Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Data-Driven Compositional Optimization in Misspecified Regimes

Shuoguang Yang, Ethan X. Fang, Uday V. Shanbhag

To cite this article:
Shuoguang Yang, Ethan X. Fang, Uday V. Shanbhag (2024) Data-Driven Compositional Optimization in Misspecified Regimes. Operations Research

Published online in Articles in Advance 16 Feb 2024

. https://doi.org/10.1287/opre.2021.0295

https://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–17
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

Data-Driven Compositional Optimization in Misspecified Regimes

Shuoguang Yang[a], **Ethan X. Fang**[b], **Uday V. Shanbhag**[c]*

[a]Department of Industrial Engineering & Decision Analytics, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR, China; [b]Department of Biostatistics & Bioinformatics, Duke University, Durham, North Carolina 27705; [c]Department of Industrial & Manufacturing Engineering, Pennsylvania State University, University Park, Pennsylvania 16803

*Corresponding author

Contact: yangsg@ust.hk, https://orcid.org/0000-0002-0915-196X (SY); xingyuan.fang@duke.edu, https://orcid.org/0000-0003-3762-9155 (EXF); udaybag@engr.psu.edu, https://orcid.org/0000-0001-5869-9561 (UVS)

Received: May 13, 2021
Revised: February 11, 2022; December 5, 2022; August 23, 2023
Accepted: August 28, 2023
Published Online in Articles in Advance: February 16, 2024

Area of Review: Optimization

https://doi.org/10.1287/opre.2021.0295

Abstract. As systems grow in size, scale, and intricacy, the challenges of misspecification become even more pronounced. In this paper, we focus on parametric misspecification in regimes complicated by risk and nonconvexity. When this misspecification may be resolved via a parallel learning process, we develop data-driven schemes for resolving a broad class of misspecified stochastic compositional optimization problems. Notably, this rather broad class of compositional problems can contend with challenges posed by diverse forms of risk, dynamics, and nonconvexity, significantly extending the reach of such avenues. Specifically, we consider the minimization of a stochastic compositional function over a closed and convex set X in a regime, where certain parameters are unknown or misspecified. Existing algorithms can accommodate settings where the parameters are correctly specified, but efficient first-order schemes are hitherto unavailable for the imperfect information compositional counterparts. Via a data-driven compositional optimization approach, we develop asymptotic and rate guarantees for unaccelerated and accelerated schemes for convex, strongly convex, and nonconvex problems in a two-level regime. Additionally, we extend the accelerated schemes to the general *T*-level setting. Notably, the nonasymptotic rate guarantees in all instances show no degradation from the rate statements obtained in a correctly specified regime. Further, under mild assumptions, our schemes achieve optimal (or near-optimal) sample complexities for general *T*-level strongly convex and nonconvex compositional problems, providing a marked improvement over prior work. Our numerical experiments support the theoretical findings based on the resolution of a misspecified three-level compositional risk-averse optimization problem.

Funding: E. X. Fang is partially supported by the National Science Foundation [Grants DMS-2230795 and DMS-2230797]. U. V. Shanbhag is partially supported by the Office of Naval Research [Grant N00014-22-1-2589] and the Department of Energy [Grant DE-SC0023303].

Supplemental Material: The e-companion is available at https://doi.org/10.1287/opre.2021.0295.

Keywords stochastic optimization · compositional optimization · misspecification · risk-averse optimization

1. Introduction

Since the seminal paper by Robbins and Monro (1951), stochastic gradient descent (SGD) methods have proven to be extraordinarily powerful in minimizing the expected value or the sum of a large number of loss functions. This avenue has been widely adopted in addressing a broad collection of problems arising in engineering, applied sciences, statistics, and machine learning, among many others (Kushner and Yin 2003, Borkar 2008, Shapiro et al. 2009). The deterministic counterpart of SGD, referred to as gradient descent (GD) methods (Beck 2017), requires computing gradients, a challenging proposition when the objective is complicated by the presence of an expectation over a general probability space. To this end, classical SGD methods iteratively update estimators by employing sampled gradients. Yet, first-order methods for deterministic optimization models are often predicated on the belief that the computation of functions and gradients is possible. Often this assumption fails to hold as seen by considering a prototypical constrained optimization problem

$$\min_{x \in X} f(x, \theta), \qquad (\mathrm{Opt}(\theta))$$

in which $f : X \times R^n \to R$ is a real-valued function, X is a suitably defined set, and θdenotes a vector of parameters.

(i) *Stochastic optimization.* Consider a setting where the parameter θis a random variable, defined as θ¢ ξ,

where $\xi : \Omega \to R^d$. Let $\rho[\bullet]$ denote a suitable risk measure. In such an instance, (Opt(θ)) is modified to allow for minimizing $\rho[f(x, \xi(\omega))]$, which reduces to $E[f(x, \xi(\omega))]$ in the risk-neutral regime. One avenue for resolving this problem in the absence of gradients of $\rho[f(x, \xi(\omega))]$ leverages the availability of an additional stochastic first-order oracle that provides sampled gradients with suitable bias properties (Shapiro et al. 2009). Such models assume relevance when contending with randomness in future demand, volatility in prices, and stochasticity in available capacities.

(ii) *Robust optimization.* Robust optimization problems emerge in settings where the parameter θ is not known but instead, the user employs a belief that θ lies in an uncertainty set T (Bertsimas and Sim 2003, 2004). This leads to an uncertain collection of problems, captured by the set $\{\min_{x\in X} f(x,\theta)\}_{\theta\in T}$, whereas a robust solution (Ben-Tal et al. 2009) requires solving the following worst-case problem:

$$\min_{x\in X} \max_{\theta\in T} f(x, \theta): \qquad \text{(ROpt)}$$

(iii) *Resolving misspecification via data-driven optimization (DDO).* A third approach inspired by classical literature on misspecification in economics (Kirman 1975, Okuguchi 1976, Okuguchi and Szidarovszky 1990) considers an avenue where θ has a true or nominal value, denoted by θ^*, that may be learned through data; for example, θ^* may represent idiosyncratic machine parameters (such as efficiencies, start-up time, shutdown times, etc.), the parameters of an inverse-demand function, or the parameters of a polynomial cost function, all of which may be learned by access to observational data. This is distinct from setting (i) because θ is by no means a random variable. Similarly, if one employs an uncertainty set T instead of leveraging data, then one expects a poorer (conservative) solution assuming that $\theta^* \in T$; however, the DDO framework is predicated on the availability of data.

There has been a recent emergence of interest in considering optimization problems where key parameters are either unknown or misspecified. For instance, inventory control problems may often be afflicted by such challenges in that the demand distribution is unknown a priori (Qi et al. 2021) or the back-ordering cost may not always be known or easily estimated (Wu et al. 2013). Similarly, in the context of Markov decision processes (MDPs), as evidenced by applications in communication networks (Dai et al. 2014) and disease progression (Zhu et al. 2019), among others, rewards and transition matrices may be either misspecified (Jiang and Shanbhag 2015, Mankowitz et al. 2020, Miyaguchi 2021) or unavailable (Sato et al. 1982; Mannor et al. 2012, 2016; Ho et al. 2018).

Naturally, we may consider an "estimate-then-optimize" approach where in the estimation stage, we learn an estimator $\hat{\theta}$ for θ^*. In the subsequent optimization stage, we minimize the resulting function $f(\bullet, \hat{\theta})$, rather than a simultaneous approach, where we generate a sequence of estimators $\{\theta_k\}$ and an adapted sequence of solution estimators $\{x_k\}$ where the sequence $\{(x_k, \theta_k)\}$ converges to a minimizer of (x^*, θ^*) and θ^*, respectively. Several distinctions exist between the two approaches.

(i) *Asymptotic convergence guarantees.* Unlike simultaneous approaches, estimate-then-optimize approaches are generally not equipped with asymptotic convergence guarantees because learning of θ^* is conducted in finite time, leading to at best an approximate solution given by $\hat{\theta}$. Consequently, the optimization phase requires the minimization of $f(\bullet, \hat{\theta})$, rather than $f(\bullet, \theta^*)$, leading to a solution $\hat{x}$. Prior efforts adopting this estimate-then-optimize approach have demonstrated the gap between $\hat{x}$ and x^* (Jiang and Shanbhag 2016, Ahmadi and Shanbhag 2020).

(ii) *Complexity of learning θ^*:* When the learning problem is a high-dimensional stochastic optimization problem, the "learning" phase in the estimate-then-optimize approach can be onerous, requiring significant effort. During this period, improved solution estimators for x^* are unavailable. However, in the simultaneous approach, one generates a sequence of estimators $\{x_k\}$ throughout the scheme.

(iii) *Serial versus parallel implementations.* The estimate-then-optimize approach is natively serial, requiring the "learning" phase to precede the optimization phase, whereas the simultaneous approach allows for leveraging the inherent parallelism and running both schemes largely in parallel, allowing for significant improvements in run-time behavior.

Accordingly, the *simultaneous* approach generates a sequence $\{(x_k, \theta_k)\}$ that converges to a minimizer of $f(\bullet, \theta^*)$ and θ^*, respectively. We draw inspiration from the DDO framework for resolving misspecification presented by Jiang and Shanbhag (2013, 2016) in the context of misspecified stochastic convex optimization and variational inequality problems. In particular, this framework considers the problem of computing

$$x^* \in \arg\min_{x\in X} E_w[f_w(x; \theta^*)] \quad \text{and}$$
$$\theta^* \in \arg\min_{\theta\in\Theta} E_\xi[\varphi_\xi(\theta)],$$

where $E_w[\bullet]$ and $E_\xi[\bullet]$ are the expectations with respect to w and ξ, respectively, and θ^* serves as the unique minimizer of a strongly convex function $E_\xi[\varphi_\xi(\bullet)]$. Yet, this framework cannot contend with three key

challenges that emerge in practical design and operational questions:

(i) *Risk aversion.* Recent weather events in Texas as well as the impact of COVID-19 on worldwide supply chains have reinforced the need for incorporating risk and reliability in decision making. When one overlays the need for knowing problem parameters (lead times, failure rates, etc.), the resulting problem becomes challenging. However, existing schemes cannot contend with a risk-averse regime when the risk-neutral measure $\mathrm{E}_w[\bullet]$ is replaced by a suitably defined risk measure $\rho_w[\bullet]$.

(ii) *Nonconvexity.* Convex formulations, although allowing for tractability, often fail to accommodate high-fidelity formulations and bring forth the need for nonconvex formulations. For instance, in some settings, cost functions may be concave whereas revenue functions might be contingent on complicated price functions, leading to a nonconvex metric. Unfortunately, prior work cannot contend with such nonconvexity.

(iii) *Multiperiod decision making.* In many settings, static optimization models alone do not suffice, particularly in the face of accommodating dynamic decision-making models. One avenue for modeling such problems is through MDPs. The cost functions and the transition matrix in an MDP are often assumed to be known, but both may require learning while resolving such a problem.

To this end, we consider extending the DDO framework to convex/nonconvex compositional regimes (Wang et al. 2017a, b; Yang et al. 2019), a framework that accommodates (i)–(iii) and considers the following problem.

The two-level data-driven compositional optimization (DDCO) problem requires an $(x^*, \theta_1^*, \theta_2^*)$ such that

$$x^* \in \arg\min_{x\in X} \mathrm{E}_v[f_v(\mathrm{E}_w[g_w(x;\theta_2^*)];\theta_1^*)],$$

$$\text{where } \theta_1^* \in \arg\min_{\theta_1\in\Theta_1} \mathrm{E}_{\xi_1}[\varphi_{\xi_1}^{(1)}(\theta_1)] \text{ and}$$

$$\theta_2^* \in \arg\min_{\theta_2\in\Theta_2} \mathrm{E}_{\xi_2}[\varphi_{\xi_2}^{(2)}(\theta_2)]. \tag{1.1}$$

In this paper, we consider the setting where $\varphi^{(1)}$ and $\varphi^{(2)}$, defined as $\varphi^{(1)}(\theta_1) \triangleq \mathrm{E}_{\xi_1}[\varphi_{\xi_1}^{(1)}(\theta_1)]$ and $\varphi^{(2)}(\theta_2) \triangleq \mathrm{E}_{\xi_2}[\varphi_{\xi_2}^{(2)}(\theta_2)]$, are both strongly convex over the sets Θ_1 and Θ_2, respectively. We focus on the scenario where the random variables v and w are independent of each other. Let $f(\bullet, \theta_1)$ and $g(\bullet, \theta_2)$ be defined as $f(y;\theta_1) \triangleq \mathrm{E}_v[f_v(y;\theta_1)]$ and $g(x;\theta_2) \triangleq \mathrm{E}_w[g_w(x;\theta_2)]$, respectively, where $\theta_1 \in \Theta_1$ and $\theta_2 \in \Theta_2$. Here $g(\bullet;\theta_2)$ is a general vector-valued function, whereas θ_1, θ_2 are exogenous parameters that have no impact on the distribution of v and w. Moreover, the objective $F(\bullet;\theta_1,\theta_2)$ is defined as $F(x;\theta_1,\theta_2) \triangleq f(g(x;\theta_2);\theta_1)$, the composition of f and g. Further, in practice, there may be multiple levels of composition, necessitating the consideration of the general T-level misspecified compositional problem, defined as

$$\min_{x\in X} F(x;\theta_1^*,\ldots,\theta_T^*)$$
$$\triangleq \mathrm{E}_{\omega_1}[f_{\omega_1}^{(1)}(\mathrm{E}_{\omega_2}[f_{\omega_2}^{(2)}(\cdots(\mathrm{E}_{\omega_T}[f_{\omega_T}^{(T)}(x;\theta_T^*)]);\cdots;\theta_2^*)];\theta_1^*)],$$
$$\text{where } \theta_j^* \in \arg\min_{\theta_j\in\Theta_j} \mathrm{E}_{\xi_j}[\varphi_{\xi_j}^{(j)}(\theta_j)], \text{ for } j = 1,2,\ldots,T.$$

Next, we discuss some motivating applications.

1.1. Motivating Applications

We present two applications of particular relevance.

(A) *Misspecified risk-averse newsvendor problems.* We begin by considering the relatively simplified misspecified variant of the newsvendor problem that determines the optimal order of a perishable item to maximize expected revenue in a single period (Arrow et al. 1951, Hadley and Whitin 1963, Porteus 1990, Arrow 2002). Suppose the unit cost of the item is given by c^* and the demand is a random variable denoted by $d : \Omega \to \mathrm{R}^+$. Furthermore, if the order quantity, denoted by x, is less than demand, then the cost of backordering is given by $b^*(d(\omega) - x)^+$ whereas the holding cost is given by $h^*(x - d(\omega))^+$. If the order quantity has to be specified before observing the demand realization, the risk-neutral newsvendor problem is given by the following:

$$\min_{x\geq 0} \mathrm{E}[c^*x + b^*(d(\omega) - x)^+ + h^*(x - d(\omega))^+]. \tag{1.2}$$

Under the caveat that the probability distribution is known and defined as $D(u) \triangleq \mathrm{P}(\omega : d(\omega) \leq u)$, it is well known that under suitable distributional assumptions (Shapiro et al. 2009), the optimal order size x^* can be expressed as

$$x^* = D^{-1}(\kappa(b^*, c^*, h^*)), \text{ where } \kappa(b^*, c^*, h^*) \triangleq \frac{(b^* - c^*)}{(b^* + h^*)}. \tag{1.3}$$

Yet such an avenue is inherently fraught with challenges from the standpoint of data:

(a) First, the distribution of d is assumed to be known as a priori, a claim that may be often violated to varying degrees; for example, whereas demand may be known to be normally distributed, its precise mean and standard deviation, denoted by μ^* and σ^*, respectively, may be unknown to the decision maker (DM). The parameters μ^* and σ^* may be estimated via the sample-mean estimator and the sample standard deviation

estimator, whose convergence guarantees to their true counterparts can be ensured under mild assumptions.

(b) Second, the parameters b^*, c^*, and h^*, although essential for solving for the optimal order quantity, may require learning through observations, where the costs can be uncertain (Wu et al. 2013).

If $\theta^* \triangleq (\mu^*, \sigma^*, b^*, c^*, h^*)$ denotes the information not known a priori, then a mean-risk variant of the *misspecified* newsvendor problem is given by the following:

$$\min_{x \geq 0} \mathrm{E}[F(x, \omega; b^*, c^*, h^*)] + \beta \rho_{\mu^*, \sigma^*}[F(x, \omega; b^*, c^*, h^*)], \quad (\mathrm{N}\ V_2(\theta^*))$$

where $F(x, \omega; b^*, c^*, h^*) \triangleq c^* x + b^*(d(\omega) - x)^+ + h^*(x - d(\omega))^+$, β denotes a risk-aversion parameter, and $\rho_{\mu^*, \sigma^*}[\cdot]$ denotes a risk measure parameterized by μ^* and σ^*. Akin to earlier, θ^* is a consequence of resolving the problem $\min_{\theta \in \Theta} \mathrm{E}_\xi[\varphi_\xi(\theta)]$. Although $\rho[\cdot]$ can be chosen in a multitude of ways, we consider the mean upper semideviation (MUS) measure, defined as follows.

Mean upper semideviation (MUS) measure: Often one is concerned with the downside risk, and one avenue for capturing this is through the mean upper semideviation measure (Shapiro et al. 2009). Assuming that ρ is the p-th order mean upper semideviation measure with $p > 0$, and letting $F(\cdot, \omega; b^*, c^*, h^*)$ be the loss function, we have that the parameterized measure ρ^{mus} is defined as

$$\rho^{\mathrm{mus}}_{\mu^*, \sigma^*}[F(x, \omega; b^*, c^*, h^*)] \triangleq [\mathrm{E}_\omega[F(x, \omega; b^*, c^*, h^*) - \mathrm{E}_\omega[F(x, \omega; b^*, c^*, h^*)]]^p_+]^{1/p}. \quad (1.4)$$

This problem falls into the class of misspecified three-level compositional optimization by setting $f^{(1)}(z) \triangleq z^{1/p}$, $f^{(2)}_\omega(x, y; b^*, c^*, h^*) \triangleq F(x, \omega; b^*, c^*, h^*) - y]^p_+$, and $f^{(3)}_\omega(x; b^*, c^*, h^*) \triangleq (x, F(x, \omega; b^*, c^*, h^*)) \in \mathbb{R}^{d+1}$.

(B) *Misspecified Markov decision processes* Markov decision problems represent an enormously powerful framework for addressing dynamic decision-making problems and have been useful in a breadth of areas including manufacturing systems, healthcare systems, and economics, among others (Ross 1983, Puterman 1994). Particularly noteworthy are efforts that apply MDP models to inventory control (Cheevaprawatdomrong and Smith 2004, Satheesh Kumar and Elango 2012, Chao 2013, Li 2013, Puranam and Katehakis 2014). Consider an MDP with continuous actions $a \in \mathrm{A}$ and discrete states $s \in \mathrm{S}$. Suppose the DM takes action a in state s, resulting in a transition to a random state s' with probability $P(s'|s, a)$, yielding a simultaneous reward $R(s, a)$. Given a discount factor γ, let $Q^*(s, a) \triangleq \mathrm{E}[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \mid s_0 = s, a_0 = a]$ be the discounted state-action function with initial state $s_0 = s$ and $a_0 = a$. Further, $Q^*(s, a)$ is defined by the Bellman equation, as observed next.

$$Q^*(s, a) = R(s, a) + \gamma \sum_{s'} P(s'|s, a) \max_{a' \in \mathrm{A}} Q^*(s', a') = R(s, a) + \gamma \mathrm{E}_{s'}\left[\max_{a' \in \mathrm{A}} Q^*(s', a')\right], a$$

where the expectation is taken over all random transitions s' with respect to the state-action pair (s, a). In more generic scenarios, at each state $s \in \mathrm{S}$, the DM is required to take an action $a \in \mathrm{A}$ based on an estimator ψ for ψ^* in the environment, where the parameter ψ^* is unknown. Such a parameter ψ^* may represent some specific characteristics of the process environment and may be estimated by the DM based on noisy observations. In most cases, these parameters serve as a (unique) solution to a stochastic optimization problem, defined as

$$\psi^* \triangleq \arg\min_{\psi \in \Psi} \mathrm{E}_\xi[\Phi(\psi, \xi)], \quad (1.5)$$

where $\Phi(\psi, \xi)$ represents the (random) estimation residual and Ψ denotes the feasible region of ψ. Yet, in many settings, neither the reward R nor the transition matrix P may be available and we denote their misspecified variants by $R(s, a; \psi)$ and $P(s'|s, a, \psi)$, respectively, where ψ^* needs to be learned and a nominal estimate of ψ^* is denoted by ψ. Let $Q^*(s, a; \psi)$ represent the optimal discounted state-action function under the correctly specified model parameter ψ^* defined in (1.5). Then the corresponding Bellman equation is given by

$$Q^*(s, a; \psi) = R(s, a; \psi) + \gamma \mathrm{E}_{s'}\left[\max_{a' \in \mathrm{A}} Q(s', a'; \psi^*)\right], a.$$

Again, the conditional expectation is taken over all random transition states s' under state-action pair (s, a) and correct model parameter ψ^*. Solving the optimal policy for the MDP is equivalent to finding a fixed point of the above Bellman equation under ψ. To obtain such a fixed point, one can solve the Bellman residual minimization problem:

$$\min_{Q \in \mathbb{R}^{|\mathrm{S}| \times |\mathrm{A}|}} \sum_{(s,a) \in \mathrm{S} \times \mathrm{A}} \left(Q(s, a; \psi) - R(s, a; \psi) - \gamma \mathrm{E}_{s'}\left[\max_{a' \in \mathrm{A}} Q(s', a'; \psi^*)\right], a\right)^2, \quad (1.6)$$

where $\psi^* \triangleq \arg\min_{\psi \in \Psi} \mathrm{E}_\xi[\Phi(\psi, \xi)]$. This problem fits into the family of misspecified two-level compositional optimization problems by choosing

$$f(y) = \|y\|^2 \text{ and the } (s, a)\text{-th entry of } g(Q; \psi^*)_{(s,a)} = Q(s, a) - \mathrm{E}\left[R(s, a; \psi) + \gamma \max_{a' \in \mathrm{A}} Q(s', a')\right], a.$$

1.2. Related Work

We briefly review related decision-making paradigms in optimization and control theory and summarize prior work on misspecified data-driven optimization and compositional gradient methods.

1.2.1. Related Decision-Making Paradigms.

(i) *Multiarmed bandit (MAB) problems.* Recall that in MAB problems, a gambler decides the selection, order, and frequency of machines to be played, given a collection of slot machines, where each machine provides a random reward from a machine-specific distribution. The goal lies in maximizing the expected sum of rewards earned through a sequence of lever pulls where the reward function needs to be learned through the process. Such a framework is particularly compelling in operations and revenue management (Besbes and Zeevi 2009, 2012). Distinct from our setting, the learning of the reward function is affected by the sequence of decisions.

(ii) *Adaptive and iterative learning control.* Adaptive control (AC) (Hovakimyan and Cao 2010) problems arise in controlling systems where the parameters are either uncertain or varying slowly in time. A prototypical example arises in flight control, where the mass of an aircraft reduces in time through the burning of fuel, requiring the controller to reflect this change. Parameter estimation to facilitate this process is carried out via least mean-squares (LMS), or its recursive variant referred to as recursive least mean-squares (RLMS), or the Kalman filter. In contrast, iterative learning control (ILC) represents a form of tracking control (Uchiyama 1978, Arimoto et al. 1984) employed for repetitive control problems in batch processes, robot arm manipulators, and reliability testing rigs. Contrary to our framework, both AC and ILC often impose stringent requirements on the model but allow for more general dynamics. In contrast, the DDCO framework is more general in terms of problem primitives, requiring fewer assumptions on noise and allowing more general constrained and nonlinear formulations, but precluding complex nonlinear dynamics, at least in current regimes.

1.2.2. Related Work on Misspecification.

In Jiang and Shanbhag (2016), it was shown that in many settings, the presence of misspecification does not adversely impact the rate statements when employing stochastic gradient descent schemes. Similar statements have been developed for gradient and extragradient schemes for misspecified deterministic convex optimization problems (Ahmadi and Shanbhag 2020), gradient-response schemes for misspecified Nash equilibrium problems (Jiang et al. 2018, Lei and Shanbhag 2020), augmented Lagrangian schemes for convex optimization with misspecified constraints (Ahmadi et al. 2016, Aybat et al. 2022), and block-coordinate schemes for stochastic nonconvex programs (Lei and Shanbhag 2020). Under the moniker of "joint estimation-optimization" (JEO), online counterparts of such schemes have been developed by Ho-Nguyen and Kılınç-Karzan (2019) and such avenues have been employed by the same authors in examining nonparametric choice models (Ho-Nguyen and Kılınç-Karzan 2021). Yet, the consideration of compositional regimes remains both open and compelling.

1.2.3. Stochastic Compositional Gradient Methods.

Some of the earliest efforts on two-level compositional problems can be traced to the work of Ermoliev (1976), where the almost sure convergence of a two-timescale scheme was provided. Wang et al. (2017a) provided the first known rate statements whereas accelerated counterparts were examined subsequently in Wang et al. (2017b). Meanwhile, variance reduction has been studied in finite-sum regimes (Lian et al. 2017) whereas the very first generalization to the multistage regime was developed by Yang et al. (2019). More recently, a single-timescale scheme was presented by Ghadimi et al. (2020) where the optimal sample complexity was proven for computing stationary points of two-level stochastic nonconvex optimization problems. In multilevel settings, algorithms achieving an $O(k^{-1/2})$ rate of convergence were developed by Zhang and Lan (2020) for convex objectives. Further, extensions to nonconvex settings were presented by Chen et al. (2021b), Ruszczynski (2021), and Balasubramanian et al. (2022) with level-independent optimal sample complexities. More recently, two-level compositional optimization was studied by Yang et al. (2022) under the decentralized setting. Yet, we believe that extensions to the misspecified regime are far from immediate and remain the core focus of this research.

1.2.4. Novelty.

In this paper, we propose the study of stochastic compositional optimization problems afflicted by misspecification. Despite the relevance of such concerns in risk-averse and dynamic decision-making problems, there are no existing schemes for resolving such problems. Following recent advances in stochastic compositional optimization (Chen et al. 2021b, Balasubramanian et al. 2022), we propose new algorithms to resolve this class of problems under a range of convexity and smoothness requirements. In particular, the proposed framework allows for deriving convergence rates in convex and nonconvex regimes via unaccelerated and accelerated schemes, where the key step lies in quantifying and controlling the bias induced by misspecification, a challenging proposition in compositional settings. To solve this issue, we decompose this problem as a coupled set of stochastic recursions whose collective convergence claims are not immediate. Distinct from the correctly specified regimes, the bias terms in the recursions are

required to be analyzed carefully and the algorithm parameters have to be selected for achieving the best possible rate. Notably, our schemes achieve the optimal level-independent rate in strongly convex regimes whereas our oracle complexity is near-optimal in nonconvex regimes. Surprisingly, it is seen that the convergence rates match their counterparts in the correctly specified regime, showing no degradation induced by misspecification.

1.3. Gaps, Contributions, and Outline

This paper is motivated by a crucial gap in addressing risk-averse, nonconvex, and dynamic decision-making problems afflicted by misspecification. Current techniques cannot contend with such considerations. We consider a compositional and nonconvex generalization where direct extensions of prior work are by no means immediate. Consequently, there is a need for a comprehensive study of how existing compositional stochastic gradient schemes can contend with misspecification and the degradation, if any, that such misspecification leads to from a rate standpoint. We summarize our major contributions next and note that numerics have been provided for on a misspecified risk-averse problem (Section 5) and a misspecified MDP (e-companion Section D.1) whereas the impact of conservatism arising from employing robust approaches has also been explored (e-companion Section D.2).

(i) *Unaccelerated schemes for two-level problems.* In Section 2, we consider two-level compositional problems complicated by misspecification, when the outer function is differentiable whereas the inner function is merely continuous. Here, we derive almost sure convergence guarantees in the convex and nonconvex regime. From a rate standpoint, the sequence is characterized by a rate of $\tilde{O}(k^{-2/5})$ in the nonconvex regime in terms of an appropriate residual function, matching the correctly specified statements developed by Yang et al. (2019), where $\tilde{O}(\cdot)$ suppresses logarithmic terms.

(ii) *Accelerated schemes for two-level problems.* When we further impose a smoothness requirement on the inner function g in Section 3, the rates for the nonconvex and strongly convex regimes are seen to improve to $\tilde{O}(k^{-4/7})$ and $O(k^{-1})$, respectively. Again, we observe that these findings do not display any degradation from their correctly specified counterparts and represent the best available rates for such problem classes.

(iii) *Extensions to multilevel problems.* Finally, when (ii) is extended to T-level compositional problems in Section 4, we obtain rate statements $\tilde{O}(k^{-4/7})$ and $O(k^{-1})$ for nonconvex and strongly convex problems, respectively, again showing no degradation of the correctly specified rates obtained by Chen et al. (2021b) and Balasubramanian et al. (2022) in the nonconvex regime.

1.3.1. Organization and Notation.

The remainder of this paper is organized into five sections. We consider the application of unaccelerated and accelerated schemes on two-level compositional problems in Sections 2 and 3, respectively. Extensions to the T-level regime are considered in Section 4 whereas preliminary numerical investigations are conducted in Section 5. The paper concludes with a brief set of remarks in Section 6.

1.3.1.1. Notation.

For $x \in \mathbb{R}^n$, we denote its transpose and Euclidean norm as x' and $\|x\|$, respectively (i.e., $\|x\| \triangleq \sqrt{x'x}$). For a function f, we denote its gradient at x by $\nabla f(x)$ and denote a sampled gradient by $\nabla f_v(x)$ if f_v is differentiable, where v represents the realization of the associated random variable; the Clarke subdifferential of f at x is denoted by $\partial f(x)$ and is equivalent to the subdifferential of f at x when f is convex. We denote an element of $\partial f(x)$ by $\tilde{\nabla} f(x)$, whereas its stochastic sample is denoted by $\tilde{\nabla} f_v(x)$. We denote the expectation of f at x by $\mathrm{E}_v[f_v(x)]$ where v denotes the associated random variable; in the absence of a subscript of E, the associated random variable will be clear. We denote "converges to" as "$\to$", and denote "with probability 1" as "w.p.1." We denote the Euclidean projection of a vector $y \in \mathbb{R}^n$ on a set $X \subset \mathbb{R}^n$ by $\Pi_X[y] \triangleq \arg\min_{x \in X} \|y - x\|^2$.

2. A Basic Algorithm

In this section, we first discuss how we sample first-order information for the functions f, g, $\varphi^{(1)}$, and $\varphi^{(2)}$ at solution x_k with possibly incorrect estimates of model parameters θ_1^k and θ_2^k. Next, we propose a scheme to resolve the misspecified compositional problem (1.1). Lastly, we derive the almost sure convergence of the sequence produced by our algorithm and derive a rate guarantee.

We first define the sampling oracle (SO). In particular, we focus on the scenario where the random variables w, v are independent of each other and assume access to the following black box SO, which independently generates stochastic first-order information upon each query. Note that here we use $\tilde{\nabla}_x g(x;\theta_2)$ to handle the possible nonsmoothness of function g, whereas we directly write the stochastic gradients as ∇f_v, $\nabla \varphi^{(1)}_{\xi_1}$, and $\nabla \varphi^{(2)}_{\xi_2}$.

Sampling oracle (SO). (i) Given $x \in X \subset \mathbb{R}^n$, $\theta_2 \in \Theta_2$, the SO returns an unbiased sampled function value $g_w(x;\theta_2) \in \mathbb{R}^m$ and an unbiased sampled Jacobian $\tilde{\nabla}_x g_w(x;\theta_2) \in \mathbb{R}^{n\times m}$ such that $\mathrm{E}[g_w(x;\theta_2)] = g(x;\theta_2)$ and $\mathrm{E}[\tilde{\nabla}_x g_w(x;\theta_2)] = \tilde{\nabla}_x g(x;\theta_2) \in \partial g(x,\theta_2)$.

(ii) Given $y \in \mathbb{R}^m$, $\theta_1 \in \Theta_1$, the SO returns an unbiased sampled gradient $\nabla_y f_v(y;\theta_1) \in \mathbb{R}^m$ such that $\mathrm{E}[\nabla_y f_v(y;\theta_1)] = \nabla_y f(y;\theta_1)$.

(iii) Given $\theta_j \in \Theta_j$ for $j = 1, 2$, the SO returns an unbiased sampled gradient $\nabla \varphi^{(j)}_{\xi_j}(\theta_j) \in \mathbb{R}^{d_j}$ such that $\mathrm{E}[\nabla \varphi^{(j)}_{\xi_j}(\theta_j)] = \nabla \varphi^{(j)}(\theta_j)$.

For notational convenience, we write $\tilde{\nabla} g_w(x;\theta_2) \triangleq \tilde{\nabla}_x g_w(x;\theta_2)$ and $\nabla f_v(y;\theta_1) \triangleq \nabla_y f_v(y;\theta_1)$. The main challenge in solving the misspecified compositional optimization problem (1.1) is that we lack access to an unbiased estimator of the true gradient of $f \circ g$; this issue is exacerbated by the presence of misspecification in that the true parameters θ_1^* and θ_2^* are not known explicitly. To be specific, the lack of an accurate estimator of $\mathbb{E}_w[g_w(x_k;\theta_2^*)]$ leads to a bias induced in evaluating $\nabla f_{v_k}(\mathbb{E}_w[g_w(x_k;\theta_2^*)];\theta_1^*)$. Moreover, the lack of knowledge of θ_1^* and θ_2^* in a misspecified regime further impacts the bias of our estimator. To address the two aforementioned issues, we develop a class of efficient algorithms to resolve the misspecified compositional problem.

Next, we propose the misspecified stochastic compositional gradient descent (m-SCGD) algorithm. The algorithm iteratively updates the solution of x_k using past queried first-order information. After the first k iterations, we let y_k denote an estimator of $\mathbb{E}_w[g_w(x_k;\theta_2^*)]$, which is the expected value of g_w under the correctly specified model, and we let θ_1^k and θ_2^k denote the estimators of the model parameters. Based on the chain rule, we update the current solution x_k using a quasistochastic gradient step defined as follows, for any $k \geq 1$:

$$x_{k+1} = \Pi_X[x_k - \alpha_k \tilde{\nabla} g_{w_k^2}(x_k;\theta_2^k)\nabla f_{v_k}(y_k;\theta_1^k)].$$

Notably, here we employ the subgradient $\tilde{\nabla} g_{w_k^2}(x_k;\theta_2^k)$ in view of the possible nonsmoothness in the inner layer function $g(\cdot;\theta)$. Because y_k serves as an estimator of $\mathbb{E}_w[g_w(x_k;\theta_2^*)]$, we then query the SO at current solution x_{k+1} with an incorrect estimate of the model parameter θ_1^k and update y_k by constructing a weighted average between prior values and the new sample $g_{w_{k+1}^1}(x_{k+1};\theta_2^k)$ returned by the SO using β_k, as defined by the following update for $k \geq 1$:

$$y_{k+1} = (1-\beta_k)y_k + \beta_k g_{w_{k+1}^1}(x_{k+1};\theta_2^k).$$

Finally, we update our estimators θ_1^k and θ_2^k using projected stochastic gradient steps with steplengths $\gamma_{1,k}$ and $\gamma_{2,k}$, respectively, as specified next for $k \geq 1$:

$$\theta_1^{k+1} = \Pi_{\Theta_1}[\theta_1^k - \gamma_{1,k}\nabla \varphi^{(1)}_{\xi_{1,k+1}}(\theta_1^k)], \text{ and}$$
$$\theta_2^{k+1} = \Pi_{\Theta_2}[\theta_2^k - \gamma_{2,k}\nabla \varphi^{(2)}_{\xi_{2,k+1}}(\theta_2^k)].$$

Algorithm 1 (m-SCGD)

Input: $x_1 \in \mathbb{R}^n$, $y_1 \in \mathbb{R}^m$, $\theta_1^1 \in \mathbb{R}^{d_{\theta_1}}$, $\theta_2^1 \in \mathbb{R}^{d_{\theta_2}}$, SO, K, step-sizes $\{\alpha_k\}$, $\{\beta_k\}$, $\{\gamma_{1,k}\}$, $\{\gamma_{2,k}\}$.

Output: The sequence $\{x_k\}_{k=1}^K$.

for $k = 1, 2, \ldots, K$ **do**

Query the SO for the sample generalized gradient $\tilde{\nabla} g_{w_k^2}(x_k;\theta_2^k)$ and sample gradient $\nabla f_{v_k}(y_k;\theta_1^k)$. Update

$$x_{k+1} = \Pi_X[x_k - \alpha_k \tilde{\nabla} g_{w_k^2}(x_k;\theta_2^k)\nabla f_{v_k}(y_k;\theta_1^k)].$$

Query the SO for the sample value of g at (x_{k+1}, θ_2^k), obtain $g_{w_{k+1}^1}(x_{k+1};\theta_2^k)$. Update

$$y_{k+1} = (1-\beta_k)y_k + \beta_k g_{w_{k+1}^1}(x_{k+1};\theta_2^k).$$

Query the SO for the sample gradients $\nabla \varphi^{(1)}_{\xi_{1,k+1}}(\theta_1^k), \nabla \varphi^{(2)}_{\xi_{2,k+1}}(\theta_2^k)$. Update

$$\theta_1^{k+1} = \Pi_{\Theta_1}[\theta_1^k - \gamma_{1,k}\nabla \varphi^{(1)}_{\xi_{1,k+1}}(\theta_1^k)],$$
$$\theta_2^{k+1} = \Pi_{\Theta_2}[\theta_2^k - \gamma_{2,k}\nabla \varphi^{(2)}_{\xi_{2,k+1}}(\theta_2^k)].$$

end for

The details are summarized in Algorithm 1. Having presented Algorithm 1, we then conduct a theoretical analysis to investigate its performance. The key question is whether and how fast the generated solution sequence converges to a solution. Intuitively, with a sufficiently large number of observations, the estimators θ_1^k and θ_2^k would induce minuscule errors. Nevertheless, it remains unclear how the random errors induced under misspecification interact with those incurred from the estimation of $\mathbb{E}_w[g_w(x_k;\theta_2^*)]$ in evaluating the sampled gradient $\nabla f_{v_k}(\mathbb{E}_w[g_w(x_k;\theta_2^*)];\theta_1^*)$, especially in the case that $\mathbb{E}_w[g_w(x_k;\theta_2^*)]$ is corrupted by misspecification arising from the parameter sequence $\{\theta_2^k\}$. Moreover, for any solution trajectory generated by our algorithm, as the parameters are misspecified in each update, it is unclear how these errors accumulate and interact with each other. To further address these questions, we first impose some smoothness and moment assumptions on the stochastic component functions.

Assumption 2.1. *Let C_g, V_g, C_f, V_f, L_f, κ_f, and κ_g be positive scalars.*

(i) *The outer function $f(\cdot;\theta_1)$ is continuously differentiable for every $\theta_1 \in \Theta_1$, the inner function $g(\cdot;\theta_2)$ is continuous for every $\theta_2 \in \Theta_2$, and the feasible set X is closed and convex. There exists at least one optimal solution to problem (1.1).*

(ii) *For every $\theta_2 \in \Theta_2$, the function $g(\cdot;\theta_2)$ is Lipschitz continuous with parameter C_g, and the random variables $g_w(x;\theta_2)$ and $\tilde{\nabla} g_w(x;\theta_2)$ have bounded second moments such that*

$$\mathbb{E}[\|\tilde{\nabla}_x g_w(x;\theta_2)\|^2] \leq C_g, \text{ and } \mathbb{E}[\|g_w(x;\theta_2)\|^2] \leq V_g, \ \forall x \in X.$$

(iii) *For every $\theta_1 \in \Theta_1$, the function $f(\cdot;\theta_1)$ is L_f-smooth such that for all $\bar{y}, y \in \mathbb{R}^m$*

$$\|\nabla_y f(y;\theta_1) - \nabla f(\bar{y};\theta_1)\| \leq L_f\|y - \bar{y}\|.$$

In addition, the random variables $\nabla_y f_v(y;\theta_1)$ and $f_v(y;\theta_1)$ have bounded second moments such that

$$\mathbb{E}[\|\nabla_y f_v(y;\theta_1)\|^2] \leq C_f \text{ and } \mathbb{E}[\|f_v(y;\theta_1)\|^2] \leq V_f, \ \forall y \in \mathbb{R}^m.$$

(iv) *For every $x \in X$, the value mapping $g(x;\cdot)$ is C_g-Lipschitz continuous in θ_2 and $\nabla_x g(x;\cdot)$ is κ_g-Lipschitz continuous in θ_2. For every $y \in \mathbb{R}^m$, the gradient mapping $\nabla_y f(y;\cdot)$ is κ_f-Lipschitz continuous in θ_1.*

We also impose the following smoothness and moment assumptions on the model parameter functions.

Assumption 2.2. *Let C_{θ_1}, C_{θ_2}, μ_{θ_1}, and μ_{θ_2} be positive scalars.*

(i) *The model parameter function $\varphi_j^{(j)}(\theta_j)$ is μ_{θ_j}-strongly convex and the feasible set Θ_j is closed and convex for $j = 1, 2$.*

(ii) *For $j = 1, 2$, and for every $\theta_j \in \Theta_j$, the random variable $\nabla\varphi_{\xi_j}^{(j)}(\theta_j)$ has a bounded second moment such that $\mathbb{E}[\|\nabla\varphi_{\xi_j}^{(j)}(\theta_j)\|^2] \le C_{\theta_j}$.*

Apart from the various assumptions on functions f, g and model parameter functions $\varphi^{(1)}$, $\varphi^{(2)}$, as well as the associated moment requirements on the associated random gradients and Jacobians, in some part of our analysis, we focus on the scenario where the overall objective function $F(x;\theta_1^*,\theta_2^*) \triangleq f(g(x;\theta_1^*,\theta_2^*))$ is L_F-smooth, as captured by the next assumption.

Assumption 2.3. *The function $F(\cdot;\theta_1^*,\theta_2^*)$ is L_F-smooth, that is, there exists $L_F > 0$ such that*

$$F(\bar{x};\theta_1^*,\theta_2^*) - F(x;\theta_1^*,\theta_2^*) \le \nabla F(x;\theta_1^*,\theta_2^*)^\top(\bar{x}-x) + \frac{L_F}{2}\|x-\bar{x}\|^2, \quad \forall x,\bar{x}\in X.$$

Here we provide an illustrative example where the nonsmoothness of g under Assumption 1 would *not* impact the smoothness of F. Consider a case where the inner function g is nonsmooth and defined as $g(x) \triangleq \max\{0, x\}$ whereas the outer level function f is defined as $f(y) \triangleq y^2$ and is clearly smooth. One may see that the overall function $f \circ g$ defined as $f(g(x)) \triangleq (\max\{0,x\})^2$ is also smooth.

Note that such an example is by no means artificial and the $\max\{\cdot, 0\}$ operator has been adopted in many risk management applications for capturing downside risk. For instance, letting $U(\cdot,\omega)$ denote a random utility function, its associated semivariance (Rigamonti and Lučivjanská 2022) is defined as

$$\mathbb{E}_\omega[\max(\mathbb{E}_\omega[U(x,\omega)] - U(x,\omega), 0)^2],$$

which captures the dispersion of the random utility $U(\cdot,\omega)$ falling below the mean. Another example is risk-averse semideviation optimization (Ruszczyński and Shapiro 2006, Ahmed et al. 2007), which also employs the $\max\{\cdot, 0\}$ operator to quantify the downside risk (see the detailed formulation in Section 5).

2.1. Almost Sure Convergence

We prove the almost sure convergence of the m-SCGD algorithm in the following theorem and provide a detailed proof in e-companion Section A.2.

Theorem 2.1 (Almost Sure Convergence). *Suppose Assumptions 2.1 and 2.2 hold. Suppose X is a convex feasible region. Let $\{(x_k, y_k, \theta_1^k, \theta_2^k)\}_{k\ge 1}$ be the sequence computed via Algorithm 1. Let the step-sizes $\{\alpha_k\}$, $\{\beta_k\}$, $\{\gamma_{1,k}\}$, and $\{\gamma_{2,k}\}$ be such that $\sum_{k=1}^{\infty}\alpha_k = \infty$, $\sum_{k=1}^{\infty}\beta_k = \infty$, $\sum_{k=1}^{\infty}\gamma_{j,k} < \infty$ for $j \in \{1, 2\}$, $\beta_k = \gamma_{2,k}$ decreasing to zero monotonically, and*

$$\sum_{k=1}^{\infty}\left(\alpha_k^2 + \beta_k^2 + \gamma_{1,k}^2 + \gamma_{2,k}^2 + \frac{\alpha_k^2}{\gamma_{2,k}} + \frac{\alpha_k^2}{\gamma_{1,k}} + \frac{\alpha_k^2}{\beta_k}\right) < \infty.$$

(a) *If $F(\cdot;\theta_1^*,\theta_2^*)$ is convex and let X^* be the set of optimal solution to problem (1.1). Then $\{x_k\}$ converges almost surely to a random point in X^*.*

(b) *Suppose Assumption 2.3 holds and $X = \mathbb{R}^n$. Then any limit point of the sequence $\{x_k\}$ is a stationary point of problem (1.1) almost surely.*

2.2. Convergence Rates

Theorem 2.1 in Section 2.1 establishes the almost sure convergence guarantees for the m-SCGD algorithm for both convex and nonconvex objectives. We now investigate the rate of convergence when we have diminishing step-size sequences of the form

$$\alpha_k = C_a k^{-a},\ \beta_k = 5k^{-b},\ \gamma_{j,k} = 2/(\mu_{\theta_j} k^c) \text{ for } j = 1, 2,$$

where $a, b, c \in (0, 1]$ are positive real numbers and $C_a > 0$ is a constant. Optimizing the rate over a, b, and c, we obtain the following result for nonconvex objectives. Notably, we distinguish the impact of misspecification on the rate.

Theorem 2.2 (Convergence Rate for Nonconvex Objectives). *Suppose Assumptions 2.1, 2.2, and 2.3 hold, and $X = \mathbb{R}^n$. Let the step-sizes be $\alpha_k = k^{-0.5}$, $\beta_k = 5k^{-0.5}$, and $\gamma_{j,k} = 2/(\mu_{\theta_j} k^c)$ for $j = 1, 2$. Let $\{(x_k, y_k, \theta_1^k, \theta_2^k)\}_{k\ge 1}$ be the sequence generated by Algorithm 1.*

(a) *For any $K > 0$, we have*

$$\frac{\sum_{k=1}^{K}\alpha_k \mathbb{E}[\|\nabla F(x_k;\theta_1^*,\theta_2^*)\|^2]}{\sum_{k=1}^{K}\alpha_k} \le O\left(\frac{\ln K}{K^{0.5}}\right) + O\left(\frac{\sum_{k=1}^{K} k^{0.5-c}}{K^{0.5}}\right).$$

(b) *Suppose $c \ge 0.5$. Then for any $K > 0$, we have*

$$\frac{\sum_{k=1}^{K}\alpha_k \mathbb{E}[\|\nabla F(x_k;\theta_1^*,\theta_2^*)\|^2]}{\sum_{k=1}^{K}\alpha_k} \le O\left(\frac{\ln K}{K^{0.5}}\right).$$

The detailed proof is provided in e-companion Section A.3.

Remark 2.1. The previous result shows that our scheme achieves a rate of $\tilde{O}(K^{-0.5})$ for nonconvex objectives, matching the $\tilde{O}(K^{-0.5})$ convergence rate for two-level basic-SCGD derived in Yang et al. (2019),

despite the misspecification of the parameters θ_1^* and θ_2^*. In conclusion, the corruption via misspecification in θ_1 and θ_2 does not adversely affect the rate of convergence under the caveat that the steplength sequence $\gamma_{i,k}$ (where $i = 1, 2$) employed for learning θ_1^* and θ_2^* diminishes suitably fast. However, the constant in the rate is indeed affected by the presence of misspecification.

3. Accelerated Schemes

In Section 2, we establish an $\tilde{\mathcal{O}}(K^{-2/3})$ rate of convergence for the misspecified stochastic composition optimization problem with a nonsmooth inner function $g(\bullet;\theta_2)$. In this section, by assuming $g(\bullet;\theta_2)$ to be smooth, we propose an accelerated gradient descent method when updating the estimator y_k. We prove the almost sure convergence of the sequence produced by the algorithm and show that the algorithm achieves a faster rate for nonconvex and strongly convex objectives.

Recall that in Section 2, at each iteration, we update our estimator y_k by the weighted average of the previous estimator and the new sample $g_{w_{k+1}^1}(x_{k+1};\theta_2^k)$ returned by SO. To accelerate the sequence, a fundamental question is how to obtain a better estimator. When the information parameter θ_2^* is perfectly known, Chen et al. (2021b) and Balasubramanian et al. (2022) proposed accelerated schemes that achieve faster rates under some mild assumptions by employing an additional linearization term. We adopt this idea in developing an accelerated method when θ_2^* is unavailable as a priori and establish convergence results. Before proceeding, in addition to Assumptions 2.1 and 2.2, we impose the following assumptions.

Assumption 3.1. *Let L_g, C_g, and C_f be positive scalars. We assume the following.*

(i) *The function $g(\bullet;\theta_2)$ is L_g-smooth for any $\theta_2 \in \Theta_2$, that is, there exists a positive scalar L_g such that*

$$\|\nabla_x g(x;\theta_2) - \nabla_x g(\bar{x};\theta_2)\| \le L_g\|x-\bar{x}\|, \quad \forall x,\bar{x}\in X,\ \theta_2\in\Theta_2.$$

(ii) *The random variables $\nabla g_w(x;\theta_2)$ and $\nabla f_v(y;\theta_1)$ have bounded fourth moments such that*

$$\mathbb{E}[\|\nabla g_w(x;\theta_2)\|^4] \le C_g^2 \text{ and } \mathbb{E}[\|\nabla f_v(y;\theta_1)\|^4] \le C_f^2,$$
$$\forall x\in X,\ y\in\mathbb{R}^n,\ \theta_1\in\Theta_1,\ \theta_2\in\Theta_2.$$

We point out that a vast class of distributions including Gaussian, uniform, and logistic, among others, are characterized by bounded fourth moments. Furthermore, this assumption is commonly employed to achieve acceleration in stochastic compositional optimization (Wang et al. 2017a, b; Yang et al. 2019). In the remainder of this section, we propose an accelerated algorithm that achieves a faster convergence rate. At the k-th iteration, we first update x_{k+1} by employing a projected gradient step and invoking the chain rule under the misspecified model that

$$x_{k+1} = \Pi_X[x_k - \alpha_k \nabla g_{w_k^1}(x_k;\theta_2^k)\nabla f_{v_k}(y_k;\theta_1^k)].$$

To achieve acceleration, we query the SO twice at x_k with misspecified parameter θ_2^k to obtain two independent samples of $\nabla g_{w_k^2}(x_k;\theta_2^k)$ and $g_{w_k^3}(x_k;\theta_2^k)$. This facilitates building an update y_{k+1} as a weighted average of the prior value, namely, y_k, and the independently sampled value $g_{w_k^3}(x_k;\theta_2^k)$ added to a conditionally independent stochastic linearization term given by $\nabla g_{w_k^2}(x_k;\theta_2^k)^\top(x_{k+1}-x_k)$, leading to the following update rule that

$$y_{k+1} = (1-\beta_k)y_k + \beta_k g_{w_k^3}(x_k;\theta_2^k) + \nabla g_{w_k^2}(x_k;\theta_2^k)^\top(x_{k+1}-x_k). \tag{3.1}$$

Finally, similar to the m-SCGD algorithm, we update the parameters θ_1^k, θ_2^k by the projected stochastic gradient method, defined as follows for $k \ge 1$:

$$\theta_1^{k+1} = \Pi_{\Theta_1}[\theta_1^k - \gamma_{1,k}\nabla\varphi^{(1)}_{\xi_{1,k+1}}(\theta_1^k)], \quad \text{and}$$
$$\theta_2^{k+1} = \Pi_{\Theta_2}[\theta_2^k - \gamma_{2,k}\nabla\varphi^{(2)}_{\xi_{2,k+1}}(\theta_2^k)].$$

The details of this misspecified accelerated stochastic compositional gradient descent (m-aSCGD) algorithm are summarized in Algorithm 2.

Algorithm 2 (m-aSCGD)

Input: $x_1\in\mathbb{R}^n$, $y_1\in\mathbb{R}^n$, $\theta_1^1\in\mathbb{R}^{d_{\theta_1}}$, $\theta_2^1\in\mathbb{R}^{d_{\theta_2}}$, SO, K, step-sizes $\{\alpha_k\}$, $\{\beta_k\}$, $\{\gamma_{1,k}\}$, $\{\gamma_{2,k}\}$.

Output: The sequence $\{x_k\}_{k=1}^K$.

for $k = 1, 2, \ldots, K$ **do**

Query the SO to obtain a sample gradient of g at x_k and θ_2^k, given by $\nabla g_{w_k^1}(x_k,\theta_2^k)$; query the SO to obtain a sample gradient of f at y_k and θ_1^k, denoted by $\nabla f_{v_k}(y_k,\theta_1^k)$. Update x_{k+1} by

$$x_{k+1} = \Pi_X[x_k - \alpha_k \nabla g_{w_k^1}(x_k;\theta_2^k)\nabla f_{v_k}(y_k;\theta_1^k)].$$

Query the SO to obtain a sample gradient $\nabla g_{w_k^2}(x_k;\theta_2^k)$ and a sample function value $g_{w_k^3}(x_k;\theta_2^k)$ at x_k and θ_2^k. Update y_{k+1} by

$$y_{k+1} = (1-\beta_k)y_k + \beta_k g_{w_k^3}(x_k;\theta_2^k) + \nabla g_{w_k^2}(x_k;\theta_2^k)^\top(x_{k+1}-x_k).$$

end for

3.1. Almost Sure Convergence

We first prove the almost sure convergence of the iterates generated by the accelerated algorithm. First, we note that when both layers f and g have Lipschitz continuous gradients with Lipschitz constants $L_f, L_g > 0$, respectively, the objective function F also has Lipschitz continuous gradients. Thus, Assumption 2.3 holds if both Assumptions 2.1 (iii) and 3.1 (i) hold. We denote the Lipschitz constant as $L_F > 0$. Theorem 3.1 shows the

almost sure convergence of the m-aSCGD algorithm, where we provide the detailed proof in e-companion Section B.1.

Theorem 3.1 (Almost Sure Convergence of m-aSCGD). *Suppose that Assumptions 2.1, 2.2, and 3.1 hold. Let* $\{(x_k, y_k, \theta_1^k, \theta_2^k)\}_{k=1}^K$ *be the sequence generated by Algorithm 2, and let the step-sizes* $\{\alpha_k\}$, $\{\beta_k\}$, $\{\gamma_{1,k}\}$, *and* $\{\gamma_{2,k}\}$ *be such that* $\sum_{k=1}^{\infty} \alpha_k = \infty$, $\sum_{k=1}^{\infty} \beta_k = \infty$, $\sum_{k=1}^{\infty} \gamma_{j,k} = \infty$ $\forall j = 1, 2$, $\alpha_k \le M\beta_k$ *for all* $k \ge 0$ *for some* $M > 0$, $\beta_k = \gamma_{2,k}$ *decreasing to zero monotonically, and*

$$\sum_{k=1}^{\infty} \left(\alpha_k^2 + \beta_k^2 + \gamma_{1,k}^2 + \gamma_{2,k}^2 + \frac{\alpha_k^2}{\gamma_{1,k}} + \frac{\alpha_k^2}{\gamma_{2,k}} + \frac{\alpha_k^2}{\beta_k}\right) < \infty.$$

(a) *Suppose that* $(F, \theta_1^*, \theta_2^*)$ *is convex, and* X^* *denotes the set of optimal solutions to problem (1.1). Then* $\{x_k\}$ *converges almost surely to a point in* X^*.

(b) *Suppose* $X = \mathbb{R}^n$. *Then any limit point of the sequence* $\{x_k\}_{k=0}^{\infty}$ *is a stationary point of (1.1) almost surely.*

3.2. Convergence Rates

We now derive the rates of convergence of the m-aSCGD algorithm under nonconvex and strongly convex settings in the following theorem, with diminishing step-size sequences taking the form

$$\alpha_k = C_a k^{-a}, \quad \beta_k = 5k^{-b}, \quad \text{and} \quad \gamma_{j,k} = 2/(\mu_{\theta_j} k^c) \quad \text{for } j = 1, 2,$$

where $a, b, c \in (0, 1]$ are positive real numbers and $C_a > 0$ is a constant.

Theorem 3.2 (Convergence Rate of m-aSCGD for Nonconvex Objectives). *Suppose that Assumptions 2.1, 2.2, and 3.1 hold, and* $X = \mathbb{R}^n$. *Let the step-sizes be* $\alpha_k = k^{-1/2}$, $\beta_k = 5k^{-1/2}$, *and* $\gamma_{j,k} = 2/(\mu_{\theta_j} k^c)$ *for* $j = 1, 2$. *Let* $\{(x_k, y_k, \theta_1^k, \theta_2^k)\}_{k=1}^K$ *be the sequence generated by Algorithm 2.*

(a) *For any* $K > 0$, *we have*

$$\frac{\sum_{k=1}^K \alpha_k \mathbb{E}[\|\nabla F(x_k; \theta_1^*, \theta_2^*)\|^2]}{\sum_{k=1}^K \alpha_k} \le O\left(\frac{\ln K}{\sqrt{K}}\right) \quad \text{if } c = 2a; \qquad O\left(\frac{1}{\sqrt{K}}\right) \quad \text{if } c > 2a.$$

(b) *Suppose* $c \ge 1/2$. *Then for any* $K > 0$, *we have*

$$\frac{\sum_{k=1}^K \alpha_k \mathbb{E}[\|\nabla F(x_k; \theta_1^*, \theta_2^*)\|^2]}{\sum_{k=1}^K \alpha_k} \le O\left(\frac{\ln K}{\sqrt{K}}\right).$$

The detailed proof is provided in e-companion Section B.2.

We conclude this section by considering the case where $F(\cdot\,; \theta_1^*, \theta_2^*)$ is σ-strongly convex. In particular, we say that a function $F(\cdot\,; \theta_1^*, \theta_2^*)$ is σ-strongly convex if there exists a constant $\sigma > 0$ such that for all $x \in X$,

$$F(x; \theta_1^*, \theta_2^*) - F(x^*; \theta_1^*, \theta_2^*) \ge \sigma \|x - x^*\|^2. \tag{3.2}$$

Then, letting the step-sizes α_k, β_k, and γ_k be defined as

$$\alpha_k = \frac{2}{\sigma(k+1)}, \quad \beta_k = 5k^{-1}, \quad \text{and} \quad \gamma_{j,k} = \frac{2}{\mu_{\theta_j} k} \quad \text{for } j = 1, 2,$$

our next theorem shows that Algorithm 2 achieves a faster rate under strong convexity.

Theorem 3.3 (Convergence Rate of m-aSCGD for Strongly Convex Objectives). *Suppose Assumptions 2.1, 2.2, and 3.1 hold, the objective function F is σ-strongly convex satisfying (3.2), and X is convex. Let* $\{(x_k, y_k, \theta_1^k, \theta_2^k)\}_{k=1}^K$ *be the sequence generated by Algorithm 2, where* $\alpha_k = 2/(\sigma(k+1))$, $\beta_k = 5k^{-1}$, *and* $\gamma_{j,k} = 2k^c/\mu_{\theta_j}$ *for* $j = 1, 2$.

(a) *For any* $K > 0$, *we have*

$$\mathbb{E}[\|x_K - x^*\|^2] \le O(K^{-1}) + \frac{2}{K(K+1)} \sum_{k=1}^K O(k^{1-c}).$$

(b) *Suppose* $c \ge 1$. *Then for any* $K > 0$, *we have*

$$\mathbb{E}[\|x_K - x^*\|^2] \le O(K^{-1}).$$

We defer the detailed proof to e-companion Section B.3.

Remark 3.1. Theorem 3.2 shows that our accelerated scheme achieves a faster rate of convergence with an additional smoothness assumption. Meanwhile, our result improves the $O(K^{-1/3})$ convergence rate obtained for nonconvex two-level *a*-SCGD developed by Yang et al. (2019) by incorporating the linearization term in the update for y_{k+1}. In fact, this scheme achieves a near-optimal rate and misspecification leads to no degradation from the optimal rate compared with that obtained with perfect knowledge of θ_1^* and θ_2^*. Moreover, the rate statement in Theorem 3.3 shows that the accelerated scheme achieves the optimal rate, improving on the rate provided in Yang et al. (2019) and displaying a corresponding lack of degradation from the rate obtained in perfectly specified strongly convex regimes.

Furthermore, it is worth pointing out that our misspecified optimization framework is able to accommodate other real-world applications where the misspecified sequence $\{(\theta_1^k, \theta_2^k)\}$ is generated through other approaches instead of SGD. In fact, based on Lemma A.11 in e-companion Section A.3, we realize that the overall convergence rate exhibits no degradation if the misspecification error $\|\theta_j^k - \theta_j^*\|^2$ diminishes to zero at a rate no slower than at which $\|y_k - g(x_k; \theta_2^*)\|^2$ decays to zero for both nonconvex and strongly convex problems when the environmental parameters $\{(\theta_1^*, \theta_2^*)\}$ are explicitly known. A slower rate will result in a build-up of error, and beyond a certain point, convergence cannot be guaranteed. The utilization of SGD in learning $\{(\theta_1^*, \theta_2^*)\}$ is to show that such a rate is indeed obtainable.

4. Multilevel Misspecified Compositional Optimization

In Sections 2 and 3, we have considered the two-level compositional problem afflicted by misspecification. As discussed in Section 1, the general T-level misspecified compositional problem assumes immense relevance. Recall that the T-level problem is of the form given by

$$\min_{x\in X} F(x;\theta_1^*,\ldots,\theta_T^*)$$

$$\equiv \mathbb{E}_{\omega_1}[f_{\omega_1}^{(1)}(\mathbb{E}_{\omega_2}[f_{\omega_2}^{(2)}(\cdots(\mathbb{E}_{\omega_T}[f_{\omega_T}^{(T)}(x;\theta_T^*)]);\cdots;\theta_2^*)];\theta_1^*)],$$

$$\text{where } \theta_j^* \in \arg\min_{\theta_j\in\Theta_j} \mathbb{E}_{\xi_j}[\varphi_{\xi_j}^{(j)}(\theta_j)], \text{ for } j=1,2,\ldots,T. \tag{4.1}$$

Akin to the two-level case, throughout this section, we assume the model parameter function $\varphi^{(j)}$ is strongly convex over a convex feasible region Θ_j, admitting a unique optimal solution θ_j^*, for $j=1,2,\ldots,T$. We focus on the scenario where the random variables $\omega_1,\ldots,\omega_T$ are independent of each other, and denote by $f^{(j)}(y_j) \equiv \mathbb{E}_{\omega_j}[f_{\omega_j}^{(j)}(y_j)]$ for $j=1,2,\ldots,T$. We assume access to the following zeroth and first-order stochastic oracle (SO):

(Sampling Oracle (SO) for T-level problem.)

- Given $x\in X\subset \mathbb{R}^{d_1}$, $\theta_T\in\Theta_T$, the SO returns an unbiased sampled function value $f_{\omega_T}^{(T)}(x;\theta_T)\in\mathbb{R}^{d_{T+1}}$ and an unbiased stochastic gradient $\nabla_x f_{\omega_T}^{(T)}(x;\theta_T)\in\mathbb{R}^{d_1\times d_{T+1}}$ such that $\mathbb{E}[f_{\omega_T}^{(T)}(x;\theta_T)]=f^{(T)}(x;\theta_T)$ and $\mathbb{E}[\nabla_x f_{\omega_T}^{(T)}(x;\theta_T)]=\nabla f^{(T)}(x;\theta_T)$.
- Given $y_j\in\mathbb{R}^{d_j}$, $\theta_j\in\Theta_j$, the SO returns an unbiased sampled function value $f_{\omega_j}^{(j)}(y_j;\theta_j)\in\mathbb{R}^{d_{j+1}}$ and an unbiased stochastic gradient $\nabla_{y_j} f_{\omega_j}^{(j)}(y_j;\theta_j)\in\mathbb{R}^{d_j\times d_{j+1}}$ for each layer $j=1,\ldots,T-1$ such that $\mathbb{E}[f_{\omega_j}^{(j)}(y_j;\theta_j)]=f^{(j)}(y_j;\theta_j)$ and $\mathbb{E}[\nabla_{y_j} f_{\omega_j}^{(j)}(y_j;\theta_j)]=\nabla f^{(j)}(y_j;\theta_j)$.
- Given $\theta_j\in\Theta_j$ for $j=1,2,\ldots,T$, the SO returns an unbiased stochastic gradient $\nabla\varphi_{\xi_j}^{(j)}(\theta_j)\in\mathbb{R}^{d_j}$ such that $\mathbb{E}[\nabla\varphi_{\xi_j}^{(j)}(\theta_j)]=\nabla\varphi^{(j)}(\theta_j)$.

For notational simplicity, we write $\nabla f_{\omega_T}^{(T)}(x;\theta_T)\equiv\nabla_x f_{\omega_T}^{(T)}(x;\theta_T)$ and $\nabla f_{\omega_j}^{(j)}(y_j;\theta_j)\equiv\nabla_{y_j} f_{\omega_j}^{(j)}(y_j;\theta_j)$. Akin to the two-level scenario, the main challenge of solving the general T-level misspecified compositional optimization problem lies in contending with the unavailability of an unbiased sample gradient. To illustrate this, under the T-stage setting with misspecified parameters $(\theta_1,\theta_2,\ldots,\theta_T)$, using the chain rule, an unbiased sampled gradient is given by

$$\nabla f_{\omega_T}^{(T)}(x;\theta_T)\nabla f_{\omega_{T-1}}^{(T-1)}(f^{(T)}(x;\theta_T);\theta_{T-1})\cdots\nabla f_{\omega_1}^{(1)}(f^{(2)}\circ\cdots\circ f^{(T)}(x;\theta_2,\ldots,\theta_T);\theta_1).$$

One can see that this unbiased sample gradient requires access to the correct parameters $(\theta_1^*,\theta_2^*,\ldots,\theta_T^*)$ and the values $f^{(T)}(x;\theta_T^*),\ldots,f^{(2)}\circ\cdots\circ f^{(T)}(x;\theta_2^*,\ldots,\theta_T^*)$. Unfortunately, the misspecification in model parameters and errors induced when estimating the functions' values result in a bias when we adopt a plugin estimator. Letting $y^{(j)}$ be an estimator for $f^{(j)}\circ\cdots\circ f^{(T)}(x;\theta_j^*,\ldots,\theta_T^*)$, when the model parameters are perfectly known, Yang et al. (2019) developed a multitimescale approach that recursively updates $y^{(j)}$. In this case, we compute a sampled gradient by

$$\nabla f_{\omega_T}^{(T)}(x;\theta_T^k)\nabla f_{\omega_{T-1}}^{(T-1)}(y^{(T-1)};\theta_{T-1}^k)\cdots\nabla f_{\omega_1}^{(1)}(y^{(1)};\theta_1^k),$$

whose bias diminishes to zero as $k\to\infty$.

Nevertheless, without prior knowledge of the model parameters, the estimators $y^{(j)}$s are also corrupted by misspecification, which makes it particularly challenging to apply gradient descent approaches. Recall that in Sections 2 and 3, for the two-level misspecified compositional problem (1.1), in iteration k, we update our estimator of the inner function value $g(x_k;\theta_2^*)$ by combining an inertial update with an accelerated linearization scheme, under a misspecified parameter θ_2^k, and establish the almost sure convergence and rates of convergence. Under the general T-level scenario, the misspecification in model parameters and inaccurately estimated values of $f^{(T)}(x;\theta_T^*),\ldots,f^{(2)}\circ\cdots\circ f^{(T)}(x;\theta_2^*,\ldots,\theta_T^*)$ further exacerbates the bias of the computed sample gradient, because of the more complicated multilevel nested structure. Therefore, it remains an open problem whether and how efficient algorithms can be developed to tackle the general multilevel misspecified compositional optimization problem (4.1).

To address this issue, throughout this section, we assume the satisfaction of the following smoothness requirements and boundedness conditions on fourth moments.

Assumption 4.1. *Let $\{C_{f_j}\}_{j=1}^T$, $\{V_{f_j}\}_{j=1}^T$, $\{L_{\theta_j}\}_{j=1}^T$, and $\{\kappa_{f_j}\}_{j=1}^T$ be positive scalars.*

(i) *For $j=1,\ldots,T$, the function $f^{(j)}(\bullet;\theta_j)$ is continuously differentiable for every $\theta_j\in\Theta_j$. The feasible set X is closed and convex, and there exists at least one optimal solution to problem (4.1).*

(ii) *For every $j=1,\ldots,T$, the random variables $f_{\omega_j}^{(j)}(y_j;\theta_j)$ and $\nabla f_{\omega_j}^{(j)}(y_j;\theta_j)$ have bounded fourth moments such that*

$$\mathbb{E}[\|f_{\omega_j}^{(j)}(y_j;\theta_j)\|^4]\le V_{f_j}^2 \text{ and } \mathbb{E}[\|\nabla f_{\omega_j}^{(j)}(y_j;\theta_j)\|^4]\le C_{f_j}^2, \quad \forall y_j\in\mathbb{R}^{d_j},\ \forall\theta_j\in\Theta_j.$$

(iii) *For every $j=1,\ldots,T$ and for every $\theta_j\in\Theta_j$, the function $f^{(j)}(\bullet,\theta_j)$ is L_{f_j}-smooth such that*

$$\|\nabla f^{(j)}(y_j;\theta_j)-\nabla f^{(j)}(\bar{y}_j;\theta_j)\|\le L_{f_j}\|y_j-\bar{y}_j\|,\quad \forall y_j,\bar{y}_j\in\mathbb{R}^{d_j}.$$

(iv) *For all $x\in X$, the value mapping $f^{(T)}(x;\bullet)$ is C_{f_T}-Lipschitz continuous in θ and the gradient mapping $\nabla f^{(T)}(x;\bullet)$ is κ_{f_T}-Lipschitz continuous in θ. For $j=2,$*

… , T−1, for all $y \in \mathbb{R}^{d_j}$, the value mapping $f^{(j)}(y_j;\cdot)$ is C_{f_j}-Lipschitz continuous in θ_j, and the gradient mapping $\nabla f^{(j)}(y_j;\cdot)$ is κ_{f_j}-Lipschitz continuous in θ_j. For all $y \in \mathbb{R}^{d_1}$, the gradient mapping $\nabla f^{(1)}(y_1;\cdot)$ is κ_{f_1}-Lipschitz continuous in θ_1.

Note that the L_j-Lipschitz continuous gradient for every layer function $f^{(j)}$ in Assumption 4.1 (iii) implies Assumption 2.3, that is, F has Lipschitz continuous gradient with parameter $L_F > 0$.

Contrary to the two-level scenario where only bounded second moments are required to ensure the efficiency of algorithms, as we have mentioned earlier, the misspecification indeed exacerbates the bias under the multilevel nested structure, especially under the accelerated scheme. For the purpose of acceleration, we also impose the following mild assumptions on the smoothness and fourth-order moments for the model parameter functions.

Assumption 4.2. *Let $\{C_{\theta_j}\}_{j=1}^T$ and $\{\mu_{\theta_j}\}_{j=1}^T$ be positive scalars.*

(i) *For $j = 1, \ldots, T$, the model parameter function $\varphi^{(j)}$ is μ_{θ_j}-strongly convex, and the feasible set Θ_j is closed and convex.*

(ii) *For $j = 1, \ldots, T$ and all $\theta \in \Theta_j$, the random variable $\nabla_{\theta_j}\varphi^{(j)}_{\xi_j}(\theta_j)$ has a bounded fourth moment such that $\mathbb{E}[\|\nabla_{\theta_j}\varphi^{(j)}_{\xi_j}(\theta_j)\|^4] \le C^2_{\theta_j}$.*

By the smoothness and the boundedness of the fourth moment of each layer $f^{(j)}$, we adopt the accelerated linearization scheme in Section 3 to develop a new algorithm. Our algorithm iteratively updates the solution x_k and the unknown model parameters θ_j^k's. In particular, at iteration k, we first update the solution x_{k+1} by a gradient step using the chain rule with misspecified model parameters $\theta_1^k, \ldots, \theta_T^k$ and estimators $y_k^{(1)}, \ldots, y_k^{(T-1)}$ via the following for $k \ge 1$.

$$x_{k+1} = \Pi_X\Big[x_k - \alpha_k \nabla f^{(T)}_{\omega^1_{T,k}}(x_k;\theta_T^k)\nabla f^{(T-1)}_{\omega^1_{T-1,k}}(y_k^{(T-1)};\theta_{T-1}^k)\cdots \nabla f^{(1)}_{\omega^1_{1,k}}(y_k^{(1)};\theta_1^k)\Big].$$

After updating x_{k+1}, we employ the accelerated linearization scheme to derive an estimator for $f^{(T)}(x_{k+1};\theta_T^*)$. Similar to the two-level scenario discussed earlier, using the misspecified parameter θ_k^T for this layer, we query the SO at x_k to obtain independent samples of the gradient $\nabla f^{(T)}_{\omega^2_{T,k}}(x_k;\theta_T^k)$ and the function value $f^{(T)}_{\omega^3_{T,k}}(x_k;\theta_T^k)$, and update $y_{k+1}^{(T-1)}$ by a weighted average of $f^{(T)}_{\omega^3_{T,k}}(x_k;\theta_T^k)$ and previous estimator $y_k^{(T-1)}$, plus an additional linearization term $\nabla f^{(T)}_{\omega^2_{T,k}}(x_k;\theta_T^k)^\top(x_{k+1}-x_k)$ such that

$$y_{k+1}^{(T-1)} = (1-\beta_k)y_k^{(T-1)} + \beta_k f^{(T)}_{\omega^3_{T,k}}(x_k;\theta_T^k) + \nabla f^{(T)}_{\omega^2_{T,k}}(x_k;\theta_T^k)^\top(x_{k+1}-x_k).$$

For the remainder of the layers, we recursively apply a similar inertial update rule based on obtaining independent samples of the gradient $\nabla f^{(j+1)}_{\omega^2_{j+1,k}}(y_k^{(j+1)};\theta_{j+1}^k)$ and function value $f^{(j+1)}_{\omega_{j,k+1}}(y_k^{(j+1)};\theta_{j+1}^k)$ with misspecified parameter θ_{j+1}^k at $y_k^{(j+1)}$, and update $y_{k+1}^{(j)}$ by

$$y_{k+1}^{(j)} = (1-\beta_k)y_k^{(j)} + \beta_k f^{(j+1)}_{\omega^3_{j+1,k}}(y_k^{(j+1)};\theta_{j+1}^k) + \nabla f^{(j+1)}_{\omega^2_{j+1,k}}(y_k^{(j+1)};\theta_{j+1}^k)^\top(y_{k+1}^{(j+1)} - y_k^{(j+1)})$$

for $j = T-2, \ldots, 1$. Finally, we update the model parameters $\theta_1^k, \theta_2^k, \ldots, \theta_T^k$ by basic SGD, as prescribed by the following update rule:

$$\theta_j^{k+1} = \Pi_{\Theta_j}[\theta_j^k - \gamma_{j,k}\nabla\varphi^{(j)}_{\xi_{j,k+1}}(\theta_j^k)], \text{ for } j = 1, 2, \ldots, T.$$

The above procedure completes the essential update steps within each iteration, which simultaneously computes the main solution x_k and update model parameters $\{\theta_j^k\}_{j=1}^T$. We provide the details of this misspecified T-level stochastic compositional gradient descent (m-TSCGD) algorithm in Algorithm 3.

Algorithm 3 (m-TSCGD)

Input: $x_1 \in \mathbb{R}^{d_T}$, $y_1^j \in \mathbb{R}^{d_j}$ for $j = T-1, \ldots, 1$, $\theta_j^1 \in \mathbb{R}^{d_{\theta_j}}$ for $j = 1, 2, \ldots, T$, SO, K, step-sizes $\{\alpha_k\}_{k=0}^K$, $\{\beta_k\}_{k=0}^K$, $\{\gamma_{j,k}\}_{k=0}^K$ for $j = 1, \ldots, T$.

Output: The sequence $\{x_k\}_{k=1}^K$.

for $k = 1, 2, \ldots, K$ **do**

Query the SO to obtain sample gradients $\nabla f^{(T)}_{\omega^1_{T,k}}(x_k;\theta_T^k)$ and $\nabla f^{(j)}_{\omega^1_{j,k}}(y_k^{(j)};\theta_j^k)$ for $j = T-1, \ldots, 1$.

Update

$$x_{k+1} = \Pi_X\Big[x_k - \alpha_k \nabla f^{(T)}_{\omega^1_{T,k}}(x_k;\theta_T^k)\nabla f^{(T-1)}_{\omega^1_{T-1,k}}(y_k^{(T-1)};\theta_{T-1}^k)\cdots \nabla f^{(1)}_{\omega^1_{1,k}}(y_k^{(1)};\theta_1^k)\Big].$$

Query the SO for $f^{(T)}_{\omega^3_{T,k}}(x_k;\theta_T^k)$ and $\nabla f^{(T)}_{\omega^2_{T,k}}(x_k;\theta_T^k)$. Update $y_{k+1}^{(T-1)}$ by

$$y_{k+1}^{(T-1)} = (1-\beta_k)y_k^{(T-1)} + \beta_k f^{(T)}_{\omega^3_{T,k}}(x_k;\theta_T^k) + \nabla f^{(T)}_{\omega^2_{T,k}}(x_k;\theta_T^k)^\top(x_{k+1}-x_k).$$

for $j = T-2, \ldots, 1$ **do**

Query the SO for $f^{(j+1)}_{\omega^3_{j+1,k+1}}(y_k^{(j+1)};\theta_{j+1}^k)$ and $\nabla f^{(j+1)}_{\omega^2_{j+1,k}}(y_k^{(j+1)};\theta_{j+1}^k)$. Update $y_{k+1}^{(j)}$ by

$$y_{k+1}^{(j)} = (1-\beta_k)y_k^{(j)} + \beta_k f^{(j+1)}_{\omega^3_{j+1,k}}(y_k^{(j+1)};\theta_{j+1}^k) + \nabla f^{(j+1)}_{\omega^2_{j+1,k}}(y_k^{(j+1)};\theta_{j+1}^k)^\top(y_{k+1}^{(j+1)} - y_k^{(j+1)}).$$

end for

for $j = 1, 2, \ldots, T$ **do**

Query the SO for the sample gradient $\nabla\varphi^{(j)}_{\xi_{j,k+1}}(\theta_j^k)$.

Update model parameters by

$$\theta_j^{k+1} = \Pi_{\Theta_j}[\theta_j^k - \gamma_{j,k}\nabla\varphi^{(j)}_{\xi_{j,k}}(\theta_j^k)].$$

endfor

end for

4.1. Convergence Rates

Recall that in Section 3.2, we establish the convergence rates of the m-aSCGD algorithm for two-level misspecified stochastic compositional problems with nonconvex and strongly convex objectives. Despite the misspecification, our rates match those provided by Chen et al. (2021b) and Balasubramanian et al. (2022) with explicit knowledge of model parameters. Under the T-level setting, with misspecification, it remains unclear if indeed Algorithm 3 produces a convergent sequence and if so, how fast this sequence converges. Furthermore, the resulting degradation in the convergence rates remains an open question. To address these issues, we derive the convergence rates of Algorithm 3 for nonconvex and strongly convex objectives and assess the impact of misspecification. We employ the following diminishing step-size sequences:

$$\alpha_k = C_a k^{-a},\ \beta_k = 5k^{-b},\ \gamma_{j,k} = 2/(\mu_{\theta_j} k^c) \text{ for } j = 1, 2, \ldots, T,$$

where $a, b, c \in (0, 1]$ are positive real numbers and $C_a > 0$ is a constant.

Now, we present the convergence guarantees of Algorithm 3 for nonconvex objectives.

Theorem 4.1 (Convergence Rate of m-TSCGD for Nonconvex Objectives). *Suppose Assumptions 4.1 and 4.2 hold, and $X = \mathbb{R}^{d_T}$. Let the step-sizes be $\alpha_k = k^{-1/2}$, $\beta_k = 5k^{-1/2}$, and $\gamma_{j,k} = 2k^{-c}/\mu_{\theta_j}$ for $j = 1, 2, \ldots, T$. Let $\{(x_k, y_k^{(T-1)}, \ldots, y_k^{(1)}, \theta_1^k, \ldots, \theta_T^k)\}_{k=1}^K$ be the sequence generated by Algorithm 3.*

(a) *For any $K > 0$, we have*

$$\frac{\sum_{k=1}^K \alpha_k \mathbb{E}[\|\nabla F(x_k; \theta_1^*, \ldots, \theta_T^*)\|^2]}{\sum_{k=1}^K \alpha_k} \le O\left(\frac{\ln K}{\sqrt{K}}\right) O\left(\frac{\sum_{k=1}^K k^{1/2-c}}{K}\right).$$

(b) *Suppose $c \ge \frac{1}{2}$. Then for any $K > 0$, we have*

$$\frac{\sum_{k=1}^K \alpha_k \mathbb{E}[\|\nabla F(x_k; \theta_1^*, \ldots, \theta_T^*)\|^2]}{\sum_{k=1}^K \alpha_k} \le O\left(\frac{\ln K}{\sqrt{K}}\right).$$

Then, the following theorem presents the convergence rate for strongly convex objectives.

Theorem 4.2 (Convergence Rate of m-TSCGD for Strongly Convex Objectives). *Suppose Assumptions 4.1 and 4.2 hold, and X is convex. Let the objective function $F(x; \theta_1^*, \ldots, \theta_T^*)$ be σ-strongly convex satisfying (3.2). Let $\{(x_k, y_k^{(T-1)}, \ldots, y_k^{(1)}, \theta_1^k, \ldots, \theta_T^k)\}_{k=1}^K$ be the sequence generated by Algorithm 3, by setting the step-sizes as $\alpha_k = 2/(\sigma(k+1))$, $\beta_k = 5k^{-1}$, and $\gamma_{j,k} = 2k^{-c}/\mu_{\theta_j}$ for $j = T, \ldots, 1$.*

(a) *For any $K > 0$, we have*

$$\mathbb{E}[\|x_K - x^*\|^2] \le O(K^{-1}) + \frac{2}{K(K+1)} \sum_{k=1}^K O(k^{1-c}).$$

(b) *Suppose $c = 1$. Then for any $K > 0$, we have*

$$\mathbb{E}[\|x_K - x^*\|^2] \le O(K^{-1}).$$

We defer the proofs for Theorems 4.1 and 4.2 to e-companion Sections C.3 and C.4, respectively.

Remark 4.1. Akin to the results in the two-level regime, we again find that the rate statements are shown to be the optimal (or near-optimal) and show no degradation from those obtained in the correctly specified regimes. Notably, these rate statements display a concrete improvement in terms of their invariance with respect to T and display an improvement over the findings in Yang et al. (2019) in terms of the actual rate as well as their dependence on T. We emphasize that the dependence on T in the rate statement appears in terms of a larger constant hidden in the $O(\cdot)$ notation.

Further, following the intuition that misspecification does not degrade the overall rate of convergence, it is also natural to conjecture that the optimal $O(K^{-1/2})$ convergence rate should be achievable for unaccelerated m-SCGD, by employing a more refined analysis based on recent advances (Chen et al. 2021a). We leave this extension for future work.

5. Numerical Experiments

In this section, we investigate the empirical performance of our proposed algorithms. We consider the misspecified risk-averse mean-deviation problem, which can be cast as a three-level compositional optimization problem. Next, we provide the detailed experimental setup and extensive numerics, whereas in e-companion Section D, we provide further numerics on a misspecified MDP and consider an uncertain portfolio optimization problem where the cost of conservatism is empirically studied.

5.1. Risk-Averse Mean-Deviation Optimization

Let $U(\bullet, \omega; \theta^*)$ be a random utility function and let θ^* be the true model parameter. We focus on the following misspecified mean-deviation risk-averse optimization problem, defined as

$$\max_\zeta \left[\mathbb{E}_\omega[U(\zeta, \omega; \theta^*)] - \mathbb{E}\left[\left(\mathbb{E}_\omega[U(\zeta, \omega; \theta^*)] - U(\zeta, \omega; \theta^*)\right)_+^p\right]^{1/p} - \frac{\sigma}{2}\|\zeta\|^2 \right], \tag{5.1}$$

where $p > 1$ is a positive scalar, and the true model parameter θ^* is the unique optimum to a μ_θ-strongly convex problem. The goal of this problem is to maximize the expected utility with mean-deviation risk and

ℓ_2-penalties under the misspecified regime. Given the true model parameter θ^*, this problem reduces to the classical mean-deviation risk-averse optimization problem examined by Ruszczyński and Shapiro (2006) and Ahmed et al. (2007). Meanwhile, this problem displays a three-level compositional structure in that it is equivalent to

$$\min_\zeta F(\zeta) \equiv f^{(1)}(\mathrm{E}_{\omega_2}[f^{(2)}_{\omega_2}(\mathrm{E}_{\omega_3}[f^{(3)}_{\omega_3}(\zeta;\theta^*)];\theta^*)]) + \frac{\sigma}{2}\|\zeta\|^2, \text{ where}$$

$$f^{(1)}((y_1,y_2)) \triangleq y_1 + y_2^{1+p},$$
$$f^{(2)}_{\omega_2}(z,\zeta;\theta^*) \triangleq (z, (z + U(\zeta,\omega_2;\theta^*))^p_+),$$
$$f^{(3)}_{\omega_3}(\zeta;\theta^*) \triangleq (U(\zeta,\omega_3;\theta^*), \zeta).$$

As pointed out by Yang et al. (2019), this problem is challenging because of the biased stochastic gradients. In this setting, only the inner functions $f^{(2)}$ and $f^{(3)}$ are complicated by uncertainty, but a bias is induced when calculating the gradient $\nabla f^{(1)}$, regardless of whether $f^{(1)}$ is deterministic or not. Therefore, solving this problem is as challenging as solving a three-level compositional problem. We tackle this misspecified compositional optimization problem by the three-level m-TSCGD algorithm (Algorithm 3), and compare its performance with the three-level a-TSCGD in Yang et al. (2019) with the true model parameter θ^*. In our experiments, we consider the scenario where θ^* is the unique optimum to the following problem

$$\theta^* \equiv \arg\min_{\theta\in\mathbb{R}^{50}} \mathrm{E}\left[(v_i - s_i^\top\theta)^2 + \frac{\mu_\theta}{2}\|\theta\|^2\right]. \tag{5.2}$$

Here we assume a linear model where $v_i \equiv s_i^\top\tilde{\theta} + \varepsilon_i$, $s_i, \tilde{\theta} \in \mathbb{R}^{50}$, and ε_i is a zero-mean noise term. We generate $\tilde{\theta}$ as per $\tilde{\theta} \sim \mathrm{Unif}[0,1]^{50}$. In each simulation iteration, s_i and ε_i are generated as per $s_i \sim \mathrm{Unif}[0,1]^{50}$ and $\varepsilon_i \sim \mathrm{N}(0,1)$, respectively. It follows that $v_i \equiv s_i^\top\tilde{\theta} + \varepsilon_i$. Note that $\mu_\theta \equiv 0.1$ to ensure the objective is μ_θ-strongly convex. Also, in our simulation, we first generate 10^6 samples and solve the batch version of the problem and take the solution as θ^*. Consider a utility function of a least-squares form given by

$$U(\zeta,\omega;\theta) \triangleq \frac{1}{2}(x_i^\top\zeta + c^\top(\theta - \theta^*))^2.$$

Here we let $\omega_i \triangleq (x_i, y_i)$, and we assume a linear model that $y_i \equiv x_i^\top\tilde{\zeta} + \varepsilon$, where $x_i, \tilde{\zeta} \in \mathbb{R}^d$ and ε_i is zero-mean. We generate $\tilde{\zeta} \sim \mathrm{Unif}[0,1]^d$ and employ it in all experiments. In each simulation iteration, we consider four settings to test the performance of our algorithm under different levels of misspecification. In each setting, we let $c \in \mathbb{R}^{50}$ and sample each of its components independently from a uniform distribution provided in Table 1; we generate x_i and ε_i as per $x_i \sim \mathrm{N}(0_{d\times 1}, 0.9\cdot I_{d\times d})$ and $\varepsilon_i \sim N(0, 0.2)$, respectively, whereby $y_i \equiv x_i^\top\tilde{\zeta} + \varepsilon_i$.

We set $p \equiv 2$, and problem (5.1) falls in the regime of convex optimization. We provide a benchmark comparison by generating 10^6 samples, solving the batch version of the problem, and taking the solution ζ^* as the optimal solution. In our simulations, at the k-iteration, we first generate one sample of $\{s_i, v_i\}$ to update our estimator of the model parameter θ_k by single-level SGD. Then, to update the solution ζ_k, we draw a random sample from $\{x_i, y_i\}$ and apply the accelerated scheme using the misspecified θ_k. We demonstrate the performance of the three-level m-TSCGD on both convex and strongly convex objectives.

5.2. Convex Objectives

We first investigate the performance of the m-TSCGD in the convex regime. We consider problem (5.1) without any ℓ_2-penalty, that is, $\sigma \equiv 0$. We set the step-sizes as $\alpha_k \equiv \min\{0.0002, 0.02k^{-1/2}\}$ and $\beta_{1,k} \equiv \beta_{2,k} \equiv \min\{0.0005, 0.05k^{-1/2}\}$ for m-TSCGD. In addition, to update the model parameters θ_k, we set the corresponding step-sizes in m-TSCGD as $\gamma_k \equiv \min\{10^{-4}, 10/k\}$. Here we let $d \equiv 100, 200$. We run both algorithms for 10^7 iterations, compare the performance of m-TSCGD over the four different misspecification setups, and plot the trajectory of the empirical log-error $\log(\mathrm{E}[\|\nabla F(\zeta_{R_k})\|^2])$ averaged over 10 independent simulations, where this metric matches that considered in Theorem 4.1, a detailed explanation for which is provided in Appendix D.1. We provide the results in Figure 1. We observe that the slopes of $\log k$ against $\log(\mathrm{E}[\|\nabla F(\zeta_{R_k})\|^2])$ are close to $-1/2$, matching the theoretical claim in Theorem 4.1 that the produced sequence converges at a rate of $\tilde{O}(k^{-1/2})$ for the three-level convex problem.

5.3. Strongly Convex Objectives

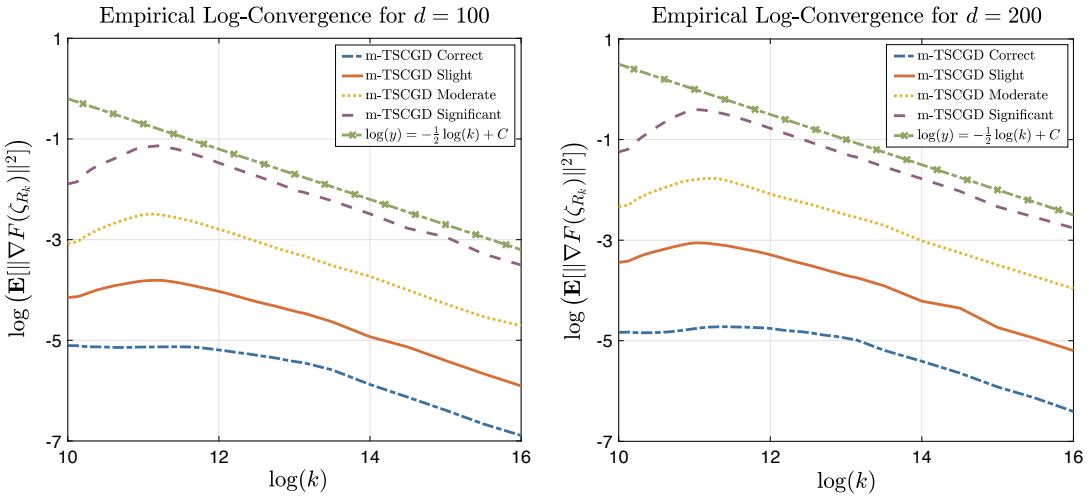
Next, we test our algorithm on σ-strongly convex problems with $\sigma \equiv 0.5$. We run m-TSCGD and a-TSCGD with step-sizes as $\alpha_k \equiv \min\{0.0002, 2/(\sigma(k+2))\}$, $\beta_{1,k} \equiv \beta_{2,k} \equiv \min\{0.0004, 4/(k+2)\}$, and $\gamma_k \equiv \min\{10^{-4}, 10/k\}$. We conduct 10 independent simulations, each of them consisting of 5×10^5 iterations, and plot the average of the log-error $\log(\|\zeta_k - \zeta^*\|^2)$. We conduct the experiments for $d \equiv 100, 200$.

In Figure 2, we observe that the slopes of the averaged log-error $\log(\|\zeta_k - \zeta^*\|^2)$ generated by misspecified algorithm approximately equal -1 over all misspecification

Table 1. Distribution of c for Different Levels of Misspecification in Risk-Averse Mean-Deviation Optimization

Setup	Significant	Moderate	Slight	Correct
Distribution of c	$c \sim \mathrm{Unif}[0,100]^{50}$	$c \sim \mathrm{Unif}[0,60]^{50}$	$c \sim \mathrm{Unif}[0,30]^{50}$	$c \equiv 0$

Figure 1. (Color online) Empirical Log-Convergence Rate of the m-TSCGD Algorithm for General Convex Objective Under Distinct Levels of Misspecification When d = 100, 200



setups, matching Theorem 4.2 where a rate of $\mathcal{O}(k^{-1})$ is derived for three-level σ-strongly convex problems. This also supports our theoretical analysis that the choice of step-sizes defined in Theorem 4.2 improves the rate from $\tilde{\mathrm{O}}(k^{-2.5})$ to $\mathrm{O}(k^{-1})$ for σ-strongly convex objectives.

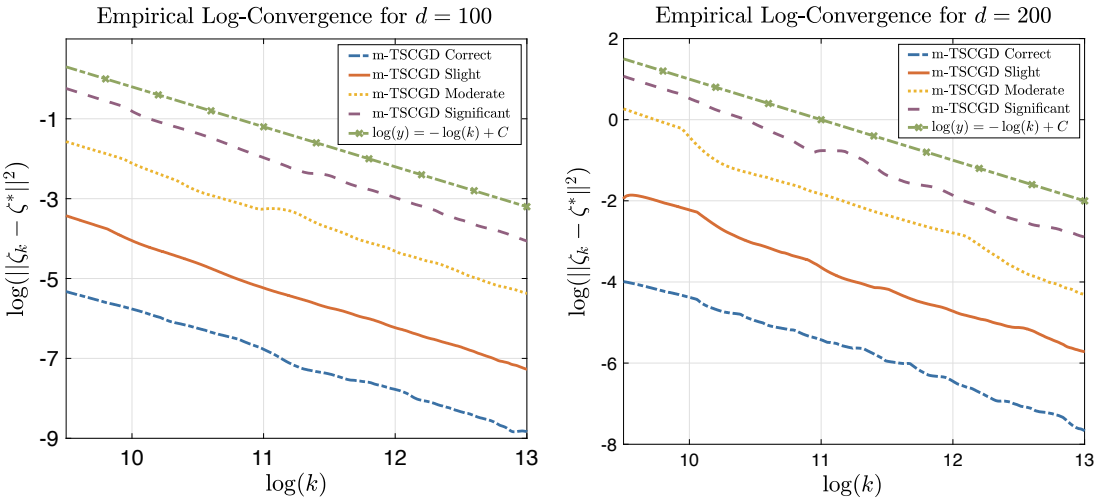
5.4. Insights from E-Companion Sections D.1 and D.2

In Section D.1, we consider the nonconvex misspecified compositional problem and observe that the empirical rates support theoretical claims, whereas in Section D.2, we observe that the impact of conservativism in robust approaches on uncertain portfolio selection problems can be significant in comparison with the proposed avenue.

6. Concluding Remarks

As systems grow in size and complexity, problem parameters are often not known a priori. One avenue for addressing the unavailability of problem parameters is to employ user-specified uncertainty sets, which have been adopted in conjunction with a robust optimization approach. An alternative approach rooted in classical research in economics emerges when there is a true or nominal value of this parameter and requires learning this parameter through available data while resolving the misspecified optimization problem. We consider the class of misspecified problems complicated by risk, nonconvexity, and, in a limited sense, dynamics. A unified framework for contending with such problems is

Figure 2. (Color online) Empirical Log-Convergence Rate of the m-TSCGD for σ-Strong Convex Objective Under Distinct Levels of Misspecification When d = 100, 200



available through compositional stochastic optimization problems. We consider the misspecified variant of such a class of problems and develop the data-driven compositional optimization schemes that can resolve misspecification while solving the original problem. Our key findings include that two-level compositional stochastic gradient schemes and their accelerated counterparts display no degradation of rate in strongly convex and nonconvex settings from their correctly specified counterparts. In fact, such statements continue to hold for accelerated schemes for the T-level multistage counterparts in strongly convex and nonconvex regimes.

Yet much remains to be investigated for this avenue of decision-making. In particular, can this framework accommodate more stylized and sophisticated learning models where smoothness and convexity concerns are weakened? To what extent can the statements be extended to misspecification in the presence of risk-afflicted constraints? Our future work will consider precisely such avenues.

Acknowledgments

The authors thank the area editor, the associate editor, and the reviewers for invaluable comments and feedback, which led to substantial improvements to earlier versions of this paper.

References

Ahmadi H, Shanbhag UV (2020) On the resolution of misspecified convex optimization and monotone variational inequality problems. *Comput. Optim. Appl.* 77(1):125–161.

Ahmadi H, Aybat NS, Shanbhag UV (2016) On the rate analysis of inexact augmented Lagrangian schemes for convex optimization problems with misspecified constraints. *2016 Amer. Control Conf.* (IEEE, New York), 4841–4846.

Ahmed S, C¸akmak U, Shapiro A (2007) Coherent risk measures in inventory problems. *Eur. J. Oper. Res.* 182(1):226–238.

Arimoto S, Kawamura S, Miyazaki F (1984) Bettering operation of robots by learning. *J. Robot. Systems* 1:123–140.

Arrow KJ (2002) The genesis of "optimal inventory policy." *Oper. Res.* 50(1):1–2.

Arrow KJ, Harris T, Marschak J (1951) Optimal inventory policy. *Econometrica* 19(3):250–272.

Aybat NS, Ahmadi H, Shanbhag UV (2022) On the analysis of inexact augmented Lagrangian schemes for misspecified conic convex programs. *IEEE Trans. Automat. Control* 67(8):3981–3996.

Balasubramanian K, Ghadimi S, Nguyen A (2022) Stochastic multilevel composition optimization algorithms with level-independent convergence rates. *SIAM J. Optim.* 32(2):519–544.

Beck A (2017) *First-Order Methods in Optimization* (SIAM, Philadelphia).

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) *Robust Optimization.* Princeton Series in Applied Mathematics (Princeton University Press, Princeton, NJ).

Bertsimas D, Sim M (2003) Robust discrete optimization and network flows. *Math. Programming* 98(1):49–71.

Bertsimas D, Sim M (2004) The price of robustness. *Oper. Res.* 52(1):35–53.

BesbesO, Zeevi AJ (2009) Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Oper. Res.* 57(6):1407–1420.

Besbes O, Zeevi AJ (2012) Blind network revenue management. *Oper. Res.* 60(6):1537–1550.

Borkar VS (2008) *Stochastic Approximation: A Dynamical Systems Viewpoint* (Hindustan Book Agency, New Delhi, India; Cambridge University Press, Cambridge, UK).

Chao GH (2013) Production and availability policies through the Markov decision process and myopic methods for contractual and selective orders. *Eur. J. Oper. Res.* 225(3):383–392.

Cheevaprawatdomrong T, Smith RL (2004) Infinite horizon production scheduling in time-varying systems under stochastic demand. *Oper. Res.* 52(1):105–115.

Chen T, Sun Y, Yin W (2021a) Closing the gap: Tighter analysis of alternating stochastic gradient methods for bilevel problems. Ranzato M, Beygelzimer A, Dauphin Y, Liang PS, Wortman Vaughan J, eds. *Adv. Neural Inform. Processing Systems 34 (NeurIPS 2021)* (Curran Associates, Red Hook, NY), 25294–25307.

Chen T, Sun Y, Yin W (2021b) Solving stochastic compositional optimization is nearly as easy as solving stochastic optimization. *IEEE Trans. Signal Processing* 69:4937–4948.

Dai W, Gai Y, Krishnamachari B (2014) Online learning for multi-channel opportunistic access over unknown Markovian channels. *2014 Eleventh Annual IEEE Internat. Conf. Sensing Comm. Networking (SECON)* (IEEE, New York), 64–71.

Ermoliev Y (1976) *Methods of Stochastic Programming.* Monographs in Optimization and OR (Nauka, Moscow).

Ghadimi S, Ruszczynski A, Wang M (2020) A single timescale stochastic approximation method for nested stochastic optimization. *SIAM J. Optim.* 30(1):960–979.

Hadley G, Whitin T (1963) *Analysis of Inventory Systems.* Prentice-Hall International Series in Management (Prentice-Hall, Upper Saddle River, NJ).

Ho CP, Petrik M, Wiesemann W (2018) Fast Bellman updates for robust MDPs. *Proc. Machine Learn. Res.* 80:1979–1988.

Ho-Nguyen N, Kilinc¸-Karzan F (2019) Exploiting problem structure in optimization under uncertainty via online convex optimization. *Math. Programming* 177(1–2):113–147.

Ho-Nguyen N, Kilinc¸-Karzan F (2021) Dynamic data-driven estimation of nonparametric choice models. *Oper. Res.* 69(4):1228–1239.

Hovakimyan N, Cao C (2010) *L1 Adaptive Control Theory: Guaranteed Robustness with Fast Adaptation.* Advances in Design and Control, vol. 21 (SIAM, Philadelphia).

Jiang H, Shanbhag UV (2013) On the solution of stochastic optimization problems in imperfect information regimes. *Proc. 2013 Winter Simulation Conf.* (IEEE Press, Piscataway, NJ), 821–832.

Jiang H, Shanbhag UV (2015) Data-driven schemes for resolving misspecified MDPs: Asymptotics and error analysis. *Proc. 2015 Winter Simulation Conf*. (IEEE Press, Piscataway, NJ), 3801–3812.

Jiang H, Shanbhag UV (2016) On the solution of stochastic optimization and variational problems in imperfect information regimes. *SIAM J. Optim.* 26(4):2394–2429.

Jiang H, Shanbhag UV, Meyn SP (2018) Distributed computation of equilibria in misspecified convex stochastic Nash games. *IEEE Trans. Automat. Control* 63(2):360–371.

Kirman AP (1975) Learning by firms about demand conditions. Day RH, Groves T, eds. *Adaptive Economic Models* (Academic Press, New York), 137–156.

Kushner H, Yin G (2003) *Stochastic Approximation and Recursive Algorithms and Applications,* Stochastic Modeling and Applied Probability, 2nd ed. (Springer, New York).

Lei J, Shanbhag UV (2020) Asynchronous schemes for stochastic and misspecified potential games and nonconvex optimization. *Oper. Res.* 68(6):1742–1766.

Li X (2013) Managing dynamic inventory systems with product returns: A Markov decision process. *J. Optim. Theory Appl.* 157(2):577–592.

Lian X, Wang M, Liu J (2017) Finite-sum composition optimization via variance reduced gradient descent. *Proc. Machine Learn. Res.* 54:1159–1167.

Mankowitz DJ, Levine N, Jeong R, Shi Y, Kay J, Abdolmaleki A, Springenberg JT, Mann T, Hester T, Riedmiller M (2020) Robust

reinforcement learning for continuous control with model misspecification. *8th Internat. Conf. Learn. Representations ICLR 2020* (International Conference on Learning Representations, Appleton, WI), 1–28.

Mannor S, Mebel O, Xu H (2012) Lightning does not strike twice: Robust MDPs with coupled uncertainty. *Proc. 29th Internat. Conf. Machine Learn.* (International Conference on Machine Learning, San Diego), 451–458.

Mannor S, Mebel O, Xu H (2016) Robust MDPs with k-rectangular uncertainty. *Math. Oper. Res.* 41(4):1484–1509.

Miyaguchi K (2021) Asymptotically exact error characterization of offline policy evaluation with misspecified linear models. Ranzato M, Beygelzimer A, Dauphin Y, Liang PS, Wortman Vaughan J, eds. *Adv. Neural Inform. Processing Systems 34 (NeurIPS 2021)* (Curran Associates, Red Hook, NY), 28573–28584.

Okuguchi K (1976) *Expectations and Stability in Oligopoly Models,* Lecture Notes in Economics and Mathematical Systems, vol. 138 (Springer-Verlag, Berlin).

Okuguchi K, Szidarovszky F (1990) *The Theory of Oligopoly with Multi-Product Firms,* Lecture Notes in Economics and Mathematical Systems, vol. 342 (Springer-Verlag, Berlin).

Porteus EL (1990) Stochastic inventory theory. Heyman DP, Sobel MJ, eds. *Stochastic Models,* Handbooks in Operations Research and Management Science, vol. 2 (Elsevier, Amsterdam), 605–652.

Puranam KS, Katehakis MN (2014) On optimal bidding and inventory control in sequential procurement auctions: The multi period case. *Ann. Oper. Res.* 217(1):447–462.

Puterman ML (1994) *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley Series in Probability and Statistics (Wiley, Hoboken, NJ).

Qi M, Grigas P, Shen M (2021) Integrated conditional estimation-optimization. Preprint, submitted October 24, https://arxiv.org/abs/2110.12351.

Rigamonti A, Lučivjanská K (2022) Mean-semivariance portfolio optimization using minimum average partial. *Ann. Oper. Res.* 1–19.

Robbins H, Monro S (1951) A stochastic approximation method. *Ann. Math. Statist.* 22(3):400–407.

Ross S (1983) *Introduction to Stochastic Dynamic Programming,* Probability and Mathematical Statistics (Academic Press, Inc., New York).

Ruszczynski A (2021) A stochastic subgradient method for nonsmooth nonconvex multilevel composition optimization. *SIAM J. Control Optim.* 59(3):2301–2320.

Ruszczyński A, Shapiro A (2006) Optimization of convex risk functions. *Math. Oper. Res.* 31(3):433–452.

Satheesh Kumar R, Elango C (2012) Markov decision processes in service facilities holding perishable inventory. *Opsearch* 49(4):348–365.

Sato M, Abe K, Takeda H (1982) Learning control of finite Markov chains with unknown transition probabilities. *IEEE Trans. Automat. Control* 27(2):502–505.

Shapiro A, Dentcheva D, Ruszczyński A (2009) *Lectures on Stochastic Programming,* MPS/SIAM Series on Optimization, vol. 9 (SIAM, Philadelphia).

Uchiyama M (1978) Formation of high-speed motion pattern of a mechanical arm by trial. *Trans. Soc. Instrument Control Engineers* 14(6):706–712.

Wang M, Fang EX, Liu H (2017a) Stochastic compositional gradient descent: Algorithms for minimizing compositions of expected-value functions. *Math. Programming* 161(1–2):419–449.

Wang M, Liu J, Fang EX (2017b) Accelerating stochastic composition optimization. *J. Machine Learn. Res.* 18(1):3721–3743.

Wu M, Zhu SX, Teunter RH (2013) Newsvendor problem with random shortage cost under a risk criterion. *Internat. J. Production Econom.* 145(2):790–798.

Yang S, Wang M, Fang EX (2019) Multilevel stochastic gradient methods for nested composition optimization. *SIAM J. Optim.* 29(1):616–659.

Yang S, Zhang X, Wang M (2022) Decentralized gossip-based stochastic bilevel optimization over communication networks. Koyejo S, Mohamed S, Agarwal A, Belgrave D, Cho K Oh A, eds. *Adv. Neural Inform. Processing Systems 35 (NeurIPS 2022)* (Neural Information Processing Systems, San Diego), 238–252.

Zhang Z, Lan G (2020) Optimal algorithms for convex nested stochastic composite optimization. Preprint, submitted November 19, https://arxiv.org/abs/2011.10076.

Zhu H, Xu T, Paschalidis IC (2019) Learning parameterized prescription policies and disease progression dynamics using Markov decision processes. *2019 Amer. Control Conf. (ACC)* (IEEE Press, Piscataway, NJ), 3438–3443.

Shuoguang Yang was an assistant professor in the Department of Industrial Engineering and Decision Analytics, The Hong Kong University of Science and Technology. His research focuses on machine learning, optimization, and revenue management.

Ethan X. Fang is an associate professor of Biostatistics & Bioinformatics and is affiliated with the Decision Sciences group at Fuqua School of Business at Duke University. His research focuses on computational and statistical perspectives of different data science problems.

Uday V. Shanbhag holds the Gary and Sheila Bello Chair in the department of Industrial and Manufacturing Engineering at the Pennsylvania State University, University Park. His interests lie in the resolution of analytical and computational questions in the context of optimization and equilibrium problems, complicated by uncertainty, nonsmoothness, nonconvexity, hierarchy, and misspecification.