Learning to Optimize with Stochastic Dominance Constraints

Hanjun Dai
Google Research

Yuan Xue
Google Cloud

Niao He
ETH Zurich

Bethany Wang
Google Cloud

Na Li
Google Research & Harvard

Dale Schuurmans
Google Research & University of Alberta

Bo Dai
Google Research & Georgia Tech

Abstract

In real-world decision-making, uncertainty is important yet difficult to handle. Stochastic dominance provides a theoretically sound approach to comparing uncertain quantities, but optimization with stochastic dominance constraints is often computationally expensive, which limits practical applicability. In this paper, we develop a simple yet efficient approach for the problem, *Light Stochastic Dominance Solver (light-SD)*, by leveraging properties of the Lagrangian. We recast the inner optimization in the Lagrangian as a learning problem for surrogate approximation, which bypasses the intractability and leads to tractable updates or even closed-form solutions for gradient calculations. We prove convergence of the algorithm and test it empirically. The proposed light-SD demonstrates superior performance on several representative problems ranging from finance to supply chain management.

1 INTRODUCTION

Decision making under uncertainty (Kochenderfer, 2015) is an ubiquitous challenge attracting research from a wide range of communities including machine learning (Sani et al., 2012; Petrik and Subramanian, 2012; Tamar et al., 2013, 2015; La and Ghavamzadeh, 2013), operations research (Delage and Mannor, 2010), economics and management science (Machina and Viscusi, 2013). To date, decision methods in machine learning primarily focus on maximizing expected return, which is only applicable to cases when the decision-maker is risk-neutral. In reality, most decision-makers are risk-sensitive – many are willing to give up some expected reward to protect against large losses, *i.e.*, demonstrating risk-aversion. Stochastic dominance (SD), introduced in (Mann and Whitney, 1947) and (Lehmann, 1955), provides a principled approach to comparing random variables, standing out as a general and flexible model for incorporating risk aversion in decision making, applied to scenarios ranging from economics (Quirk and Saposnik, 1962; Rothschild and Stiglitz, 1970, 1971), finance (Dentcheva and Ruszczyński, 2006), path planning (Nie et al., 2012), to control and reinforcement learning (Dentcheva and Ruszczyński, 2008; Haskell and Jain, 2013). Despite the elegance of the concept, an efficient computational recipe for ensuring stochastic dominance between general distributions remains lacking, due to the continuum nature of the criterion. In this paper, we consider a longstanding challenge (Levy, 1992):

*Develop a **simple yet efficient** algorithm for optimization with **general** stochastic dominance constraints.*

There have been many attempts to handle stochastic dominance constraints in optimization. Ogryczak and Ruszczyński (1999, 2001, 2002) established consistency between stochastic dominance and mean-risk surrogates. Therefore, several mean-risk optimization formulations have been proposed as surrogates for evaluating stochastic dominance conditions. Although the proposed mean-risk models can be efficiently solved, these surrogates only provide a necessary condition and are unable to model the full spectrum of risk-averse preferences, and thus might lead to inferior solutions (Ogryczak and Ruszczyński, 2001).

To bypass suboptimality from surrogates, research on original optimization with stochastic dominance was considered by (Dentcheva and Ruszczyński, 2003). Existing methods for such problems can generally be categorized by the underlying stochastic dominance formulations considered, based on k-th order distribution functions, utility functions, or the Strassen theorem perspective, respectively. Based on the direct comparison over distribution functions, Dentcheva and Ruszczyński (2003, 2004a) derived a linear programming (LP) formulation for second-order stochas-

tic dominance constraints. Similarly, Noyan et al. (2006) derived a mixed-integer programming (MIP) formulation of first-order stochastic dominance constraints, which was further adopted in a cutting plane algorithm in (Rudolf and Ruszczyński, 2008). Instead of stochastic dominance based on k-th order distribution functions, (Luedtke, 2008) established a more compact LP and MIP for second-/first-order stochastic dominance constraints based on an equivalent reformulation through Strassen's theorem (Strassen, 1965). Alternatively, stochastic dominance can also be characterized from a dual view, which leads to an implementation through robust optimization with respect to utility functions (Post, 2003; Armbruster and Delage, 2015).

Even though these ideas and formulations are inspiring, they primarily focus on random variables with *finite* support. It is always possible to approximate the constraints over continuous random variables by applying existing methods with the empirical distribution over samples (Hu et al., 2012; Haskell and Toriello, 2018), but the size of the resulting optimization increases *at least quadratically* with respect to the number of discretization bins, thus making the memory and computation costs unaffordable if one seeks a high-quality solution with a fine discretization.

In this paper, we introduce a novel algorithm, *Light Stochastic Dominance Solver (light-SD)*, as a viable solution to the question. The proposed light-SD is able to handle general stochastic dominance constraints while maintaining memory and computational efficiency, thus enabling applicability to large-scale practical problems. Our development starts with the well-established Lagrangian of the optimization with stochastic dominance constraints, which induces a special structure in the dual functions and establishes a connection to utility functions. Although such a connection has been noted in the literature, it has remained unclear how this can be exploited for an efficient algorithm, due to the difficulty in optimizing with an *intractable* expectation over an *infinite* number of *unbounded* functions. We overcome these difficulties through learning for surrogate approximation. Specifically, we approximate the intractable expectation via samples, and optimize with design special parametrized dual functions for surrogate approximation. These two strategies pave the way for a simple and efficient algorithm.

The reminder of the paper is organized as follows. First, we provide the necessary background on optimization with stochastic dominance constraints and its corresponding primal-dual form in Section 2. In Section 3.1 we design an efficient stochastic gradient descent algorithm applicable to general stochastic dominance constraints with general random variables. By scrutinizing the conditions on the dual functions, we introduce a specially tailored parametrization for dual functions in Section 3.2. We provide the theoretical analysis of the proposed light-SD in Section 3.3. Finally, in Section 5, we conduct an empirical evaluation on synthetic benchmarks and real-world problems, and find that

the proposed light-SD significantly improves performance in terms of both memory and computation.

Our implementation of light-SD will be released at https://github.com/google-research/google-research/tree/master/lightsd.

2 PRELIMINARIES

In this section, we introduce the stochastic dominance concept and specify the optimization problem under stochastic dominance constraints with concrete applications. Then, we provide the corresponding primal-dual reformulation, which inspires our algorithm in next section.

2.1 Optimization with Stochastic Dominance Constraints

First-order dominance: Let X denote a real valued random variable. The corresponding distribution function is defined as $F(X; \eta) = P(X \le \eta)$ for $\eta \in \mathbb{R}$. We say that a random variable X dominates in the *first order* a random variable Y if the following holds:

$$X \succeq_1 Y \text{ if } F(X; \eta) \le F(Y; \eta), \quad \forall \eta \in \mathbb{R}. \tag{1}$$

First-order stochastic dominance is describing that, when both X and Y are the outcome of two portfolios, $X \succeq_1 Y$ implies $P(X > \eta) \ge P(Y > \eta)$, which means for any possible return threshold η, the portfolio X gives at least as high a probability of exceeding the threshold as Y.

Second-order dominance: We say X dominates Y in the *second order* if the following holds:

$$X \succeq_2 Y \text{ if } \int_{-\infty}^{\eta} F(X; \alpha) d\alpha \le \int_{-\infty}^{\eta} F(Y; \alpha) d\alpha, \forall \eta \in \mathbb{R}. \tag{2}$$

Second-order dominance has significant practical value: it is closely linked to risk-adverse decision making (Ogryczak and Ruszczyński, 1999, 2001, 2002), and robust optimization w.r.t. the family of concave, non-decreasing utility functions (Armbruster and Delage, 2015). In particular, second-order dominance can be equivalently defined by:

$$X \succeq_2 Y \Leftrightarrow E[(\eta - X)_+] \le E[(\eta - Y)_+], \forall \eta \in \mathbb{R}. \tag{3}$$

High-order dominance: The concept of stochastic dominance has been extended to high-order. Specifically, one can recursively define the functions

$$F_k(X; \eta) = \int_{-\infty}^{\eta} F_{k-1}(X; \alpha) d\alpha = \frac{1}{(k-1)!} E\left[(\eta - X)_+^{k-1}\right],$$
$$\tag{4}$$

Then, X dominates Y in the *kth-order* if $F_k(X; \eta) \le F_k(Y; \eta)$ for $\forall \eta \in \mathbb{R}$. By definition, kth-order dominance implies $(k+1)$st-order dominance if the random variables F_{k+1} are well-defined.

We are interested in the following optimization problem:

$$\max_{z \in \mathcal{Z}} f(z), \quad \text{s.t.} \quad g(z, \xi) \succeq_k Y, \tag{5}$$

where $f(\cdot): \mathbb{R}^d \to \mathbb{R}$ is the objective function, $\mathbb{R}^d$

denotes the feasible set of decision variables, Y denotes the reference random variable, and $g(\cdot,\cdot): \mathbb{R}^d \times \mathbb{R}^p \to \mathbb{R}$ is a given mapping expressing a deterministic outcome depending on the decision z and random variable ξ.

The optimization (5) can be easily generalized with multiple stochastic dominance constraints, i.e., $g_i(z,\xi) \succeq_k Y_i$, $\forall i = 1,\ldots,h$, with each random outcome $g_i(z,\xi)$ stochastic dominant w.r.t. the corresponding references Y_i; and multivariate stochastic dominance constraints (Dentcheva and Ruszczyński, 2009).

Below we provide several concrete applications of (5).

- **Portfolio Optimization:** Consider a set of d financial assets, whose returns can be captured by a random variable $\xi \in \mathbb{R}^d$. The optimization problem seeks a portfolio of assets formed by an allocation vector $z \in \mathbb{R}^d_+, |z| = 1$. The return of the portfolio $X = \xi^\top z$ is also a random variable. The objective of portfolio optimization is to maximize the expected portfolio return, subject to the constraint that the return dominates a reference return Y:

$$\max_{z\in\mathcal{Z}} \; \mathbb{E}\left[\xi^\top z\right], \quad \text{s.t.} \quad \xi^\top z \succeq_k Y, \quad (6)$$

Y can come from a reference portfolio or market index. Under second-order dominance, this formulation guarantees a risk-averse decision maker will not prefer Y to the optimal solution of (6).

- **Optimal Transportation:** Consider a set of m regions with demand expressed by a m-dimensional random variable $\xi \in \mathbb{R}^m$. This demand is to be fulfilled by n warehouses with inventories given by an n-dimensional random variable $Y \in \mathbb{R}^n$ (capturing uncertainty in supplier lead time and replenishment availability). The optimal transport problem seeks a transport map, represented as a decision variable $z \in \mathbb{R}^{m \times n}_+$, where z_{ij} is the ratio of demand ξ_i being fulfilled by warehouse j. The decision must satisfy $\forall i, \sum_j z_{ij} = 1$. Let $h_{ij} \in \mathbb{R}^{m \times n}_+$ be the transport cost to fulfill a unit demand from region i by warehouse j. The objective is to minimize expected transport cost, subject to the constraint that the demand ξ is dominated by the warehouse supply Y:

$$\max_{z\in\mathcal{Z}} \; \mathbb{E}\left[\sum_i \sum_j h_{ij} z_{ij} \xi_i\right], \quad (7)$$

$$\text{s.t.} \; Y_j \succeq_k \sum_i \xi_i z_{ij}, \forall j \quad (8)$$

A standard approach is to rewrite the first-/second-order stochastic dominance constraints as binary/linear constraints with $O(M^2)$ auxiliary variables, where M denotes the support size for discrete variables or the number of discretization bins for continuous variables. Please refer to Appendix A for details of the constraint reformulation. Clearly, the computational cost quickly becomes unacceptable for large-scale problems. Such computational difficulty is a major bottleneck in applying stochastic dominance in practice,

making an efficient and general algorithm an urgent need.

2.2 Primal-Dual Formulation

It is common when solving a stochastic constrained optimization to consider a relaxation where $\eta \in [a,b]$ (Dentcheva and Ruszczyński, 2003). Given a continuous F_k defined in (4) over $[a,b]$, we define the Lagrangian $\mathcal{L}^k(z,\mu): \mathbf{rca}([a,b]) \to \mathbb{R}$ of (5) and obtain:

$$\max_z \min_{\mu\in\mathbf{rca}([a,b])} \mathcal{L}^k(z,\mu) =$$
$$f(z) - \int_a^b \left[F_k(g(z,\xi);\eta) - F_k(Y;\eta)\right] d\mu(\eta), \quad (9)$$

where $\mathbf{rca}([a,b])$ are the regular, countably additive measures on $[a,b]$, i.e., the dual space of the continuous function space given by the Riesz representation theorem. Under some regularity conditions, the $\mathcal{L}^k(z,\mu)$ in (9) can be equivalently reformulated as

$$\max_z \min_{u\in\mathcal{U}_k} L(z,u) = f(z) + \mathbb{E}\left[u(g(z,\xi))\right] - \mathbb{E}\left[u(Y)\right], \quad (10)$$

where $\mathcal{U}_k := \left\{ u(\cdot) \; \middle| \; \int_a^b F_k(X;\eta)\, d\mu(\eta) = \mathbb{E}\left[u(X)\right], \forall \mu \in \mathbf{rca}([a,b]) \right\}$

denotes the feasible set of dual functions for the k-th order dominance constraints. Obviously, the major difference between k-th order stochastic dominance mainly lies in the construction of the dual functions.

Theorem 1 (Theorem 1 (Dentcheva and Ruszczyński, 2004b)). *Under regularity conditions on the solution to (5) under a first-order dominance condition, $\mathcal{U}_1$ in (10) is the set of nondecreasing and left continuous functions over $[a,b]$.*

Similarly, one can characterize the dual function for the second-order dominance constraints.

Theorem 2 (Theorem 4.2 (Dentcheva and Ruszczyński, 2003)). *Under some regularity conditions, the Lagrangian can be reformulated as (10) with $u(\cdot) \in \mathcal{U}_2$, $\mathcal{U}_2$ is the set of concave and nondecreasing functions over $[a,b]$.*

This primal-dual view connects stochastic dominance with the utility function perspective (Von Neumann and Morgenstern, 2004). In fact, instead of fully specifying the utility function, (10) can be understood as a robust optimization with respect to all possible utilities, ensuring for the worst-case utility function in the set that the solution is still preferable to the reference strategy, therefore, demonstrating the benefits of the stochastic dominance comparison.

Although the duality relation and optimality condition through the Lagrangian (9) and (10) has been established in (Dentcheva and Ruszczyński, 2003) (please refer to Appendix B for completeness), the Lagrangian forms are still intractable in general due to

- **i),** the *intractable* expectation in the optimization,

- **ii),** the *infinite* number of *unbounded* functions in dual feasible set $\mathcal{U}_k$.

3   LIGHT STOCHASTIC DOMINANCE SOLVER

In this section, we exploit machine learning techniques to bypass the two intractabilities of Lagrangian (10) discussed in previous section. These machine learning techniques eventually lead to a simple yet efficient algorithm for general optimization problems with stochastic dominance, while obtaining rigorous guarantees.

3.1   Stochastic Approximation

We first introduce the stochastic approximation scheme for the Lagrangian (10) to handle the intractable expectation. Specifically, we define

$$\hat{L}(z, u) = f(z) + \frac{1}{N} \sum_{i=1}^{N} \left(u(g(z, \xi_i)) - u(y_i) \right) \quad (11)$$

with $\{\xi_i\}_{i=1}^{N} \sim p(\xi)$ and $\{y_j\}_{j=1}^{N} \sim p(y)$. Consider a stochastic approximation algorithm that alternates between solving $\hat{u} = \arg\min_{u \in U} \hat{L}(z, u)$, and updating z with a stochastic gradient $\nabla_z \hat{L}(z, \hat{u})$, which leads to Algorithm 1.

Algorithm 1 Light Stochastic Dominance Solver (light-SD)

1: Initialize z randomly, set number of total iterations T.
2: **for** iteration $t = 1, \ldots, T$ **do**
3:    Sample Sample $\{\xi_i, y_i\}_{i=1}^{n}$.
4:    Compute $\hat{u} = \arg\min_{u \in U} \hat{L}(z, u)$.
5:    $z \leftarrow z + \gamma_t \nabla_z \hat{L}(z, \hat{u})$.
6: **end for**

In Algorithm 1, we avoid the first difficulty by using the stochastic approximation $\hat{L}(z, u)$ to replace intractable expectation. However, optimization over an unbounded function space U is in general intractable. Next, we design the special dual parametrization, which induces an efficient solver for $\hat{u}$ in Section 3.2 to bypass the optimization over an infinite dual feasible set.

3.2   Dual Parametrization

To bypass the computational intractable and PAC unlearnable optimization over infinitely many unbounded functions in the dual feasible set, we introduce bounded parameterized dual functions, thus reducing the semi-infinite programming problem to a finite parameter optimization. However, arbitrary parameterizations will generate functions outside the feasible set of utility functions, which breaks the requirements of stochastic dominance with specific order.

In this section, inspired by the proofs of Theorem 1 and Theorem 2, we design specific parametrizations that cover the corresponding *bounded* utility function sets, which balances the tradeoff between sample complexity and relaxation of the optimization. Meanwhile, for the commonly used first and second order stochastic dominance constraints, our

parametrization induces an efficient dual solution, which further reduces the computation complexity. We also introduce the generic parametrization of the dual for high-order dominance problems.

3.2.1   Parameterization of U_1

Following Theorem 1, one straightforward idea is to parametrize the whole family of non-decreasing functions, *e.g.*, (Sill, 1997; Dugas et al., 2009; Wehenkel and Louppe, 2019), which, however, is unbounded function space, and thus, not only computationally intractable but also sample complexity unbounded. Therefore, we do not directly work with all non-decreasing functions, but a restricted dual function space as a surrogate:

$$U_1 := \left\{ u(x) = \mathbb{E}_{p(\eta)} \left[\mathbf{1} \left(x \leqslant \eta \right) \right], p(\eta) \in \mathcal{P}([a, b]) \right\}, \quad (12)$$

where $\mathcal{P}([a, b])$ is the set of distribution over $[a, b]$. This dual actually can be understood as seeking $\eta \in [a, b]$, such that $F(X; \eta) > F(Y; \eta)$. This bounded function space relaxes the original optimization, but makes the problem statistical learnable as we discussed in Appendix B.1.

With this dual parametrization, the optimal dual function can be easily obtained. For stochastic approximation with finite samples $X \sim P(X)$, let us define

$$h_X(\eta) = \sum_{x \in X} \mathbf{1} \left(\eta > x \right). \quad (13)$$

We overload the notation a bit to denote $h_{\{g\}}(\eta) = \sum_{i=1}^{N} \mathbf{1} \left(\eta > g(z, \xi_i) \right)$.

Given sample $\{\xi_i, y_i\}_{i=1}^{N}$, we define

$$\mu^*(\{\xi\}, \{y\}) = \left\{ \eta : \eta \in \mathbb{R} \text{ and } h_{\{g\}}(\eta) > h_{\{y\}}(\eta) \right\},$$

which is the set of the values of η that violate the constraints. We omit the arguments of μ^* when appropriate. Based on the understanding of dual functions in (12), the optimal dual function is constructed by putting probability mass on the step-function with threshold in the set μ^*, *i.e.*,

$$U_1^* = \left\{ u^*(x) := \mathbf{1}(\eta > x), \eta \in \mu^* \right\}. \quad (14)$$

and optimal dual u^*, where $\eta^* = \arg\max_{\in U_1} h_{\{g\}}(\eta) - h_{\{y\}}(\eta)$. When $\{\xi, y\}_{i=1}^{N}$ has no duplication, we can use any convex combination in U_1^*.
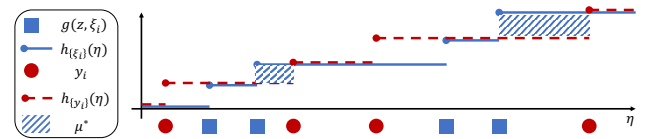
Figure 1: *Designed dual function for first order constraints.*

The relations among samples $g(z, \xi_i), y_i$, the corresponding $h_{\{g\}}, h_{\{y\}}$ and the resulting u^* can be depicted in Figure 1.

Although the above design of the functional space U_1 is not exhaustive, it suffices to show that the dominance constraints are satisfied on the current batch of samples $\{\xi_i\}$ and $\{y_i\}$ only when U_1^* is an empty set for this batch of samples.

Practical implementation. Note that although μ may contain infinitely many values, only a finite number of them would result in distinct gradients with respect to z. Particularly, a finite set μ overlapped with the actual samples would be sufficient:

$$\mu = \mu \cap \{g(z, \xi_i) \cup \{y_i\}\}. \tag{15}$$

Furthermore, note that the above function μ is a step function which is not continuous, and thus, causes difficulty for gradient-based optimization. To resolve this issue, inspired from neural network approximation (Barron, 1993), we approximate the function with a smoothed basis, *i.e.*,

$$\mathcal{U}_1 = \left\{ u_\eta(x) := \tanh\left(\frac{x - \eta}{\tau}\right), \eta \in \mu \right\}. \tag{16}$$

With the above construction, we exploit all functions in $\mathcal{U}_1$ to obtain $\hat u(\cdot) = \sum_{u \in \mathcal{U}_1} u(\cdot)$, upon which we can have the stochastic gradient for the constrained optimization as

$$\nabla_z \hat L(z, \hat u) = \nabla_z f(z) + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{|\mu|} \sum_{u \in \mathcal{U}_1} \nabla_z u(g(z, \xi_i)). \tag{17}$$

3.2.2 Parameterization of $\mathcal{U}_2$

Similar to the above case, we design the dual function

$$\mathcal{U}_2 := \left\{ u(x) = \mathbb{E}_{p(\cdot)} \left[(\eta - x)_+, p(\eta) \in (a, b) \right] \right\}, \tag{18}$$

which essentially seeks η such that $F_2(X; \eta) > F_2(Y; \eta)$. Please refer to Appendix B.2 for the discussion about the approximation induce by boundedness.

We follow the same paradigm to first define the function

$$h_X(\eta) := \sum_{x \in X} (\eta - x)_+ \tag{19}$$

and correspondingly the overloaded definition of $h_{f_i g}(\eta) := \sum_{i=1}^{N} (\eta - g(z, \xi_i))_+$. The set of η that would violate the second order stochastic dominance constraints is defined similarly where

$$\mu = \left\{ \eta : \eta \in \mathbb{R} \text{ and } h_{f_i g}(\eta) > h_{y_i}(\eta) \right\}.$$

In this case, the basis of constructing optimal dual function can be derived as

$$\mathcal{U}_2 = \left\{ \hat u(x) := (\eta - x)_+, \eta \in \mu \right\}, \tag{20}$$

and we can obtain the optimal dual by putting mass on the $\eta = \arg\max_{\eta \in \mathcal{U}_2} h_{f_g}(\eta) - h_y(\eta)$.

Figure 2: *Designed dual function (slope is rescaled for visualization) for second order constraints.*
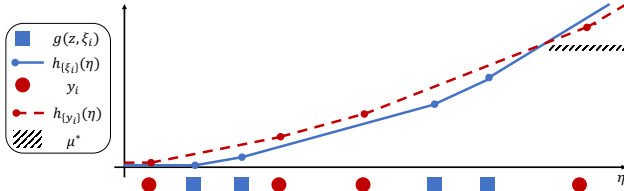
The relations among samples $g(z, \xi_i), y_i$, the corresponding $h_{f_i g}, h_{fy_i g}$ and the resulting u can be depicted in Figure 2.

One can also see through Figure 1 and Figure 2 that, the first order stochastic dominance constraints are strictly stronger

Table 1: *Computation and Memory Cost per Iteration.*

Algorithms	# of Variables	Memory	Computation
light-SD	(N)	(N)	$\Theta(N)$
LP-based	(N^2)	(N^2)	$O(N^3)$

than the second order ones, as μ here is a subset of which in the first order case.

Practical implementation. Similar to the first order stochastic dominance, here only a finite subset of μ is needed to obtain the stochastic gradients, which is analogous to the previous case in (15). Therefore, we use $\hat u(x) = \frac{1}{|k - k|} \sum (\eta - x)_+$, which is a ReLU network, and the gradient w.r.t. x can also be calculated with (17) using the obtained $\hat u$. Although we use the average $\hat u$, instead of the optimal dual solution with η. It is easy to check $\hat u$ is a psuedo-gradient, and still preserves convergence (Poljak and Tsypkin, 1973; Yang et al., 2019).

3.2.3 Parameterization of $\mathcal{U}_k$

In our paper we mainly focus on the first and second order SD constraints, as these are most practically useful ones. For completeness we present the derivations of higher-order SD constraints here. The derivation of the Lagrangian for (5) with high-order dominance constraints also suggests a neural network architecture for parameterizing functions in $\mathcal{U}_k$, as shown in Appendix B.2. Specifically, we consider the function parametrization as:

$$u(x) = \int_a^x \int_a^{x_{k-2}} \cdots \int_a^{x_{k-1}} \phi(t) dt + \beta \, dx_{k-2} \ldots dx_1, \tag{21}$$

where $\phi(t) > 0$ is a bounded non-negative function and can be easily satisfied by applying activation functions like `relu` or `softplus` to the output layer. Since the integration is one-dimensional, the numerical quadrature can be used in the forward-pass. This parametrization directly follows the derivation of dual function for high-order stochastic dominance, therefore, satisfying the requirements.

For higher-order SD we can no longer achieve a closed-form solution even with finite samples. We consider gradient descent for seeking the optimal inner solution. Specifically, the gradient calculation can be easily implemented by exploiting the linearity of integration and differentiation, *i.e.*,

$$\nabla u(x) = \int_a^x \int_a^{x_{k-2}} \cdots \int_a^{x_{k-1}} \nabla \phi(t) dt + \beta \, dx_{k-2} \ldots dx_1.$$

The gradient calculation is quite similar to the forward-pass through numerical quadrature layer, but with gradients w.r.t. the parameters of first layer.

3.3 Theoretical Analysis

The proposed light-SD can be understood as optimization with a learned surrogate approximation. In fact, for second-

order case, we can characterize the global optimal convergence with the learned surrogate.

Theorem 3. *If f and g are both concave, $\ell(z) := \min_{u \in U_2([a;b])} L(z, u)$ is concave with respect to z. With stepsize $\gamma_t = O\left(\frac{1}{\sqrt{T}}\right)$, we define $z_T = \frac{\sum_{t=1}^{T} \gamma_t z_t}{\sum_{t=1}^{T} \gamma_t}$ where $\{z_t\}_{t=1}^{T}$ are the iterates from light-SD, and $z^* = \arg\max_{z} \ell(z)$ is the optimal solution. Under the assumption that $\|\nabla_z f(z)\|_2^2 \leq C_f^2$ and $\|\nabla_z g(z, \xi)\|_2^2 \leq C_g^2$, with probability $1 - \delta$, we have*

$$\mathbb{E}[\ell(z^*) - \ell(z_T)] = O\left(\frac{(|a| + |b|) + \sqrt{\frac{\log(1/\delta)}{N}}}{\sqrt{N}} + \frac{1}{\sqrt{T}}\right).$$

The theorem implies that besides the standard convergence rate for stochastic gradient for convex function, *i.e.*, $O\left(\frac{1}{\sqrt{T}}\right)$, there is an extra error $O\left(\frac{(|a| + |b|) + \sqrt{\frac{\log(1/\delta)}{N}}}{\sqrt{N}}\right)$, which comes from optimization using the learned surrogate with N-finite samples. To balance these two errors, one can use T samples in each iteration, which leads to total error in rate $\Theta\left(\frac{1}{\sqrt{T}}\right)$. With fewer samples in each batch, *e.g.*, $N \sim O(T^\alpha)$ with $\alpha \in (0, 1)$, we obtain $\Theta\left(T^{-\frac{\alpha}{2}}\right)$.

The proof is obtained by characterizing the approximation error induced in each step in the mini-batch stochastic gradient descent. Due to the space limitation, we omit the details. Please refer to Appendix B.3 for the complete proof. The proof in fact can be of independent interests. It improves the results from (Nouiehed et al., 2019) by exploiting stochastic gradient to make the computation tractable. Meanwhile, we relax the strongly convex requirement. Comparing to (Hu et al., 2021), the proof considers the mini-batch in the algorithm, while the algorithm proposed in (Hu et al., 2012) requires full batch updates.

The proposed light-SD is efficient in terms of both memory and computation cost, bypasses notorious difficulties in realizing stochastic dominance constraints, as discussed in Appendix A. We compare the computation and memory cost with the existing LP-based algorithm in Table 1. With an interior-point solver for LP (Nesterov and Nemirovskii, 1994), the proposed light-SD reduces both computation and memory cost to linear w.r.t. N, therefore, is scalable for practical problems.

Although we focused on SD over scalar random variable, light-SD can be extended to multivariate SD constraints (Dentcheva and Ruszczyński, 2009), where one extra random projection will be introduced to project multivariate random variables to scalar.

4 RELATED WORK

This work bridges several topics, including robust optimization, decision making under uncertainty, ML for optimization, and optimization with approximated gradients.

ML for optimization. Leveraging machine learning to help optimization has raised a lot of interest in recent years. Some representative works include leveraging reinforcement learning (Khalil et al., 2017; Bello et al., 2016), unsupervised learning (Karalias and Loukas, 2020), and learning guided search (Li et al., 2018) for combinatorial optimization and stochastic optimization (Dai et al., 2021). One main goal has been to leverage the generalization ability of neural networks to help solve new instances from the same distribution better or faster. While these successes inspired and motivated this work, one key distinction is that we solve individual problems with guarantees and without meta-training. This principle is similar to model-based black-box optimization (Snoek et al., 2012; Papalexopoulos et al., 2022), where surrogate models are estimated iteratively and used to guide the optimization. To the best of our knowledge, this is the first stochastic approach for solving optimization problems with stochastic dominance constraints that is both scalable and achieves provable global convergence.

Robust optimization/decision making under uncertainty. Robust optimization and decision making under uncertainty have been longstanding topics in artificial intelligence. For example, utility elicitation (Chajewska et al., 2000; Boutilier et al., 2006) considers the problem of maximizing expected utility under feasibility constraints without precise knowledge of the utility function. A standard approach is to minimize mini-max regret subject to constraints on possible utility functions, thereby achieving worst case robustness to the true underlying utility. Much of the work in this area has considered structured problem formulations using graphical models, to simply the constraints and utility function forms. Here we handle more flexible formulations through neural network parameterizations of the dual functions.

5 EXPERIMENTS

We evaluate the proposed light-SD using two practical problem formulations, namely portfolio optimization and the stochastic optimal transport problem as defined in Section 2.

Baselines: We compare light-SD against a list of prominent algorithms. Most such algorithms convert SD constraints into either a LP (SDLP, cSSD1, cSSD2) or a MIP formulation (cFSD, FDMIP).

- **SDLP** (Dentcheva and Ruszczyński, 2003) handles 2nd-order SD constraints by converting them into linear constraints via sampling. It scales poorly since the number of constraints and variables grows quadratically with respect to the sample size.

- **cSSD1** (Luedtke, 2008) reduces the growth of the number of constraints to a linear scale, but the number of variables still grows quadratically.

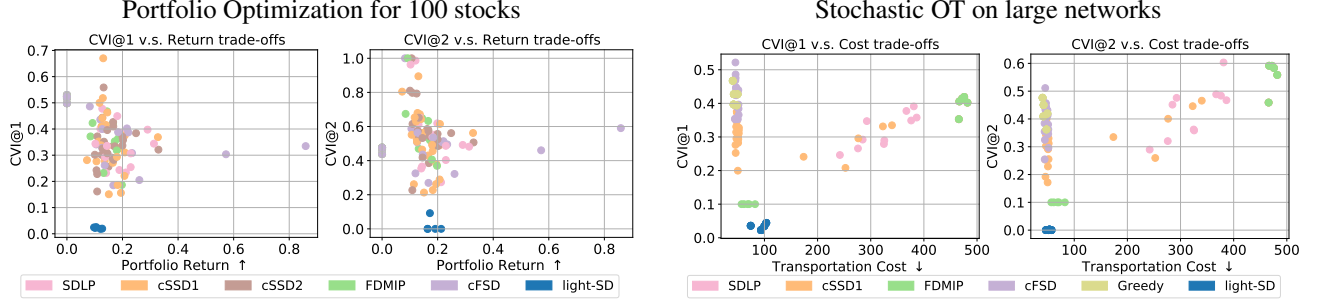- **cSSD2** (Luedtke, 2008) has the same theoretical complex-

Figure 3: Each dot in the above figure represents the objective and CVI of corresponding solution obtained by different methods over multiple random seeds. Generally light-SD achieves better objective value (*i.e.*, higher return or lower cost) and lower CVI compared to alternative methods.
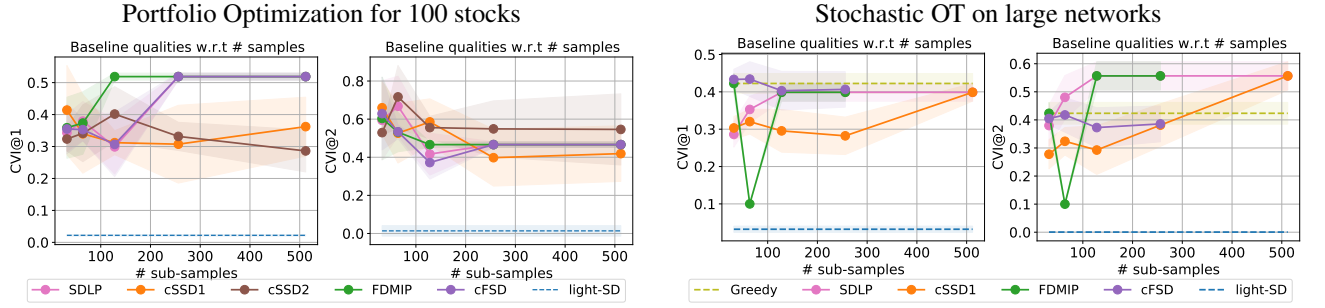

Figure 4: Solution quality w.r.t. different number of samples for baseline methods.

ity as cSSD1 but may yield many fewer nonzero elements in the LP formulation.

- **FDMIP** (Noyan et al., 2006) handles 1st-order SD using a MIP formulation. The number of variables and constraints both grow quadratically. MIPs are also harder to solve than LPs of the same scale.

- **cFSD** (Luedtke, 2008) handles 1st-order SD and improves the MIP formulation by reducing the number of constraints to a linear scale.

- **Greedy** sets a lower-bound for the minimization objective without considering the SD constraints.

Since the 1st-order SD conditions also imply satisfaction of the 2nd-order conditions, the minimization objective achieved by FDMIP or cFSD upper-bounds that of the 2nd-order formulations. Meanwhile, given that the 2nd-order constraints are necessary for the 1st-order constraints, we leverage the 2nd-order formulations as approximations to the 1st-order SD problems, following (Noyan et al., 2006). As a result, we compare all the above methods to light-SD for both 1st and 2nd-order SD constrained optimizations.

Evaluation Metrics: We evaluate algorithm performance by solving the stochastic optimization problem and evaluating their solution quality with respect to both the objective value and the compliance to the stochastic dominance constraints. Specifically we have:

- *Optimality Ratio* (**obj-ratio**) is $\frac{jObjective \; Objective \; j}{Objective}$.

When the optimal solution is known we evaluate how closely a solution approximates the optimal value.

- *Objective*. When the optimal solution is unknown, we can still compare the objective values of different solutions (*i.e.*, the portfolio return or the transportation cost) with higher (lower) value for maximization (minimization) problem indicating better algorithm quality.

- *Constraint Violation Index for k-th order* (**CVI@k**). Based on the fact that $F_X^k(\eta) = \frac{1}{(k \; 1)!} jj \max(0, \eta \; X) jj_k^k \; \frac{1}{1}$ (Prop. 1 (Ogryczak and Ruszczyński, 2001)), we can empirically compare $F_X^k(\eta)$ and $F_Y^k(\eta)$ for a given η. To evaluate how faithful a solution satisfies the k-th order stochastic dominance constraint for $\eta \; 2 \; [a, b]$, we introduce constraint violation index, **CVI@k** as
$$\mathbb{E}_{U(a;b)} \; | \; \mathbb{E} \; (\eta \; X)_+^k \; \frac{1}{} > \mathbb{E} \; (\eta \; Y)_+^k \; \frac{1}{}$$
to measure the degree of constraint violation within $[a, b]$.

For all methods we run for at most 1 hour with CPUs and compare their final solutions. light-SD can be further accelerated with P100 GPUs. For each configuration we run with 5 random seeds and report the corresponding mean and standard deviations of the evaluation metrics. Full experimental results are included in Appendix C.

5.1 Portfolio Optimization

We consider the portfolio optimization problem as in Eq. (6) with two experiment setups.
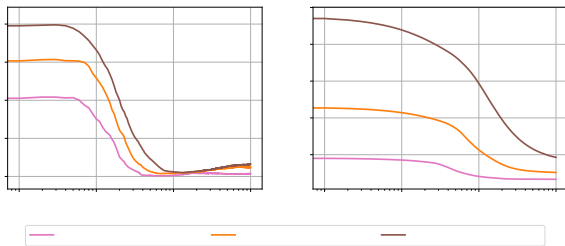
Figure 5: Convergence of light-SD w.r.t CVI and transportation cost under different SD constraints.

Table 2: *Obj-ratio and CVI compared to optimal solution.*

| | light-SD | | 2nd-order LP | |
Constraints	obj-ratio	CVI	obj-ratio	CVI
1st Order	**0.19%**	**5.37%**	3.36%	26.27%
2nd Order	0.00	0.00	0.00	0.00

Table 3: *light-SD running on CPUs or GPUs for 1 hour.*

| | CVI@2 | | Cost | |
	GPU	CPU	GPU	CPU
Small Network	0.04%	0.05%	31.64	31.64
Mid Network	0.00%	0.00%	43.17	43.31
Large Network	0.05%	0.08%	52.35	60.53

Small-Scale Experiment: We first verify the correctness of light-SD using a small-scale portfolio optimization problem identical to the one used in (Dentcheva and Ruszczyński, 2003). In this setup, eight assets including NASDAQ, S&P 500, U.S. long-term bonds, etc, are considered to form a portfolio whose performance is measured in terms of *yearly* returns. The distributions of the yearly returns are estimated using the historical data of these eight assets over 22 years (Table 8.1 in (Dentcheva and Ruszczyński, 2003)).

The reference policy is an equally weighted portfolio with an expected return of 10.6%. In this simple setting we only have 22 observations in total, so both the baseline solvers are able to achieve the optimal solution under the 2nd and 1st SD constraints, respectively. So we primarily evaluate the quality of light-SD and compare with 2nd-order LP relaxation for the 1st-order MIP constraints. Specifically under the 1st and 2nd order constraints, the optimal policy would yield 10.65% and 11.00% expected returns, respectively. We can see in Table 2 that light-SD achieves a better error ratio than the 2nd-order LP approximation for the 1st-order constraints, while being almost perfect in solving with the 2nd-order constraints.

Large-Scale Experiment: For the large-scale experiment, we sample different sets of stocks listed on NASDAQ and consider their daily returns from Jan 1, 2015 to May 1, 2022 for assessment. Concretely, we evaluate portfolios consisting of $\{20, 50, 100\}$ stocks. For each portfolio size setting, we first fit a density model $p(\xi)$ using kernel density estimation, which serves as the distribution for random variable

ξ. We train all the methods using samples from $p(\xi)$, and evaluate against the actual daily returns.

Since the baseline methods only work for finite samples from $p(\xi)$, we vary the number of samples within $\{32, 64, 128, 256, 512\}$ and evaluate the solution against the actual daily returns from Jan 1, 2015. For the constrained optimization, we report the trade-off between objective optimality and constraint satisfaction in Figure 3. From the top row we can see that light-SD achieves almost zero CVI for most random seeds, and relatively high portfolio returns. For some rare cases, the baseline methods yield higher returns but also a much higher CVI, due to the insufficiency of approximating constraints with a finite sample. The computational complexity of the traditional baseline methods generally limits their applicability to optimizing portfolios with more granular returns.

To provide a more intuitive understanding of how good the portfolio obtained by light-SD compared to the reference policy, we further report standard deviation, sharp ratio and largest drawback in Table 4. Overall it indeed improves the return while reduces the risks in terms of the variance and worst case performance. Please refer to Appendix C for more details about the problem settings and the full results.

5.2 Stochastic Optimal Transport

We consider the optimal transport problem as defined in Eq. (7), with three different scales of network structures. In the experiment, demands and supplies are synthetically generated. To make the problem more realistic, we assume $p(\xi)$ is a multi-modal distribution where each mode corresponds to a multivariate Gaussian. The mean configuration of each mode is sampled from a Poisson distribution with mean equals to 10.0 for each region and covariance a random positive-definite matrix. The supply distribution is configured similarly, but with a larger expected supply that can cover the expected total demand.

The general experimental setting is similar to portfolio optimization, with one subtle but major difference – here we are seeking a transport plan where the total demand attributed to each warehouse is *dominated by* the supply at that warehouse. The change of the dominance direction makes little difference for light-SD, but causes major problems for the

baseline approaches. For example, the constraints of cSSD become quadratic instead of linear, thus making cSSD nontrivial to solve. We report the trade-offs between transportation cost and CVI in Figure 3 on large networks with 100 regions/nodes. In almost all cases it appears that light-SD achieves a clear win over the baselines. The baselines either struggle with nding feasible solutions or suffer from high transportation costs. Our hypothesis is that the multimodality of $p(\cdot)$ makes the required number of samples grows signi cantly relative to the single mode case, making the baselines very dif cult to scale up.

5.3 Ef ciency

CPU or GPU: The above results of light-SD are obtained using a P100 GPU. As the LP/MIP baselines cannot leverage the advantages of GPUs, we here include extra results of running light-SD on the same CPU-only machines.

light-SD is able to nish the solving for each of the instances of each portfolio optimization settings in 30 mins. For stochastic OT the computation is more intense, and we report the quality comparison between CPU/GPU based light-SD in Table 3. Overall running light-SD would take a bit longer and thus obtain slightly worse results than on GPU. But the results are still signi cantly better than all other baselines.

Baseline sample ef ciency: We plot the CVI of the baseline methods under different sample sizes (from 32 up to 512) in Figure 4. A maximum of 512 samples is selected as limits modern solvers to take less than 1 hour. light-SD is also plotted as a line in Figure 4 as a comparison, since it is not affected by nite samples. For portfolio optimization, more samples generally lead to better CVI satisfaction in the baseline methods. For the stochastic OT problem, this observation does not hold for growing sample size, due to the dif culty of the multi-modal distribution estimation.

Convergence of light-SD: We visualize the convergence of CVI and the objective value with light-SD in Figure 5. For both problems we can see that light-SD converges in around 10^4 steps with a batch size of 512. Moreover, the stochastic OT problem takes more steps in nding a feasible solution compared to the portfolio optimization problem. This in turn justi es that with nite samples, it is hard for the baseline approaches to nd a near optimal solution.

6 CONCLUSION

We introduce the stochastic dominance concept to the ML community as a principled way to handle uncertainty comparisons. On the other hand, we exploit ML techniques to provide an ef cient new way to handle the constraints, solving a long-standing OR problem. Speci cally, by exploiting stochastic approximation and special dual parametriztion, we bypass the dif culties in the Lagrangian of optimiza-

tion with stochastic dominance constraint, and achieve a simple yet ef cient algorithm. The proposed light-SD is empirically scalable and theoretically guaranteed.

References

Mykel J Kochenderfer. Decision making under uncertainty: theory and application. MIT press, 2015.

Amir Sani, Alessandro Lazaric, and Rémi Munos. Risk-aversion in multi-armed bandits. Advances in Neural Information Processing Systems, 25, 2012.

Marek Petrik and Dharmashankar Subramanian. An approximate solution method for large risk-averse markov decision processes. arXiv preprint arXiv:1210.4901, 2012.

Aviv Tamar, Dotan Di Castro, and Shie Mannor. Temporal difference methods for the variance of the reward to go. In International Conference on Machine Learning, pages 495–503. PMLR, 2013.

Aviv Tamar, Yonatan Glassner, and Shie Mannor. Optimizing the cvar via sampling. In Twenty-Ninth AAAI Conference on Arti cial Intelligence, 2015.

Prashanth La and Mohammad Ghavamzadeh. Actor-critic algorithms for risk-sensitive mdps. Advances in neural information processing systems, 26, 2013.

Erick Delage and Shie Mannor. Percentile optimization for markov decision processes with parameter uncertainty. Operations research, 58(1):203–213, 2010.

Mark Machina and Kip Viscusi. Handbook of the Economics of Risk and Uncertainty. Newnes, 2013.

Henry B Mann and Donald R Whitney. On a test of whether one of two random variables is stochastically larger than the other. The Annals of Mathematical Statistics, pages 50–60, 1947.

Erich Leo Lehmann. Ordered families of distributions. The Annals of Mathematical Statistics, pages 399–419, 1955.

James P Quirk and Rubin Saposnik. Admissibility and measurable utility functions. The Review of Economic Studies, 29(2):140–146, 1962.

Michael Rothschild and Joseph E Stiglitz. Increasing risk: I. a de nition. Journal of Economic theory, 2(3):225–243, 1970.

Michael Rothschild and Joseph E Stiglitz. Increasing risk ii: Its economic consequences. Journal of Economic theory, 3(1):66–84, 1971.

Darinka Dentcheva and Andrzej Ruszczyński. Portfolio optimization with stochastic dominance constraints. Journal of Banking & Finance, 30(2):433–451, 2006.

Yu Marco Nie, Xing Wu, and Tito Homem-de Mello. Optimal path problems with second-order stochastic dominance constraints. Networks and Spatial Economics, 12(4):561–587, 2012.

Darinka Dentcheva and Andrzej Ruszczyński. Stochastic dynamic optimization with discounted stochastic dominance constraints. SIAM journal on control and optimization, 47(5):2540–2556, 2008.

William B Haskell and Rahul Jain. Stochastic dominance constrained markov decision processes. SIAM Journal on Control and Optimization, 51(1):273–303, 2013.

Haim Levy. Stochastic dominance and expected utility: Survey and analysis. Management science, 38(4):555–593, 1992.

W odzimierz Ogryczak and Andrzej Ruszczyński. From stochastic dominance to mean-risk models: Semideviations as risk measures. European journal of operational research, 116(1):33–50, 1999.

W odzimierz Ogryczak and Andrzej Ruszczyński. On consistency of stochastic dominance and mean–semideviation models. Mathematical Programming, 89(2):217–232, 2001.

W odzimierz Ogryczak and Andrzej Ruszczyński. Dual stochastic dominance and related mean-risk models. SIAM Journal on Optimization, 13(1):60–78, 2002.

Darinka Dentcheva and Andrzej Ruszczyński. Optimization with stochastic dominance constraints. SIAM Journal on Optimization, 14(2):548–566, 2003.

Darinka Dentcheva and Andrzej Ruszczyński. Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. Mathematical Programming, 99(2):329–350, 2004a.

Nilay Noyan, Gabor Rudolf, and Andrzej Ruszczyński. Relaxations of linear programming problems with rst order stochastic dominance constraints. Operations Research Letters, 34(6):653–659, 2006.

Gábor Rudolf and Andrzej Ruszczyński. Optimization problems with second order stochastic dominance constraints: duality, compact formulations, and cut generation methods. SIAM Journal on Optimization, 19(3):1326–1343, 2008.

James Luedtke. New formulations for optimization under stochastic dominance constraints. SIAM Journal on Optimization, 19(3):1433–1450, 2008.

Volker Strassen. The existence of probability measures with given marginals. The Annals of Mathematical Statistics, 36(2):423–439, 1965.

Thierry Post. Empirical tests for stochastic dominance ef ciency. The Journal of Finance, 58(5):1905–1931, 2003.

Benjamin Armbruster and Erick Delage. Decision making under uncertainty when preference information is incomplete. Management science, 61(1):111–128, 2015.

Jian Hu, Tito Homem-de Mello, and Sanjay Mehrotra. Sample average approximation of stochastic dominance constrained programs. Mathematical programming, 133(1): 171–201, 2012.

William B Haskell and Alejandro Toriello. Modeling stochastic dominance as in nite-dimensional constraint systems via the strassen theorem. Journal of Optimization Theory and Applications, 178(3):726–742, 2018.

Darinka Dentcheva and Andrzej Ruszczyński. Optimization with multivariate stochastic dominance constraints. Mathematical Programming, 117(1):111–127, 2009.

Darinka Dentcheva and Andrzej Ruszczyński. Semi-in nite probabilistic optimization: rst-order stochastic dominance constrain. Optimization, 53(5-6):583–601, 2004b.

J. Von Neumann and O. Morgenstern. Theory of Games and Economic Behavior. Princeton classic editions. Princeton University Press, 2004. ISBN 9780691119939.

Joseph Sill. Monotonic networks. Advances in neural information processing systems, 10, 1997.

Charles Dugas, Yoshua Bengio, François Bélisle, Claude Nadeau, and René Garcia. Incorporating functional knowledge in neural networks. Journal of Machine Learning Research, 10(6), 2009.

Antoine Wehenkel and Gilles Louppe. Unconstrained monotonic neural networks. Advances in neural information processing systems, 32, 2019.

Andrew R Barron. Universal approximation bounds for superpositions of a sigmoidal function. IEEE Transactions on Information theory, 39(3):930–945, 1993.

BT Poljak and Ya Z Tsypkin. Pseudogradient adaptation and training algorithms. Automation and remote control, 34:45–67, 1973.

Yingxiang Yang, Haoxiang Wang, Negar Kiyavash, and Niao He. Learning positive functions with pseudo mirror descent. Advances in Neural Information Processing Systems, 32, 2019.

Maher Nouiehed, Maziar Sanjabi, Tianjian Huang, Jason D Lee, and Meisam Razaviyayn. Solving a class of nonconvex min-max games using iterative rst order methods. Advances in Neural Information Processing Systems, 32, 2019.

Yifan Hu, Xin Chen, and Niao He. On the bias-variance-cost tradeoff of stochastic optimization. Advances in Neural Information Processing Systems, 34:22119–22131, 2021.

Yurii Nesterov and Arkadii Nemirovskii. Interior-point polynomial algorithms in convex programming. SIAM, 1994.

Elias Khalil, Hanjun Dai, Yuyu Zhang, Bistra Dilkina, and Le Song. Learning combinatorial optimization algorithms over graphs. *Advances in neural information processing systems*, 30, 2017.

Irwan Bello, Hieu Pham, Quoc V Le, Mohammad Norouzi, and Samy Bengio. Neural combinatorial optimization with reinforcement learning. *arXiv preprint arXiv:1611.09940*, 2016.

Nikolaos Karalias and Andreas Loukas. Erdos goes neural: an unsupervised learning framework for combinatorial optimization on graphs. *Advances in Neural Information Processing Systems*, 33:6659–6672, 2020.

Zhuwen Li, Qifeng Chen, and Vladlen Koltun. Combinatorial optimization with graph convolutional networks and guided tree search. *Advances in neural information processing systems*, 31, 2018.

Hanjun Dai, Yuan Xue, Zia Syed, Dale Schuurmans, and Bo Dai. Neural stochastic dual dynamic programming. *arXiv preprint arXiv:2112.00874*, 2021.

Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25, 2012.

Theodore P Papalexopoulos, Christian Tjandraatmadja, Ross Anderson, Juan Pablo Vielma, and David Belanger. Constrained discrete black-box optimization using mixed-integer programming. In *International Conference on Machine Learning*, pages 17295–17322. PMLR, 2022.

Urszula Chajewska, Daphne Koller, and Ronald Parr. Making rational decisions using adaptive utility elicitation. In *AAAI*, pages 363–369, 2000.

Craig Boutilier, Relu Patrascu, Pascal Poupart, and Dale Schuurmans. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence*, 170(8-9):686–713, 2006.

J Frédéric Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems*. Springer Science & Business Media, 2013.

Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. MIT press, 2018.

Appendix

A   Existing Stochastic Dominance Reformulations

We review the existing representative reformulations of stochastic dominance from distribution function, utility function and Strassen theorem, respectively. We mainly focus on first-/second-order stochastic dominance constraints with discrete random variables as the original papers specified.

Denote $g(z; \theta) = W$, and the support for W and Y is $\{g(z; \theta_i) = w_i\}_{i=1}^{M}$ and $\{y_k\}_{k=1}^{K}$. Without loss of generality, we assume $\{y_k\}_{k=1}^{K}$ are ordered $y_1 \leqslant y_2 \leqslant \cdots \leqslant y_K$. We denote the distributions of W and Y are described as
$$P(W = w_i) = p_i; \quad \text{and} \quad P(Y = y_k) = q_k.$$

Distribution function reformulation: Based on the stochastic dominance condition definition through 2nd distribution functions, Dentcheva and Ruszczyński (2004a) reformulates

$$W \succeq_2 Y \iff \exists s \in R_+^{MK}; \text{ s.t.:} \quad \begin{array}{l} \sum_{i=1}^{M} p_i s_{ik} \leqslant \sum_{j=1}^{K} q_j (y_k - y_j)_+; \quad k = 1, \ldots, K \\ s_{ik} + w_i \geqslant y_k; \quad i = 1, \ldots, M; k = 1, \ldots, K \end{array} \tag{22}$$

Similarly, Noyan et al. (2006) reformulates first order stochastic dominance as

$$W \succeq_1 Y \iff \exists \beta \in \{0,1\}^{MK}; \text{ s.t.:} \quad \begin{array}{l} \sum_{i=1}^{M} p_i \beta_{ik} \leqslant \sum_{j=1}^{k-1} q_j; \quad k = 1, \ldots, K \\ w_i + C \beta_{ik} \geqslant y_k; \quad i = 1, \ldots, M; k = 1, \ldots, K \end{array} \tag{23}$$

where C is a sufficient large scalar.

Utility function reformulation: In Armbruster and Delage (2015), the property that the nondecreasing and convex function for discrete variables are piece-wise linear functions has been exploited, then, the second order stochastic dominance can be reformulated as

$$W \succeq_2 Y \iff \exists (v; u) \in R^M; (\lambda; \mu) \in R^K; \text{ s.t.:} \left\{ \begin{array}{l} \sum_i p_i (v_i w_i + u_i) \geqslant \sum_k q_k \lambda_k \\ y_k v_i + u_i \geqslant \lambda_j; \, \forall i = 1, \ldots, M; k = 1, \ldots, K \\ (\mu_{k+1} - \mu_k) \geqslant \mu_{k+1} (y_{k+1} - y_k); \\ (\lambda_{k+1} - \lambda_k) \leqslant \mu_k (y_{k+1} - y_k); \, \forall k = 1, \ldots, K \\ (v; \mu) \geqslant 0 \end{array} \right. \tag{24}$$

We emphasize that this utility function reformulation of stochastic dominance is highly related to the proposed method. However, the major difference is that we introduce neural network to parametrize the dual function space, which enables the efficient stochastic gradient calculation, while the reformulation (24) is nonparametric, relying on all the samples in the dataset.

Strassen theorem reformulation: Luedtke (2008) exploits the Strassen theorem, which leads to several equivalent reformulations of first and second stochastic dominance. We list the one representative for each condition below, respectively,

$$W \succeq_2 Y \iff \exists \pi \in R_+^{MK}; \text{ s.t.:} \left(\begin{array}{l} \sum_{k=1}^{K} y_j \pi_{ik} \leqslant w_i; \sum_{k=1}^{K} \pi_{ik} = 1; i = 1, \ldots, M \\ \sum_{i=1}^{M} p_i \pi_{ik} = q_k; k = 1, \ldots, K \end{array} \right) \tag{25}$$

$$W \succeq_1 Y \iff \exists \pi \in \{0,1\}^{MK}; \text{ s.t.:} \left(\begin{array}{l} \sum_{k=1}^{K} y_j \pi_{ik} \leqslant w_i; \sum_{k=1}^{K} \pi_{ik} = 1; i = 1, \ldots, M \\ \sum_{i=1}^{M} p_i \sum_{j=1}^{k-1} \pi_{ik} = \sum_{j=1}^{k-1} q_j; k = 2, \ldots, K \end{array} \right) \tag{26}$$

12

B Details of Proofs

In this section, for completeness, we provide the proofs for Theorem 1 from (Dentcheva and Ruszczyński, 2004b) and Lagrangian for high-order, from which Theorem 2 is a corollary from (Dentcheva and Ruszczyński, 2003). The major proofs of these two theorems are largely the same, only the conditions and assumptions are different.

In fact, the neural parametrization we proposed can be inspired by the proofs of these theorems.

B.1 Proof for Theorem 1

We first specify the assumptions with the full theorem.

Assumption 1. There is an optimal solution z to (5).

Assumption 2. There exists $z^0 \in \mathcal{Z}$, such that

$$\max_{\eta \in [a;b]} \left\{ F(g(z;\omega);\eta) + \int_a^\eta D_z \mu_z(\eta) d\eta - F(Y;\eta) \right\} < 0.$$

Assumption 3. The reference variable Y has continuous CDF.

Theorem 4 (Theorem 1 (Dentcheva and Ruszczyński, 2004b)) Denote the probability density $g(z;\omega)$ as $\mu_z(\eta)$, which is continuously differentiable w.r.t z and its derivative $D_z \mu_z(\eta)$ is bounded, under Assumption 1, 2, and 3 with $k = 1$, we can represent the Lagrangian of (5) as (10) where $u(\eta) \in U_1$, where U_1 is the set of nondecreasing and left continuous functions over $[a;b]$.

Proof. Under the Assumptions, the KKT optimality condition is that there exists a non-negative $\mu \in rca([a;b])$ and $\eta_0 \geq 0$, such that

$$\langle D_z \mu(z;\omega); x - z \rangle = 0; \quad \forall x \in \mathcal{Z}; \quad \int_a^b (F(g(z;\omega);\eta) - F(Y;\eta)) d\mu(\eta) = 0:$$

We extend the measure μ to the whole real line by setting 0 outside of $[a;b]$, then, we have

$$\int_a^b F(Y;\eta) d\mu(\eta) = \int_a^b P(Y \leq \eta) d\mu(\eta) = \int_a^b \int_{-\infty}^\eta dP_Y(y) d\mu(\eta)$$
$$= \int_{-\infty}^b \int_y^b d\mu(\eta) dP_Y(y) = \int_{-\infty}^b \mu([y;b]) dP_Y(y) = \mathbb{E}[\mu([y;b])];$$

where the second line of the equation comes from Fubini's theorem.

Define $u(\eta) = \mu([\eta;b])$, due to the non-negativity of μ, the $u(\eta)$ in U_1. Therefore, we conclude that $\ell(z;\omega) = f(z) + \int_a^b [F_1(g(z;\omega);\eta) - F_1(Y;\eta)] d\mu(\eta)$ can be reformulated as $\ell_s(z;u)$ in (10) without loss of optimality.

$\square$

Remark (Surrogate with the augmented probability measure dual for U_1): For each positive $\mu(\eta) \in rca([a;b])$, we can define the normalized $\hat{\mu}(\eta) = \frac{\mu(\eta)}{\int \mu(\eta)d\eta} \in \Delta([a;b])$, which is a probability distribution. We augment the probability measure space with zero measure, denoted as $\Delta_0([a;b]) := \Delta([a;b]) \cup 0$. It is straightforward to check that if the non-negative $\mu(\eta)$ satisfies the KKT conditions, we can find $\hat{\mu}(\eta) = \begin{cases} \frac{\mu(\eta)}{\int \mu(\eta)d\eta}; & \mu(\eta) \neq 0 \\ 0; & \mu(\eta) = 0 \end{cases}$ such that $\hat{\mu}(\eta) \in \Delta_0([a;b])$ also satisfies KKT conditions. Specifically, we have the surrogate objective as

$$\min_{\hat{\mu}(\eta) \in \Delta([a;b]) \cup 0} \ell(z;\hat{\mu}) = f(z) + \lambda \int_a^b (F_k(g(z;\omega);\eta) - F_k(Y;\eta)) d\hat{\mu}(\eta) = \begin{cases} f(z); & \text{if } F_1(g(z;\omega);\eta) \leq F_1(Y;\eta); \quad \forall \eta \\ f(z) + \lambda C(z); & \text{o.t.} \end{cases};$$

where $C(z) > 0$ is a function of $\hat{\mu}_z(\eta)$, indicating the level of violation of stochastic dominance constraint at z, and $\lambda > 0$ is the weight. If $\|f\|_{k_1} \leq C$, we set $\lambda > \frac{2C}{\varepsilon}$, we always have

$$f(z) + \lambda C(z) \geq f(z^*); \quad \forall z \text{ s.t.: } C(z) > \varepsilon;$$

which means we relax the optimization by increasing the feasible set of stochastic dominance constraints with ε.

We will consider this relaxed problem instead the original one as a surrogate. The suboptimality gap of the relaxed problem solution is controlled by ε. Without loss of generality, we set $\varepsilon = 1$ for simplicity in our discussion. Practically, the ε can be tuned.

B.2 Proof for Theorem 2

We first specify the Lagrangian for high-order dominance constraint.

Theorem 5 (Theorem 7.1 Dentcheva and Ruszczyński (2003)) Under the assumption that $\exists z \in Z$ such that $\inf_{\eta \in [a;b]} (F_k(Y;\eta) - F_k(g(z;\theta);\eta)) > 0$, for $k > 2$, we can represent the Lagrangian (5) as (10) with $u(\cdot) \in U_k$, where

$$U_k = \left\{ u(\cdot) : [a;b] \to R \mid u^{(k-1)}(t) = (-1)^k \mu(t) \right\};$$

where $u^{(i)}$ denotes the ith derivative and $\mu(\cdot)$ is a non-negative, nonincreasing, left-continuous, and bounded function.

Proof. We introduce the Lagrangian for the optimization (5) with infinite constraints as

$$\mathcal{L}(z;\mu) = f(z) - \int_a^b (F_k(g(z;\theta);\eta) - F_k(Y;\eta)) d\mu(\eta);$$

where $\mu \in rca([a;b])$. Under the assumptions, by the KKT condition in abstract space (Bonnans and Shapiro, 2013, Theorem 3.4), there exists a non-negative measure $\mu \in rca([a;b])$ such that

$$\mathcal{L}(z^\star;\mu) = \max_{z \in Z} \mathcal{L}(z;\mu); \quad \int_a^b F_k(g(z;\theta);\eta) - F_k(Y;\eta) d\mu(\eta) = 0:$$

We extend the measure μ to the whole real line by setting $\mu = 0$ outside of $[a;b]$, then, we have

$$\int_a^b F_k(Y;\eta) d\mu(\eta) = \int_a^b \int_{-\infty}^\eta F_{k-1}(Y;t) dt\, d\mu(\eta) \tag{27}$$

$$= \int_{-\infty}^b \int_t^b d\mu(\eta) F_{k-1}(Y;t) dt = \int_{-\infty}^b \mu([t;b]) F_{k-1}(Y;t) dt: \tag{28}$$

Denote $u(\cdot) \in U_k$ such that $u^{(k-1)}(t) = (-1)^k \mu([t;b])$, $u^{(i)}(b) = 0$, for $i = 1;\ldots;k-2$, then, we can rewrite

$$\int_a^b F_k(Y;\eta) d\mu(\eta) = (-1)^k \int_{-\infty}^b F_{k-1}(Y;t) du^{(k-1)}(t):$$

We further apply integration by parts $k-1$ times,

$$(-1)^k \int_{-\infty}^b F_{k-1}(Y;t) du^{(k-1)}(t) = \int_{-\infty}^b u(t) dF(Y;t) = E[u(Y)] + C;$$

Therefore, we conclude the Lagrangian can be reformulated as $\mathcal{L}(z,u)$ in (10) with $u_k \in U_k$ without loss of optimality. $\square$

Proof of Theorem 2. Theorem 2 can be directly obtained from Theorem 5 with $k = 2$. $\square$

Remark (Surrogate with the augmented probability measure dual for U_2): Similarly, we also define the normalized $\tilde{\mu}(\cdot) \in \mathcal{P}([a;b])$, which is a probability distribution, for each positive $\mu \in rca([a;b])$, and keep $\tilde{\mu}(\cdot) = 0$ if $\mu(\cdot) = 0$ and the KKT conditions still preserved for the corresponding $\mu \in \mathcal{P}([a;b])$. We consider

$$\min_{\tilde{\mu}(\cdot) \in \mathcal{P}([a;b]), \lambda \geq 0} \ell(z;\tilde{\mu}) = \begin{cases} f(z); & \text{if } F_2(g(z;\theta);\eta) \leq F_2(Y;\eta); \ \forall \eta \\ f(z) - \lambda C(z); & \text{o.t.} \end{cases};$$

where $C(z) > 0$ is a function of $\tilde{\mu}_z(\cdot)$, indicating the level of violation of stochastic dominance constraint at z, and $\lambda > 0$ is the weight. Similar to the relaxation of the first-order stochastic dominance constraint, we also obtain an approximate solution whose suboptimality gap is controlled by λ. Without loss of generality, we set $\lambda = 1$ for simplicity in our discussion. Practically, the λ can be tuned.

Remark (High-order dual parametrization): Since $\mu > 0$, the $(k-1)$st derivative of u is monotone. We can exploit the proposed monotonic neural network U_m for modeling u^{k-1}, therefore, we have

$$u(x) = \int_a^x \int_a^{x_{k-2}} \cdots \int_a^{x_{k-1}} u_m(t) dt + \cdots dx_{k-2} \cdots dx_1:$$

As we discussed in Section 3.2.3, one can exploit $|\cdot|$ or softplus to obtain a non-negative neural network for $u_m(t)$ parametrization. We emphasize that the obtain non-negative neural network is bounded for learnability, but implicitly introduces approximation error comparing to the unbounded dual space.

Remark (Trade-off in sample complexity and approximation error): Based on our discussion for the parametrization of U_1, U_2 and U_k, the scale of the dual function is either controlled by ν or the neural network parametrization, which balances the tradeoff between learnability and approximation error. Specifically, with bounded dual function parametrization, we can achieve efficient sample complexity for some approximation solution; while with unbounded dual function parametrization, the approximation error becomes zero, but the problem becomes unlearnable, i.e., the sample complexity is infinite. We can always tune in practice to achieve the delicate balance.

B.3 Proof for Theorem 3

Proof. We prove the claims sequentially.

Concavity. The concavity of $L(z; u)$ w.r.t. z is straightforward by applying the composition rule. Specifically, U_2 is always nondecreasing and concave, therefore $u(g(z; \cdot))$ is concave w.r.t. z. Obviously, $L(z; u) = f(z) + E[u(g(z; \cdot))] - E_y[u(y)]$ is concave.

Global Convergence. Recall the notations,

$$L(z; u) = f(z) + E[u(g(z; \cdot))] - E_y[u(y)];$$

$$\hat{L}(z; u) = f(z) + \frac{1}{N}\sum_{i=1}^{N}(u(g(z; \cdot_i))) - u(y_i));$$

for coherent with literature, we define the

$$\ell(z; u) = -L(z; u); \quad \text{and} \quad \hat{\ell}(z; u) = -\hat{L}(z; u):$$

Then, the problem is equivalently considering

$$\min_z \max_{u\in U_2} \ell(z; u); \quad \text{and} \quad \min_z \max_{u\in U_2} \hat{\ell}(z; u):$$

We denote

$$\ell(z) = \max_{u\in U_2} \ell(z; u); \quad \text{with } z^* = \arg\min_z \ell(z); \quad u^* = \arg\max_{u\in U_2} \ell(z^*; u)$$

and

$$\hat{\ell}(z) = \max_{u\in U_2} \hat{\ell}(z; u) \quad \text{with } \hat{u}_z = \arg\max_{u\in U_2} \hat{\ell}(z; u):$$

We also denote $A_t = \frac{1}{2}\|z_t - z^*\|_2^2$, and $a_t = E[A_t]$, then we have the recursion as

$$\begin{aligned}
A_{t+1} &= \frac{1}{2}\|z_{t+1} - z^*\|_2^2 = \frac{1}{2}\left\|z_t - \eta_t \nabla_z \hat{\ell}(z_t) - z^*\right\|_2^2 \\
&= A_t + \frac{1}{2}\eta_t^2 \left\|\nabla_z \hat{\ell}(z_t)\right\|_2^2 - \eta_t \nabla_z \hat{\ell}(z_t)^\top (z_t - z^*):
\end{aligned} \tag{29}$$

By convexity of $\hat{\ell}(z)$, we have

$$\nabla_z \hat{\ell}(z_t)^\top (z_t - z^*) \geq \hat{\ell}(z_t) - \hat{\ell}(z^*): \tag{30}$$

Combining with (29), this implies that

$$E[\ell(z_t) - \ell(z^*)] \leq \frac{a_t - a_{t+1}}{\eta_t} + \frac{\eta_t}{2}E\left\|\nabla_z \hat{\ell}(z_t)\right\|_2^2 + \underbrace{E\left[\hat{\ell}(z_t) - \ell(z_t)\right]}_{h} + \underbrace{E\left[\ell(z^*) - \hat{\ell}(z^*)\right]}_{i}: \tag{31}$$

Note that for any z, $E\left[\ell(z) - \hat{\ell}(z)\right] \leq \epsilon_{stat}(z) := E\left[\sup_{u\in U_2}|\hat{\ell}(z; u) - \ell(z; u)|\right]$. Let $\epsilon_{stat} := \sup_z \epsilon_{stat}(z)$.

By telescoping the sum and invoking convexity of $\ell(z)$, we further have

$$E[\ell(z_T) - \ell(z^*)] \leq \frac{1}{T}\sum_{i=1}^{T} E[\ell(z_t) - \ell(z^*)] \tag{32}$$

$$\leq \frac{1}{T}\sum_{i=1}^{T}\frac{\eta_t}{2}E\left\|\nabla_z \hat{\ell}(z_t)\right\|_2^2 + 2\epsilon_{stat} + \frac{1}{T}\sum_{i=1}^{T} a_t\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t+1}}\right) + \frac{a_1}{T\eta_1}: \tag{33}$$

where $z_T = \frac{1}{T}\sum_{i=1}^{T} z_t$. Setting stepsize $\eta_t = O\left(\frac{1}{\sqrt{T}}\right)$, and under the assumption that $\left\|\nabla_z \hat{\ell}(z)\right\|_2 \leq C_f + C_g$, we achieve

the convergence rate,

$$E[\ell(z_T) - \ell(z^*)] \le \frac{1}{T}\sum_{i=1}^{T}(2\epsilon_{stat}) + \frac{a_1}{\sqrt{T}} + \frac{C_f + C_g}{\sqrt{2T}}. \tag{34}$$

The statistical error can be bound in standard way. We first calculate the Rademacher complexity of U_2 which is formed as the convex combination of ReLU functions, i.e.,

$$U_2 = \{u(x) = E_{\sigma(\cdot)}(\sigma - x)_+ ; \sigma(\cdot) \in ([\cdot; a; b])\}.$$

The empirical Rademacher complexity of U_2 with N samples, denoted as $R_N(U_2)$, is equivalent to the that of ReLU (Mohri et al., 2018, Lemma 7.4), i.e.,

$$R_N(U_2) = \frac{4(|a| + |b|)}{\sqrt{N}};$$

which immediately implies that with probability γ (Mohri et al., 2018, Theorem 3.3),

$$\epsilon_{stat} = O\left(\frac{(|a| + |b|) + \sqrt{\log(1/\gamma)}}{\sqrt{N}}\right).$$

Combining the ϵ_{stat} with (34), we conclude the proof that with probability γ, we have

$$E[\ell(z_T) - \ell(z^*)] = O\left(\frac{(|a| + |b|) + \sqrt{\log(1/\gamma)}}{\sqrt{N}} + \sqrt{\frac{1}{T}}\right). \tag{35}$$

$\square$

C   More experimental details

C.1   Large scale portfolio optimization setup

We collect the daily stock prices of over 1,500 stocks listed on NASDAQ during Jan 1, 2015 - May 1, 2022. Then we calculate the daily return (i.e., the relative gain/loss of each day compared to the day before) of each stock, and compute the mean/standard deviations of the returns per stock. We select these stocks in a way that no stock would dominate the other (i.e., no stock would have higher mean return and lower standard deviation than other stocks), so as to avoid trivial solutions.

In reality we would only have a estimated distribution of future returns, so to evaluate the quality of different policies using historical replay, we first fit a generative distribution over the returns of selected stocks, then use the actual daily return as the evaluation set to report the mean return and violation index (CVI) of stochastic dominance. For simplicity we use KDE with factorized Gaussian distribution assumption, where each dimension has a standard deviation of 0.01 (given that the mean return is around 0.1%).

C.2   Full experimental results

In the main paper we present the experimental results for the largest problem configurations. Here for the completeness, we include the results of both 1st and 2nd SD constrained optimizations for all the problem scales.

Objective / CVI trade-offs   Figure 6 and Figure 7 shows scatter plot of objective / CVI trade-offs under 1st/2nd order stochastic dominance constraints, respectively. We can see light-SD achieved almost perfect CVI, while also retaining low transportation cost (which is slightly higher than the lowerbound obtained by greedy-based approaches, where the greedy one totally ignores the SD constraints).

Sample complexity for baseline methods Figure 8 and Figure 9 shows that the LP/MIP based approaches are not sample efficient enough. Compared to the second-order cases, first-order stochastic dominance is even harder (especially given that the LP based approach are just solving relaxation of the first-order SD constraints).

Convergence behaviors Figure 10 and Figure 11 shows the convergence of light-SD with first/second order SD constraints.

Evaluation on more metrics for the Portfolio Optimization setting

Portfolio Optimization

Stochastic OT

Figure 6: Each dot in the above figure represents the objective and CVI@1 (for first-order stochastic dominance constraints) of corresponding solution obtained by different methods over multiple random seeds. Generally light-SD achieves better objective value (i.e., higher return or lower cost) and lower CVI compared to alternative methods.

Portfolio Optimization

Stochastic OT

Figure 7: Each dot in the above figure represents the objective and CVI@2 of corresponding solution obtained by different methods over multiple random seeds. Generally light-SD achieves better objective value (i.e., higher return or lower cost) and lower CVI compared to alternative methods.
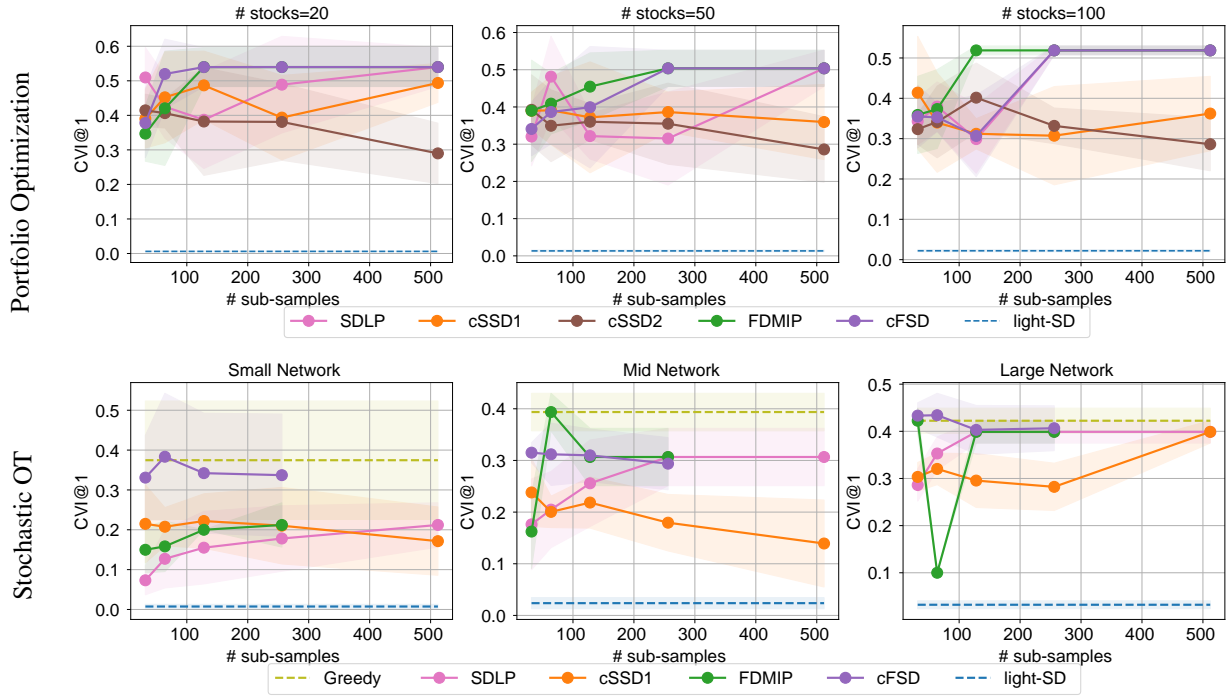
Figure 8: Solution quality w.r.t different number of samples for baseline methods on optimization with first order stochastic dominance constraints.
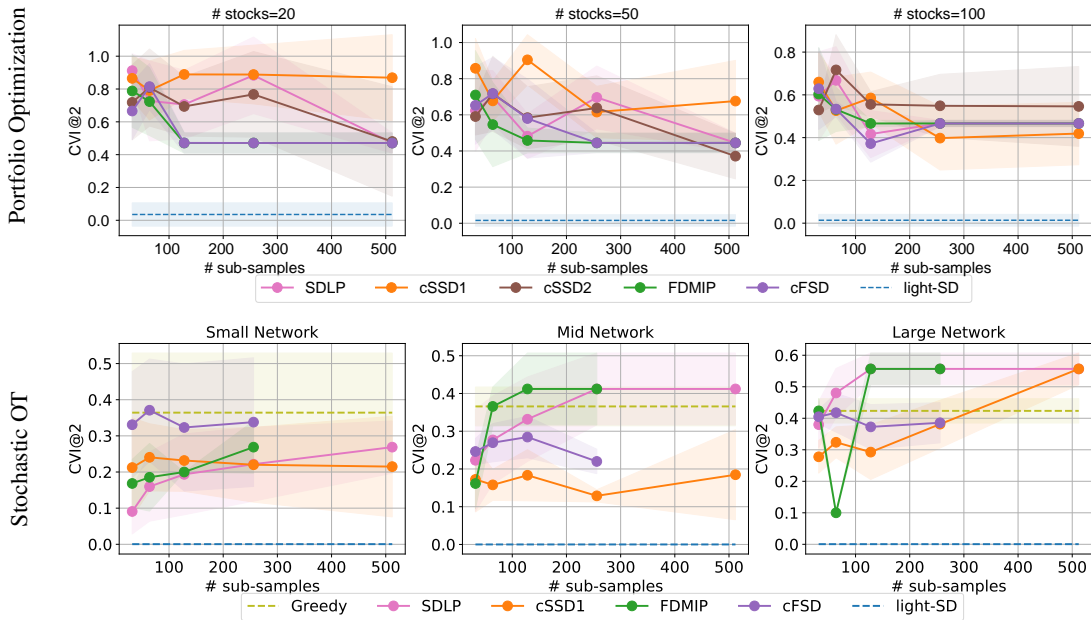

Figure 9: Solution quality w.r.t different number of samples for baseline methods on optimization with second order stochastic dominance constraints.
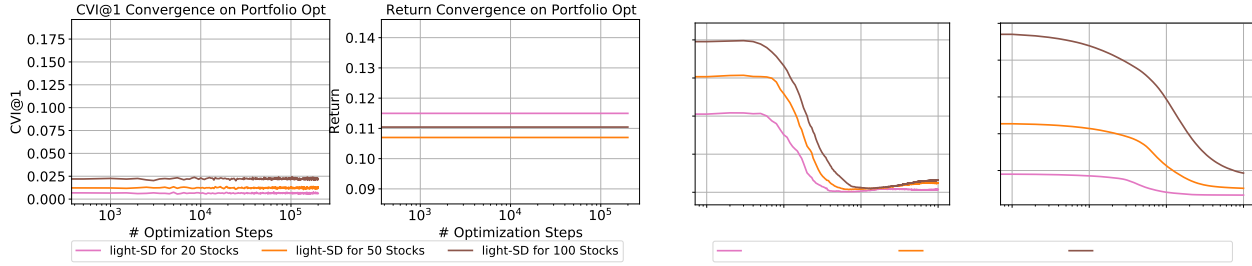
Figure 10: Convergence of light-SD w.r.t CVI@1 and objectives under different problem settings.
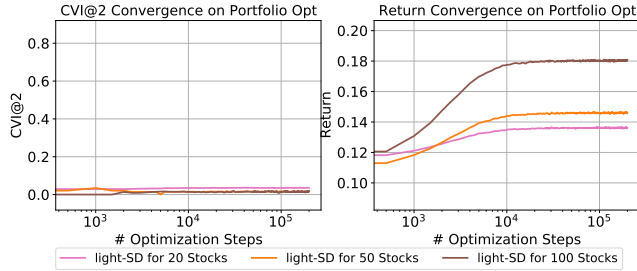

Figure 11: Convergence of light-SD w.r.t CVI@2 and objectives under different problem settings.

More evaluation metrics for portfolio optimization In Table 4 we present more metrics in terms of the performance of portfolio optimization. As the goal of optimization under stochastic dominance constraints is to minimize the risks, we provide more intuitive metrics in Table 4 to demonstrate the effectiveness of light-SD. We can see that overall it reduces the variance of the return, and improve the worst case performance. These are indicators of policies with lower risks. Both the expected returns and Sharpe ratios are also improved.

Table 4: Comparing light-SD against reference policy for portfolio optimization, with more evaluation metrics.

	Standard Deviation		SharpeRatio		Largest Drawback	
	Reference	light-SD	Reference	light-SD	Reference	light-SD
10-stocks	1.67	**1.61**	1.25	**1.29**	-9.98	**-9.70**
20-stocks	1.38	**1.34**	1.23	**1.51**	-9.74	**-9.01**
50-stocks	1.23	**1.16**	1.38	**1.90**	-8.94	**-7.94**
100-stocks	**1.12**	1.16	1.48	**2.34**	-8.89	**-7.49**