0.124**	0.073	0.038
	$N = 4$	0.304	0.270	0.209	**0.124**	0.073	0.038
	$N = 5$	0.305	0.270	0.209	**0.124**	0.073	0.038
	$N = 1$	0.315	0.276	0.214	0.135	0.083	**0.034**
	$N = 2$	0.307	0.272	0.210	**0.124**	0.073	**0.034**
$\mathcal{A}_{\mathrm{DNN}}$	$N = 3$	0.301	0.269	0.209	**0.124**	**0.072**	**0.034**
	$N = 4$	**0.300**	**0.267**	**0.208**	**0.124**	**0.072**	**0.034**
	$N = 5$	**0.300**	**0.267**	0.209	**0.124**	**0.072**	0.035

TABLE 1. Numerical results for the optimal tracking of a fractional Brownian motion with different Hurst parameters H using strategies in $\mathcal{A}_{\mathrm{lin}}$ and $\mathcal{A}_{\mathrm{DNN}}$ ($I = 2$, $q = 30 + \eta_{2,N}$), with various signature truncation levels N. Presented are the estimated expected costs. The fixed model parameters are $y_0 = 0$, $T = 1$, and $\kappa = 0.1$. An overall time discretization of $\Delta t = 10^{-3}$ was used for the calculation of signatures, Y^U, and the cost functional. The number of training and testing paths was 2^{19} and 2^{20}, respectively. The Monte Carlo resampling error was below 0.0002. The first row presents the continuous-time optimal values calculated using (5.5).

We have tested our approximation method against the optimal values obtained from Theorem 5.2 for several choices of Hurst parameters H using strategies from $\mathcal{A}_{\text{lin}}$ and $\mathcal{A}_{\text{DNN}}$ with different signature truncation levels N. The remaining parameters were fixed to $y_0 = 0$, $T = 1$ and $\kappa = 0.1$. The outcomes of these numerical test are collected in Table 1. We notice that using our method we can reach reasonably accurate approximations of the continuous time optimum. Overall the performance is improved when increasing the truncation level and when going from linear strategies in $\mathcal{A}_{\text{lin}}$ to deep strategies in $\mathcal{A}_{\text{DNN}}$. It turns out that already beyond truncation levels $N = 3$ or $N = 4$ there is no further improvement and for cases $H \geq 1/2$ even level $N = 2$ seems to be sufficient.

Let us comment on a few more aspects in detail:

Due to the Markovianity in the Brownian monition case $H = 1/2$, there exists an optimal control in feedback form $U_t^* = \alpha(t, X_t, Y_t^{U^*})$ for a suitable deterministic function α (cf. [BSV17]). It turns out that to learn this control in open loop form $U_t^* = \theta(X|_{[0,t]})$ we only need level $N = 2$ log-signatures in $\mathcal{A}_{\text{DNN}}$ and level $N = 3$ signatures in $\mathcal{A}_{\text{lin}}$ to achieve sufficient accuracy. The second special case is $H = 1.0$, where the trajectories of X are straight lines starting at the origin with a standard normal distributed slope X_1 which becomes known right after the start. In this case the optimal solution is of the form $U_t^* = \mathbf{1}_{\{t>0\}} X_1 f(t) = \mathbf{1}_{\{t>0\}} \frac{X_t}{t} f(t)$, where $f : [0, T] \to \mathbb{R}$ is a suitable nonlinear deterministic function (cf. [BSV17]). While with deep neural networks this control can easily be learned as a function of the first level signature (t, X_t), the linear model needs a much higher truncation level to approximate the nonlinear function $f(t)/t$ with sufficient precision. This carries over to an explanation for the general outperforming of $\mathcal{A}_{\text{DNN}}$ strategies over those from $\mathcal{A}_{\text{lin}}$ when comparing identical signature truncation levels.

The method turned out to be efficient for this case study despite the fact that the cost functional and the drift coefficient are unbounded over the the set of admissible controls, contrary to assumptions (2.2) and (4.2) needed for the applicability of Theorem 4.7. This carries little surprise as we expect that our approximation results carries over to the case of L^2-type cost functionals (see paragraph on the *boundedness of the cost functional* in Section 2).

5.3. **Case study 2: Non-Markovian optimal execution.** As a second benchmark, we consider an instance of the optimal order execution problem of [KLPA20]. There X describes the fluctuations in the fundamental price of some financial asset. The controller seeks to unwind an initial inventory $q_0 > 0$ of shares over a (typically short) period $[0, T]$ by choosing trading rates $U = (U_t)_{t \in [0,T]}$ so as to maximize on average the proceeds $W_T^U := \int_0^T X_t U_t dt$ while taking into account market impact costs $\kappa \int_0^T U_t^2 dt$ from her sales. Any inventory remaining $Q_T^U := q_0 - \int_0^T U_t dt$ at time T is marked to market at $Q_T X_T$ and penalized by $\kappa_T Q_T^2$.

In the framework we study above this corresponds to the choice

$$Y_0^U = (0, q_0, x_0), \quad dY_t^U = (X_t U_t, -U_t, 0)dt + (0, 0, 1)dX_t, \quad t \in [0, T],$$

i.e., $Y^U := (W^U, Q^U, X)$ with cost functional

$$L(\mathbb{Y}^U, U) = -W_T^U - Q_T^U X_T + \int_0^T \kappa U_t^2 dt + \kappa_T (Q_T^U)^2.$$

In our numerical experiments, we followed [KLPA20, Section 5.4], and chose $X = x_0 + \sigma\xi$ with $x_0 = 1.0$, $\sigma = 0.02$ and ξ a fractional Brownian motion. We tested our methodology for a range of Hurst-parameters H and signature truncation levels N. The remaining parameters are given by $\kappa = 10^{-3}$, $\kappa_T = 10^{-1}$, $q_0 = 1$, and $T = 1$. For comparison, we calculated benchmarks with the method proposed in [KLPA20]. As described in Section 5.1, this method is based on a transformation to a deterministic optimization problem in terms of the expected signature.

	H	1/16	1/8	1/4	1/2	3/4	7/8
	$N = 1$	0.03	0.02	0.01	0.00	0.02	0.01
$\mathcal{A}_{\mathrm{lin}}$	$N = 2$	1.36	0.85	0.30	0.00	0.16	0.43
using [KLPA20]	$N = 3$	2.46	1.45	0.44	0.00	0.20	0.41
	$N = 4$	2.61	1.51	0.47	0.00	0.20	0.44
	$N = 1$	0.03	0.02	0.01	0.00	0.00	0.01
	$N = 2$	1.35	0.85	0.30	0.00	0.16	0.33
$\mathcal{A}_{\mathrm{lin}}$	$N = 3$	2.46	1.45	0.44	0.00	0.19	0.41
	$N = 4$	2.55	1.51	0.47	0.00	0.20	0.40
	$N = 5$	2.59	1.54	0.47	0.00	0.20	0.41
	$N = 1$	0.03	0.02	0.01	0.00	0.02	0.07
	$N = 2$	1.36	0.87	0.30	0.00	0.20	0.43
$\mathcal{A}_{\mathrm{DNN}}$	$N = 3$	2.53	1.47	0.44	0.00	0.20	0.44
	$N = 4$	2.61	1.53	0.48	0.00	0.21	0.44
	$N = 5$	2.69	1.54	0.49	0.00	0.21	0.43

TABLE 2. Numerical study of the optimal execution problem with fractional price fluctuations for different Hurst parameters H using strategies in $\mathcal{A}_{\mathrm{lin}}$ and $\mathcal{A}_{\mathrm{DNN}}$ ($I = 2$, $q = 30 + \eta_{2,N}$), with various signature truncation levels N. Presented are **relative improvements in percent** w.r.t. the TWAP strategy $J(U^{\circ}) \approx 0.9990$. The fixed model parameters are $q_0 = 1.0$, $T = 1$, $\kappa = 0.001$, $\kappa_T = 0.1$ and $\sigma = 0.02$. An overall time discretization of $\Delta t = 10^{-2}$ was used. The number of training and testing paths was 2^{19} and 2^{22}, respectively. The Monte Carlo resampling error was below 0.001. The first block contains benchmarks obtained with the linearization method from [KLPA20].

Table 2 collects the results of all tested scenarios. For better comparison we do not present the total expected costs, but the relative improvement compared to a trading strategy liquidating at a suitably chosen constant speed $U^{\circ} \equiv u \in \mathbb{R}$, also called the *time weighted average price* (TWAP) strategy [CJP15, Section 6.3]. The corresponding optimal rate and associated costs turn out to be

$$u = q_0 \frac{\kappa_T}{\kappa + T\kappa_T}, \qquad J(U^{\circ}) = x_0 q_0 - q_0^2 \frac{\kappa\kappa_t}{\kappa + T\kappa_T}.$$

Similar to the previous case study, we find an overall improvement when increasing the signature truncation level N and when moving from strategies in $\mathcal{A}_{\mathrm{lin}}$ to strategies in $\mathcal{A}_{\mathrm{DNN}}$. We also verify that the results of $\mathcal{A}_{\mathrm{lin}}$ strategies trained with our Monte Carlo method are close to those obtained with the approach proposed in [KLPA20]. The slightly better performance of the latter method is expected due to the availability of a more efficient optimization procedure when utilizing the quadratic convex structure of the problem. Nevertheless, using deep signature strategies in $\mathcal{A}_{\mathrm{DNN}}$, we are able to match and, in certain situations, surpass these benchmarks when comparing the same signature truncation levels.

APPENDIX A. APPROXIMATION OF PROGRESSIVELY MEASURABLE CONTROLS

In this section we relate canonical filtration on Ω_T^p to the Borel σ-algebras of the restricted path space Ω_t^p. To this end we define the coordinate process $\mathbb{Z}$ on Ω_T^p by

$$\mathbb{Z} : [0, T] \times \Omega_T^p \to G^{\lfloor p \rfloor}(V) : \quad (t, \mathbf{x}) \mapsto \mathbb{Z}_t(\mathbf{x}) := \mathbf{x}(t).$$

Lemma A.1. *For all $t \in [0, T]$ it holds $\sigma(\mathbb{Z}_s \mid 0 \le s \le t) = \rho_t^{-1}(\mathcal{B}(\Omega_t^p))$, where*

$$\rho_t : \Omega_T^p \to \Omega_t^p \quad \mathbf{x} \mapsto \mathbf{x}|_{[0,t]}.$$

Proof. Let $t \in [0, T]$. Clearly $\sigma(\mathbb{Z}_s \mid 0 \le s \le t) = \sigma(\mathcal{C}_t^T)$, where

$$\mathcal{C}_t^T := \left\{ \left\{ \mathbf{x} \in \Omega_T^p \mid (\mathbf{x}(t_1), \ldots, \mathbf{x}(t_n)) \in B \right\} \,\Big|\, t_1, \ldots, t_n \in [0, t], \quad B \in \mathcal{B}((G^{\lfloor p \rfloor}(V))^n) \right\}$$

are the cylinder subsets of Ω_T^p restricted to the time interval $[0, t]$. One observes that $\mathcal{C}_t^T = \rho_t^{-1}(\mathcal{C}_t)$ with $\mathcal{C}_t := \mathcal{C}_t^t \subseteq 2^{\Omega_t^p}$. By the continuity of the map $\Omega_t^p \to G^{\lfloor p \rfloor}(V) : \mathbf{x} \mapsto \mathbf{x}(s)$ for all $s \in [0, t]$ it follows that $\mathcal{C}_t \subset \mathcal{B}(\Omega_t^p)$ and thus

$$\sigma(\mathbb{X}_s \mid 0 \le s \le t) = \sigma(\mathcal{C}_t^T) = \rho_t^{-1}(\sigma(\mathcal{C}_t)) \subseteq \rho_t^{-1}(\mathcal{B}(\Omega_t^p)).$$

Conversely, define the system of dissections $\{t_k^n \mid k = 0, \ldots, n\} := \{tk/n \mid k = 0, \ldots, n\}$ of the interval $[0, t]$ and consider the geodesic interpolation map

$$\phi_n : (G^{\lfloor p \rfloor}(V))^n \to \Omega_t^p, \quad (\eta_1, \ldots, \eta_n) \mapsto \sum_{k=1}^n \mathbf{1}_{[t_{k-1}^n, t_k^n)} \Upsilon^{\eta_{k-1}, \eta_k} \left(\frac{(\cdot) - t_{k-1}^n}{t_k^n - t_{k-1}^n} \right),$$

where $\Upsilon^{a,b} : [0, 1] \to G^{\lfloor p \rfloor}(V)$ denotes the standard geodesic in $G^{\lfloor p \rfloor}(V)$ connecting the points $a, b \in G^{\lfloor p \rfloor}(V)$ (see [FV10, Theorem 7.32, Proposition 7.42]). One readily observes that ϕ_n is continuous. Indeed, this continuity is obvious when $G^{\lfloor p \rfloor}(V)$ is equipped with the geodesic distance, i.e. the Carnot-Caratheodory metric. In Section 3 we have equipped $G^{\lfloor p \rfloor}(V)$ with the (inhomogeneous) subspace topology of $T^{\lfloor p \rfloor}(V)$. This is consistent because the induced topology on $G^{\lfloor p \rfloor}(V)$ and Ω_t^p coincides with the one induced by the Carnot-Caratheodory metric (see [FV10, Section 8.1.3]).

For ease of notion we define for any $\mathbf{x} \in \Omega_t^p$ the geodesic interpolation on $\{t_k^n\}$ by $\mathbf{x}^{(n)} := \phi_n(\mathbf{x}(t_1^n), \ldots, \mathbf{x}(t_n^n))$. From Wiener's characterization of geometric rough paths (see [FV10, Theorem 8.23]) it thus follows that

$$\lim_{n \to \infty} d_{p-\mathrm{var};[0,t]}(\mathbf{x}^{(n)}, \mathbf{x}) = 0.$$

Hence for any $\mathbf{x}_0 \in \Omega_t^p$ and $r > 0$ we have

$$\begin{aligned}
B_r(\mathbf{x}_0) :=& \left\{ \mathbf{x} \in \Omega_t^p \,\Big|\, d_{p-\mathrm{var};[0,t]}(\mathbf{x}_0, \mathbf{x}) \le r \right\} \\
=& \left\{ \mathbf{x} \in \Omega_T^p \,\Big|\, \lim_{n \to \infty} d_{p-\mathrm{var};[0,t]}(\mathbf{x}_0, \mathbf{x}^{(n)}) \le r \right\} \\
=& \bigcap_{k=1}^\infty \bigcup_{m=1}^\infty \bigcap_{n=m}^\infty \left\{ \mathbf{x} \in \Omega_t^p \,\Big|\, d_{p-\mathrm{var};[0,t]}(\mathbf{x}_0, \mathbf{x}^{(n)}) \le r + \frac{1}{k} \right\} \\
=& \bigcap_{k=1}^\infty \bigcup_{m=1}^\infty \bigcap_{n=m}^\infty \left\{ \mathbf{x} \in \Omega_t^p \,\Big|\, (\mathbf{x}(t_1^n), \ldots, \mathbf{x}(t_n^n)) \in \phi_n^{-1}(B_{r+\frac{1}{k}}(\mathbf{x}_0)) \right\}
\end{aligned}$$

Since by continuity of ϕ_n it holds $\phi_n^{-1}(B_r(\mathbf{x}_0)) \in \mathcal{B}((G^{\lfloor p \rfloor}(V))^n)$, we see that the above right-hand side is the limes inferior of cylinder sets and thus in $\sigma(\mathcal{C}_t)$. This proves that $\mathcal{B}(\Omega_t^p) \subseteq \sigma(\mathcal{C}_t)$ and thus finally $\mathcal{B}(\Omega_t^p) = \sigma(\mathcal{C}_t^T) = \sigma(\mathbb{Z}_s \mid 0 \le s \le t)$. $\square$

References

[AC20] Andrew L. Allan and Samuel N. Cohen. Pathwise stochastic control with applications to robust filtering. *Ann. Appl. Probab.*, 30(5):2274–2310, 2020.

[BB23a] Christian Bayer and Simon Breneis. Efficient option pricing in the rough Heston model using weak simulation schemes, 2023.

[BB23b] Christian Bayer and Simon Breneis. Weak Markovian approximations of rough Heston. *arXiv preprint arXiv:2309.07023*, 2023.

[BCJ19] Sebastian Becker, Patrick Cheridito, and Arnulf Jentzen. Deep optimal stopping. *Journal of Machine Learning Research*, 20(74):1–25, 2019.

[BD20] Nicole Bäuerle and Sascha Desmettre. Portfolio optimization in fractional and rough Heston models. *SIAM Journal on Financial Mathematics*, 11(1):240–273, 2020.

[BEST23] Christian Bayer, Martin Eigel, Leon Sallandt, and Philipp Trunschke. Pricing high-dimensional Bermudan options with hierarchical tensor formats. *SIAM Journal on Financial Mathematics*, 14(2):383–406, 2023.

[BFRS16] Christian Bayer, Peter K. Friz, Sebastian Riedel, and John Schoenmakers. From rough path estimates to multilevel Monte Carlo. *SIAM J. Numer. Anal.*, 54(3):1449–1483, 2016.

[BGLY16] Horatio Boedihardjo, Xi Geng, Terry Lyons, and Danyu Yang. The signature of a rough path: uniqueness. *Adv. Math.*, 293:720–737, 2016.

[BHRS23] Christian Bayer, Paul P. Hager, Sebastian Riedel, and John Schoenmakers. Optimal stopping with signatures. *Ann. Appl. Probab.*, 33(1):238–273, 2023.

[BHT19] Christian Bayer, Juho Häppölä, and Raúl Tempone. Implied stopping rules for American basket options from Markovian projection. *Quantitative Finance*, 19(3):371–390, 2019.

[BPS23] Christian Bayer, Luca Pelizzari, and John Schoenmakers. Primal and dual optimal stopping with signatures. *arXiv preprint arXiv:2312.03444*, 2023.

[BSV17] Peter Bank, H. Mete Soner, and Moritz Voß. Hedging with temporary price impact. *Math. Financ. Econ.*, 11(2):215–239, 2017.

[BV18] Peter Bank and Moritz Voß. Linear quadratic stochastic control problems with stochastic terminal constraint. *SIAM J. Control Optim.*, 56(2):672–699, 2018.

[CF10] Rama Cont and David Fournie. A functional extension of the Itô formula. *Comptes Rendus Mathematique*, 348(1-2):57–61, 2010.

[CF19] Ilya Chevyrev and Peter K. Friz. Canonical RDEs and general semimartingales as rough paths. *Ann. Probab.*, 47(1):420–463, 2019.

[CGMSF23] Christa Cuchiero, Guido Gazzani, Janka Möller, and Sara Svaluto-Ferro. Joint calibration to SPX and VIX options with signature-based models, 2023.

[CGSF23] Christa Cuchiero, Guido Gazzani, and Sara Svaluto-Ferro. Signature-based models: theory and calibration. *SIAM J. Financ. Math.*, 14(3):910–957, 2023.

[Che54] Kuo-Tsai Chen. Iterated integrals and exponential homomorphisms. *Proc. Lond. Math. Soc. (3)*, 4:502–512, 1954.

[Che57] Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.

[CHT24] Prakash Chakraborty, Harsha Honnappa, and Samy Tindel. Pathwise relaxed optimal control of rough differential equations, 2024.

[CJP15] Á. Cartea, S. Jaimungal, and J. Penalva. *Algorithmic and High-Frequency Trading*. Cambridge University Press, 2015.

[CM24] Christa Cuchiero and Janka Möller. Signature methods in stochastic portfolio theory, 2024.

[CO22] Ilya Chevyrev and Harald Oberhauser. Signature moments to characterize laws of stochastic processes. *J. Mach. Learn. Res.*, 23(1), jan 2022.

[CPASB20] Álvaro Cartea, Imanol Perez Arribas, and Leandro Sánchez-Betancourt. Optimal execution of foreign securities: A double-execution problem. *Available at SSRN 3562251*, 2020.

[CPSFar] C. Cuchiero, F. Primavera, and S. Svaluto-Ferro. Universal approximation theorems for continuous functions of càdlàg paths and Lévy-type signature models. *Finance Stoch.*, to appear.

[CST23] Christa Cuchiero, Philipp Schmocker, and Josef Teichmann. Global universal approximation of functional input maps on weighted spaces. *arXiv preprint arXiv:2306.03303*, 2023.

[DFG17] Joscha Diehl, Peter K Friz, and Paul Gassiat. Stochastic control with rough paths. *Applied Mathematics & Optimization*, 75:285–315, 2017.

[DNT12] Aurélien Deya, Andreas Neuenkirch, and Samy Tindel. A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(2):518–550, 2012.

[Dup19]   Bruno Dupire. Functional Itô calculus. *Quant. Finance*, 19(5):721–729, 2019.

[EHJ17]   Weinan E, Jiequn Han, and Arnulf Jentzen. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Commun. Math. Stat.*, 5(4):349–380, 2017.

[FGS17]   Giorgio Fabbri, Fausto Gozzi, and Andrzej Swiech. Stochastic optimal control in infinite dimension. *Probability and Stochastic Modelling. Springer*, 2017.

[FH20]   Peter K. Friz and Martin Hairer. *A course on rough paths. With an introduction to regularity structures*. Universitext. Cham: Springer, 2nd updated edition edition, 2020.

[FHW23]   Owen Futter, Blanka Horvath, and Magnus Wiese. Signature trading: A path-dependent extension of the mean-variance framework with exogenous signals. *arXiv preprint arXiv:2308.15135*, 2023.

[FR14]   Peter K. Friz and Sebastian Riedel. Convergence rates for the full Gaussian rough paths. *Ann. Inst. Henri Poincaré Probab. Stat.*, 50(1):154–194, 2014.

[FS06]   Wendell Helms Fleming and Halil Mete Soner. *Controlled Markov processes and viscosity solutions*. Stochastic modelling and applied probability. Springer, New York, NY, 2006.

[FS17]   Peter K. Friz and Atul Shekhar. General rough integration, Lévy rough paths and a Lévy-Kintchine-type formula. *Ann. Probab.*, 45(4):2707–2765, 2017.

[FV10]   Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Theory and applications.

[GM05]   Emmanuel Gobet and Remi Munos. Sensitivity analysis using Itô–Malliavin calculus and martingales, and application to stochastic optimal control. *SIAM J. Control Optim.*, 43(5):1676–1713, 2005.

[GPW23]   M. Germain, H. Pham, and X. Warin. *Neural Networks-Based Algorithms for Stochastic Control and PDEs in Finance*, page 426–452. Cambridge University Press, 2023.

[HE16]   Jiequn Han and Weinan E. Deep learning approximation for stochastic control problems, 2016.

[HFH+23]   Melker Höglund, Emilio Ferrucci, Camilo Hernández, Aitor Muguruza Gonzalez, Cristopher Salvi, Leandro Sánchez-Betancourt, and Yufei Zhang. A neural RDE approach for continuous-time non-markovian stochastic control problems. In *ICML Workshop on New Frontiers in Learning, Control, and Dynamical Systems*, 2023.

[HH21]   Jiequn Han and Ruimeng Hu. Recurrent neural networks for stochastic control problems with delay. *Math. Control Signals Syst.*, 33(4):775–795, 2021.

[HL10]   Ben Hambly and Terry Lyons. Uniqueness for the signature of a path of bounded variation and the reduced path group. *Ann. Math. (2)*, 171(1):109–167, 2010.

[HPBL21]   Côme Huré, Huyên Pham, Achref Bachouch, and Nicolas Langrené. Deep neural networks algorithms for stochastic control problems on finite horizon: convergence analysis. *SIAM J. Numer. Anal.*, 59(1):525–557, 2021.

[HPW20]   Côme Huré, Huyên Pham, and Xavier Warin. Deep backward schemes for high-dimensional nonlinear PDEs. *Math. Comput.*, 89(324):1547–1579, 2020.

[KB14]   Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[KLPA20]   Jasdeep Kalsi, Terry Lyons, and Imanol Perez Arribas. Optimal execution with rough path signatures. *SIAM J. Financial Math.*, 11(2):470–493, 2020.

[KMFL20]   Patrick Kidger, James Morrill, James Foster, and Terry Lyons. Neural controlled differential equations for irregular time series. In *Advances in Neural Information Processing Systems*, volume 33, pages 6696–6707. Curran Associates, Inc., 2020.

[KS97]   V.B. Kolmanovskiĭ and L.E. Shaĭkhet. *Control of Systems with Aftereffect*. Translations of Mathematical Monographs. American Mathematical Soc., 1997.

[LCL07]   Terry J. Lyons, Michael Caruana, and Thierry Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004.

[LLPS93]   Moshe Leshno, Vladimir Ya. Lin, Allan Pinkus, and Shimon Schocken. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks*, 6(6):861–867, 1993.

[LOSar]   Dorival Leão, Alberto Ohashi, , and Francys de Souza. Solving non-Markovian stochastic control problems driven by Wiener functionals. *Ann. Appl. Probab.*, to appear.

[Lyo98]   Terry J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoam.*, 14(2):215–310, 1998.

[MSKF21] James Morrill, Cristopher Salvi, Patrick Kidger, and James Foster. Neural rough differential equations for long time series. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 7829–7838. PMLR, 18–24 Jul 2021.

[NP00] Carl J. Nuzman and H. Vincent Poor. Linear estimation of self-similar processes via Lamperti's transformation. *J. Appl. Probab.*, 37(2):429–452, 2000.

[PGM+19] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[Pro05] P.E. Protter. *Stochastic Integration and Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2005.

[RG20] Jeremy Reizenstein and Benjamin Graham. Algorithm 1004: The iisignature library: Efficient calculation of iterated-integral signatures and log signatures. *ACM Transactions on Mathematical Software (TOMS)*, 2020.

[RR22] M. Redmann and S. Riedel. Runge-Kutta methods for rough differential equations. *J. Stoch. Anal.*, 3(4):Art. 6, 24, 2022.

[Rud87] Walter Rudin. *Real and Complex Analysis, 3rd Ed.* McGraw-Hill, Inc., USA, 1987.

[RW20] Johannes Ruf and Weiguan Wang. Neural networks for option pricing and hedging: a literature review. *Journal of Computational Finance*, 24(1):1–46, 2020.

[SS18] Justin Sirignano and Konstantinos Spiliopoulos. DGM: a deep learning algorithm for solving partial differential equations. *J. Comput. Phys.*, 375:1339–1364, 2018.

[Wiś94] Andrzej Wiśniewski. The structure of measurable mappings on metric spaces. *Proc. Amer. Math. Soc.*, 122(1):147–150, 1994.

PETER BANK, INSTITUT FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT BERLIN, GERMANY
Email address: `bank@math.tu-berlin.de`

CHRISTIAN BAYER, WEIERSTRASS INSTITUTE, BERLIN, GERMANY
Email address: `christian.bayer@wias-berlin.de`

PAUL PETER HAGER, INSTITUT FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT BERLIN, GERMANY
Email address: `hager@math.tu-berlin.de`

SEBASTIAN RIEDEL, FAKULTÄT FÜR MATHEMATIK UND INFORMATIK, FERNUNIVERSITÄT HAGEN, GERMANY
Email address: `sebastian.riedel@fernuni-hagen.de`

TOBIAS NAUEN, DEUTSCHES FORSCHUNGSZENTRUM FÜR KÜNSTLICHE INTELLIGENZ, KAISERSLAUTERN, GERMANY
Email address: `tobias_christian.nauen@dfki.de`