

Multistage Conditional Compositional Optimization

Buse Şen[†], Yifan Hu[‡], and Daniel Kuhn[†]

[†]EPFL, Switzerland

buse.sen@epfl.ch

daniel.kuhn@epfl.ch

[‡]Rutgers University, USA

yifan.hu@rutgers.edu

April 15, 2026

Abstract

We introduce Multistage Conditional Compositional Optimization (MCCO) as a new paradigm for decision-making under uncertainty that combines aspects of multistage stochastic programming and conditional stochastic optimization. MCCO minimizes a nest of conditional expectations and nonlinear cost functions. It has numerous applications and arises, for example, in optimal stopping, linear–quadratic regulator problems, distributionally robust contextual bandits, as well as in problems involving dynamic risk measures. The naïve nested sampling approach for MCCO suffers from the curse of dimensionality familiar from scenario tree-based multistage stochastic programming, that is, its scenario complexity grows exponentially with the number of nests. We develop new multilevel Monte Carlo techniques for MCCO whose scenario complexity grows only polynomially with the desired accuracy.

1 Introduction

Multistage stochastic optimization constitutes a central and still largely unresolved challenge in the field of optimization under uncertainty. Efficient exact solution methods are out of reach even if the probability distribution of the random problem parameters is given in closed form. Indeed, even linear two-stage stochastic programs with independent and uniformly distributed random parameters are #P-hard [20, 34]. Consequently, multistage stochastic programs are often studied in a black-box framework, where the decision-maker has no analytical distributional information and aims to construct an ϵ -optimal solution relying only on oracles that supply samples of the random problem parameters. In this framework, multistage stochastic programs are often addressed with the *Sample Average Approximation* (SAA), which builds a scenario tree using conditional sampling and solves a finite-dimensional approximate stochastic program on this empirically generated tree. However, the scenario complexity of SAA for obtaining an ϵ -optimal solution to a T -stage linear stochastic program is bounded above by $\mathcal{O}(\epsilon^{-2T})$ [56, 57], and this bound cannot be significantly improved by any SAA-type method even if the random parameters are serially independent [52, Proposition 2]. Hence, the size of this scenario tree grows exponentially with the number of stages. We are not aware of any other black-box method that overcomes this curse of dimensionality.

With this paper we hope to lay the groundwork for a new generation of black-box methods for multistage stochastic optimization whose scenario complexity depends on the desired accuracy ϵ only through $(1/\epsilon)^p$, where the exponent p is independent of T . To this end, we study *Multistage Conditional Compositional Optimization* (MCCO) problems, which are reminiscent of multistage stochastic programs. In fact, we

will show that MCCO encapsulates various stochastic dynamic control problems with serially dependent disturbances as special cases. In addition, we will argue that MCCO is highly expressive and even includes problem instances that cannot be reformulated as any risk-neutral multistage stochastic program. The main contribution of this paper is to propose new multilevel Monte Carlo methods for MCCO problems with $T \geq 2$ stages whose scenario complexity scales as $\mathcal{O}(\epsilon^{-2})$ under standard regularity conditions.

The MCCO problems to be investigated in this paper can be represented as

$$\min_{x \in \mathcal{X}} F(x), \quad (\text{MCCO})$$

where $\mathcal{X} \subseteq \mathbb{R}^d$ denotes the feasible set, and where the objective function $F : \mathbb{R}^d \rightarrow \mathbb{R}$ is defined as a nest of T conditional expectations and nonlinear integrands of the form

$$F(x) := \mathbb{E} \left[f_1 \left(\xi_1, \mathbb{E}_1 \left[f_2 \left(\xi_2, \dots, \mathbb{E}_{T-2} \left[f_{T-1} \left(\xi_{T-1}, \mathbb{E}_{T-1} [f_T(\xi_T, x)] \right) \right] \dots \right) \right] \right) \right]. \quad (\text{MCCE})$$

Here, ξ_1, \dots, ξ_T represents a stochastic process, $\mathbb{E}[\cdot]$ denotes unconditional expectation, and $\mathbb{E}_t[\cdot]$ denotes expectation conditional on the σ -algebra \mathcal{F}_t generated by the time- t observation history $\xi_{[t]} := (\xi_1, \dots, \xi_t)$. The integrands $f_t : \mathbb{R}^{m_t} \times \mathbb{R}^{d_t} \rightarrow \mathbb{R}^{d_{t-1}}$ with $d_0 = 1$ and $d_T = d$ can be interpreted as uncertainty-affected cost functions. A key assumption throughout the paper is that the random vectors ξ_t are *dependent* and that the integrands f_t are *nonlinear*. Hence, evaluating the objective function at a given $x \in \mathcal{X}$ is challenging, a problem which we henceforth refer to as *Multistage Conditional Compositional Estimation* (MCCE).

MCCO reduces to classical (single-stage) stochastic optimization when $T = 1$, to conditional stochastic optimization when $T = 2$ [18, 29, 37, 39–41], to multilevel compositional stochastic optimization when the random vectors ξ_t are mutually independent [4, 14, 17, 21, 44, 54, 69–71] and to MCCE when the feasible set \mathcal{X} is a singleton [51, 61, 72]. As we allow T to be any integer and the random vectors ξ_t to be dependent, MCCO strictly generalizes all of these more basic decision-making paradigms.

Both MCCE and MCCO arise in many applications relevant for operations research, economics or engineering. For example, the problem of computing credit valuation adjustments can be naturally formulated as an instance of MCCE [28]. MCCE also emerges in optimal stopping problems, which encompass applications such as option pricing in finance [72] or inventory replenishment in operations management [50]. In addition, we will show that linear quadratic regulator problems and general finite-state stochastic control problems can be reformulated as instances of MCCE. Remarkably, these reformulations are possible even when the disturbances are serially dependent, as is the case when the underlying system is in a transient regime or is subject to distribution shifts. Finally, MCCO naturally arises in distributionally robust contextual bandits [59] or nested risk minimization [67]. All of these examples, which will be presented in detail in Section 2, underscore the strong ties between MCCE, MCCO and multistage stochastic optimization. However, MCCO is not subsumed by multistage stochastic optimization. Indeed, the infimum of any risk-neutral multistage stochastic program is concave in the joint distribution of the random problem parameters, whereas $F(x)$ clearly fails to be concave in the distribution of $\xi_{[T]}$ for some choices of the integrands f_t .

Much like in multistage stochastic optimization, efficient exact solution methods for MCCE and MCCO are out of reach even if the probability distribution of the underlying stochastic process is given in closed form. Indeed, solving MCCE exactly is #P-hard even if $T = 1$, ξ_1 follows the uniform distribution on the unit hypercube in \mathbb{R}^{m_1} and f_1 is a piecewise affine function with only two pieces [34, Corollary 1]. We thus adopt a black-box model, that is, we assume access to T oracles (“black boxes”) that one can use to generate any number of independent samples from the marginal distribution of ξ_1 and the conditional distribution of ξ_{t+1} given $\xi_{[t]}$, $t \in [T - 1]$, respectively. However, we do not presume to know these distributions. In this situation, we can resort to sampling methods for solving MCCE and MCCO approximately.

In this paper we will design and analyze function value estimators $\widehat{F}(x)$ for $F(x)$ that attain a specified mean squared error with a minimum number of scenarios (*i.e.*, sample paths), and we will refer to this number as the estimator’s mean squared error-based scenario complexity. Given a function value estimator $\widehat{F}(x)$, one can attempt to solve the following sample-based approximation of the intractable MCCO problem.

$$\min_{x \in \mathcal{X}} \widehat{F}(x) \quad (1)$$

We will show that the minimizers of problem (1) are near-optimal in (MCCO) with high probability. However, computing these minimizers can be challenging in the absence of restrictive convexity and monotonicity conditions. Thus, we will also design and analyze sample-efficient estimators for $\nabla F(x)$ that can be employed within stochastic gradient descent-type methods for solving problem (MCCO) to stationarity.

The SAA estimator is arguably the simplest estimator for $F(x)$. It is obtained by replacing all conditional expectations in (MCCE) with conditional sample averages. This amounts to approximating the stochastic process of the disturbances $\xi_t, t \in [T]$, by a scenario tree. The number of samples used to approximate $\mathbb{E}_t[\cdot]$ coincides with the scenario tree’s branching factor at stage t . We will show that the scenario complexity of an SAA estimator with optimally tuned branching factors amounts to $\mathcal{O}(\epsilon^{-2T})$ if all integrands f_t are Lipschitz continuous and to $\mathcal{O}(\epsilon^{-(T+1)})$ if the integrands are additionally smooth. In both cases, the SAA estimator suffers from a curse of dimensionality. To design more efficient estimators, we leverage *Multilevel Monte Carlo* (MLMC) methods, which have originally been popularized in computational finance and uncertainty quantification [23, 24]. A basic MLMC estimator can be constructed as follows. For any level $\ell \in \mathbb{N}_0$, let \widehat{Y}_ℓ denote the SAA estimator of $F(x)$ with a uniform branching factor 2^ℓ across all nodes of the scenario tree, and set $\widehat{Y}_{-1} = 0$. By the law of large numbers, \widehat{Y}_ℓ converges almost surely to $F(x)$ as ℓ grows. Hence, $F(x)$ admits the telescoping sum representation $F(x) = \mathbb{E}[\sum_{\ell=0}^{\infty} (\widehat{Y}_\ell - \widehat{Y}_{\ell-1})]$. If λ is a discrete random variable independent of all samples and if $\mathbb{P}(\lambda = \ell) = q(\ell) > 0$ for all \mathbb{N}_0 , then the law of total expectation implies

$$F(x) = \mathbb{E} \left[\sum_{\ell=0}^{\infty} q(\ell) \frac{\widehat{Y}_\ell - \widehat{Y}_{\ell-1}}{q(\ell)} \right] = \mathbb{E} \left[\frac{\widehat{Y}_\lambda - \widehat{Y}_{\lambda-1}}{q(\lambda)} \right]. \quad (2)$$

The inverse propensity weighted telescoping term $\widehat{F}(x) = (\widehat{Y}_\lambda - \widehat{Y}_{\lambda-1})/q(\lambda)$ associated with level λ constitutes a randomized MLMC estimator in the spirit of [11, 53]. By construction, it satisfies $\mathbb{E}[\widehat{F}(x)] = F(x)$ and is thus unbiased. The choice of the distribution for λ is critical for controlling its computational cost. To see this, note that the SAA estimator \widehat{Y}_ℓ requires $2^{\ell(T-1)}$ scenarios—one for each leaf node of the underlying scenario tree—implying that its computational cost grows quickly with ℓ . Drawing λ from a geometric distribution, for instance, ensures that small (inexpensive) levels ℓ are sampled with high probability, while large (expensive) levels are sampled rarely. This trick allows us to reduce the expected computational cost of $\widehat{F}(x)$ while preserving its unbiasedness. To further ensure that the variance of $\widehat{F}(x)$ remains bounded, we reuse all samples of the level $\ell - 1$ scenario tree when constructing the level ℓ tree. Hence, \widehat{Y}_ℓ and $\widehat{Y}_{\ell-1}$ are strongly positively correlated, and the variance of $\widehat{F}(x)$ remains bounded thanks to a control variate effect.

In this paper, we introduce new MLMC estimators for $F(x)$ and $\nabla F(x)$ whose construction departs from that of the naïve MLMC estimator described above in three key respects. First, we truncate the telescoping sum in (2) at a maximum level M in order to control the estimator’s variance [10]. Truncation introduces a bias, but the bias decays exponentially with M . Hence, selecting M becomes a matter of balancing bias and variance. Second, we adopt a stagewise recursive construction employing antithetic sampling [27], thereby improving the estimator’s computational cost. Third, we average multiple independent and identically distributed (i.i.d.) copies of the same estimator to reduce the variance. We will prove that the scenario complexity of the resulting MLMC estimator for $F(x)$ amounts to $\mathcal{O}(\log(\epsilon^{-1})^{2T-2} \epsilon^{-2})$ if all integrands f_t are Lipschitz continuous and to $\mathcal{O}(\epsilon^{-2})$ if the integrands are additionally smooth. Under an additional Hölder continuity assumption, analogous results are obtained for our MLMC estimator of $\nabla F(x)$.

The main contributions of this paper can be summarized as follows.

1. **SAA value estimator:** We provide a refined scenario complexity analysis of the SAA estimator for $F(x)$. First, we prove that the mean squared error-based scenario complexity of this estimator scales as $\mathcal{O}(\epsilon^{-2T})$ and $\mathcal{O}(\epsilon^{-(T+1)})$ for nonsmooth and smooth integrands, respectively. Under a standard sub-Gaussianity assumption, we further show that, for nonsmooth and smooth integrands respectively, $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-2T})$ and $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-(T+1)})$ scenarios suffice to ensure that the estimator differs from $F(x)$ at most by ϵ , uniformly across all $x \in \mathcal{X}$ and with high probability. All bounds are presented explicitly without any unknown constants. Thus, they provide finite-sample guarantees.
2. **MLMC value estimator:** We propose a new recursive truncated MLMC estimator for $F(x)$ together with sharp¹ scenario complexity guarantees. First, we prove that the mean squared error-based scenario complexity of this estimator scales as $\mathcal{O}(\log(\epsilon^{-1})^{2T-2}\epsilon^{-2})$ and $\mathcal{O}(\epsilon^{-2})$ for nonsmooth and smooth integrands, respectively. Under a standard sub-Gaussianity assumption, we further show that, for nonsmooth and smooth integrands respectively, $\mathcal{O}(\log(\epsilon^{-1})^T\epsilon^{-2})$ and $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-2})$ scenarios suffice to ensure that the estimator differs from $F(x)$ at most by ϵ , uniformly across all $x \in \mathcal{X}$ and with high probability. Consequently, we provide the first scenario complexity results for MCCE with nonsmooth integrands that eliminate the usual exponential dependence on T in ϵ^{-1} .
3. **MCCO paradigm:** We introduce MCCO as a new paradigm for optimization under uncertainty and show that it unifies and generalizes numerous traditional decision-making models from the literature.
4. **MLMC gradient estimator:** We propose new recursive truncated as well as *untruncated* MLMC estimators for $\nabla F(x)$ together with sharp scenario complexity guarantees. Under the assumption that the integrands are smooth and have Hölder continuous Hessians, we prove that the mean squared error-based scenario complexities of both estimators scale as $\mathcal{O}(\epsilon^{-2})$. If the Hölder continuity assumption is relaxed, the scenario complexity increases to $\mathcal{O}(\log(\epsilon^{-1})^{2T-2}\epsilon^{-2})$ for the truncated and diverges for the untruncated estimator. We develop the first MLMC gradient estimators for MCCO with $T \geq 2$.
5. **SGD algorithm:** We integrate our new MLMC gradient estimators into a stochastic gradient descent (SGD) algorithm for finding a stationary point of problem (MCCO). We prove that if all integrands are smooth and have Hölder continuous Hessians, then this algorithm needs $\mathcal{O}(\epsilon^{-4})$ scenarios to reach an ϵ -stationary point both when equipped with the truncated and untruncated variant of the estimator.
6. **Numerical experiments:** We validate our theoretical results with numerical experiments revolving around Bermudan option pricing (for MCCE) and distributionally robust contextual bandits (for MCCO). These experiments confirm that the proposed MLMC estimators achieve the same mean square error as an SAA estimator using a significantly smaller number of scenarios.

Literature Review MCCE with dependent disturbances has been widely studied in the literature. However, most works focus on problems with $T = 2$ stages, which arise in derivative pricing [13, 33] or risk management [12, 24, 26, 32]. For $T = 2$, it is well known that SAA estimators typically achieve a root mean squared error of ϵ using $\mathcal{O}(\epsilon^{-3})$ scenarios, while MLMC estimators require only $\mathcal{O}(\epsilon^{-2})$ scenarios [25]. The literature on MCCE with $T > 2$ stages is far more limited. It is shown in [51] that the scenario complexity of SAA estimators amounts to $\mathcal{O}(\epsilon^{-2T})$ for nonsmooth integrands and to $\mathcal{O}(\epsilon^{-(T+1)})$ for smooth ones. This work, however, provides only asymptotic rates without explicit constants. Moreover, it only studies mean squared errors but provides no uniform convergence analysis, which would be necessary to ensure that the minimizers of (1) approximate those of (MCCO) with high probability. To the best of our knowledge, only two papers develop MLMC estimators for MCCE with $T > 2$. The first focuses on

¹Some of these guarantees are sharp only up to poly-logarithmic factors in ϵ^{-1} .

optimal stopping problems and imposes restrictive assumptions on the integrands [72]. The second primarily considers smooth integrands. While it also discusses nonsmooth integrands, the corresponding MLMC estimators have infinite variance, making problem (1) numerically unstable [61]. No prior work develops MLMC estimators for MCCE with $T > 2$ stages and general nonsmooth integrands.

To our best knowledge, all existing work on MCCO assumes that $T = 2$. In this case, MCCO reduces to conditional stochastic optimization, which arises in many applications. In machine learning, for instance, it appears in reinforcement learning [19, 39], model-agnostic meta-learning, robust supervised learning or invariant learning [39, 41]. In statistics and econometrics, applications include instrumental variable regression [8, 15, 29, 36, 49, 60] or Bayesian experimental design [5, 30, 31]. The simplest approach to MCCO is to approximate $F(x)$ by an SAA estimator $\hat{F}(x)$ and to solve (1). For $T = 2$, the number of scenarios needed to guarantee that the resulting minimizer of (1) constitutes an ϵ -optimal solution to (MCCO) with high probability amounts to $\mathcal{O}(\epsilon^{-4})$ if the integrands are nonsmooth and can be reduced to $\mathcal{O}(\epsilon^{-3})$ if the integrands are smooth [39]. However, solving (1) may be computationally infeasible without restrictive convexity or monotonicity assumptions. To address this, subsequent work develops SAA estimators for $\nabla F(x)$ that can be employed within a stochastic gradient descent algorithm for solving problem (MCCO) to stationarity. For smooth integrands, this method requires $\mathcal{O}(\epsilon^{-6})$ scenarios to find an ϵ -stationary point, or $\mathcal{O}(\epsilon^{-5})$ with standard variance reduction techniques [40]. MLMC estimators for $\nabla F(x)$ introduced in [42] further reduce this to $\mathcal{O}(\epsilon^{-4})$ scenarios. MCCO with $T > 2$ stages has not yet been investigated.

The remainder of the paper is structured as follows. Section 2 illustrates motivating applications of MCCE and MCCO. Sections 3 and 4 then introduce and analyze SAA and MLMC estimators for $F(x)$ and $\nabla F(x)$, respectively. Numerical results are presented in Section 5, and all proofs as well as some auxiliary technical results are relegated to the appendix.

Notation The entrywise p -norm of a matrix $A \in \mathbb{R}^{n \times m}$ is denoted by $\|A\|_p := (\sum_{j=1}^m \sum_{i=1}^n |A_{ij}|^p)^{1/p}$. If $p = 2$, $\|A\|_p$ reduces to the Frobenius norm. We use \mathbb{N}_0 and \mathbb{N} to denote the sets of natural numbers with and without 0, respectively, and for any $n \in \mathbb{N}$, we set $[n] = \{1, 2, \dots, n\}$. The transpose of the Jacobian of a differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is denoted by $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$. If $m = 1$, ∇f thus reduces to the gradient of f . All (in)equalities involving random variables are understood to hold almost surely. The variance of a random variable ξ is denoted by $\mathbb{V}(\xi)$. We designate estimators by the superscript ‘ $\hat{\cdot}$ ’ and define $\mathbb{C}(\hat{F})$ as the number of scenarios (sample paths) needed to construct an estimator \hat{F} . The truncated geometric distribution with rate $r \in (0, 1)$ and truncation point $M \in \mathbb{N}$ is denoted by $\text{Geo}(r|M)$. Thus, a random variable λ governed by $\text{Geo}(r|M)$ satisfies $\mathbb{P}(\lambda = \ell) = r(1-r)^\ell / (1 - (1-r)^{M+1})$ for every $\ell = 0, \dots, M$. If $a(\epsilon)$ and $b(\epsilon)$ are real-valued functions of a tolerance $\epsilon > 0$, then $a(\epsilon) = \mathcal{O}(b(\epsilon))$ means that there exist $C > 0$ and $\bar{\epsilon} > 0$ such that $a(\epsilon) \leq Cb(\epsilon)$ for all $\epsilon \in (0, \bar{\epsilon})$. Note that $a(\epsilon)$ and $b(\epsilon)$ could depend on other variables. However, $\mathcal{O}(\cdot)$ captures only their asymptotic dependence on ϵ .

2 Motivating Applications

Instances of (MCCE) and (MCCO) with $T > 2$ stages arise in numerous applications. Examples include optimal stopping, linear-quadratic regulator and finite-state stochastic control problems for (MCCE) as well as policy learning in contextual bandits and risk measurement with nested risk measures for (MCCO).

Optimal Stopping Suppose that an agent observes a sequence of random variables ξ_t , $t \in [T]$, which are revealed one at a time, and let $g_t : \mathbb{R}^{m_t} \rightarrow \mathbb{R}$, $t \in [T]$, be Borel measurable payoff functions. At each time $t \in [T]$, the agent must decide whether to stop observing, in which case an immediate payoff $g_t(\xi_t)$ is received, or to continue observing, in which case no immediate payoff is received. The process must stop at time T at the latest. If the goal is to maximize the expected payoff, then the optimal stopping time can be

found via dynamic programming [72]. To this end, use $U_t(\xi_{[t]})$ to denote the agent's expected future payoff conditional on all observations up to time t . One readily verifies that $U_T(\xi_{[T]}) = g_T(\xi_T)$ and

$$U_t(\xi_{[t]}) = \max \left\{ g_t(\xi_t), \mathbb{E}_t [U_{t+1}(\xi_{[t+1]})] \right\} \quad \forall t \in [T-1].$$

Unraveling this recursion yields

$$\mathbb{E} [U_1(\xi_1)] = \mathbb{E} \left[\max \left\{ g_1(\xi_1), \mathbb{E}_1 \left[\max \left\{ g_2(\xi_2), \dots, \mathbb{E}_{T-1} [g_T(\xi_T)] \dots \right\} \right] \right\} \right].$$

Thus, the problem of computing the expected value of the optimal stopping time can be viewed as an instance of (MCCE) with $f_T(\xi_T, x_T) = g_T(\xi_T)$ and $f_t(\xi_t, x_t) = \max\{g_t(\xi_t), x_t\}$ for all $t \in [T-1]$. Optimal stopping problems emerge in various applications across many domains [16]. For example, they arise in the pricing of American-style options [72] or in sequential hypothesis testing [47]. In addition, they also arise in supply chain management [63] to determine the timing of inventory replenishment [50], and in game theory to model the reasoning about the probabilistic strategies of players [51].

Linear-Quadratic Regulator Consider the finite-horizon stochastic linear-quadratic regulator problem

$$\begin{aligned} \min_{a,s} \quad & \mathbb{E} \left[\sum_{t=1}^{T-1} \left(s_t^\top Q s_t + a_t^\top R a_t \right) + s_T^\top P_T s_T \right] \\ \text{s.t.} \quad & s_{t+1} = A s_t + B a_t + \xi_{t+1} \quad \forall t \in [T-1] \\ & a_t \text{ is } \mathcal{F}_t\text{-measurable} \quad \forall t \in [T-1], \end{aligned} \quad (3)$$

where $s_t \in \mathbb{R}^m$ and $a_t \in \mathbb{R}^n$ denote the state and the control action at time t . We assume that the system matrices $A \in \mathbb{R}^{m \times m}$ and $B \in \mathbb{R}^{m \times n}$ as well as the symmetric positive semidefinite cost matrices $Q \in \mathbb{R}^{m \times m}$, $R \in \mathbb{R}^{n \times n}$ and $P_T \in \mathbb{R}^{m \times m}$ are deterministic and that R is invertible. We also assume that $s_1 = \xi_1$ and that $\{\xi_t\}_{t \in \mathbb{N}}$ is a (possibly serially dependent) stochastic process with finite second moments taking values in \mathbb{R}^m . Control problems of the form (3) are ubiquitous in economics [35], engineering [9] or neuroscience [62], among many others. Traditionally, the stochastic disturbances are assumed to be i.i.d. However, in an era where control systems must navigate continuously shifting environments and interact with adaptive, data-driven agents, models that abandon the fragile i.i.d. noise assumption are essential for ensuring robustness under real-world nonstationarity and persistent distribution shift.

The Q-function at time t associated with problem (3) is defined via the Bellman-type recursion

$$Q_t(s, a) = s^\top Q s + a^\top R a + \mathbb{E}_t \left[\min_{a' \in \mathbb{R}^n} Q_{t+1}(A s + B a + \xi_{t+1}, a') \right] \quad (4)$$

with terminal value $Q_T(s, a) = s^\top P_T s$ (which is deterministic and independent of a); see also [9, § 4.1]. Due to the serial dependencies of the noise process, $Q_t(s, a)$ is an \mathcal{F}_t -measurable random object for any fixed (s, a) . But we notationally suppress the dependence of $Q_t(s, a)$ on $\xi_{[t]}$ to avoid clutter. The optimal value of (3) can be expressed in terms of the time-1 Q-function as

$$\mathbb{E} \left[\min_{a \in \mathbb{R}^n} Q_1(\xi_1, a) \right]. \quad (5)$$

The following lemma shows that the Q-functions are quadratic in the states and actions with random coefficients. To keep the paper self-contained, we sketch the proof of this elementary result in Appendix B.

Lemma 2.1. For each $t \in [T]$, the Q -function is quadratic in (s, a) and can be represented as

$$Q_t(s, a) = \begin{bmatrix} s \\ a \end{bmatrix}^\top \begin{pmatrix} Q_t^{ss} & Q_t^{sa} \\ (Q_t^{sa})^\top & Q_t^{aa} \end{pmatrix} \begin{bmatrix} s \\ a \end{bmatrix} + 2 \begin{bmatrix} b_t^s \\ b_t^a \end{bmatrix}^\top \begin{bmatrix} s \\ a \end{bmatrix} + c_t \quad \mathbb{P}\text{-a.s.}, \quad (6a)$$

where the \mathcal{F}_t -measurable coefficient matrices and vectors are defined through the backward recursions

$$\begin{aligned} Q_t^{ss} &= \mathbb{E}_t [Q + A^\top P_{t+1} A] \\ Q_t^{sa} &= \mathbb{E}_t [A^\top P_{t+1} B] \\ Q_t^{aa} &= \mathbb{E}_t [R + B^\top P_{t+1} B] \\ b_t^s &= \mathbb{E}_t [A^\top P_{t+1} \xi_{t+1} + A^\top g_{t+1}] \\ b_t^a &= \mathbb{E}_t [B^\top P_{t+1} \xi_{t+1} + B^\top g_{t+1}] \\ c_t &= \mathbb{E}_t [\xi_{t+1}^\top P_{t+1} \xi_{t+1} + 2 g_{t+1}^\top \xi_{t+1} + d_{t+1}] \end{aligned} \quad (6b)$$

for all $t \in [T - 1]$ initialized by

$$Q_T^{ss} = P_T, \quad Q_T^{sa} = 0, \quad Q_T^{aa} = 0, \quad b_T^s = 0, \quad b_T^a = 0 \quad \text{and} \quad c_T = 0. \quad (6c)$$

The underlying auxiliary variables are defined by

$$P_t = Q_t^{ss} - Q_t^{sa} (Q_t^{aa})^{-1} (Q_t^{sa})^\top, \quad g_t = b_t^s - Q_t^{sa} (Q_t^{aa})^{-1} b_t^a \quad \text{and} \quad d_t = c_t - (b_t^a)^\top (Q_t^{aa})^{-1} b_t^a \quad (6d)$$

for all $t \in [T - 1]$, while P_T is the terminal state cost matrix, $g_T = 0$ and $d_T = 0$.

For all $t \in [T]$ we define an aggregate coefficient vector encoding the time- t Q -function as

$$x_t = (Q_t^{ss}, Q_t^{sa}, Q_t^{aa}, b_t^s, b_t^a, c_t) \in \mathbb{R}^d,$$

where the components are the \mathcal{F}_t -measurable matrices and vectors appearing in (6a), and where $d = m^2 + mn + n^2 + m + n + 1$. Also, for all $t = 2, \dots, T$ we define $f_t : \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ through

$$f_t(\xi_t, x_t) = \left(Q + A^\top P_t A, A^\top P_t B, R + B^\top P_t B, A^\top P_t \xi_t + A^\top g_t, B^\top P_t \xi_t + B^\top g_t, \xi_t^\top P_t \xi_t + 2g_t^\top \xi_t + d_t \right),$$

where P_t, g_t and d_t are defined as in (6d) and thus constitute rational functions of x_t . In addition, we define $f_1 : \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}$ through

$$f_1(\xi_1, x_1) = \xi_1^\top \left(Q_1^{ss} - Q_1^{sa} (Q_1^{aa})^{-1} (Q_1^{sa})^\top \right) \xi_1 + 2 \left(b_1^s - Q_1^{sa} (Q_1^{aa})^{-1} b_1^a \right)^\top \xi_1 + \left(c_1 - (b_1^a)^\top (Q_1^{aa})^{-1} b_1^a \right).$$

The backward recursions (6b) for the Q -function coefficients now take the compact form

$$x_t = \mathbb{E}_t [f_{t+1}(\xi_{t+1}, x_{t+1})] \quad \forall t \in [T - 1]$$

with terminal condition $x_T = (P_T, 0, 0, 0, 0, 0)$. Using similar arguments as in the proof of Lemma 2.1, one readily verifies that the optimal value (5) of the control problem (3) can be represented compactly as $\mathbb{E}[f_1(\xi_1, x_1)]$. In view of the above recursion, this implies that the problem of computing the optimal value of (3) is equivalent to solving problem (MCCE), where x is identified with the deterministic vector x_T .

A similar reasoning as shown here can be used to reduce linear exponential quadratic regulator and risk-sensitive control problems [68] to instances of (MCCE). We omit the derivations for brevity.

Stochastic Control with Finite State and Action Spaces We use the same symbols as in the last example to denote states, actions and disturbances and consider the finite-horizon stochastic control problem

$$\begin{aligned} \min_{a,s} \quad & \mathbb{E} \left[\sum_{t=1}^{T-1} c_t(s_t, a_t) + c_T(s_T) \right] \\ \text{s.t.} \quad & a_t \in \mathcal{A}_t(\xi_t) \quad \forall t \in [T-1] \\ & s_{t+1} = \psi(s_t, a_t, \xi_{t+1}) \quad \forall t \in [T-1] \\ & a_t \text{ is } \mathcal{F}_t\text{-measurable} \quad \forall t \in [T-1]. \end{aligned} \quad (7)$$

All states and actions take values in finite sets \mathcal{S} and \mathcal{A} , respectively, while $\{\xi_t\}_{t \in \mathbb{N}}$ represents a generic (possibly serially dependent) stochastic process of disturbances taking values in \mathbb{R}^m . The system evolves according to a generic state transition function $\psi : \mathcal{S} \times \mathcal{A} \times \mathbb{R}^m \rightarrow \mathcal{S}$, and at each time t , the controller must select an action from a generic random and time-varying feasible set $\mathcal{A}_t(\xi_t) \subseteq \mathcal{A}$. The stage-costs $c_t : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$, $t \in [T-1]$, and the terminal cost $c_T : \mathcal{S} \rightarrow \mathbb{R}$ are also captured by generic functions.

If the disturbances are serially independent, then problem (7) reduces to a Markov decision process, a problem class that occupies central stage in operations research and machine learning [9]. Allowing the disturbances to display serial correlations enlarges this problem class significantly.

The time- t Q-function of problem (7) is defined by the Bellman-type recursion

$$Q_t(s, a) = c_t(s, a) + \mathbb{E}_t \left[\min_{a' \in \mathcal{A}_{t+1}(\xi_{t+1})} Q_{t+1}(\psi(s, a', \xi_{t+1}), a') \right] \quad (8)$$

with terminal condition $Q_T(s, a) = c_T(s)$. As in the previous example, $Q_t(s, a)$ is an \mathcal{F}_t -measurable random object for any fixed (s, a) , but we suppress its dependence on $\xi_{[t]}$ to avoid clutter. The optimal value of problem (7) can be expressed in terms of the time-1 Q-function as

$$\mathbb{E} \left[\min_{a \in \mathcal{A}_1(\xi_1)} Q_1(\xi_1, a) \right]. \quad (9)$$

For all $t \in [T]$ we define $x_t \in \mathbb{R}^d$ with $d = |\mathcal{S}| \times |\mathcal{A}|$ as the vectorized time- t Q-function with components $(x_t)_{(s,a)} = Q_t(s, a)$, $(s, a) \in \mathcal{S} \times \mathcal{A}$. Also, for all $t = 2, \dots, T$, we define $f_t : \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ through

$$(f_t(\xi_t, x_t))_{(s,a)} = c_{t-1}(s, a) + \min_{a' \in \mathcal{A}_t(\xi_t)} x_t(\psi(s, a, \xi_t), a') \quad \forall (s, a) \in \mathcal{S} \times \mathcal{A}.$$

In addition, we define $f_1 : \mathbb{R}^m \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ through $f_1(\xi_1, x_1) = \min_{a' \in \mathcal{A}_1(\xi_1)} x_1(s, a')$, where we use again the convention that $s = \xi_1$. With this notation, the Bellman recursions (8) can be written compactly as

$$x_t = \mathbb{E}_t[f_{t+1}(\xi_{t+1}, x_{t+1})] \quad \forall t \in [T-1],$$

with terminal condition $(x_T)_{(s,a)} = c_T(s)$ for every $s \in \mathcal{S}$ and $a \in \mathcal{A}$. The optimal value (9) of problem (7) thus coincides with $\mathbb{E}[f_1(\xi_1, x_1)]$, and computing this value is equivalent to solving problem (MCCE).

Distributionally Robust Policy Learning for Offline Contextual Bandits The contextual bandit problem is a sequential decision problem under uncertainty. In each round, an agent first observes a context $c \in \mathcal{C}$, then selects an action $a \in \mathcal{A}$ (metaphorically thought of as an arm of a slot machine) and finally incurs a random cost $y_a \in \mathcal{Y}$ associated with the chosen action. The distribution $\mathbb{Q}_{y_a|c}$ of y_a conditional on c is unknown to the agent. However, the pairs (c, y) with $y = (y_a)_{a \in \mathcal{A}}$ are known to be independent and identically distributed across rounds. The crux of multi-armed bandit problems is that, in each round, the agent observes only the cost of the chosen action but not those of the other actions. Thus, each action $a \in \mathcal{A}$ must be selected repeatedly for each given c in order to learn the distribution $\mathbb{Q}_{y_a|c}$.

Consider now a parametric family of policies π_θ , $\theta \in \Theta$, which map any context $c \in \mathcal{C}$ to a probability distribution over the actions $a \in \mathcal{A}$. We will henceforth assume that \mathcal{A} is finite. Thus, given a context c , the policy π_θ selects action a with probability $\pi_\theta(a|c)$. The agent aims to find $\theta \in \Theta$ such that π_θ minimizes the average cost. If the joint distribution $\mathbb{Q}_{(c,y)}$ of c and y was known, this could be achieved by solving

$$\min_{\theta \in \Theta} \int_{\mathcal{C} \times \mathcal{Y}} \sum_{a \in \mathcal{A}} y_a \pi_\theta(a|c) d\mathbb{Q}_{(c,y)}(c, y) = \min_{\theta \in \Theta} \int_{\mathcal{C}} \sum_{a \in \mathcal{A}} \pi_\theta(a|c) \int_{\mathcal{Y}} y_a d\mathbb{Q}_{y|c}(y) d\mathbb{Q}_c(c).$$

In off-policy learning, however, $\mathbb{Q}_{(c,y)}$ is unknown, and the agent has only access to context, action and cost data collected under a fixed behavioral policy. Any estimator $\mathbb{P}_{(c,y)}$ constructed from this data invariably differs from $\mathbb{Q}_{(c,y)}$. To combat the detrimental effects of estimation errors, it has been proposed to solve the following distributionally robust variant of the off-policy learning problem [59].

$$\min_{\theta \in \Theta} \sup_{\substack{\mathbb{Q}_c \in \mathcal{P}(\mathcal{C}): \\ d(\mathbb{Q}_c, \mathbb{P}_c) \leq r_c}} \int_{\mathcal{C}} \sum_{a \in \mathcal{A}} \pi_\theta(a|c) \sup_{\substack{\mathbb{Q}_{y|c} \in \mathcal{P}(\mathcal{Y}): \\ d(\mathbb{Q}_{y|c}, \mathbb{P}_{y|c}) \leq r_y}} \int_{\mathcal{Y}} y_a d\mathbb{Q}_{y|c}(y) d\mathbb{Q}_c(c) \quad (10)$$

Here, we use the 2-Wasserstein metric $d(\cdot, \cdot)$ to measure distances between probability distributions. Note that the distributionally robust off-policy learning problem replaces the expectations with respect to the unknown distributions \mathbb{Q}_c and $\mathbb{Q}_{y|c}$ by worst-case expectations over all distributions in an r_c -neighborhood around \mathbb{P}_c and an r_y -neighborhood around $\mathbb{P}_{y|c}$, respectively. By [46, Proposition 6.20], the maximization problem over $\mathbb{Q}_{y|c}$ in (10) can be solved in closed form, and its supremum equals $\int_{\mathcal{Y}} y_a d\mathbb{P}_{y|c}(y) + r_y$. In addition, by [46, Theorem 4.18], the maximization problem over \mathbb{Q}_c can be dualized and thus be converted to a minimization problem over a scalar dual decision variable λ . Hence, problem (10) can be recast as

$$\min_{\theta \in \Theta, \lambda \geq 0} \int_{\mathcal{C}} \sup_{u \in \mathcal{C}} \left(\sum_{a \in \mathcal{A}} \pi_\theta(a|u) \int_{\mathcal{Y}} y_a d\mathbb{P}_{y|c=u}(y) + r_y + r_c^2 \lambda - \lambda \|u - c'\|_2^2 \right) d\mathbb{P}_c(c'). \quad (11)$$

Note that (11) involves a maximization problem over all contexts $u \in \mathcal{C}$, which may be difficult to solve because the expected reward of action a conditional on $c = u$ is a measurable function of u and because maximizing a generic measurable function is hard. As a remedy, the supremum over all $u \in \mathcal{C}$ can be approximated by a softmax function involving the uniform distribution \mathbb{P}_u on \mathcal{C} . Thus, (11) simplifies to

$$\min_{\theta \in \Theta, \lambda \geq 0} \int_{\mathcal{C}} \frac{1}{\mu} \log \left(\int_{\mathcal{C}} \exp \left(\mu \left(\sum_{a \in \mathcal{A}} \pi_\theta(a|u) \int_{\mathcal{Y}} y_a d\mathbb{P}_{y|c=u}(y) + r_y + r_c^2 \lambda - \lambda \|u - c'\|_2^2 \right) \right) d\mathbb{P}_u(u) \right) d\mathbb{P}_c(c'), \quad (12)$$

where the parameter $\mu > 0$ controls the smoothness of the softmax function. For further details see [59]. Problem (12) is readily recognized as an instance of (MCCO) with $T = 3$, $x = (\theta, \lambda)$, $\mathcal{X} = \Theta \times \mathbb{R}_+$, $\xi_1 = c'$, $\xi_2 = u$ and $\xi_3 = (c', u, y)$. The integrands are given by $f_1(\xi_1, x_1) = \frac{1}{\mu} \log(x_1)$, $f_2(\xi_2, x_2) = \exp(\mu x_2)$ and

$$f_3(\xi_3, x) = \sum_{a \in \mathcal{A}} \pi_\theta(a|u) y_a + r_y + r_c^2 \lambda - \lambda \|u - c'\|_2^2.$$

Variants of the distributionally robust off-policy learning problem (10) use the 2-Wasserstein distance with a Kullback-Leibler regularization term [3] or the Sinkhorn divergence [65] instead of the plain 2-Wasserstein distance to measure the discrepancy of different distributions. These alternative formulations also give rise to instances of MCCO with $T = 3$. For further details we refer to [65, Appendix D].

Nested Risk Measures Consider a stream of random costs ξ_t , for $t \in [T]$, that are p -integrable for some exponent $p \geq 1$ (i.e., $\mathbb{E}[\|\xi_t\|_2^p] < \infty$ for all $t \in [T]$). In this situation, it makes sense to assign the total cost $\sum_{t=1}^T \xi_t$ a risk that depends on the temporal structure of the underlying stochastic process. This can be done by using a nested dynamic risk measure $\rho = \rho_0 \circ \dots \circ \rho_{T-1}$, where $\rho_t : \mathcal{L}_{t+1}^p \rightarrow \mathcal{L}_t^p$ is a conditional risk mapping and \mathcal{L}_t^p represents the family of all p -integrable functions of the history $\xi_{[t]}$ of costs up to time t . The defining properties of a conditional risk mapping are translation invariance, convexity and monotonicity; see [55] for details. By exploiting these properties, the risk of the total cost can then be represented as

$$\rho \left[\sum_{t=1}^T \xi_t \right] = \rho_0 \left[\xi_1 + \rho_1 \left[\xi_2 + \dots + \rho_{T-1} \left[\xi_T \right] \dots \right] \right].$$

A widely studied choice is to set ρ_t to a conditional entropic risk measure, defined as

$$\rho_t(\xi_{t+1}) = \frac{1}{\mu_t} \log \left(\mathbb{E}_t[\exp(-\mu_t \xi_{t+1})] \right),$$

where $\mu_t > 0$ represents the risk aversion level at time t . Notably, the risk aversion level may vary over time. Such scenarios arise, for instance, in pension planning, where the goal is to assess the risk of a pension contract. In this context, tolerance for uncertainty in cash flows typically decreases over time as retirement approaches. It is now easy to see that evaluating $\exp(\mu_0 \rho(\sum_{t=1}^T \xi_t))$ reduces to an instance of (MCCE), with $f_t(\xi_t, x_t) = \exp(-\mu_{t-1}(\xi_t + \mu_t^{-1} \log(x_t)))$ for $t \in [T-1]$ and $f_T(\xi_T, x) = \exp(-\mu_{T-1} \xi_T)$. Similarly, optimization problems involving nested risk measures can often be reformulated as instances of (MCCO). Such problems arise, for example, in Bayesian distributionally robust optimization [58] and in the analysis of Bayesian risk Markov decision processes [48], which emerge in betting or control problems [48, 67]. Solving Bayesian risk Markov decision processes presents significant challenges due to the complexity of computing nested risk functionals [48]. To address these difficulties, approximation techniques are commonly used. However, traditional estimation and optimization methods can be computationally expensive. This paper proposes cost-efficient methods to mitigate this issue and resolves an open question raised in [48].

3 Numerical Solution of MCCE Problems

We begin by presenting the fundamental assumptions that underpin this paper.

Assumption 3.1 (Basic Regularity Conditions).

- (i) The integrand $f_t(\xi_t, x_t)$ is Borel-measurable in ξ_t , and $\mathbb{E}[\|f_t(\xi_t, 0)\|_2] < \infty$ for every $t \in [T]$.
- (ii) The feasible set \mathcal{X} is convex and compact. Thus, it has a finite diameter $D_{\mathcal{X}}$.

Neither F nor the underlying integrands $f_t, t \in [T]$, are assumed to display any convexity properties. However, we assume these functions to be Lipschitz continuous, which is common practice in sample complexity analysis (see e.g., [57]). Sometimes, we additionally assume the integrands to be smooth.

Assumption 3.2 (Lipschitz Continuity). $f_t(\xi_t, x_t)$ is L_t -Lipschitz continuous in x_t uniformly for all ξ_t , i.e.,

$$\|f_t(\xi_t, x_t) - f_t(\xi_t, y_t)\|_2 \leq L_t \|x_t - y_t\|_2 \quad \forall x_t, y_t \in \mathbb{R}^{d_t}, \forall \xi_t \in \mathbb{R}^{m_t}.$$

Assumption 3.3 (Smoothness). $f_t(\xi_t, x_t)$ is S_t -smooth in x_t for across all ξ_t , i.e.,

$$\|\nabla_{x_t} f_t(\xi_t, x_t) - \nabla_{x_t} f_t(\xi_t, y_t)\|_2 \leq S_t \|x_t - y_t\|_2 \quad \forall x_t, y_t \in \mathbb{R}^{d_t}, \forall \xi_t \in \mathbb{R}^{m_t}.$$

Note that if Assumption 3.3 holds, then the fundamental theorem of calculus implies that

$$\left\| f_t(\xi_t, x_t) - f_t(\xi_t, y_t) - \nabla_{x_t} f_t(\xi_t, y_t)^\top (x_t - y_t) \right\|_2 \leq \frac{S_t}{2} \|x_t - y_t\|_2^2 \quad \forall x_t, y_t \in \mathbb{R}^{d_t}, \forall \xi_t \in \mathbb{R}^{m_t}. \quad (13)$$

Assumptions 3.1 and 3.2 ensure that (MCCE) and (MCCO) are well-defined. Indeed, they imply that all conditional expectations in the definition of $F(x)$ exist and are finite, which is an immediate consequence of Lemma C.2 in the appendix. Thus, problem (MCCE) is well-defined. In addition, Assumptions 3.1 and 3.2 imply that $F(x)$ is Lipschitz continuous with Lipschitz constant $\prod_{t=1}^T L_t$. This is an immediate corollary of Lemma C.3 in the appendix. By Assumption 3.1 (ii), we can thus invoke Weierstrass' extreme value theorem to conclude that the minimum in (MCCO) is attained. Hence, problem (MCCO) is well-defined.

Problem (MCCE) is challenging due to the inherent difficulty of high-dimensional integration. Consequently, attempting to compute $F(x)$ exactly is overly ambitious. In this section we thus pursue the more modest but achievable goal of approximating $F(x)$ with an estimator $\widehat{F}(x)$ constructed from samples. Specifically, we will construct an SAA estimator in Section 3.1 and an MLMC estimator in Section 3.2. We will show that both estimators converge to $F(x)$ in mean squared error and in probability, uniformly over all $x \in \mathcal{X}$. Moreover, we will demonstrate that the MLMC estimator outperforms the SAA estimator in terms of both mean squared error-based and high-probability scenario complexity.

3.1 SAA Estimator

Perhaps the most straightforward estimator for $F(x)$ is the SAA estimator defined below.

Definition 3.4 (SAA Estimator). *We construct a forest of n_1 scenario trees as follows. The root nodes of the trees are indexed by $i_1 \in [n_1]$. The children of i_1 are indexed by $i_{[2]} = (i_1, i_2)$ for $i_2 \in [n_2]$, and the children of $i_{[2]}$ are indexed by $i_{[3]} = (i_{[2]}, i_3)$ for $i_3 \in [n_3]$, etc. Thus, the multi-index $i_{[t]} = (i_1, i_2, \dots, i_t)$ of a stage- t node encodes the tree that accommodates the node $i_{[t]}$ as well as the path from the root node in that tree to the node $i_{[t]}$; see Figure 1. The SAA estimator for $F(x)$ is then given by*

$$\widehat{F}(x) := \frac{1}{n_1} \sum_{i_1=1}^{n_1} \widehat{F}_{i_1}(x), \quad (14a)$$

where the term corresponding to the i_1 -th scenario tree is set to

$$\widehat{F}_{i_1}(x) := f_1 \left(\xi_1^{i_1}, \frac{1}{n_2} \sum_{i_2=1}^{n_2} f_2 \left(\xi_2^{i_{[2]}}, \dots, \frac{1}{n_T} \sum_{i_T=1}^{n_T} f_T \left(\xi_T^{i_{[T]}}, x \right) \right) \dots \right) \quad \forall i_1 \in [n_1], \quad (14b)$$

and where the sample $\xi_t^{i_{[t]}}$ associated with node $i_{[t]}$ is constructed by the forward recursion

$$\xi_t^{i_{[t]}} \stackrel{i.i.d.}{\sim} \mathbb{P}_{\xi_t} \left(\cdot \mid \xi_1 = \xi_1^{i_1}, \dots, \xi_{t-1} = \xi_{t-1}^{i_{[t-1]}} \right) \quad \forall i_1 \in [n_1], \dots, \forall i_t \in [n_t], \forall t \in [T].$$

We introduce some key notations used throughout the paper. Given any $t \in \mathbb{N}_0$ with $t+1 \in [T]$ and any Borel-measurable function $\psi(\xi_{[t+1]})$, we henceforth use $\widehat{\mathbb{E}}_t[\psi(\xi_{[t+1]})]$ as shorthand for the sample average of $\psi(\xi_{[t+1]})$ over n_{t+1} samples ξ_{t+1}^i , $i \in [n_{t+1}]$, drawn independently from the marginal distribution of ξ_1 (if $t = 0$) or the conditional distribution of ξ_{t+1} given $\xi_{[t]}$ (if $t > 0$). That is, we set

$$\widehat{\mathbb{E}}_t[\psi(\xi_{[t+1]})] := \frac{1}{n_{t+1}} \sum_{i=1}^{n_{t+1}} \psi(\xi_{[t]}, \xi_{t+1}^i). \quad (15)$$

Below, we sometimes use $\widehat{\mathbb{E}}[\cdot]$ as shorthand for $\widehat{\mathbb{E}}_0[\cdot]$. Using these conventions, the SAA estimator of Definition 3.4 can be recast more compactly as

$$\widehat{F}(x) = \widehat{\mathbb{E}} \left[f_1 \left(\xi_1, \widehat{\mathbb{E}}_1 \left[f_2 \left(\xi_2, \dots, \widehat{\mathbb{E}}_{T-1} \left[f_T \left(\xi_T, x \right) \dots \right] \right) \dots \right] \right) \right].$$

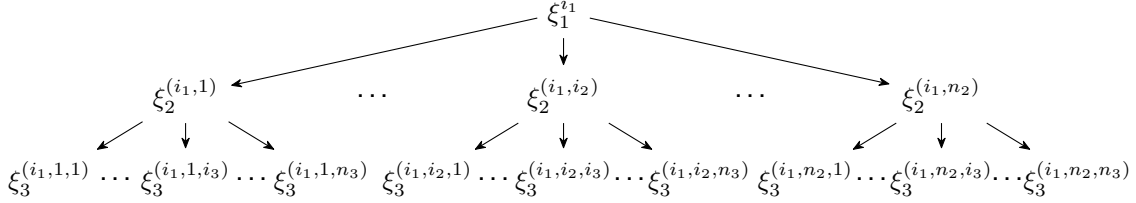


Figure 1: Visualization of the i_1 -th scenario tree underlying the SAA estimator when $T = 3$.

Note that the SAA estimator of Definition 3.4 requires $\mathbb{C}(\widehat{F}(x)) = \prod_{t=1}^T n_t$ scenarios. In the following we will prove that, as the sample sizes n_t , $t \in [T]$, tend to infinity, the estimator $\widehat{F}(x)$ converges in mean squared error and in probability to $F(x)$ uniformly across all $x \in \mathcal{X}$. To this end, we first rewrite the mean squared error of $\widehat{F}(x)$ with respect to $F(x)$ as the sum of the squared bias and the variance of $\widehat{F}(x)$.

$$\text{MSE}(\widehat{F}(x)) = \mathbb{E} \left[(\widehat{F}(x) - F(x))^2 \right] = \left(\mathbb{E}[\widehat{F}(x)] - F(x) \right)^2 + \mathbb{V}(\widehat{F}(x)) \quad (16)$$

This formula shows that, for any fixed $\epsilon > 0$, we can drive the mean squared error of $\widehat{F}(x)$ below ϵ^2 by ensuring that $|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \epsilon/\sqrt{2}$ and $\mathbb{V}(\widehat{F}(x)) \leq \epsilon^2/2$. In the following, we construct bounds on the bias and the variance of the SAA estimator and analyze their dependence on the sample sizes n_t , $t \in [T]$. These bounds are obtained under the assumption that all integrands have finite variance.

Assumption 3.5. *Each integrand $f_t(\xi_t, x_t)$ has a finite conditional variance in the sense that*

$$\sigma_t^2 := \sup_{x_t \in \mathbb{R}^{d_t}} \text{ess sup} \mathbb{E}_{t-1} \left[\|f_t(\xi_t, x_t) - \mathbb{E}_{t-1}[f_t(\xi_t, x_t)]\|_2^2 \right] < \infty \quad \forall t \in [T].$$

We are now ready to derive an upper bound on the bias of the SAA estimator. This upper bound is expressed in terms of product Lipschitz constants of the form $L_{[t]} := \prod_{s=1}^t L_s$ for $t \in [T]$ and $L_{[0]} = 1$.

Lemma 3.6. *If Assumptions 3.1, 3.2 and 3.5 hold, then the bias of the SAA estimator satisfies*

$$\left| \mathbb{E}[\widehat{F}(x)] - F(x) \right| \leq \begin{cases} \sum_{t=1}^{T-1} \frac{L_{[t]} \sigma_{t+1}}{\sqrt{n_{t+1}}} & \text{if Assumption 3.3 does not hold,} \\ \sum_{t=1}^{T-1} \frac{L_{[t-1]} S_t \sigma_{t+1}}{2n_{t+1}} & \text{if Assumption 3.3 holds.} \end{cases}$$

Lemma 3.6 shows that the bias of the SAA estimator decays faster with the sample sizes if the integrands are smooth (*i.e.*, if Assumption 3.3 holds). If $T = 2$, then Lemma 3.6 recovers the bounds of [39, Lemma 3.1] for conditional stochastic optimization. Next, we bound the variance of the SAA estimator.

Lemma 3.7. *If Assumptions 3.1, 3.2 and 3.5 hold, then the variance of the SAA estimator satisfies*

$$\mathbb{V}(\widehat{F}(x)) \leq \frac{1}{n_1} \left(\sum_{t=1}^{T-1} \frac{L_{[t]} \sigma_{t+1}}{\sqrt{n_{t+1}}} + \sigma_1 \right)^2.$$

Remark 3.8. *The variance bound of Lemma 3.7 can be strengthened when the integrands are smooth. However, the improved bound is cumbersome and does not decay faster with the sample sizes. To keep our analysis transparent, we therefore use the simpler bound of Lemma 3.7 even if the integrands are smooth.*

Combining Lemmas 3.6 and 3.7 with (16) shows that the mean squared error of $\widehat{F}(x)$ satisfies

$$\text{MSE}(\widehat{F}(x)) \leq \begin{cases} \left(\sum_{t=1}^{T-1} \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}} \right)^2 + \frac{1}{n_1} \left(\sum_{t=1}^{T-1} \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}} + \sigma_1 \right)^2 & \text{if Assumption 3.3 does not hold,} \\ \left(\sum_{t=1}^{T-1} \frac{L_{[t-1]}S_t\sigma_{t+1}^2}{2n_{t+1}} \right)^2 + \frac{1}{n_1} \left(\sum_{t=1}^{T-1} \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}} + \sigma_1 \right)^2 & \text{if Assumption 3.3 holds.} \end{cases}$$

We define the mean squared error-based scenario complexity of any estimator for $F(x)$ as the total number of scenarios needed in order to ensure a root mean squared error of at most ϵ uniformly across all $x \in \mathcal{X}$. The above bound enables us to characterize the mean squared error-based scenario complexity of the SAA estimator. To this end, recall that the construction of $\widehat{F}(x)$ requires a total of $\prod_{t=1}^T n_t$ scenarios. The next theorem shows that setting the sample sizes to $n_t = \mathcal{O}(1/\epsilon^2)$ for nonsmooth integrands and $n_t = \mathcal{O}(1/\epsilon)$ for smooth integrands, for all $t = 2, \dots, T$, ensures that the bias is bounded by $\epsilon/\sqrt{2}$. Moreover, choosing $n_1 = \mathcal{O}(1/\epsilon^2)$ guarantees that the standard deviation is bounded by $\epsilon/\sqrt{2}$. This implies that the mean squared error-based scenario complexity of the SAA estimator is at most $\mathcal{O}(\epsilon^{-2T})$ for nonsmooth integrands and $\mathcal{O}(\epsilon^{-(T+1)})$ for smooth integrands. The proof of this theorem is straightforward and therefore omitted.

Theorem 3.9 (Mean Squared Error-Based Scenario Complexity of the SAA Estimator). *Suppose that Assumptions 3.1, 3.2 and 3.5 hold. Assume also that $n_1 = \lceil 1 + 2\sqrt{2}\sigma_1/\epsilon + 2\sigma_1^2/\epsilon^2 \rceil$ and that*

$$n_t = \begin{cases} \left\lceil \left(\frac{\sqrt{2}L_{[t-1]}\sigma_t(T-1)}{\epsilon} \right)^2 \right\rceil & \text{if Assumption 3.3 does not hold,} \\ \left\lceil \frac{\sqrt{2}L_{[t-2]}S_{t-1}\sigma_t^2(T-1)}{2\epsilon} \right\rceil & \text{if Assumption 3.3 holds,} \end{cases}$$

for all $t = 2, \dots, T$. Then, the SAA estimator achieves a root mean squared error of at most ϵ uniformly across all $x \in \mathcal{X}$. Thus, the mean squared error-based scenario complexity of the SAA estimator amounts at most to $\mathcal{O}(\epsilon^{-2T})$ if Assumption 3.3 does not hold and to $\mathcal{O}(\epsilon^{-(T+1)})$ if Assumption 3.3 holds.

Theorem 3.9 shows that although smoothness improves the exponent in the accuracy parameter ϵ , the scenario complexity of the SAA estimator still exhibits an exponential dependence on the horizon T in both the smooth and nonsmooth cases. This highlights a fundamental limitation of nested estimation. Achieving high accuracy becomes increasingly sample-intensive as T grows. It is important to note that these guarantees hold only *in expectation*; that is, the root mean squared error of the SAA estimator is at most ϵ on average over repeated independent simulation runs. To obtain stronger guarantees that hold for a *single* run with high confidence, we next turn to high-probability bounds. These results are based on concentration inequalities and therefore require an additional standard sub-Gaussianity assumption.

Definition 3.10 (Sub-Gaussian Random Vector). *An m -dimensional random vector ξ with finite mean is called sub-Gaussian with variance proxy $\zeta^2 > 0$ if*

$$\mathbb{E}[\exp(\lambda^\top(\xi - \mathbb{E}[\xi]))] \leq \exp(\|\lambda\|_2^2 \zeta^2 / 2) \quad \forall \lambda \in \mathbb{R}^m.$$

Assumption 3.11 (Sub-Gaussianity of $\widehat{F}_{i_1}(x)$). *For each $i_1 \in [n_1]$, the estimator $\widehat{F}_{i_1}(x)$ corresponding to the i_1 -th scenario tree in Definition 3.4 is sub-Gaussian with variance proxy $\zeta^2 > 0$ for any fixed $x \in \mathcal{X}$.*

Assumption 3.11 is hard to verify in general, but Remark 3.12 below outlines conditions when it holds.

Remark 3.12. *Assumption 3.11 holds under either of the following easily verifiable conditions.*

- (i) If ξ_t is sub-Gaussian with variance proxy ζ_t^2 and $f_t(\xi_t, x_t)$ is jointly L_t -Lipschitz continuous in ξ_t and x_t for all $t \in [T]$, then Assumption 3.5 is automatically satisfied [64, Exercise 2.5]. In addition, one readily verifies that $\widehat{F}_{i_1}(x)$ is sub-Gaussian. Indeed, $\widehat{F}_{i_1}(x)$ can be constructed by iteratively averaging and concatenating sub-Gaussian random variables and transforming them under Lipschitz continuous functions. By Lemma C.5 in the appendix, all of these operations preserve sub-Gaussianity.
- (ii) If there exists a constant $M > 0$ such that $|f_1(\xi_1, x_1)| \leq M$ for all $\xi_1 \in \mathbb{R}^{m_1}$ and $x_1 \in \mathbb{R}^{d_1}$, then [64, Example 2.4] readily implies that $\widehat{F}_{i_1}(x)$ is sub-Gaussian with variance proxy $\zeta^2 = 4M^2$.

We now establish the uniform convergence of the SAA estimator to the true objective function.

Lemma 3.13. *Suppose that Assumptions 3.1, 3.2, 3.5 and 3.11 hold and that $\epsilon > 0$ is a prescribed tolerance. Assume also that the stagewise sample sizes satisfy*

$$n_t = \begin{cases} \left\lceil \left(\frac{4L_{[t-1]}\sigma_t(T-1)}{\epsilon} \right)^2 \right\rceil & \text{if Assumption 3.3 does not hold,} \\ \left\lceil \frac{2L_{[t-2]}S_{t-1}\sigma_t^2(T-1)}{\epsilon} \right\rceil & \text{if Assumption 3.3 holds,} \end{cases}$$

for all $t = 2, \dots, T$. Then, we have

$$\mathbb{P} \left(\sup_{x \in \mathcal{X}} |\widehat{F}(x) - F(x)| > \epsilon \right) \leq 2 \lceil 8L_{[T]}D_{\mathcal{X}}/\epsilon + 1 \rceil^d \exp \left(- \frac{n_1\epsilon^2}{32\zeta^2} \right).$$

Recall that ζ^2 is the variance proxy of $\widehat{F}_{i_1}(x)$, d and $D_{\mathcal{X}}$ are the dimension and the diameter of \mathcal{X} , respectively, and $L_{[T]}$ is the Lipschitz constant of $F(x)$. Lemma 3.13 establishes a high probability bound on the worst possible difference between the SAA estimator and the true objective function $F(x)$ uniformly across all $x \in \mathcal{X}$. It asserts that the probability that this difference exceeds any given threshold $\epsilon > 0$ decays exponentially fast with ϵ and with the sample size n_1 .

The uniform convergence result of Lemma 3.13 not only characterizes the number of scenarios needed to approximately solve the estimation problem (MCCE), which is the central focus of this section, but also inspires a straightforward method to approximately solve the optimization problem (MCCO). Indeed, instead of minimizing the function $F(x)$, whose values and gradients are difficult to compute exactly, one may minimize the SAA estimator $\widehat{F}(x)$, that is, one may solve an instance of (1). Below, we will use Lemma 3.13 to show that the exact minimizers of the SAA problem (1) yield near-optimal solutions to (MCCO) with high probability. The practicality of this approach ultimately depends on the tractability of (1).

Remark 3.14. *Global optimization of the SAA estimator $\widehat{F}(x)$ is tractable in settings where $\widehat{F}(x)$ is convex and admits an efficient separation oracle. A sufficient condition for convexity of both $F(x)$ and $\widehat{F}(x)$ is that the integrand $f_t(\xi_t, x_t)$ is convex in x_t for every $t \in [T]$ and non-decreasing in x_t for every $t \in [T-1]$. These conditions ensure that the global minimum of (1) can be found efficiently. However, they are restrictive.*

We now show that the minimizers of (1) are near-optimal in (MCCO) with high probability.

Corollary 3.15. *Assume that all conditions of Lemma 3.13 hold for some $\epsilon > 0$, and let x^* and \widehat{x}^* be minimizers of (MCCO) and the SAA problem (1), respectively. Then, we have*

$$\mathbb{P}(F(\widehat{x}^*) - F(x^*) > \epsilon) \leq 4 \lceil 8L_{[T]}D_{\mathcal{X}}/\epsilon + 1 \rceil^d \exp \left(- \frac{n_1\epsilon^2}{128\zeta^2} \right).$$

We define the high-probability scenario complexity of a method for solving (MCCO) as the total number of scenarios that this method needs to identify an ϵ -optimal solution of (MCCO) with probability $1 - \beta$.

Theorem 3.16 (High-Probability Scenario Complexity of the SAA Estimator). *Suppose that all conditions of Lemma 3.13 hold for some $\epsilon > 0$ and that*

$$n_1 = \left\lceil \frac{128\zeta^2}{\epsilon^2} (d \log(\lceil 8L_{[T]}D_X/\epsilon + 1 \rceil) + \log(4/\beta)) \right\rceil$$

for some $\beta \in (0, 1)$. Then any solution of the SAA problem (1) constitutes an ϵ -optimal solution of (MCCO) with probability at least $1 - \beta$. Thus, the high-probability scenario complexity of the SAA estimator equals $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-2T})$ if Assumption 3.3 does not hold and $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-(T+1)})$ if Assumption 3.3 holds.

Theorem 3.16 generalizes [39, Corollary 4.2] to general $T > 2$. It shows that, like the mean squared error-based scenario complexity, the high-probability scenario complexity of the SAA method grows exponentially with T . This aligns with the observation in multistage stochastic programming that SAA suffers from a curse of dimensionality [52, 56]. Note also that the exponential growth rate is reduced by a factor of two if the integrands are smooth in the sense of Assumption 3.3. In contrast to Theorem 3.9, the bound on n_1 depends now on the dimension d of the decision variable and the variance proxy ζ^2 of the tree-wise SAA estimator. This captures the cost of enhancing guarantees from convergence in expectation to convergence with high probability. If the SAA estimator $\widehat{F}(x)$ further satisfies the Hölderian error bound condition proposed in [39, Assumption 4.1], then the bound on n_1 in Theorem 3.16 can be made independent of the dimension d . In the special case where $T = 2$, the scenario complexity bounds of Theorem 3.16 reduce to the known bounds for conditional stochastic optimization reported in [39]. These bounds scale quartically with the inverse accuracy $1/\epsilon$ for nonsmooth integrands and cubically for smooth integrands, demonstrating the efficiency of the SAA method when T is small. However, for larger values of T , the method becomes increasingly inefficient, underscoring the need for more advanced and scalable approaches to solve (MCCO).

3.2 MLMC Estimator

We now develop a recursive MLMC estimator for the nested expectation problem (MCCE). Our approach generalizes the unbiased MLMC method of [11] to arbitrary depths $T \geq 2$, but introduces a controlled bias to simultaneously guarantee finite variance and computational cost under weaker conditions (e.g., Lipschitz continuity). This contrasts with the unbiased recursive MLMC estimator of [61], which requires smoothness to ensure a finite variance. To this end, we need additional notation and definitions. For any $t + 1 \in [T]$ and Borel-measurable function $\psi(\xi_{[t+1]})$, we continue to use the empirical conditional expectation operator $\widehat{\mathbb{E}}_t[\psi(\xi_{[t+1]})]$ with n_{t+1} samples defined in (15). We further introduce empirical conditional expectation operators that run over 2^ℓ samples as well as over the even or odd indices in $[2^\ell]$. In other words, we define

$$\begin{aligned} \widehat{\mathbb{E}}_t^\ell[\psi(\xi_{[t+1]})] &:= \frac{1}{2^\ell} \sum_{i=1}^{2^\ell} \psi(\xi_{[t]}, \xi_{t+1}^i) & \forall \ell \in \mathbb{N}_0, \\ \widehat{\mathbb{E}}_t^{\ell,o}[\psi(\xi_{[t+1]})] &:= \frac{1}{2^{\ell-1}} \sum_{i=1}^{2^{\ell-1}} \psi(\xi_{[t]}, \xi_{t+1}^{2i-1}) & \forall \ell \in \mathbb{N}, \\ \widehat{\mathbb{E}}_t^{\ell,e}[\psi(\xi_{[t+1]})] &:= \frac{1}{2^{\ell-1}} \sum_{i=1}^{2^{\ell-1}} \psi(\xi_{[t]}, \xi_{t+1}^{2i}) & \forall \ell \in \mathbb{N}, \end{aligned}$$

where the samples ξ_{t+1}^i , $i \in [2^\ell]$, are drawn independently from the marginal distribution of ξ_1 (if $t = 0$) or the conditional distribution of ξ_{t+1} given $\xi_{[t]}$ (if $t > 0$). By construction, we thus have

$$\widehat{\mathbb{E}}_t^\ell[\psi(\xi_{[t+1]})] = \frac{1}{2} \widehat{\mathbb{E}}_t^{\ell,o}[\psi(\xi_{[t+1]})] + \frac{1}{2} \widehat{\mathbb{E}}_t^{\ell,e}[\psi(\xi_{[t+1]})]. \quad (17)$$

Definition 3.17 (MLMC Estimator). *For any $t \in [T - 1]$, select a rate parameter $r_t \in (0, 1)$ and truncation point $M_t \in \mathbb{N}$, and let the random variable λ_t follow the truncated geometric distribution $\text{Geo}(r_t | M_t)$. Assume that λ_t is independent of all other random objects, and set $q_t(\ell) := \mathbb{P}(\lambda_t = \ell)$ for $\ell = 0, \dots, M_t$ and $t \in [T - 1]$. We then construct a sequence of estimators $\widehat{H}_t(x)$, $t \in [T]$, using the following backward recursion. We initialize the recursion by setting $\widehat{H}_T(x) := f_T(\xi_T, x)$. Next, for any $t \in [T - 1]$ we set*

$$\widehat{H}_t(x) := \frac{1}{q_t(\lambda_t)} \left(\widehat{h}_t^{\lambda_t}(x) - \frac{1}{2} \widehat{h}_t^{\lambda_t, e}(x) - \frac{1}{2} \widehat{h}_t^{\lambda_t, o}(x) \right), \quad (18)$$

where

$$\widehat{h}_t^\ell(x) := f_t\left(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}(x)]\right) \quad \forall \ell = 0, \dots, M_t$$

as well as

$$\widehat{h}_t^{\ell, e}(x) := f_t\left(\xi_t, \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}(x)]\right) \quad \text{and} \quad \widehat{h}_t^{\ell, o}(x) := f_t\left(\xi_t, \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}(x)]\right) \quad \forall \ell \in [M_t].$$

We also set $\widehat{h}_t^{0, e}(x) := \widehat{h}_t^{0, o}(x) := 0$. The MLMC estimator for $F(x)$ is then defined as $\widehat{F}(x) := \widehat{\mathbb{E}}[\widehat{H}_1(x)]$.

The MLMC estimator is built on a forest of n_1 scenario trees that display a random branching structure with 2^{λ_t} branches per stage- t node, where λ_t is sampled from a truncated geometric distribution. For this reason we will sometimes refer to λ_t as the log-branching factor of stage t . If $M_t = \infty$ for all $t \in [T - 1]$ and the integrands f_t map to \mathbb{R} , then our MLMC estimator reduces to the untruncated MLMC estimator proposed in [61]. If additionally $T = 2$, then it further reduces to the estimator discussed in [66, § 3.2].

Working with random (log-)branching factors offers significant advantages. We will show that the bias of the MLMC estimator is of the same order as that of the SAA estimator with a large fixed branching factor 2^{M_t} at each stage $t \in [T - 1]$. Since the MLMC estimator is constructed on a scenario tree with a random branching factor 2^{λ_t} that never exceeds but is typically much smaller than 2^{M_t} at every stage, it achieves this bias using substantially fewer samples in total. As λ_t concentrates near zero, the branching factor 2^{λ_t} often equals 1, so the MLMC estimator requires only a single sample in most branches. Consequently, the scenario tree grows more slowly with the time horizon T , making the evaluation of the MLMC estimator significantly more efficient than that of the SAA estimator with comparable bias.

We now prove that the MLMC estimator requires fewer samples than the SAA estimator to achieve a mean squared error no greater than ϵ^2 . To do so, we derive upper bounds on the variance, bias, and expected sampling cost of the MLMC estimator in Lemmas 3.19, 3.20, and 3.22, respectively. These bounds depend on the estimator's hyperparameters, that is, the number n_1 of scenario trees as well as the truncation points M_t and the rates r_t that define the truncated geometric distributions of the log-branching factors λ_t , $t \in [T - 1]$. We then tune these hyperparameters to ensure that the bias and the variance of the MLMC estimator are bounded by $\epsilon/\sqrt{2}$. This calibration guarantees an overall mean squared error of at most ϵ^2 , enabling a fair comparison of the expected sampling costs of the optimally tuned SAA and MLMC estimators.

From this point onward, we denote the p -th moment of $\widehat{H}_t(x)$ conditional on time $t - 1$ information by

$$\mu_t^p(x) := \mathbb{E}_{t-1}[\|\widehat{H}_t(x)\|_p^p], \quad (19)$$

where $p \geq 1$. To analyze the MLMC estimator, we also define the conditional moment bound of order p as

$$\bar{\mu}_T^p := \sup_{x \in \mathbb{R}^d} \text{ess sup} \mathbb{E}_{T-1}[\|f_T(\xi_T, x)\|_p^p]. \quad (20)$$

By using Jensen's inequality, one readily verifies that if $\bar{\mu}_T^p < \infty$, then $\bar{\mu}_T^{p'} < \infty$ for every $p' \in [1, p]$. The following assumption is crucial for our subsequent results.

Assumption 3.18. *The conditional moment bound $\bar{\mu}_T^p$ of the integrand $f_T(\xi_T, x)$ is finite up to order $p = 2$ if Assumption 3.3 does not hold and up to order $p = 2^T$ if Assumption 3.3 holds.*

The following lemma establishes bounds on $\mu_t^2(x)$ and $\mu_t^{2^t}(x)$ for all $t \in [T]$, corresponding to the cases of nonsmooth and smooth integrands, respectively. In particular, for $t = 1$, these bounds can be used to derive upper bounds on the unconditional variance of the MLMC estimator in both settings. For any $t \in [T]$ and $s \in [t+1]$ we henceforth set $L_{[s:t]} := \prod_{i=s}^t L_i$ if $s < t+1$; $:= 1$ if $s = t+1$; see also Lemma C.3 in the appendix. In particular, note that $L_{[1:t]}$ coincides with the Lipschitz constant $L_{[t]}$ defined before Lemma 3.6.

Lemma 3.19. *Suppose that Assumptions 3.1, 3.2 and 3.18 hold. Then, for all $x \in \mathcal{X}$ and $t \in [T]$, the conditional moments of $\widehat{H}_t(x)$ satisfy*

$$\begin{aligned} \mu_t^2(x) &\leq C_t \prod_{s=t}^{T-1} \left(\sum_{\ell=0}^{M_s} \frac{1}{2^\ell q_s(\ell)} \right) && \text{if Assumption 3.3 does not hold,} \\ \mu_t^{2^t}(x) &\leq D_t \prod_{s=t}^{T-1} \left(\sum_{\ell=0}^{M_s} \frac{1}{2^{2^s \ell} q_s(\ell)^{2^s-1}} \right) && \text{if Assumption 3.3 holds,} \end{aligned}$$

where

$$C_t := \bar{\mu}_T^2 (2B_2)^{T-t} L_{[t:T-1]}^2 \quad \text{and} \quad D_t := \bar{\mu}_T^{2^T} \prod_{s=t}^{T-1} \left(\left(\frac{3S_s}{2} \right)^{2^s} d_s^{2^s-1} B_{2^{s+1}} \right),$$

and where B_2 and $B_{2^{s+1}}$ are universal constants defined in Lemma A.4 in the appendix.

Lemma 3.19 allows us to bound the variance of the MLMC estimator $\widehat{F}(x) = \widehat{\mathbb{E}}[\widehat{H}_1(x)]$. Indeed, as $\widehat{F}(x)$ constitutes an average of n_1 independent copies of $\widehat{H}_1(x)$, its variance admits the bound

$$\mathbb{V}(\widehat{F}(x)) = \frac{1}{n_1} \mathbb{V}(\widehat{H}_1(x)) \leq \frac{\mu_1^2(x)}{n_1}, \quad (21)$$

while Lemma 3.19 provides a bound on $\mu_1^2(x)$. In fact, Lemma 3.19 constructs bounds on all conditional moments $\mu_t^2(x)$ (for nonsmooth integrands that may violate Assumption 3.3) and $\mu_t^{2^t}(x)$ (for smooth integrands that satisfy Assumption 3.3) by backward recursion on t . Note that if the integrands are nonsmooth, then the upper bound on $\mu_1^2(x)$ is finite only if $f_T(\xi_T, x)$ has a finite second moment. If the integrands are smooth, on the other hand, then the upper bound on $\mu_1^2(x)$ is finite only if $f_T(\xi_T, x)$ has a finite moment of order 2^T . The requirement that moments of order higher than two of $f_T(\xi_T, x)$ be finite arises as an artifact of the bound's construction, which involves iterative application of the inequality (13) for smooth functions. This inequality is conservative because it neglects the Lipschitz continuity of $f_T(\xi_T, x)$ with respect to x . Indeed, the Lipschitz continuity ensures that the global growth of the left hand side of (13) is at most linear. In spite of this, we do not attempt to improve the conditional moment bounds for smooth integrands in Lemma 3.19. Deriving such bounds is cumbersome and unlikely to improve our scenario complexity results below. Moreover, to establish high-probability scenario complexity guarantees for the MLMC estimator, we will anyway assume that $f_T(\xi_T, x)$ is sub-Gaussian, which implies that all moments of $f_T(\xi_T, x)$ are finite.

Lemma 3.19 readily extends to the untruncated MLMC estimator. In addition, if Assumption 3.3 holds while $d_t = 1$ and $M_t = \infty$ for all $t \in [T-1]$, then our upper bound on the 2^t -th conditional moment of \widehat{H}_t coincides with the 2^t -th unconditional moment of the untruncated MLMC estimator analyzed in [61].

The next result uses Lemma 3.19 to derive bounds on the bias of the MLMC estimator.

Lemma 3.20. *If Assumptions 3.1, 3.2 and 3.18 hold, then the bias of the MLMC estimator satisfies*

$$\left| \mathbb{E}[\widehat{F}(x)] - F(x) \right| \leq \begin{cases} \sum_{t=1}^{T-1} \frac{L_{[t]} \mathbb{E}[\mu_{t+1}^2(x)]^{\frac{1}{2}}}{2^{M_t/2}} & \text{if Assumption 3.3 does not hold,} \\ \sum_{t=1}^{T-1} \frac{L_{[t-1]} S_t \mathbb{E}[\mu_{t+1}^2(x)]}{2^{M_t+1}} & \text{if Assumption 3.3 holds.} \end{cases}$$

Lemma 3.20 shows that the bias of the MLMC estimator decays faster with the truncation points if the integrands are smooth (*i.e.*, if Assumption 3.3 holds) and if $\mathbb{E}[\mu_{t+1}^2(x)]$ is finite. Lemma C.6 in the appendix establishes a stronger result than Lemma 3.20 in that it provides a bound on the bias of $\widehat{H}_t(x)$ conditional on time $t - 1$ information for every $t \in [T]$. As $\mathbb{E}[\widehat{F}(x)] = \mathbb{E}[\widehat{H}_1(x)]$, Lemma 3.20 follows immediately from Lemma C.6 for $t = 1$, and therefore its proof is omitted for brevity.

Lemmas 3.20 and 3.6 reveal that the bias of the MLMC estimator with random branching factors 2^{λ_t} for $t \in [T - 1]$ is of the same order of magnitude as that of the SAA estimator with deterministic branching factors $n_{t+1} = 2^{M_t}$ for $t \in [T - 1]$. In addition, Lemma 3.20 generalizes [41, Proposition 4.1], which focuses on conditional stochastic programs with $T = 2$, and [61, Theorem 2.2], which focuses on untruncated and thus unbiased MLMC estimators with $M_t = \infty$ for all $t \in [T - 1]$. Note that when $T = 2$, the bias of the SAA and MLMC estimators can be analyzed in the same manner. However, when $T > 2$, the bias of the MLMC estimator requires a more involved analysis due to its recursive construction, and it also involves a bound on the variance of $\widehat{H}_t(x)$.

Remark 3.21. *The bias bounds stated in Lemma 3.20 are derived by first expressing the bias of $\widehat{F}(x)$ as a telescoping sum. Then, Jensen's inequality is applied to bound the absolute value of this sum by the sum of the absolute values of its terms. A tighter bias bound can be obtained using an inductive proof technique that, at each step, leverages a bias-variance decomposition. If the integrands are only known to be Lipschitz continuous (*i.e.*, if Assumption 3.3 does not hold), then this refined technique yields the bias bound*

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \left(\sum_{t=1}^{T-1} \frac{L_{[t]}^2 \mathbb{E}[\mu_{t+1}^2(x)]}{2^{M_t}} \right)^{1/2}.$$

*This bound on the bias of $\widehat{F}(x)$ matches the 2-norm of the vector with entries $L_{[t]} \mathbb{E}[\mu_{t+1}^2(x)]^{1/2} / 2^{M_t/2}$, $t \in [T - 1]$. Thus, it strengthens the corresponding bias bound of Lemma 3.20, which matches the 1-norm of the same vector. In addition, if the integrands are known to be smooth (*i.e.*, if Assumption 3.3 holds), it is still possible to recursively construct a bias bound that improves upon the one given in Lemma 3.20. However, this refined bound does not admit a simple closed-form expression. As the recursively constructed bounds (for nonsmooth as well as smooth integrands) exhibit the same asymptotic dependence on the truncation points M_t , $t \in [T - 1]$, we choose to work with the simpler and more interpretable bounds of Lemma 3.20.*

The next lemma provides an explicit formula for the expected sampling cost of the MLMC estimator.

Lemma 3.22. *The expected number of scenarios needed to construct the MLMC estimator amounts to*

$$\mathbb{E}[\mathbb{C}(\widehat{F}(x))] = n_1 \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell.$$

Lemmas 3.19, 3.20 and 3.22 provide bounds on the variance, the bias and the expected computational cost of the MLMC estimator, respectively. They hold for any choices of the estimator's hyperparameters, that is, the number n_1 of scenario trees as well as the rate parameter r_t and truncation point M_t of the geometric distribution governing the log-branching factor λ_t for every $t \in [T - 1]$. These hyperparameters should be tuned to guarantee a mean squared error of at most ϵ^2 at minimal computational cost. To achieve this, we select sufficiently large truncation points M_t , $t \in [T - 1]$, to ensure that the bias bound of Lemma 3.20 is at most $\epsilon/\sqrt{2}$. We also select a sufficiently large number n_1 of scenario trees to ensure that the variance is at most $\epsilon^2/2$. In view of (21), this is achieved by setting $n_1 \approx 2\mu_1^2(x)/\epsilon^2$, which implies that the expected sampling cost satisfies

$$\mathbb{C}(\widehat{F}(x)) \approx \frac{2\mu_1^2(x)}{\epsilon^2} \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell;$$

see Lemma 3.22. Recall that we only have access to upper bounds on $\mu_1^2(x)$ that depend indirectly on the rate parameters r_t through the probabilities $q_t(\ell)$. We thus construct approximate upper bounds on $\mathbb{C}(\widehat{F}(x))$ uniformly across all $x \in \mathcal{X}$ for both nonsmooth and smooth integrands by replacing $\mu_1^2(x)$ on the right hand side of the above expression with the respective bounds given in Lemma 3.19. We denote the resulting upper bounds by $c_{\text{ns}}(r)$ and $c_s(r)$ for nonsmooth and smooth integrands, respectively, where $r = (r_1, \dots, r_{T-1})$. Ideally, we should choose $r \in (0, 1)^{T-1}$ so as to minimize these upper bounds. However, the exact minimizers are difficult to obtain in closed form. The next assumption provides near-optimal choices for r .

Assumption 3.23 (Rate Parameters). *For any $t \in [T - 1]$ we set*

$$r_t = \begin{cases} 2^{-1} & \text{if Assumption 3.3 does not hold,} \\ 1 - 2^{-1-2^{-t}} & \text{if Assumption 3.3 holds.} \end{cases}$$

Remark 3.24. *Lemma C.7 in the appendix shows that if the integrands are nonsmooth (Assumption 3.3 does not hold), then the rate parameters of Assumption 3.23 exactly minimize the approximate upper bound $c_{\text{ns}}(r)$ on the expected sampling cost over all $r \in (0, 1)$. If $M_t = \infty$ for all $t \in [T - 1]$, however, then the choice $r_t = 2^{-1}$ leads to infinite variance and cost bounds. In fact, no value of r_t can simultaneously ensure finite second moment and finite expected sampling cost; see also [61]. These findings underscore the merits of truncated MLMC estimators. If the integrands are smooth (Assumption 3.3 holds), then the rate parameters of Assumption 3.23 exactly minimize the approximate upper bound $c_s(r)$ only for $T = 2$. For $T > 2$, the rate parameters in Assumption 3.23 jointly control both the geometric growth of the second moment and the expected sampling cost of $\widehat{F}(x)$. Specifically, Lemma 3.19 implies that μ_1^2 is bounded in M_t whenever $r_t < 1 - 2^{-1-(2^t-1)^{-1}}$, while Lemma 3.22 implies that the expected sampling cost is bounded in M_t whenever $r_t > 2^{-1}$ for all $t \in [T - 1]$. These bounds remain valid in the special case when $M_t = \infty$ for all $t \in [T - 1]$ addressed in [61]. The choice $r_t = 1 - 2^{-1-2^{-t}}$ for all $t \in [T - 1]$ satisfies both of these strict inequalities and provides a balanced trade-off between the geometric growth of the level- ℓ contributions to the variance and the expected sampling cost. Hence, this choice is natural.*

The MLMC estimator is fully determined by the number n_1 of scenario trees as well as the rates r_t and the truncation points M_t that define the truncated geometric distributions of the log-branching factors λ_t for all $t \in [T - 1]$. The next theorem specifies a particular choice of these hyperparameters that guarantees the mean squared error of the MLMC estimator to be bounded by ϵ^2 . It also characterizes the estimator's mean squared error-based scenario complexity, which we define as the expected value of its sampling cost.

Theorem 3.25 (Mean Squared Error-Based Scenario Complexity of the MLMC Estimator). *Suppose that Assumptions 3.1, 3.2, 3.18 and 3.23 hold and that $\epsilon > 0$ is a given tolerance. Set $n_1 = \lceil \sup_{x \in \mathcal{X}} 2\mu_1^2(x)/\epsilon^2 \rceil$, and define the truncation points M_t , $t \in [T - 1]$, through the backward recursion*

$$M_t = \left\lceil 2 \log_2 \left(\sqrt{2} L_{[t]} \sqrt{C_{t+1}} \left(\prod_{s=t+1}^{T-1} \sqrt{Z_s} \sqrt{M_s + 1} \right) (T - 1) / \epsilon \right) \right\rceil$$

if Assumption 3.3 does not hold and

$$M_t = \left\lceil \log_2 \left(\sqrt{2} L_{[t-1]} S_t D_{t+1}^{2^{-t}} \left(\prod_{s=t+1}^{T-1} Z_s^{\frac{2^s-1}{2^t}} \left(\frac{1-2^{-\frac{M_s+1}{2^s}}}{1-2^{-\frac{1}{2^s}}} \right)^{\frac{1}{2^t}} \right) (T - 1) / (2\epsilon) \right) \right\rceil$$

if Assumption 3.3 holds. Here, C_t and D_t are defined as in Lemma 3.19, and $Z_t := (1 - (1 - r_t)^{M_t+1})/r_t$. Then, the MLMC estimator achieves a root mean squared error of at most ϵ uniformly across all $x \in \mathcal{X}$. Thus, the mean squared error-based scenario complexity of the MLMC estimator amounts at most to $\mathcal{O}(\log(\epsilon^{-1})^{2(T-1)} \epsilon^{-2})$ if Assumption 3.3 does not hold and to $\mathcal{O}(\epsilon^{-2})$ if Assumption 3.3 holds.

Theorems 3.9 and 3.25 show that the MLMC estimator substantially outperforms the SAA estimator in terms of mean squared error-based scenario complexity. Specifically, for nonsmooth integrands, the MLMC estimator achieves a scenario complexity of $\mathcal{O}(\log(\epsilon^{-1})^{2(T-1)}\epsilon^{-2})$, representing a significant improvement over the $\mathcal{O}(\epsilon^{-2T})$ complexity of the SAA estimator. Similarly, for smooth integrands, the MLMC estimator improves the scenario complexity from $\mathcal{O}(\epsilon^{-(T+1)})$ for SAA to $\mathcal{O}(\epsilon^{-2})$. The MLMC estimator of Definition 3.17, which discards all terms \widehat{h}_t^ℓ , $\widehat{h}_t^{\ell,e}$ and $\widehat{h}_t^{\ell,o}$ beyond the truncation point M_t for each $t \in [T-1]$, also outperforms the *untruncated* MLMC estimator proposed in [61]. While the untruncated estimator achieves a scenario complexity of $\mathcal{O}(\epsilon^{-2})$ for smooth integrands, it suffers from infinite scenario complexity for nonsmooth integrands. In contrast, our MLMC estimator avoids this divergence by using truncation.

Note that while the scenario complexity of our MLMC estimator scales gracefully with $1/\epsilon$, it exhibits exponential dependence on T when the integrands are nonsmooth. As a result, the expected computational cost grows rapidly with T . This scaling behavior reflects the inherent difficulty of evaluating nested compositions of conditional expectations and nonsmooth functions. Nevertheless, for applications characterized by small values of T , as discussed in Section 2, the MLMC estimator can be computed to a moderate relative accuracy (e.g., 10% or 1%) at a practical computational cost. We emphasize, however, that Theorem 3.25 provides the first finite scenario complexity guarantee for problem (MCCE) with nonsmooth integrands that avoids exponential terms of the order ϵ^{-T} . As these scenario complexity guarantees pertain only to expected errors, it is expedient to complement them with bounds that hold with high probability. As in Section 3.1, such bounds are available for estimators with finite second moments under a sub-Gaussianity assumption.

Assumption 3.26 (Sub-Gaussianity of $\widehat{H}_1(x)$). *The estimator $\widehat{H}_1(x)$, which was introduced in Definition 3.17, is sub-Gaussian with variance proxy $\zeta^2 > 0$ for any fixed $x \in \mathcal{X}$.*

Remark 3.27 below outlines easily verifiable conditions under which Assumption 3.26 holds.

Remark 3.27. *Assumption 3.26 holds under the same conditions that imply Assumption 3.11 in Remark 3.12. Indeed, the sub-Gaussianity of $\widehat{H}_t(x)$ can be shown by backward recursion on $t \in [T-1]$, using the assumed sub-Gaussianity of ξ_t together with Lemma C.5, which ensures that sub-Gaussianity is preserved under finite sums, concatenations, Lipschitz continuous transformations and mixtures.*

Using Assumption 3.26, we now demonstrate that if the rates r_t and truncation points M_t for all $t \in [T-1]$ are chosen as in Theorem 3.25, then the worst-case absolute error of the MLMC estimator is bounded by ϵ with high probability, where the failure probability tends to 0 as n_1 increases. In the remainder of this section, we use $L' \in \mathbb{R}$ to denote the Lipschitz constant of the MLMC estimator $\widehat{F}(x)$ with respect to x , which can be shown to exist by using similar arguments as in Lemma C.3. Details are omitted for brevity.

Lemma 3.28. *Suppose that Assumptions 3.1, 3.2, 3.23 and 3.26 hold and that $\epsilon > 0$. Assume also that the truncation points M_t , $t \in [T-1]$ satisfy*

$$M_t = \left\lceil 2 \log_2 \left(4L_{[t]} \sqrt{C_{t+1}} \left(\prod_{s=t+1}^{T-1} \sqrt{Z_s} \sqrt{M_s + 1} \right) (T-1) / \epsilon \right) \right\rceil$$

if Assumption 3.3 does not hold and

$$M_t = \left\lceil \log_2 \left(2L_{[t-1]} S_t D_{t+1}^{2-t} \left(\prod_{s=t+1}^{T-1} Z_s^{\frac{2^s-1}{2^t}} \left(\frac{1-2^{-\frac{M_s+1}{2^s}}}{1-2^{-\frac{1}{2^s}}} \right)^{\frac{1}{2^t}} \right) (T-1) / \epsilon \right) \right\rceil$$

if Assumption 3.3 holds. Then, we have

$$\mathbb{P} \left(\sup_{x \in \mathcal{X}} |\widehat{F}(x) - F(x)| > \epsilon \right) \leq 2 \lceil 8L'D_{\mathcal{X}}/\epsilon + 1 \rceil^d \exp \left(-\frac{n_1 \epsilon^2}{32\zeta^2} \right).$$

Recall that ζ^2 is the variance proxy of $\widehat{H}_1(x)$ (see Assumption 3.26), while d and $D_{\mathcal{X}}$ are the dimension and the diameter of \mathcal{X} , respectively. The proof of Lemma 3.28 parallels that of Lemma 3.13 and is thus omitted. Lemma 3.28 readily implies that the minimizers of problem (1), which optimizes the MLMC estimator over \mathcal{X} , are near-optimal in (MCCO) with high probability. Unfortunately, these minimizers are often difficult to compute even if $F(x)$ is a convex function.

Remark 3.29. *Recall that $F(x)$ as well as the SAA estimator from Section 3.1 are convex if the integrand $f_t(\xi_t, x_t)$ is convex in x_t for every $t \in [T]$ and non-decreasing in x_t for every $t \in [T - 1]$. Perhaps surprisingly, however, these conditions fail to guarantee that the MLMC estimator is convex. To see this, assume for example that $T = 2$, $f_1(\xi_1, x_1) = \exp(x_1)$ and $f_2(\xi_2, x_2) = \xi_2 + x_2$. As $f_t(\xi_t, x_t)$ is convex and non-decreasing in x_t for every $t \in [T]$, we conclude that $F(x)$ and its SAA estimator are convex. Next, we construct the MLMC estimator with $n_1 = 1$ scenario trees in the event $\lambda_1 = 1$. This estimator requires a sample ξ_1^1 from the marginal distribution of ξ_1 and $2^{\lambda_1} = 2$ samples ξ_2^1 and ξ_2^2 from the conditional distribution of ξ_2 given $\xi_1 = \xi_1^1$. By Definition 3.17, the estimator can be represented as*

$$\begin{aligned}\widehat{F}(x) &= \frac{1}{q_1(\lambda_1)} \left[\exp((\xi_2^1 + \xi_2^2)/2 + x) - \frac{1}{2} \exp(\xi_2^1 + x) - \frac{1}{2} \exp(\xi_2^2 + x) \right] \\ &= \frac{\exp(x)}{q_1(1)} \left[\exp((\xi_2^1 + \xi_2^2)/2) - \frac{1}{2} \exp(\xi_2^1) - \frac{1}{2} \exp(\xi_2^2) \right],\end{aligned}$$

where the second equality holds because we condition on the event $\lambda_1 = 1$. Assuming that $\xi_2^1 \neq \xi_2^2$, Jensen's inequality implies that the last term in square brackets is negative, which in turn implies that $\widehat{F}(x)$ is nonconvex. Hence, minimizing the MLMC estimator over \mathcal{X} may, with positive probability, lead to an intractable nonconvex optimization problem. We emphasize, however, that the MLMC estimator constitutes an unbiased estimator for the SAA estimator with branching factor 2^{M_1} , which is convex in this example.

We now demonstrate that the minimizers of the MLMC problem (1) are, regardless of whether they can be computed efficiently, near-optimal in (MCCO) with high probability.

Corollary 3.30. *Assume that all conditions of Lemma 3.28 hold for some $\epsilon > 0$, and let x^* and \widehat{x}^* be minimizers of (MCCO) and the MLMC problem (1), respectively. Then, we have*

$$\mathbb{P}(F(\widehat{x}^*) - F(x^*) > \epsilon) \leq 4 \lceil 8L'D_{\mathcal{X}}/\epsilon + 1 \rceil^d \exp\left(-\frac{n_1 \epsilon^2}{128\zeta^2}\right).$$

The proof of Corollary 3.30 parallels that of Corollary 3.15 and is thus omitted. As the computational cost of the MLMC estimator is *random*, we generalize our definition of the high-probability scenario complexity as the *expected* number of scenarios used by a method that is guaranteed to find an ϵ -optimal solution of problem (MCCO) with probability $1 - \beta$. The next result follows immediately from Corollary 3.30.

Theorem 3.31 (High-Probability Scenario Complexity of the MLMC Estimator). *Suppose that all conditions of Lemma 3.28 hold for some $\epsilon > 0$ and that*

$$n_1 = \left\lceil \frac{128\zeta^2}{\epsilon^2} (d \log(\lceil 8L'D_{\mathcal{X}}/\epsilon + 1 \rceil) + \log(4/\beta)) \right\rceil$$

for some $\beta \in (0, 1)$. Then any solution of the MLMC problem (1) is an ϵ -optimal solution of (MCCO) with probability at least $1 - \beta$. Thus, the high-probability scenario complexity of the MLMC estimator equals $\mathcal{O}(\log(\epsilon^{-1})^T \epsilon^{-2})$ if Assumption 3.3 does not hold and $\mathcal{O}(\log(\epsilon^{-1}) \epsilon^{-2})$ if Assumption 3.3 holds.

Theorems 3.16 and 3.31 show that the MLMC estimator significantly outperforms the SAA estimator in terms of high-probability scenario complexity. Indeed, it improves the high-probability scenario complexity from $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-2T})$ to $\mathcal{O}(\log(\epsilon^{-1})^T\epsilon^{-2})$ for nonsmooth integrands and from $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-(T+1)})$ to $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-2})$ for smooth integrands. In contrast, the *untruncated* MLMC estimator studied in [61] lacks a finite Lipschitz constant. Indeed, as is evident from (18), the Lipschitz modulus of $\widehat{H}_t(x)$ scales with $1/q(\lambda_t)$. As the random variable λ_t has unbounded support when $M_t = \infty$, $1/q(\lambda_t)$ can be arbitrarily large. Hence, $\widehat{H}_t(x)$ admits no Lipschitz constant that applies uniformly across all realizations of λ_t . Our bound on the high-probability scenario complexity of the untruncated MLMC estimator thus diverges.

Note that the high-probability guarantees of Theorem 3.31 for smooth integrands require a larger number n_1 of scenario trees than the mean squared error-based scenario complexity results in Theorem 3.25. Specifically, the high-probability scenario complexity incurs an additional multiplicative factor of $\log(\epsilon^{-1})$. Despite this increase, the ϵ -scaling of the MLMC estimator remains better than that of the SAA estimator.

For non-smooth integrands, the high-probability scenario complexity $\mathcal{O}(\log(\epsilon^{-1})^T\epsilon^{-2})$ even grows slower with ϵ^{-1} than the mean squared error-based sample complexity $\mathcal{O}(\log(\epsilon^{-1})^{2(T-1)}\epsilon^{-2})$ from Theorem 3.25. This phenomenon arises because of the sub-Gaussian tail assumption on the estimator $\widehat{H}_1(x)$ in Theorem 3.31, which yields tighter control over the worst-case behavior than variance-based bounds.

Finally, we emphasize that while the scenario complexity of the truncated MLMC estimator grows gracefully with ϵ^{-1} for (moderate) fixed values of T , it usually grows exponentially with T . Such exponential growth appears to be unavoidable even for smooth integrands, in which case the scenario complexity can be shown to include a constant term that is exponential in T . In addition, the untruncated MLMC estimator from [61] as well as the SAA estimator from Section 3.1 also display an exponential cost growth in T .

4 Gradient Methods for MCCO

From Section 3 we know that the minimizers of problem (1), equipped with the SAA or the MLMC estimator, are near-optimal in (MCCO) with high probability. Despite their favorable statistical properties, computing these minimizers can be challenging in the absence of restrictive convexity and monotonicity conditions. This motivates us in this section to develop efficient estimators for $\nabla F(x) \in \mathbb{R}^d$ that can be employed within first-order optimization methods for solving problem (MCCO) to stationarity.

We aim to recursively construct MLMC gradient estimators using similar ideas as in Section 3.2. These estimators are again based on a forest of scenario trees with a random branching structure, and the underlying log-branching factors are modeled by truncated geometric distributions. The resulting estimators are biased, but the bias can be reduced by increasing the truncation points. In the special case where the truncation points are set to infinity, we obtain *untruncated*, and therefore *unbiased*, estimators. Both the truncated and untruncated gradient estimators proposed here are novel for problems with more than two stages ($T > 2$). Truncated as well as untruncated MLMC gradient estimators for two-stage problems are discussed in [29, 41]. As will become clear below, the generalization to $T \geq 2$ poses significant technical challenges.

Throughout, we work under Assumptions 3.2 and 3.3, which guarantee the existence and well-definedness of the gradients. We first show that the gradient $\nabla F(x)$ can be constructed via the backward recursion

$$G_t(x) := \mathbb{E}_{t-1} [G_{t+1}(x)\nabla_{x_t} f_t(\xi_t, F_{t+1}^T(x))] \in \mathbb{R}^{d \times d_{t-1}} \quad \forall t \in [T-1] \quad (22)$$

initialized by $G_T(x) := \mathbb{E}_{T-1}[\nabla_x f_T(\xi_T, x)] \in \mathbb{R}^{d \times d_{T-1}}$, where $F_{t+1}^T(x)$ represents the nest of the last $T-t$ conditional expectations and nonlinear integrands in (MCCO). Hence, we have $F(x) = F_1^T(x)$; see also Definition C.1 in the appendix. All gradients in (22) exist by Assumption 3.3. Moreover, since $\nabla_{x_t} f_t(\xi_t, x_t)$ is uniformly bounded under Assumption 3.2, the same holds for $G_t(x)$. Consequently, the conditional expectation in (22) is well-defined. We can now prove that $\nabla F(x) = G_1(x)$.

Lemma 4.1. *If Assumptions 3.1, 3.2 and 3.3 hold, then $\nabla F(x) = G_1(x)$ for all $x \in \mathbb{R}^d$.*

Lemma 4.1 shows that the gradient $\nabla F(x)$ admits a recursive representation analogous to that of $F(x)$ itself. This observation allows us to construct gradient estimators using methods similar to those in Section 3. In particular, following the analogy with Definition 3.17, we can define an MLMC estimator for $\nabla F(x)$.

Definition 4.2 (MLMC Gradient Estimator). *For any $t \in [T - 1]$, select a rate parameter $r_t \in (0, 1)$ and truncation point $M_t \in \mathbb{N}$, and let λ_t be a random variable that follows the truncated geometric distribution $\text{Geo}(r_t | M_t)$. Assume that λ_t is independent of all other random objects, and set $q_t(\ell) := \mathbb{P}(\lambda_t = \ell)$ for $\ell = 0, \dots, M_t$ and $t \in [T - 1]$. We construct gradient estimators $\widehat{G}_t(x)$, $t \in [T]$, using the following backward recursion. We initialize the recursion by $\widehat{G}_T(x) := \nabla_x f_T(\xi_T, x)$. Next, for any $t \in [T - 1]$, we set*

$$\widehat{G}_t(x) := \frac{1}{q_t(\lambda_t)} \left(\widehat{g}_t^{\lambda_t}(x) - \frac{1}{2} \widehat{g}_t^{\lambda_t, e}(x) - \frac{1}{2} \widehat{g}_t^{\lambda_t, o}(x) \right), \quad (23)$$

where

$$\widehat{g}_t^\ell(x) := \widehat{\mathbb{E}}_t^\ell[\widehat{G}_{t+1}(x)] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}(x)]) \quad \forall \ell = 0, \dots, M_t$$

as well as

$$\begin{aligned} \widehat{g}_t^{\ell, e}(x) &:= \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{G}_{t+1}(x)] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}(x)]) \quad \forall \ell \in [M_t], \\ \widehat{g}_t^{\ell, o}(x) &:= \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{G}_{t+1}(x)] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}(x)]) \quad \forall \ell \in [M_t]. \end{aligned}$$

We also set $\widehat{g}_t^{0, e}(x) := \widehat{g}_t^{0, o}(x) := 0$. The MLMC estimator for $\nabla F(x)$ is then defined as $\widehat{G}(x) := \widehat{\mathbb{E}}[\widehat{G}_1(x)]$.

A special case arises when the truncation points in Definition 4.2 are set to $M_t = \infty$ for all $t \in [T - 1]$. In this case, we refer to the resulting estimator as the *untruncated* MLMC gradient estimator. Recall that $\widehat{H}_t(x)$ was introduced in Definition 3.17 and can be viewed as an estimator for the conditional risk mapping $F_t^T(x)$ from Definition C.1. Similarly, $\widehat{G}_t(x)$ constitutes an estimator for $G_t(x)$. Crucially, the construction above reveals an important structural feature, namely that the gradient estimator $\widehat{G}_t(x)$ depends on the function value estimator $\widehat{H}_{t+1}(x)$ as an argument at each stage, and both sequences of estimators must therefore be maintained and propagated simultaneously. This coupling is a novel structural feature of our construction, *absent* from classical MLMC gradient estimators. Both $\widehat{G}_t(x)$ and $\widehat{H}_t(x)$ are built on the same forest of n_1 scenario trees displaying a random branching structure with 2^{λ_t} branches per stage- t node, where λ_t is sampled from a truncated geometric distribution. Note that the forest of trees is defined implicitly through the nesting of the empirical conditional expectations $\widehat{\mathbb{E}}[\cdot]$ and $\widehat{\mathbb{E}}^{\lambda_t}[\cdot]$ for $t \in [T - 1]$, which are used both for the construction of the function value and gradient estimators. This implies that $\widehat{G}_t(x)$ and $\widehat{H}_t(x)$, $t \in [T - 1]$, are constructed from the same samples and are therefore correlated. Reusing the same samples across both estimators reduces the sampling cost, but it complicates the analysis of the trade-off between variance and computational cost. Our analysis uses the following Hölder condition.

Assumption 4.3 (Hölder Continuous Hessians). *The Hessian matrix $\nabla_{x_t}^2 f_{t,i}(\xi_t, x_t)$ is ρ_t -Hölder continuous in x_t across all ξ_t . That is, for all $t \in [T - 1]$, there exist $R_t > 0$ and $\rho_t \in (1 - 2^{1-t}, 1]$ such that*

$$\|\nabla_{x_t}^2 f_{t,i}(\xi_t, x_t) - \nabla_{x_t}^2 f_{t,i}(\xi_t, y_t)\|_2 \leq R_t \|x_t - y_t\|_2^{\rho_t} \quad \forall x_t, y_t \in \mathbb{R}^{d_t}, \forall \xi_t \in \mathbb{R}^{m_t}, \forall i \in [d_{t-1}].$$

Note that when $\rho_t = 1$, Assumption 4.3 requires the Hessian matrices of the components of $f_t(\xi_t, x_t)$ to be Lipschitz continuous in x_t . Assumption 4.3 has been used in [29] to control the variance and the sampling costs of MLMC gradient estimators for conditional stochastic optimization problems, which arise as special cases of (MCCO) with $T = 2$ nests. It implies via the fundamental theorem of calculus that

$$\|\nabla_{x_t} f_{t,i}(\xi_t, x_t) - \nabla_{x_t} f_{t,i}(\xi_t, y_t) - \nabla_{x_t}^2 f_{t,i}(\xi_t, y_t)(x_t - y_t)\|_2 \leq \frac{R_t}{\rho_t + 1} \|x_t - y_t\|_2^{\rho_t + 1} \quad (24)$$

for all $x_t, y_t \in \mathbb{R}^{d_t}$ and $\xi_t \in \mathbb{R}^{m_t}$.

In order to derive bounds on the mean squared error-based scenario complexity of the MLMC gradient estimator, we first analyze its variance, bias, and sampling cost. In analogy with (19), for any $p \geq 1$, we define the p -th conditional moment of $\widehat{G}_t(x)$ given the information available up to time $t - 1$ by

$$\nu_t^p(x) := \mathbb{E}_{t-1} \left[\left\| \widehat{G}_t(x) \right\|_2^p \right]. \quad (25)$$

Recall that $\widehat{G}_t(x) \in \mathbb{R}^{d \times d_{t-1}}$ is a matrix, and thus $\|\widehat{G}_t(x)\|_2$ stands for its Frobenius norm. Note also that the p -th conditional moment of $\widehat{H}_t(x)$ in (19) was defined with respect to the vector p -norm. To facilitate the analysis of the MLMC gradient estimator, we define the conditional moment bound of order p as

$$\bar{\nu}_T^p := \sup_{x \in \mathbb{R}^d} \text{ess sup} \mathbb{E}_{T-1} \left[\left\| \nabla_x f_T(\xi_T, x) \right\|_p^p \right].$$

The following assumption is crucial for our subsequent results.

Assumption 4.4. *The conditional moment bound $\bar{\mu}_T^p$ of the integrand $f_T(\xi_T, x)$ is finite up to order $p = 2^T(\rho_{T-1} + 1)$, and the conditional moment bound $\bar{\nu}_T^p$ of the gradient $\nabla f_T(\xi_T, x)$ is finite up to order $p = 2^T$.*

Note that Assumption 4.4 implies Assumption 3.18 because $\rho_{T-1} > 0$. The following lemma establishes a bound on $\nu_t^{2^t}(x)$ for all $t \in [T]$. In particular, for $t = 1$, this bound can be used to derive an upper bound on the unconditional variance of the MLMC gradient estimator.

Lemma 4.5. *Suppose that Assumptions 3.1, 3.2, 3.3, 4.3 and 4.4 hold. Then, for every $t \in [T - 1]$ there exists E_t that depends only on T , the moment bounds from Assumption 4.4, the dimensions d_s and smoothness constants S_s for $s = t + 1, \dots, T - 1$, and the Hölder constants R_s and ρ_s for $s = t, \dots, T$ such that*

$$\nu_t^{2^t}(x) \leq E_t \prod_{s=t}^{T-1} \max \left\{ \sum_{\ell=0}^{M_s} \frac{1}{2^{(2^s-1)(\rho_s+1)\ell} q_s(\ell)^{2^s-1}}, \sum_{\ell=0}^{M_s} \frac{1}{2^{2^s(\rho_{s-1}+1)\ell} q_s(\ell)^{2^s(\rho_{s-1}+1)-1}} \right\}$$

for all $x \in \mathcal{X}$ and $t \in [T]$. Here, we adopt the convention $\rho_0 = 0$ for notational convenience.

The exact closed-form expressions for the constants E_t are cumbersome and thus omitted from the statement of Lemma 4.5. Instead, an explicit recursive definition is given in the proof. Lemma 4.5 allows us to bound the variance of the MLMC gradient estimator $\widehat{G}(x) = \widehat{\mathbb{E}}[\widehat{G}_1(x)]$. Indeed, as $\widehat{G}(x)$ constitutes an average of n_1 independent copies of $\widehat{G}_1(x)$, its variance admits the bound

$$\mathbb{V}(\widehat{G}(x)) = \frac{1}{n_1} \mathbb{V}(\widehat{G}_1(x)) \leq \frac{\nu_1^2(x)}{n_1}, \quad (26)$$

while Lemma 4.5 provides a bound on $\nu_1^2(x)$. Note that this bound is finite only if $\nabla f_T(\xi_T, x)$ admits finite moments of order 2^T , and $f_T(\xi_T, x)$ admits finite moments of order $2^T(\rho_{T-1} + 1)$. We further emphasize that Lemma 4.5 also applies to the *untruncated* MLMC gradient estimator obtained by setting $M_t = \infty$ for all $t \in [T - 1]$. In this case, however, the resulting variance bound remains finite only if the rate parameters $r_t, t \in [T - 1]$, are sufficiently large to ensure convergence of all infinite series appearing in the bound.

Lemma 4.5 is closely related to Lemma 3.19, as both establish bounds on conditional moments of MLMC estimators by means of backward recursions in time. The key differences lie in the underlying assumptions and in the complexity of the resulting recursions. For function-value estimators, Lemma 3.19 requires only Lipschitz continuity of the integrands, yielding relatively simple product-form bounds that admit closed-form expressions. If the integrands are also smooth, the bounds strengthen while retaining a simple closed form. By contrast, for gradient estimators, Lemma 4.5 requires not only smoothness of the integrands but also Hölder continuity of their Hessian matrices, which introduces the Hölder exponents ρ_t

into the analysis and makes the proof technically more demanding. Nevertheless, both lemmas serve the same purpose. They provide explicit upper bounds on conditional moments that can be used to control the growth of the estimator's variance and to establish scenario complexity guarantees. When $\rho_t = 1$ for all $t \in [T - 1]$ (i.e., when all Hessian matrices are Lipschitz continuous), the second term in the max operator of the moment bound in Lemma 4.5 dominates the first term for all $s \in [T - 1]$ —recall that $\rho_0 = 0$ —which implies that the conditional moment bounds for the gradient estimator impose stricter requirements on the rate parameters r_s governing $q_s(\ell)$ than those for the function-value estimator in Lemma 3.19.

Next, we analyze the bias of the MLMC gradient estimator using the moment bounds established in Lemmas 3.19 and 4.5. This analysis is cumbersome because \widehat{g}_t^ℓ and its variants based on only even- or odd-indexed samples, $\widehat{g}_t^{\ell,e}$ and $\widehat{g}_t^{\ell,o}$, respectively (see Definition 4.2), exhibit a product structure. More precisely, any bias in \widehat{H}_{t+1} is propagated through the gradient mapping $\nabla_{x_t} f_t(\xi_t, \cdot)$, and its magnitude is thus controlled by the smoothness of f_t . Furthermore, the biases of \widehat{G}_{t+1} and \widehat{H}_{t+1} interact multiplicatively, giving rise to cross terms that must be carefully controlled. This interplay significantly complicates the analysis compared with the case of the function value estimators studied in Section 3.2.

Lemma 4.6. *If Assumptions 3.1, 3.2, 3.3 and 4.4 hold, the bias of the MLMC gradient estimator satisfies*

$$\|\mathbb{E}[\widehat{G}(x)] - G_1(x)\|_2 \leq \sum_{t=1}^{T-1} L_{[t-1]} S_t \mathbb{E}[\nu_{t+1}^2(x)]^{\frac{1}{2}} \left(\frac{\mathbb{E}[\mu_{t+1}^2(x)]^{\frac{1}{2}}}{2^{M_t/2}} + \sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}[(\mu_{s+1}^2(x))^2]^{\frac{1}{2}}}{2^{M_s+1}} \right).$$

As in the case of the MLMC function value estimator discussed in Section 3.2, Lemma 4.6 shows that the MLMC gradient estimator is unbiased if $M_t = \infty$ for all $t \in [T - 1]$ provided that all expectations in the bias bound are finite. Hence, the *untruncated* MLMC gradient estimator is *unbiased*. We now quantify the expected sampling cost of the MLMC gradient estimator.

Lemma 4.7. *The expected number of scenarios needed to construct the MLMC gradient estimator equals*

$$\mathbb{E}[\mathbb{C}(\widehat{G}(x))] = n_1 \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell.$$

The proof of Lemma 4.7 widely parallels that of Lemma 3.22 and is thus omitted. As in Section 3.2, we now tune the hyperparameters n_1 , r_t and M_t of the MLMC estimator to guarantee a mean squared error of at most ϵ^2 at minimal computational cost. To achieve this, we select sufficiently large truncation points M_t , $t \in [T - 1]$, to ensure that the bias bound of Lemma 4.6 is at most $\epsilon/\sqrt{2}$. We also select a sufficiently large number $n_1 \approx 2\nu_1^2(x)/\epsilon^2$ of scenario trees to ensure that the variance is at most $\epsilon^2/2$; see (26). By Lemma 4.7, the expected sampling cost can thus be approximated by

$$\mathbb{C}(\widehat{G}(x)) \approx \frac{2\nu_1^2(x)}{\epsilon^2} \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell.$$

We can construct an approximate upper bound on $\mathbb{C}(\widehat{G}(x))$ uniformly across all $x \in \mathcal{X}$ by replacing $\nu_1^2(x)$ with the respective bound given in Lemma 4.5. The resulting upper bound $c_h(r)$ depends on the rate parameters $r = (r_1, \dots, r_{T-1})$ of the truncated geometric distributions used to model the log-branching factors. Here, the subscript “h” indicates that the bound was obtained under the assumption of Hölder continuous Hessians. Ideally, $r \in (0, 1)^{T-1}$ should be selected to minimize the given upper bound. As in Section 3.2, closed-form minimizers are elusive. The next assumption specifies a simple near-optimal choice for r .

Assumption 4.8 (Rate Parameters). *For any $t \in [T - 1]$, we select r_t satisfying*

$$2^{-1} < r_t < \min \left\{ 1 - 2^{-\frac{2^{t-1}(\rho_t+1)}{2^t-1}}, 1 - 2^{-\frac{2^t(\rho_{t-1}+1)}{2^t(\rho_{t-1}+1)-1}} \right\}.$$

A rate parameter r_t that satisfies both strict inequalities is guaranteed to exist whenever Assumption 4.3 holds. Indeed, in this case the Hölder exponent satisfies $\rho_t > 1 - 2^{1-t}$, which guarantees that the first term in the minimum is strictly greater than 2^{-1} . The second term is strictly greater than 2^{-1} for any $\rho_t \geq 0$. Assumption 4.8 ensures that the given rate parameters jointly control the geometric growth of the second moment as well as the expected sampling cost of $\widehat{G}(x)$. Specifically, Lemma 4.5 implies that ν_1^2 remains bounded in M_t if r_t obeys the given upper bound in Assumption 4.8, and Lemma 4.7 implies that the expected sampling cost is bounded in M_t if $r_t > 2^{-1}$ for all $t \in [T - 1]$. Note also that if $M_t = \infty$ for all $t \in [T - 1]$, then Assumption 4.8 guarantees that the infinite sums in Lemmas 4.5 and 4.7 converge. Hence, the variance and the expected sampling cost of the untruncated MLMC gradient estimator are finite.

The MLMC gradient estimator is fully characterized by the number n_1 of scenario trees as well as the rates r_t and truncation points M_t governing the geometric distributions of the log-branching factors λ_t for $t \in [T - 1]$. The following theorem presents a specific configuration of these hyperparameters that ensures the root mean squared error of the truncated MLMC gradient estimator to be bounded by ϵ .

Theorem 4.9 (Mean Squared Error-Based Scenario Complexity of the MLMC Gradient Estimator). *Let Assumptions 3.1, 3.2, 3.3, 4.3, 4.4 and 4.8 hold, and let $\epsilon > 0$ be a given tolerance. Then, for every $t \in [T - 1]$ there exists W_t that depends only on T , the moment bounds from Assumption 4.4, the dimensions d_s and smoothness constants S_s for $s = t + 1, \dots, T - 1$, the Lipschitz constants L_s and the rate parameters r_s for $s \in [T - 1]$, and the Hölder constants R_s and ρ_s for $s = t, \dots, T$ such that the following holds. If*

$$n_1 = \lceil \sup_{x \in \mathcal{X}} 2\nu_1^2(x)/\epsilon^2 \rceil \quad \text{and} \quad M_t = \lceil 2 \log_2(W_t/\epsilon) \rceil \quad \forall t \in [T - 1],$$

then the MLMC gradient estimator achieves a root mean squared error of at most ϵ uniformly for all $x \in \mathcal{X}$. Thus, its mean squared error-based scenario complexity is at most $\mathcal{O}(\epsilon^{-2})$.

The exact closed-form expressions for the constants W_t are cumbersome and thus omitted from the statement of Theorem 4.9. The explicit construction of these constants will become clear in the proof.

Theorem 4.9 establishes that, if the integrands have Hölder continuous Hessians, then the MLMC gradient estimator achieves the optimal scenario complexity of $\mathcal{O}(\epsilon^{-2})$. This may be surprising because the MLMC gradient estimator has a significantly more complex structure than the MLMC function value estimator, implying that it is more difficult to control its variance; see also the discussion after Lemma 4.5. We emphasize that Theorem 4.9 remains valid for the *untruncated* MLMC gradient estimator, which is obtained by setting $M_t = \infty$ for all $t \in [T - 1]$. This result is new for $T > 2$. When $T = 2$, Theorem 4.9 recovers the $\mathcal{O}(\epsilon^{-2})$ scenario complexity for estimating the gradients of conditional stochastic optimization problems established in [29, Theorem 2]. The Hölder continuity of the Hessians is a crucial assumption for Theorem 4.9. Indeed, if the integrands are merely Lipschitz continuous and smooth, then *no* choice of r_t , $t \in [T - 1]$, guarantees that the variance and the expected sampling cost remain simultaneously bounded [41]. Hence, the scenario complexity of the *untruncated* MLMC gradient estimator diverges.

Remark 4.10. *The MLMC gradient estimator introduced in Definition 4.2 employs the same set of samples to construct both $\widehat{G}_t(x)$ and $\widehat{H}_t(x)$. This shared use of samples reduces the expected sampling cost. However, it also induces correlation between $\widehat{G}_t(x)$ and $\widehat{H}_t(x)$, which complicates the analysis of the scenario complexity and necessitates the assumption that the integrands possess Hölder continuous Hessians. Alternatively, one could construct a simpler MLMC gradient estimator that employs independent sets of samples to compute the components $\widehat{G}_t(x)$ and $\widehat{H}_t(x)$. The independence between these components simplifies the moment bounds in Lemma 4.5 and, in fact, eliminates the need to establish moment bounds of orders $p > 2$. Consequently, it suffices to derive a second-order moment bound analogous to that obtained in Lemma 3.19 for nonsmooth integrands. However, this alternative estimator entails a higher expected sampling cost. One can show that its mean squared error-based scenario complexity amounts to $\mathcal{O}(\log(\epsilon^{-1})^2(T-1)\epsilon^{-2})$, which*

matches that of the function value estimator for nonsmooth integrands from Theorem 3.25. The proof follows similar arguments to those used in Theorem 4.9, albeit in a simpler form. Detailed derivations are therefore omitted for brevity. We emphasize that, unless the integrands have Hölder continuous Hessians, the truncation points of the simpler MLMC gradient estimator cannot be sent to infinity without causing either the variance or the expected sampling cost to diverge. Hence, in this simpler setting, it is impossible to construct an untruncated unbiased estimator without imposing additional regularity conditions.

We can now show that the MLMC gradient estimator introduced in Definition 4.2 can be incorporated into an SGD algorithm (see Algorithm 1) to efficiently compute an ϵ -stationary point of problem (MCCO). In the following, we use $S_{[T]} := \sum_{t=1}^T L_{[1:t-1]} S_t L_{[t+1:T]}^2$ as a measure for the smoothness of $F(x)$, where $L_{[s:t]}$ for $t \in [T]$ and $s \in [T+1]$ is defined as before Lemma 3.19. For more details see Lemma C.3.

Theorem 4.11. *Let Assumptions 3.1, 3.2, 3.3, 4.3, 4.4 and 4.8 hold, and set $\bar{\nu}_1^2 := \sup_{x \in \mathcal{X}} \nu_1^2(x)$. Assume that Algorithm 1 uses a constant stepsize $\eta = \bar{\nu}_1 \sqrt{n_1/K}$ and uses the MLMC gradient estimator with $n_1 = \mathcal{O}(1)$ and $M_t, t \in [T-1]$, defined as in Theorem 4.9. If Algorithm 1 is initialized at $x_1 \in \mathcal{X}$ and runs over*

$$K \geq \frac{\bar{\nu}_1^2}{n_1 \epsilon^4} \left(2 \left[F(x_1) - \min_{x \in \mathcal{X}} F(x) \right] + S_{[T]} \right)^2$$

iterations, then it outputs an ϵ -stationary point \bar{x}_K with $\mathbb{E}[\|\nabla F(\bar{x}_K)\|_2^2] \leq \epsilon^2$.

The last part of the proof of Theorem 4.9 shows that if $n_1 = \mathcal{O}(1)$, then the expected sampling cost of constructing a single MLMC gradient estimator is independent of ϵ (i.e., $\mathcal{O}(1)$). Theorem 4.11 thus implies that Algorithm 1 requires $\mathcal{O}(\epsilon^{-4})$ scenarios to compute an ϵ -stationary point of (MCCO). Theorem 4.11 generalizes a known scenario complexity result for conditional stochastic optimization problems, which correspond to instances of (MCCO) with $T = 2$ stages [41, Corollary 4.1]. When employing untruncated MLMC gradient estimators, obtained by setting $M_t = \infty$ for all $t \in [T-1]$, a general complexity result for constrained nonconvex stochastic optimization with unbiased gradient oracles [22, Corollary 4] implies that an ϵ -stationary point of problem (MCCO) can still be computed in $\mathcal{O}(\epsilon^{-4})$ iterations. This convergence rate is known to be optimal for unbiased gradient oracles in the absence of variance reduction [2]. Variance-reduction techniques can further improve the optimal rate for unbiased gradient oracles to $\mathcal{O}(\epsilon^{-3})$ [2].

Algorithm 1 Projected SGD

Input: Number of iterations K , stepsize schedule $\{\eta_k\}_{k=1}^K$, initial iterate x_1

for $k = 1, \dots, K$ **do**

 Construct an estimator $\widehat{G}(x_k)$ for the gradient $\nabla F(x_k)$

 Update $x_{k+1} = \Pi_{\mathcal{X}}(x_k - \eta_k \widehat{G}(x_k))$, where $\Pi_{\mathcal{X}}$ is the Euclidean projection on \mathcal{X}

end for

Output: \bar{x}_K chosen uniformly at random from $\{x_k\}_{k=1}^K$

5 Numerical Experiments

We now assess the benefits of our MLMC function value and gradient estimators for solving MCCE and MCCO problems, and we experimentally validate our theoretical results. All experiments are conducted on a MacBook Pro equipped with an Apple M1 Max chip and 32 GB RAM, running macOS Sonoma 14.5. Code and data for reproducing the numerical experiments are available at <https://github.com/busesen/MCCO>.

MCCE with a Known Ground-Truth Solution The first experiment is centered around a synthetic instance of (MCCE), originally introduced in [61, § 3.1], with $T = 3$ stages and integrands $f_1(\xi_1, x_1) = \sin(\xi_1 + x_1)$, $f_2(\xi_2, x_2) = \sin(\xi_2 - x_2)$ and $f_3(\xi_3, x) = \xi_3$. The disturbances follow a Markovian structure with $\xi_1 \sim \mathcal{N}(\pi/2, 1)$, $\xi_2|\xi_1 \sim \mathcal{N}(\xi_1, 1)$ and $\xi_3|\xi_2 \sim \mathcal{N}(\xi_2, 1)$. This problem has the closed-form solution $F(x) = \exp(-1/2) \approx 0.6065$, against which the accuracy of different estimators can be assessed.

We compare two SAA estimators in the sense of Definition 3.4 with two MLMC estimators in the sense of Definition 3.17. The first SAA estimator (SAA1) employs a uniform branching structure with $n_1 = n_2 = n_3$, while the second SAA estimator (SAA2) emphasizes the first stage by setting $n_1 = n_2^2 = n_3^2$; see [51, 61] for motivation and discussion. The first MLMC estimator, taken from [61], is untruncated, uses the recommended rate parameters $r_1 = 0.74$ and $r_2 = 0.60$, and sets $M_1 = M_2 = \infty$, resulting in an expected sampling cost of 4.6250 per each of the n_1 underlying scenario trees. The second MLMC estimator is truncated, adopts the rates $r_1 = 1 - 2^{-3/2}$ and $r_2 = 1 - 2^{-5/4}$ proposed in Assumption 3.23 for smooth integrands, and sets $M_1 = 6$ and $M_2 = 5$, yielding an expected sampling cost of 4.7674 per scenario tree. The truncation points are chosen to satisfy $M_1 > M_2$, in accordance with Lemma 3.28, while maintaining an expected sampling cost per tree comparable to that of the untruncated MLMC estimator.

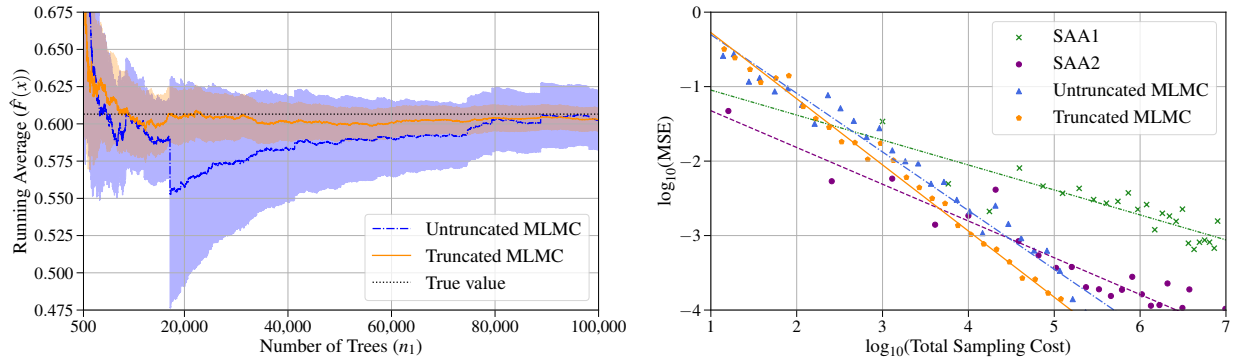
For each of the four estimators, we consider a sequence of instances indexed by n_1 , yielding decreasing mean squared errors as n_1 increases. For the SAA estimators, the associated scenario trees become increasingly bushy, and arguments analogous to those in Theorem 3.9 show that their mean squared errors converge to 0, implying asymptotic unbiasedness. By Lemma 3.20, the untruncated MLMC estimator is unbiased for every fixed n_1 , and similar reasoning as in Theorem 3.25 establishes that its mean squared error, which coincides with its variance, also vanishes as n_1 grows. In contrast, the truncated MLMC estimator exhibits a nonvanishing bias as n_1 grows. Thus, increasing n_1 serves only to reduce its variance.

Figure 2 (a) displays the two MLMC estimators (lines) together with their 95% confidence intervals (shaded regions) as functions of n_1 . Since any MLMC estimator $\widehat{F}(x)$ is formed by averaging n_1 independent tree-wise estimators with finite second moments (see Definition 3.17 and Lemma 3.19), the central limit theorem implies approximate normality, and we therefore report confidence bounds of the form $\widehat{F}(x) \pm 1.96 \widehat{\sigma}_{n_1} / \sqrt{n_1}$, where $\widehat{\sigma}_{n_1}$ denotes the empirical standard deviation of the n_1 tree-wise estimators. For $n_1 < 500$ the variance becomes excessively large, and the plot is therefore restricted to $n_1 \geq 500$. Although unbiased, the untruncated MLMC estimator exhibits substantially larger variance across all values of n_1 and is notably fragile, as illustrated by the discontinuity at $n_1 = 17,212$. Such jumps are caused by the occasional sampling of large log-branching factors, which can occur because the underlying geometric distribution is *not* truncated.¹ The figure shows the first of 10 independent simulation runs, with similarly pronounced discontinuities observed in 4 out of the 10 runs for the untruncated estimator. Overall, truncation yields markedly more stable MLMC estimators with reduced variance, and although it introduces bias, this bias remains negligible at least up to $n_1 = 10^5$. The two SAA estimators are known to converge slower to $F(x)$ than the MLMC estimators and are thus not shown in Figure 2 (a); see, e.g., [61, § 3].

Figure 2 (b) illustrates the trade-off between mean squared error and sampling cost for the four estimators. As before, different instances of each estimator are parametrized by n_1 . Since the truncated MLMC estimator is biased, its mean squared error is eventually dominated, for sufficiently large n_1 , by the squared bias. This causes the error to saturate as n_1 (and thus the total sampling cost) increases. By Lemma 3.20, the squared bias is bounded above by $(2^{-7} + 2^{-6})^2 \approx 5.4932 \times 10^{-4}$, implying that saturation occurs below a log-mean squared error of $\log_{10}(5.4932 \times 10^{-4}) \approx -3.2602$. It actually lies outside of the plotting range. The untruncated MLMC estimator is unbiased for any n_1 , yet it attains a larger mean squared error for all shown computational budgets due to its substantially higher variance. Although for sufficiently large n_1 this variance can be reduced below the squared bias of the truncated MLMC estimator, so that the

¹The discontinuity in Figure 2 (a) occurs because a large log-branching factor equal to $\ell = 10$ was sampled in the 17,212-th tree, which incurs a sampling cost that scales with 2^ℓ and also introduces a large importance weight $1/q_t(\ell)$ in (18).

untruncated estimator may eventually outperform its truncated counterpart with fixed truncation levels, the required sampling costs typically exceed practical computational limits. Moreover, with sufficiently large computational budgets, one could jointly increase n_1 as well as the truncation points, thereby reducing both variance and bias of the truncated estimator and preserving its advantage. The mean squared errors of the two MLMC estimators decay at comparable empirical rates of -0.8707 and -0.7871 for the truncated and untruncated variants, respectively, in agreement with the theoretical n_1^{-1} rate for smooth integrands (see Theorem 3.25 and [61, Corollary 2.3]). In contrast, the SAA estimators perform substantially worse, with decay rates of -0.3352 for SAA1 and -0.4936 for SAA2, consistent with the theoretical rates $n_1^{-1/3}$ and $n_1^{-1/2}$, respectively (see Theorem 3.9 and [51, Appendix G]). Taken together, these findings indicate that the truncated MLMC estimator performs best when computational resources are limited.



(a) MLMC estimators with 95% confidence intervals vs n_1 . (b) Mean squared error (MSE) vs total sampling cost.

Figure 2: Left panel: dependence of the untruncated and truncated MLMC estimators on n_1 . Right panel: trade-off between the mean squared error and the total sampling cost of all MLMC and SAA estimators. Each marker corresponds to an average of 10 independent estimators of the same type for a fixed n_1 .

Pricing Bermudan Options Consider a Bermudan basket option that grants its holder the right to sell an equally weighted basket of $m = 5$ assets at a fixed strike price $K = 100$ on one of $T = 4$ predetermined exercise dates $t \in [T]$. We use $\xi_t \in \mathbb{R}^m$ to denote the random vector of asset prices at time t and assume that the risk-free interest rate amounts to $\gamma = 0.05$. Exercising the option at time $t \in [T]$ yields an immediate payoff $g_t(\xi_t) = \max\{0, K - \frac{1}{m} e_m^\top \xi_t\}$, where $e_m \in \mathbb{R}^m$ denotes the vector of ones. Under the risk-neutral measure, the asset prices follow independent geometric Brownian motions with drift γ and volatility $\sigma = 0.2$. Hence, the asset prices at the exercise dates satisfy $\xi_{t+1} = \text{diag}(\varepsilon_t) \xi_t$, where the disturbances $\varepsilon_t \in \mathbb{R}^m$, $t \in [T-1]$, are serially independent and identically distributed with $\ln(\varepsilon_t) \sim \mathcal{N}((\gamma - \sigma^2/2) e_m, \sigma^2 I_m)$. We also assume that $\xi_1 = 100 e_m$ is deterministic. The price of the option is given by the risk-neutral expected net present value of the payoffs under the optimal exercise strategy, which can be found by solving an optimal stopping problem as described in Section 2. If x_t is the expected value of the option at time $t+1$ conditional on time- t information, then we have $x_{t-1} = \mathbb{E}_{t-1}[\max\{g_t(\xi_t), e^{-\gamma} x_t\}]$ for every $t \in [T]$. Thus, computing today's price x_0 amounts to solving the instance of (MCCE) with $f_t(\xi_t, x_t) = \max\{g_t(\xi_t), e^{-\gamma} x_t\}$ for all $t \in [T-1]$ and $f_T(\xi_T, x) = g_T(\xi_T)$. For a more detailed discussion of Bermudan options see [43].

The option pricing problem under consideration does not admit a closed-form solution and must therefore be tackled using sampling-based methods. Approximate prices obtained with classical approaches are reported in [43, Table 4]; for instance, least squares Monte Carlo and the stochastic grid method yield estimates of 2.163(0.001) and 2.141(0.008), respectively, where numbers in parentheses denote standard errors and do not account for bias. In contrast, a policy improvement method [7] produces a 95% confidence in-

terval of [2.154, 2.164] for the true option value. All of these classical methods are specifically designed for option pricing and do not readily extend to general MCCE problems. Furthermore, as noted in [72], their computational cost typically scales superlinearly with the dimension m , which limits their applicability to high-dimensional basket options, where m can be as large as 5,000 [6]. In [72] an untruncated MLMC estimator is shown to produce an unbiased price estimate of 2.161 for the Bermudan option under consideration. Further numerical experiments suggest that the computational cost of this estimator scales sublinearly with m . However, the analysis of this estimator heavily exploits the special structure of optimal stopping problems. As for any MLMC estimator, its efficiency depends critically on the rate parameters r_t governing the geometric distributions of the log-branching factors. If $r_t \leq 0.5$, then the estimator’s expected cost diverges. In addition, there exists a positive tolerance $\delta \leq 10^{-4}$ such that the estimator’s variance cannot be guaranteed to be finite for any $r_t \geq 0.5 + \delta$ [72, Theorem 2]. Within the theoretically admissible narrow range of r_t close to 0.5, the untruncated MLMC estimator typically incurs high computational costs. For instance, setting $r_t = 0.5001$ for $t \in [T - 1]$ ($\delta = 10^{-4}$) results in an expected cost of 2500.5^{T-1} per tree.

We now compute the option price using three truncated MLMC estimators with $M_t = 9, 10, 11$, as proposed in this paper, and two untruncated MLMC estimators with $M_t = \infty$ proposed in [72]. The selected truncation points for the truncated estimators empirically provide the best balance between bias and computational cost. For the truncated MLMC estimators, we adopt a heuristic proposed in [72] to choose identical rate parameters $r_t = r$ for all geometric distributions governing the log-branching factors. As discussed after Lemma 3.22, the rate r should approximately minimize the expected sampling cost while ensuring that the estimator’s variance does not exceed $\epsilon^2/2$. This is achieved by minimizing the work-normalized second moment [11], defined as $\mu_1^2(x) \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell$ (see also the proof of Lemma C.7). We approximate this quantity for each $r \in \{0.51, 0.52, \dots, 0.70\}$ by estimating the expectation in the definition (19) of $\mu_1^2(x)$ via a sample average over 10^6 independent copies of $\hat{H}_1(x)$, each computed as in Definition 3.17 but from a simplified and computationally cheap option pricing problem in which the average asset price $\frac{1}{m} e^\top \xi_t$ is replaced by a standard normal distribution. Next, we fit a piecewise linear convex function of r to the resulting approximate work-normalized second moments via a least squares procedure and select the rate parameter for all geometric distributions as a minimizer of this function. This yields $r = 0.59$ for the truncation points $M_t = 9$ and $M_t = 11$ and $r = 0.58$ for $M_t = 10$. These rates are consistent with the theory of Section 3.2, that is, they yield estimators with finite biases, variances and expected costs. As in [72], we use the same heuristic approach to find a near-optimal rate of $r = 0.60$ for the first untruncated MLMC estimator. However, this rate is inconsistent with the theory of [72], that is, for this choice of r the variance of the estimator is not guaranteed to be finite. The rate of the second untruncated MLMC estimator is set to $r = 0.5001$, which is the largest rate that guarantees the estimator to have a finite variance. Table 1 reports the option prices predicted by all truncated and untruncated MLMC estimators when the number of scenario trees is set to $n_1 = 5 \times 10^6$. Even though they only use half as many scenario trees as the methods in [7, 72], our truncated MLMC estimators provide comparable 95% confidence intervals. Note that the untruncated MLMC estimator with $r = 0.6$ offers no confidence interval because its variance is not guaranteed to be finite, while the untruncated MLMC estimator with $r = 0.5001$ suffers from an excessive sampling cost per tree.

Contextual Bandits In the last experiment we solve an instance of the off-policy learning problem (12) borrowed from [59], which determines one of two treatments to be administered to hospital patients suffering from acute ischaemic stroke. Each patient is characterized by a context vector $c \in \mathcal{C} \subseteq \mathbb{Z}_+^6$. The first context feature $c_1 \in \{0, 1, 2\}$ represents the patient’s level of consciousness (0: fully conscious, 1: drowsy, 2: unconscious), while $c_2 \in \{0, \dots, 4\}$, $c_3, c_4 \in \{0, 1\}$, $c_5 \in \{0, \dots, 5\}$ and $c_6 \in \{0, \dots, 3\}$ capture other categorical features such as the patient’s age etc. The overall context space \mathcal{C} is thus finite, and its cardinality equals 1,440. As in [59], we assume that there are two actions $a \in \mathcal{A} = \{1, 2\}$ (1: prescribing both aspirin

M_t	r_t	Cost per tree	Estimate (se)	Confidence interval
9	0.5900	22.6084	2.1684(0.0076)	[2.1536, 2.1832]
10	0.5800	29.5795	2.1562(0.0072)	[2.1421, 2.1703]
11	0.5900	26.3283	2.1641(0.0080)	[2.1485, 2.1797]
∞	0.5001	1.5634×10^{10}	2.1547(0.0073)	[2.1403, 2.1690]
∞	0.6000	27.0000	2.1589(-)	—

Table 1: MLMC estimators for pricing Bermudan basket options: expected costs per tree, estimated option prices with standard errors (se) and 95% confidence intervals for different truncation points and rates.

and heparin, 2: not administering any treatment), and we parametrize the context-dependent policy π_θ with $\theta \in [0, 1]^2$, where the agent selects action 1 with probability θ_1 if $c_1 \neq 0$ and with probability θ_2 if $c_1 = 0$.

We will solve the resulting instance of (MCCO) with a variant of the Adam optimizer [45] equipped with the MLMC gradient estimator of Definition 4.2. This allows us to assess our MLMC gradient estimator in an interesting scenario where we have no theoretical results guaranteeing its superiority over basic SAA estimators. Indeed, as explained in Section 2, the integrands f_1 , f_2 and f_3 involve logarithms as well as quadratic and exponential terms. Thus, they fail to satisfy the Lipschitz continuity and smoothness conditions detailed in Assumptions 3.2 and 3.3, which are needed for all theoretical results in Section 4.

We prescribe the joint distribution of c , u and y under \mathbb{P} synthetically, which allows us to solve problem (12) exactly. The resulting ground-truth solution serves as a benchmark against which any data-driven policies can be compared. First, we assume that the context c as well as the auxiliary random vector u follow independent uniform distributions on \mathcal{C} . Next, the mean of the cost vector y conditional on c is set to

$$\mathbb{E}[y|c] = \begin{cases} (3.0 + 5.0c_5 + p(c), 5.5 + 1.0c_5 + p(c)) & \text{if } c_1 = 0, \\ (1.7 + 3.5c_5 + p(c), 3.0 + 1.0c_5 + p(c)) & \text{if } c_1 \neq 0, \end{cases}$$

where $p(c) = 2.4 + 1.92(c_5/5 - 2.5)^2$ if $c_2 = 4$, $c_3 = 1$, $c_4 = 1$ and $c_6 = 3$; $p(c) = 0$ otherwise. Conditional on c , the actual cost vector y follows a bivariate log-normal distribution, that is, $\log(y)$ follows the normal distribution with mean $\log(\mathbb{E}[y|c]) - \frac{1}{2}\text{diag}(\Sigma)$ and covariance matrix Σ , where the diagonal and off-diagonal elements of Σ are set to 5.0 and 2.5, respectively. The heterogeneity of $\mathbb{E}[y|c]$ ensures that the exact optimal solution of problem (12) displays a sufficient level of complexity, whereas the positive correlation between the costs of the two actions captures the effect of an unobserved confounder that simultaneously affects both treatment outcomes. Finally, we set the smoothness temperature in (12) to $\mu = 2$, and we set the ambiguity radii to $r_c = 0.4$ and $r_y = 0.15$. This choice roughly accounts for a covariate shift whereby the subpopulation of patients with $c_1 = 0$ is over-sampled by 50%.

By leveraging known properties of log-normal distributions, the innermost integral in (12) can be evaluated analytically. As \mathcal{C} is finite, the objective of (12) thus reduces to a standard log-sum-exp function. Hence, the exact global minimizer $(\lambda^*, \theta_1^*, \theta_2^*) \approx (11.829, 0.589, 0.713)$ of problem (12) can be computed efficiently to any desired accuracy with any off-the-shelf convex optimization solver.

We now address problem (12) with oracle-based methods that have only sample access to \mathbb{P} . Specifically, we solve (12) with a variant of the Adam optimizer using learning rates of 0.75 for λ and 0.025 for θ_1 and θ_2 , alongside default decay rate parameters. To prevent Adam’s momentum from accumulating against a hard boundary and stalling convergence, we enforce the $\lambda \geq 0$ constraint via a softplus reparameterization rather than a standard projection. Concretely, we set $\lambda = \log(1 + \exp(\lambda'))$ for some $\lambda' \in \mathbb{R}$ and run Adam entirely on λ' . However, we simply project θ_1 and θ_2 to $[0, 1]$ in each iteration of the algorithm. We also add the L_2 regularization term $0.005 \cdot \|\theta\|_2^2$ to the objective of (12) to enhance stability. As stochastic gradient estimators we use the SAA-type estimator proposed in [59] (adapted to account for stochastic costs) as well as the MLMC gradient estimator of Definition 4.2. As the Lipschitz continuity and smoothness conditions re-

quired to control the bias and variance of the MLMC estimator fail to hold in this example, we clip the norms of the truncated MLMC gradient estimators with respect to λ , θ_1 and θ_2 at 100, 50 and 50, respectively. Note that problem (12) as described in [59] constitutes an instance of (MCCO) with $T = 3$. However, the numerical experiments in [59] focus on a simplified model with deterministic costs and $T = 2$ —presumably due to the hardness of the original three-stage problem. In contrast, our experiments faithfully capture the three-stage nature of the more challenging original problem with stochastic costs.

Figure 3 illustrates the convergence of λ , θ_1 and θ_2 as a function of the cumulative number of scenarios over 2,000 Adam iterations and for different choices of the estimators’ hyperparameters. Solid lines and shaded regions represent means as well as corresponding 95% confidence intervals obtained from 20 independent simulation runs, whereas dotted lines represent ground-truth minimizers. We observe that Adam converges significantly faster when equipped with MLMC estimators instead of SAA estimators. Recall from Lemma 4.6 that the truncated MLMC estimator is biased, and note that gradient clipping introduces an additional bias. Nevertheless, MLMC estimators lead to less biased optimizers than SAA estimators. Even though Assumptions 3.2 and 3.3 fail to hold, MLMC estimators thus enable faster and tighter convergence to the true optimal solution than SAA estimators equipped with a comparable computational budget.

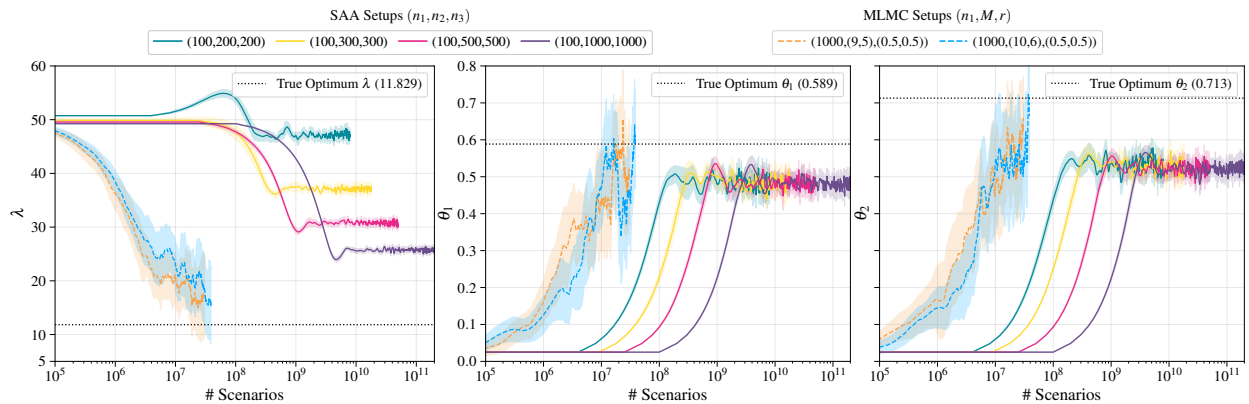


Figure 3: Convergence of the decision variables λ (left), θ_1 (middle), and θ_2 (right) generated by Adam. Results show mean trajectories with 95% confidence intervals based on 20 independent simulation runs.

Acknowledgments This work was supported as a part of NCCR Automation, a National Centre of Competence in Research, funded by the Swiss National Science Foundation (grant number 51NF40_225155). We thank Guanyang Wang for providing the code for the experiments in [61], as well as Yi Shen and Pan Xu for the code corresponding to [59]. We also thank Guanyang Wang, Lie He and Niao He for helpful discussions.

References

- [1] Charalambos D Aliprantis and Kim C Border. *Infinite Dimensional Analysis: A Hitchhiker’s Guide*. Springer, 2006.
- [2] Yossi Arjevani, Yair Carmon, John C Duchi, Dylan J Foster, Nathan Srebro, and Blake Woodworth. Lower bounds for non-convex stochastic optimization. *Mathematical Programming*, 199(1–2):165–214, 2023.

- [3] Waïss Azizian, Franck Iutzeler, and Jérôme Malick. Regularization for Wasserstein distributionally robust optimization. *ESAIM: Control, Optimisation and Calculus of Variations*, 29:1–33, 2023.
- [4] Krishnakumar Balasubramanian, Saeed Ghadimi, and Anthony Nguyen. Stochastic multilevel composition optimization algorithms with level-independent convergence rates. *SIAM Journal on Optimization*, 32(2):519–544, 2022.
- [5] Joakim Beck, Ben Mansour Dia, Luis Espath, and Raúl Tempone. Multilevel double loop Monte Carlo and stochastic collocation methods with importance sampling for Bayesian optimal experimental design. *International Journal for Numerical Methods in Engineering*, 121(15):3482–3503, 2020.
- [6] Sebastian Becker, Patrick Cheridito, Arnulf Jentzen, and Timo Welti. Solving high-dimensional optimal stopping problems using deep learning. *European Journal of Applied Mathematics*, 32(3):470–514, 2021.
- [7] Christian Bender, Anastasia Kolodko, and John Schoenmakers. Policy iteration for American options: Overview. *Monte Carlo Methods and Applications*, 12(5):347–362, 2006.
- [8] Andrew Bennett, Nathan Kallus, and Tobias Schnabel. Deep generalized method of moments for instrumental variable analysis. In *Advances in Neural Information Processing Systems*, pages 3564–3574, 2019.
- [9] Dimitri P Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, 3rd edition, 1995.
- [10] Jose Blanchet, Donald Goldfarb, Garud Iyengar, Fengpei Li, and Chaoxu Zhou. Unbiased simulation for optimizing stochastic function compositions. *arXiv:1711.07564*, 2017.
- [11] Jose H Blanchet and Peter W Glynn. Unbiased Monte Carlo for optimization and functions of expectations via multi-level randomization. In *Winter Simulation Conference*, pages 3656–3667, 2015.
- [12] Mark Broadie, Yiping Du, and Ciamac C Moallemi. Efficient risk estimation via nested sequential simulation. *Management Science*, 57(6):1172–1194, 2011.
- [13] Karolina Bujok, Ben M Hambly, and Christoph Reisinger. Multilevel simulation of functionals of Bernoulli random variables with application to basket credit derivatives. *Methodology and Computing in Applied Probability*, 17:579–604, 2015.
- [14] Tianyi Chen, Yuejiao Sun, and Wotao Yin. Solving stochastic compositional optimization is nearly as easy as solving stochastic optimization. *IEEE Transactions on Signal Processing*, 69:4937–4948, 2021.
- [15] Xuxing Chen, Abhishek Roy, Yifan Hu, and Krishnakumar Balasubramanian. Stochastic optimization algorithms for instrumental variable regression with streaming data. In *Advances in Neural Information Processing Systems*, pages 26510–26542, 2024.
- [16] Dragos Florin Ciocan and Velibor V Mišić. Interpretable optimal stopping. *Management Science*, 68(3):1616–1638, 2022.
- [17] Weilin Cong, Rana Forsati, Mahmut Kandemir, and Mehrdad Mahdavi. Minimal variance sampling with provable guarantees for fast training of graph neural networks. In *International Conference on Knowledge Discovery & Data Mining*, pages 1393–1403, 2020.

- [18] Bo Dai, Niao He, Yunpeng Pan, Byron Boots, and Le Song. Learning from conditional distributions via dual embeddings. In *Artificial Intelligence and Statistics*, pages 1458–1467, 2017.
- [19] Bo Dai, Albert Shaw, Lihong Li, Lin Xiao, Niao He, Zhen Liu, Jianshu Chen, and Le Song. SBEDD: Convergent reinforcement learning with nonlinear function approximation. In *International Conference on Machine Learning*, pages 1125–1134, 2018.
- [20] Martin Dyer and Leen Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106(3):423–432, 2006.
- [21] Hongchang Gao. Decentralized multi-level compositional optimization algorithms with level-independent convergence rate. In *International Conference on Artificial Intelligence and Statistics*, pages 4402–4410, 2024.
- [22] Saeed Ghadimi, Guanghui Lan, and Hongchao Zhang. Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization. *Mathematical Programming*, 155(1):267–305, 2016.
- [23] Michael B Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.
- [24] Michael B Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, 2015.
- [25] Michael B Giles. MLMC for nested expectations. In Josef Dick, Frances Y. Kuo, and Henryk Woźniakowski, editors, *Contemporary Computational Mathematics: A Celebration of the 80th Birthday of Ian Sloan*, pages 425–442. Springer, 2018.
- [26] Michael B Giles and Abdul-Lateef Haji-Ali. Multilevel nested simulation for efficient risk estimation. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2):497–525, 2019.
- [27] Michael B Giles and Lukasz Szpruch. Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation. *Annals of Applied Probability*, 24(4):1585–1620, 2014.
- [28] Michael B Giles, Abdul-Lateef Haji-Ali, and Jonathan Spence. Efficient risk estimation for the credit valuation adjustment. *arXiv:2301.05886*, 2023.
- [29] Takashi Goda and Wataru Kitade. Constructing unbiased gradient estimators with finite variance for conditional stochastic optimization. *Mathematics and Computers in Simulation*, 204:743–763, 2023.
- [30] Takashi Goda, Tomohiko Hironaka, and Takeru Iwamoto. Multilevel Monte Carlo estimation of expected information gains. *Stochastic Analysis and Applications*, 38(4):581–600, 2020.
- [31] Takashi Goda, Tomohiko Hironaka, Wataru Kitade, and Adam Foster. Unbiased MLMC stochastic gradient-based optimization of Bayesian experimental designs. *SIAM Journal on Scientific Computing*, 44(1):A286–A311, 2022.
- [32] Michael B Gordy and Sandeep Juneja. Nested simulation in portfolio risk measurement. *Management Science*, 56(10):1833–1848, 2010.
- [33] Abdul-Lateef Haji-Ali and Jonathan Spence. Nested multilevel Monte Carlo with biased and antithetic sampling. *arXiv:2308.07835*, 2023.
- [34] Grani A Hanasusanto, Daniel Kuhn, and Wolfram Wiesemann. A comment on “computational complexity of stochastic programming problems”. *Mathematical Programming*, 159(1–2):557–569, 2016.

- [35] Lars Peter Hansen and Thomas J Sargent. *Robustness*. Princeton University Press, 2008.
- [36] Jason Hartford, Greg Lewis, Kevin Leyton-Brown, and Matt Taddy. Deep IV: A flexible approach for counterfactual prediction. In *International Conference on Machine Learning*, pages 1414–1423, 2017.
- [37] Lie He and Shiva Kasiviswanathan. Debiasing conditional stochastic optimization. In *Advances in Neural Information Processing Systems*, pages 78846–78893, 2023.
- [38] Roger A Horn and Charles R Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [39] Yifan Hu, Xin Chen, and Niao He. Sample complexity of sample average approximation for conditional stochastic optimization. *SIAM Journal on Optimization*, 30(3):2103–2133, 2020.
- [40] Yifan Hu, Siqi Zhang, Xin Chen, and Niao He. Biased stochastic first-order methods for conditional stochastic optimization and applications in meta learning. In *Advances in Neural Information Processing Systems*, pages 2759–2770, 2020.
- [41] Yifan Hu, Xin Chen, and Niao He. On the bias-variance-cost tradeoff of stochastic optimization. In *Advances in Neural Information Processing Systems*, pages 22119–22131, 2021.
- [42] Yifan Hu, Jie Wang, Xin Chen, and Niao He. Multi-level Monte-Carlo gradient methods for stochastic optimization with biased oracles. *arXiv:2408.11084*, 2024.
- [43] Shashi Jain and Cornelis W. Oosterlee. Pricing high-dimensional Bermudan options using the stochastic grid method. *International Journal of Computer Mathematics*, 89(9):1186–1211, 2012.
- [44] Wei Jiang, Bokun Wang, Yibo Wang, Lijun Zhang, and Tianbao Yang. Optimal algorithms for stochastic multi-level compositional optimization. In *International Conference on Machine Learning*, pages 10195–10216, 2022.
- [45] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations*, 2015.
- [46] Daniel Kuhn, Soroosh Shafiee, and Wolfram Wiesemann. Distributionally robust optimization. *Acta Numerica*, 34:579–804, 2025.
- [47] Tze Leung Lai. Optimal stopping and sequential tests which minimize the maximum expected sample size. *Annals of Statistics*, pages 659–673, 1973.
- [48] Yifan Lin, Yuxuan Ren, and Enlu Zhou. Bayesian risk Markov decision processes. In *Advances in Neural Information Processing Systems*, pages 17430–17442, 2022.
- [49] Krikamol Muandet, Arash Mehrjou, Si Kai Lee, and Anant Raj. Dual instrumental variable regression. In *Advances in Neural Information Processing Systems*, pages 2710–2721, 2020.
- [50] Emin Ozyoruk, Nesim Kohen Erkip, and Çağın Ararat. End-of-life inventory management problem: Results and insights. *International Journal of Production Economics*, 243:108313, 2022.
- [51] Tom Rainforth, Rob Cornish, Hongseok Yang, Andrew Warrington, and Frank Wood. On nesting Monte Carlo estimators. In *International Conference on Machine Learning*, pages 4267–4276, 2018.
- [52] Marcus de Mendes C. R. Reaiche. A note on sample complexity of multistage stochastic programs. *Operations Research Letters*, 44(4):430–435, 2016.

- [53] Chang-han Rhee and Peter W Glynn. Unbiased estimation with square root convergence for SDE models. *Operations Research*, 63(5):1026–1043, 2015.
- [54] Andrzej Ruszczyński. A stochastic subgradient method for nonsmooth nonconvex multilevel composition optimization. *SIAM Journal on Control and Optimization*, 59(3):2301–2320, 2021.
- [55] Andrzej Ruszczyński and Alexander Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31(3):544–561, 2006.
- [56] Alexander Shapiro. On complexity of multistage stochastic programs. *Operations Research Letters*, 34(1):1–8, 2006.
- [57] Alexander Shapiro and Arkadi Nemirovski. On complexity of stochastic programming problems. In Vaithilingam Jeyakumar and Alexander Rubinov, editors, *Continuous Optimization*, pages 111–146. Springer, Boston, MA, 2005.
- [58] Alexander Shapiro, Enlu Zhou, and Yifan Lin. Bayesian distributionally robust optimization. *SIAM Journal on Optimization*, 33(2):1279–1304, 2023.
- [59] Yi Shen, Pan Xu, and Michael M Zavlanos. Wasserstein distributionally robust policy evaluation and learning for contextual bandits. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856. Featured Certification.
- [60] Rahul Singh, Maneesh Sahani, and Arthur Gretton. Kernel instrumental variable regression. In *Advances in Neural Information Processing Systems*, pages 4593–4605, 2019.
- [61] Yasa Syed and Guanyang Wang. Optimal randomized multilevel Monte Carlo for repeatedly nested expectations. In *International Conference on Machine Learning*, pages 33343–33364, 2023.
- [62] Emanuel Todorov and Michael I. Jordan. Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5(11):1226–1235, 2002.
- [63] Nicky D Van Foreest and Onur A Kilic. An intuitive approach to inventory control with optimal stopping. *European Journal of Operational Research*, 311(3):921–924, 2023.
- [64] Martin J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.
- [65] Jie Wang, Rui Gao, and Yao Xie. Sinkhorn distributionally robust optimization. *Operations Research*, 2025. Forthcoming.
- [66] Tianze Wang and Guanyang Wang. Unbiased Multilevel Monte Carlo methods for intractable distributions: MLMC meets MCMC. *Journal of Machine Learning Research*, 24(249):1–40, 2023.
- [67] Yuhao Wang and Enlu Zhou. Bayesian risk-averse Q-learning with streaming observations. In *Advances in Neural Information Processing Systems*, pages 75967–75992, 2024.
- [68] Peter Whittle. *Risk-Sensitive Optimal Control*. Wiley, 1990.
- [69] Tesi Xiao, Krishnakumar Balasubramanian, and Saeed Ghadimi. A projection-free algorithm for constrained stochastic multi-level composition optimization. In *Advances in Neural Information Processing Systems*, pages 19984–19996, 2022.

- [70] Shuoguang Yang, Mengdi Wang, and Ethan X Fang. Multilevel stochastic gradient methods for nested composition optimization. *SIAM Journal on Optimization*, 29(1):616–659, 2019.
- [71] Junyu Zhang and Lin Xiao. Multilevel composite stochastic optimization via nested variance reduction. *SIAM Journal on Optimization*, 31(2):1131–1157, 2021.
- [72] Zhengqing Zhou, Guanyang Wang, Jose H Blanchet, and Peter W Glynn. Unbiased optimal stopping via the MUSE. *Stochastic Processes and their Applications*, 166:104088, 2023.

Appendix

A Auxiliary Results

The following lemma establishes a uniform deviation bound based on covering numbers. It is a standard result in stochastic programming, and we include a concise proof to keep the presentation self-contained.

Lemma A.1 ([56, Proof of Theorem 1]). *Consider the problem of minimizing an L -Lipschitz continuous function $F : \mathbb{R}^d \rightarrow \mathbb{R}$ over a compact set $\mathcal{X} \subseteq \mathbb{R}^d$ with diameter $D_{\mathcal{X}}$. Let $\widehat{F} : \mathbb{R}^d \rightarrow \mathbb{R}$ be an L -Lipschitz continuous random function that approximates F . Then, for every $v > 0$, there is a v -net $\mathcal{X}_v \subseteq \mathcal{X}$ such that $|\mathcal{X}_v| \leq \lceil 2D_{\mathcal{X}}/v + 1 \rceil^d$ and $\min_{x' \in \mathcal{X}_v} \|x - x'\|_2 \leq v$ for every $x \in \mathcal{X}$. If $\epsilon > 0$ and $v = \frac{\epsilon}{4L}$, then we have*

$$\mathbb{P}\left(\sup_{x \in \mathcal{X}} |\widehat{F}(x) - F(x)| > \epsilon\right) \leq \sum_{x' \in \mathcal{X}_v} \mathbb{P}\left(|\widehat{F}(x') - F(x')| > \frac{\epsilon}{2}\right).$$

Proof. The set \mathcal{X} is contained in a ball of radius $D_{\mathcal{X}}$ and can thus be covered by $\lceil 2D_{\mathcal{X}}/v + 1 \rceil^d$ smaller balls of diameter v [64, Lemma 5.7]. By construction, the centroids of these smaller balls, projected onto \mathcal{X} , constitute a v -net \mathcal{X}_v . That is, for every $x \in \mathcal{X}$, there is an $x' \in \mathcal{X}_v$ in its closed v -neighborhood. Assume now that $v = \frac{\epsilon}{4L}$ for some $\epsilon > 0$. Then, for every $x \in \mathcal{X}$ and $x' \in \operatorname{argmin}_{x' \in \mathcal{X}_v} \|x - x'\|_2$, we have

$$|\widehat{F}(x) - F(x)| \leq |\widehat{F}(x) - \widehat{F}(x')| + |\widehat{F}(x') - F(x')| + |F(x') - F(x)| \leq |\widehat{F}(x') - F(x')| + \frac{\epsilon}{2},$$

where the second inequality follows from the L -Lipschitz continuity of F and \widehat{F} and from the observation that $\|x' - x\|_2 \leq v$. As the above bound holds for every $x \in \mathcal{X}$, we readily obtain

$$\sup_{x \in \mathcal{X}} |\widehat{F}(x) - F(x)| \leq \frac{\epsilon}{2} + \max_{x' \in \mathcal{X}_v} |\widehat{F}(x') - F(x')|.$$

This further implies via the union bound that

$$\mathbb{P}\left(\sup_{x \in \mathcal{X}} |\widehat{F}(x) - F(x)| > \epsilon\right) \leq \mathbb{P}\left(\max_{x' \in \mathcal{X}_v} |\widehat{F}(x') - F(x')| > \frac{\epsilon}{2}\right) \leq \sum_{x' \in \mathcal{X}_v} \mathbb{P}\left(|\widehat{F}(x') - F(x')| > \frac{\epsilon}{2}\right).$$

Hence, the claim follows. \square

The following matrix norm bounds will be used repeatedly in proving the main results of this paper.

Lemma A.2 ([38, § 5.4.P3]). *If $A \in \mathbb{R}^{m \times n}$ and $1 \leq p < q \leq \infty$, then $\|A\|_q \leq \|A\|_p \leq (mn)^{\frac{1}{p} - \frac{1}{q}} \|A\|_q$.*

Lemma A.3 ([38, § 5.6.0.2]). *If $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times k}$, then $\|AB\|_2 \leq \|A\|_2 \|B\|_2$.*

Finally, we state a well-known corollary of the Marcinkiewicz-Zygmund inequality as an auxiliary lemma, which is useful for deriving variance bounds of the estimators proposed in the main text.

Lemma A.4. *Assume that Y and Z are random matrices of arbitrary dimensions with finite p -th moments for some $p \geq 2$, and let Z_i , $i \in [n]$, be independent samples from the conditional distribution $\mathbb{P}_{Z|Y}$. Then there exists a constant B_p depending only on p such that*

$$\mathbb{E}_{Z|Y} \left[\left\| \frac{1}{n} \sum_{i=1}^n Z_i - \mathbb{E}_{Z|Y}[Z] \right\|_p^p \right] \leq \frac{B_p}{n^{p/2}} \mathbb{E}_{Z|Y} [\|Z\|_p^p].$$

The proof parallels that of [72, Corollary 2], generalized to matrices. Details are omitted for brevity.

B Proofs of Section 2

Proof of Lemma 2.1. We prove the claim by backward induction on t . At $t = T$ we have $Q_T(s, a) = s^\top P_T s$ by assumption, which is of the form (6a) if the coefficients satisfy (6c). This establishes the base step. As for the induction step, assume that (6a) holds for Q_{t+1} . Thus, we find

$$\min_{a' \in \mathbb{R}^n} Q_{t+1}(s', a') = s'^\top P_{t+1} s' + 2 g_{t+1}^\top s' + d_{t+1},$$

with the Schur-complement expressions for P_{t+1} , g_{t+1} and d_{t+1} given in (6d). Substituting the above formula with $s' = As + Ba + \xi_{t+1}$ into the Bellman recursion (4) and expanding the quadratic and linear terms shows that $Q_t(s, a)$ is again a quadratic function in (s, a) . Collecting coefficients of $s^\top s$, $s^\top a$, $a^\top a$, s , a and the constant term then yields exactly the expressions (6b). Thus the quadratic structure of the Q-function is preserved backward in time, completing the induction. \square

The proof of Lemma 2.1 readily reveals that the optimal control at time t is given by

$$a_t^* = -(Q_t^{aa})^{-1} ((Q_t^{sa})^\top s + b_t^a)$$

and thus constitutes an affine function of the state s with \mathcal{F}_t -measurable random coefficients.

C Proofs of Section 3

To streamline the proofs in Section 3, we first introduce some key notation and conventions that will be useful for our derivations. Throughout this discussion we use \mathcal{L}_t as shorthand for the space $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{d_t})$ of all \mathbb{P} -integrable d_t -dimensional random vectors on $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition C.1. *For any $t \in [T]$ and $s \in [t+1]$, we define the conditional risk mapping $F_s^t : \mathcal{L}_t \rightarrow \mathcal{L}_{s-1}$ via*

$$F_s^t(z_t) := \begin{cases} \mathbb{E}_{s-1} [f_s(\xi_s, \dots, \mathbb{E}_{t-1} [f_t(\xi_t, z_t)] \dots)] & \text{if } s < t+1, \\ z_t & \text{if } s = t+1. \end{cases}$$

Note that F_s^t maps d_t -dimensional random vectors to d_{s-1} -dimensional random vectors. By construction, we have $F^t(x) = F_1^T(x)$ for every (deterministic) vector $x \in \mathbb{R}^d$. If Assumptions 3.1 and 3.2 hold, then $F_s^t(z_t)$ is well-defined and constitutes indeed an element of \mathcal{L}_{s-1} for every $z_t \in \mathcal{L}_t$.

Lemma C.2. *If Assumptions 3.1 and 3.2 hold, then $F_s^t(z_t) \in \mathcal{L}_{s-1}$ for all $z_t \in \mathcal{L}_t$, $t \in [T]$ and $s \in [t+1]$.*

Proof. We fix $t \in [T]$ and $z_t \in \mathcal{L}_t$ and prove the claim by induction on s . If $s = t + 1$, then we have $F_s^t(z_t) = z_t \in \mathcal{L}_t$, and thus the base case is trivially true. As for the induction step corresponding to any $s \in [t]$, assume that $F_{s+1}^t(z_t) \in \mathcal{L}_s$. Then, $f_s(\xi_s, F_{s+1}^t(z_t))$ is \mathcal{F} -measurable thanks to [1, Lemma 4.51] and because measurability is preserved under Borel-measurable transformations. Also, we have

$$\begin{aligned} \mathbb{E} [\|F_s^t(z_t)\|_2] &= \mathbb{E} [\|\mathbb{E}_{s-1}[f_s(\xi_s, F_{s+1}^t(z_t))]\|_2] \\ &\leq \mathbb{E} [\|f_s(\xi_s, F_{s+1}^t(z_t))\|_2] \\ &\leq \mathbb{E} [\|f_s(\xi_s, F_{s+1}^t(z_t)) - f_s(\xi_s, 0)\|_2 + \|f_s(\xi_s, 0)\|_2] \\ &\leq L_s \mathbb{E} [\|F_{s+1}^t(z_t)\|_2] + \mathbb{E} [\|f_s(\xi_s, 0)\|_2] < \infty, \end{aligned}$$

where the first inequality follows from Jensen's inequality and the law of total expectation, while the second and the third inequalities follow from the triangle inequality and Assumption 3.2, respectively. The resulting bound is finite by the induction hypothesis, which assumes that $F_{s+1}^t(z_t) \in \mathcal{L}_s$, and by Assumption 3.1. We may thus conclude that $F_s^t(z_t) \in \mathcal{L}_{s-1}$. This observation completes the induction step. \square

Lemma C.2 ensures that the function $F(x)$ is well-defined. If Assumptions 3.1 and 3.2 hold, then all conditional risk mappings of Definition C.1 are Lipschitz continuous. If additionally Assumption 3.3 holds, then the conditional risk mappings are also Lipschitz smooth.

Lemma C.3. Fix any $t \in [T]$ and $s \in [t + 1]$. If Assumptions 3.1 and 3.2 hold, then F_s^t is an $L_{[s:t]}$ -Lipschitz continuous function on the space \mathcal{L}_t with $L_{[s:t]} := \prod_{i=s}^t L_i$ if $s < t + 1$; $:= 1$ if $s = t + 1$. This means that

$$\|F_s^t(z_t) - F_s^t(z'_t)\|_2 \leq L_{[s:t]} \mathbb{E}_{s-1} [\|z_t - z'_t\|_2] \quad \forall z_t, z'_t \in \mathcal{L}_t.$$

If additionally Assumption 3.3 holds, then F_s^t is an $S_{[s:t]}$ -smooth function on the space \mathcal{L}_t with $S_{[s:t]} := \sum_{i=s}^t S_i L_{[s:i-1]} L_{[i+1:t]}^2$ if $s < t + 1$; $:= 0$ if $s = t + 1$. This means that

$$\|\nabla F_s^t(z_t) - \nabla F_s^t(z'_t)\|_2 \leq S_{[s:t]} \mathbb{E}_{s-1} [\|z_t - z'_t\|_2] \quad \forall z_t, z'_t \in \mathcal{L}_t.$$

Proof. We fix $t \in [T]$ and use backward induction on s to prove that F_s^t is $L_{[s:t]}$ -Lipschitz continuous. In the base case corresponding to $s = t + 1$, F_s^t collapses to the identity mapping on \mathcal{L}_t , and thus the claim is trivially true. As for the induction step corresponding to any $s \in [t]$, assume that F_{s+1}^t is $L_{[s+1:t]}$ -Lipschitz continuous. Since $F_s^t(z_t) = \mathbb{E}_{s-1}[f_s(\xi_s, F_{s+1}^t(z_t))]$, we thus have

$$\begin{aligned} \|F_s^t(z_t) - F_s^t(z'_t)\|_2 &\leq \mathbb{E}_{s-1} [\|f_s(\xi_s, F_{s+1}^t(z_t)) - f_s(\xi_s, F_{s+1}^t(z'_t))\|_2] \\ &\leq L_s \mathbb{E}_{s-1} [\|F_{s+1}^t(z_t) - F_{s+1}^t(z'_t)\|_2] \\ &\leq L_s L_{[s+1:t]} \mathbb{E}_{s-1} [\|z_t - z'_t\|_2] = L_{[s:t]} \mathbb{E}_{s-1} [\|z_t - z'_t\|_2] \end{aligned}$$

for any $z_t, z'_t \in \mathcal{L}_t$. Here, the first two inequalities follow from Jensen's inequality and from Assumption 3.2, respectively, while the third inequality exploits the induction hypothesis and the law of total expectation. This completes the induction step, and thus the claim follows.

Suppose now that Assumption 3.3 holds, too. One readily verifies that the smoothness constants obey the recursion $S_{[s:t]} = L_s S_{[s+1:t]} + S_s L_{[s+1:t]}^2$ for all $s \in [t]$ and $t \in [T]$. In the remainder we fix any $t \in [T]$ and use backward induction on s to prove that F_s^t is $S_{[s:t]}$ -smooth. In the base case corresponding to $s = t + 1$, the claim is trivially true because F_s^t is the identity mapping on \mathcal{L}_t . As for the induction step corresponding to any $s \in [t]$, assume that F_{s+1}^t is $S_{[s+1:t]}$ -smooth. By the definition of $F_s^t(z_t)$, we have

$$\nabla F_s^t(z_t) = \nabla \mathbb{E}_{s-1} [f_s(\xi_s, F_{s+1}^t(z_t))] = \mathbb{E}_{s-1} [\nabla_{z_t} f_s(\xi_s, F_{s+1}^t(z_t))].$$

Note that the gradient may be interchanged with the conditional expectation by the dominated convergence theorem, which applies thanks to Assumption 3.2. Jensen's inequality and the chain rule then imply that

$$\begin{aligned} \|\nabla F_s^t(z_t) - \nabla F_s^t(z'_t)\|_2 &\leq \mathbb{E}_{s-1} [\|\nabla_{z_t} f_s(\xi_s, F_{s+1}^t(z_t)) - \nabla_{z_t} f_s(\xi_s, F_{s+1}^t(z'_t))\|_2] \\ &= \mathbb{E}_{s-1} [\|\nabla F_{s+1}^t(z_t) \nabla_{x_s} f_s(\xi_s, F_{s+1}^t(z_t)) - \nabla F_{s+1}^t(z'_t) \nabla_{x_s} f_s(\xi_s, F_{s+1}^t(z'_t))\|_2]. \end{aligned}$$

Next, we may use the triangle inequality to conclude that

$$\begin{aligned} \|\nabla F_s^t(z_t) - \nabla F_s^t(z'_t)\|_2 &\leq \mathbb{E}_{s-1} [\|(\nabla F_{s+1}^t(z_t) - \nabla F_{s+1}^t(z'_t)) \nabla_{x_s} f_s(\xi_s, F_{s+1}^t(z_t))\|_2] \\ &\quad + \mathbb{E}_{s-1} [\|\nabla F_{s+1}^t(z'_t) (\nabla_{x_s} f_s(\xi_s, F_{s+1}^t(z_t)) - \nabla_{x_s} f_s(\xi_s, F_{s+1}^t(z'_t)))\|_2] \\ &\leq L_s \mathbb{E}_{s-1} [\|\nabla F_{s+1}^t(z_t) - \nabla F_{s+1}^t(z'_t)\|_2] \\ &\quad + S_s \mathbb{E}_{s-1} [\|\nabla F_{s+1}^t(z'_t)\|_2 \|F_{s+1}^t(z_t) - F_{s+1}^t(z'_t)\|_2] \\ &\leq (L_s S_{[s+1:t]} + S_s L_{[s+1:t]}^2) \|z_t - z'_t\|_2 = S_{[s:t]} \|z_t - z'_t\|_2, \end{aligned}$$

where the second inequality follows from Assumptions 3.2 and 3.3 and from the Cauchy-Schwarz inequality. The third inequality exploits the induction hypothesis, whereby F_{s+1}^t is $S_{[s+1:t]}$ -smooth, as well as the first part of the proof, whereby F_{s+1}^t is $L_{[s+1:t]}$ -Lipschitz continuous. Finally, the equality follows from the recursive formula for $S_{[s:t]}$ established earlier. Hence, $F_s^t(z_t)$ is indeed $S_{[s:t]}$ -smooth in z_t . \square

In analogy to Definition C.1, we can now introduce *empirical* conditional risk mappings on \mathcal{L}_t .

Definition C.4. For any $t \in [T]$ and $s \in [t+1]$, we define the conditional risk mapping $\widehat{F}_s^t : \mathcal{L}_t \rightarrow \mathcal{L}_{s-1}$ via

$$\widehat{F}_s^t(z_t) := \begin{cases} \widehat{\mathbb{E}}_{s-1} [f_s(\xi_s, \dots, \widehat{\mathbb{E}}_{t-1} [f_t(\xi_t, z_t)] \dots)] & \text{if } s < t+1, \\ z_t & \text{if } s = t+1. \end{cases}$$

Here, $\widehat{\mathbb{E}}_{t-1}[\cdot]$, $t \in [T]$, denote the empirical conditional expectation operators defined in (15).

Note that $\widehat{F}(x) = \widehat{F}_1^T(x)$ for every (deterministic) vector $x \in \mathbb{R}^d$. If Assumptions 3.1 and 3.2 hold, one can proceed as in Lemmas C.2 and C.3 to show that the conditional risk mapping \widehat{F}_s^t is well-defined and $L_{[s:t]}$ -Lipschitz continuous for any $t \in [T]$ and $s \in [t+1]$. In fact, the proofs simplify because \mathbb{P} -integrability and Lipschitz continuity are trivially preserved under finite sums. Details are omitted for brevity. We are now ready to prove the formal results of Section 3.1.

Proof of Lemma 3.6. Suppose first that only Assumptions 3.1, 3.2 and 3.5 hold. By the construction of the conditional risk mappings in Definitions C.1 and C.4, we have

$$F(x) = F_1^T(x) = F_1^T(\widehat{F}_{T+1}^T(x)) \quad \text{and} \quad \mathbb{E}[\widehat{F}(x)] = \mathbb{E}[\widehat{F}_1^T(x)] = F_1^1(\widehat{F}_2^T(x))$$

for all $x \in \mathbb{R}^d$. We can thus represent the bias of the SAA estimator $\widehat{F}(x)$ as a telescoping sum and then use the triangle inequality to conclude that

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \sum_{t=1}^{T-1} \left| F_1^t(\widehat{F}_{t+1}^T(x)) - F_1^{t+1}(\widehat{F}_{t+2}^T(x)) \right|. \quad (27)$$

The t -th term in the resulting upper bound satisfies

$$\begin{aligned} \left| F_1^t(\widehat{F}_{t+1}^T(x)) - F_1^{t+1}(\widehat{F}_{t+2}^T(x)) \right| &= \left| F_1^t(\widehat{F}_{t+1}^T(x)) - F_1^t(\mathbb{E}_t[\widehat{F}_{t+1}^T(x)]) \right| \\ &\leq L_{[t]} \mathbb{E} [\|\widehat{F}_{t+1}^T(x) - \mathbb{E}_t[\widehat{F}_{t+1}^T(x)]\|_2] \\ &\leq L_{[t]} \mathbb{E} [\|\widehat{F}_{t+1}^T(x) - \mathbb{E}_t[\widehat{F}_{t+1}^T(x)]\|_2^2]^{\frac{1}{2}}, \end{aligned}$$

where the two inequalities follow from the Lipschitz continuity of F_1^t established in Lemma C.3 and from Jensen's inequality, respectively. By the definition of \widehat{F}_t^T , we then have

$$\begin{aligned} & L_{[t]} \mathbb{E} \left[\left\| \widehat{F}_{t+1}^T(x) - \mathbb{E}_t[\widehat{F}_{t+1}^T(x)] \right\|_2^2 \right]^{\frac{1}{2}} \\ &= L_{[t]} \mathbb{E} \left[\left\| \mathbb{E}_t[f_{t+1}(\xi_{t+1}, \widehat{F}_{t+2}^T(x))] - \mathbb{E}_t[f_{t+1}(\xi_{t+1}, \widehat{F}_{t+2}^T(x))] \right\|_2^2 \right]^{\frac{1}{2}} \\ &= \frac{L_{[t]}}{\sqrt{n_{t+1}}} \mathbb{E} \left[\left\| f_{t+1}(\xi_{t+1}, \widehat{F}_{t+2}^T(x)) - \mathbb{E}_t[f_{t+1}(\xi_{t+1}, \widehat{F}_{t+2}^T(x))] \right\|_2^2 \right]^{\frac{1}{2}} \leq \frac{L_{[t]} \sigma_{t+1}}{\sqrt{n_{t+1}}}, \end{aligned}$$

where the second equality holds because, conditional on $\xi_{[t]}$, the variance of an average of n_{t+1} i.i.d. random variables coincides with the variance of a single random variable divided by n_{t+1} . The inequality in the last line follows from Assumption 3.5. Substituting the resulting estimate into (27) yields

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \sum_{t=1}^{T-1} \frac{L_{[t]} \sigma_{t+1}}{\sqrt{n_{t+1}}}.$$

Suppose now that Assumption 3.3 holds, too. Hence, the t -th term in (27) admits the tighter estimate

$$\begin{aligned} & \left| F_1^t(\widehat{F}_{t+1}^T(x)) - F_1^{t+1}(\widehat{F}_{t+2}^T(x)) \right| \\ &= \left| F_1^{t-1}(\mathbb{E}_{t-1}[f_t(\xi_t, \widehat{F}_{t+1}^T(x))]) - F_1^{t-1}(\mathbb{E}_{t-1}[f_t(\xi_t, \mathbb{E}_t[\widehat{F}_{t+1}^T(x)])]) \right| \quad (28) \\ &\leq L_{[t-1]} \mathbb{E} \left[\left\| \mathbb{E}_{t-1}[f_t(\xi_t, \widehat{F}_{t+1}^T(x)) - f_t(\xi_t, \mathbb{E}_t[\widehat{F}_{t+1}^T(x)])] \right\|_2 \right] \end{aligned}$$

where the inequality follows from the Lipschitz continuity of F_1^{t-1} established in Lemma C.3.

$$\mathbb{E}_{t-1} \left[\nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{F}_{t+1}^T(x)])^\top (\widehat{F}_{t+1}^T(x) - \mathbb{E}_t[\widehat{F}_{t+1}^T(x)]) \right] = 0$$

as $\nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{F}_{t+1}^T(x)])$ is \mathcal{F}_t -measurable. The conditional expectation in the last line of (28) thus satisfies

$$\begin{aligned} & \left\| \mathbb{E}_{t-1} \left[f_t(\xi_t, \widehat{F}_{t+1}^T(x)) - f_t(\xi_t, \mathbb{E}_t[\widehat{F}_{t+1}^T(x)]) \right] \right\|_2 \\ &= \left\| \mathbb{E}_{t-1} \left[f_t(\xi_t, \widehat{F}_{t+1}^T(x)) - f_t(\xi_t, \mathbb{E}_t[\widehat{F}_{t+1}^T(x)]) + \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{F}_{t+1}^T(x)])^\top (\widehat{F}_{t+1}^T(x) - \mathbb{E}_t[\widehat{F}_{t+1}^T(x)]) \right] \right\|_2 \\ &\leq \frac{S_t}{2} \mathbb{E}_{t-1} \left[\left\| \widehat{F}_{t+1}^T(x) - \mathbb{E}_t[\widehat{F}_{t+1}^T(x)] \right\|_2^2 \right] \\ &= \frac{S_t}{2} \mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[f_{t+1}(\xi_{t+1}, \widehat{F}_{t+2}^T(x))] - \mathbb{E}_t[f_{t+1}(\xi_{t+1}, \widehat{F}_{t+2}^T(x))] \right\|_2^2 \right] \leq \frac{S_t \sigma_{t+1}^2}{2n_{t+1}}. \end{aligned}$$

Here, the first inequality follows from the bound in (13), which applies thanks to Assumption 3.3. The second inequality holds again because, conditional on $\xi_{[t]}$, the variance of an average of n_{t+1} i.i.d. random variables coincides with the variance of a single random variable divided by n_{t+1} . Substituting the above estimate into (28) and combining the resulting bound with (27) finally yields

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \sum_{t=1}^{T-1} \frac{L_{[t-1]} S_t \sigma_{t+1}^2}{2n_{t+1}}.$$

This observation completes the proof. \square

Proof of Lemma 3.7. Since $F(x)$ is deterministic, the variance of the SAA estimator $\widehat{F}(x)$ satisfies

$$\mathbb{V}(\widehat{F}(x)) = \mathbb{V}(\widehat{F}(x) - F(x)) \leq \mathbb{E}\left[(\widehat{F}(x) - F(x))^2\right].$$

To bound the variance of $\widehat{F}(x)$, it thus suffices to bound the second moment of the estimation error $\widehat{F}(x) - F(x)$. By the construction of the conditional risk mappings in Definitions C.1 and C.4, we have

$$F(x) = F_1^T(x) = \widehat{F}_1^0(F_1^T(x)) \quad \text{and} \quad \widehat{F}(x) = \widehat{F}_1^T(x) = \widehat{F}_1^T(F_{T+1}^T(x))$$

for all $x \in \mathbb{R}^d$. We can thus represent the estimation error as the telescoping sum

$$\widehat{F}(x) - F(x) = \sum_{t=0}^{T-1} \left(\widehat{F}_1^t(F_{t+1}^T(x)) - \widehat{F}_1^{t+1}(F_{t+2}^T(x)) \right).$$

Therefore, the variance of $\widehat{F}(x)$ satisfies

$$\begin{aligned} \mathbb{V}(\widehat{F}(x))^{\frac{1}{2}} &\leq \mathbb{E}\left[\sum_{t=0}^{T-1} \left(\widehat{F}_1^t(F_{t+1}^T(x)) - \widehat{F}_1^{t+1}(F_{t+2}^T(x)) \right)^2\right]^{\frac{1}{2}} \\ &\leq \sum_{t=0}^{T-1} \mathbb{E}\left[\left(\widehat{F}_1^t(F_{t+1}^T(x)) - \widehat{F}_1^{t+1}(F_{t+2}^T(x)) \right)^2\right]^{\frac{1}{2}} \\ &= \sum_{t=0}^{T-1} \mathbb{E}\left[\left(\widehat{F}_1^t\left(\mathbb{E}_t[f_{t+1}(\xi_{t+1}, F_{t+2}^T(x))]\right) - \widehat{F}_1^{t+1}\left(\mathbb{E}_t[f_{t+1}(\xi_{t+1}, F_{t+2}^T(x))]\right) \right)^2\right]^{\frac{1}{2}}, \end{aligned}$$

where the second inequality follows from the Minkowski inequality. From now on, we use the notational shorthand $A_t := f_t(\xi_t, F_{t+1}^T(x))$ for all $t \in [T]$. As \widehat{F}_1^t is $L_{[t]}$ -Lipschitz continuous, we thus find

$$\begin{aligned} \mathbb{V}(\widehat{F}(x))^{\frac{1}{2}} &\leq \sum_{t=0}^{T-1} L_{[t]} \mathbb{E}\left[\left(\widehat{\mathbb{E}}_t\left[\left\|\widehat{\mathbb{E}}_t[A_{t+1}] - \mathbb{E}_t[A_{t+1}]\right\|_2\right]\right)^2\right]^{\frac{1}{2}} \\ &\leq \sum_{t=0}^{T-1} L_{[t]} \mathbb{E}\left[\left\|\widehat{\mathbb{E}}_t[A_{t+1}] - \mathbb{E}_t[A_{t+1}]\right\|_2^2\right]^{\frac{1}{2}} \\ &\leq \sum_{t=0}^{T-1} \frac{L_{[t]}}{\sqrt{n_{t+1}}} \mathbb{E}\left[\left\|A_{t+1} - \mathbb{E}_t[A_{t+1}]\right\|_2^2\right]^{\frac{1}{2}} \leq \sum_{t=0}^{T-1} \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}}. \end{aligned}$$

Here, the second inequality exploits Jensen's inequality and the observation that $\mathbb{E}[\widehat{\mathbb{E}}[\cdot]] = \mathbb{E}[\cdot]$. The third inequality follows from the formula for the (conditional) variance of a sample average, and the last inequality follows from Assumption 3.5. Squaring both sides of the above inequality then yields the basic bound

$$\mathbb{V}(\widehat{F}(x)) \leq \left(\sum_{t=0}^{T-1} \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}} \right)^2. \quad (29)$$

This basic bound can be strengthened by recalling that $\widehat{F}(x) = \frac{1}{n_1} \sum_{i_1=1}^{n_1} \widehat{F}_{i_1}(x)$, where $\widehat{F}_{i_1}(x)$, $i_1 \in [n_1]$, constitute independent and identically distributed random variables that depend on the samples of the i_1 -th scenario tree used in the construction of the SAA estimator; see Definition 3.4. Thus, we have

$$\mathbb{V}(\widehat{F}(x)) = \frac{1}{n_1^2} \sum_{i_1=1}^{n_1} \mathbb{V}(\widehat{F}_{i_1}(x)) \leq \frac{1}{n_1} \left(\sum_{t=1}^{T-1} \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}} + \sigma_1 \right)^2$$

where the inequality uses (29) with $n_1 = 1$ to bound the variance of $\widehat{F}_{i_1}(x)$. Hence, the claim follows. \square

The next lemma shows that sub-Gaussianity is preserved under averaging and concatenation as well as under transformations with Lipschitz continuous functions and mixtures. Among other things, this lemma enables us to identify easily verifiable conditions under which Assumption 3.11 holds.

Lemma C.5 (Operations Preserving Sub-Gaussianity). *The following hold.*

- (i) *If z_i , $i \in [n]$, are i.i.d. sub-Gaussian random vectors with common variance proxy ζ^2 , then their average $\frac{1}{n} \sum_{i=1}^n z_i$ constitutes another sub-Gaussian random vector with variance proxy ζ^2/n .*
- (ii) *If z_1 and z_2 are (possibly dependent) sub-Gaussian random vectors valued in \mathbb{R}^{m_1} and \mathbb{R}^{m_2} with respective variance proxies ζ_1^2 and ζ_2^2 , then (z_1, z_2) is sub-Gaussian with variance proxy $2 \max\{\zeta_1^2, \zeta_2^2\}$.*
- (iii) *If z is an m -dimensional sub-Gaussian random vector with variance proxy ζ^2 , and if $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is an L -Lipschitz continuous function with $\|f(z) - f(z')\|_2 \leq L\|z - z'\|_2$ for all $z, z' \in \mathbb{R}^m$, then $f(z)$ is a sub-Gaussian random vector with variance proxy that grows monotonically with $L^2 m^2 \zeta^2$.*
- (iv) *If z_i , $i \in [n]$ are (possibly dependent) sub-Gaussian random vectors valued in \mathbb{R}^m with respective variance proxies ζ_i^2 , $i \in [n]$, and if α is a random index independent of all other random objects with $\mathbb{P}(\alpha = i) = w_i$ for all $i \in [n]$, then the mixture z_α is a sub-Gaussian random vector with variance proxy $\max_{i \in [n]} \{\zeta_i^2\} + \max_{i \in [n]} \{\|\mathbb{E}[z_i] - \mathbb{E}[z_\alpha]\|_2\}$.*

Proof. Assertion (i) follows from an immediate generalization of [64, Exercise 2.13].

As for assertion (ii), observe that

$$\begin{aligned} \mathbb{E}[\exp((\lambda_1, \lambda_2)^\top((z_1, z_2) - \mathbb{E}[(z_1, z_2)]))] &= \mathbb{E}[\exp(\lambda_1^\top(z_1 - \mathbb{E}[z_1])) \exp(\lambda_2^\top(z_2 - \mathbb{E}[z_2]))] \\ &\leq \mathbb{E}[\exp(2\lambda_1^\top(z_1 - \mathbb{E}[z_1]))]^{1/2} \mathbb{E}[\exp(2\lambda_2^\top(z_2 - \mathbb{E}[z_2]))]^{1/2} \\ &\leq \exp(\|\lambda_1\|_2^2 \zeta_1^2) \exp(\|\lambda_2\|_2^2 \zeta_2^2) \leq \exp(\|(\lambda_1, \lambda_2)\|_2^2 \max\{\zeta_1^2, \zeta_2^2\}) \end{aligned}$$

for all $\lambda_1 \in \mathbb{R}^{m_1}$ and $\lambda_2 \in \mathbb{R}^{m_2}$. Here, the three inequalities follow from Hölder's inequality, the sub-Gaussianity of z_1 and z_2 and the monotonicity of the exponential function, respectively. This shows that the combined random vector (z_1, z_2) is indeed sub-Gaussian with variance proxy $2 \max\{\zeta_1^2, \zeta_2^2\}$.

As for assertion (iii), let z' be an independent copy of z . Thus, $-z'$ has the same variance proxy as z , which implies via [64, Exercise 2.13] that $z - z'$ is sub-Gaussian with variance proxy $2\zeta^2$. Hence, we find

$$\begin{aligned} \mathbb{E}[\exp(\lambda^\top(f(z) - \mathbb{E}[f(z)]))] &= \mathbb{E}[\exp(\lambda^\top(\mathbb{E}[f(z) - f(z')|z]))] \\ &\leq \mathbb{E}[\exp(\lambda^\top(f(z) - f(z')))] \\ &\leq \mathbb{E}[\exp(\|\lambda\|_2 \|f(z) - f(z')\|_2)] \\ &\leq \mathbb{E}[\exp(L\|\lambda\|_2 \|z - z'\|_2)], \end{aligned}$$

where we have used Jensen's inequality, the Cauchy-Schwarz inequality and the L -Lipschitz continuity of f . Since $\|\cdot\|_2 \leq \|\cdot\|_1$ and as the exponential function is monotonically increasing, we also find

$$\exp(\|z - z'\|_2) \leq \exp(\|z - z'\|_1) = \prod_{i=1}^m \exp(|z_i - z'_i|).$$

Combining the above estimates and applying the multi-variable Hölder inequality then yields

$$\begin{aligned} \mathbb{E}[\exp(\lambda^\top(f(z) - \mathbb{E}[f(z)]))] &\leq \mathbb{E}\left[\prod_{i=1}^m \exp(L\|\lambda\|_2 |z_i - z'_i|)\right] \\ &\leq \prod_{i=1}^m \mathbb{E}[\exp(Lm\|\lambda\|_2 |z_i - z'_i|)]^{1/m}. \end{aligned} \tag{30}$$

As $z^{1/m}$ increases with $z \geq 0$, an upper bound on (30) follows by bounding the underlying expectations as

$$\begin{aligned} \mathbb{E}[\exp(Lm \|\lambda\|_2 |z_i - z'_i|)] &\leq \mathbb{E}[\exp(Lm \|\lambda\|_2 (z_i - z'_i))] + \mathbb{E}[\exp(Lm \|\lambda\|_2 (z'_i - z_i))] \\ &\leq 2 \exp(\|\lambda\|_2^2 L^2 m^2 \zeta^2). \end{aligned}$$

Here, the two inequalities hold because $\exp(|a|) \leq \exp(a) + \exp(-a)$ for any $a \in \mathbb{R}$ and because $z_i - z'_i$ is sub-Gaussian with variance proxy $2\zeta^2$. Substituting this estimate into (30) shows that

$$\mathbb{E}[\exp(\lambda^\top (f(z) - \mathbb{E}[f(z)]))] \leq 2 \exp(\|\lambda\|_2^2 L^2 m^2 \zeta^2). \quad (31)$$

It remains to be shown that the factor 2 can be absorbed in the variance proxy. For ease of notation, we henceforth use $\psi(\lambda)$ as shorthand for the left hand side of (31). That is, $\psi(\lambda)$ denotes the moment-generating function of $f(z) - \mathbb{E}[f(z)]$. We must now prove that there exists $\varsigma > 0$ such that $\psi(\lambda) \leq \exp(\|\lambda\|_2^2 \varsigma^2 / 2)$ for all $\lambda \in \mathbb{R}^n$. To this end, note that (31) ensures that $\psi(\lambda)$ is analytic on a neighborhood of 0 [64, Section 2]. Thus, there exists $\varsigma_1^2 > 0$ with $\nabla^2 \psi(0) \preceq \varsigma_1^2 I_m$. In addition, we have $\psi(0) = 1$ and $\nabla \psi(0) = 0$. Next, define $\phi_1(\lambda) := \exp(\|\lambda\|_2^2 \varsigma_1^2 / 2)$, and note that $\phi_1(0) = 1$, $\nabla \phi_1(0) = 0$ and $\nabla^2 \phi_1(0) = \varsigma_1^2 I_m$. Thus, Taylor's theorem guarantees that there exists $\epsilon > 0$ such that $\psi(\lambda) \leq \phi_1(\lambda)$ whenever $\|\lambda\|_2 \leq \epsilon$. Define now $\varsigma_2^2 := 2(\log(2)/\epsilon^2 + L^2 m^2 \zeta^2)$, and set $\phi_2(\lambda) := \exp(\|\lambda\|_2^2 \varsigma_2^2 / 2)$. By (31), we then have

$$\psi(\lambda) \leq \exp(\log(2) + \|\lambda\|_2^2 L^2 m^2 \zeta^2) = \exp\left(\|\lambda\|_2^2 (\log(2)/\|\lambda\|_2^2 + L^2 m^2 \zeta^2)\right) < \phi_2(\lambda)$$

whenever $\|\lambda\|_2 > \epsilon$. We may thus conclude that

$$\psi(\lambda) \leq \max\{\phi_1(\lambda), \phi_2(\lambda)\} = \exp(\|\lambda\|_2^2 \max\{\varsigma_1^2, \varsigma_2^2\} / 2) \quad \forall \lambda \in \mathbb{R}^m.$$

Consequently, $f(z)$ is sub-Gaussian with variance proxy $\max\{\varsigma_1^2, \varsigma_2^2\}$.

As for assertion (iv), observe that

$$\begin{aligned} \mathbb{E}[\exp(\lambda^\top (z_\alpha - \mathbb{E}[z_\alpha]))] &= \sum_{i=1}^n w_i \mathbb{E}[\exp(\lambda^\top (z_i - \mathbb{E}[z_\alpha]))] \\ &= \sum_{i=1}^n w_i \exp(\lambda^\top (\mathbb{E}[z_i] - \mathbb{E}[z_\alpha])) \mathbb{E}[\exp(\lambda^\top (z_i - \mathbb{E}[z_i]))], \end{aligned}$$

where the first equality holds because α is independent of z_i , $i \in [n]$, and the second equality centers each z_i around its own mean. The sub-Gaussianity of each z_i implies that $\mathbb{E}[\exp(\lambda^\top (z_i - \mathbb{E}[z_i]))] \leq \exp(\|\lambda\|_2^2 \zeta_i^2 / 2)$. Factoring out a uniform bound involving $\zeta_{\max}^2 := \max_{i \in [n]} \{\zeta_i^2\}$, we obtain

$$\mathbb{E}[\exp(\lambda^\top (z_\alpha - \mathbb{E}[z_\alpha]))] \leq \exp(\|\lambda\|_2^2 \zeta_{\max}^2 / 2) \sum_{i=1}^n w_i \exp(\lambda^\top (\mathbb{E}[z_i] - \mathbb{E}[z_\alpha])).$$

Here, the sum can be interpreted as the moment generating function of a discrete random variable Y taking values $y_i = \lambda^\top (\mathbb{E}[z_i] - \mathbb{E}[z_\alpha])$ with probabilities w_i . Note that $\mathbb{E}[Y] = \sum w_i \lambda^\top (\mathbb{E}[z_i] - \mathbb{E}[z_\alpha]) = 0$. Note also that $|y_i| \leq \|\lambda\|_2 R$ for $R := \max_{i \in [n]} \{\|\mathbb{E}[z_i] - \mathbb{E}[z_\alpha]\|_2\}$, that is, the variable Y is bounded. Hoeffding's Lemma (see, e.g., [64, Example 2.4]) thus implies that $\mathbb{E}[\exp(Y)] \leq \exp(\|\lambda\|_2^2 R^2 / 2)$. In summary, we find

$$\mathbb{E}[\exp(\lambda^\top (z_\alpha - \mathbb{E}[z_\alpha]))] \leq \exp(\|\lambda\|_2^2 (\zeta_{\max}^2 + R^2) / 2),$$

which establishes that z_α is sub-Gaussian with variance proxy $\zeta_{\max}^2 + R^2$. \square

Proof of Lemma 3.13. Assumptions 3.1 and 3.2 ensure that \mathcal{X} is compact and that both F and \widehat{F} are $L_{[T]}$ -Lipschitz continuous; see Lemma C.3 as well as the discussion after Definition C.4. Next, set $v = \epsilon/(4L_{[T]})$. By Lemma A.1, there exists a v -net $\mathcal{X}_v \subseteq \mathcal{X}$ with $|\mathcal{X}_v| \leq \lceil 2D_{\mathcal{X}}/v + 1 \rceil^d$ such that

$$\mathbb{P}\left(\sup_{x \in \mathcal{X}} |\widehat{F}(x) - F(x)| > \epsilon\right) \leq \sum_{x' \in \mathcal{X}_v} \mathbb{P}\left(|\widehat{F}(x') - F(x')| > \frac{\epsilon}{2}\right). \quad (32)$$

Recall that the SAA estimator is given by $\widehat{F}(x) = \frac{1}{n_1} \sum_{i_1=1}^{n_1} \widehat{F}_{i_1}(x)$, where the $\widehat{F}_{i_1}(x)$, $i_1 \in [n_1]$, can be interpreted as independent and identically distributed estimators for $F(x)$, each of which is constructed from a single scenario tree; see Definition 3.4. In the following we introduce the auxiliary bias function

$$\mu(x) := \mathbb{E}[\widehat{F}_{i_1}(x) - F(x)] = \mathbb{E}[\widehat{F}(x) - F(x)].$$

Given any $\epsilon > 0$, we now seek sufficiently large sample sizes to ensure that $|\mu(x')| \leq \frac{\epsilon}{4}$ for every $x' \in \mathcal{X}_v$. If the integrands are Lipschitz continuous but not smooth (*i.e.*, only Assumption 3.2 holds), then we have

$$n_{t+1} = \left\lceil \left(\frac{4L_{[t]}\sigma_{t+1}(T-1)}{\epsilon} \right)^2 \right\rceil \quad \forall t \in [T-1] \quad \implies \quad \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}} \leq \frac{\epsilon}{4(T-1)} \quad \forall t \in [T-1],$$

which ensures via Lemma 3.6 that

$$|\mu(x')| \leq \sum_{t=1}^{T-1} \frac{L_{[t]}\sigma_{t+1}}{\sqrt{n_{t+1}}} \leq \frac{\epsilon}{4} \quad \forall x' \in \mathcal{X}_v.$$

Similarly, if the integrands are Lipschitz continuous and smooth (*i.e.*, both Assumptions 3.2 and 3.3 hold), then we have

$$n_{t+1} = \left\lceil \frac{2L_{[t-1]}S_t\sigma_{t+1}^2(T-1)}{\epsilon} \right\rceil \quad \forall t \in [T-1] \quad \implies \quad \frac{L_{[t-1]}S_t\sigma_{t+1}^2}{2n_{t+1}} \leq \frac{\epsilon}{4(T-1)} \quad \forall t \in [T-1],$$

which ensures via Lemma 3.6 that

$$|\mu(x')| \leq \sum_{t=1}^{T-1} \frac{L_{[t-1]}S_t\sigma_{t+1}^2}{2n_{t+1}} \leq \frac{\epsilon}{4} \quad \forall x' \in \mathcal{X}_v.$$

In both cases, we may conclude that

$$\begin{aligned} \mathbb{P}\left(\widehat{F}(x') - F(x') > \frac{\epsilon}{2}\right) &\leq \mathbb{P}\left(\widehat{F}(x') - F(x') > \frac{\epsilon}{4} + \mu(x')\right) \\ &= \mathbb{P}\left(\frac{1}{n_1} \sum_{i_1=1}^{n_1} (\widehat{F}_{i_1}(x') - F(x') - \mu(x')) > \frac{\epsilon}{4}\right) \leq \exp\left(-\frac{n_1\epsilon^2}{32\zeta^2}\right) \end{aligned}$$

for all $x' \in \mathcal{X}_v$. The second inequality follows from a standard Hoeffding bound [64, Proposition 2.5], which applies because the i.i.d. centered random variables $\widehat{F}_{i_1}(x) - F(x) - \mu(x)$, $i \in [n_1]$, are sub-Gaussian with variance proxy ζ^2 thanks to Assumption 3.11. Following a similar reasoning, one can also show that

$$\mathbb{P}\left(F(x') - \widehat{F}(x') > \frac{\epsilon}{2}\right) \leq \exp\left(-\frac{n_1\epsilon^2}{32\zeta^2}\right)$$

for all $x' \in \mathcal{X}_v$. Combining the above probability bounds with (32) finally yields

$$\mathbb{P}\left(\sup_{x \in \mathcal{X}} |\widehat{F}(x) - F(x)| > \epsilon\right) \leq 2|\mathcal{X}_v| \exp\left(-\frac{n_1\epsilon^2}{32\zeta^2}\right) \leq 2 \lceil 8L_{[T]}D_{\mathcal{X}}/\epsilon + 1 \rceil^d \exp\left(-\frac{n_1\epsilon^2}{32\zeta^2}\right),$$

and thus the claim follows. \square

Proof of Corollary 3.15. Problems (MCCO) and (1) are solvable by virtue of Weierstrass' extreme value theorem, which applies thanks to Assumptions 3.1 and 3.2 and Lemma C.3. Hence, x^* and \hat{x}^* exist. Elementary manipulations then show that

$$\begin{aligned} \mathbb{P}(F(\hat{x}^*) - F(x^*) > \epsilon) &= \mathbb{P}\left(F(\hat{x}^*) - \widehat{F}(\hat{x}^*) + \widehat{F}(\hat{x}^*) - \widehat{F}(x^*) + \widehat{F}(x^*) - F(x^*) > \epsilon\right) \\ &\leq \mathbb{P}\left(F(\hat{x}^*) - \widehat{F}(\hat{x}^*) > \epsilon/2 \text{ or } \widehat{F}(x^*) - F(x^*) > \epsilon/2\right) \\ &\leq \mathbb{P}\left(F(\hat{x}^*) - \widehat{F}(\hat{x}^*) > \epsilon/2\right) + \mathbb{P}\left(\widehat{F}(x^*) - F(x^*) > \epsilon/2\right) \\ &\leq 4 \lceil 8L_{[T]}D_{\mathcal{X}}/\epsilon + 1 \rceil^d \exp\left(-\frac{n_1\epsilon^2}{128\zeta^2}\right), \end{aligned}$$

where the first inequality holds because $\widehat{F}(\hat{x}^*) - \widehat{F}(x^*) \leq 0$. The second and the third inequalities follow from the union bound and from Lemma 3.13, respectively. \square

Proof of Theorem 3.16. The condition on n_1 ensures via Corollary 3.15 that $\mathbb{P}(F(\hat{x}^*) - F(x^*) > \epsilon) \leq \beta$ for any minimizers x^* and \hat{x}^* of problems (MCCO) and (1), respectively. Hence, any minimizer of problem (1) is an ϵ -optimal solution to (MCCO) with probability at least $1 - \beta$. The scenario complexity of the SAA estimator is thus obtained by multiplying the given expressions for the sample sizes $n_t, t \in [T]$. \square

Next, we provide the proofs of all formal results in Section 3.2.

Proof of Lemma 3.19. Since x is fixed, we simplify notation by suppressing the dependence on x for all functions of x throughout the proof. Suppose first that only Assumptions 3.1, 3.2 and 3.18 hold. We prove the upper bound on μ_t^2 by backward induction on t . As for the base case corresponding to $t = T$, note that

$$\mu_T^2 = \mathbb{E}_{T-1}[\|\widehat{H}_T\|_2^2] = \mathbb{E}_{T-1}[\|f_T(\xi_T, x)\|_2^2] \leq \bar{\mu}_T^2,$$

where the inequality follows from the definition of $\bar{\mu}_T^2$. As for the induction step corresponding to any t with $t + 1 \in [T]$, we use the law of total expectation and the definition of $\widehat{H}_t(x)$ in (18) to obtain

$$\begin{aligned} \mu_t^2 &= \mathbb{E}_{t-1}[\|\widehat{H}_t\|_2^2] = \sum_{\ell=0}^{M_t} q_t(\ell) \mathbb{E}_{t-1} \left[\frac{1}{q_t(\lambda_t)^2} \|\widehat{h}_t^{\lambda_t} - \frac{1}{2}\widehat{h}_t^{\lambda_t, e} - \frac{1}{2}\widehat{h}_t^{\lambda_t, o}\|_2^2 \middle| \lambda_t = \ell \right] \\ &= \sum_{\ell=0}^{M_t} \frac{1}{q_t(\ell)} \mathbb{E}_{t-1} \left[\|\widehat{h}_t^\ell - \frac{1}{2}\widehat{h}_t^{\ell, e} - \frac{1}{2}\widehat{h}_t^{\ell, o}\|_2^2 \right]. \end{aligned} \tag{33}$$

In the second line we have eliminated the condition on the event $\lambda_t = \ell$, which is allowed because λ_t is independent of all other random objects. Next, define $\Delta^\ell := \widehat{h}_t^\ell - \frac{1}{2}\widehat{h}_t^{\ell, e} - \frac{1}{2}\widehat{h}_t^{\ell, o}$ to avoid clutter, and note that

$$\begin{aligned} \|\Delta^\ell\|_2 &\leq \frac{L_t}{2} (\|\widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}] - \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]\|_2 + \|\widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}] - \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]\|_2) \\ &= \frac{L_t}{2} \|\widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}] - \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}]\|_2 \\ &\leq \frac{L_t}{2} (\|\widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2 + \|\widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2), \end{aligned}$$

where the first inequality holds because of Assumption 3.2, which ensures that $f_t(\xi_t, x_t)$ is L_t -Lipschitz continuous in x_t , and because of the triangle inequality. The equality follows from the antithetic sample

average identity (17), and the second inequality follows again from the triangle inequality. As $(a + b)^2 \leq 2a^2 + 2b^2$ for all $a, b \in \mathbb{R}$, the conditional second moment of Δ^ℓ admits the following upper bound.

$$\begin{aligned} \mathbb{E}_{t-1}[\|\Delta^\ell\|_2^2] &\leq \frac{L_t^2}{2} \left(\mathbb{E}_{t-1}[\|\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^2] + \mathbb{E}_{t-1}[\|\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^2] \right) \\ &= L_t^2 \mathbb{E}_{t-1}[\|\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^2] \leq \frac{B_2 L_t^2}{2^{\ell-1}} \mathbb{E}_{t-1}[\|\widehat{H}_{t+1}\|_2^2] = \frac{B_2 L_t^2}{2^{\ell-1}} \mathbb{E}_{t-1}[\mu_{t+1}^2]. \end{aligned}$$

Here, the first equality holds because the even and the odd sample averages of \widehat{H}_t have the same distribution conditional on \mathcal{F}_t , and the second inequality follows from Lemma A.4. The second equality follows from the definition of μ_{t+1}^2 and the law of total expectation. Substituting this bound back into (33) yields

$$\begin{aligned} \mu_t^2 &\leq 2B_2 L_t^2 \mathbb{E}_{t-1}[\mu_{t+1}^2] \sum_{\ell=0}^{M_t} \frac{1}{q_t(\ell)2^\ell} \\ &\leq 2B_2 L_t^2 \bar{\mu}_T^2 \prod_{s=t+1}^{T-1} \left(2B_2 L_s^2 \sum_{\ell=0}^{M_s} \frac{1}{2^s q_s(\ell)} \right) \sum_{\ell=0}^{M_t} \frac{1}{q_t(\ell)2^\ell} = \bar{\mu}_T^2 \prod_{s=t}^{T-1} \left(2B_2 L_s^2 \sum_{\ell=0}^{M_s} \frac{1}{2^s q_s(\ell)} \right), \end{aligned}$$

where the second inequality follows from the induction hypothesis. This establishes the first claim.

Suppose now that Assumption 3.3 holds, too. In this case, we prove the upper bound on $\mu_t^{2^t}$ by backward induction on t . The base case corresponding to $t = T$ is trivially true because

$$\mu_T^{2^T} = \mathbb{E}_{T-1}[\|\widehat{H}_T\|_{2^T}^2] = \mathbb{E}_{T-1}[\|f_T(\xi_T, x)\|_{2^T}^2] \leq \bar{\mu}_T^{2^T},$$

where the inequality uses the definition of $\bar{\mu}_T^{2^T}$. As for the induction step corresponding to any t with $t+1 \in [T]$, we can use the law of total expectation and define Δ^ℓ as in the first part of the proof to demonstrate that

$$\mu_t^{2^t} = \mathbb{E}_{t-1}[\|\widehat{H}_t\|_{2^t}^2] = \sum_{\ell=0}^{M_t} \frac{\mathbb{E}_{t-1}[\|\Delta^\ell\|_{2^t}^2]}{q_t(\ell)2^{t-1}}. \quad (34)$$

Next, we can use the antithetic sample average identity (17) to reformulate Δ^ℓ as

$$\begin{aligned} \Delta^\ell &= \widehat{h}_t^\ell - f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}])^\top (\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \\ &\quad - \frac{1}{2} \left[\widehat{h}_t^{\ell,e} - f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}])^\top (\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \right] \\ &\quad - \frac{1}{2} \left[\widehat{h}_t^{\ell,o} - f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}])^\top (\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \right]. \end{aligned}$$

Taking norms on both sides and replacing \widehat{h}_t^ℓ , $\widehat{h}_t^{\ell,e}$ and $\widehat{h}_t^{\ell,o}$ by their definitions then yields

$$\begin{aligned} \|\Delta^\ell\|_{2^t} \leq \|\Delta^\ell\|_2 &\leq \frac{S_t}{2} \left\| \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^2 \\ &\quad + \frac{S_t}{4} \left\| \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^2 + \frac{S_t}{4} \left\| \widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^2. \end{aligned}$$

Here, the two inequalities follow from Lemma A.2 and from a combination of the triangle inequality and

the inequality (13) for smooth functions, respectively. By the Minkowski inequality, we then obtain

$$\begin{aligned}
& \mathbb{E}_{t-1} [\|\Delta^\ell\|_2^{2t}]^{\frac{1}{2t}} \\
& \leq \frac{S_t}{2} \left(\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{2^{t+1}} \right] \right)^{\frac{1}{2t}} + \frac{S_t}{4} \left(\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{2^{t+1}} \right] \right)^{\frac{1}{2t}} \\
& \quad + \frac{S_t}{4} \left(\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{2^{t+1}} \right] \right)^{\frac{1}{2t}} \\
& = \frac{S_t}{2} \left(\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{2^{t+1}} \right] \right)^{\frac{1}{2t}} + \frac{S_t}{2} \left(\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{2^{t+1}} \right] \right)^{\frac{1}{2t}},
\end{aligned}$$

where the equality holds because the even and the odd sample averages of \widehat{H}_t have the same distribution conditional on \mathcal{F}_t . Applying Lemma A.2 with $p = 2$, $q = 2^{t+1}$, $m = 1$ and $n = d_t$ to the two differences of conditional expectations in the above expression then yields

$$\begin{aligned}
& \mathbb{E}_{t-1} [\|\Delta^\ell\|_2^{2t}]^{\frac{1}{2t}} \\
& \leq \frac{S_t}{2} \left((d_t)^{2t-1} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_{2^{t+1}}^{2^{t+1}} \right] \right)^{\frac{1}{2t}} \\
& \quad + \frac{S_t}{2} \left((d_t)^{2t-1} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_{2^{t+1}}^{2^{t+1}} \right] \right)^{\frac{1}{2t}} \\
& \leq \frac{S_t}{2} \left((d_t)^{2t-1} \frac{B_{2^{t+1}}}{2^{2t\ell}} \mathbb{E}_{t-1} \left[\left\| \widehat{H}_{t+1} \right\|_{2^{t+1}}^{2^{t+1}} \right] \right)^{\frac{1}{2t}} + \frac{S_t}{2} \left((d_t)^{2t-1} \frac{B_{2^{t+1}}}{2^{2t(\ell-1)}} \mathbb{E}_{t-1} \left[\left\| \widehat{H}_{t+1} \right\|_{2^{t+1}}^{2^{t+1}} \right] \right)^{\frac{1}{2t}} \\
& = \frac{3S_t}{2} \left((d_t)^{2t-1} \frac{B_{2^{t+1}}}{2^{2t\ell}} \mathbb{E}_{t-1} \left[\left\| \widehat{H}_{t+1} \right\|_{2^{t+1}}^{2^{t+1}} \right] \right)^{\frac{1}{2t}} = \frac{3S_t}{2} \left((d_t)^{2t-1} \frac{B_{2^{t+1}}}{2^{2t\ell}} \mathbb{E}_{t-1} \left[\mu_{t+1}^{2^{t+1}} \right] \right)^{\frac{1}{2t}}.
\end{aligned}$$

Here, the second inequality follows from Lemma A.4 with $p = 2^{t+1}$ and $n = 2^\ell$ (for sample averages involving even *as well as* odd samples) or $n = 2^{\ell-1}$ (for sample averages involving only even *or* odd samples). Next, by raising both sides of the resulting inequality to the 2^t -th power, we find

$$\begin{aligned}
\mathbb{E}_{t-1} [\|\Delta^\ell\|_2^{2t}] & \leq \left(\frac{3S_t}{2} \right)^{2^t} (d_t)^{2t-1} B_{2^{t+1}} \frac{1}{2^{2t\ell}} \mathbb{E}_{t-1} \left[\mu_{t+1}^{2^{t+1}} \right] \\
& \leq \left(\frac{3S_t}{2} \right)^{2^t} (d_t)^{2t-1} B_{2^{t+1}} \frac{1}{2^{2t\ell}} \bar{\mu}_T^{2^T} \prod_{s=t+1}^{T-1} \left(\left(\frac{3S_s}{2} \right)^{2^s} d_s^{2^s-1} B_{2^{s+1}} \sum_{\ell'=0}^{M_s} \frac{1}{2^{2^s \ell'} q_s(\ell')^{2^s-1}} \right),
\end{aligned}$$

where the second inequality uses the induction hypothesis. Substituting this back into (34) finally yields

$$\mu_t^{2^t} \leq \bar{\mu}_T^{2^T} \prod_{s=t}^{T-1} \left(\left(\frac{3S_s}{2} \right)^{2^s} d_s^{2^s-1} B_{2^{s+1}} \sum_{\ell=0}^{M_s} \frac{1}{2^{2^s \ell} q_s(\ell)^{2^s-1}} \right).$$

This observation completes the proof. \square

The random variable $\widehat{H}_t(x)$ represents an estimator for the conditional risk mapping $F_t^T(x)$ introduced in Definition C.1. Lemma C.6 below bounds the conditional bias of the estimator $\widehat{H}_t(x)$ for any $t \in [T]$. As $\mathbb{E}[\widehat{F}(x)] = \mathbb{E}[\widehat{H}_1(x)]$ for any $x \in \mathcal{X}$, Lemma 3.20 follows immediately from Lemma C.6. Its proof is thus omitted for brevity. We point out that the stronger result of Lemma C.6 will also be used in Section 4.

Lemma C.6. *If Assumptions 3.1, 3.2 and 3.18 hold and if $t \in [T]$, then we have*

$$\left\| \mathbb{E}_{t-1}[\widehat{H}_t(x)] - F_t^T(x) \right\|_2 \leq \begin{cases} \sum_{s=t}^{T-1} \frac{L_{[t:s]} \mathbb{E}_{t-1}[\mu_{s+1}^2(x)]^{\frac{1}{2}}}{2^{M_s/2}} & \text{if Assumption 3.3 does not hold,} \\ \sum_{s=t}^{T-1} \frac{L_{[t:s-1]} S_s \mathbb{E}_{t-1}[\mu_{s+1}^2(x)]}{2^{M_{s+1}}} & \text{if Assumption 3.3 holds.} \end{cases}$$

Proof. We simplify notation by suppressing the dependence on x for all functions of x throughout the proof. Suppose first that only Assumptions 3.1 and 3.2 hold. By the construction of the conditional risk mappings F_s^t in Definition C.1 and the estimators \widehat{H}_s in Definition 3.17, we have

$$F_t^T = F_t^{T-1}(\mathbb{E}_{T-1}[\widehat{H}_T]) \quad \text{and} \quad \mathbb{E}_{t-1}[\widehat{H}_t] = F_t^{t-1}(\mathbb{E}_{t-1}[\widehat{H}_t]).$$

The triangle inequality thus implies that the squared conditional bias of \widehat{H}_t admits the estimate

$$\|\mathbb{E}_{t-1}[\widehat{H}_t] - F_t^T\|_2 \leq \sum_{s=t}^{T-1} \|F_t^{s-1}(\mathbb{E}_{s-1}[\widehat{H}_s]) - F_t^s(\mathbb{E}_s[\widehat{H}_{s+1}])\|_2. \quad (35)$$

To reformulate the resulting upper bound, observe that for any $s \in [T-1]$, we have

$$\begin{aligned} \mathbb{E}_{s-1}[\widehat{H}_s] &= \mathbb{E}_{s-1} \left[\frac{1}{q_s(\lambda_s)} \left(\widehat{h}_s^{\lambda_s} - \frac{1}{2} \widehat{h}_s^{\lambda_s, e} - \frac{1}{2} \widehat{h}_s^{\lambda_s, o} \right) \right] \\ &= \sum_{\ell=0}^{M_s} \mathbb{E}_{s-1} \left[\widehat{h}_s^\ell - \frac{1}{2} \widehat{h}_s^{\ell, e} - \frac{1}{2} \widehat{h}_s^{\ell, o} \right] = \mathbb{E}_{s-1}[\widehat{h}_s^{M_s}] = \mathbb{E}_{s-1}[f_s(\xi_s, \widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}])], \end{aligned} \quad (36)$$

where the first equality follows from the definition of \widehat{H}_s in (18), and the second uses the law of total expectation along with the independence of λ_s from all other random objects. The third and fourth equalities hold because $\widehat{h}_s^{\ell, e}$ and $\widehat{h}_s^{\ell, o}$ share the distribution of $\widehat{h}_s^{\ell-1}$ conditional on \mathcal{F}_s for all $\ell \in [M_s]$ and because $\widehat{h}_s^{0, e} = \widehat{h}_s^{0, o} = 0$, respectively. Hence, the s -th term in the sum on the right hand side of (35) satisfies

$$\begin{aligned} &\|F_t^{s-1}(\mathbb{E}_{s-1}[\widehat{H}_s]) - F_t^s(\mathbb{E}_s[\widehat{H}_{s+1}])\|_2 \\ &= \|F_t^{s-1}(\mathbb{E}_{s-1}[f_s(\xi_s, \widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}])]) - F_t^{s-1}(\mathbb{E}_{s-1}[f_s(\xi_s, \mathbb{E}_s[\widehat{H}_{s+1}])])\|_2 \\ &\leq L_{[t:s]} \mathbb{E}_{t-1} [\|\widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}] - \mathbb{E}_s[\widehat{H}_{s+1}]\|_2] \\ &\leq L_{[t:s]} \mathbb{E}_{t-1} [\|\widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}] - \mathbb{E}_s[\widehat{H}_{s+1}]\|_2^2]^{\frac{1}{2}} \\ &\leq \frac{L_{[t:s]}}{2^{M_s/2}} \mathbb{E}_{t-1} [\|\widehat{H}_{s+1} - \mathbb{E}_s[\widehat{H}_{s+1}]\|_2^2]^{\frac{1}{2}} \leq \frac{L_{[t:s]}}{2^{M_s/2}} \mathbb{E}_{t-1} [\mu_{s+1}^2]^{\frac{1}{2}}, \end{aligned}$$

where the first inequality follows from Lemma C.3, which guarantees that $F_t^{s-1}(z_{s-1})$ is $L_{[t:s-1]}$ -Lipschitz continuous in z_{s-1} , and from Assumption 3.2, which ensures that $f_s(\xi_s, x_s)$ is L_s -Lipschitz continuous in x_s . The second and third inequalities follow from Jensen's inequality and the fact that the variance of the average of i.i.d. random variables equals the variance of one of these random variables divided by the sample size. The last inequality holds because the variance of any random variable is bounded by its second moment. The claim then follows by substituting the emerging bound into (35).

If Assumption 3.3 holds, too, then the s -th term on the right hand side of (35) satisfies

$$\begin{aligned} &\|F_t^{s-1}(\mathbb{E}_{s-1}[\widehat{H}_s]) - F_t^s(\mathbb{E}_s[\widehat{H}_{s+1}])\|_2 \\ &= \|F_t^{s-1}(\mathbb{E}_{s-1}[f_s(\xi_s, \widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}])]) - F_t^{s-1}(\mathbb{E}_{s-1}[f_s(\xi_s, \mathbb{E}_s[\widehat{H}_{s+1}])])\|_2 \\ &\leq L_{[t:s-1]} \mathbb{E}_{t-1} [\|\mathbb{E}_{s-1}[f_s(\xi_s, \widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}])] - f_s(\xi_s, \mathbb{E}_s[\widehat{H}_{s+1}])\|_2]. \end{aligned} \quad (37)$$

As $\nabla_{x_s} f_s(\xi_s, \mathbb{E}_s[\widehat{H}_{s+1}])$ is \mathcal{F}_s -measurable and as $\mathbb{E}_s[\widehat{\mathbb{E}}_s^{M_s}[\cdot]] = \mathbb{E}_s[\cdot]$, we may conclude that

$$\mathbb{E}_{s-1} [\nabla_{x_s} f_s(\xi_s, \mathbb{E}_s[\widehat{H}_{s+1}])^\top (\widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}] - \mathbb{E}_s[\widehat{H}_{s+1}])] = 0.$$

Hence, the last expression in (37) equals

$$\begin{aligned}
& L_{[t:s-1]} \mathbb{E}_{t-1} \left[\left\| \mathbb{E}_{s-1} [f_s(\xi_s, \widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}]) - f_s(\xi_s, \mathbb{E}_s[\widehat{H}_{s+1}])] \right. \right. \\
& \quad \left. \left. - \nabla_{x_s} f_s(\xi_s, \mathbb{E}_s[\widehat{H}_{s+1}])^\top (\widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}] - \mathbb{E}_s[\widehat{H}_{s+1}]) \right\|_2 \right] \\
& \leq \frac{L_{[t:s-1]} S_s}{2} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_s^{M_s}[\widehat{H}_{s+1}] - \mathbb{E}_s[\widehat{H}_{s+1}] \right\|_2^2 \right] \\
& \leq \frac{L_{[t:s-1]} S_s}{2^{M_s+1}} \mathbb{E}_{t-1} \left[\left\| \widehat{H}_{s+1} - \mathbb{E}_s[\widehat{H}_{s+1}] \right\|_2^2 \right] \leq \frac{L_{[t:s-1]} S_s}{2^{M_s+1}} \mathbb{E}_{t-1} [\mu_{s+1}^2].
\end{aligned}$$

Here, the first inequality follows by applying Jensen's inequality to move the conditional expectation $\mathbb{E}_{s-1}[\cdot]$ outside the norm, and then using the smoothness inequality (13), which holds by Assumption 3.3. The second inequality holds because the variance of an average of i.i.d. samples equals the variance of a single sample divided by the sample size. The third inequality follows from the fact that the variance of any random variable is bounded by its second moment, together with the definition of μ_{s+1}^2 in (19). Substituting the emerging bound into (35) yields the desired result. \square

Proof of Lemma 3.22. The MLMC estimator of Definition 3.17 is built on n_1 independent scenario trees, and thus, we have $\mathbb{E}[\mathbb{C}(\widehat{F}(x))] = n_1 \mathbb{E}[\mathbb{C}(\widehat{H}_1(x))]$. Each tree accommodates T stages, 2^{λ_t} branches per stage- t node for every $t \in [T-1]$ and one sample per node. As each log-branching factor λ_t , $t \in [T-1]$, is independent of all other random objects, the expected number of scenarios per tree is given by

$$\mathbb{E}[\mathbb{C}(\widehat{H}_1(x))] = \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell.$$

Hence, the claim follows. \square

The following lemma shows that if the integrands fail to be smooth, then the rate parameters in Assumption 3.23 are optimal in a precise sense, as detailed in Remark 3.24.

Lemma C.7. *If Assumptions 3.1, 3.2 and 3.18 hold, then the choice $r = (r_1, \dots, r_{T-1})$ with $r_t = 1/2$ for all $t \in [T-1]$ minimizes the approximate sampling cost $c_{\text{ns}}(r)$ over all $r \in (0, 1)^{T-1}$.*

Proof of Lemma C.7. We adopt a similar proof strategy as in [11, § 3], which proposes an *untruncated* MLMC estimator for a subclass of (MCCE) problems with $T = 2$. Using the definition of $c_{\text{ns}}(r)$ and the uniform upper bound on $\mu_t^2(x)$ for nonsmooth integrands established in Lemma 3.19, we obtain

$$c_{\text{ns}}(r) = \frac{2}{\epsilon^2} \left(\prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} \frac{1}{q_t(\ell) 2^\ell} \right) \left(\prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell \right) = \frac{2}{\epsilon^2} \prod_{t=1}^{T-1} \left(\sum_{\ell=0}^{M_t} \frac{1}{q_t(\ell) 2^\ell} \right) \left(\sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell \right).$$

As $q_t(\ell) = r_t(1-r_t)^\ell / (1 - (1-r_t)^{M_t+1})$ is independent of $r_{t'}$ for all $t' \neq t$, we can find the optimal r_t by minimizing the logarithm of the t -th term in the resulting product over $r_t \in (0, 1)$. Thus, we must solve

$$\min_{r_t \in (0, 1)} \log \left(\sum_{\ell=0}^{M_t} \frac{1}{2^\ell q_t(\ell)} \right) + \log \left(\sum_{\ell=0}^{M_t} 2^\ell q_t(\ell) \right),$$

where $q_t(\ell)$ is interpreted as a function of r_t . This minimization problem can be simplified by representing the rate as $r_t = 1 - 2^{-b_t}$ using an auxiliary decision variable $b_t \in (0, \infty)$. Under this re-parametrization

we have $q_t(\ell) = 2^{-b_t \ell} / Z_t$, and the normalization constant amounts to $Z_t = (1 - 2^{-b_t(M_t+1)}) / (1 - 2^{-b_t})$. Replacing $q_t(\ell)$ with this formula allows us to reformulate the above minimization problem equivalently as

$$\min_{b_t \in (0, \infty)} \log \left(\sum_{\ell=0}^{M_t} 2^{(b_t-1)\ell} \right) + \log \left(\sum_{\ell=0}^{M_t} 2^{(1-b_t)\ell} \right).$$

The objective function of the re-parametrized optimization problem is convex in b_t because logarithms of sums of exponential functions are known to be convex. An optimal b_t can thus be found from the problem's first-order optimality condition. Indeed, the gradient of the objective function vanishes if and only if

$$\frac{\sum_{\ell=0}^{M_t} \ell 2^{(b_t-1)\ell}}{\sum_{\ell'=0}^{M_t} 2^{(b_t-1)\ell'}} = \frac{\sum_{\ell=0}^{M_t} \ell 2^{(1-b_t)\ell}}{\sum_{\ell'=0}^{M_t} 2^{(1-b_t)\ell'}}.$$

Clearly, the above equation is solved by $b_t = 1$ for every $t \in [T-1]$, which implies that $c_{\text{ns}}(r)$ is minimized by setting $r_t = 1 - 2^{-b_t} = 1/2$ for every $t \in [T-1]$. Hence, the claim follows. \square

Proof of Theorem 3.25. To guarantee a root mean squared error of at most ϵ , we require both the bias and the standard deviation of the MLMC estimator to be bounded by $\epsilon/\sqrt{2}$ for any $x \in \mathcal{X}$. Suppose first that Assumption 3.3 does not hold (*i.e.*, the integrands are only Lipschitz continuous). Hence, the bias satisfies

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \sum_{t=1}^{T-1} \frac{L_{[t]} \mathbb{E}[\mu_{t+1}^2(x)]^{\frac{1}{2}}}{2^{M_t/2}}$$

thanks to Lemma 3.20. Lemma 3.19 further implies that

$$\mathbb{E}[\mu_{t+1}^2(x)]^{\frac{1}{2}} \leq \sqrt{C_{t+1}} \prod_{s=t+1}^{T-1} \left(\sum_{\ell=0}^{M_s} \frac{1}{2^\ell q_s(\ell)} \right)^{\frac{1}{2}} = \sqrt{C_{t+1}} \prod_{s=t+1}^{T-1} \sqrt{Z_s} \sqrt{M_s + 1} \quad \forall t \in [T-1],$$

where the equality holds thanks to Assumption 3.23, which implies that $r_t = 1/2$. Indeed, for this choice of the rate parameter r_t we have $q_t(\ell) = 2^{-\ell} / Z_t$, and the corresponding normalization constant simplifies to $Z_t = 2(1 - 2^{-(M_t+1)})$. Combining the above estimates then yields

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \sum_{t=1}^{T-1} \frac{L_{[t]} \sqrt{C_{t+1}}}{2^{M_t/2}} \prod_{s=t+1}^{T-1} \sqrt{Z_s} \sqrt{M_s + 1}.$$

Recalling from the theorem statement that the t -th truncation point is given by

$$M_t = \left\lceil 2 \log_2 \left(\sqrt{2} L_{[t]} \sqrt{C_{t+1}} \left(\prod_{s=t+1}^{T-1} \sqrt{Z_s} \sqrt{M_s + 1} \right) (T-1) / \epsilon \right) \right\rceil$$

for every $t \in [T-1]$, we finally obtain the desired estimate $|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \epsilon/\sqrt{2}$ for every $x \in \mathcal{X}$. Similarly, the variance of the MLMC estimator satisfies

$$\mathbb{V}(\widehat{F}(x)) = \mathbb{V}(\widehat{\mathbb{E}}[\widehat{H}_1(x)]) = \frac{1}{n_1} \mathbb{V}(\widehat{H}_1(x)) \leq \frac{\mu_1^2(x)}{n_1} \leq \frac{\epsilon^2}{2}$$

for every $x \in \mathcal{X}$, where the two equalities exploit the definitions of $\widehat{F}(x)$ and $\widehat{\mathbb{E}}[\cdot]$, respectively. The first inequality holds because the variance of any random variable is bounded by its second moment, and the second inequality follows from our choice of $n_1 = \lceil \sup_{x \in \mathcal{X}} 2\mu_1^2/\epsilon^2 \rceil$. In summary, the above reasoning implies that the root mean squared error of $\widehat{F}(x)$ is bounded by ϵ uniformly across all $x \in \mathcal{X}$.

Next, Lemma 3.22 implies that the expected sampling cost of the MLMC estimator amounts to

$$\mathbb{E}[\mathbb{C}(\widehat{F}(x))] = n_1 \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell = n_1 \prod_{t=1}^{T-1} Z_t^{-1} (M_t + 1),$$

where the second equality follows from Assumption 3.23. It remains to analyze how the resulting expression scales with respect to $1/\epsilon$. The recursive definition of the truncation points readily implies that $M_t = \mathcal{O}(\log(\epsilon^{-1}))$ for all $t \in [T-1]$. By Lemma 3.19 and the definition of n_1 , we further have

$$n_1 \leq \frac{2C_1}{\epsilon^2} \prod_{s=1}^{T-1} \left(\sum_{\ell=0}^{M_s} \frac{1}{2^{\ell q_s(\ell)}} \right) + 1 = \frac{2C_1}{\epsilon^2} \prod_{t=1}^{T-1} Z_t (M_t + 1) + 1 = \mathcal{O}(\log(\epsilon^{-1})^{T-1} \epsilon^{-2}).$$

Hence, the expected sampling cost of the MLMC estimator satisfies $\mathbb{E}[\mathbb{C}(\widehat{F}(x))] \leq \mathcal{O}(\log(\epsilon^{-1})^{2(T-1)} \epsilon^{-2})$.

Suppose now that Assumption 3.3 holds, too. In this case, Lemma 3.20 and Jensen's inequality imply that the bias of $\widehat{F}(x)$ satisfies

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \sum_{t=1}^{T-1} \frac{L_{[t-1]} S_t \mathbb{E}[\mu_{t+1}^2(x)]}{2^{M_t+1}} \leq \sum_{t=1}^{T-1} \frac{L_{[t-1]} S_t \mathbb{E}[\mu_{t+1}^{2^{t+1}}(x)]^{\frac{1}{2^t}}}{2^{M_t+1}}.$$

Lemma 3.19 and Assumption 3.23 further imply that

$$\mathbb{E}[\mu_{t+1}^{2^{t+1}}(x)]^{\frac{1}{2^t}} \leq \left(D_{t+1} \prod_{s=t+1}^{T-1} \sum_{\ell=0}^{M_s} Z_s^{2^s-1} 2^{-\frac{\ell}{2^s}} \right)^{\frac{1}{2^t}} = D_{t+1}^{2^{-t}} \prod_{s=t+1}^{T-1} Z_s^{\frac{2^s-1}{2^t}} \left(\frac{1 - 2^{-\frac{M_s+1}{2^s}}}{1 - 2^{-\frac{1}{2^s}}} \right)^{\frac{1}{2^t}} \quad \forall t \in [T-1].$$

Indeed, setting the t -th rate parameter to $r_t = 1 - 2^{-1-2^{-t}}$ implies that $q_t(\ell) = 2^{-(1+2^{-t})\ell}/Z_t$ with normalization constant $Z_t = (1 - 2^{-(1+2^{-t})(M_t+1)})/(1 - 2^{-1-2^{-t}})$. Combining these estimates then yields

$$|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \sum_{t=1}^{T-1} \frac{L_{[t-1]} S_t D_{t+1}^{2^{-t}}}{2^{M_t+1}} \prod_{s=t+1}^{T-1} Z_s^{\frac{2^s-1}{2^t}} \left(\frac{1 - 2^{-\frac{M_s+1}{2^s}}}{1 - 2^{-\frac{1}{2^s}}} \right)^{\frac{1}{2^t}}.$$

Recall now from the theorem statement that the t -th truncation point is given by

$$M_t = \left\lceil \log_2 \left(\sqrt{2} L_{[t-1]} S_t D_{t+1}^{2^{-t}} \left(\prod_{s=t+1}^{T-1} Z_s^{\frac{2^s-1}{2^t}} \left(\frac{1 - 2^{-\frac{M_s+1}{2^s}}}{1 - 2^{-\frac{1}{2^s}}} \right)^{\frac{1}{2^t}} \right) (T-1)/(2\epsilon) \right) \right\rceil$$

for every $t \in [T-1]$. This choice then yields the desired bias estimate $|\mathbb{E}[\widehat{F}(x)] - F(x)| \leq \epsilon/\sqrt{2}$ for every $x \in \mathcal{X}$. To bound the variance of the MLMC estimator, we may use the exact same reasoning as in the first part of the proof. Thus, we may conclude that $\mathbb{V}(\widehat{F}(x)) \leq \epsilon^2/2$ uniformly across all $x \in \mathcal{X}$.

In addition, we may use a similar reasoning as in the first part of the proof to conclude that

$$\mathbb{E}[\mathbb{C}(\widehat{F}(x))] = n_1 \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} Z_t^{-1} 2^{-2^{-t}\ell} = n_1 \prod_{t=1}^{T-1} Z_t^{-1} \frac{1 - 2^{-\frac{M_t+1}{2^t}}}{1 - 2^{-\frac{1}{2^t}}}.$$

The recursive definition of the truncation points readily implies that $M_t = \mathcal{O}(\log(\epsilon^{-1}))$ for all $t \in [T-1]$. Furthermore, the definition of n_1 implies via Lemma 3.19 that

$$n_1 \leq \frac{2D_1}{\epsilon^2} \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} Z_t^{2^t-1} 2^{-2^{-t}\ell} + 1 = \frac{2D_1}{\epsilon^2} \prod_{t=1}^{T-1} Z_t^{2^t-1} \frac{1 - 2^{-\frac{M_t+1}{2^t}}}{1 - 2^{-\frac{1}{2^t}}} + 1 = \mathcal{O}(\epsilon^{-2}).$$

Hence, the expected sampling cost of the MLMC estimator satisfies $\mathbb{E}[\mathbb{C}(\widehat{F}(x))] \leq \mathcal{O}(\epsilon^{-2})$. \square

Proof of Theorem 3.31. The condition on n_1 ensures via Corollary 3.30 that $\mathbb{P}(F(\widehat{x}^*) - F(x^*) > \epsilon) \leq \beta$ for any minimizers x^* and \widehat{x}^* of problems (MCCO) and (1), respectively. Hence, any minimizer of problem (1) constitutes an ϵ -optimal solution to (MCCO) with probability at least $1 - \beta$.

By Lemma 3.22, the expected sampling cost of the MLMC estimator amounts to

$$\mathbb{E}[\mathbb{C}(\widehat{F}(x))] = n_1 \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell.$$

By using a similar reasoning as in the proof of Theorem 3.25, one can show that the product term is of order $\mathcal{O}(\log(\epsilon^{-1})^{T-1})$ for Lipschitz continuous integrands and of order $\mathcal{O}(1)$ for smooth integrands. Moreover, the number n_1 of scenario trees appearing in the statement of the corollary is of order $\mathcal{O}(\log(\epsilon^{-1})\epsilon^{-2})$. The desired scenario complexity bounds then follow by multiplying these estimates. \square

D Proofs of Section 4

Proof of Lemma 4.1. We use backward induction on t to prove that $\nabla F_t^T(x) = G_t(x)$ for all $x \in \mathbb{R}^d$ and $t \in [T]$. As for the base case corresponding to $t = T$, fix any $x \in \mathbb{R}^d$, and observe that

$$\nabla F_T^T(x) = \nabla \mathbb{E}_{T-1}[f_T(\xi_T, x)] = \mathbb{E}_{T-1}[\nabla_x f_T(\xi_T, x)] = G_T(x).$$

Here, the first equality follows from the construction of $F_T^T(x)$ in Definition C.1. Next, the gradient may be interchanged with the conditional expectation by the dominated convergence theorem, which applies thanks to Assumptions 3.2 and 3.3. The resulting expression coincides with the definition of $G_T(x)$.

As for the induction step corresponding to any $t \in [T - 1]$, assume that $\nabla F_{t+1}^T(x) = G_{t+1}(x)$ for all $x \in \mathbb{R}^d$. Next, fix an arbitrary $x \in \mathbb{R}^d$, and observe that

$$\begin{aligned} \nabla F_t^T(x) &= \nabla_{x_t} \mathbb{E}_{t-1}[f_t(\xi_t, F_{t+1}^T(x))] \\ &= \mathbb{E}_{t-1}[\nabla F_{t+1}^T(x) \nabla_{x_t} f_t(\xi_t, F_{t+1}^T(x))] \\ &= \mathbb{E}_{t-1}[G_{t+1}(x) \nabla_{x_t} f_t(\xi_t, F_{t+1}^T(x))] = G_t(x). \end{aligned}$$

Here, the first equality follows from the definition of F_t^T , while the second equality follows from the dominated convergence theorem and the chain rule. The third and the fourth equalities then exploit the induction hypothesis and the definition of G_t , respectively. The claim ultimately follows because $F(x) = F_1^T(x)$. \square

Proof of Lemma 4.5. Throughout the proof we fix x , and for simplicity we suppress the dependence on x in all functions. We then find an upper bound on $\nu_t^{2^t}$ and a recursive formula for E_t by backward induction on t . As for the base case corresponding to $t = T$, the definition of \widehat{G}_T and Assumption 4.4 imply that

$$\nu_T^{2^T} = \mathbb{E}_{T-1}[\|\widehat{G}_T\|_2^{2^T}] \leq \bar{\nu}_T^{2^T} =: E_T < \infty.$$

As for the induction step corresponding to any t with $t + 1 \in [T]$, we introduce the auxiliary variables

$$\mathcal{S}_s := \sum_{\ell=0}^{M_s} \frac{1}{2^{(2^{s-1})(\rho_s+1)\ell} q_s(\ell)^{2^s-1}} \quad \text{and} \quad \mathcal{S}'_s := \sum_{\ell=0}^{M_s} \frac{1}{2^{2^s(\rho_{s-1}+1)\ell} q_s(\ell)^{2^s(\rho_{s-1}+1)-1}} \quad \forall s \in [T-1], \quad (38)$$

which allows us to express the induction hypothesis compactly as

$$\nu_{t+1}^{2^{t+1}} = \mathbb{E}_t[\|\widehat{G}_{t+1}\|_2^{2^{t+1}}] \leq E_{t+1} \prod_{s=t+1}^{T-1} \max\{\mathcal{S}_s, \mathcal{S}'_s\}. \quad (39)$$

In order to bound $\nu_t^{2^t}$, we use the law of total expectation and the definition of \widehat{G}_t in (23) to obtain

$$\begin{aligned}\nu_t^{2^t} &= \mathbb{E}_{t-1}[\|\widehat{G}_t\|_2^{2^t}] = \sum_{\ell=0}^{M_t} q_t(\ell) \mathbb{E}_{t-1} \left[\frac{1}{q_t(\lambda_t)^{2^t}} \left\| \widehat{g}_t^{\lambda_t} - \frac{1}{2} \widehat{g}_t^{\lambda_t, e} - \frac{1}{2} \widehat{g}_t^{\lambda_t, o} \right\|_2^{2^t} \middle| \lambda_t = \ell \right] \\ &= \sum_{\ell=0}^{M_t} \frac{1}{q_t(\ell)^{2^t-1}} \mathbb{E}_{t-1} \left[\left\| \widehat{g}_t^\ell - \frac{1}{2} \widehat{g}_t^{\ell, e} - \frac{1}{2} \widehat{g}_t^{\ell, o} \right\|_2^{2^t} \right].\end{aligned}\tag{40}$$

In the second line we have eliminated the condition on the event $\lambda_t = \ell$, which is allowed because λ_t is independent of all other random objects. Next, we define $\Delta^\ell := \widehat{g}_t^\ell - \frac{1}{2} \widehat{g}_t^{\ell, e} - \frac{1}{2} \widehat{g}_t^{\ell, o}$ and

$$A_t^\ell := \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \frac{1}{2} \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}]) - \frac{1}{2} \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}])$$

to avoid clutter. By the construction of \widehat{g}_t^ℓ , $\widehat{g}_t^{\ell, e}$ and $\widehat{g}_t^{\ell, o}$ in Definition 4.2, we thus have

$$\begin{aligned}\Delta^\ell &= \widehat{\mathbb{E}}_t^\ell[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \frac{1}{2} \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}]) \\ &\quad - \frac{1}{2} \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}]) \\ &= \left(\widehat{\mathbb{E}}_t^\ell[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right) \left[\nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right] \\ &\quad - \frac{1}{2} \left(\widehat{\mathbb{E}}_t^{\ell, e}[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right) \left[\nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right] \\ &\quad - \frac{1}{2} \left(\widehat{\mathbb{E}}_t^{\ell, o}[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right) \left[\nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right] + \mathbb{E}_t[\widehat{G}_{t+1}] A_t^\ell,\end{aligned}$$

where the second equality is obtained by adding terms that sum to 0 and using the antithetic sample average identity (17). Next, we observe that Jensen's inequality yields $|\sum_{i=1}^4 a_i|^p \leq 4^{p-1} \sum_{i=1}^4 |a_i|^p$ for all $a_1, \dots, a_4 \in \mathbb{R}$ and $p \geq 1$. Hence, the $(p = 2^t)$ -th conditional moment of Δ^ℓ obeys the estimate

$$\begin{aligned}\mathbb{E}_{t-1}[\|\Delta^\ell\|_2^{2^t}] &\leq 4^{2^t-1} \mathbb{E}_{t-1} \left[\left\| \left(\widehat{\mathbb{E}}_t^\ell[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right) \left[\nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right] \right\|_2^{2^t} \right] \\ &\quad + \frac{4^{2^t-1}}{2^{2^t}} \mathbb{E}_{t-1} \left[\left\| \left(\widehat{\mathbb{E}}_t^{\ell, e}[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right) \left[\nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right] \right\|_2^{2^t} \right] \\ &\quad + \frac{4^{2^t-1}}{2^{2^t}} \mathbb{E}_{t-1} \left[\left\| \left(\widehat{\mathbb{E}}_t^{\ell, o}[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right) \left[\nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right] \right\|_2^{2^t} \right] \\ &\quad + 4^{2^t-1} \mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[\widehat{G}_{t+1}] A_t^\ell \right\|_2^{2^t} \right].\end{aligned}$$

To bound the first three terms above, we recall that $2ab \leq a^2 + b^2$ for all $a, b \in \mathbb{R}$. Hence, we obtain

$$\begin{aligned}\mathbb{E}_{t-1}[\|\Delta^\ell\|_2^{2^t}] &\leq \frac{4^{2^t-1}}{2} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_2^{2^{t+1}} + \left\| \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2^{2^{t+1}} \right] \\ &\quad + \frac{4^{2^t-1}}{2^{2^t+1}} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_2^{2^{t+1}} + \left\| \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, e}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2^{2^{t+1}} \right] \\ &\quad + \frac{4^{2^t-1}}{2^{2^t+1}} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_2^{2^{t+1}} + \left\| \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell, o}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2^{2^{t+1}} \right] \\ &\quad + 4^{2^t-1} \mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[\widehat{G}_{t+1}] A_t^\ell \right\|_2^{2^t} \right].\end{aligned}\tag{41}$$

The remainder of the proof proceeds in three steps, where we analyze separately the terms appearing in (41). In Step (i), we bound the 2^t -th conditional moment of $\mathbb{E}_t[\widehat{G}_{t+1}] A_t^\ell$, which appears on the last line of (41). In Step (ii), we bound the 2^{t+1} -th conditional moments of the three gradient-difference terms involving \widehat{H}_{t+1} . Finally, in Step (iii), we bound the 2^{t+1} -th conditional moments of the three value-difference terms involving \widehat{G}_{t+1} . To complete the induction step, we then substitute all obtained bounds into (41) and (40).

Step (i) To bound the 2^t -th conditional moment of $\mathbb{E}_t[\widehat{G}_{t+1}] A_t^\ell$, we first observe that

$$\mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[\widehat{G}_{t+1}] A_t^\ell \right\|_2^{2^t} \right] \leq \mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_2^{2^t} \left\| A_t^\ell \right\|_2^{2^t} \right], \quad (42)$$

where the inequality follows from the submultiplicativity of the Frobenius norm established in Lemma A.3. We emphasize that entry-wise matrix p -norms are not submultiplicative for $p > 2$. This is our reason for using the Frobenius norm (corresponding to $p = 2$) in the definition of the p -th conditional moment of $\widehat{G}_t(x)$ in (25). Next, we bound $\|A_t^\ell\|_2^{2^t}$. Jensen's inequality readily implies that

$$\|A_t^\ell\|_2^{2^t} = \left(\sum_{i=1}^{d_{t-1}} \|A_t^\ell e_i\|_2^2 \right)^{2^{t-1}} \leq (d_{t-1})^{2^{t-1}-1} \sum_{i=1}^{d_{t-1}} \|A_t^\ell e_i\|_2^{2^t}, \quad (43)$$

where e_i stands for the i -th standard basis vector in $\mathbb{R}^{d_{t-1}}$. By the definition of A_t^ℓ , we can rewrite $A_t^\ell e_i$ as

$$\begin{aligned} & \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \frac{1}{2} \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}]) - \frac{1}{2} \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}]) \\ = & \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \nabla_{x_t} f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t}^2 f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) (\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \\ & + \frac{1}{2} \left[\nabla_{x_t} f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}]) - \nabla_{x_t}^2 f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) (\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \right] \\ & + \frac{1}{2} \left[\nabla_{x_t} f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}]) - \nabla_{x_t}^2 f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) (\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \right], \end{aligned}$$

where the equality follows from the antithetic sample average identity (17). The norm of $A_t^\ell e_i$ thus satisfies

$$\begin{aligned} \|A_t^\ell e_i\|_2 \leq & \left\| \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \nabla_{x_t} f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right. \\ & \left. - \nabla_{x_t}^2 f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) (\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2 \\ & + \frac{1}{2} \left\| \nabla_{x_t} f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}]) \right. \\ & \left. - \nabla_{x_t}^2 f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) (\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2 \\ & + \frac{1}{2} \left\| \nabla_{x_t} f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) - \nabla_{x_t} f_{t,i}(\xi_t, \widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}]) \right. \\ & \left. - \nabla_{x_t}^2 f_{t,i}(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) (\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2 \end{aligned}$$

by virtue of the triangle inequality. Next, we bound the three norm terms by using the estimate (24), which is allowed because the Hessians are Hölder continuous thanks to Assumption 4.3. We therefore obtain

$$\begin{aligned} \|A_t^\ell e_i\|_2 \leq & \frac{R_t}{\rho_t + 1} \left\| \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{\rho_t+1} + \frac{R_t}{2(\rho_t + 1)} \left\| \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{\rho_t+1} \\ & + \frac{R_t}{2(\rho_t + 1)} \left\| \widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{\rho_t+1}. \end{aligned}$$

Two successive applications of Jensen's inequality yield $|a + b + c|^p \leq 2^{p-1}(|a|^p + 2^{p-1}(|b|^p + |c|^p))$ for all $a, b, c \in \mathbb{R}$ and $p \geq 1$. By raising both sides of the previous inequality to the power $p = 2^t$ we thus find

$$\begin{aligned} \|A_t^\ell e_i\|_2^{2^t} &\leq \left(\frac{R_t}{\rho_t + 1}\right)^{2^t} 2^{2^t-1} \left[\|\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} \right. \\ &\quad \left. + \frac{1}{2} \|\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} + \frac{1}{2} \|\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} \right] \\ &\leq \left(\frac{R_t}{\rho_t + 1}\right)^{2^t} 2^{2^t-1} (d_t)^{2^{t-1}(\rho_t+1)-1} \left[\|\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} \right. \\ &\quad \left. + \frac{1}{2} \|\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} + \frac{1}{2} \|\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} \right], \end{aligned}$$

where the second inequality follows from Lemma A.2 with $p = 2$, $q = 2^t(\rho_t + 1)$, $m = 1$ and $n = d_t$. Substituting the above bound on $\|A_t^\ell e_i\|_2^{2^t}$ back into (43) then results in the estimate

$$\|A_t^\ell\|_2^{2^t} \leq \left(\frac{R_t}{\rho_t + 1}\right)^{2^t} 2^{2^t-1} (d_{t-1})^{2^{t-1}} (d_t)^{2^{t-1}(\rho_t+1)-1} \bar{A}_t^\ell, \quad (44)$$

where we have used the shorthand

$$\begin{aligned} \bar{A}_t^\ell &= \|\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} \\ &\quad + \frac{1}{2} \|\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)} + \frac{1}{2} \|\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^t(\rho_t+1)}. \end{aligned}$$

Substituting (44) into (42) and applying the Cauchy-Schwarz and Jensen inequalities then yields

$$\begin{aligned} &\mathbb{E}_{t-1} \left[\|\mathbb{E}_t[\widehat{G}_{t+1}]\|_2^{2^t} \|A_t^\ell\|_2^{2^t} \right] \\ &\leq \left(\frac{R_t}{\rho_t + 1}\right)^{2^t} 2^{2^t-1} (d_{t-1})^{2^{t-1}} (d_t)^{2^{t-1}(\rho_t+1)-1} \mathbb{E}_{t-1} \left[\|\mathbb{E}_t[\widehat{G}_{t+1}]\|_2^{2^t} \bar{A}_t^\ell \right] \\ &\leq \left(\frac{R_t}{\rho_t + 1}\right)^{2^t} 2^{2^t-1} (d_{t-1})^{2^{t-1}} (d_t)^{2^{t-1}(\rho_t+1)-1} \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_2^{2^{t+1}} \right]^{\frac{1}{2}} \mathbb{E}_{t-1} \left[(\bar{A}_t^\ell)^2 \right]^{\frac{1}{2}}. \end{aligned} \quad (45)$$

Note that \widehat{G}_{t+1} and \widehat{H}_{t+1} are constructed from the same sets of samples, which implies that $\mathbb{E}_t[\widehat{G}_{t+1}]$ and \bar{A}_t^ℓ are dependent. To further simplify the bound in (45), we use the Minkowski inequality to obtain

$$\begin{aligned} &\mathbb{E}_{t-1} \left[(\bar{A}_t^\ell)^2 \right]^{\frac{1}{2}} \\ &\leq \mathbb{E}_{t-1} \left[\|\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} \\ &\quad + \frac{1}{2} \mathbb{E}_{t-1} \left[\|\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} + \frac{1}{2} \mathbb{E}_{t-1} \left[\|\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} \\ &\leq (d_t)^{\frac{1}{2}} \left(\mathbb{E}_{t-1} \left[\|\widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} \right. \\ &\quad \left. + \frac{1}{2} \mathbb{E}_{t-1} \left[\|\widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} + \frac{1}{2} \mathbb{E}_{t-1} \left[\|\widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}]\|_2^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} \right). \end{aligned}$$

Here, the second inequality follows from Lemma A.2 with $p = 2^t(\rho_t + 1)$, $q = 2^{t+1}(\rho_t + 1)$, $m = 1$ and $n = d_t$. By using Lemma A.4 with $p = 2^{t+1}(\rho_t + 1)$ and $n = 2^\ell$ or $n = 2^{\ell-1}$, we further obtain

$$\begin{aligned} \mathbb{E}_{t-1} \left[(\bar{A}_t^\ell)^2 \right]^{\frac{1}{2}} &\leq (d_t)^{\frac{1}{2}} \frac{B_{2^{t+1}(\rho_t+1)}^{\frac{1}{2}}}{2^{(2^{t-1})(\rho_t+1)\ell}} \mathbb{E}_{t-1} \left[\|\widehat{H}_{t+1}\|_{2^{t+1}(\rho_t+1)}^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} \\ &\quad + (d_t)^{\frac{1}{2}} \frac{B_{2^{t+1}(\rho_t+1)}^{\frac{1}{2}}}{2^{(2^{t-1})(\rho_t+1)(\ell-1)}} \mathbb{E}_{t-1} \left[\|\widehat{H}_{t+1}\|_{2^{t+1}(\rho_t+1)}^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}} \\ &= \left(1 + 2^{2^{t-1}(\rho_t+1)}\right) (d_t)^{\frac{1}{2}} \frac{B_{2^{t+1}(\rho_t+1)}^{\frac{1}{2}}}{2^{(2^{t-1})(\rho_t+1)\ell}} \mathbb{E}_{t-1} \left[\|\widehat{H}_{t+1}\|_{2^{t+1}(\rho_t+1)}^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}}. \end{aligned}$$

Substituting the resulting expression into (45) and combining it with (42) finally yields

$$\begin{aligned} \mathbb{E}_{t-1} \left[\|\mathbb{E}_t[\widehat{G}_{t+1}] A_t^\ell\|_2^{2^t} \right] &\leq \left(\frac{R_t}{\rho_t + 1} \right)^{2^t} 2^{2^{t-1}} (d_{t-1})^{2^{t-1}} (d_t)^{2^{t-1}(\rho_t+1)-1} \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_2^{2^{t+1}} \right]^{\frac{1}{2}} \\ &\quad \times \left(1 + 2^{2^{t-1}(\rho_t+1)}\right) (d_t)^{\frac{1}{2}} \frac{B_{2^{t+1}(\rho_t+1)}^{\frac{1}{2}}}{2^{(2^{t-1})(\rho_t+1)\ell}} \mathbb{E}_{t-1} \left[\|\widehat{H}_{t+1}\|_{2^{t+1}(\rho_t+1)}^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}}. \end{aligned}$$

We have thus established an upper bound on the last term in (41), which depends on conditional moments of \widehat{G}_{t+1} and \widehat{H}_{t+1} and decays geometrically with ℓ . This concludes Step (i).

Step (ii) To bound the gradient difference terms in (41) depending on \widehat{H}_{t+1} , we note that

$$\begin{aligned} \mathbb{E}_{t-1} \left[\left\| \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2^{2^{t+1}} \right] &\leq S_t^{2^{t+1}} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^{2^{t+1}} \right] \\ &\leq S_t^{2^{t+1}} (d_t)^{2^{t-1}} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_{2^{t+1}}^{2^{t+1}} \right] \leq S_t^{2^{t+1}} (d_t)^{2^{t-1}} \frac{B_{2^{t+1}}}{2^{2^t \ell}} \mathbb{E}_{t-1} \left[\|\widehat{H}_{t+1}\|_{2^{t+1}}^{2^{t+1}} \right], \end{aligned}$$

where the first inequality exploits Assumption 3.3, and the second inequality follows from Lemma A.2 with $p = 2$, $q = 2^{t+1}$, $m = 1$ and $n = d_t$. The third inequality follows from the law of total expectation and from Lemma A.4 with $p = 2^{t+1}$ and $n = 2^\ell$. A similar reasoning implies that

$$\begin{aligned} \mathbb{E}_{t-1} \left[\left\| \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell,e}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2^{2^{t+1}} \right] &\leq S_t^{2^{t+1}} (d_t)^{2^{t-1}} \frac{B_{2^{t+1}}}{2^{2^t(\ell-1)}} \mathbb{E}_{t-1} \left[\|\widehat{H}_{t+1}\|_{2^{t+1}}^{2^{t+1}} \right], \\ \mathbb{E}_{t-1} \left[\left\| \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{\ell,o}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, \mathbb{E}_t[\widehat{H}_{t+1}]) \right\|_2^{2^{t+1}} \right] &\leq S_t^{2^{t+1}} (d_t)^{2^{t-1}} \frac{B_{2^{t+1}}}{2^{2^t(\ell-1)}} \mathbb{E}_{t-1} \left[\|\widehat{H}_{t+1}\|_{2^{t+1}}^{2^{t+1}} \right]. \end{aligned}$$

As the empirical averages here are taken only over samples with even or odd indices, we set $n = 2^{\ell-1}$ when applying Lemma A.4. In summary, we have found upper bounds on the gradient difference terms in (41), which depend on conditional moments of \widehat{H}_{t+1} and decay geometrically with ℓ . This concludes Step (ii).

Step (iii) To bound the terms in (41) depending on \widehat{G}_{t+1} , we note that

$$\begin{aligned} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_2^{2^{t+1}} \right] &\leq (d_t d)^{2^{t-1}} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^\ell[\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_{2^{t+1}}^{2^{t+1}} \right] \\ &\leq (d_t d)^{2^{t-1}} \frac{B_{2^{t+1}}}{2^{2^t \ell}} \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_{2^{t+1}}^{2^{t+1}} \right] \\ &\leq (d_t d)^{2^{t-1}} \frac{B_{2^{t+1}}}{2^{2^t \ell}} \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_2^{2^{t+1}} \right], \end{aligned}$$

where the first inequality follows from Lemma A.2 with $p = 2$, $q = 2^{t+1}$, $m = d_t$ and $n = d$, while the second inequality follows from the law of total expectation and from Lemma A.4 with $p = 2^{t+1}$ and $n = 2^\ell$. The third inequality follows from Jensen's inequality (or from Lemma A.2 with $p = 2$ and $q = 2^{t+1}$). A similar reasoning involving Lemma A.4 with $p = 2^{t+1}$ and $n = 2^{\ell-1}$ implies that

$$\begin{aligned}\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell, e} [\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_2^{2^{t+1}} \right] &\leq (d_t d)^{2^t-1} \frac{B_{2^{t+1}}}{2^{2^t(\ell-1)}} \mathbb{E}_{t-1} \left[\left\| \widehat{G}_{t+1} \right\|_2^{2^{t+1}} \right], \\ \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{\ell, o} [\widehat{G}_{t+1}] - \mathbb{E}_t[\widehat{G}_{t+1}] \right\|_2^{2^{t+1}} \right] &\leq (d_t d)^{2^t-1} \frac{B_{2^{t+1}}}{2^{2^t(\ell-1)}} \mathbb{E}_{t-1} \left[\left\| \widehat{G}_{t+1} \right\|_2^{2^{t+1}} \right].\end{aligned}$$

These bounds depend on conditional moments of \widehat{G}_{t+1} and decay geometrically with ℓ , completing Step (iii).

Having analyzed all terms in (41) in Steps (i), (ii) and (iii), we are now ready to combine the resulting estimates to a bound on the conditional moments of Δ^ℓ . For notational convenience, we define the constants

$$\begin{aligned}Y_t &:= 3 \cdot 2^{2^{t+1}-3} (d_t d)^{2^t-1} B_{2^{t+1}}, \\ Y'_t &:= 3 \cdot 2^{2^{t+1}-3} S_t^{2^{t+1}} (d_t)^{2^t-1} B_{2^{t+1}}, \\ Y''_t &:= 2^{2^{t+1}+2^t-3} \left(\frac{R_t}{\rho_t+1} \right)^{2^t} (d_{t-1})^{2^t-1} (d_t)^{2^t-1} (\rho_t+1)^{-\frac{1}{2}} \left(1 + 2^{2^t-1} (\rho_t+1) \right) B_{2^{t+1}(\rho_t+1)}^{\frac{1}{2}},\end{aligned}$$

which are independent of ℓ . Substituting the bounds from Steps (i), (ii) and (iii) into (41) then yields

$$\begin{aligned}\mathbb{E}_{t-1} \left[\left\| \Delta^\ell \right\|_2^{2^t} \right] &\leq \frac{Y_t}{2^{2^t \ell}} \mathbb{E}_{t-1} \left[\left\| \widehat{G}_{t+1} \right\|_2^{2^{t+1}} \right] + \frac{Y'_t}{2^{2^t \ell}} \mathbb{E}_{t-1} \left[\left\| \widehat{H}_{t+1} \right\|_{2^{t+1}}^{2^{t+1}} \right] \\ &\quad + \frac{Y''_t}{2^{(2^{t-1})(\rho_t+1)\ell}} \mathbb{E}_{t-1} \left[\left\| \widehat{G}_{t+1} \right\|_2^{2^{t+1}} \right]^{\frac{1}{2}} \mathbb{E}_{t-1} \left[\left\| \widehat{H}_{t+1} \right\|_{2^{t+1}(\rho_t+1)}^{2^{t+1}(\rho_t+1)} \right]^{\frac{1}{2}}.\end{aligned}\tag{47}$$

Next, we bound the three conditional expectations appearing in (47). By Lemma 3.19, under Assumption 3.3, and using the fact that $2^{s-1}(\rho_s+1) \leq 2^s$, the second conditional expectation satisfies

$$\mathbb{E}_{t-1} \left[\left\| \widehat{H}_{t+1} \right\|_{2^{t+1}}^{2^{t+1}} \right] \leq D_{t+1} \prod_{s=t+1}^{T-1} \left(\sum_{\ell=0}^{M_s} \frac{1}{2^{2^s \ell} q_s(\ell)^{2^s-1}} \right) \leq D_{t+1} \prod_{s=t+1}^{T-1} \mathcal{S}_s,$$

where

$$D_{t+1} = \bar{\mu}_T^{2^T} \prod_{s=t+1}^{T-1} \left(\left(\frac{3\mathcal{S}_s}{2} \right)^{2^s} d_s^{2^s-1} B_{2^{s+1}} \right)$$

and \mathcal{S}_s is defined in (38). By adapting Lemma 3.19 in the obvious way, one can further show that

$$\mathbb{E}_t \left[\left\| \widehat{H}_{t+1} \right\|_{2^{t+1}(\rho_t+1)}^{2^{t+1}(\rho_t+1)} \right] \leq D'_{t+1} \prod_{s=t+1}^{T-1} \mathcal{S}'_s,$$

where

$$D'_{t+1} := \bar{\mu}_T^{2^T(\rho_{T-1}+1)} \prod_{s=t+1}^{T-1} \left(\left(\frac{3\mathcal{S}_s}{2} \right)^{2^s(\rho_{s-1}+1)} d_s^{2^s(\rho_{s-1}+1)-1} B_{2^{s+1}(\rho_{s-1}+1)} \right)$$

and \mathcal{S}'_s is defined in (38). Moreover, by the induction hypothesis, the conditional moment of \widehat{G}_{t+1} in (47) satisfies the bound (39). Substituting all of these conditional moment bounds into (47) then yields

$$\begin{aligned}\mathbb{E}_{t-1} \left[\left\| \Delta^\ell \right\|_2^{2^t} \right] &\leq \frac{Y_t E_{t+1}}{2^{2^t \ell}} \prod_{s=t+1}^{T-1} \max \{ \mathcal{S}_s, \mathcal{S}'_s \} + \frac{Y'_t D_{t+1}}{2^{2^t \ell}} \prod_{s=t+1}^{T-1} \mathcal{S}_s \\ &\quad + \frac{Y''_t \sqrt{E_{t+1} D'_{t+1}}}{2^{(2^{t-1})(\rho_t+1)\ell}} \prod_{s=t+1}^{T-1} \max \{ \mathcal{S}_s, \mathcal{S}'_s \}^{\frac{1}{2}} (\mathcal{S}'_s)^{\frac{1}{2}} \leq \frac{E_t}{2^{(2^{t-1})(\rho_t+1)\ell}} \prod_{s=t+1}^{T-1} \max \{ \mathcal{S}_s, \mathcal{S}'_s \},\end{aligned}$$

where the second inequality follows from the trivial observations that $2^{t-1}(\rho_t + 1) \leq 2^t$ and that both \mathcal{S}_s as well as \mathcal{S}'_s are dominated by $\max\{\mathcal{S}_s, \mathcal{S}'_s\}$. The inequality also exploits the (recursive) definition of E_t as

$$E_t := Y_t E_{t+1} + Y'_t D_{t+1} + Y''_t \sqrt{E_{t+1} D'_{t+1}}.$$

Replacing the conditional moment $\mathbb{E}_{t-1}[\|\Delta^\ell\|_2^{2^t}]$ in (40) with the above bound finally yields

$$\nu_t^{2^t} = \sum_{\ell=0}^{M_t} \frac{\mathbb{E}_{t-1}[\|\Delta^\ell\|_2^{2^t}]}{q_t(\ell)^{2^t-1}} \leq E_t \mathcal{S}_t \prod_{s=t+1}^{T-1} \max\{\mathcal{S}_s, \mathcal{S}'_s\} \leq E_t \prod_{s=t}^{T-1} \max\{\mathcal{S}_s, \mathcal{S}'_s\}.$$

This completes the induction step. By their recursive definition and thanks to Assumption 4.4, the constants $E_t, t \in [T-1]$, are finite whenever E_T is finite. Hence, the claim follows. \square

Proof of Lemma 4.6. We simplify notation by suppressing the dependence on x for all functions of x in this proof. In the following we will show that the conditional bias of the estimator $\widehat{G}_t(x)$ obeys the estimate

$$\begin{aligned} & \|\mathbb{E}_{t-1}[\widehat{G}_t] - G_t\|_2 \\ & \leq \sum_{s=t}^{T-1} L_{[t:s-1]} \mathcal{S}_s \mathbb{E}_{t-1}[\nu_{s+1}^2]^{\frac{1}{2}} \left(\frac{\mathbb{E}_{t-1}[\mu_{s+1}^2]^{\frac{1}{2}}}{2^{M_s/2}} + \sum_{k=s+1}^{T-1} \frac{L_{[s+1:k-1]} \mathcal{S}_k \mathbb{E}_{t-1}[(\mu_{k+1}^2)^2]^{\frac{1}{2}}}{2^{M_{k+1}}} \right) \end{aligned} \quad (48)$$

for all $t \in [T]$. As $\mathbb{E}[\widehat{G}] = \mathbb{E}[\widehat{G}_1]$, the claim then follows immediately. We prove (48) by backward induction on t . As for the base case corresponding to $t = T$, note that \widehat{G}_T constitutes an unbiased estimator for G_T . Thus, the left hand side of (48) evaluates to 0, and the claim holds trivially (in fact, the right hand side of (48) vanishes, too). As for the induction step corresponding to any t with $t+1 \in [T]$, assume that

$$\|\mathbb{E}_t[\widehat{G}_{t+1}] - G_{t+1}\|_2 \leq \sum_{s=t+1}^{T-1} L_{[t+1:s-1]} \mathcal{S}_s \mathbb{E}_t[\nu_{s+1}^2]^{\frac{1}{2}} \left(\frac{\mathbb{E}_t[\mu_{s+1}^2]^{\frac{1}{2}}}{2^{M_s/2}} + \sum_{k=s+1}^{T-1} \frac{L_{[s+1:k-1]} \mathcal{S}_k \mathbb{E}_t[(\mu_{k+1}^2)^2]^{\frac{1}{2}}}{2^{M_{k+1}}} \right).$$

In addition, note that

$$\begin{aligned} \mathbb{E}_{t-1}[\widehat{G}_t] &= \mathbb{E}_{t-1} \left[\frac{1}{q_t(\lambda_t)} \left(\widehat{g}_t^{\lambda_t} - \frac{1}{2} \widehat{g}_t^{\lambda_t, e} - \frac{1}{2} \widehat{g}_t^{\lambda_t, o} \right) \right] = \sum_{\ell=0}^{M_t} \mathbb{E}_{t-1} \left[\widehat{g}_t^\ell - \frac{1}{2} \widehat{g}_t^{\ell, e} - \frac{1}{2} \widehat{g}_t^{\ell, o} \right] \\ &= \mathbb{E}_{t-1}[\widehat{g}_t^{M_t}] = \mathbb{E}_{t-1} \left[\widehat{\mathbb{E}}_t^{M_t}[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}]) \right], \end{aligned} \quad (49)$$

where the first equality follows from the definition of \widehat{G}_t in (23), and the second uses the law of total expectation along with the independence of λ_t from all other random objects. The third equality holds because $\widehat{g}_t^{\ell, e}$ and $\widehat{g}_t^{\ell, o}$ share the distribution of $\widehat{g}_t^{\ell-1}$ conditional on \mathcal{F}_t for all $\ell \in [M_t]$ and because $\widehat{g}_t^{0, e} = \widehat{g}_t^{0, o} = 0$. Finally, the fourth equality follows from Definition 4.2. Next, we introduce the shorthands

$$\begin{aligned} A_1 &:= \left\| \mathbb{E}_{t-1} \left[\widehat{\mathbb{E}}_t^{M_t}[\widehat{G}_{t+1}] \left(\nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) \right) \right] \right\|_2 \\ A_2 &:= \left\| \mathbb{E}_{t-1} \left[\left(\mathbb{E}_t[\widehat{G}_{t+1}] - G_{t+1} \right) \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) \right] \right\|_2 \end{aligned}$$

to simplify notation. Then, we have

$$\begin{aligned} \|\mathbb{E}_{t-1}[\widehat{G}_t] - G_t\|_2 &= \left\| \mathbb{E}_{t-1} \left[\widehat{\mathbb{E}}_t^{M_t}[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}]) \right] - \mathbb{E}_{t-1} \left[G_{t+1} \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) \right] \right\|_2 \\ &\leq \left\| \mathbb{E}_{t-1} \left[\widehat{\mathbb{E}}_t^{M_t}[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}]) - \widehat{\mathbb{E}}_t^{M_t}[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) \right] \right\|_2 \\ &\quad + \left\| \mathbb{E}_{t-1} \left[\widehat{\mathbb{E}}_t^{M_t}[\widehat{G}_{t+1}] \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) - G_{t+1} \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) \right] \right\|_2 = A_1 + A_2, \end{aligned} \quad (50)$$

where the first equality exploits (49) and the definition of G_t in (22). The second equality follows from the law of total expectation and the elementary identity $\mathbb{E}_t[\widehat{\mathbb{E}}_t^{M_t}[\cdot]] = \mathbb{E}_t[\cdot]$. Next, we separately bound the terms A_1 and A_2 . As for A_1 , we use the Cauchy-Schwarz and Jensen inequalities to obtain

$$\begin{aligned} A_1 &\leq \mathbb{E}_{t-1} \left[\widehat{\mathbb{E}}_t^{M_t} \left[\|\widehat{G}_{t+1}\|_2 \right] \left\| \nabla_{x_t} f_t(\xi_t, \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}]) - \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) \right\|_2 \right] \\ &\leq S_t \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_2 \left\| \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}] - F_{t+1}^T \right\|_2 \right] \\ &\leq S_t \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_2^2 \right]^{\frac{1}{2}} \mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}] - F_{t+1}^T \right\|_2^2 \right]^{\frac{1}{2}} \\ &= S_t \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_2^2 \right]^{\frac{1}{2}} \left(\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^2 \right] + \mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[\widehat{H}_{t+1}] - F_{t+1}^T \right\|_2^2 \right] \right)^{\frac{1}{2}}. \end{aligned}$$

Here, the second inequality follows from the smoothness of the integrand f_t , as stipulated in Assumption 3.3. Note that we have also removed the first sample average $\widehat{\mathbb{E}}_t^{M_t}[\cdot]$. This is permissible because the second norm term does not depend on ξ_{t+1} and because the underlying samples are independent and identically distributed conditional on the information available at time $t - 1$. The third inequality and the equality follow from the Cauchy-Schwarz inequality and from a standard bias-variance decomposition of the mean squared error. By Lemma C.6, the bias term admits the following upper bound.

$$\mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[\widehat{H}_{t+1}] - F_{t+1}^T \right\|_2^2 \right] \leq \mathbb{E}_{t-1} \left[\left(\sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}_t[\mu_{s+1}^2]}{2^{M_s+1}} \right)^2 \right]$$

In addition, the variance term satisfies

$$\mathbb{E}_{t-1} \left[\left\| \widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}] - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^2 \right] = 2^{-M_t} \mathbb{E}_{t-1} \left[\left\| \widehat{H}_{t+1} - \mathbb{E}_t[\widehat{H}_{t+1}] \right\|_2^2 \right] \leq 2^{-M_t} \mathbb{E}_{t-1}[\mu_{t+1}^2],$$

where the equality holds because the variance of the sample average $\widehat{\mathbb{E}}_t^{M_t}[\widehat{H}_{t+1}]$ equals the variance of \widehat{H}_{t+1} divided by the sample size 2^{M_t} . The inequality follows from the observation that the variance of \widehat{H}_{t+1} is bounded below by its second moment and from the definition of μ_{t+1}^2 in (19). Substituting the bounds on these mean and the variance terms into the above bound on A_1 then yields

$$\begin{aligned} A_1 &\leq S_t \mathbb{E}_{t-1} \left[\|\widehat{G}_{t+1}\|_2^2 \right]^{\frac{1}{2}} \left(\mathbb{E}_{t-1} \left[\left(\sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}_t[\mu_{s+1}^2]}{2^{M_s+1}} \right)^2 \right] + \frac{\mathbb{E}_{t-1}[\mu_{t+1}^2]}{2^{M_t}} \right)^{\frac{1}{2}} \\ &\leq S_t \mathbb{E}_{t-1}[\nu_{t+1}^2]^{\frac{1}{2}} \left(\mathbb{E}_{t-1} \left[\left(\sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}_t[\mu_{s+1}^2]}{2^{M_s+1}} \right)^2 \right]^{\frac{1}{2}} + \frac{\mathbb{E}_{t-1}[\mu_{t+1}^2]^{\frac{1}{2}}}{2^{M_t/2}} \right), \end{aligned}$$

where the second inequality holds because of the definition of ν_{t+1}^2 in (25) and because $(a^2 + b^2)^{\frac{1}{2}} \leq |a| + |b|$ for all $a, b \in \mathbb{R}$. The Minkowski and Jensen inequalities further imply that

$$\begin{aligned} \mathbb{E}_{t-1} \left[\left(\sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}_t[\mu_{s+1}^2]}{2^{M_s+1}} \right)^2 \right]^{\frac{1}{2}} &\leq \sum_{s=t+1}^{T-1} \mathbb{E}_{t-1} \left[\frac{L_{[t+1:s-1]}^2 S_s^2 \mathbb{E}_t[\mu_{s+1}^2]^2}{2^{2M_s+2}} \right]^{\frac{1}{2}} \\ &\leq \sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}_{t-1}[(\mu_{s+1}^2)^2]^{\frac{1}{2}}}{2^{M_s+1}}. \end{aligned}$$

Using this inequality to simplify that above bound on A_1 yields

$$A_1 \leq S_t \mathbb{E}_{t-1}[\nu_{t+1}^2]^{\frac{1}{2}} \left(\sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}_{t-1}[(\mu_{s+1}^2)^2]^{\frac{1}{2}}}{2^{M_s+1}} + \frac{\mathbb{E}_{t-1}[\mu_{t+1}^2]^{\frac{1}{2}}}{2^{M_t/2}} \right).$$

As for the term A_2 , we use the Cauchy-Schwarz and Jensen inequalities to obtain

$$\begin{aligned}
A_2 &\leq \mathbb{E}_{t-1} \left[\left\| \nabla_{x_t} f_t(\xi_t, F_{t+1}^T) \right\|_2 \left\| \mathbb{E}_t[\widehat{G}_{t+1}] - G_{t+1} \right\|_2 \right] \\
&\leq L_t \mathbb{E}_{t-1} \left[\left\| \mathbb{E}_t[\widehat{G}_{t+1}] - G_{t+1} \right\|_2 \right] \\
&\leq L_t \mathbb{E}_{t-1} \left[\sum_{s=t+1}^{T-1} L_{[t+1:s-1]} S_s \mathbb{E}_t[\nu_{s+1}^2]^{\frac{1}{2}} \left(\frac{\mathbb{E}_t[\mu_{s+1}^2]^{\frac{1}{2}}}{2^{M_s/2}} + \left(\sum_{k=s+1}^{T-1} \frac{L_{[s+1:k-1]} S_k \mathbb{E}_t[(\mu_{k+1}^2)^2]^{\frac{1}{2}}}{2^{M_k+1}} \right) \right) \right],
\end{aligned}$$

where the second inequality follows from Assumption 3.2, and the third inequality exploits the induction hypothesis. Jensen's inequality together with the law of total expectations further yields

$$A_2 \leq L_t \sum_{s=t+1}^{T-1} L_{[t+1:s-1]} S_s \mathbb{E}_{t-1}[\nu_{s+1}^2]^{\frac{1}{2}} \left(\frac{\mathbb{E}_{t-1}[\mu_{s+1}^2]^{\frac{1}{2}}}{2^{M_s/2}} + \left(\sum_{k=s+1}^{T-1} \frac{L_{[s+1:k-1]} S_k \mathbb{E}_{t-1}[(\mu_{k+1}^2)^2]^{\frac{1}{2}}}{2^{M_k+1}} \right) \right).$$

Combining the bounds on A_1 and A_2 with (50) and noting that $L_{[t:t-1]} = 1$, we finally obtain

$$\begin{aligned}
&\left\| \mathbb{E}_{t-1}[\widehat{G}_t] - G_t \right\|_2 \\
&\leq S_t \mathbb{E}_{t-1}[\nu_{t+1}^2]^{\frac{1}{2}} \left(\frac{\mathbb{E}_{t-1}[\mu_{t+1}^2]^{\frac{1}{2}}}{2^{M_t/2}} + \left(\sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}_{t-1}[(\mu_{s+1}^2)^2]^{\frac{1}{2}}}{2^{M_s+1}} \right) \right) \\
&\quad + L_t \sum_{s=t+1}^{T-1} L_{[t+1:s-1]} S_s \mathbb{E}_{t-1}[\nu_{s+1}^2]^{\frac{1}{2}} \left(\frac{\mathbb{E}_{t-1}[\mu_{s+1}^2]^{\frac{1}{2}}}{2^{M_s/2}} + \left(\sum_{k=s+1}^{T-1} \frac{L_{[s+1:k-1]} S_k \mathbb{E}_{t-1}[(\mu_{k+1}^2)^2]^{\frac{1}{2}}}{2^{M_k+1}} \right) \right) \\
&= \sum_{s=t}^{T-1} L_{[t:s-1]} S_s \mathbb{E}_{t-1}[\nu_{s+1}^2]^{\frac{1}{2}} \left(\frac{\mathbb{E}_{t-1}[\mu_{s+1}^2]^{\frac{1}{2}}}{2^{M_s/2}} + \sum_{k=s+1}^{T-1} \frac{L_{[s+1:k-1]} S_k \mathbb{E}_{t-1}[(\mu_{k+1}^2)^2]^{\frac{1}{2}}}{2^{M_k+1}} \right).
\end{aligned}$$

This completes the induction step, and thus the claim follows. \square

Proof of Theorem 4.9. To bound the root mean squared error of $\widehat{G}(x)$ by ϵ , it suffices to bound both the bias and the standard deviation of $\widehat{G}(x)$ by $\epsilon/\sqrt{2}$ for any $x \in \mathcal{X}$. As for the bias, Lemma 4.6 implies that

$$\begin{aligned}
\left\| \mathbb{E}[\widehat{G}(x)] - G_1(x) \right\|_2 &\leq \sum_{t=1}^{T-1} L_{[t-1]} S_t \mathbb{E}[\nu_{t+1}^2(x)]^{\frac{1}{2}} \left(\frac{\mathbb{E}[\mu_{t+1}^2(x)]^{\frac{1}{2}}}{2^{M_t/2}} + \sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}[(\mu_{s+1}^2(x))^2]^{\frac{1}{2}}}{2^{M_s+1}} \right) \\
&\leq \sum_{t=1}^{T-1} L_{[t-1]} S_t \mathbb{E}[\nu_{t+1}^{2^{t+1}}(x)]^{\frac{1}{2^{t+1}}} \left(\frac{\mathbb{E}[\mu_{t+1}^{2^{t+1}}(x)]^{\frac{1}{2^{t+1}}}}{2^{M_t/2}} + \sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s \mathbb{E}[(\mu_{s+1}^{2^{s+1}}(x))^2]^{\frac{1}{2^{s+1}}}}{2^{M_s+1}} \right),
\end{aligned}$$

where the last expression follows from repeated application of Jensen's inequality and from the definitions of the conditional moments $\mu_t^p(x)$ and $\nu_t^p(x)$ in (19) and (25), respectively. Lemma 4.5 further implies that

$$\mathbb{E}[\nu_{t+1}^{2^{t+1}}(x)]^{\frac{1}{2^{t+1}}} \leq E_{t+1}^{\frac{1}{2^{t+1}}} \prod_{s=t+1}^{T-1} \max\{\mathcal{S}_s, \mathcal{S}'_s\}^{\frac{1}{2^{t+1}}},$$

where the auxiliary variables \mathcal{S}_s and \mathcal{S}'_s are defined as in (38) in the proof of Lemma 4.5. Thus, we have

$$\mathcal{S}_s = \sum_{\ell=0}^{M_s} \frac{1}{2^{(2^s-1)(\rho_s+1)\ell} q_s(\ell)^{2^s-1}} = \sum_{\ell=0}^{M_s} \frac{Z_s^{2^s-1}}{2^{(2^s-1)(\rho_s+1)\ell} (1-r_s)^{(2^s-1)\ell}}.$$

Here, the second equality holds because $q_s(\ell) = (1-r_s)^\ell/Z_s$, and $Z_s := (1-(1-r_s)^{M_s+1})/r_s$ is shorthand for the normalization constant of the truncated geometric distribution $\text{Geo}(r_s|M_s)$. Note that $Z_s \leq 1/r_s$. The first upper bound on r_s in Assumption 4.8 thus ensures that \mathcal{S}_s remains bounded as M_s tends to infinity. Consequently, there exists a constant $\bar{\mathcal{S}}_s \geq \mathcal{S}_s$ that is independent of M_s . Similarly, one can show that there exists a constant $\bar{\mathcal{S}}'_s \geq \mathcal{S}'_s$ that is independent of M_s . In summary, we may then conclude that

$$\mathbb{E}[\nu_{t+1}^{2^{t+1}}(x)]^{\frac{1}{2^{t+1}}} \leq E_{t+1}^{\frac{1}{2^{t+1}}} \prod_{s=t+1}^{T-1} \max\{\bar{\mathcal{S}}_s, \bar{\mathcal{S}}'_s\}^{\frac{1}{2^{t+1}}}$$

uniformly across all $x \in \mathcal{X}$ and $M_s, s = t+1, \dots, T-1$. As $\rho_s \leq 1$, Lemma 3.19 further implies that

$$\mu_{t+1}^{2^{t+1}}(x) \leq D_{t+1} \prod_{s=t+1}^{T-1} \bar{\mathcal{S}}_s$$

uniformly across all $x \in \mathcal{X}$ and $M_s, s = t+1, \dots, T-1$. Combining the above estimates then yields

$$\begin{aligned} \|\mathbb{E}[\widehat{G}(x)] - G_1(x)\|_2 &\leq \sum_{t=1}^{T-1} L_{[t-1]} S_t E_{t+1}^{\frac{1}{2^{t+1}}} \left(\prod_{s=t+1}^{T-1} \max\{\bar{\mathcal{S}}_s, \bar{\mathcal{S}}'_s\}^{\frac{1}{2^{t+1}}} \right) \\ &\times \left(\frac{D_{t+1}^{\frac{1}{2^{t+1}}} \prod_{s=t+1}^{T-1} \bar{\mathcal{S}}_s^{\frac{1}{2^{t+1}}}}{2^{M_t/2}} + \sum_{s=t+1}^{T-1} \frac{L_{[t+1:s-1]} S_s D_{s+1}^{\frac{1}{2^s}} \prod_{s'=s+1}^{T-1} \bar{\mathcal{S}}_{s'}^{\frac{1}{2^{s'}}}}{2^{M_s+1}} \right). \end{aligned}$$

By construction, all constants in the resulting bound are independent of the truncation points. Since $M_s \geq 0$, we can lower bound the denominator 2^{M_s+1} by $2^{M_s/2}$ and introduce auxiliary constants $W'_t \geq 0$ that are independent of the truncation points to obtain a simpler but relaxed bias bound of the form

$$\|\mathbb{E}[\widehat{G}(x)] - G_1(x)\|_2 \leq \sum_{t=1}^{T-1} \frac{W'_t}{2^{M_t/2}}.$$

Note that W'_t absorbs all dependencies of the bias on the Lipschitz, smoothness and Hölder constants, as well as on the dimensions, moment bounds and rate parameters. Setting $W_t = W'_t \sqrt{2}(T-1)$, it is now evident that the bias drops below $\epsilon/\sqrt{2}$ uniformly across all $x \in \mathcal{X}$ if $M_t = \lceil 2 \log_2(W_t/\epsilon) \rceil$ for all $t \in [T-1]$.

Next, we show that the variance of the MLMC gradient estimator is bounded by $\epsilon^2/2$. From the first part of the proof we know that $\nu_1^2(x)$ is bounded above by $E_1 \prod_{t=1}^{T-1} \max\{\bar{\mathcal{S}}_t, \bar{\mathcal{S}}'_t\}$ uniformly across all $x \in \mathcal{X}$. Thus, both $\sup_{x \in \mathcal{X}} \nu_1^2(x)$ and n_1 are finite. Consequently, we may conclude that

$$\mathbb{V}(\widehat{G}(x)) \leq \frac{\nu_1^2(x)}{n_1} \leq \frac{\epsilon^2}{2}$$

uniformly across all $x \in \mathcal{X}$, where the two inequalities follow from (26) and the definition of n_1 .

Finally, the expected sampling cost of the MLMC gradient estimator is given by

$$\mathbb{E}[\mathbb{C}(\widehat{G}(x))] = n_1 \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} q_t(\ell) 2^\ell = n_1 \prod_{t=1}^{T-1} \sum_{\ell=0}^{M_t} \frac{2^\ell (1-r_t)^\ell}{Z_t} = n_1 \prod_{t=1}^{T-1} \frac{1 - (2-2r_t)^{M_t+1}}{Z_t(2r_t-1)},$$

where the first equality can be proved as in Lemma 3.22. As $2^{-1} < r_t < 1$ by virtue of Assumption 4.8, the resulting expression remains bounded as M_t tends to infinity.

In summary, the above reasoning implies that the root mean squared error of the MLMC gradient estimator is bounded by ϵ uniformly across all $x \in \mathcal{X}$ if $n_1 = \mathcal{O}(\epsilon^{-2})$ and $M_t = \mathcal{O}(\log(\epsilon^{-1}))$ for all $t \in [T-1]$, in which case the expected sampling cost satisfies $\mathbb{E}[\mathbb{C}(\widehat{G}(x))] \leq \mathcal{O}(\epsilon^{-2})$. Hence, the claim follows. \square

Proof of Theorem 4.11. Under the stated assumptions, Lemmas 4.5, 4.6, and 4.7, together with Theorem 4.9, are applicable. It is therefore easy to verify that the claim is a direct consequence of [41, Theorem C.1]. \square