Numerical Methods for Convex Multistage Stochastic Optimization

Other titles in Foundations and Trends® in Optimization

A Tutorial on Hadamard Semidifferentials
Kenneth Lange
ISBN: 978-1-63828-348-5

Massively Parallel Computation: Algorithms and Applications
Sungjin Im, Ravi Kumar, Silvio Lattanzi, Benjamin Moseley and Sergei Vassilvitskii
ISBN: 978-1-63828-216-7

Acceleration Methods
Alexandre d'Aspremont, Damien Scieur and Adrien Taylor
ISBN: 978-1-68083-928-9

Atomic Decomposition via Polar Alignment: The Geometry of Structured Optimization
Zhenan Fan, Halyun Jeong, Yifan Sun and Michael P. Friedlander
ISBN: 978-1-68083-742-1

Optimization Methods for Financial Index Tracking: From Theory to Practice
Konstantinos Benidis, Yiyong Feng and Daniel P. Palomar
ISBN: 978-1-68083-464-2

Numerical Methods for Convex Multistage Stochastic Optimization

Guanghui Lan
Georgia Institute of Technology
george.lan@isye.gatech.edu

Alexander Shapiro
Georgia Institute of Technology
ashapiro@isye.gatech.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Optimization

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

G. Lan and A. Shapiro. *Numerical Methods for Convex Multistage Stochastic Optimization*. Foundations and Trends® in Optimization, vol. 6, no. 2, pp. 63–144, 2024.

ISBN: 978-1-63828-351-5

Foundations and Trends® in Optimization

Volume 6, Issue 2, 2024

Editorial Scope

Foundations and Trends® in Optimization publishes survey and tutorial articles in the following topics:

- algorithm design, analysis, and implementation (especially, on modern computing platforms
- models and modeling systems, new optimization formulations for practical problems
- applications of optimization in machine learning, statistics, and data analysis, signal and image processing, computational economics and finance, engineering design, scheduling and resource allocation, and other areas

Information for Librarians

Foundations and Trends® in Optimization, 2024, Volume 6, 4 issues. ISSN paper version 2167-3888. ISSN online version 2167-3918. Also available as a combined paper and online subscription.

Contents

Numerical Methods for Convex Multistage Stochastic Optimization

Guanghui Lan and Alexander Shapiro

Georgia Institute of Technology, USA; george.lan@isye.gatech.edu, ashapiro@isye.gatech.edu

ABSTRACT

Optimization problems involving sequential decisions in a stochastic environment were studied in Stochastic Programming (SP), Stochastic Optimal Control (SOC) and Markov Decision Processes (MDP). In this monograph, we mainly concentrate on SP and SOC modeling approaches. In these frameworks, there are natural situations when the considered problems are convex. The classical approach to sequential optimization is based on dynamic programming. It has the problem of the so-called "curse of dimensionality", in that its computational complexity increases exponentially with respect to the dimension of state variables. Recent progress in solving convex multistage stochastic problems is based on cutting plane approximations of the cost-to-go (value) functions of dynamic programming equations. Cutting plane type algorithms in dynamical settings is one of the main topics of this monograph. We also discuss stochastic approximation type methods applied to multistage stochastic optimization problems. From the computational complexity point of view, these two types of methods seem to be complementary to each other. Cutting plane type methods can handle multistage problems with a large number of stages

Guanghui Lan and Alexander Shapiro (2024), "Numerical Methods for Convex Multistage Stochastic Optimization", Foundations and Trends® in Optimization: Vol. 6, No. 2, pp 63–144. DOI: 10.1561/2400000044.

but a relatively smaller number of state (decision) variables. On the other hand, stochastic approximation type methods can only deal with a small number of stages but a large number of decision variables.

Keywords: Stochastic programming, Stochastic optimal control, Markov decision process, Dynamic programming, Risk measures, Stochastic dual dynamic programming, Stochastic approximation method, Cutting plane algorithm.

AMS subject classifications: 65K05, 90C15, 90C39, 90C40.

1

Introduction

Traditionally different communities of researchers dealt with optimization problems involving uncertainty, modeled in stochastic terms, using different terminology and modeling frameworks. In this respect we can point to the fields of Stochastic Programming (SP), Stochastic Optimal Control (SOC) and Markov Decision Processes (MDP). Historically the developments in SP on the one hand, and SOC and MDP on the other, went along different directions with different modeling frameworks and solution methods. SOC is an interesting model since it can not only be naturally written in the MDP terms, but also can be formulated in the SP framework. In this monograph we mainly concentrate on SP approaches, and often specialize them to SOC whenever possible to demonstrate some basic ideas that can potentially bridge these three communities.

In these modeling frameworks mentioned above, there exist some natural situations when the considered problems are *convex*. An optimization problem is said to be convex if both its objective function and feasible set are convex. It is well-known that convexity provides the main apparatus for the development of efficient numerical algorithms for continuous optimization [45], [46]. The main goal of this work is to present some recent developments in numerical approaches to solve convex optimization problems involving sequential decision making. Note that we do not intend to give a comprehensive review of the subject with a complete list of references. *Rather the aim is to present a certain point of view about some recent developments in solving* convex *multistage stochastic programming problems.*

Stochastic Programming (SP) has a long history. Two stage stochastic programming (with recourse) was introduced in Dantzig [10] and Beale [2], and was intrinsically connected with linear programming. From the beginning SP aimed at numerical solutions. Until about twenty years ago, the modeling approach to two and multistage SP was predominately based on construction of scenarios represented by scenario trees. This approach allows one to formulate the so-called deterministic equivalent optimization problem with the number of decision variables more or less proportional to the number of scenarios. When the deterministic equivalent could be represented as a linear program, such problems were considered to be numerically solvable. Because of that, the topic of SP was often viewed as a large scale linear programming. Further discussion and development of this approach can be found in Birge [6] and references therein.

From the point of view of the scenarios construction approach there is not much difference between two stage and multistage SP. In both cases the numerical effort in solving the deterministic equivalent is more or less proportional to the number of generated scenarios. This view on SP started to change with developments of randomization methods and the sample complexity theory [68]. From the perspective of solving the deterministic equivalent problem, even two stage linear stochastic programs are computationally intractable; their computational complexity is #P-hard for a sufficiently high accuracy, implying that they are at least as hard as NP problems (cf., [14], [23]). On the other hand, under reasonable assumptions, the number of randomly generated scenarios (by Monte Carlo sampling techniques), which are required to solve two stage SP problems with accuracy $\varepsilon > 0$ and high probability is of order $O(\varepsilon^{-2})$, see [68, Section 5.3]. While randomization methods were reasonably successful in solving two stage problems, the situation is different as far as multistage SP is concerned. The number of scenarios needed to solve multistage SP problems grows exponentially with the increase of the number of stages, see [70] and [68, Section 5.8.2].

Classical approach to sequential optimization is based on *dynamic programming* [3]. Dynamic programming also has a long history and is at the heart of the SOC and MDP modeling. It has the problem of the so-called "Curse of Dimensionality", a term coined by Bellman [3]. Its computational complexity increases exponentially with respect to (w.r.t.) the dimension of state variables. There is a large literature intending to deal with this problem by using various approximations of dynamic programming equations (see [56] and the references therein). Most of these methods are heuristics and often do not give verifiable guarantees for the accuracy of obtained solutions. There exist some developments on approximate dynamic programming with performance guarantees, e.g. those based on fitted value/policy iteration [41]–[43] and policy gradient methods [30], [32]. However, these performance guarantees often depend on an unknown function approximation error associated with the expressiveness of a given function class used to approximate the cost-to-go (value) functions.

Recent progress in solving convex multistage SP problems is based on cutting plane approximations of the cost-to-go functions of dynamic programming equations. These methods allow to estimate the error of the computed solution. Cutting plane type algorithms in dynamical settings is one of the main topics of this work. In particular, Stochastic Dual Dynamic Programming (SDDP), an algorithm first introduced by Pereira and Pinto [47] that builds upon the nested decomposition algorithm of Birge [5], has been a popular cutting plane method for multistage SP. Its convergence properties have been extensively studied in the literature (see, e.g., [12], [19], [24], [31], [49], [65], [77]). In this monograph, we will discuss cutting plane algorithms in the frameworks of SP and SOC with a focus on their associated rate of convergence. Moreover, we will also present extensions of stochastic approximation (a.k.a. stochastic gradient descent) type methods [29], [44], [45], [57] for multistage stochastic optimization, referred to as dynamic stochastic approximation in [34]. From the computational complexity point of view, these two types of methods seem to be complimentary to each other in the following sense. Specifically, certain variants of cutting plane methods have a computational complexity that grows mildly (linearly or quadratically) w.r.t. the number of stages, but exponentially w.r.t. the dimension of decision variables. On the other hand, the computational complexity for dynamic stochastic approximation methods increases exponentially w.r.t. the number of stages, but only mildly depends on the dimension of decision variables. Therefore, cutting plane type methods can handle multistage problems with a large number of stages, but a relatively small number of state (decision) variables. On the other hand, stochastic approximation type methods can only deal with a small number of stages, but a large number of decision variables. These methods share the following common features: (a) both methods utilize the convex structure of the cost-to-go (value) functions of dynamic programming equations, (b) both methods do not require explicit discretization of the state space, (c) both methods guarantee the convergence to the global optimality, (d) rates of convergence for both methods have been established.

It is worth noting a few alternative numerical methods for solving convex multistage SP problems that will not be covered in detail in this monograph. Firstly, the progressive hedging algorithm by Rockafellar and Wets [58] is a well-known scenario-based decomposition method, which basically applies the alternating direction method of multipliers (ADMM) to handle linear non-anticipativity constraints in the randomly generated sample average approximation problem [68]. In fact, one can also apply other primal-dual first-order optimization methods to handle these linear constraints (see, e.g., Chapter 3 of [29]). However, the size of the decomposition problem, i.e., the number of decision variables and linear constraints, will grow exponentially with the number of stages. Hence, these methods can only be applied to problems with a small number of stages. In addition, different from SA method, these decomposition methods would require the scenario tree to be generated and saved in the computer memory. Secondly, some advanced cutting plane methods, e.g., those based on bundle level method [28], [36], [37], can be used for solving two-stage SP problems efficiently. However, their extensions to multistage SP appear to be nontrivial.

This monograph is organized as follows. SP and SOC models will be first discussed in Sections 2 and 3, respectively. In Section 4, we present risk averse and distributionally robust SP and SOC models. Sections 5 and 6, respectively, are dedicated to cutting plane methods and their rates of convergence. In Section 7, we review some recent progress on SA methods for multistage stochastic optimization. This work concludes with a brief summary and possible future research directions in Section 8. Readers certainly do not need to strictly follow the above outline. For example, beginners can skip the more technically involved discussion of risk averse models in Section 4, and move directly to algorithmic studies in their first pass through this work. It should be pointed out that we attempt to cover the fundamental models (SP, SOC, and risk aversion) in earlier sections, and discuss numerical methods in later sections. However, we also cover some other models in later sections, including infinite horizon models, periodic models, and hierarchical models, since the development of these models was inspired by the studies on numerical methods for multistage SP.

We use the following notation and terminology throughout the monograph. For $a \in \mathbb{R}$ we denote $[a]_+ := \max\{0, a\}$. Unless stated otherwise $\|\cdot\|$ denotes Euclidean norm in $\mathbb{R}^n$. By $\text{dist}(x, S) := \inf_{y \in S} \|x - y\|$ we denote the distance from a point $x \in \mathbb{R}^n$ to a set $S \subset \mathbb{R}^n$. We write $x^\top y$ or $\langle x, y \rangle$ for the scalar product $\sum_{i=1}^n x_i y_i$ of vectors $x, y \in \mathbb{R}^n$. It is said that a set $S \subset \mathbb{R}^n$ is polyhedral if it can be represented by a finite number of affine constraints, it is said that a function $f : \mathbb{R}^n \to \mathbb{R}$ is polyhedral if it can be represented as maximum of a finite number of affine functions. For a process $\xi_1, \xi_2, \ldots$, we denote by $\xi_{[t]} = (\xi_1, \ldots, \xi_t)$ its history up to time t. By $\mathbb{E}_{|X}[\,\cdot\,]$ we denote the conditional expectation, conditional on random variable (random vector) X. We use the same notation ξ_t viewed as a random vector or as a vector variable, the particular meaning will be clear from the context. For a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, by $L_p(\Omega, \mathcal{F}, \mathbb{P})$, $p \in [1, \infty)$, we denote the space of random variables $Z : \Omega \to \mathbb{R}$ having finite p-th order moment, i.e., such that $\int |Z|^p d\mathbb{P} < \infty$. Equipped with norm $\|Z\|_p := (\int |Z|^p d\mathbb{P})^{1/p}$, $L_p(\Omega, \mathcal{F}, \mathbb{P})$ becomes a Banach space. The dual of $\mathcal{Z} := L_p(\Omega, \mathcal{F}, \mathbb{P})$ is the space $\mathcal{Z}^* = L_q(\Omega, \mathcal{F}, \mathbb{P})$ with $q \in (1, \infty]$ such that $1/p + 1/q = 1$.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, 1999, pp. 203–228.

[2] E. M. L. Beale, "On minimizing a convex function subject to linear inequalities," *Journal of the Royal Statistical Society, Series B*, vol. 17, 1955, pp. 173–184.

[3] R. Bellman, *Dynamic Programming*. Princton, NJ: Princeton University Press, 1957.

[4] D. Bertsekas and S. Shreve, *Stochastic Optimal Control, The Discrete Time Case*. Academic Press, New York, 1978.

[5] J. Birge, "Decomposition and partitioning methods for multistage stochastic linear programs," *Operations Research*, vol. 33, 1985, pp. 989–1007.

[6] R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, 2nd. New York: Springer, 2011.

[7] J. F. Bonnans, Z. Cen, and T. Christel, "Energy contracts management by stochastic programming techniques," *Annals of Operations Research*, vol. 200, 2012, pp. 199–222.

[8] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, ser. Springer Series in Operations Research. Springer, 2000.

[9] B. da Costa and V. Leclére, "Dual SDDP for risk-averse multistage stochastic programs," *Operations Research Letters*, vol. 51, 2023, pp. 332–337.
[10] G. Dantzig, "Linear programming under uncertainty," *Management Science*, vol. 1, 1955, pp. 197–206.
[11] L. Ding, S. Ahmed, and A. Shapiro, "A python package for multistage stochastic programming," *Optimization online*, 2019. URL: http://www.optimization-online.org/DB_FILE/2019/05/7199.pdf.
[12] C. J. Donohue and J. R. Birge, "The abridged nested decomposition method for multistage stochastic linear programs with relatively complete recourse," *Algorithmic Operations Research*, vol. 1, no. 1, 2006.
[13] O. Dowson and L. Kapelevich, "SDDP.jl: A Julia Package for Stochastic Dual Dynamic Programming," *Informs Journal on Computing*, vol. 33, 2021, pp. 27–33.
[14] M. Dyer and L. Stougie, "Computational complexity of stochastic programming problems," *Mathematical Programming*, vol. 106, 2006, pp. 423–432.
[15] H. Föllmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, 2nd. Walter de Gruyter, Berlin, 2004.
[16] C. Füllner and S. Rebennack, "Stochastic dual dynamic programming and its variants," *Optimization online*, 2021.
[17] S. Ghadimi, A. Ruszczynski, and M. Wang, "A single timescale stochastic approximation method for nested stochastic optimization," *SIAM Journal on Optimization*, vol. 30, no. 1, 2020, pp. 960–979.
[18] P. Girardeau, V. Leclére, and A. B. Philpott, "On the convergence of decomposition methods for multistage stochastic convex programs," *Mathematics of Operations Research*, vol. 40, 2016, pp. 130–145.
[19] V. Guigues, *Inexact cuts in deterministic and stochastic dual dynamic programming applied to linear optimization problems*, 2018. arXiv: 1801.04243 [math.OC].

[20] V. Guigues and M. Bandarra, “Single cut and multicut sddp with cut selection for multistage stochastic linear programs: Convergence proof and numerical experiments,” 2019.

[21] V. Guigues, A. Shapiro, and Y. Cheng, “Duality and sensitivity analysis of multistage linear stochastic programs,” *European Journal of Operational Research*, vol. 308, 2023, pp. 752–767.

[22] V. Guigues, A. Shapiro, and Y. Cheng, “Risk-averse stochastic optimal control: An efficiently computable statistical upper bound,” *Operations Research Letters*, vol. 51, 2023, pp. 393–400.

[23] G. Hanasusanto, D. Kuhn, and W. Wiesemann, “A comment on computational complexity of stochastic programming problems,” *Mathematical Programming*, vol. 159, 2015, pp. 557–569.

[24] M. Hindsberger and A. Philpott, “Resa: A method for solving multistage stochastic linear programs,” *Journal of Applied Operational Research*, vol. 6, no. 1, 2014, pp. 2–15.

[25] C. Ju and G. Lan, “Dual dynamic programming for stochastic programs over an infinite horizon,” *arXiv*, 2023.

[26] J. Kelley, “The cutting-plane method for solving convex programs,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, 1960, pp. 703–712.

[27] G. Lan, “An optimal method for stochastic composite optimization,” *Mathematical Programming*, vol. 133(1), 2012, pp. 365–397.

[28] G. Lan, “Bundle-level type methods uniformly optimal for smooth and nonsmooth convex optimization,” *Mathematical Programming*, vol. 149(1-2), 2015, pp. 1–45.

[29] G. Lan, *First-order and Stochastic Optimization Methods for Machine Learning*. Switzerland AG: Springer Nature, 2020.

[30] G. Lan, “Policy mirror descent for reinforcement learning: Linear convergence, new sampling complexity, and generalized problem classes,” 2021.

[31] G. Lan, “Complexity of stochastic dual dynamic programming,” *Mathematical Programming*, vol. 191, 2022, pp. 717–754.

[32] G. Lan, “Policy optimization over general state and action spaces,” *arXiv preprint arXiv:2211.16715*, 2022.

[33] G. Lan, A. S. Nemirovski, and A. Shapiro, "Validation analysis of mirror descent stochastic approximation method," *Mathematical Programming*, vol. 134, 2012, pp. 425–458.

[34] G. Lan and Z. Zhou, "Dynamic stochastic approximation for multi-stage stochastic optimization," *Mathematical Programming*, vol. 187, 2021, pp. 487–532.

[35] V. Leclére, P. Carpentier, J.-P. Chancelier, A. Lenoir, and F. Pacaud, "Exact converging bounds for Stochastic Dual Dynamic Programming via Fenchel duality.," *SIAM J. Optimization*, vol. 30, 2020, pp. 1223–1250.

[36] C. Lemaréchal, A. Nemirovskii, and Y. Nesterov, "New variants of bundle methods," *Mathematical Programming*, vol. 69, 1995, pp. 111–147.

[37] J. Liang and R. D. Monteiro, "A proximal bundle variant with optimal iteration-complexity for a large range of prox stepsizes," *SIAM Journal on Optimization*, vol. 31, no. 4, 2021, pp. 2955–2986.

[38] N. Löhndorf and A. Shapiro, "Modeling time-dependent randomness in stochastic dual dynamic programming," *European Journal of Operational Research*, vol. 273, 2019, pp. 650–661.

[39] N. Löhndorf, D. Wozabal, and S. Minner, "Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming," *Operational Research*, vol. 61, 2013, pp. 810–823.

[40] V. L. de Matos, A. B. Philpott, and E. Finardi, "Improving the performance of stochastic dual dynamic programming," *Journal of Computational and Applied Mathematics*, vol. 290, 2015, pp. 196–208.

[41] R. Munos, "Error bounds for approximate policy iteration," in *19th International Conference on Machine Learning*, pp. 560–567, 2003.

[42] R. Munos, "Error bounds for approximate value iteration," in *20th national conference on Artificial intelligence*, pp. 1006–1011, 2005.

[43] R. Munos and C. Szepesvári, "Finite-time bounds for fitted value iteration," *Journal of Machine Learning Research*, 2008, pp. 815–857.

[44] A. S. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, "Robust stochastic approximation approach to stochastic programming," *SIAM J. Optimization*, vol. 19, 2009, pp. 1574–1609.

[45] A. S. Nemirovski and D. Yudin, *Problem complexity and method efficiency in optimization*, ser. Wiley-Interscience Series in Discrete Mathematics. John Wiley, XV, 1983.

[46] Y. Nesterov and A. Nemirovski, *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia: SIAM, 1994.

[47] M. Pereira and L. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical programming*, vol. 52, no. 1-3, 1991, pp. 359–375.

[48] G. Pflug, "Some remarks on the value-at-risk and the conditional value-at-risk," in *Probabilistic Constrained Optimization: Methodology and Applications, S. Uryasev (Ed.)* Kluwer Academic Publishers, Norwell, MA, 2000.

[49] A. Philpott, V. d. Matos, and E. Finardi, "On solving multistage stochastic programs with coherent risk measures," *Operations Research*, vol. 61, 2013, pp. 957–970.

[50] A. Philpott, F. Wahid, and F. Bonnans, "A mixed integer dynamic approximation scheme," *Mathematical Programming*, vol. 181, 2020, pp. 19–50.

[51] A. Philpott and V. de Matos, "Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion," *European Journal of Operational Research*, vol. 218, 2012, pp. 470–483.

[52] A. Philpott, V. de Matos, and E. Finardi, "On solving multistage stochastic programs with coherent risk measures," *Operations Research*, vol. 61, no. 4, 2013, pp. 957–970.

[53] A. Pichler and A. Shapiro, "Mathematical foundations of distributionally robust multistage optimization," *SIAM J. Optimization*, vol. 31, 2021, pp. 3044–3067.

[54] B. Polyak, "New stochastic approximation type procedures," *Automat. i Telemekh.*, vol. 7, 1990, pp. 98–107.

[55] B. Polyak and A. Juditsky, "Acceleration of stochastic approximation by averaging," *SIAM J. Control and Optimization*, vol. 30, 1992, pp. 838–855.

[56] W. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, 2nd. New York: John Wiley and Sons, 2011.

[57] H. Robbins and S. Monro, “A stochastic approximation method,” *Annals of Mathematical Statistics*, vol. 22, 1951, pp. 400–407.

[58] R. T. Rockafellar and R. J.-B. Wets, “Scenarios and policy aggregation in optimization under uncertainty,” *Mathematics of Operations Research*, vol. 16, no. 1, 1991, pp. 119–147.

[59] R. T. Rockafellar, *Conjugate Duality and Optimization*. Philadelphia: Society for Industrial and Applied Mathematics, 1974.

[60] R. Rockafellar and S. Uryasev, “Conditional value-at-risk for general loss distributions,” *J. of Banking and Finance*, vol. 26, no. 7, 2002, pp. 1443–1471.

[61] R. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[62] A. Ruszczyński and A. Shapiro, “Conditional risk mappings,” *Mathematics of Operations Research*, vol. 31, 2006, pp. 544–561.

[63] A. Ruszczyński and A. Shapiro, “Optimization of convex risk functions,” *Mathematics of Operations Research*, vol. 31, 2006, pp. 433–452.

[64] A. Shapiro, “On complexity of multistage stochastic programs,” *Operations Research Letters*, vol. 34, 2006, pp. 1–8.

[65] A. Shapiro, “Analysis of stochastic dual dynamic programming method,” *European Journal of Operational Research*, vol. 209, 2011, pp. 63–72.

[66] A. Shapiro and Y. Cheng, “Central limit theorem and sample complexity of stationary stochastic programs,” *Operations Research Letters*, vol. 49, 2021, pp. 676–681.

[67] A. Shapiro and Y. Cheng, “Dual bounds for periodical stochastic programs,” *Operations Research*, 2021.

[68] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, third. Philadelphia: SIAM, 2021.

[69] A. Shapiro and L. Ding, “Periodical multistage stochastic programs,” *SIAM J. Optimization*, vol. 30, 2020, pp. 2083–2102.

[70] A. Shapiro and A. Nemirovski, "On complexity of stochastic programming problems," in *Continuous Optimization: Current Trends and Applications: Current Trends and Applications*, V. Jeyakumar and A. Rubinov, Eds., Springer, 2005, pp. 111–144.

[71] A. Shapiro, W. Tekaya, J. da Costa, and M. P. Soares, "Risk neutral and risk averse stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 224, 2013, pp. 375–391.

[72] P. Shinde, I. Kouveliotis-Lysikatos, and M. Amelin, "Multistage stochastic programming for vpp trading in continuous intraday electricity markets," *IEEE Transactions on Sustainable Energy*, vol. 13, 2022, pp. 1037–1048.

[73] M. Wang, E. X. Fang, and H. Liu, "Stochastic compositional gradient descent: Algorithms for minimizing compositions of expected-value functions," *Mathematical Programming*, vol. 161, no. 1-2, 2017, pp. 419–449.

[74] S. Zhang and X. Sun, "Stochastic dual dynamic programming for multistage stochastic mixed-integer nonlinear optimization," *Mathematical Programming*, vol. 196, 2022, pp. 935–985.

[75] Z. Zhang and G. Lan, "Optimal methods for convex nested stochastic composite optimization," *Mathematical Programming*, 2020.

[76] P. Zipkin, *Foundation of inventory management*. McGraw-Hill, 2000.

[77] J. Zou, S. Ahmed, and X. A. Sun, "Stochastic dual dynamic integer programming," *Mathematical Programming*, vol. 175, 2019, pp. 461–502.