

STOCHASTIC DECOMPOSITION ALGORITHMS FOR RISK-AVERSE MULTISTAGE
STOCHASTIC LINEAR PROGRAMS AND APPLICATIONS

A Dissertation

by

PRASAD PARAB

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Chair of Committee, Lewis Ntaimo
Committee Members, David Larson
 Erick Moreno-Centeno
 Natarajan Gautam
Head of Department, Mark Lawley

December 2019

Major Subject: Industrial Engineering

Copyright 2019 Prasad Parab

ABSTRACT

Mean-risk stochastic linear programming provides a framework for controlling cost variability in problems involving sequential decision making under uncertainty. It goes beyond the classical expected value framework by including risk measures in the objective function and aims at controlling cost variability in the solution. This allows for modeling risk averseness in variety of applications such as long-term financial planning, scheduling of power systems, supply chain management and portfolio optimization.

In this dissertation, we derive stochastic decomposition algorithms for solving mean-risk two-stage stochastic linear programs (MR-SLP) and mean-risk multistage stochastic linear programs (MR-MSLP) with *deviation* and *quantile* risk measures. Stochastic decomposition(SD) is a type of internal sampling method and at every iteration of algorithm only one linear problem is solved for approximating the recourse function. A salient feature of the SD algorithm is that the number of samples is not fixed a priori, which allows to obtain good candidate solutions early in the procedure.

We also report on a computational study to evaluate the empirical performance of the SD algorithms for MR-SLP and MR-MSLP with *expected excess*(EE), *quantile deviation*(QDEV) and *conditional value-at-risk*(CVaR) as risk measures. The goal of the study was to analyze for a given instance how SD algorithm performs across different levels of risk, investigate the effect of different risk measures and understand when it is appropriate to use the risk-averse approach.

For MR-SLP, the SD algorithm is implemented and applied to standard test instances and it shows that the risk measure QDEV has more impact on expected cost and the cost associated with extreme scenarios compared to the impact of CVaR and EE. We also observed that for higher target values, the risk measure EE becomes effective only for a relatively small number of scenarios and has little to no-effect on the optimal solution for small values of the risk trade-off factor. The computational study also demonstrates that under risk aversion the rate of convergence of SD algorithm remains consistent as opposed to sample average approximation approach.

For the multistage case, the SD algorithm is applied to an instance of long-term hydrothermal

scheduling (LTHS) and it shows that the risk trade-off factor has a significant impact on the solution and the risk measure *conditional value-at-risk* exhibits a better control over the extreme scenarios at lower values of risk trade-off factor. The study also shows that the risk-neutral approach is still appropriate for the LTHS problem.

DEDICATION

To my parents and teachers:

For always supporting me, believing in me and guiding me.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude towards my family for the never ending support I have always received from them. I would like to remember my mother Sumitra, the discipline and values she instilled in me made it possible to finish this arduous task successfully. My father Dashrath, my brother Vinayak, my uncle Ganpat and aunt Vidya deserves special thanks for their endless encouragement throughout my graduate studies.

I am also thankful to the several members of the Aggie family for providing a home away from home. First and foremost I am extremely grateful to Dr. Ntaimo, for being an excellent advisor and mentor to me, and for patiently spending countless hours in guiding me throughout this process. I sincerely appreciate the help I received from Dr. Pagnoncelli and Maryam Khatami for completion of my research work. I also want to thank my committee members Dr. Gautam, Dr. Larson and Dr. Moreno-Centeno.

I am grateful to my fellow graduate students and especially my friends Jianguye Gong, Kumaresh Kasi, Kevin Layer, Haochen Luo, Niranjan Rajghatta and Brittany Segundo for making the graduate school memorable. I am also very grateful to all the students, whom I had the privilege to teach as a graduate instructor and which further enriched my experience at Texas A & M. Finally, I would like to thank the Industrial and Systems Engineering Department for providing financial support and making this endeavor possible.

NOMENCLATURE

LP	Linear Programming
SLP	Stochastic Linear Programming
MR-SLP	Mean-Risk Two-Stage Stochastic Linear Programming
MSLP	Multistage Stochastic Linear Programming
MR-MSLP	Mean-Risk Multistage Stochastic Linear Programming
EE	Expected Excess
QDEV	Quantile Deviation
CVaR	Conditional Value at Risk
DEP	Deterministic Equivalent Formulation
SD	Stochastic Decomposition
MSD	Multistage Stochastic Decomposition
SAA	Sample Average Approximation
SDDP	Stochastic Dual Dynamic Programming
LTHS	Long-Term Hydrothermal Scheduling

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGMENTS	v
NOMENCLATURE	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	x
LIST OF TABLES	xii
1. INTRODUCTION.....	1
2. LITERATURE REVIEW	4
2.1 Risk Measures and MR-SLP	4
2.2 MR-MSLP	6
3. MEAN-RISK STOCHASTIC LINEAR PROGRAMS.....	9
3.1 Two-Stage MR-SLP.....	9
3.2 Risk Measures	10
3.2.1 Expected Excess	10
3.2.2 Quantile Deviation	12
3.2.3 Conditional Value-at-Risk.....	12
3.3 Properties of Coherent Risk Measures	13
3.4 Deterministic Equivalent Formulation of MR-SLP.....	15
3.4.1 MR-SLP with EE.....	15
3.4.2 MR-SLP with QDEV.....	15
3.4.3 MR-SLP with CVaR.....	16
3.5 MR-SLP	17
3.6 Deterministic Equivalent Formulation of MR-MSLP	19
3.6.1 MR-MSLP with EE	19
3.6.2 MR-MSLP with QDEV	20
3.6.3 MR-MSLP with CVaR	22
4. STOCHASTIC DECOMPOSITION ALGORITHM.....	24

4.1	Stochastic Decomposition Algorithm for EE	24
4.1.1	Decomposition for MR-SLP with EE	24
4.1.2	SD Algorithm for MR-SLP with EE	26
4.2	Stochastic Decomposition for QDEV	29
4.2.1	Decomposition for MR-SLP with QDEV	29
4.2.2	SD Algorithm for MR-SLP with QDEV	30
4.3	Stochastic Decomposition for CVaR	32
4.3.1	Decomposition for MR-SLP with CVaR	33
4.3.2	SD Algorithm for MR-SLP with CVaR	34
4.4	Proof of Convergence	37
4.4.1	Convergence Proof for SD-EE	37
5.	COMPUTATIONAL RESULTS	49
5.1	Test Instances	49
5.2	Results	51
5.2.1	Results for SD-EE	52
5.2.2	Results for SD-QDEV	53
5.2.3	Results for SD-CVaR	54
5.3	Impact of Risk Measure on Optimal Solution	54
5.4	Expected Cost and Risk Trade-Off Factor	56
5.5	Mean-Risk SD versus Mean-Risk SAA	57
6.	MULTISTAGE STOCHASTIC DECOMPOSITION ALGORITHM	61
6.1	MSD for EE	62
6.1.1	Decomposition for MR-MSLP with EE	62
6.1.2	MSD Algorithm for MR-MSLP with EE	66
6.2	MSD for QDEV	69
6.2.1	Decomposition for MR-MSLP with QDEV	69
6.2.2	MSD Algorithm for MR-MSLP with QDEV	72
6.3	MSD for CVaR	75
6.3.1	Decomposition for MR-MSLP with CVaR	76
6.3.2	MSD Algorithm for MR-MSLP with CVaR	79
7.	LONG-TERM HYDROTHERMAL SCHEDULING	83
7.1	LTHS Problem Formulation	83
7.1.1	LTHS Formulation with EE	84
7.1.2	LTHS Formulation with QDEV	85
7.1.3	LTHS Formulation with EECVaR	86
7.2	Computational Results	87
7.2.1	Risk-Neutral MSLP	88
7.2.2	MR-MSLP with EE	90
7.2.3	MR-MSLP with QDEV	92
7.2.4	MR-MSLP with CVaR	93

7.2.5	Discussion	95
7.2.6	Impact of Risk Measures on Extreme Scenarios	96
8.	CONCLUSIONS AND FUTURE RESEARCH	100
8.1	Conclusions.....	100
8.2	Future Research	101
	REFERENCES	103
	APPENDIX A. COMPUTATIONAL RESULTS TABLES.....	108
A.1	EE.....	108
A.2	QDEV.....	118
A.3	CVaR.....	128
	APPENDIX B. CONVERGENCE PROOFS FOR SD	143
B.1	Convergence Proof for SD-QDEV	143
B.2	Convergence Proof for SD-CVaR	156
	APPENDIX C. LTHS: GRAPHS	169
C.1	MR-MSLP with EE	169
C.2	MR-MSLP with QDEV	172
C.3	MR-MSLP with CVaR.....	175
C.4	Impact of Risk Measures	178
	APPENDIX D. PROOF FOR PROPERTIES OF RISK MEASURES	181
D.1	Expected Excess.....	181
D.2	Quantile Deviation	182
D.3	Conditional Value-at-Risk	185

LIST OF FIGURES

FIGURE	Page
5.1 Expected Cost $\mathbb{E}[f(x, \tilde{\omega})]$	56
5.2 Cost of Worst Case Scenarios	58
6.1 Scenario Tree	61
7.1 Inflow per Stage	88
7.2 Results for Risk-Neutral MSLP for LTHS	89
7.3 Results for MR-MSLP with EE for LTHS ($\lambda = 0.1$).....	91
7.4 Results for MR-MSLP with QDEV for LTHS ($\lambda = 0.1$)	93
7.5 Results for MR-MSLP with CVaR for LTHS ($\lambda = 0.1$).....	94
7.6 Cost of Operation per Stage	97
7.7 Impact of Risk Measures on Extreme Scenarios ($\lambda = 0.1$)	99
C.1 Results for MR-MSLP with EE for LTHS ($\lambda = 0.25$)	169
C.2 Results for MR-MSLP with EE for LTHS ($\lambda = 0.5$).....	170
C.3 Results for MR-MSLP with EE for LTHS ($\lambda = 1$)	171
C.4 Results for MR-MSLP with QDEV for LTHS ($\lambda = 0.25$)	172
C.5 Results for MR-MSLP with QDEV for LTHS ($\lambda = 0.5$)	173
C.6 Results for MR-MSLP with QDEV for LTHS ($\lambda = 1$).....	174
C.7 Results for MR-MSLP with CVaR for LTHS ($\lambda = 0.25$)	175
C.8 Results for MR-MSLP with CVaR for LTHS ($\lambda = 0.5$).....	176
C.9 Results for MR-MSLP with CVaR for LTHS ($\lambda = 1$).....	177
C.10 Impact of Risk Measures on Extreme Scenarios ($\lambda = 0.25$)	178
C.11 Impact of Risk Measures on Extreme Scenarios ($\lambda = 0.5$)	179

C.12 Impact of Risk Measures on Extreme Scenarios 180

LIST OF TABLES

TABLE	Page
3.1 Properties Satisfied by Risk Measures	14
3.2 Properties Satisfied by Mean-Risk Measures	14
5.1 Standard Test Problems	50
5.2 Optimal Solution for Instance <i>pgp2e</i>	55
5.3 Optimal Values for Instance <i>pgp2e</i>	55
5.4 Average CPU Time (Sec)	59
5.5 Number of Iterations	60
7.1 Average Cost of Operation per Stage	95
7.2 Average Cost per Stage for Extreme Scenarios	98
A.1 Results for SD-EE for <i>pgp2</i> ($\psi = 450.00$).	108
A.2 Results for SAA-EE for <i>pgp2</i> ($\psi = 450.00$).	109
A.3 Results for SD-EE for <i>pgp2e</i> ($\psi = 420.00$).	110
A.4 Results for SAA-EE for <i>pgp2e</i> ($\psi = 420.00$).	111
A.5 Results for SD-EE for <i>gbd</i> ($\psi = 1672.33$).	112
A.6 Results for SAA-EE for <i>gbd</i> ($\psi = 162.73$).	113
A.7 Results for SD-EE for <i>LandS</i> ($\psi = 227.40$).	114
A.8 Results for SAA-EE for <i>LandS</i> ($\psi = 227.40$).	115
A.9 Results for SD-EE for <i>storm</i> ($\psi = 15580000.00$).	116
A.10 Results for SAA-EE for <i>storm</i> ($\psi = 15580000.00$).	117
A.11 Results for SD-QDEV for <i>pgp2</i>	118
A.12 Results for SAA-QDEV for <i>pgp2</i>	119

A.13 Results for SD-QDEV for <i>pgp2e</i>	120
A.14 Results for SAA-QDEV for <i>pgp2e</i>	121
A.15 Results for SD-QDEV for <i>gbd</i>	122
A.16 Results for SAA-QDEV for <i>gbd</i>	123
A.17 Results for SD-QDEV for <i>LandS</i>	124
A.18 Results for SAA-QDEV for <i>LandS</i>	125
A.19 Results for SD-QDEV for <i>storm</i>	126
A.20 Results for SAA-QDEV for <i>storm</i>	127
A.21 Results for SD-CVaR for <i>pgp2</i>	128
A.22 Results for SAA-CVaR for <i>pgp2</i>	129
A.23 Results for SD-CVaR for <i>pgp2e</i>	130
A.24 Results for SAA-CVaR for <i>pgp2e</i>	131
A.25 Results for SD-CVaR for <i>gbd</i>	132
A.26 Results for SAA-CVaR for <i>gbd</i>	133
A.27 Results for SD-CVaR for <i>LandS</i>	134
A.28 Results for SAA-CVaR for <i>LandS</i>	135
A.29 Results for SD-CVaR for <i>storm</i>	136
A.30 Results for SAA-CVaR for <i>storm</i>	137
A.31 Results for SD-EE for <i>pgp2</i> ($\psi = 557.00$).	138
A.32 Results for SD-EE for <i>pgp2e</i> ($\psi = 645.00$).	139
A.33 Results for SD-EE for <i>gbd</i> ($\psi = 3100.00$).	140
A.34 Results for SD-EE for <i>LandS</i> ($\psi = 320.00$).....	141
A.35 Results for SD-EE for <i>storm</i> ($\psi = 16095000.00$).....	142

1. INTRODUCTION

Stochastic linear programming (SLP) provides a framework for modeling problems involving sequential decision making under uncertainty. Hence, it is commonly used for modeling purposes in variety of applications such as long-term financial planning [1], scheduling of power systems [2, 3, 4], supply chain management [5] and others. However, most of the literature in the field of stochastic programming is based on the classical expected value framework to quantify variability of random variables. The expected value framework implies risk neutrality, which may not be a suitable approach in certain applications such as financial planning, portfolio management, energy planning etc. Lack of ability to provide hedging against extreme scenarios or controlling variability of scenario costs has led to a constant criticism of risk-neutral SLP approach. Therefore, risk measures have been introduced into the SLP models, where risk arising due to the inherent uncertainty in a problem is reflected by including a dispersion statistic along with expectation in the objective function of SLP.

This research work makes several contributions to the literature on mean-risk stochastic programming. First of all, this work makes initial efforts towards applying stochastic decomposition (SD) to mean-risk two-stage stochastic linear programs (MR-SLPs) with a goal of solving large-scale instances. Solving large-scale MR-SLPs is challenging and this work devises a SD approach [6] towards alleviating this challenge. SD is a streamlined interior sampling approach that allows for generating one sample at a time in the course of an algorithm until enough samples have been generated. Second, this work considers risk measures belonging to two different categories, deviation and quantile risk measures and we have devised SD algorithm for each one. In particular, we consider deviation risk measure *expected excess* [7] and quantile risk measures *quantile deviation* [8] and *conditional value-at-risk* [9]. Another contribution of this work is a computational study of an implementation of the SD algorithms applied to standard test instances and contrasting it against exterior sampling method SAA. The study shows promising results that demonstrate the suitability of the SD algorithms towards solving MR-SLPs and also helps in understanding the effect of risk

measures on the optimal solution.

This work also makes contributions to the literature on mean-risk multistage stochastic linear programs(MR-MSLP). Solving large-scale instances of MR-MSLP is challenging due to the problem size resulting from the combinatorial explosion of number of stages and scenarios per stage. The main objective of this work is to derive a mean-risk MSD approach towards alleviating this challenge. The subgradients in MSD are generated using one scenario of a random outcome per iteration and therefore it has an advantage of not having to process every potential random outcome, which can be a major drawback in case of large scale instances. For quantifying risk in our formulation, we again consider the deviation risk measure *expected excess* [7] and quantile risk measures *quantile deviation* [8] and *conditional value-at-risk* [9]. For each one of these three risk measures, we present the detailed problem formulation at every stage of a MSLP and explain the decomposition approach. Then we derive a mean-risk MSD algorithm. We also extend the asymptotic convergence results of risk-neutral MSD to risk-averse cases and we address the issues arising due to introduction of risk variables.

Finally, we report on a computational study based on application of mean-risk MSD to long-term hydrothermal scheduling (LTHS), which is a widely studied application in MSLP [10, 4, 11, 12]. We performed extensive computations using different risk levels for each one of the three risk measures and we present the details of impact this has over the optimal solution and each decision variable over the planning horizon. The result help in assessing the suitability of the three risk measures in context of MSLP.

The dissertation is organized into eight sections. Section 2 provides a brief review of relevant literature on stochastic programming, covering the topics of risk-aversion, two-stage stochastic programming and multistage stochastic programming. Section 3 introduces the risk measures and deterministic equivalent formulation for MR-SLP. The decomposition approach, two-stage SD algorithm and the convergence proof for risk-averse SD algorithm is outlined in Section 4. The details of computational study for two-stage SLP along with the standard test instances used and the computational results are provided in Section 5. Section 6 contains the decomposition approach

for MR-MSLP and the detailed MSD algorithms. Section 7 includes introduction to the long-term hydrothermal scheduling problem and computational results of MSD algorithm. Finally, in Section 8 a succinct summary of findings, along with conclusion and future research is provided.

2. LITERATURE REVIEW

In this section, research literature relevant to our work is discussed. In the first section, an overview of literature dealing with algorithms, formulation and computational studies associated with risk measures and two-stage SLP is described. The second section provides a brief review of the literature related to MR-MSLP and the challenges associated with decomposition and optimization of MSLPs due to their large scale.

2.1 Risk Measures and MR-SLP

A risk measure in principle, can be any function that complements the expected value of a SLP instance by avoiding solutions with high variability, or ones that deviate significantly from a target. Various risk measures have been defined in [7, 13] such that the risk formulations maintain a block angular structure and therefore are amenable to decomposition. Based on the computation of risk, the risk measures are broadly classified in two different categories *deviation* and *quantile* risk measures. The deviation measure calculates risk based on deviation of a random variable from a preselected target or the mean. Some common examples of deviation risk measures are *expected excess* and *absolute semi-deviation*. In *expected excess*, deviation of outcomes from a pre-selected target is used to calculate risk [7]; whereas in *absolute semideviation*, risk is measured by the deviation from the expected value [14].

In contrast to *deviation* risk measures, *quantile* risk measures use a quantile of the probability distribution to compute risk. Examples of *quantile* risk measures are *Value-at-Risk*, *excess probability*, *quantile deviation* and *conditional value-at-risk*. The most popular of those measures is the *value-at-risk* (VaR), which has been extensively used in finance and banking [15, 16]. The VaR is not tractable given its lack of convexity [17], which made it unsuitable for optimization when losses are not normally distributed [18]. As a result *Conditional Value-at-Risk* (CVaR) was proposed in [9], which is a coherent risk measure and also assesses the extent of losses beyond the VaR by computing conditional expectations of values above the VaR. *Quantile deviation* (QDEV),

unlike other risk measures computes two-sided deviation of scenario costs from a quantile, and an optimization procedure has been proposed in [8] to compute QDEV. In risk measure *excess probability*, risk is calculated as probability of a random variable exceeding a given target [19].

Risk aversion has a crucial role in optimization under uncertainty. Two-stage stochastic linear programs with risk aversion provide a better mechanism in coping with losses, controlling variability and handling extreme scenarios. Structural and algorithmic properties of two-stage stochastic linear programs with deviation measures such as continuity, differentiability, convexity and stability are derived in [20]. The continuity properties of the recourse function and stability results for the optimal solutions of excess probability are derived and discussed in [21]. A min-max model that is equivalent to a mean-risk model with quantile deviation risk measure is considered in [22]. The approach includes variance reduction techniques that enhance the rate of convergence as well as a certificate of optimality and statistical stopping criteria for an iterative algorithm. Various risk measures have been defined in [7, 13] such that the problem formulation maintains a dual block angular structure, which is amenable to decomposition.

A number of sampling-based algorithms have been proposed in the literature. Depending on the sampling strategy used, these algorithms can be broadly classified in two categories, *exterior* and *interior* sampling methods. In exterior methods, sampling is used to generate a fixed number of scenarios, and the resulting problem is often solved using some decomposition scheme. An example of exterior sampling is the sample average approximation (SAA) scheme, described in [23, 22].

Several approaches have been proposed to solve MR-SLPs with both *deviation* and *quantile* risk measures using the external sampling schemes. An algorithmic treatment based on variations of the L-shaped method has been proposed in [20]. In [21] the authors developed an algorithm for mean-risk problems and presented computational results for standard test instances. A decomposition-based parametric cutting plane algorithm to generate mean-risk efficient frontiers for two classes of mean-risk models has been proposed in [24]. The focus of this work is on convexity properties and subgradient decomposition. An aggregate optimality cut and separate

cut subgradient-based algorithm is proposed in [25], using absolute semideviation and quantile deviation as risk measures.

In interior sampling methods, sampling is done during the course of execution of the algorithm. SD is a prototype example of an interior sampling scheme. Other examples can be found in a comprehensive survey of Monte Carlo sampling-based methods [26]. The description and convergence properties of SD algorithm are discussed in [6] and [27]. A detailed survey of MR-SLP, applications, algorithms, and risk measures is presented in [28].

2.2 MR-MSLP

MSLP provides a framework for modeling problems involving sequential decision making under uncertainty and hence MSLP is commonly used for modeling purposes in variety of applications such as long-term financial planning [1], scheduling of power systems [2, 3, 4], supply chain management [5] and others. MSLPs are difficult to optimize due to their large scale nature. Therefore, various decomposition algorithms that use deterministic approximations of the recourse function (expected-cost-to-go function) have been devised.

Two decomposition and partitioning methods for optimizing MSLP were first proposed by Birge [29]. The first method is an outer linearization decomposition approach which extends the L-shaped method to multistage. The second method is a piecewise partitioning strategy which determined the optimal first stage solutions by partitioning the feasible region. Rockafellar and Mets [3] have applied the principle of progressive hedging to generate improving sequence of decision policies for optimizing multistage problems. In their method different scenarios are bundled together based on available information and iterative adjustments are made to decisions to arrive at an optimal policy.

Unfortunately, multistage problems become unwieldy as the number of scenarios at every stage increases or as the number of stages grows. Such an exponential growth, makes the MSLP computationally intractable and difficult to optimize even after the application of decomposition methods. Therefore, to address this difficulty various sampling approaches have been developed for optimizing large scale MSLPs. A sampling based method popularly known as stochastic dual dynamic

programming (SDDP) was proposed by Pereira and Pinto [2]. In the SDDP method, using the same principle as that of the L-shaped algorithm, piecewise linear functions are obtained from the dual solutions of the MSLP at each stage to approximate the recourse function. The SDDP method, does not require state discretization and thus avoids the combinatorial explosion resulting from large number of stages and scenarios. Another sampling approach for optimizing MSLPs, called multistage stochastic decomposition (MSD), was devised by Sen and Zhou [30]. MSD approach is an extension of the sample regularized version of stochastic decomposition algorithm [6, 27] to the multistage setting . The MSD algorithm shares some of the recursive features of ADP and has been shown to asymptotically converge to an optimal solution.

Majority of the research work in the field of MSLP deals with the risk-neutral cases, where the goal is to optimize the expected cost function. Although risk-neutral formulations provide an optimal policy, they fail to control variability among different scenarios. Also, risk-neutral formulations cannot deal with extreme losses. Therefore, risk aversion has a crucial role in MSLP because it provides a suitable mechanism for handling losses, controlling variability and hedging against extreme scenarios. Hence, MR-MSLP has applications in variety of domains such as long term scheduling of power plant operations [12], multistage asset allocation [31] and others.

The pioneering work in the field of risk-aversion by Artzner et al. [17] proposed an axiomatic theory of risk. According to the authors a risk measure is considered to be coherent if it satisfies the properties of *positive homogeneity*, *translation invariance*, *monotonicity* and *convexity*. Although a desired property, coherent risk measures are not necessarily adequate in addressing risks for certain problems [32].

Another important concept of MR-MSLP is *time consistency*. A MR-MSLP formulation is said to be *time consistent* if its optimal decisions do not depend on scenarios which are improbable to occur. We provide a formal definition of *time consistency* in Section 3.5. Risk-neutral MSLP formulation always satisfy *time consistency*, but this may not be necessarily true for MR-MSLP. Various concepts and approaches have been proposed [33, 34, 35, 36] to make MR-MSLP time consistent. It should be noted that lack of *time consistency* does not imply that the corresponding

optimal policies are unsuitable.

Owing to the difficulty in defining risk across different stages and the issue of time consistency, MR-MSLPs are even more difficult to optimize than the MSLPs. Philpott and De Matos [12] have considered a time consistent formulation of coherent risk measure *Conditional Value-at-Risk* (CVaR) for modeling risk into a MSLPs to optimize a hydro-thermal scheduling problem. The scheduling problem was then solved using SDDP algorithm. The connections between minmax models, risk-aversion and nested formulations in the context of multistage setting has also been studied [37]. A scenario decomposition method for solving MR-MSLP problems along with the convergence proof was proposed by Ruszczyński et al. [38]. In this approach, bundles of risk neutral approximations were constructed and the method was applied to a risk averse inventory and assembly problem.

In this work, we extend the MSD for risk-neutral MSLP to MR-MSLP and apply it to the long-term hydrothermal scheduling. MSD involves approximating the cost-to-go function and the optimal decisions in sequential manner. At every iteration, a sample path traversing the entire length of scenario tree from root node to a terminal node is generated. Then, the LPs associated with the nodes on this path are solved to update the recourse function approximations. Hence, unlike SDDP, MSD does not require traversing the entire tree in one single iteration.

3. MEAN-RISK STOCHASTIC LINEAR PROGRAMS

In this section, we begin by defining a two-stage MR-SLP and brief explanation of decision variables, constraints and properties. Then we present definitions and deterministic equivalent formulations of each of the risk measure considered in this work. In the end, we provide the stagewise formulations of MR-MSLP using recursion.

3.1 Two-Stage MR-SLP

A two-stage MR-SLP can be formulated as follows:

$$\begin{aligned} \text{Min} \quad & \mathbb{E}[f(x, \tilde{\omega})] + \lambda \rho[f(x, \tilde{\omega})] \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{3.1}$$

where $x \in \mathbb{R}_+^{n_1}$ is a vector of decision variables and $X \subseteq \mathbb{R}^{n_1}$ is a compact set. The use of a *risk measure* $\rho : \mathcal{F} \mapsto \mathbb{R}$ characterizes the so-called *mean-risk* stochastic program, where \mathcal{F} is the space of real random cost variables. The weighting factor $\lambda > 0$ quantifies the trade-off between expected cost and risk. A risk measure can, in principle, be any function that complements the expected value, for instance by avoiding solutions with high variability, or ones that deviate significantly from a target. Such specifications have to take into account tractability, also it is desirable that the MR-SLP maintains suitable properties such as convexity and block angular structure to allow the formulation to be amenable to optimization methods.

For a given $x \in X$ the real random cost variable $f(x, \tilde{\omega})$ is given by

$$f(x, \tilde{\omega}) := c^\top x + h(x, \tilde{\omega}),$$

with $\{f(x, \tilde{\omega})\}_{x \in X} \subseteq \mathcal{F}$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The mapping $\mathbb{E} : \mathcal{F} \mapsto \mathbb{R}$ denotes the expected value, where \mathcal{F} is the space of all real random cost variables on Ω satisfying

$\mathbb{E}[|f(\tilde{\omega})|] < \infty$. For a realization (scenario) ω of $\tilde{\omega}$, $\omega \in \Omega$, the recourse function $h(x, \omega)$ is given by

$$\begin{aligned} h(x, \omega) := & \text{Min} \quad q^\top y(\omega) & (3.2) \\ \text{s.t.} \quad & Wy(\omega) \geq r(\omega) - T(\omega)x, \\ & y(\omega) \geq 0. \end{aligned}$$

In problem (3.2), $q \in \mathbb{R}^{n_2}$ is the second-stage cost vector, $y(\omega) \in \mathbb{R}^{n_2}$ is the recourse decision vector, $W \in \mathbb{R}^{m_2 \times n_2}$ is the recourse matrix, $T(\omega) \in \mathbb{R}^{m_2 \times n_1}$ is the technology matrix, and $r(\omega) \in \mathbb{R}^{m_2}$ is the right hand side vector.

We consider MR-SLP under the following assumptions:

- (A1) The sets X and Ω are compact.
- (A2) For any given $x \in X$, $\mathbb{E}[|f(x, \omega)|] < \infty$.
- (A3) Recourse matrix W is fixed (fixed recourse).
- (A4) For all $x \in X$, there exist a constant L such that $L \leq \mathbb{E}[f(x, \tilde{\omega})]$.

Assumptions (A1) and (A2) guarantee the existence of an optimal solution. Satisfying assumption (A3) guarantees relatively complete recourse which enables having a dual feasible set for the sub-problem. In SD, since optimality cuts are generated using one scenario at a time, assumption (A4) is required to make sure that the optimal solution is not cut-off by the optimality cuts.

3.2 Risk Measures

Let us begin with the mathematical definition and the interpretation of each risk measure. In the definitions, $\max\{a, 0\}$ denotes the *maximum* operator applied to $a \in \mathbb{R}$ and 0, while Min and Max refer to *minimization* and *maximization* of a real-valued objective function, respectively.

3.2.1 Expected Excess

Given a target $\psi \in \mathbb{R}$, expected excess (EE) [7] is defined as:

$$\phi_{EE_\psi}(x) := \mathbb{E}[\max\{f(x, \tilde{\omega}) - \psi, 0\}].$$

It reflects the expected value of the excess over the target $\psi \in \mathbb{R}$, that is, if function $f(x, \omega)$ represents costs, then EE computes average losses greater than the threshold ψ . Setting $\rho := \phi_{EE_\psi}$ in formulation (3.1) we obtain the mean-risk formulation with EE as risk measure:

$$\mathbf{Min}_{x \in X} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \phi_{EE_\psi}(x). \quad (3.3)$$

EE is used to evaluate and minimize the risk of not achieving a given performance target for every scenario. For example, it can be used in portfolio optimization problems in which some desired financial return is expected [39, 40].

Selection of an appropriate target ψ is crucial for risk measure EE. For every realization $\omega \in \Omega$, if $\psi \leq f(x, \omega)$ or if $\psi \geq f(x, \omega)$, then the optimal decision vector x will be same as that of the risk-neutral case.

LEMMA 3.2.1. *If for risk measure EE, the target $\psi \in \mathbb{R}$ is such that $\psi \leq f(x, \omega)$ for every realization $\omega \in \Omega$. Then, the optimal decision vector x will be same as that of the risk-neutral case.*

Proof. If $\psi \leq f(x, \omega)$ for every $\omega \in \Omega$, then

$$\max\{f(x, \omega) - \psi, 0\} = f(x, \omega) - \psi \quad (3.4)$$

Substituting result 3.4 in the formulation 3.3, we get

$$\begin{aligned} \mathbf{Min}_{x \in X} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \mathbb{E}[\max\{f(x, \tilde{\omega}) - \psi, 0\}] &= \mathbf{Min}_{x \in X} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \mathbb{E}[f(x, \tilde{\omega}) - \psi] \\ &= \mathbf{Min}_{x \in X} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \mathbb{E}[f(x, \tilde{\omega})] - \lambda \mathbb{E}[\psi] \\ &= -\lambda \psi + (1 + \lambda) \mathbf{Min}_{x \in X} \quad \mathbb{E}[f(x, \tilde{\omega})] \end{aligned}$$

□

LEMMA 3.2.2. *If for risk measure EE, the target $\psi \in \mathbb{R}$ is such that $\psi \geq f(x, \omega)$ for every realization $\omega \in \Omega$. Then, the optimal decision vector x will be same as that of the risk-neutral case.*

Proof. If $\psi \geq f(x, \omega)$ for every $\omega \in \Omega$, then

$$\max\{f(x, \omega) - \psi, 0\} = 0 \quad (3.5)$$

Substituting result 3.5 in the formulation 3.3, we get

$$\begin{aligned} \underset{x \in X}{\text{Min}} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \mathbb{E}[\max\{f(x, \tilde{\omega}) - \psi, 0\}] &= \underset{x \in X}{\text{Min}} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \mathbb{E}[0] \\ &= \underset{x \in X}{\text{Min}} \quad \mathbb{E}[f(x, \tilde{\omega})] \end{aligned}$$

□

3.2.2 Quantile Deviation

Given $\epsilon_1, \epsilon_2 > 0$, let $\alpha = \epsilon_2 / (\epsilon_1 + \epsilon_2)$. The quantile deviation (QDEV) [41] risk measure is defined as follows:

$$\phi_{QDEV_{\epsilon_1, \epsilon_2}}(x) := \underset{\psi}{\text{Min}} \{ \mathbb{E}[\epsilon_1 \max\{\psi - f(x, \tilde{\omega}), 0\}] + \epsilon_2 \max\{f(x, \tilde{\omega}) - \psi, 0\} \}.$$

The QDEV captures the average two-sided deviation from the α -quantile. Setting $\rho := \phi_{QDEV_{\epsilon_1, \epsilon_2}}$ in formulation (3.1) we obtain the mean-QDEV as follows:

$$\underset{x \in X}{\text{Min}} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \phi_{QDEV_{\epsilon_1, \epsilon_2}}(x). \quad (3.6)$$

3.2.3 Conditional Value-at-Risk

Given $\alpha \in (0, 1)$, the conditional value-at-risk (CVaR $_{\alpha}$) [9] is defined as:

$$\phi_{CVaR_{\alpha}}(x) := \underset{\psi}{\text{Min}} \{ \psi + \frac{1}{1 - \alpha} \mathbb{E}[\max\{f(x, \tilde{\omega}) - \psi, 0\}] \}.$$

It can be shown that at optimality ψ is an alpha-quantile of the distribution of $f(x, \tilde{\omega})$ [9]. If $f(x, \tilde{\omega})$ represents costs, then CVaR reflects the average losses higher than the α -quantile. Setting $\rho := \phi_{CVaR_\alpha}$ in formulation (3.1) we obtain the mean-CVaR risk measure as follows:

$$\text{Min}_{x \in X} \quad \mathbb{E}[f(x, \tilde{\omega})] + \lambda \phi_{CVaR_\alpha}(x). \quad (3.7)$$

The CVaR is widely used in the field of financial risk management to evaluate market risk or credit risk of a portfolio [42, 18, 43].

3.3 Properties of Coherent Risk Measures

Let \mathcal{F} be set of all real-valued random variables. For a random variable S defined on probability space $(\Omega, \mathcal{A}, \mathcal{P})$, we will have $S \in \mathcal{F}$. A risk function ρ is said to be coherent if it satisfies all the following properties:

- **Translation invariance:** If $a \in \mathbb{R}$ and $S \in \mathcal{F}$, then

$$\rho(S + a) = a + \rho(S).$$
- **Positive homogeneity:** If $c > 0$ and $S \in \mathcal{F}$, then

$$\rho(cS) = c\rho(S).$$
- **Monotonicity:** If $S_1, S_2 \in \mathcal{F}$ and $S_1 \leq S_2$, then

$$\rho(S_1) \leq \rho(S_2).$$
- **Convexity:** If $S_1, S_2 \in \mathcal{F}$ and $\lambda \in (0, 1)$, then

$$\rho(\lambda S_1 + (1 - \lambda)S_2) \leq \lambda\rho(S_1) + (1 - \lambda)\rho(S_2).$$

EE is not a coherent risk measure, nor is the objective function value of formulation (3.3), for a fixed value of x . While EE satisfies properties of monotonicity and convexity, it fails to satisfy the properties of *translation invariance* and *positive homogeneity*. Like EE, QDEV is not a coherent risk measure as it violates *translation invariance* and *monotonicity*. This is also true for the objective function of formulation (3.6). Among the three risk measures considered in this

work, only CVaR satisfies all the four properties required to be a coherent risk measure and the objective function of formulation (3.7) is coherent under special condition. Table 3.1 provides the list of properties satisfied by each risk measure and and table 3.2 provides the list of properties satisfied by each mean-risk formulation. The proof of properties associated with each risk measure is provided in AppendixA. For the Mean-CVaR to be coherent we can redefined (3.7) as follows: $\text{Min}(1 - \lambda)\mathbb{E}[f(x, \tilde{\omega})] + \lambda\rho[f(x, \tilde{\omega})]$, where ρ is the risk measure and λ is the trade-off value.

Table 3.1: Properties Satisfied by Risk Measures

Risk Measure	Translational Invariance	Positive Homogeneity	Monotonicity	Convexity
Expected Excess	X	X	✓	✓
Absolute Semi-Deviation	X	✓	X	✓
Quantile Deviation	X	✓	X	✓
CVaR	✓	✓	✓	✓

Table 3.2: Properties Satisfied by Mean-Risk Measures

Risk Measure	Translational Invariance	Positive Homogeneity	Monotonicity	Convexity
Expected Excess	X	X	✓	✓
Absolute Semi-Deviation	✓	✓	X	✓
Quantile Deviation	✓	✓	X	✓
CVaR*	✓	✓	✓	✓

3.4 Deterministic Equivalent Formulation of MR-SLP

Let us now establish the deterministic equivalent formulation for all of the risk measures using formulation (3.1) and the definitions provided in subsection 3.2. The following formulations are then used to explain the decomposition approach used in SD.

3.4.1 MR-SLP with EE

For a discrete random variable $\tilde{\omega}$, the deterministic equivalent formulation for mean EE is given by the following proposition:

PROPOSITION 3.4.1. *Given $\lambda \geq 0$ and a target $\psi \in \mathbb{R}$, problem (3.1) with $\rho = \phi_{EE_\psi}$ is equivalent to the following formulation [19]:*

$$\begin{aligned}
 \underset{x,y,\nu}{\text{Min}} \quad & c^\top x + \mathbb{E}[q^\top y(\tilde{\omega}) + \lambda \nu(\tilde{\omega})] & (3.8) \\
 \text{s.t.} \quad & T(\omega)x + Wy(\omega) \geq r(\omega), & \forall \omega \in \Omega \\
 & -c^\top x - q^\top y(\omega) + \nu(\omega) \geq -\psi, & \forall \omega \in \Omega \\
 & x \in X, y(\omega) \in \mathbb{R}_+, \nu(\omega) \in \mathbb{R}_+, & \forall \omega \in \Omega,
 \end{aligned}$$

where decision variable $\nu(\omega)$ measures the excess above the target ψ for scenario ω .

Formulation (3.8), is a MR-SLP with dual block angular structure since the recourse decision vector $y(\omega)$ and decision variable $\nu(\omega)$ do not appear in the constraint $x \in X$. It is therefore amenable to Benders decomposition [44].

3.4.2 MR-SLP with QDEV

For a discrete random variable $\tilde{\omega}$, the deterministic equivalent formulation for mean QDEV is given by the following proposition:

PROPOSITION 3.4.2. *Given $\epsilon_1, \epsilon_2 > 0, \lambda \in [0, \frac{1}{\epsilon_1}]$, let $\alpha = \epsilon_2 / (\epsilon_1 + \epsilon_2)$, problem (3.1) with*

$\rho = \phi_{QDEV_{\epsilon_1, \epsilon_2}}$ is equivalent to the following formulation [24]:

$$\begin{aligned}
\text{Min}_{x, \psi, y, \nu} \quad & (1 - \lambda\epsilon_1)c^\top x + \lambda\epsilon_1\psi + (1 - \lambda\epsilon_1)\mathbb{E}[q(\tilde{\omega})^\top y(\tilde{\omega})] + \lambda(\epsilon_1 + \epsilon_2)\mathbb{E}[\nu(\tilde{\omega})] & (3.9) \\
\text{s.t.} \quad & T(\omega)x + Wy(\omega) \geq r(\omega), & \forall \omega \in \Omega \\
& -c^\top x - q^\top y(\omega) + \psi + \nu(\omega) \geq 0, & \forall \omega \in \Omega \\
& x \in X, \psi \in \mathbb{R}, y(\omega) \in \mathbb{R}_+, \nu(\omega) \in \mathbb{R}_+, & \forall \omega \in \Omega,
\end{aligned}$$

where ψ is the α -quantile of $f(x, \tilde{\omega})$ at optimality.

Formulation (3.9) has a dual block angular structure and is therefore amenable to Benders decomposition.

3.4.3 MR-SLP with CVaR

For a discrete random variable $\tilde{\omega}$, the deterministic equivalent formulation for mean CVaR is given by the following proposition:

PROPOSITION 3.4.3. *Given $\lambda \geq 0$ and $\alpha \in (0, 1)$, problem (3.1) with $\rho = \phi_{CVaR_\alpha}$ is equivalent to the following formulation [19]:*

$$\begin{aligned}
\text{Min}_{x, \psi, y, \nu} \quad & c^\top x + \mathbb{E}[q^\top y(\tilde{\omega})] + \lambda\psi + \frac{\lambda}{1 - \alpha}\mathbb{E}[\nu(\tilde{\omega})] & (3.10) \\
\text{s.t.} \quad & T(\omega)x + Wy(\omega) \geq r(\omega), & \forall \omega \in \Omega \\
& -c^\top x - q^\top y(\omega) + \psi + \nu(\omega) \geq 0, & \forall \omega \in \Omega \\
& x \in X, \psi \in \mathbb{R}, y(\omega) \in \mathbb{R}_+, \nu(\omega) \in \mathbb{R}_+, & \forall \omega \in \Omega.
\end{aligned}$$

where ψ is a first stage decision variable. At optimality its value corresponds to the α -quantile of $f(x, \tilde{\omega})$.

Formulation (3.10), also has a dual block angular structure and is therefore amenable to Benders decomposition.

3.5 MR-SLP

To define a MR-MSLP with T stages, let $t \in \{1, 2, \dots, T\}$ be the stage index and consider a stochastic process $\tilde{\omega} = (\tilde{\omega}_1, \tilde{\omega}_2, \dots, \tilde{\omega}_T)$ and a decision process $x = (x_1, x_2, \dots, x_T)$. The component x_1 is a nonrandom vector-valued decision variable and $\tilde{\omega}_1$ is deterministic. The rest of the components x_2, \dots, x_T of x and $\tilde{\omega}_2, \dots, \tilde{\omega}_T$ of $\tilde{\omega}$ are random vectors, not necessarily of the same dimension, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The $\tilde{\omega}_t \in \Omega_t$, where $\Omega_t \subseteq \Omega$ is the sample space of random outcomes at stage t and $\Omega_1 \subseteq \Omega_2 \dots \subseteq \Omega_t \subseteq \Omega_{t+1} \dots \subseteq \Omega_T \subseteq \Omega$. The sequential decision and stochastic data process is

$$x_1, \tilde{\omega}_2, x_2(x_1, \tilde{\omega}_2), \dots, x_T(x_{T-1}, \tilde{\omega}_2, \dots, \tilde{\omega}_T).$$

The decisions made at a given stage do not depend on any specific future outcomes of the stochastic data or on future decisions, that is, the decision process is *nonanticipative*. Mathematically, $\mathcal{A}_t \subseteq \mathcal{A}$ is a σ -field generated by $\omega_{[t]} := (\omega_1, \dots, \omega_t)$ of the stochastic process $\tilde{\omega}$ that includes the stochastic data up to stage t , and $\mathcal{A}_1 = \{\emptyset, \Omega\}$ is the trivial σ -field. Let $\tilde{\omega}_{[t+1]}$ denote the stochastic process at stage $t+1$ given $\omega_{[t]}$. Since the decision x_t at stage t depends only on the available information, it means that it is \mathcal{A}_t -measurable.

We have $x_{[t]} := (x_1, \dots, x_t)$ to be the sequence of decisions at stages $1, \dots, t$ and \mathbb{P}_t the marginal distribution of ω_t . The stochastic process is considered *stagewise independent* if the process $\tilde{\omega}_t$ is independent of $\omega_{[t-1]}$, for all $t = 2, \dots, T$. For $1 \leq t_1 < t_2 \leq T$, let the set $X_{t_1}^*$ be the set of optimal solutions for stages $t \in \{t_1, \dots, T\}$ of MR-SLP conditional on a realization $\omega_{[t_1]}$. Then, the MR-SLP is said to be time consistent if $X_{t_2}^* \subset X_{t_1}^*$ is an set of optimal solution conditional on realization $\omega_{[t_2]}$ given realization $\omega_{[t_1]}$.

Let $\rho_t : \mathcal{F} \mapsto \mathbb{R}$ be a risk measure at stage t , where \mathcal{F} is set of real random cost variables defined on probability space $(\Omega, \mathcal{A}, \mathbb{P})$. A risk measure is a function that complements the expected value, for instance by avoiding solutions with high variability, or ones that deviate significantly from a target. Let $\lambda_t > 0$ be the weighting factor that quantifies the trade-off between expected cost and

risk. Then a MR-MSLP with $T \geq 2$ stages, can be formally stated as follows:

$$\begin{aligned} \text{MSLP1 : } \quad & \underset{x_1}{\text{Min}} \quad c_1^\top x_1 + \mathbb{E}[f_2(x_1, \tilde{\omega}_2)] + \lambda_2 \rho_2[f_2(x_1, \tilde{\omega}_2)] \\ & \text{s.t.} \quad x_1 \in X_1, \end{aligned} \quad (3.11)$$

where $x_1 \in \mathbb{R}_+^{n_1}$ is a vector of first-stage decision variables, $c_1 \in \mathbb{R}^{n_1}$ is the cost vector associated with the first-stage and $X_1 \subseteq \mathbb{R}^{n_1}$ is a compact set of feasible first-stage decisions. It is desirable that the MSLP1 maintains suitable properties such as convexity and dual block angular structure to allow the formulation to be amenable to decomposition methods.

Let $x_t(\omega_t) \in \mathbb{R}_+^{n_t}$ be decision vector, $c_t \in \mathbb{R}^{n_t}$ be the cost vector and $X_t(x_{t-1}, \omega_{[t]}) \subseteq \mathbb{R}^{n_t}$ be set of feasible decisions at stage $t \in \{2, \dots, T-1\}$ for realization ω_t . Then the real random cost variable $f_t(x_{t-1}, \omega_{[t]})$ at stage t is defined recursively as

$$\begin{aligned} f_t(x_{t-1}, \omega_{[t]}) := & \underset{x_t(\omega_t) \in X_t(x_{t-1}, \omega_{[t]})}{\text{Min}} \quad c_t^\top x_t(\omega_t) + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]}) \mid \omega_{[t]}] \\ & + \lambda_{t+1} \rho_{t+1}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]}) \mid \omega_{[t]}], \end{aligned} \quad (3.12)$$

where $\mathbb{E}[\cdot \mid \omega_{[t]}]$ is the conditional expectation and $\rho_{t+1}[\cdot \mid \omega_{[t]}]$ is the conditional risk measure. For the final stage $t = T$ and a given feasible solution $x_{T-1} \in X_{T-1}(x_{T-2}, \omega_{[T-1]})$, the real random cost variable $f_T(x_{T-1}, \omega_{[T]})$ is defined as

$$f_T(x_{T-1}, \omega_{[T]}) := \underset{x_T(\omega_T) \in X_T(x_{T-1}, \omega_{[T]})}{\text{Min}} \quad c_T^\top x_T(\omega_T). \quad (3.13)$$

For MSLP1 to be well-defined, we consider it under the following assumptions:

- (A1)** The set of first stage decisions X_1 and Ω are compact.
- (A2)** At every stage, for any given $x_t(\omega_t) \in X_t(x_{t-1}, \omega_{[t]})$, $\mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]})] < \infty$.
- (A3)** At every stage recourse matrix W_t is fixed (fixed recourse).
- (A4)** For all $x_t(\omega_t) \in X_t(x_{t-1}, \omega_{[t]})$, there exists a constant L_t such that $L_t \leq \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]})]$.

(A5) The stochastic process $\tilde{\omega}$ has finite support.

Assumptions (A1) and (A2) guarantee the existence of an optimal solution in MSD. Satisfying assumption (A3) guarantees relatively complete recourse at every stage, which enables having a dual feasible set for the subproblem. Since in MSD subgradients are generated using one scenario at a time, assumption (A4) is required to make sure that the optimal solution is not cut-off by the optimality cuts.

3.6 Deterministic Equivalent Formulation of MR-MSLP

In this section we define the mean-risk formulation for MSLP using the definitions of risk measures provided in Subsection 3.2. The formulations are defined recursively for each stage of MSLP and they are used to explain the derivation of MSD algorithms for MR-MSLP.

3.6.1 MR-MSLP with EE

For a scenario $\omega \in \Omega$, the deterministic equivalent formulation for MR-MSLP with EE as risk measure and T stages is given by the following proposition:

PROPOSITION 3.6.1. *Given $\lambda_t \geq 0$ and a target $\psi_t \in \mathbb{R}$, problem (3.11) with $\rho = \phi_{EE\psi}$ is equivalent to the following formulation:*

At stage $t = 1$,

$$\begin{aligned} \text{Min} \quad & c_1^\top x_1 + \mathbb{E}[f_2(x_1, \tilde{\omega}_2)] & (3.14) \\ \text{s.t.} \quad & W_1 x_1 \geq r_1, \\ & x_1 \in X_1. \end{aligned}$$

For stages $t = \{2, 3, \dots, T - 1\}$, $f_t(x_{t-1}, \omega_{[t]})$ is defined recursively as follows:

$$\begin{aligned}
f_t(x_{t-1}, \omega_{[t]}) &:= \text{Min} \quad c_t^\top x_t(\omega_t) + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]}) \mid \omega_{[t]}] + \lambda_t \nu_t(\omega_t) & (3.15) \\
\text{s.t.} \quad & W_t x_t(\omega_t) \geq r_t(\omega_t) - T_t(\omega_t) x_{t-1}(\omega_{t-1}), & \forall \omega \in \Omega \\
& -c_t^\top x_t(\omega_t) - \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]}) \mid \omega_{[t]}] \\
& \quad + \nu_t(\omega_t) \geq -\psi_t + c_{t-1}^\top x_{t-1}(\omega_{t-1}), & \forall \omega \in \Omega \\
& x_t(\omega_t) \in X_t(x_{t-1}, \omega_{[t]}), \nu_t(\omega_t) \geq 0. & \forall \omega \in \Omega,
\end{aligned}$$

For terminal stage $t = T$,

$$\begin{aligned}
f_T(x_{T-1}, \omega_{[T]}) &:= \text{Min} \quad c_T^\top x_T(\omega_T) + \lambda_T \nu_T(\omega_T) & (3.16) \\
\text{s.t.} \quad & W_T x_T(\omega_T) \geq r_T(\omega_T) - T_T(\omega_T) x_{T-1}(\omega_{T-1}), & \forall \omega \in \Omega \\
& -c_T^\top x_T(\omega_T) + \nu_T(\omega_T) \geq -\psi_T + c_{T-1}^\top x_{T-1}(\omega_{T-1}), & \forall \omega \in \Omega \\
& x_T(\omega_T) \in X_T(x_{T-1}, \omega_{[T]}), \nu_T(\omega_T) \geq 0, & \forall \omega \in \Omega,
\end{aligned}$$

where decision variable $\nu_t(\omega_t)$ measures the excess above the target ψ_t in stage t for scenario ω .

Formulations (3.14 -3.16) have a dual block angular structure, since the decision vector $x_{t+1}(\omega_{t+1})$ and the variable $\nu_{t+1}(\omega_{t+1})$ do not appear in the constraint $x_t \in X_t(x_{t-1}, \omega_{[t]})$. Therefore, this formulation is amenable to Benders decomposition [44].

3.6.2 MR-MSLP with QDEV

For a specified quantile α and a given scenario $\omega \in \Omega$, the MR-MSLP formulation with QDEV can be recursively defined as follows:

PROPOSITION 3.6.2. *Given $\epsilon_1, \epsilon_2 > 0$, $\lambda_t \in [0, 1/\epsilon_1]$, let $\alpha = \epsilon_2/(\epsilon_1 + \epsilon_2)$, problem (3.11) with $\rho = \phi_{QDEV_{\epsilon_1, \epsilon_2}}$ is equivalent to the following formulation:*

At stage $t = 1$,

$$\begin{aligned}
\text{Min} \quad & (1 - \lambda_2 \epsilon_1) c_1^\top x_1 + \lambda_2 \epsilon_1 \psi_1 + \mathbb{E}[f_2(x_1, \tilde{\omega}_2)] & (3.17) \\
\text{s.t.} \quad & W_1 x_1 \geq r_1, \\
& x_1 \in X_1, \psi_1 \in \mathbb{R}.
\end{aligned}$$

For stages $t = \{2, 3, \dots, T - 1\}$, the function $f_t(x_{t-1}, \omega_{[t]})$ is defined recursively as follows:

$$\begin{aligned}
f_t(x_{t-1}, \omega_{[t]}) := \text{Min} \quad & (1 - \lambda_t \epsilon_1) [(1 - \lambda_{t+1} \epsilon_1) c_t^\top x_t(\omega_t) + \lambda_{t+1} \epsilon_1 \psi_t(\omega_t) \\
& + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]}) \mid \omega_{[t]}]] + \lambda_t (\epsilon_1 + \epsilon_2) \nu_t(\omega_t) & (3.18)
\end{aligned}$$

$$\begin{aligned}
\text{s.t.} \quad & W_t x_t(\omega_t) \geq r_t(\omega_t) - T_t(\omega) x_{t-1}(\omega_{t-1}), & \forall \omega \in \Omega \\
& - (1 - \lambda_{t+1} \epsilon_1) c_t^\top x_t(\omega_t) - \lambda_{t+1} \epsilon_1 \psi_t(\omega_t) - \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]}) \mid \omega_{[t]}] \\
& + \nu_t(\omega_t) \geq -\psi_{t-1}(\omega_{t-1}) + c_{t-1}^\top x_{t-1}(\omega_{t-1}), & \forall \omega \in \Omega \\
& & (3.19)
\end{aligned}$$

$$x_t(\omega) \in X_t(x_{t-1}, \omega_{[t]}), \psi_t(\omega_t) \in \mathbb{R}, \nu_t(\omega_t) \geq 0, \quad \forall \omega \in \Omega.$$

Constraint (3.19) is used to compute the excess of current and future cost at stage t over quantile $\psi_{t-1}(\omega)$ determined at stage $t - 1$. At terminal stage $t = T$,

$$\begin{aligned}
f_T(x_{T-1}, \omega_{[T]}) := \text{Min} \quad & (1 - \lambda_T \epsilon_1) c_T^\top x_T(\omega_T) + \lambda_T (\epsilon_1 + \epsilon_2) \nu_T(\omega_T) & (3.20) \\
\text{s.t.} \quad & W_T x_T(\omega_T) \geq r_T(\omega_T) - T_T(\omega_T) x_{T-1}(\omega_{T-1}), & \forall \omega \in \Omega \\
& - c_T^\top x_T(\omega_T) + \nu_T(\omega_T) \geq -\psi_{T-1}(\omega_{T-1}) + c_{T-1}^\top x_{T-1}(\omega_{T-1}), & \forall \omega \in \Omega \\
& x_T(\omega_T) \in X_T(x_{T-1}, \omega_{[T]}), \nu_T(\omega_T) \geq 0, & \forall \omega \in \Omega,
\end{aligned}$$

where ψ_t is the α -quantile of $f_t(x_{t-1}, \tilde{\omega}_{[t]})$ at optimality.

Formulations (3.17-3.20) have dual block angular structure and therefore, are amenable to Benders decomposition.

3.6.3 MR-MSLP with CVaR

For a discrete random variable $\tilde{\omega}$, the deterministic equivalent formulation for MR-MSLP with CVaR is given by the following proposition [12]:

PROPOSITION 3.6.3. *Given $\lambda \geq 0$ and $\alpha \in (0, 1)$, problem (3.11) with $\rho = \phi_{\text{CVaR}_\alpha}$ is equivalent to the following formulation:*

At stage $t = 1$,

$$\begin{aligned} \text{Min} \quad & c_1^\top x_1 + \lambda_2 \psi_1 + \mathbb{E}[f_2(x_1, \tilde{\omega}_2)] & (3.21) \\ \text{s.t.} \quad & W_1 x_1 \geq r_1, \\ & x_1 \in X_1, \psi_1 \in \mathbb{R}. \end{aligned}$$

For stages $t = \{2, 3, \dots, T-1\}$, the function $f_t(x_{t-1}, \omega_{[t]})$ is defined as follows:

$$\begin{aligned} \text{Min} \quad & c_t^\top x_t(\omega_t) + \lambda_{t+1} \psi_t(\omega_t) + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]} \mid \omega_{[t]})] + \frac{\lambda_t}{1-\alpha} \nu_t(\omega_t) & (3.22) \\ \text{s.t.} \quad & W_t x_t(\omega_t) \geq r_t(\omega_t) - T_t(\omega_t) x_{t-1}(\omega_{t-1}), & \forall \omega \in \Omega \\ & -c_t^\top x_t(\omega_t) - \lambda_{t+1} \psi_t(\omega_t) - \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t+1]} \mid \omega_{[t]})] + \\ & \quad \nu_t(\omega_t) \geq -\psi_{t-1}(\omega_{t-1}) + c_{t-1}^\top x_{t-1}(\omega_{t-1}), & \forall \omega \in \Omega \\ & x_t(\omega_t) \in X_t(x_{t-1}, \omega_{[t]}), \psi_t(\omega_t) \in \mathbb{R}, \nu_t(\omega_t) \geq 0. & \forall \omega \in \Omega \end{aligned}$$

At terminal stage $t = T$,

$$\begin{aligned} f_T(x_{T-1}, \omega_{[T]}) := \text{Min} \quad & c_T^\top x_T(\omega_T) + \frac{\lambda_T}{1-\alpha} \nu_T(\omega_T) & (3.23) \\ \text{s.t.} \quad & W_T x_T(\omega_T) \geq r_T(\omega_T) - T_T(\omega_T) x_{T-1}(\omega_{T-1}), & \forall \omega \in \Omega \\ & -c_T^\top x_T(\omega_T) + \nu_T(\omega_T) \geq -\psi_{T-1}(\omega_{T-1}) + c_{T-1}^\top x_{T-1}(\omega_{T-1}), & \forall \omega \in \Omega \\ & x_T(\omega_T) \in X_T(x_{T-1}, \omega_{[T]}), \nu_T(\omega_T) \geq 0, & \forall \omega \in \Omega, \end{aligned}$$

where ψ_t is a decision variable at stage t and at optimality its value corresponds to the α -quantile of $f_t(x_{t-1}, \tilde{\omega}_{[t]})$ [9].

Formulations (3.21-3.23), have dual block angular structure and are amenable to Benders decomposition. Using the definitions and the deterministic equivalent formulations that are presented in section 3.4, in the next section we provide details of the SD approach and outline the risk-averse SD algorithm for MR-SLP.

4. STOCHASTIC DECOMPOSITION ALGORITHM

Now that in Section 3 we have presented the definitions and the formulations of MR-SLP, in this section we present a detailed decomposition approach for each risk measure and we follow it by outlining mean-risk SD algorithms for MR-SLP along with the proof of convergence to optimality.

4.1 Stochastic Decomposition Algorithm for EE

We begin by presenting the details of our decomposition approach for MR-SLP with EE. We first decompose problem (3.8), defined in Section 3, into *master problem* and *subproblem*. This is possible due to the dual block angular structure of the formulation. The fundamental idea of the SD algorithm involves solving one *subproblem* and one *master problem* at each iteration. The *subproblem* is then used to recursively update a piecewise linear approximation of the recourse function at each iteration k , while the *master problem* is used to generate successive iterates $\{x^k\}$.

4.1.1 Decomposition for MR-SLP with EE

The deterministic equivalent formulation (3.8) has a dual block angular structure and therefore, we can decompose the MR-SLP with EE as follows:

$$\begin{aligned} \text{Min}_x \quad & c^\top x + \mathbb{E}[h(x, \tilde{\omega})] & (4.1) \\ \text{s.t.} \quad & x \in X, \end{aligned}$$

where for a realization $\omega \in \Omega$, we have

$$h(x, \omega) := \text{Min}_{y, \nu} \quad q^\top y(\omega) + \lambda \nu(\omega) \tag{4.2}$$

$$\text{s.t.} \quad Wy(\omega) \geq r(\omega) - T(\omega)x, \quad \forall \omega \in \Omega, \tag{4.3}$$

$$-q^\top y(\omega) + \nu(\omega) \geq c^\top x - \psi, \quad \forall \omega \in \Omega, \tag{4.4}$$

$$y(\omega) \geq 0, \nu(\omega) \geq 0, \quad \forall \omega \in \Omega. \tag{4.5}$$

Notice that unlike function $h(x, \tilde{\omega})$ define in (3.2), $h(x, \tilde{\omega})$ in (4.2) includes a weighted deviation variable $\nu(\omega)$. Given an iterate x^k at iteration k and an outcome ω^k , the dual of problem (4.2)-(4.5) can be stated as follows:

$$\begin{aligned}
h(x^k, \omega^k) &:= \text{Max} \quad \pi(\omega^k)^\top (r(\omega^k) - T(\omega^k)x) + \phi(\omega^k)(c^\top x^k - \psi) & (4.6) \\
&\text{s.t.} \quad \pi(\omega^k)^\top W - \phi(\omega^k)q \leq q \\
&\quad \phi(\omega^k) \leq \lambda \\
&\quad \pi(\omega^k) \geq 0, \phi(\omega^k) \geq 0,
\end{aligned}$$

where, $\pi(\omega^k)$ and $\phi(\omega^k)$ are the dual multipliers associated with constraints (4.3) and (4.4) respectively.

The solution of dual problem (4.6) at iteration k , together with the dual solutions of all the past iterations can be used to obtain a piecewise linear approximation of the expected recourse function $\mathbb{E}[h(x, \tilde{\omega})]$, which can be defined as follows:

$$\eta^k(x) := \max\{\alpha_t^k + (\beta_t^k)^\top x \mid t = 1, \dots, k\}, \quad (4.7)$$

where k is the current iteration number, t is the counter for all the iterations performed up to k , α_t^k is the cut constant and β_t^k is the cut coefficient generated for t -th iteration at iteration k . Therefore, the peicewise linear approximation of objective function at iteration k is

$$F_k(x) := c^\top x + \eta^k,$$

where η^k is an auxiliary variable that describes the outer approximation of the recourse function.

Now using the cut approximation $\eta^k(x)$ we can define a master program as follows:

$$\begin{aligned}
& \underset{x, \eta^k}{\text{Min}} && c^\top x + \eta^k && (4.8) \\
& \text{s.t.} && x \in X \\
& && \eta^k - (\beta_t^k)^\top x \geq \alpha_t^k, \quad \forall t = 1, \dots, k.
\end{aligned}$$

4.1.2 SD Algorithm for MR-SLP with EE

Using the master problem and sub problems defined in Section 4.1.1, the SD algorithm for mean EE can be stated as follows:

SD-EE Algorithm

Step 0: Initialization.

Set $k \leftarrow 0$; $V^0 \leftarrow \emptyset$; $U^0 \leftarrow \emptyset$; $\eta^0 = -\infty$; $x^0 \in X$; $\bar{x}^0 \in X$; $\psi \in \mathbb{R}$ given (EE target); $L \in \mathbb{R}$ given (lower bound) and $\delta \in (0, 1)$.

Step 1: Generate a Scenario.

$k \leftarrow k + 1$. Randomly generate scenario ω^k of $\tilde{\omega}$, independent of previously generated scenarios.

Step 2: Determining Cut Approximation $\eta^k(x)$.

a. Solve dual problem (4.2) for scenario ω^k

$$\begin{aligned}
& (\pi^k(\omega^k), \phi^k(\omega^k)) \in \underset{\pi, \phi}{\text{argmax}} \{ (\pi^k)^\top (r(\omega^k) - T(\omega^k)x^k) + \phi^k (c^\top x - \psi) \\
& \quad | \pi^\top W - q\phi^k \leq q, \phi^k \leq \lambda, \pi^k, \phi^k \geq 0 \}.
\end{aligned}$$

b. Update sets V^k and U^k .

$$V^k \leftarrow V^{k-1} \cup \pi^k(\omega^k).$$

$$U^k \leftarrow U^{k-1} \cup \phi^k(\omega^k).$$

c. Determine the coefficients of the k -th cutting plane.

$$(\pi_t^k, \phi_t^k) \in \operatorname{argmax} \{ (\pi)^\top (r(\omega^t) - T(\omega^t)x^k) + \phi(c^\top x^k - \psi) \mid \pi \in V^k, \phi \in U^k \}, \forall t = \{1, \dots, k-1\}.$$

$$\alpha_k^k = \frac{1}{k} \sum_{t=1}^k \{ (\pi_t^k)^\top r(\omega^t) - \phi_t^k \psi \}.$$

$$(\beta_k^k)^\top = \frac{1}{k} \sum_{t=1}^k \{ (\pi_t^k)^\top (-T(\omega^t)) + \phi_t^k c^\top \}.$$

d. Update coefficients of all previously generated cuts.

$$\alpha_t^k \leftarrow \left(\frac{k-1}{k}\right)\alpha_t^{k-1} + \left(\frac{1}{k}\right)L, \quad \forall t = 1, \dots, k-1.$$

$$\beta_t^k \leftarrow \left(\frac{k-1}{k}\right)\beta_t^{k-1}, \quad \forall t = 1, \dots, k-1.$$

e. Add updated cuts to master problem.

$$\eta^k \geq \alpha_t^k + (\beta_t^k)^\top x, \quad \forall t = 1, \dots, k.$$

$$\Rightarrow \eta^k - (\beta_t^k)^\top x \geq \alpha_t^k, \quad \forall t = 1, \dots, k.$$

Step 3. Updating Incumbent Solution.

$$\text{if } F_k(x^k) - F_k(\bar{x}^{k-1}) < \delta[F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1})],$$

$$\bar{x}^k \leftarrow x^k$$

$$\bar{x}^k \leftarrow \bar{x}^{k-1}.$$

Step 4. Solve Master Problem.

$$\operatorname{Min}_{x, \eta^k} c^\top x + \eta^k$$

$$\text{s.t. } x \in X,$$

$$\eta^k - (\beta_t^k)^\top x \geq \alpha_t^k, \quad \forall t = \{1, \dots, k\},$$

to get the new candidate solution x^{k+1} .

Step 5. Termination Criterion.

One of the following criteria can be used to terminate the algorithm. If the selected

criterion is not satisfied, then return to step 1 and repeat the steps.

a. Termination criterion based on incumbent test:

Let m_k be the number of times the incumbent solution changes up to iteration k and let $\{\bar{x}^{k_n}\}_{k_n=1}^{m_k}$ a subsequence of $\{x^n\}_{n=1}^k$ be a collection of all incumbent solutions up to iteration k . Define

$$\gamma^k = \frac{1}{k} \sum_{t=1}^k F_t(\bar{x}^t) \text{ and } \bar{\gamma}^k = \frac{1}{m_k} \sum_{n=1}^{m_k} F_{k_n}(\bar{x}^{k_n}).$$

Then terminate the algorithm for a large enough k , if

$$\left| \frac{F_k(\bar{x}^k) - \gamma^{k-1}}{\bar{\gamma}^{k-1}} \right| \leq \epsilon, \text{ where } \epsilon > 0 \text{ is a given tolerance level.}$$

b. Termination criteria based on objective value:

For a large enough k , terminate the algorithm if

$$F_{k-1}(\bar{x}^{k-1}) - F_{k-1}(x^k) \leq \epsilon, \text{ where } \epsilon > 0 \text{ is a given tolerance level.}$$

REMARK 4.1.1. In step 2 of the proposed SD-EE algorithm, we generate a random scenario ω^k independent of previously generated scenarios based on the probability distribution of Ω , which can be done using Monte Carlo. In each iteration, at step 2 of the algorithms we obtain a solution of one dual subproblem and at step 4 we have the solution of the master problem. The former is used to recursively update a piecewise linear approximation of the recourse function, while the latter is used to generate successive candidate solutions. Together they provide us with a valid lower bound on the objective function.

After a moderate number of iterations, it is expected that the frequency of updating the incumbent solution as described in step 3 will decrease. The termination criterion described in step 5(a) makes use of this fact to terminate the algorithm once a stable solution is reached. The termination criterion described in step 5(b) is based on assumption that after a large enough number of iterations the algorithm will stabilize, and therefore will reach the accumulation point of the

incumbent solution.

4.2 Stochastic Decomposition for QDEV

Following the similar approach as that of EE, we begin by decomposing the deterministic equivalent problem (3.9) of risk measure QDEV. The problem has a dual block angular structure and therefore can be decomposed into *master problem* and *subproblem*. Unlike EE, in risk measure QDEV an α -quantile needs to be estimated in the first stage of SLP for computation of risk.

4.2.1 Decomposition for MR-SLP with QDEV

The deterministic equivalent formulation (3.9), of the MR-SLP with QDEV is decomposed as follows:

$$\begin{aligned} \text{Min}_{x, \psi} \quad & (1 - \lambda\epsilon_1)c^\top x + \lambda\epsilon_1\psi + \mathbb{E}[h(x, \psi, \tilde{\omega})] & (4.9) \\ \text{s.t.} \quad & x \in X, \end{aligned}$$

where

$$h(x, \psi, \omega) := \text{Min}_{y, \nu} \quad (1 - \lambda\epsilon_1)q^\top y(\omega) + \lambda(\epsilon_1 + \epsilon_2)\nu(\omega) \quad (4.10)$$

$$\text{s.t.} \quad Wy(\omega) \geq r(\omega) - T(\omega)x, \quad \forall \omega \in \Omega, \quad (4.11)$$

$$-q^\top y(\omega) + \nu(\omega) \geq c^\top x - \psi, \quad \forall \omega \in \Omega, \quad (4.12)$$

$$y(\omega) \geq 0, \nu(\omega) \geq 0, \quad \forall \omega \in \Omega. \quad (4.13)$$

Given an iterate x^k at iteration k and an outcome ω^k , we solve the dual of problem (4.10)-(4.13) in order to obtain the dual multipliers $\pi(\omega^k)$ and $\phi(\omega^k)$ associated with constraints (4.11) and (4.12) respectively.

The dual solution of (4.10)-(4.13) at iteration k , together with the dual solutions of all the past iterations can be used to obtain a piecewise linear approximation of the expected recourse function

$\mathbb{E}[h(x, \psi, \tilde{\omega})]$, which in the case of QDEV is defined as

$$\eta^k(x) := \max\{\alpha_t^k + (\beta_t^k)^\top x + \gamma_t^k \psi \mid t = 1, \dots, k\}. \quad (4.14)$$

Therefore, the peicewise linear approximation of objective function at iteration k is

$$F_k(x, \psi) := (1 - \lambda\epsilon_1)c^\top x + \lambda\epsilon_1\psi + \eta^k$$

and the master program for MR-SLP with QDEV can be defined as follows:

$$\begin{aligned} \text{Min}_{x, \psi, \eta^k} \quad & (1 - \lambda\epsilon_1)c^\top x + \lambda\epsilon_1\psi + \eta^k \\ \text{s.t.} \quad & x \in X \\ & \eta^k - (\beta_t^k)^\top x - \gamma_t^k \psi \geq \alpha_t^k, \quad \forall t = 1, \dots, k. \end{aligned} \quad (4.15)$$

4.2.2 SD Algorithm for MR-SLP with QDEV

Using the master problem and sub problems defined in section 4.2.1, the SD algorithm for mean QDEV can be stated as follows:

SD-QDEV Algorithm

Step 0: Initialization.

Set $k \leftarrow 0$; $V^0 \leftarrow \emptyset$; $U^0 \leftarrow \emptyset$; $\eta^0 = -\infty$; $x^0 \in X$; $\bar{x}^0 \in X$; $\psi \in \mathbb{R}$ make suitable assumption (α quantile); $L \in \mathbb{R}$ given (lower bound), $\epsilon_1 = 1 - \alpha$, $\epsilon_2 = \alpha$ and $\delta \in (0, 1)$.

Step 1: Generate a Scenario.

$k \leftarrow k + 1$. Randomly generate scenario ω^k of $\tilde{\omega}$, independent of previously generated scenarios.

Step 2: Determining Cut Approximation $\eta^k(x)$.

a. Solve dual problem (4.10) for scenario ω^k

$$(\pi^k(\omega^k), \phi^k(\omega^k)) \in \operatorname{argmax}\{(\pi^k)^\top (r(\omega^k) - T(\omega^k)x^k) + \phi^k(c^\top x - \psi) \\ |\pi^\top W - q\phi^k \leq q, \phi^k \leq \lambda(\epsilon_1 + \epsilon_2), \pi^k, \phi^k \geq 0\}$$

b. Update sets V^k and U^k

$$V^k \leftarrow V^{k-1} \cup \pi^k(\omega^k).$$

$$U^k \leftarrow U^{k-1} \cup \phi^k(\omega^k).$$

c. Determine the coefficients of the k^{th} cutting plane.

$$(\pi_t^k, \phi_t^k) \in \operatorname{argmax}\{(\pi)^\top (r(\omega^t) - T(\omega^t)x^k) + \phi(c^\top x^k - \psi) \\ |\pi \in V^k, \phi \in U^k\}, \forall t = \{1, \dots, k-1\}.$$

$$\alpha_k^k = \frac{1}{k} \sum_{t=1}^k \{(\pi_t^k)^\top r(\omega^t)\}.$$

$$(\beta_k^k)^\top = \frac{1}{k} \sum_{t=1}^k \{(\pi_t^k)^\top (-T(\omega^t)) + \phi_t^k c^\top\}.$$

$$\gamma_t^k = \frac{1}{k} \sum_{t=1}^k -\phi_t^k.$$

d. Update coefficients of all previously generated cuts.

$$\alpha_t^k \leftarrow (k-1/k)\alpha_t^{k-1} + (1/k)L, \quad \forall t = 1, \dots, k-1.$$

$$\beta_t^k \leftarrow (k-1/k)\beta_t^{k-1}, \quad \forall t = 1, \dots, k-1.$$

$$\gamma_t^k \leftarrow (k-1/k)\gamma_t^{k-1}, \quad \forall t = 1, \dots, k-1.$$

e. Add Updated Cuts to Master Problem.

$$\eta^k \geq \alpha_t^k + (\beta_t^k)^\top x + \gamma_t^k \psi, \quad \forall t = 1, \dots, k.$$

$$\Rightarrow \eta^k - (\beta_t^k)^\top x - \gamma_t^k \psi \geq \alpha_t^k, \quad \forall t = 1, \dots, k.$$

Step 3.Updating Incumbent Solution.

$$\text{if } F_k(x^k) - F_k(\bar{x}^{k-1}) < \delta[F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1})]$$

$$\bar{x}^k \leftarrow x^k.$$

else

$$\bar{x}^k \leftarrow \bar{x}^{k-1}.$$

Step 4. Solve the Master Problem.

$$\begin{aligned} \text{Min} \quad & (1 - \lambda\epsilon_1)c^\top x + \lambda\epsilon_1\psi + (1 - \lambda\epsilon_1)\eta^k \\ \text{s.t.} \quad & x \in X, \\ & \eta^k - (\beta_t^k)^\top x - \gamma_t^k\psi \geq \alpha_t^k, \quad \forall t = \{1, \dots, k\}, \end{aligned}$$

to get the new candidate solution x^{k+1} .

Step 5. Termination Criterion.

If the following criterion is not satisfied then starting from step 1 repeat all the steps.

Termination criteria based on objective value:

For a large enough k , terminate the algorithm if

$$F_{k-1}(\bar{x}^{k-1}) - F_{k-1}(x^k) \leq \epsilon,$$

where $\epsilon > 0$ is a given tolerance level.

REMARK 4.2.1. For termination of SD-QDEV, the criterion 5(a) described in SD-EE algorithm can also be used.

4.3 Stochastic Decomposition for CVaR

Similar to QDEV, an α -quantile usually referred as *value-at-risk* is used for computing risk in CVaR. The decision variable ψ in the first stage of SLP is used to estimate the value of α -quantile and at optimality ψ becomes equal to the *value-at-risk*. The deterministic equivalent problem (3.9) of mean-CVaR has a dual block angular structure and therefore it is amenable to decomposition.

4.3.1 Decomposition for MR-SLP with CVaR

Using the deterministic equivalent formulation (3.10), we can decompose the two-stage MR-SLP with CVaR, as follows:

$$\begin{aligned} \text{Min}_{x, \psi} \quad & c^\top x + \lambda \psi + \mathbb{E}[h(x, \psi, \tilde{\omega})] \\ \text{s.t.} \quad & x \in X, \end{aligned} \tag{4.16}$$

where

$$h(x, \psi, \omega) := \text{Min}_{y, \nu} \quad q^\top y(\omega) + \frac{\lambda}{1 - \alpha} \nu(\omega) \tag{4.17}$$

$$\text{s.t.} \quad Wy(\omega) \geq r(\omega) - T(\omega)x, \quad \forall \omega \in \Omega, \tag{4.18}$$

$$-q^\top y(\omega) + \nu(\omega) \geq c^\top x - \psi, \quad \forall \omega \in \Omega, \tag{4.19}$$

$$y(\omega) \geq 0, \nu(\omega) \geq 0, \quad \forall \omega \in \Omega. \tag{4.20}$$

Given an iterate x^k at iteration k and an outcome ω^k , we solve the dual of problem (4.17)-(4.20) to obtain the dual multipliers $\pi(\omega^k)$ and $\phi(\omega^k)$ associated with constraints (4.18) and (4.19) respectively.

The dual solution of (4.17)-(4.20) at iteration k , together with the dual solutions of all the past iterations can be used to obtain a piecewise linear approximation of the expected recourse function $\mathbb{E}[h(x, \tilde{\omega})]$, which in this case is defined as

$$\eta^k(x) := \max\{\alpha_t^k + (\beta_t^k)^\top x + \gamma_t^k \psi \mid t = 1, \dots, k\}. \tag{4.21}$$

Therefore, the peicwise linear approximation of objective function is

$$F_k(x, \psi) := c^\top x + \lambda \psi + \eta^k$$

and the master problem for MR-SLP with CVaR is defined as follows:

$$\begin{aligned}
& \underset{x, \psi, \eta^k}{\text{Min}} && c^\top x + \lambda \psi + \eta^k && (4.22) \\
& \text{s.t.} && x \in X \\
& && \eta^k - (\beta_t^k)^\top x - \gamma_t^k \psi \geq \alpha_t^k, \quad \forall t = 1, \dots, k.
\end{aligned}$$

4.3.2 SD Algorithm for MR-SLP with CVaR

Using the master problem and sub problems defined in section [4.3] we propose the following algorithm:

SD-CVaR

Step 0: Initialization.

Set $k \leftarrow 0$; $V^0 \leftarrow \emptyset$; $U^0 \leftarrow \emptyset$; $\eta^0 = -\infty$; $x^0 \in X$; $\bar{x}^0 \in X$; $\psi \in \mathbb{R}$ make suitable assumption (α quantile); $L \in \mathbb{R}$ given (lower bound) and $\delta \in (0, 1)$.

Step 1: Generate a Scenario.

$k \leftarrow k + 1$. Randomly generate scenario ω^k of $\tilde{\Omega}$, independent of previously generated scenarios.

Step 2: Determining Cut Approximation $\eta^k(x)$.

a. Solve dual problem (4.17) for scenario ω^k

$$\begin{aligned}
& (\pi^k(\omega^k), \phi^k(\omega^k)) \in \operatorname{argmax}\{(\pi^k)^\top (r(\omega^k) - T(\omega^k)x^k) + \phi^k(c^\top x - \psi) \\
& \quad |\pi^\top W - q\phi^k \leq q, \phi^k \leq \frac{\lambda}{1-\alpha}, \pi^k, \phi^k \geq 0\}
\end{aligned}$$

b. Update sets V^k and U^k

$$V^k \leftarrow V^{k-1} \cup \pi^k(\omega^k).$$

$$U^k \leftarrow U^{k-1} \cup \phi^k(\omega^k).$$

c. Determine the coefficients of the k^{th} cutting plane.

$$(\pi_t^k, \phi_t^k) \in \max \{(\pi)^\top (r(\omega^t) - T(\omega^t)x^k) + \phi(c^\top x^k - \psi) \mid \pi \in V^k, \phi \in U^k\}, \forall t = \{1, \dots, k-1\}.$$

$$\alpha_k^k = \frac{1}{k} \sum_{t=1}^k \{(\pi_t^k)^\top r(\omega^t)\}.$$

$$(\beta_k^k)^\top = \frac{1}{k} \sum_{t=1}^k \{(\pi_t^k)^\top (-T(\omega^t)) + \phi_t^k c^\top\}.$$

$$\gamma_k^k = \frac{1}{k} \sum_{t=1}^k -\phi_t^k.$$

d. Update coefficients of all previously generated cuts.

$$\alpha_t^k \leftarrow (k-1/k)\alpha_t^{k-1} + (1/k)L, \quad \forall t = 1, \dots, k-1.$$

$$\beta_t^k \leftarrow (k-1/k)\beta_t^{k-1}, \quad \forall t = 1, \dots, k-1.$$

$$\gamma_t^k \leftarrow (k-1/k)\gamma_t^{k-1}, \quad \forall t = 1, \dots, k-1.$$

e. Add Updated Cuts to Master Problem.

$$\eta^k \geq \alpha_t^k + (\beta_t^k)^\top x + \gamma_t^k \psi, \quad \forall t = 1, \dots, k.$$

$$\Rightarrow \eta^k - (\beta_t^k)^\top x - \gamma_t^k \psi \geq \alpha_t^k, \quad \forall t = 1, \dots, k.$$

Step 3.Updating Incumbent Solution.

$$\text{if } F_k(x^k) - F_k(\bar{x}^{k-1}) < \delta[F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1})]$$

$$\bar{x}^k \leftarrow x^k.$$

else

$$\bar{x}^k \leftarrow \bar{x}^{k-1}.$$

Step 4. Solve Master Problem.

$$\begin{aligned}
\text{Min} \quad & c^\top x + \lambda\psi + \eta^k \\
\text{s.t.} \quad & x \in X, \\
& \eta^k - (\beta_t^k)^\top x - \gamma_t^k \psi \geq \alpha_t^k, \quad \forall t = \{1, \dots, k\},
\end{aligned}$$

to get the new candidate solution x^{k+1} .

Step 5. Termination Criterion.

If the following criterion is not satisfied then starting from step 1 repeat all the steps.

Termination criteria based on objective value:

For a large enough k , terminate the algorithm if

$$F_{k-1}(\bar{x}^{k-1}) - F_{k-1}(x^k) \leq \epsilon,$$

where $\epsilon > 0$ is a given tolerance level.

REMARK 4.3.1. For termination of SD-CVaR, we can also employ the criterion 5(a) described in SD-EE algorithm.

4.4 Proof of Convergence

In this section we extend the convergence results of the risk-neutral SD algorithm described in [6] to the risk averse SD algorithms. For each risk-measure we first show that the mean-risk formulation satisfy assumption (A1)-(A5), even after the addition of dispersion variables. Then using the fact that MR-SLP has relatively complete recourse and set of first-stage decision variables is compact we show that the risk averse SD converges to an optimal solution.

4.4.1 Convergence Proof for SD-EE

For SD-EE, we prove that the approximation of the recourse function generated during the execution of the algorithm uniformly converges and the accumulation point of the candidate solutions is an optimal solution. We begin by proving that the variable $\nu(\omega)$ used for computing the excess over the target ψ for realization ω is always finite for $x \in X$ and therefore the assumption **A2** holds for two-stage SLP with EE.

COROLLARY 4.4.1. *Suppose that assumptions A1-A2 hold, then for any $x \in X$ and $\omega \in \Omega$ the dispersion statistic $\mathbb{E}[\nu(\tilde{\omega})] < \infty$.*

Proof. By assumption (A1), $X \subseteq \mathbb{R}_+^{n_1}$ is a compact set and from Equation (3.1) we have cost vector $c \in \mathbb{R}^{n_1}$. Therefore the elements of set $\{c^\top x^k\}_{k=1}^\infty$ will always be finite.

From assumption (A2), for any given $x \in X$ we have $\mathbb{E}[f(x, \tilde{\omega})] < \infty$. Hence from Equation (3.2), for any $\omega \in \Omega$, we have $q^\top y(\omega) < \infty$.

Since target $\psi \in \mathbb{R}$, and in Proposition 3.4.1 we define $\nu(\omega)$ as:

$$-c^\top x^k - q^\top y(\omega) + \nu(\omega) \geq \psi \quad \forall \omega \in \Omega. \quad (4.23)$$

Therefore, we have $\mathbb{E}[\nu(\tilde{\omega})] < \infty$. □

In the following statements for denoting limits we use \lim , for denoting upper limits and lower limits we use $\overline{\lim}$ and $\underline{\lim}$ respectively. The functions $h_k(x, \omega)$ and $h(x, \omega)$ are defined as follows:

$$h_k(x, \omega) = \max \{(\pi)^\top [r(\omega^k) - T(\omega^k)x^k] + \phi[c^\top x - \psi] \mid \pi \in V^k, \phi \in U^k\},$$

$h(x, \omega) = \max \{(\pi)^\top [r(\omega^k) - T(\omega^k)x^k] + \phi[c^\top x - \psi] \mid \pi \in V, \phi \in U\}$, where set V and U are collections of all dual vertices of the subproblem.

LEMMA 4.4.2. *Suppose that assumptions A1-A2 hold, then the sequence $\{h_k\}_{k=1}^\infty$ of functions $h_k(x, \omega)$, converges uniformly on $X \times \Omega$.*

Proof. Note that the set $V^k \subseteq V^{k+1} \subseteq V$ and set $U^k \subseteq U^{k+1} \subseteq U$, This implies that $h_k(x, \omega) \leq h_{k+1}(x, \omega) \leq h(x, \omega)$ for all k and for all $(x, \omega) \in X \times \Omega$. Since $\{h_k\}_{k=1}^\infty$ increases monotonically and is bounded from above by the function $h(x, \omega)$, it follows that $\{h_k\}_{k=1}^\infty$ converges pointwise to some function $g(x, \omega) \leq h(x, \omega)$. Since set $V^k \subseteq V^{k+1} \subseteq V$ and set $U^k \subseteq U^{k+1} \subseteq U$ for all k ,

$$\bar{V} = \lim_{k \rightarrow \infty} V_k \subseteq V, \quad (4.24)$$

and

$$\bar{U} = \lim_{k \rightarrow \infty} U_k \subseteq U. \quad (4.25)$$

By assumption (A2), elements of V and U are finite and so are \bar{V} and \bar{U} , hence

$$\begin{aligned} g(x, \omega) &= \lim_{k \rightarrow \infty} h_k(x, \omega) \\ &= \lim_{k \rightarrow \infty} \{ \text{Max} \quad \{ \pi^\top [r(\omega) - T(\omega)x] + \phi[c^\top x - \psi] \mid \pi \in V_k, \phi \in U_k \} \} \\ &= \text{Max} \quad \{ \pi^\top [r(\omega) - T(\omega)x] + \phi[c^\top x - \psi] \mid \pi \in \bar{V}, \phi \in \bar{U} \}. \end{aligned} \quad (4.26)$$

Therefore, from the statements (B.2), (B.3) and (B.4), we can conclude that $\{h_k\}_{k=1}^\infty$ converges uniformly to the function $g(x, \omega)$, since $\{h_k\}_{k=1}^\infty$ is a monotone sequence of continuous functions and $X \times \Omega$ is a compact set. \square

THEOREM 4.4.3. *Let $\{x^{k_n}\}_{n=1}^\infty$ be an infinite subsequence of $\{x^k\}_{k=1}^\infty$. Suppose that assumptions A1-A4 hold and if $x^{k_n} \rightarrow \hat{x}$, then with probability one*

$$\frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi) \rightarrow \mathbb{E}[h(\hat{x}, \tilde{\omega})].$$

Proof. From the equation (4.2) and step 2 of the algorithm, we know that

$$h_{k_n}(x^{k_n}, \omega^t) = \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi)$$

and

$$\frac{1}{k_n} \sum_{t=1}^{k_n} h_{k_n}(x^{k_n}, \omega^t) = \frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi).$$

By Lemma 4.4.2, there exists a function $g(x, \omega) \leq h(x, \omega)$ such that $\{h_{k_n}\}_{n=1}^\infty$ converges uniformly to $g(x, \omega)$. Thus, since we have

$$\frac{1}{k_n} \sum_{t=1}^{k_n} [h_{k_n}(x^{k_n}, \omega^t) - g(\hat{x}, \omega^t)] \rightarrow 0 \quad \text{and} \quad \frac{1}{k_n} \sum_{t=1}^{k_n} h(x, \omega^t) \rightarrow \mathbb{E}[h(x, \tilde{\omega})],$$

it is sufficient to show that $g(\hat{x}, \omega^t) = h(\hat{x}, \omega^t)$ with probability one. Since $h(x, \omega)$ is a continuous function and $\{h_{k_n}\}_{n=1}^\infty$ is a uniformly convergent sequence of continuous function, for every $\epsilon > 0$ there exist $\delta > 0$ and $N < \infty$ such that

$$|(\hat{x}, \omega^t) - (x, \omega)| < \delta \Rightarrow |h(\hat{x}, \omega^t) - h(x, \omega)| < \frac{\epsilon}{3} \quad \forall n \geq N$$

and

$$|h_{k_n}(\hat{x}, \omega^t) - h_{k_n}(x, \omega)| < \frac{\epsilon}{3} \quad \forall n \geq N.$$

Thus, since $x^{k_n} \rightarrow \hat{x}$, for every $\epsilon > 0$ there exist a further subsequence $\{(x^{k'_n}, \omega^{k'_n})\}_{n=1}^\infty$ and $K < \infty$ such that

$$|h(\hat{x}, \omega^t) - h(\hat{x}, \omega^{k'_n})| < \epsilon/3,$$

$$|h(\hat{x}, \omega^{k'_n}) - h(x^{k'_n}, \omega^{k'_n})| < \epsilon/3$$

and

$$|h_{k'_n}(x^{k'_n}, \omega^{k'_n}) - h_{k'_n}(x^{k'_n}, \omega^t)| < \epsilon/3,$$

for all $k'_n \geq K$. By construction we have $h_{k'_n}(x^{k'_n}, \omega^{k'_n}) = h(x^{k'_n}, \omega^{k'_n})$. Thus, for every $\epsilon > 0$ there exist a subsequence $\{x^{k'_n}\}_{n=1}^\infty$ and $K < \infty$ such that

$$\begin{aligned} |h(\hat{x}, \omega^t) - h_{k'_n}(x^{k'_n}, \omega^t)| &\leq |h(\hat{x}, \omega^t) - h(\hat{x}, \omega^{k'_n})| \\ &\quad + |h(\hat{x}, \omega^{k'_n}) - h(x^{k'_n}, \omega^{k'_n})| \\ &\quad + |h(x^{k'_n}, \omega^{k'_n}) - h_{k'_n}(x^{k'_n}, \omega^t)| < \epsilon, \end{aligned}$$

for all $k'_n \geq K$. Hence, by the uniqueness of the sequential limit, it follows that $g(\hat{x}, \omega^t) = h(\hat{x}, \omega^t)$. Therefore by probability one, we have

$$\frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi) \rightarrow \mathbb{E}[h(\hat{x}, \tilde{\omega})].$$

Also since $h(x, \omega^t) = \operatorname{argmax}\{\pi(r(\omega^t) - T(\omega^t)x) + \phi(c^\top x - \psi) | \pi \in V, \phi \in U\}$, $V^k \subseteq V$, and $U^k \subseteq U \forall k$, it follows that

$$\begin{aligned} c^\top x + \frac{1}{k_n} \sum_{t=1}^{k_n} h(x, \omega^t) &\geq c^\top x + \frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi) \\ &= c^\top x + \alpha_{k_n}^{k_n} + \beta_{k_n}^{k_n} x, \quad x \in X. \end{aligned}$$

□

THEOREM 4.4.4. *Suppose that assumptions A1-A4 hold, then there exists a subsequence $\{x^{k_n}\}_{n=1}^{\infty}$ of $\{x^k\}_{k=1}^{\infty}$, such that $\lim_{n \rightarrow \infty} [F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n})] = 0$, with probability one.*

Proof. If assumption (A2) is satisfied, then for every $\omega \in \Omega$ there exist $M(\omega) \in \mathbb{R}_+$, such that $|h(x^1, \omega) - h(x^2, \omega)| \leq M(\omega) \|x^1 - x^2\|$ for all $x^1, x^2 \in X$. Let $\epsilon > 0$ be given, let $M = \mathbb{E}[M(\omega)]$, let $r = \frac{\epsilon}{2M}$, and let $B_r(x)$ denote an open ball of radius r centered at x . Then $\cup_{x \in X} B_r(x)$ is an open cover of X . Since X is a compact set, there exist $N_\epsilon \leq \infty$ and $\{x_i\}_{i=1}^{N_\epsilon} \subset X$ such that $X \subset \cup_{i=1}^{N_\epsilon} B_r(x_i)$. Moreover, since $\{x^k\} \subset X$, it follows that each iterate is contained in one or more of the open balls $\{B_r(x_i)\}_{i=1}^{N_\epsilon}$. Thus, there exist two sequence of indices, $\{k_n\}$ and $\{t_n\}$ such that

$$0 < k_n - t_n \leq N_\epsilon + 1 \quad \text{and} \quad \|x^{k_n} - x^{t_n}\| < r.$$

By assumption (A1), we know that X is a compact set. Thus, without loss of generality we may assume that

$$\lim_{n \rightarrow \infty} x^{k_n} = \hat{x}_k \quad \text{and} \quad \lim_{n \rightarrow \infty} x^{t_n} = \hat{x}_t,$$

where \hat{x}_k and \hat{x}_t are accumulation points of sequences x^{k_n} and x^{t_n} respectively.

Now, in iteration k the cutting plane generated during iteration t appears as (please refer to step 2(d) of the SD-EE algorithm)

$$\alpha_t^k + \beta_t^k x = \frac{t}{k}(\alpha_t^t + \beta_t^t x) + \frac{k-t}{k} L. \quad (4.27)$$

As per the step 2 of algorithm we have

$$\eta_{k-1}(x^k) = \text{Max}\{\alpha_t^{k-1} + \beta_t^{k-1} x^k | t = 1, \dots, k-1\}. \quad (4.28)$$

Therefore, using equation (4.27) we can rewrite equation (4.28) as follows

$$\eta_{k-1}(x^k) \geq \frac{t}{k-1}(\alpha_t^t + \beta_t^t x^k) + (1 - \frac{t}{k-1})L \quad \forall t = 1, \dots, k-1.$$

Thus,

$$\begin{aligned} \eta_{k_n-1}(x^{k_n}) &\geq \frac{t_n}{k_n-1}(\alpha_{t_n}^{t_n} + \beta_{t_n}^{t_n} x^{k_n}) + (1 - \frac{t_n}{k_n-1})L \\ &= \frac{t_n}{k_n-1}(\alpha_{t_n}^{t_n} + \beta_{t_n}^{t_n} x^{t_n}) + \frac{t_n}{k_n-1}\beta_{t_n}^{t_n}(x^{k_n} - x^{t_n}) + (1 - \frac{t_n}{k_n-1})L \\ &= \frac{t_n}{k_n-1}\eta_{t_n}(x^{t_n}) + \frac{t_n}{k_n-1}\beta_{t_n}^{t_n}(x^{k_n} - x^{t_n}) + (1 - \frac{t_n}{k_n-1})L, \end{aligned} \quad (4.29)$$

where the last equality follows the fact that $\eta_k(x^k) = \alpha_k^k + \beta_k^k x^k$ for all k . Furthermore by definition,

$$\begin{aligned} F_k(x^k) - F_{k-1}(x^k) &= c^\top x^k + \eta_k(x^k) - c^\top x^k - \eta_{k-1}(x^k) \\ &= \eta_k(x^k) - \eta_{k-1}(x^k). \end{aligned} \quad (4.30)$$

Using equations (4.29) and (4.30) we have

$$F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \leq \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n-1}(\eta_{t_n}(x^{t_n}) + \beta_{t_n}^{t_n}(x^{k_n} - x^{t_n})) - (1 - \frac{t_n}{k_n-1})L.$$

By construction,

$$0 < k_n - t_n \leq N_\epsilon + 1 < \infty,$$

we have

$$\begin{aligned}\|\beta_{t_n}^{t_n}\| &\leq \frac{1}{t_n} \sum_{t=1}^{t_n} M(\omega^t), \\ \lim_{n \rightarrow \infty} \frac{t_n}{k_n - 1} &= 1 \quad \text{and} \\ \lim_{n \rightarrow \infty} \left(1 - \frac{t_n}{k_n - 1}\right)L &= 0.\end{aligned}$$

Moreover, Theorem 4.4.3 ensures that $\eta_{k_n}(x^{k_n}) \rightarrow \mathbb{E}[h(\hat{x}_k, \tilde{\omega})]$ and $\eta_{t_n}(x^{t_n}) \rightarrow \mathbb{E}[h(\hat{x}_t, \tilde{\omega})]$ with probability one. It follows that

$$\begin{aligned}0 &\leq \underline{\lim}_{k \rightarrow \infty} F_k(x^k) - F_{k-1}(x^k) \\ &\leq \underline{\lim}_{n \rightarrow \infty} F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \\ &\leq \underline{\lim}_{n \rightarrow \infty} \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n - 1} (\eta_{t_n}(x^{t_n}) + \|\beta_{t_n}^{t_n}\| \|x^{k_n} - x^{t_n}\|) - \left(1 - \frac{t_n}{k_n - 1}\right)L. \\ &\leq \underline{\lim}_{n \rightarrow \infty} \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n - 1} (\eta_{t_n}(x^{t_n}) + \frac{1}{t_n} \sum_{t=1}^{t_n} M(\omega^t) \|x^{k_n} - x^{t_n}\|) - \left(1 - \frac{t_n}{k_n - 1}\right)L. \\ &= \mathbb{E}[h(\hat{x}_k, \tilde{\omega})] - \mathbb{E}[h(\hat{x}_t, \tilde{\omega})] + M \|\hat{x}_k - \hat{x}_t\| \\ &\leq |\mathbb{E}[h(\hat{x}_k, \tilde{\omega})] - \mathbb{E}[h(\hat{x}_t, \tilde{\omega})]| + M \|\hat{x}_k - \hat{x}_t\| \\ &\leq 2M \|\hat{x}_k - \hat{x}_t\| \\ &\leq 2M \left(\frac{\epsilon}{2M}\right) \\ &= \epsilon.\end{aligned}$$

Thus, for every $\epsilon > 0$, $0 \leq \lim_{n \rightarrow \infty} F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \leq \epsilon$ and hence the result. □

THEOREM 4.4.5. *There exist a subsequence $\{x^{k_n}\}_{n=1}^{\infty}$ of $\{x^k\}_{k=1}^{\infty}$, such that every accumulation point of $\{x^{k_n}\}_{n=1}^{\infty}$ is an optimal solution x^* with probability one.*

Proof. From Theorem 4.4.3, we know that there exist subsequence $\{x^{k_n}\}_{n=1}^{\infty}$ such that $\lim_{n \rightarrow \infty} h_{k_n}(x^{k_n}) -$

$h_{k_{n-1}}(x^{k_n}) = 0$. Let $\{x^{k_n}\}_{n \in N}$ be a subsequence such that $\lim_{n \in N} x^{k_n} = \hat{x}$. By assumption (A1) we will always have accumulation point $\hat{x} \in \mathcal{X}$ thus for an optimal solution x^* we have

$$F(x^*) \leq F(\hat{x}), \quad (4.31)$$

where $F(x) = c^\top x + \mathbb{E}[h(x, \omega)]$, and also by construction we have

$$\lim_{k \in K} F_k(x^*) \leq c^\top x^* + \mathbb{E}[h(x^*, \omega)] = F(x^*). \quad (4.32)$$

As per the step 4 of algorithm we know that x^k minimizes F_{k-1} , therefore

$$F_{k-1}(x^k) \leq F_{k-1}(x^*). \quad (4.33)$$

Now using Theorem 4.4.3 and result $\lim_{n \in N} F_{k_n}(x^{k_n}) = F(\hat{x})$ we get $\lim_{n \in N} F_{k_{n-1}}(x^{k_n}) = F(\hat{x})$, with probability one. Combining equations (4.31), (4.32) and (4.33) we have

$$F(x^*) \leq F(\hat{x}) = \lim_{n \in N} F_{k_{n-1}}(x^{k_n}) \leq \lim_{k \in K} F_k(x^*) \leq F(x^*).$$

Hence accumulation point of subsequence $\{x^{k_n}\}_{n=1}^\infty$ is an optimal solution x^* with probability one. □

We have now proved that the SD algorithm will generate an optimal solution with probability one. Next we need to show that with probability one, there exists at least one optimal accumulation point of the sequence of incumbents $\{\bar{x}^k\}$. We will use Theorems 4.4.3 and 4.4.4 along with the incumbent test used in step 3 of algorithm to prove this result. We begin by noting that for all k ,

$$c^\top \bar{x}^k + \alpha_{i_k}^k + \beta_{i_k}^k x \leq F_k(\bar{x}^k) \leq c^\top \bar{x}^k + \frac{1}{k} \sum_{t=1}^k h(\bar{x}^k, \omega^t).$$

Since $h(x, \omega)$ is continuous in x for all $\omega \in \Omega$, if $\{\bar{x}^{k_n}\}_{n=1}^\infty$ is a subsequence such that $\bar{x}^{k_n} \rightarrow \bar{x}$,

$h(\bar{x}^{k_n}, \omega^t) \rightarrow h(\bar{x}, \omega^t)$ for all t . Thus,

$$\begin{aligned} \underline{\lim}_{n \rightarrow \infty} c^\top \bar{x}^{k_n} + \alpha_{i_k}^{k_n} + \beta_{i_k}^{k_n} x &\leq \underline{\lim}_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) \\ &\leq \overline{\lim}_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) \\ &\leq \overline{\lim}_{n \rightarrow \infty} c^\top \bar{x}^{k_n} + \lambda\psi + \frac{1}{k_n} \sum_{t=1}^{k_n} h(\bar{x}^{k_n}, \omega^t). \end{aligned}$$

With probability one, both the upper and lower limits described above are $F(\bar{x})$ (please see Theorem 4.4.3). Thus, it follows that

$$\lim_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) = F(\bar{x}).$$

Similarly, if $\{\bar{x}^{k_n}\}_{n=1}^\infty \rightarrow \bar{x}$, then the nature of the update mechanism described in step 2 of algorithm and Theorem 4.4.4 ensures that

$$\lim_{n \rightarrow \infty} F_{k_n+1}(\bar{x}^{k_n}) = F(\bar{x}).$$

The above results can be summarized and formerly stated in the following corollary.

COROLLARY 4.4.6. *Let $\{\bar{x}^k\}$ denote the sequence of incumbent solutions, and let $\{\bar{x}^{k_n}\}_{n=1}^\infty$ be an infinite subsequence such that $\{\bar{x}^{k_n}\} \rightarrow \bar{x}$. If the assumptions (A1)-(A4) hold, then with probability one*

$$\lim_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) = \lim_{n \rightarrow \infty} F_{k_n+1}(\bar{x}^{k_n}) = F(\bar{x}).$$

To establish that an optimal accumulation point of the incumbent sequence exists, next we explore the implication of the incumbent test described in step 3 of the algorithm.

LEMMA 4.4.7. *Suppose that assumption A1-A4 hold. Let $\theta^k = F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1})$, and let $\{k_n\}_{n \in \mathbb{N}}$ represent the sequence of iterations at which the incumbent is changed. If N is finite,*

then $\overline{\lim}_{k \rightarrow \infty} \theta^k = 0$, with probability one. Otherwise, $\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} = 0$ with probability one.

Proof. By definition, $\theta^k = F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1}) \leq 0$ for all k . If N is a finite set, there exist \bar{x} and $K < \infty$ such that $\bar{x}^k = \bar{x}$ for all $k \geq K$ and thus

$$F_k(x^k) - F_k(\bar{x}) \geq \delta[F_{k-1}(x^k) - F_{k-1}(\bar{x})] = \delta\theta \quad \forall k \geq K.$$

By Theorem 4.4.4 and Corollary 4.4.6, there exist a subsequence indexed by set \mathcal{K} such that

$$\begin{aligned} \lim_{k \in \mathcal{K}} x^k &= \hat{x} \\ \lim_{k \in \mathcal{K}} F_k(x^k) &= f(\hat{x}), \quad \lim_{k \in \mathcal{K}} F_k(\bar{x}) = f(\bar{x}), \\ \lim_{k \in \mathcal{K}} F_{k-1}(x^k) &= f(\hat{x}), \quad \lim_{k \in \mathcal{K}} F_{k-1}(\bar{x}) = f(\bar{x}), \end{aligned}$$

with probability one. Thus,

$$\begin{aligned} \lim_{k \in \mathcal{K}} \{F_k(x^k) - F_k(\bar{x})\} &\geq \delta[\lim_{k \in \mathcal{K}} \{F_{k-1}(x^k) - F_{k-1}(\bar{x})\}] \\ &\Rightarrow F(\hat{x}) - F(\bar{x}) \geq \delta[F(\hat{x}) - F(\bar{x})] \\ &= \lim_{k \in \mathcal{K}} \delta\theta^k, \end{aligned}$$

with probability one. Now since $\delta \in (0, 1)$ and $\theta^k \leq 0$ for all k , it follows that $F(\hat{x}) - F(\bar{x}) = 0$ and thus $\lim_{k \in \mathcal{K}} \delta\theta^k \leq 0$, with probability one. Now suppose N is not a finite set. By hypothesis,

$$F_{k_n}(x^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) < \delta[F_{k_n-1}(x^{k_n}) - F_{k_n-1}(\bar{x}^{k_n-1})] = \delta\theta^{k_n} \leq 0 \quad \forall n.$$

By definition of the subsequence $\{k_n\}$, we note that $\bar{x}^{k_n-1} = \bar{x}^{k_{n-1}}$. Therefore

$$F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) \leq \delta \theta^{k_n} \leq 0 \quad \forall n.$$

Thus,

$$\begin{aligned} \frac{1}{m} \sum_{n=1}^m \{F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1})\} &\leq \frac{\delta}{m} \sum_{n=1}^m \theta^{k_n} \leq 0 \quad \forall m \\ \Rightarrow \frac{1}{m} \left\{ \left(\sum_{n=1}^{m-1} F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) \right) + (F_{k_m}(\bar{x}^{k_m}) - F_{k_1}(\bar{x}^{k_0})) \right\} \\ &\leq \frac{\delta}{m} \sum_{n=1}^m \theta^{k_n} \leq 0 \quad \forall m. \end{aligned}$$

Assumptions (A1)-(A4) ensure that there exist $M < \infty$, such that

$$|F_{k_m}(\bar{x}^{k_m}) - F_{k_1}(\bar{x}^{k_0})| < M \quad \forall m.$$

Thus, since $\bar{x}^{k_n} = \bar{x}^{k_{n+1}-1}$, the left hand side converges to zero with probability one as m approaches ∞ . Thus,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} = 0.$$

□

THEOREM 4.4.8. *Suppose that assumptions A1-A4 are satisfied. Let $\{\bar{x}^k\}_{k=1}^{\infty}$ represent the sequence of incumbents and let X^* represent set of optimal solutions. Then there exist a subsequence $\{\bar{x}^k\}_{k \in K}$ for which every accumulation point is contained in X^* , with probability one.*

Proof. Let $\{k_n\}_{n \in \mathbb{N}}$ represent the sequence of iterations at which the incumbent is changed. Note

that if N is infinite set,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} \leq \overline{\lim}_{n \rightarrow \infty} \theta^{k_n} \leq 0.$$

Thus, as a result of lemma 4.4.7, whether N is finite or infinite, there exist a subsequence indexed by set \mathcal{K} such that

$$\lim_{k \in \mathcal{K}} \theta^{k+1} = 0.$$

Note that

$$\theta^{k+1} = F_k(x^{k+1}) - F_k(\bar{x}^k) \leq F_k(x^*) - F_k(\bar{x}^k) \quad \forall k \in \mathcal{K}.$$

Thus as a result of Corollary 4.4.6, it follows that if \bar{x} is an accumulation point of $\{\bar{x}^k\}_{k \in \mathcal{K}}$, then

$$\begin{aligned} F(\bar{x}) &\leq \overline{\lim}_{k \in \mathcal{K}} F_k(x^*) \\ &\leq c^\top x^* + \lim_{k \in \mathcal{K}} \frac{1}{k} \sum_{t=1}^k h(x^*, \omega^t) \\ &\leq F(x^*), \end{aligned}$$

and thus, $\bar{x} \in X^*$, with probability one. □

The convergence proofs for risk measure QDEV and CVaR follow similar pattern as that of EE, so to avoid distraction from the foregoing discussion we present the detailed proofs for SD-QDEV and SD-CVaR in Appendix B.

5. COMPUTATIONAL RESULTS

In this section we present the details of our computational study. The study is based on standard test instances and it was performed to gain understanding of the empirical behavior of the SD algorithms. Specifically, the study was carefully designed to achieve the following goals:

- Validate the convergence of SD algorithms to an optimal solution computationally.
- Compare the performance of SD algorithm with SAA approach.
- Study how the optimal solution for a given test instance varies with the risk measure used.
- Assess the impact of risk measures on the cost associated with the in-sample ‘extreme’ scenarios.

The SD and SAA algorithms were implemented in C++ using the IBM CPLEX Callable Library version 12.8 [45] in the Microsoft Visual Studio 2017 environment. We used the object-oriented approach for coding the algorithms by creating our own *classes* and *methods* to interact with CPLEX Callable Library’s method. We created five major classes, namely *LObjectClass*, *Reader*, *Master*, *Sublp* and *Algorithm*. The class *LObjectClass* is a superclass and its properties are inherited by rest of the classes with exception of *Algorithm* class. The *Reader* class reads the SLP data from test instances stored in SMPS (Stochastic mathematical Programming Society) Format. The master problem and subproblem aspects are dealt by *Master* and *Sublp* class respectively. Finally, the execution of algorithm is implemented in *Algorithm* class. All the experiments were conducted on a computer workstation running Intel Xeon 2.40GHz with dual processors and 12GB RAM.

5.1 Test Instances

We use standard test instances from the SP literature, which are described in [27] and [46] and are known as *pgp2* [27], *pgp2e* [25], *gbd* [47], *LandS* [48], and *storm* [49]. The characteristics of these test instances are summarized in Table 5.1. The columns of the table are instance name, application, number of scenarios, number of first-stage constraints and variables, number of

second-stage constraints and variables, and the optimal objective value [25]. For instances with a very large number of scenarios the optimal value cannot be obtained, and a 95% confidence interval of optimal objective value is reported [46].

Table 5.1: Standard Test Problems

Name	Application	Scenarios	First-Stage (Cons., Vars.)	Second-Stage (Cons., Vars.)	Optimal Objective Value
pgp2	Power Generation Planning	576	(2, 4)	(7, 12)	447.32
pgp2e	Power Generation Planning	576	(2, 4)	(7, 12)	413.94
gbd	Aircraft Allocation	6.5×10^5	(4, 17)	(5, 10)	(1648.76, 1657.50)
LandS	Electricity Planning	10^6	(2, 4)	(7, 12)	(225.20, 226.24)
storm	Cargo Flight Scheduling	6×10^{81}	(185, 121)	(528, 1259)	(15496242.30, 15500000.30)

Test instance *pgp2* deals with electrical capacity expansion problem. It is a model so as to select optimal cost strategy for investing in different sources of electricity such as gas-fired, coal-fired, and nuclear generators. The stochastic elements in this problem are the power generation costs and regional demands. The first-stage variables model the capital cost (\$/Kw) per year based on type of acquired generator. The second-stage decision variables determine how much each generator should be used to satisfy the regional power demand. Instance *pgp2e* is a modified version of *pgp2*. All the random variables associated with *pgp2* have a marginal uniform distribution, which is not suitable for studying the empirical behavior of mean-risk measures. Hence, the instance *pgp2e* was generated by skewing the marginal distributions of the random variables thus making the instance suitable for mean-risk measures. The details of the modification process are given in [25].

Test instance *gbd* is an aircraft allocation problem. The objective is to maximize profit by allocating four different types of aircraft to five different routes. The cost involved are related to the operation of an aircraft and the bumping of passengers. Randomness comes from the uncertain demand associated with each route. The problem *LandS* deals with electrical investment planning. In the first-stage a decision is made regarding the capacities of four new technologies. In the second-stage, decisions concern the amount of electricity produced through the four technologies

working under three different modes. Randomness represents the demand of electricity for each mode. Finally, the instance *storm* deals with planning the allocation of aircraft routes, and was used by the U.S. military during the Gulf War of 1991. Randomness arises from the uncertain demand of cargo delivery. The first-stage of this instance deals with scheduling of flight routes so as to minimize the cost of scheduled flights and cargo handling costs. The second-stage deals with satisfying the cargo delivery demand and minimizing the penalty associated with the unmet demand.

5.2 Results

To draw insights from the results regarding the risk-neutral versus risk-averse cases, we present plots of the optimal value versus the risk trade-off parameter λ for each of the three risk measures. For every test instance, we ran the algorithm for $\lambda \in \{0, 0.1, 0.2, \dots, 1\}$ to trace the efficient frontier of each problem. For QDEV and CVaR risk measures, we set the parameter $\alpha = 0.95$ and for EE we set the target ψ closer to the α -quantile values of QDEV and CVaR. This was done to provide a basis for comparison among the three risk measures. To evaluate the performance of the algorithm for each risk measure, we recorded computational time (CPU), number of iterations and objective value at termination. We should point out that, for the SD algorithm, the number of samples is equal to the number of iterations at termination.

Since the SD algorithm involves sampling, we replicate every standard instance to get a better statistical estimation of the optimal value. For each instance, we perform 30 replications for every combination of the risk measure and the risk trade-off parameter λ . For each of the 30 replications, in order to compute the value of statistics which are not directly available at the termination of algorithm, we fix the optimal solutions and generate 5000 independent samples, uncorrelated to any previously generated ones. In the end, by using results from each of the 30 replications, we record following statistics for our analysis: the objective value at termination, the expected cost, the expected cost of the risk deviation statistic, the upper bound value, the CPU time in seconds and number of iterations.

Next, we report and discuss the computational results. We place most of the tables in the

Appendix A to provide continuity to the foregoing discourse. We report the computational results tables in the following test instance order: *pgp2*, *pgp2e*, *gbd*, *lands* and *storm*. The names of the columns of each table are as follows: λ (risk trade-off parameter), ‘Obj Val’ (optimal objective value), $\mathbb{E}[f(x, \tilde{\omega})]$ (expected cost), $\mathbb{E}[\nu(\tilde{\omega})]$ (expected cost of the deviation statistic), ‘UB’ (upper bound value), ‘CPU’ (computational time taken in seconds), and ‘Iteration’ (number of iterations, which is equal to the number of samples taken). For each λ value in the table, the first row records the average value of a statistic mentioned in the column for all 30 replications of the instance for the given risk measure and the second row records the standard deviation.

5.2.1 Results for SD-EE

The results for SD-EE and SAA-EE are reported in Table A.1 and Table A.2 for *pgp2*, Table A.3 and Table A.4 for *pgp2e*, Table A.5 and Table A.6 for *gbd*, Table A.7 and Table A.8 for *lands*, and Table A.9 and Table A.10 for *storm*, respectively. All the tables are included in Section A.1 of Appendix A. The first and second rows of each table ($\lambda = 0$) show the results for the risk-neutral case. Notice that for all the test instances the optimal objective values of SD-EE correspond very closely to the values from the literature listed in Table 5.1 and to the values reported for SAA-EE.

The most important feature of the SD-EE algorithm is that it takes a relatively small fraction of the total number of scenarios to converge to an optimal solution, which is particularly salient in the last three instances of Table 5.1. Unlike SAA, we do not need to specify the number of scenarios to sample in SD-EE: we just set the tolerance level for the termination criteria and the algorithm only samples the required number of scenarios to satisfy the criteria. This feature guarantees a desired level of accuracy in solution, and at the same time offers a competitive computational time, which is an advantage over the SAA approach.

In the tables, we can see that the objective values increase with the value of λ . This is expected and is an indication of the increasing cost of being more risk-averse. The change in the value of expected cost for different values of λ is an indication in the change of the first-stage decision to adapt to the risk aversion level. For example, we see that for *pgp2* the value remains almost constant for most λ values, an indication that the risk-neutral approach is sufficient, whereas for

pgp2e that is not the case.

The target value ψ for each instance was carefully set through experimentation. By making several runs for different target values, we observed that for higher ψ values (which correspond to fewer scenarios exceeding the target), the risk measure EE becomes effective only for a relatively small number of scenarios and has little to no-effect on the solution for smaller values. Finally, we should point out that another attractive feature of the SD-EE algorithm is that the computation time required for convergence is relatively small since the algorithm typically requires few scenarios to reach termination.

5.2.2 Results for SD-QDEV

The results of SD-QDEV algorithm for instances *pgp2*, *pgp2e*, *gbd*, *lands* and *storm* are summarized in Appendix A in Tables A.11, A.13, A.15, A.17 and A.19, respectively and the results of SAA-QDEV are summarized in Tables A.12, A.14, A.16, A.18 and A.19. As mentioned earlier, for this algorithm we set the α value to 0.95, which results in $\epsilon_1 = 1$ and $\epsilon_2 = \alpha\epsilon_1/(1 - \alpha) = 19$. As in the results for SD-EE, for all test instances the risk-neutral optimal values are very similar to the optimal objective values from the literature listed in Table 5.1 and to the values reported for SAA-QDEV. For instances with a large number of scenarios, the SD-QDEV algorithm uses only a small fraction before convergence is achieved.

We can see from the tables that the objective values increase with λ . This is more pronounced than in the SD-EE case. We also see that with SD-QDEV, the expected cost $\mathbb{E}[f(x, \tilde{\omega})]$ generally increases with the risk trade-off factor, an indication that QDEV is more adept to changes in the level of risk-averseness than EE. This is especially true for *gbd*, where there is a 30% increase in the expected cost. The reader should recall that with QDEV, we penalize deviations from the α -quantile both above and below, and this affords this risk measure more ‘flexibility’. We also see that the SD-QDEV algorithm takes relatively small computation time to terminate, which is comparable to that of SD-EE in general.

5.2.3 Results for SD-CVaR

The results for the SD-CVaR and SAA-CVaR algorithms are reported in Appendix A in Tables A.21-A.29 and Tables A.22-A.30 for instances *pgp2*, *pgp2e*, *gbd*, *lands* and *storm*, respectively. Recall that, as with SD-QDEV, we set the α value to 0.95. As with the other two risk measures, the risk-neutral optimal values ($\lambda = 0$) are similar to the optimal objective values reported for the SAA-CVaR. It can also be observed that for all instances the objective value increases with λ .

Unlike the other quantile risk measures QDEV, with CVaR we can see that the expected cost $\mathbb{E}[f(x, \tilde{\omega})]$ does not change by much as λ increases. For the CVaR it seems that more scenarios are needed in order to be above the VaR or the α value needs to be set to a lower value than 0.95, to have any measurable impact on the optimal first-stage solution. Also, CVaR has more impact on instances having skewed marginal distributions compared to the instances with uniform marginal distribution. This is true for instance *pgp2e* which was generated by skewing the marginal distribution of instance *pgp2*. For *pgp2e*, we see a noticeable increase in the expected cost (about 2%) from the risk-neutral case to the risk-averse case whereas for *pgp2* the increase in the expected cost is less pronounced ($< 0.5\%$).

5.3 Impact of Risk Measure on Optimal Solution

To go deeper in the computational findings, in Table 5.2 we report the first-stage optimal solution for one replication of instance *pgp2e* for each combination of the risk measure and different values of trade-off factor. Recall from Section 5.1 that instance *pgp2e* deals with electrical capacity expansion. The first-stage decisions in this instance correspond to electricity generated from gas fired power plants x_1 , coal fired power plants x_2 , nuclear power plants x_3 and other generators x_4 . For EE, we set the target $\psi = 420$. For QDEV and CVaR we set the parameter $\alpha = 0.95$.

In Table 5.3, we report the first-stage cost $c^\top x$ and the expected cost $\mathbb{E}[f(x, \tilde{\omega})]$ of instance *pgp2e* based on optimal solutions reported in Table 5.2. For each λ we fix the optimal solution obtained in Table 5.2 and generate a random sample of size 5000. The expected cost $\mathbb{E}[f(x, \tilde{\omega})]$ is estimated by averaging out the values of $f(x, \tilde{\omega})$ obtained.

Table 5.2: Optimal Solution for Instance *pgp2e*

λ	EE				QDEV				CVaR			
	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
0.0	3.603	4.832	2.921	10.207	3.603	4.832	2.921	10.207	3.603	4.832	2.921	10.207
0.1	3.472	7.342	2.020	8.415	4.812	3.184	2.849	11.663	3.939	4.785	2.914	9.411
0.2	0.000	10.260	3.207	7.077	2.769	6.545	3.623	9.661	0.000	9.110	5.307	6.920
0.3	5.027	3.588	0.000	11.898	5.493	8.144	3.233	5.872	0.000	5.949	4.412	11.383
0.4	2.184	6.676	4.791	6.333	2.435	6.092	4.664	9.211	0.000	8.160	5.445	8.823
0.5	3.971	2.640	4.282	9.617	1.172	6.751	7.129	7.449	1.708	9.384	1.455	9.596
0.6	0.887	5.988	4.357	10.299	0.000	7.858	6.744	8.050	5.242	5.248	3.224	8.434
0.7	3.093	4.620	3.708	9.401	1.358	6.394	7.317	7.432	1.664	6.080	5.568	8.671
0.8	0.000	6.114	5.927	9.351	1.221	6.495	7.369	7.402	1.479	8.488	6.438	5.653
0.9	0.442	7.288	4.275	8.807	2.262	6.891	6.898	6.461	4.174	6.450	5.445	5.783
1.0	1.087	6.372	3.322	9.697	2.872	5.403	6.803	7.436	0.000	10.094	7.444	4.436

Table 5.3: Optimal Values for Instance *pgp2e*

λ	EE		QDEV		CVaR	
	$c^\top x$	$\mathbb{E}[f(x, \tilde{\omega})]$	$c^\top x$	$\mathbb{E}[f(x, \tilde{\omega})]$	$c^\top x$	$\mathbb{E}[f(x, \tilde{\omega})]$
0	177.827	417.567	177.827	417.567	177.827	417.567
0.1	168.919	418.291	185.965	423.512	175.964	418.763
0.2	165.601	418.879	189.44	423.213	190.207	420.870
0.3	146.774	422.894	198.887	430.373	180.527	420.340
0.4	183.223	421.389	196.881	423.582	197.168	426.345
0.5	184.398	420.411	217.731	431.749	163.630	423.942
0.6	182.290	418.821	211.210	429.390	191.346	425.318
0.7	179.003	418.443	220.000	432.918	200.314	424.233
0.8	193.740	421.024	220.000	432.876	211.134	429.333
0.9	176.674	418.002	220.000	433.74	208.717	429.534
1	166.808	417.928	220.000	433.504	216.381	432.258

We can see from Table 5.2 that for all the three risk measures, the power generated from gas fired power plant x_1 and other generators x_4 decreases with increasing λ (as we become more risk-averse), while the power generated from nuclear power plant x_3 increases. This trend is moderate for risk measure EE, but significant in the case of CVaR. Also, as the value of λ increases, the power generated from coal fired power plant x_2 increases slightly for risk measure QDEV and EE, and more than doubles for CVaR. From Table 5.3, it can be inferred that the impact of risk measure QDEV and CVaR on the first-stage cost and the expected cost for $pgp2e$ is more pronounced compared to the impact of EE, since in the cases of QDEV and CVaR there is an increase of 25% in first-stage cost as the value of λ increases from 0 to 1. These results demonstrate the different nature of these mean-risk measures when applied to the same test instance.

5.4 Expected Cost and Risk Trade-Off Factor

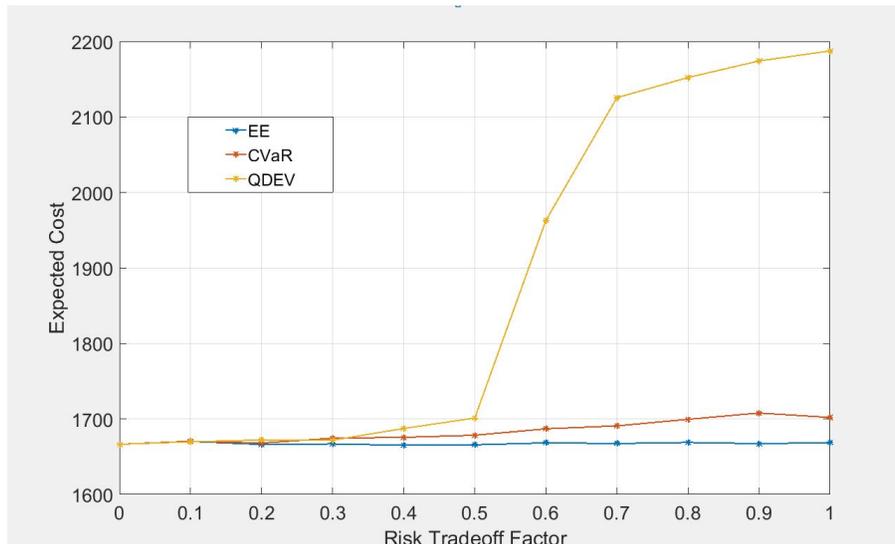


Figure 5.1: Expected Cost $\mathbb{E}[f(x, \tilde{\omega})]$

Understanding the effect of a risk measure and the risk trade-off factor λ on the expected cost and on the “extreme” scenarios is important for risk management purposes. This also helps in deciding when it is appropriate to use a given risk measure for a given problem. It is also important

to know the effect of the λ and how one can select which level to use in practice. In Figure 5.1 we plot the expected cost versus λ instance *gbd* for all three risk measures. We selected instance *gbd* because the marginal distributions of the random variables in this instance are not uniform nor normally distributed and therefore, the effect of a risk measure on the expected cost is significant.

We can observe from Figure 5.1 that as the value of λ increases, the expected cost for EE stays almost constant, it increases slightly for CVaR and has a major increase for QDEV, in particular after $\lambda = 0.5$. In Figure 5.2 we plot the expected cost for different λ values for a selected percentage of worst-case ('extreme') scenarios. We see from the figure that the cost associated with the worst-case scenarios for risk measure QDEV decreases as we become more and more risk-averse. The expected cost of the worst-case scenarios drops by almost 10% as the value of λ increases from 0 to 1. This effect is less evident in the case of CVaR (2.5%) and is almost insignificant in case of EE (< 1%). This indicates that for a given value of risk trade-off factor, the risk measure QDEV has more impact on expected cost and the cost associated with 'extreme' scenarios compared to EE and CVaR.

5.5 Mean-Risk SD versus Mean-Risk SAA

To understand the impact of risk aversion on performance of different sampling approaches, we compared the performance of SD and SAA approach under risk-neutral and risk-averse cases. We implement the SAA approach described in [50] for each of the three risk measures EE, QDEV and CVaR. For each instance, we collect a sample of one thousand random scenarios using Monte Carlo simulation to implement SAA and for each instance we perform thirty replications of SAA.

Table 5.4 summarizes the average computation time in seconds for SD and SAA approach with risk measures EE, QDEV and CVaR for test instances *pgp2*, *pgp2e*, *gbd*, *LandS* and *storm*. For each instance the row $\lambda = 0$ lists the average computation time for risk-neutral case, whereas row $\lambda > 0$ lists the average computation time under risk aversion for each risk measure..

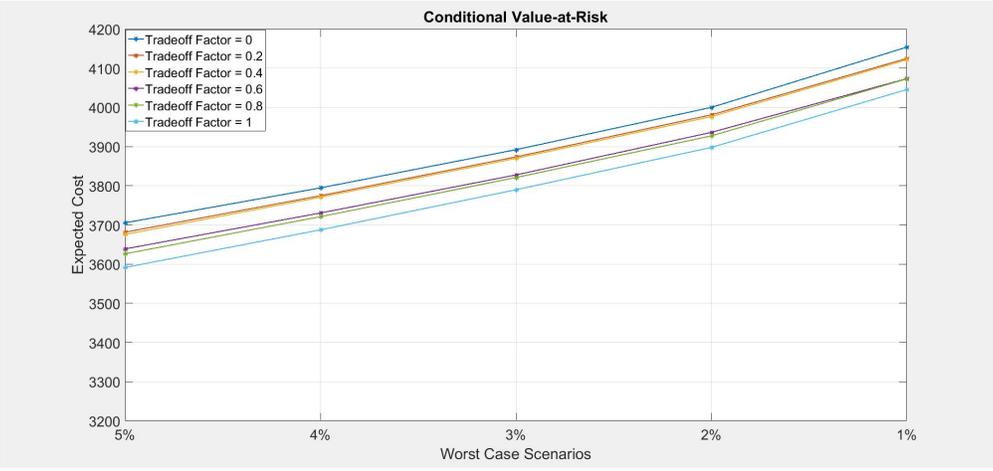
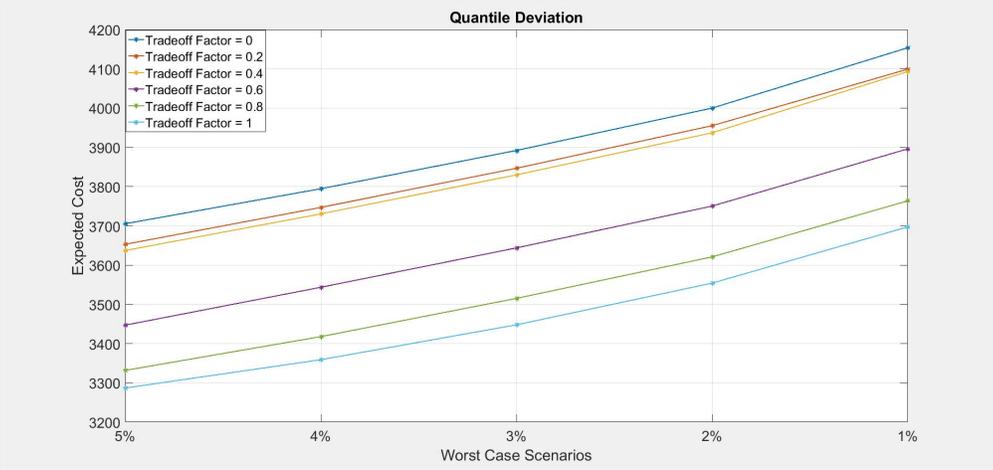
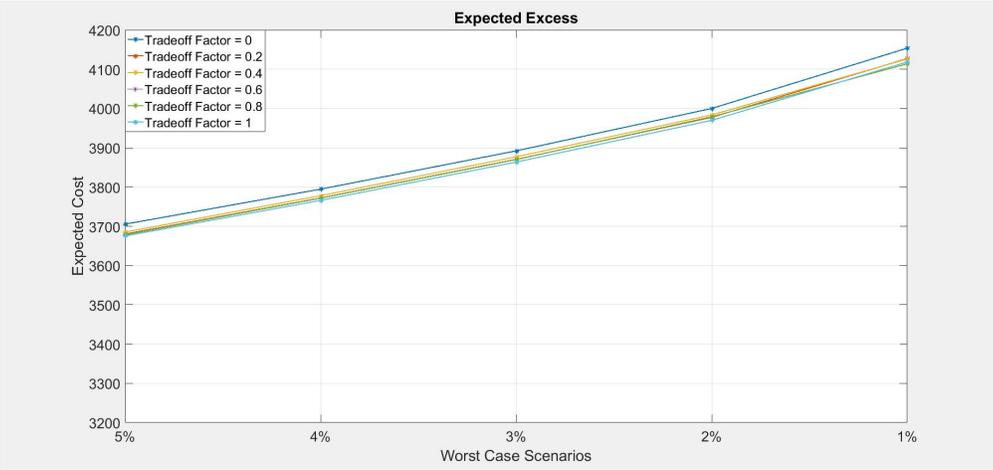


Figure 5.2: Cost of Worst Case Scenarios

Table 5.4: Average CPU Time (Sec)

	λ	SD-EE	SAA-EE	SD-QDEV	SAA-QDEV	SD-CVaR	SAA-CVaR
pgp2	0	4.89	4.21	3.88	4.47	4.35	4.09
	> 0	4.74	4.20	2.90	5.80	3.64	5.92
pgp2e	0	3.79	5.44	4.84	5.60	4.30	5.43
	> 0	4.49	5.46	2.80	7.03	3.13	8.37
gbd	0	3.43	3.83	3.42	4.13	4.10	4.15
	> 0	4.10	4.32	3.44	7.53	3.62	8.41
LandS	0	2.96	4.43	3.14	4.45	3.50	4.30
	> 0	3.67	4.32	2.39	6.54	2.83	6.81
storm	0	129.92	62.63	207.32	87.01	191.40	69.01
	> 0	161.00	63.78	170.54	101.86	225.41	98.94

For quantile risk measures QDEV and CVaR, the computation time of SAA approach increases under risk aversion. Specifically for instances with skewed distribution of second stage random variables such as *pgp2e* and *gbd*, the computation time under risk aversion with CVaR and QDEV increases by around 75% and 50% respectively. In contrast to SAA, the performance of SD algorithm remains consistent under risk aversion.

Table 5.5 summarizes the average number of iterations required for the SD and SAA approach to converge with risk measures EE, QDEV and CVaR for each of the five test instances. One should note that in SD algorithm at each iteration only one LP is solved, whereas in case of SAA approach LPs associated with all of the one thousand scenarios from the sample are solved. The average number of iterations for SAA approach increases under risk aversion with quantile risk measures QDEV and CVaR.

Table 5.5: Number of Iterations

	λ	SD-EE	SAA-EE	SD-QDEV	SAA-QDEV	SD-CVaR	SAA-CVaR
pgp2	0	790.17	31.43	758.70	29.67	742.00	30.40
	> 0	778.43	31.09	635.81	42.45	705.19	43.91
pgp2e	0	744.30	39.90	796.03	40.40	781.97	38.83
	> 0	774.60	39.61	629.94	49.57	671.70	58.32
gbd	0	677.13	30.40	661.60	30.13	710.27	30.57
	> 0	713.42	33.78	719.00	51.33	648.02	55.10
LandS	0	670.60	31.63	709.17	30.83	726.97	31.13
	> 0	725.95	31.11	607.67	45.88	657.64	48.62
storm	0	702.20	51.17	759.83	70.39	739.53	52.13
	> 0	714.16	53.23	630.77	79.06	796.39	76.91

6. MULTISTAGE STOCHASTIC DECOMPOSITION ALGORITHM

In this section we extend the non-regularized mean-risk two-stage SD algorithm to the multi-stage setting. The fundamental idea of MSD algorithm is to solve only one LP at each stage per iteration of the algorithm to generate cut approximations. For each iteration, a sample path ω starting from the first stage to the terminal stage is randomly generated. Figure 6.1, shows a three stage scenario tree and an example of a sample path $\omega^2 = (\omega_2^2, \omega_3^2)$, represented by the black nodes.

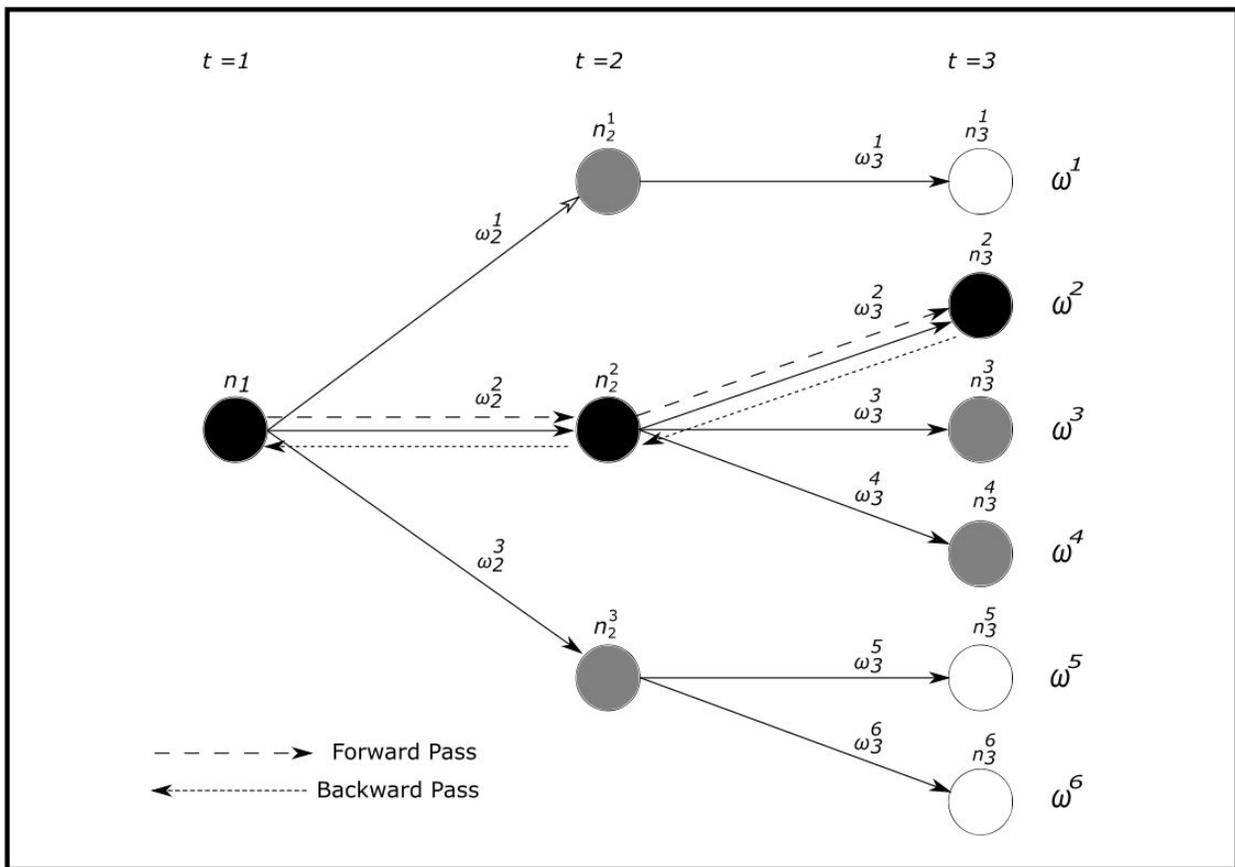


Figure 6.1: Scenario Tree

Except the root node n_1 , each node on the sample path has sibling nodes that share a common ancestor. For example, node n_2^2 on sample path ω^2 has two sibling nodes n_2^1 and n_2^3 , with one com-

mon ancestor node n_1 . After a sample path is randomly generated, the MSD algorithm executes the *forward pass* and then the *backward pass*.

During the *forward pass*, at each stage of the scenario tree an LP associated with the node on the path is solved and the solution associated with the node is updated. In the *backward pass*, the solutions associated with the nodes on sample path and the sibling nodes are used to compute the approximation of the recourse function $\mathbb{E}[f_t(x_{t-1}, \tilde{\omega}_{[t]})]$ at every stage. Next, we derive the decomposition approach for MR-MSLP with EE, QDEV and CVaR, respectively.

6.1 MSD for EE

Using the recursive formulations for MR-MSLP with EE from Proposition 3.6.1, in this section we present our detailed decomposition approach for risk measure EE. The decomposition approach at the terminal stage T is same as that of the MR-SLP, whereas for the non-terminal stages we use approximation of associated LPs. We also provide formulas for computing cut coefficients and cut constant at each stage and we follow it with detailed MSD-EE algorithm.

6.1.1 Decomposition for MR-MSLP with EE

For a given sample path ω^k at iteration k , the forward pass generates the decision vector $x_t(\omega_t^k)$ for all stages $t \in \{1, \dots, T-1\}$. In the backward pass, starting from the terminal stage the approximation η of the recourse function is computed for every stage except for the terminal stage.

Terminal Stage

Using the iterate $x_{T-1}(\omega_{T-1}^k)$ generated during the forward pass and the formulation (3.16), we have the following deterministic LP at stage T on sample path ω^k :

$$f_T^k(x_{T-1}, \omega_T^k) := \text{Min} \quad c_T^\top x_T(\omega_T^k) + \lambda_T \nu_T(\omega_T^k) \quad (6.1a)$$

$$\text{s.t.} \quad W_T x_T(\omega_T^k) \geq r_T(\omega_T^k) - T_T(\omega_T^k) x_{T-1}(\omega_{T-1}^k), \quad (6.1b)$$

$$-c_T^\top x_T(\omega_T^k) + \nu_T(\omega_T^k) \geq -\psi_T(\omega_T^k) + c_{T-1}^\top x_{T-1}(\omega_{T-1}^k), \quad (6.1c)$$

$$x_T(\omega^k) \in X_T, \nu_T(\omega^k) \geq 0.$$

Using the iterate $x_{T-1}(\omega_T^k)$ for iteration k and realization ω^k , we solve the dual of problem (6.1a-6.1c) to obtain optimal dual multipliers $\pi_T^k(\omega_T^k)$ and $\phi_T^k(\omega_T^k)$ associated with constraints (6.1b) and (6.1c), respectively. This dual solution, along with all the dual solutions generated at previous iterations at stage T for sibling nodes are used to compute the lower bounding function $\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]}) \leq \mathbb{E}[f_T(x_{T-1}, \tilde{\omega}_{[T]})]$ at stage $T - 1$, as follows:

$$\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]}) := \max\{\alpha_{T_i}^k + (\beta_{T_i}^k)^\top x_{T-1}(\omega_{T-1}^k) \mid i = 1, \dots, k\}. \quad (6.2)$$

The scalar $\alpha_{T_k}^k$ and the cut coefficients $\beta_{T_k}^k$ are calculated as follows:

$$\alpha_{T_k}^k = \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top r_T(\omega_T^i) - \phi_{T_i}^k \psi_T\}, \quad (6.3)$$

$$(\beta_{T_k}^k)^\top = \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top (-T_T(\omega_T^i)) + \phi_{T_i}^k c_{T-1}^\top\}, \quad (6.4)$$

$$\begin{aligned} (\pi_{T_i}^k, \phi_{T_i}^k) \in \operatorname{argmax}\{ & (\pi_T)^\top (r_T(\omega_T^i) - T_T(\omega_T^i) x_{T-1}^k(\omega_{T-1}^k)) + \phi_T (c_{T-1}^\top x_{T-1}^k(\omega_{T-1}^k) - \psi_T(\omega_T^i)) \\ & \mid \pi_T \in V_T^k, \phi_T \in U_T^k\}, \forall i = \{1, \dots, k-1\}, \end{aligned} \quad (6.5)$$

where k is the current iteration index, i is the enumerator of all the iterations and V_T^k and U_T^k are sets of all the optimal dual multipliers π_T and ϕ_T respectively, generated at the terminal stage t up to iteration k . The approximation function $\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]})$ provides a lower bounding affine function for $\mathbb{E}[f_T^k(x_{T-1}, \omega_{[T]}^k)]$, and using the approximation $\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]})$ we can redefine function

$f_{T-1}^k(x_{T-2}, \omega_{[T-1]}^k)$ at stage $T - 1$ as follows:

$$\begin{aligned}
f_{T-1}^k(x_{T-2}, \omega_{[T-1]}^k) &:= \text{Min} \quad c_{T-1}^\top x_{T-1}(\omega_{T-1}^k) + \eta_T^k + \lambda_{T-1} \nu_{T-1}(\omega_{T-1}^k) & (6.6) \\
\text{s.t.} \quad & W_{T-1} x_{T-1}(\omega_{T-1}^k) \geq r_{T-1}(\omega_{T-1}^k) - T_{T-1}(\omega_{T-1}^k) x_{T-2}(\omega_{T-2}), \\
& - c_{T-1}^\top x_{T-1}(\omega_{T-1}^k) - \eta_T^k + \nu_{T-1}(\omega_{T-1}^k) \geq -\psi_{T-1}(\omega_{T-1}^k) \\
& \quad + c_{T-2}^\top x_{T-2}(\omega_{T-2}^k), \\
& - (\beta_{Ti}^k)^\top x_{T-1}(\omega_{T-1}^k) + \eta_T^k \geq \alpha_{Ti}^k, \quad \forall i = 1, \dots, k \\
& x_{T-1}(\omega_{T-1}^k) \in X_{T-1}(x_{T-2}, \omega_{[T-1]}^k), \nu_{T-1}(\omega_{T-1}^k) \geq 0.
\end{aligned}$$

Non-Terminal Stage

Unlike the terminal stage, the LPs associated with the non-terminal stages are not deterministic, due to the fact that the value of $f_{t+1}(x_t, \omega_{[t]})$ in problem (3.15) and in problem (3.16) is based on realization ω . Let p_t^{kn} denote empirical frequency for node n , that is the probability that reflects the number of times node n at stage t has been visited up to iteration k given that it's parent node is visited. Therefore, since the LPs associated with non-terminal stages are not deterministic, for approximating $\eta_t^k(x_t, \tilde{\omega}_{[t+1]})$ at a non-terminal stages t , we use the cut approximation η_{t+1}^k and the empirical frequencies p_t^{kn} . From equation (3.15) the objective function for a non-terminal stage t at iteration k with cut approximation η_{t+1}^k is defined as follows:

$$\begin{aligned}
f_t^k(x_{t-1}, \omega_{[t]}^k) &:= \text{Min} \quad c_t^\top x_t(\omega_t^k) + \eta_{t+1}^k + \lambda_t \nu_t(\omega_t^k), & (6.7) \\
\text{s.t.} \quad & W_t x_t(\omega_t^k) \geq r_t(\omega_t^k) - T_t(\omega_t^k) x_{t-1}(\omega_{t-1}^k), \\
& - c_t^\top x_t(\omega_t^k) - \eta_{t+1}^k + \nu_t(\omega_t^k) \geq -\psi_t + c_{t-1}^\top x_{t-1}(\omega_{t-1}^k), \\
& - (\beta_{(t+1)i}^k)^\top x_t(\omega_t^k) + \eta_{t+1}^k \geq \alpha_{(t+1)i}^k, \quad \forall i = 1, \dots, k \\
& x_t(\omega_t^k) \in X_t(x_{t-1}, \omega_{[t]}^k), \nu_t(\omega_t^k) \geq 0, \eta_{t+1}^k \in \mathbb{R},
\end{aligned}$$

where the approximation of the recourse function $\eta_{t+1}^k(x_t, \tilde{\omega}_{[t+1]})$ is linear and convex and is defined

as follows:

$$\eta_{t+1}^k(x_t, \tilde{\omega}_{[t+1]}) := \max\{\alpha_{(t+1)i}^k + (\beta_{(t+1)i}^k)^\top x_t(\omega_t^k) \mid i = 1, \dots, k\}. \quad (6.8)$$

Now to obtain a lower bounding function $\eta_t^k(x_{t-1}, \tilde{\omega}_{[t]}) \leq \mathbb{E}[f_t(x_{t-1}, \tilde{\omega}_{[t]})]$, for all sibling nodes n at stage t , let

$$(\hat{\alpha}_{(t+1)}^{kn}, \hat{\beta}_{(t+1)}^{kn}) \in \operatorname{argmax}\{\alpha_{(t+1)i}^k + (\beta_{(t+1)i}^k)^\top x_t^n \mid i = 1, \dots, k\}, \quad (6.9)$$

where x_t^n is incumbent solution at node n . Then, the weighted average of recourse function approximation $\eta_{t+1}^k(x_t, \tilde{\omega}_{[t+1]})$ at stage t for iteration k can be stated using the empirical frequencies and set (6.9) as follows

$$\eta_{t+1}^k(x_t, \tilde{\omega}_{[t+1]}) := \sum_n p_t^{kn} (\hat{\alpha}_{(t+1)}^{kn} + (\hat{\beta}_{(t+1)}^{kn})^\top x_t). \quad (6.10)$$

Next, by substituting η_{t+1}^k from problem (6.7) with result from equation (6.10) we have following approximation at stage t :

$$F_t^k(x_{t-1}, \omega_{[t]}^k) := \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn} + \operatorname{Min}(c_t + \sum_n p_t^{kn} \hat{\beta}_{(t+1)}^{kn})^\top x_t + \lambda_t \nu_t \quad (6.11a)$$

$$\text{s.t. } W_t x_t \geq r_t(\omega_t^k) - T_t(\omega_t^k) x_{t-1}(\omega_{t-1}^k), \quad (6.11b)$$

$$\begin{aligned} - (c_t + \sum_n p_t^{kn} \hat{\beta}_{(t+1)}^{kn})^\top x_t + \nu_t &\geq -\psi_t \\ &+ c_{t-1}^\top x_{t-1}(\omega_{t-1}^k) + \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn}, \end{aligned} \quad (6.11c)$$

$$x_t \in X_t(x_{t-1}, \omega_{[t]}), \nu_t \geq 0.$$

Given the iterate $x_{t-1}(\omega_{t-1}^k)$ for iteration k and realization ω^k , we solve the dual of problem (6.11a-6.11c) to obtain the dual multipliers π_t^k and ϕ_t^k associated with constraints (6.11b) and (6.11c), respectively. Using these dual multipliers π_t^k and ϕ_t^k , we compute the scalar α_t^k and coef-

ficients β_t^k for the peicewise linear approximation η_t^k of $\mathbb{E}[f_t(x_{t-1}, \tilde{\omega}_{[t]})]$ at stage $t - 1$, as follows:

$$\alpha_t^k = (\pi_t^k)^\top r_t(\omega_t^k) - \phi_t^k \psi_t + (1 + \phi_t^k) \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn} \quad (6.12)$$

$$(\beta_t^k)^\top = (\pi_t^k)^\top (-T_t(\omega_t^k)) + \phi_t^k c_{t-1}^\top. \quad (6.13)$$

6.1.2 MSD Algorithm for MR-MSLP with EE

Let us continue to denote the algorithm iteration index by k and the iterates for stage t by $x_t^k(\omega_t^k)$. Furthermore, let incumbent solution and the candidate solution for stage $t = 1$ at iteration k be denoted by \bar{x}_1^k and x_1^k , respectively. A node under consideration on the sample path ω^k will be denoted by n , while the immediate ancestor of node n will be denoted by n_a . We use sets V_T^k and U_T^k to store all the dual variables π^k and ϕ^k generated at the terminal stage T up to iteration k . Based on results derived in Section 6.1.1, the MSD algorithm for MR-MSLP with EE can be stated as follows:

MSD-EE Algorithm

Step 0: Initialization.

Set $k \leftarrow 0$. For the terminal nodes, set $V_n^0 \leftarrow \emptyset$ and $U_n^0 \leftarrow \emptyset$. Set $\eta_n^0(x) \leftarrow -\infty$, $x_1^0 \in X_1$ and $\bar{x}_1^0 \in X$. Set target $\psi_t \in \mathbb{R}$ for $t = 1, \dots, T - 1$, and choose $\delta \in (0, 1)$. Set lower bound $L_t \in \mathbb{R}$ for each stage t .

Step 1: Generate Sample Path.

Set $k \leftarrow k + 1$. Randomly generate a sample path $\omega^k \in \tilde{\omega}$, independent of any previously generated sample paths.

Step 2: Forward Recursion.

2.1 If all the nodes on sample path ω^k are visited in previous iterations, then the approximation η_t^{k-1} defined in (6.7) will be available for all nodes n . Starting from the stage $t = 2$

and using \bar{x}_1^k , solve problem (6.7) to optimize $f_t^{k-1}(x_{t-1}, \omega_{[t]}^k)$ to obtain new incumbent solution $x_t^k(\omega_t^k)$ for all the nodes associated with stages $t = 2, \dots, T - 1$.

2.2 If some nodes on the sample path ω^k have not been visited during previous iterations, then perform step 2.1 up to a previously unseen node is reached. Then obtain initial feasible solution $x_t^k(\omega_t^k)$ for LPs associated with all the remaining non-terminal nodes on the sample path.

Step 3: Determine Cut Approximation η_t^k .

Begin backward recursion from the terminal node and trace back the sample path to the root node.

a. Solve the dual problem (6.1a)-(6.1c) for the terminal node on the sample path ω^k to obtain the dual multipliers:

$$(\pi_T^k(\omega^k), \phi_T^k(\omega^k)) \in \text{Max}\{(\pi_T^k)^\top (r_T(\omega^k) - T_T(\omega^k)x_{T-1}(\omega^k)) + \phi_T^k(c_{T-1}^\top x_{T-1}(\omega^k) - \psi_T) \\ | \pi_T^\top W_T - c_T \phi_T^k \leq c_T, \phi_T^k \leq \lambda_T, \pi_T^k \geq 0, \phi_T^k \geq 0\}.$$

b. Update sets V_T^k and U_T^k :

$$V_T^k \leftarrow V_T^{k-1} \cup \pi_T^k(\omega^k).$$

$$U_T^k \leftarrow U_T^{k-1} \cup \phi_T^k(\omega^k).$$

c. Determine the coefficients of the k -th cutting plane for node n_a :

$$(\pi_{T_i}^k, \phi_{T_i}^k) \in \text{argmax}\{(\pi_T^k)^\top (r_T(\omega^i) - T_T(\omega^i)x_{T-1}(\omega^i)) + \phi_T^k(c_{T-1}^\top x_{T-1}(\omega^i) - \psi_T) \\ | \pi_T \in V_T^k, \phi_T \in U_T^k\}, \forall i \in \{1, 2, \dots, k-1\}.$$

$$\alpha_{T_k}^k \leftarrow \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top r_T(\omega^i) - \phi_{T_i}^k \psi_T\}.$$

$$(\beta_{T_k}^k)^\top \leftarrow \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top (-T_T(\omega^i)) + \phi_{T_i}^k c^\top\}.$$

d. Determine the cut coefficients of the k -th cutting plane for all nodes on sample path ω^k , starting from stage $t = T - 2, T - 3, \dots, 1$ by solving dual of problem (6.11a)-(6.11c) to obtain α_t^k and $(\beta_t^k)^\top$ using equation (6.12) and (6.13), respectively.

e. Update coefficients of all previously generated cuts associated with each node n on sample path ω^k .

$$\alpha_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)\alpha_{ni}^{k-1} + \left(\frac{1}{k}\right)L_n, \quad \forall i = 1, \dots, k-1.$$

$$\beta_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)\beta_{ni}^{k-1}, \quad \forall i = 1, \dots, k-1.$$

f. Update $f_t^k(x_t, \omega_{[t]}^k)$ for each node n on sample path ω^k .

$$\eta_n^k - (\beta_{ni}^k)^\top x_{n-} \geq \alpha_{ni}^k, \quad \forall i = 1, \dots, k.$$

g. Update approximation for each node n not on the sample path ω^k .

$$f_n^k(x_t, \omega^k) \leftarrow f_{n-1}^k(x_t, \omega^k)$$

Step 4. Update Incumbent Solution.

if $f_1^k(x_1^k) - f_1^k(\bar{x}_1^{k-1}) < \delta[f_1^{k-1}(x_1^k) - f_1^{k-1}(\bar{x}_1^{k-1})]$,

$$\bar{x}_1^k \leftarrow x_1^k$$

else

$$\bar{x}_1^k \leftarrow \bar{x}_1^{k-1}.$$

Step 5. Solve Master Problem.

$$\text{Min}_{x, \eta^k} \quad c_1^\top x_1 + \eta_2^k$$

$$\text{s.t.} \quad x_1 \in X_1,$$

$$\eta_2^k - (\beta_{2i}^k)^\top x_1 \geq \alpha_{2i}^k, \quad \forall i = \{1, \dots, k\},$$

to get the new candidate solution x_1^{k+1} .

Step 6. Termination Criterion.

If the following criterion is not satisfied, return to step 1. (*Termination criteria based on objective value*):

For a large enough k , terminate the algorithm if

$$f_1^{k-1}(\bar{x}_1^{k-1}) - f_1^{k-1}(x_1^k) \leq \epsilon, \text{ where } \epsilon > 0 \text{ is a given tolerance level.}$$

REMARK 6.1.1. In addition to the termination criteria defined in step 6 of MSD-EE algorithm, certain characteristics of the MSD algorithm can be used to design additional termination criteria. For example, it is known that the incumbent solution \bar{x}_1^k changes only finitely often. Therefore if m_k is the number of times the incumbent solution changes up to iteration k . Then eventually $k - m_k$ will increase without bound. Therefore, a termination criterion based on incumbent objective value can be defined as follows:

Let $\{\bar{x}_1^{k_n}\}_{k_n=1}^{m_k}$ a subsequence of $\{x_1^n\}_{n=1}^k$, be a collection of all incumbent solutions up to iteration k . Define

$$\gamma^k = \frac{1}{k} \sum_{t=1}^k f_t(\bar{x}_1^t) \text{ and } \bar{\gamma}^k = \frac{1}{m_k} \sum_{n=1}^{m_k} f_{k_n}(\bar{x}_1^{k_n}).$$

Terminate the algorithm for a large enough k , if

$$|(f_k(\bar{x}_1^k) - \gamma^{k-1})/\bar{\gamma}^{k-1}| \leq \epsilon, \text{ where } \epsilon > 0 \text{ is a given tolerance level.}$$

6.2 MSD for QDEV

In this section we present our detailed decomposition approach for MR-MSLP with QDEV using the recursive formulations described in Proposition 3.6.1. At the terminal stage T , the problem is deterministic and the dual solutions can be used for generating MSD cuts. Whereas, at the non-terminal stages we solve an approximation of the associated LP to generate cuts. Next, using the described decomposition approach we present the detailed MSD-QDEV algorithm.

6.2.1 Decomposition for MR-MSLP with QDEV

Similar to the decomposition approach of EE, in case of QDEV for a given sample path ω^k at iteration k , the forward pass generates the decision vector $x_t^k(\omega_t^k)$ for all stages $t \in \{1, \dots, T-1\}$. Then, the approximation η_t of the recourse function is computed at each stage during the backward pass, beginning from the terminal stage. Unlike EE, in QDEV we also need to determine the α -quantile ψ_t for all non-terminal nodes to compute variable ν_t .

Terminal Stage

From the iterate $x_{T-1}(\omega_{T-1}^k)$ and using the formulation (3.20), we have the following deterministic linear problem for the terminal node at stage T on sample path ω^k :

$$f_T^k(x_{T-1}, \omega_{[T]}^k) := \text{Min} \quad (1 - \lambda_T \epsilon_1) c_T^\top x_T(\omega_T^k) + \lambda_T (\epsilon_1 + \epsilon_2) \nu_T(\omega_T^k) \quad (6.14a)$$

$$\text{s.t.} \quad W_T x_T(\omega_T^k) \geq r_T(\omega_T^k) - T_T(\omega_T^k) x_{T-1}(\omega_{T-1}^k), \quad (6.14b)$$

$$- c_T^\top x_T(\omega_T^k) + \nu_T(\omega_T^k) \geq c_{T-1}^\top x_{T-1}(\omega_{T-1}^k) - \psi_{T-1}(\omega_{T-1}^k), \quad (6.14c)$$

$$x_T(\omega_T^k) \in X_T(x_{T-1}, \omega_{[T]}), \nu_T(\omega_T^k) \geq 0.$$

Using the iterate $x_{T-1}(\omega_{T-1}^k)$ for iteration k and realization ω^k , we solve the dual of problem (6.14a-6.14c) to obtain dual multipliers $\pi_T^k(\omega_T^k)$ and $\phi_T^k(\omega_T^k)$ associated with constraints (6.14b) and (6.14c), respectively. This dual solution, along with all the dual solutions generated in past iterations at stage T are used to compute the approximation $\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]})$ at stage $T - 1$, defined as follows:

$$\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]}) := \max\{\alpha_{T_i}^k + (\beta_{T_i}^k)^\top x_{T-1}(\omega_{T-1}^k) + \gamma_{T_i}^k \psi_{T-1}(\omega_{T-1}^k) \mid i = 1, \dots, k\}. \quad (6.15)$$

The cut scalar $\alpha_{T_k}^k$ and the cut coefficients $\beta_{T_k}^k$ and $\gamma_{T_k}^k$ are calculated as follows:

$$\alpha_{T_k}^k = \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top r_T(\omega^i)\}, \quad (6.16)$$

$$(\beta_{T_k}^k)^\top = \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top (-T_T(\omega^i)) + \phi_{T_i}^k (c_{T-1}^\top)\}, \quad (6.17)$$

$$\gamma_{T_k}^k = \frac{1}{k} \sum_{i=1}^k -\phi_{T_i}^k, \quad (6.18)$$

$$\begin{aligned} (\pi_{T_i}^k, \phi_{T_i}^k) \in \text{argmax} \{ & (\pi_T)^\top (r_T(\omega^i) - T_T(\omega^i) x_{T-1}^k(\omega^k)) + \phi_T (c_{T-1}^\top x_{T-1}^k(\omega^k) - \psi_{T-1}(\omega^k)) \\ & \mid \pi_T \in V_T^k, \phi_T \in U_T^k, \forall i = \{1, \dots, k-1\} \}. \end{aligned} \quad (6.19)$$

Non-Terminal Stage

The value of function $f_{t+1}(x_t, \omega_{[t+1]})$ in problems (3.17) and (3.18) is based on realization ω_{t+1} . Therefore, the LPs associated with the non-terminal stages are not deterministic. Hence, for computation of approximation function at a non-terminal stages t , we use the cut approximation η_{t+1}^k and the empirical frequencies p_t^{kn} . From equation (3.18) the objective function for a non-terminal stage t at iteration k with cut approximation η_{t+1}^k is defined as follows:

$$\begin{aligned}
f_t^k(x_{t-1}, \omega_{[t]}^k) := \text{Min} \quad & (1 - \lambda_t \epsilon_1) [(1 - \lambda_{t+1} \epsilon_1) c_t^\top x_t(\omega_t^k) + \lambda_{t+1} \epsilon_1 \psi_t(\omega_t^k) + \eta_{t+1}^k] \\
& + \lambda_t (\epsilon_1 + \epsilon_2) \nu_t(\omega_t^k) \tag{6.20} \\
\text{s.t.} \quad & W_t x_t(\omega_t^k) \geq r_t(\omega_t^k) - T_t(\omega_t^k) x_{t-1}(\omega_{t-1}^k), \\
& - (1 - \lambda_{t+1} \epsilon_1) c_t^\top x_t(\omega_t^k) - \lambda_{t+1} \epsilon_1 \psi_t(\omega_t^k) - \eta_{t+1}^k \\
& \quad + \nu_t(\omega_t^k) \geq c_{t-1}^\top x_{t-1}(\omega_{t-1}^k) - \psi_{t-1}(\omega_{t-1}^k), \\
& - (\beta_{(t+1)i}^k)^\top x_t(\omega_t^k) - \gamma_{(t+1)i}^k \psi_t(\omega_t^k) + \eta_{t+1}^k \geq \alpha_{(t+1)i}^k, \quad \forall i = 1, \dots, k \\
& x_t(\omega_t^k) \in X_t(x_{t-1}, \omega_{[t]}), \psi_t(\omega_t^k) \in \mathbb{R}, \nu_t(\omega_t^k) \geq 0, \eta_{t+1}^k \in \mathbb{R},
\end{aligned}$$

where the approximation $\eta_{t+1}^k(x_t, \tilde{\omega}_{[t+1]})$ is linear and convex and is defined as follows:

$$\eta_{t+1}^k(x_t, \tilde{\omega}_{[t+1]}) := \max\{\alpha_{(t+1)i}^k + (\beta_{(t+1)i}^k)^\top x_t(\omega_t^k) + \gamma_{(t+1)i}^k \psi_t(\omega_t^k) \mid i = 1, \dots, k\}.$$

Let for a sibling node n at stage t

$$(\hat{\alpha}_{(t+1)}^{kn}, \hat{\beta}_{(t+1)}^{kn}, \hat{\gamma}_{(t+1)}^{kn}) \in \text{argmax}\{\alpha_{(t+1)i}^k + (\beta_{(t+1)i}^k)^\top x_t^n + \gamma_{(t+1)i}^k \psi_t^n \mid i = 1, \dots, k\}. \tag{6.21}$$

As in case of EE, substituting set (6.21) and empirical probabilities in formulation (6.20) we have

following approximation at stage t :

$$\begin{aligned}
F_t^k(x_{t-1}, \omega_{[T]}^k) := & (1 - \lambda_t \epsilon_1) \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn} + \text{Min} \quad (1 - \lambda_t \epsilon_1) [(1 - \lambda_{t+1} \epsilon_1) c_t \\
& + \sum_n p_t^{kn} \hat{\beta}_{(t+1)}^{kn}]^\top x_t + (\lambda_{t+1} \epsilon_1 + \sum_n p_t^{kn} \hat{\gamma}_{(t+1)}^{kn}) \psi_t] \\
& + \lambda_t (\epsilon_1 + \epsilon_2) \nu_t
\end{aligned} \tag{6.22a}$$

$$\text{s.t.} \quad W_t x_t \geq r_t(\omega_{T-1}^k) - T_t(\omega^k) x_{t-1}(\omega^k), \tag{6.22b}$$

$$\begin{aligned}
& - ((1 - \lambda_{t+1} \epsilon_1) c_t + \sum_n p_t^{kn} \hat{\beta}_{(t+1)}^{kn})^\top x_t - (\lambda_{t+1} \epsilon_1 \\
& + \sum_n p_t^{kn} \hat{\gamma}_{(t+1)}^{kn}) \psi_t + \nu_t \geq -\psi_{t-1}(\omega_{t-1}^k) \\
& + c_{t-1}^\top x_{t-1}(\omega_{t-1}^k) + \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn},
\end{aligned} \tag{6.22c}$$

$$x_t \in X_t(x_{t-1}, \omega_{[t]}), \psi_t \in \mathbb{R}, \nu_t \geq 0.$$

Given the iterate $x_{t-1}(\omega_t^k)$ for iteration k and realization ω^k , we solve the dual of problem (6.22a-6.22c) to obtain the dual multipliers π_t^k and ϕ_t^k associated with constraints (6.22b) and (6.22c), respectively. Using the dual multipliers π_t^k and ϕ_t^k , we compute the scalar α_t^k and the cut coefficients β_t^k and γ_t^k for the linear approximation η_t^k at stage $t - 1$, as follows:

$$\alpha_t^k = (\pi_t^k)^\top r_t(\omega^k) + (1 - \lambda_t \epsilon_1 + \phi_t^k) \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn} \tag{6.23}$$

$$(\beta_t^k)^\top = (\pi_t^k)^\top (-T_t(\omega^k)) + \phi_t^k c_{t-1}^\top, \tag{6.24}$$

$$\gamma_t^k = -\phi_t^k. \tag{6.25}$$

6.2.2 MSD Algorithm for MR-MSLP with QDEV

We continue to denote the algorithm iteration index by k and the iterates for stage t by $x_t^k(\omega_t^k)$, while the immediate ancestor of node n will be denoted by n_a . We use sets V_T^k and U_T^k to store all the dual variables π^k and ϕ^k generated at the terminal stage T up to iteration k . Based on results

derived in Section 6.2.1, the MSD algorithm for MR-MSLP with QDEV is stated as follows:

MSD-QDEV Algorithm

Step 0: Initialization.

Set $k \leftarrow 0$. For the terminal nodes, set $V_n^0 \leftarrow \emptyset$ and $U_n^0 \leftarrow \emptyset$. Set $\eta_n^0(x) \leftarrow -\infty$, $x_1^0 \in X_1$ and $\bar{x}_1^0 \in X$. Assume α -quantile $\psi_t \in \mathbb{R}$ for the stage t , and $\delta \in (0, 1)$. Set lower bound $L_t \in \mathbb{R}$ for each stage t .

Step 1: Generate Sample Path.

Set $k \leftarrow k + 1$. Randomly generate a sample path $\omega^k \in \tilde{\omega}$, independent of any previously generated sample paths.

Step 2: Forward Recursion.

2.1 If all the nodes on sample path ω^k are visited in previous iterations, then the approximation η_t^{k-1} defined in (6.20) will be available for all nodes n . Starting from the stage $t = 2$ and using \bar{x}_1^k , solve problem (6.20) to optimize $f_t^{k-1}(x_{t-1}, \omega_{[t]}^k)$ to obtain new incumbent solution $x_t^k(\omega^k)$ for all the nodes associated with stages $t = 2, \dots, T - 1$.

2.2 If some nodes on the sample path ω^k have not been visited during previous iterations, then perform step 2.1 up to a previously unseen node is reached. Then obtain initial feasible solution $x_t^k(\omega^k)$ for LPs associated with all the remaining non-terminal nodes on the sample path.

Step 3: Determine Cut Approximation η_t^k .

Begin backward recursion from the terminal node and trace back the sample path to the root node.

- a. Solve the dual problem (6.14a)-(6.14c) for the terminal node on the sample path ω^k to

obtain the dual solution:

$$(\pi_T^k(\omega^k), \phi_T^k(\omega^k)) \in \text{Max}\{(\pi_T^k)^\top (r_T(\omega^k) - T_T(\omega^k)x_{T-1}(\omega^k)) + \phi_T^k(c_{T-1}^\top x_{T-1}(\omega^k) - \psi_T) \\ | \pi_T^\top W_T - c_T \phi_T^k \leq c_T, \phi_T^k \leq \lambda_T(\epsilon_1 + \epsilon_2), \pi_T^k \geq 0, \phi_T^k \geq 0\}.$$

b. Update sets V_T^k and U_T^k :

$$V_T^k \leftarrow V_T^{k-1} \cup \pi_T^k(\omega^k).$$

$$U_T^k \leftarrow U_T^{k-1} \cup \phi_T^k(\omega^k).$$

c. Determine the coefficients of the k -th cutting plane for node n_a :

$$(\pi_{Ti}^k, \phi_{Ti}^k) \in \text{argmax}\{(\pi_T)^\top (r_T(\omega^i) - T_T(\omega^i)x_{T-1}(\omega^i)) + \phi_T(c_{T-1}^\top x_{T-1}(\omega^i) - \psi_T) \\ | \pi_T \in V_T^k, \phi_T \in U_T^k\}, \forall i \in \{1, 2, \dots, k-1\}.$$

$$\alpha_{Tk}^k \leftarrow \frac{1}{k} \sum_{i=1}^k \{(\pi_{Ti}^k)^\top r_T(\omega^i)\}.$$

$$(\beta_{Tk}^k)^\top \leftarrow \frac{1}{k} \sum_{i=1}^k \{(\pi_{Ti}^k)^\top (-T_T(\omega^i)) + \phi_{Ti}^k c^\top\}.$$

$$\gamma_{Tk}^k \leftarrow \frac{1}{k} \sum_{i=1}^k -\phi_{Ti}^k.$$

d. Determine the cut coefficients of the k -th cutting plane for all nodes on sample path

ω^k , starting from stage $t = T-2, T-3, \dots, 1$ by solving dual of problem

(6.22a)-(6.22c) to obtain α_t^k , $(\beta_t^k)^\top$ and γ_t^k using equation (6.23) and (6.24),

respectively.

e. Update coefficients of all previously generated cuts associated with each node n on

sample path ω^k .

$$\alpha_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)\alpha_{ni}^{k-1} + \left(\frac{1}{k}\right)L_n, \quad \forall i = 1, \dots, k-1.$$

$$\beta_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)\beta_{ni}^{k-1}, \quad \forall i = 1, \dots, k-1.$$

$$\gamma_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)(\beta_{ni}^{k-1})_0, \quad \forall i = 1, \dots, k-1.$$

f. Update $f_t^k(x_t, \omega^k)$ for each node n on sample path ω^k .

$$\eta_n^k - (\beta_{ni}^k)^\top x_{n_a} - \gamma_{ni}^k \psi_{n_a} \geq \alpha_{ni}^k, \quad \forall i = 1, \dots, k.$$

g. Update approximation for each node n not on the sample path ω^k .

$$f_n^k(x_t, \omega^k) \leftarrow f_{n-1}^k(x_t, \omega^k)$$

Step 4. Update Incumbent Solution.

$$\text{if } f_1^k(x_1^k) - f_1^k(\bar{x}_1^{k-1}) < \delta[f_1^{k-1}(x_1^k) - f_1^{k-1}(\bar{x}_1^{k-1})],$$

$$\bar{x}_1^k \leftarrow x_1^k$$

else

$$\bar{x}_1^k \leftarrow \bar{x}_1^{k-1}.$$

Step 5. Solve Master Problem.

$$\begin{aligned} \text{Min}_{x, \eta^k} \quad & (1 - \lambda_2 \epsilon_1) c_1^\top x_1 + \lambda_2 \epsilon_1 \psi_1 + \eta_2^k \\ \text{s.t.} \quad & x_1 \in X_1, \\ & \eta_2^k - (\beta_{2i}^k)^\top x_1 - \gamma_{2i}^k \psi_1 \geq \alpha_{2i}^k, \quad \forall i = \{1, \dots, k\}, \end{aligned}$$

to get the new candidate solution x_1^{k+1} .

Step 6. Termination Criterion.

If the following criterion is not satisfied, return to step 1.

For a large enough k , terminate the algorithm if

$$f_1^{k-1}(\bar{x}_1^{k-1}) - f_1^{k-1}(x_1^k) \leq \epsilon, \text{ where } \epsilon > 0 \text{ is a given tolerance level.}$$

6.3 MSD for CVaR

Following the same approach as for the MR-MSLP with EE and QDEV, we present our detailed decomposition approach and MSD algorithm for CVaR using the recursive formulations of MR-MSLP with CVaR from Proposition 3.6.1. AT the stage T , the cut coefficients are computed using the same approach as two-stage MR-SLP and for the non-terminal stages we solve an approximation of LPs to generate MSD cuts.

6.3.1 Decomposition for MR-MSLP with CVaR

We follow the same decomposition approach as that of QDEV for CVaR. For a given sample path ω^k at iteration k , the forward pass generates the decision vector $x_t(\omega_t^k)$ for all stages $t \in \{1, \dots, T-1\}$. Then, starting from the terminal stage the cut coefficients η for the approximation of the recourse function is computed at each stage during the backward pass. We also need to determine the α -quantile ψ at each stage for computation of CVaR.

Terminal Stage

Given the decision vector $x_{T-1}(\omega_{T-1}^k)$ and α -quantile $\psi_{T-1}(\omega_{T-1}^k)$, generated during the forward pass at stage $T-1$ and the formulation (3.23), we have the following LP for the terminal node at stage T for sample path ω^k :

$$f_T^k(x_{T-1}, \omega_{[T]}^k) := \text{Min} \quad c_T^\top x_T(\omega_T^k) + \frac{\lambda_T}{1-\alpha} \nu_T(\omega_T^k) \quad (6.26a)$$

$$\text{s.t.} \quad W_T x_T(\omega_T^k) \geq r_T(\omega_T^k) - T_T(\omega_T^k) x_{T-1}(\omega_{T-1}^k), \quad (6.26b)$$

$$-c_T^\top x_T(\omega_T^k) + \nu_T(\omega_T^k) \geq -\psi_{T-1}(\omega_{T-1}^k) + c_{T-1}^\top x_{T-1}(\omega_{T-1}^k), \quad (6.26c)$$

$$x_T(\omega_T^k) \in X_T(x_{T-1}, \omega_{[T]}), \nu_T(\omega_T^k) \geq 0.$$

Given the iterate $x_{T-1}(\omega^k)$ and $\psi_{T-1}(\omega^k)$ for iteration k and realization ω^k , we solve the dual of problem (6.26a-6.26c) to obtain dual multipliers $\pi_T^k(\omega^k)$ and $\phi_T^k(\omega^k)$ associated with constraints (6.14b) and (6.14c), respectively. The dual solution of problem (6.26a-6.26c) at iteration k , along with the dual solutions of all the past iterations generated at stage T are then used to compute the approximation $\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]})$ of $\mathbb{E}[f_T(x_{T-1}, \tilde{\omega}_{[T]})]$ at stage $T-1$, as follows:

$$\eta_T^k(x_{T-1}, \tilde{\omega}_{[T]}) := \max\{\alpha_{T_i}^k + (\beta_{T_i}^k)^\top x_{T-1}(\omega_T^k) + \gamma_{T_i}^k \psi_{T-1}(\omega_T^k) \mid i = 1, \dots, k\}. \quad (6.27)$$

The scalar α_{Tk}^k and the cut coefficients β_{Tk}^k and γ_{Tk}^k are defined as follows:

$$\alpha_{Tk}^k = \frac{1}{k} \sum_{i=1}^k \{(\pi_{Ti}^k)^\top r_T(\omega_T^i)\}, \quad (6.28)$$

$$(\beta_{Tk}^k)^\top = \frac{1}{k} \sum_{i=1}^k \{(\pi_{Ti}^k)^\top (-T_T(\omega_T^i)) + \phi_{Ti}^k(c_{T-1}^\top)\}, \quad (6.29)$$

$$\gamma_{Tk}^k = \frac{1}{k} \sum_{i=1}^k -\phi_{Ti}^k, \quad (6.30)$$

$$\begin{aligned} (\pi_{Ti}^k, \phi_{Ti}^k) \in \operatorname{argmax} \{ & (\pi_T)^\top (r_T(\omega_T^i) - T_T(\omega_T^i)x_{T-1}^k(\omega_{T-1}^k)) + \phi_T(c_{T-1}^\top x_{T-1}^k(\omega_{T-1}^k) \\ & - \psi_{T-1}(\omega_{T-1}^k)) \mid \pi_T \in V_T^k, \phi_T \in U_T^k\}, \forall i = \{1, \dots, k-1\}. \end{aligned} \quad (6.31)$$

Non-Terminal Stage

The value of $f_{t+1}(x_t, \omega_{[t+1]})$ associated with CVaR, defined in problem (3.21) and in problem (3.22) is based on realization ω_{t+1} . Therefore, just like EE and QDEV the LPs associated with the non-terminal stages of CVaR are not deterministic. Hence, for computation of approximation function at a non-terminal stages t , we use the cut approximation η_{t+1}^k and the empirical frequencies p_t^{kn} . From equation (3.22) the objective function for a non-terminal stage t at iteration k with cut approximation η_{t+1}^k is defined as follows:

$$\begin{aligned} f_t^k(x_{t-1}, \omega_{[t]}^k) := \operatorname{Min} \quad & c_t^\top x_t(\omega_t^k) + \lambda_{t+1} \psi_t(\omega_t^k) + \eta_{t+1}^k + \frac{\lambda_t}{1-\alpha} \nu_t(\omega_t^k) \quad (6.32) \\ \text{s.t.} \quad & W_t x_t(\omega_t^k) \geq r_t(\omega_t^k) - T_t(\omega_t^k) x_{t-1}(\omega_{t-1}^k), \\ & -c_t^\top x_t(\omega_t^k) - \lambda_{t+1} \psi_t(\omega_t^k) - \eta_{t+1}^k + \nu_t(\omega_t^k) \geq -\psi_{t-1}(\omega_{t-1}^k) + c_{t-1}^\top x_{t-1}(\omega_{t-1}^k), \\ & -(\beta_{(t+1)i}^k)^\top x_t(\omega_t^k) - \gamma_{(t+1)i}^k \psi_t(\omega_t^k) + \eta_{t+1}^k \geq \alpha_{(t+1)i}^k, \quad \forall i = 1, \dots, k \\ & x_t(\omega_t^k) \in X_t(x_{t-1}, \omega_{[t]}^k), \psi_t(\omega_t^k) \in \mathbb{R}, \nu_t(\omega_t^k) \geq 0, \eta_{t+1}^k \in \mathbb{R}, \end{aligned}$$

where the approximation of the recourse function $\eta_{t+1}^k(x_t)$ is linear and convex and is defined as

follows:

$$\eta_{t+1}^k(x_t, \tilde{\omega}_{[t+1]}) := \max\{\alpha_{(t+1)i}^k + (\beta_{(t+1)i}^k)^\top x_t(\omega_t^k) + \gamma_{(t+1)i}^k \psi_t(\omega_t^k) \mid i = 1, \dots, k\}.$$

For all sibling nodes n at stage t , let

$$(\hat{\alpha}_{(t+1)}^{kn}, \hat{\beta}_{(t+1)}^{kn}, \hat{\gamma}_{(t+1)}^{kn}) \in \operatorname{argmax}\{\alpha_{(t+1)i}^k + (\beta_{(t+1)i}^k)^\top x_t^n + \gamma_{(t+1)i}^k \psi_t^n \mid i = 1, \dots, k\}. \quad (6.33)$$

Following the similar approach as that of EE, substituting set (6.33) and the empirical probabilities in formulation (6.32), we have following approximation at stage t :

$$\begin{aligned} F_t^k(x_{t-1}, \omega_{[t]}^k) := & \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn} + \operatorname{Min} \left(c_t + \sum_n p_t^{kn} \hat{\beta}_{(t+1)}^{kn} \right)^\top x_t \\ & + (\lambda_{t+1} + \sum_n p_t^{kn} \hat{\gamma}_{(t+1)}^{kn}) \psi_t + \frac{\lambda_t}{1-\alpha} \nu_t \end{aligned} \quad (6.34a)$$

$$\text{s.t. } W_t x_t \geq r_t(\omega_t^k) - T_t(\omega_t^k) x_{t-1}(\omega_{t-1}^k), \quad (6.34b)$$

$$\begin{aligned} - \left(c_t + \sum_n p_t^{kn} \hat{\beta}_{(t+1)}^{kn} \right)^\top x_t - (\lambda_{t+1} + \sum_n p_t^{kn} \hat{\gamma}_{(t+1)}^{kn}) \psi_t \\ + \nu_t \geq -\psi_{t-1}(\omega_{t-1}^k) + c_{t-1}^\top x_{t-1}(\omega_{t-1}^k) + \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn}, \end{aligned} \quad (6.34c)$$

$$x_t \in X_t(x_{t-1}, \omega_{[t]}), \psi_t \in \mathbb{R}, \nu_t \geq 0.$$

Given the iterate $x_{t-1}(\omega_{t-1})$ for iteration k and realization ω^k , we solve the dual of problem (6.34a-6.34c) to obtain the dual multipliers π_t^k and ϕ_t^k associated with constraint (6.34b) and (6.34c), respectively. Using the dual multipliers π_t^k and ϕ_t^k , we then compute the scalar α_t^k and the

cut coefficients β_t^k and γ_t^k for the linear approximation η_t^k at stage $t - 1$, as follows:

$$\alpha_t^k = (\pi_t^k)^\top r_t(\omega^k) + (1 + \phi_t^k) \sum_n p_t^{kn} \hat{\alpha}_{(t+1)}^{kn} \quad (6.35)$$

$$(\beta_t^k)^\top = (\pi_t^k)^\top (-T_t(\omega^k)) + \phi_t^k c_{t-1}^\top, \quad (6.36)$$

$$\gamma_t^k = -\phi_t^k. \quad (6.37)$$

6.3.2 MSD Algorithm for MR-MSLP with CVaR

We use sets V_T^k and U_T^k to store all the dual variables π^k and ϕ^k generated at the terminal stage T up to iteration k . We continue to denote the iterates for stage t by $x_t^k(\omega_t^k)$, while the immediate ancestor of node n will be denoted by n_a . Based on results derived in Section 6.3.1, the MSD algorithm for MR-MSLP with CVaR is stated as follows:

MSD-CVaR Algorithm

Step 0: Initialization.

Set $k \leftarrow 0$. For the terminal nodes, set $V_n^0 \leftarrow \emptyset$ and $U_n^0 \leftarrow \emptyset$. Set $\eta_n^0(x) \leftarrow -\infty$, $x_1^0 \in X_1$ and $\bar{x}_1^0 \in X$. Initialize α -quantile $\psi_t \in \mathbb{R}$ for $t = 1, \dots, T - 1$, and choose $\delta \in (0, 1)$. Set lower bound $L_t \in \mathbb{R}$ for each stage t .

Step 1: Generate Sample Path.

Set $k \leftarrow k + 1$. Randomly generate a sample path $\omega^k \in \tilde{\omega}$, independent of any previously generated sample paths.

Step 2: Forward Recursion.

2.1 If all the nodes on sample path ω^k are visited in previous iterations, then the approximation η_t^{k-1} defined in (6.32) will be available for all nodes n . Starting from the stage $t = 2$ and using \bar{x}_1^k , solve problem (6.32) to optimize $f_t^{k-1}(x_{t-1}, \omega_{[t]}^k)$ to obtain new incumbent

solution $x_t^k(\omega^k)$ for all the nodes associated with stages $t = 2, \dots, T - 1$.

2.2 If some nodes on the sample path ω^k have not been visited during previous iterations, then perform step 2.1 up to a previously unseen node is reached. Then obtain initial feasible solution $x_t^k(\omega^k)$ for LPs associated with all the remaining non-terminal nodes on the sample path.

Step 3: Determine Cut Approximation η_t^k .

Begin backward recursion from the terminal node and trace back the sample path to the root node.

a. Solve the dual problem (6.26a)-(6.26c) for the terminal node on the sample path ω^k :

$$(\pi_T^k(\omega^k), \phi_T^k(\omega^k)) \in \text{Max}\{(\pi_T^k)^\top (r_T(\omega^k) - T_T(\omega^k)x_{T-1}(\omega^k)) + \phi_T^k(c_{T-1}^\top x_{T-1}(\omega^k) - \psi_T) \\ | \pi_T^\top W_T - c_T \phi_T^k \leq c_T, \phi_T^k \leq \lambda_T/(1 - \alpha), \pi_T^k \geq 0, \phi_T^k \geq 0\}.$$

b. Update sets V_T^k and U_T^k :

$$V_T^k \leftarrow V_T^{k-1} \cup \pi_T^k(\omega^k).$$

$$U_T^k \leftarrow U_T^{k-1} \cup \phi_T^k(\omega^k).$$

c. Determine the coefficients of the k -th cutting plane for node n_a :

$$(\pi_{T_i}^k, \phi_{T_i}^k) \in \text{argmax}\{(\pi_T)^\top (r_T(\omega^i) - T_T(\omega^i)x_{T-1}(\omega^i)) + \phi_T(c_{T-1}^\top x_{T-1}(\omega^i) - \psi_T) \\ | \pi_T \in V_T^k, \phi_T \in U_T^k\}, \forall i \in \{1, 2, \dots, k-1\}.$$

$$\alpha_{T_k}^k \leftarrow \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top r_T(\omega^i)\}.$$

$$(\beta_{T_k}^k)^\top \leftarrow \frac{1}{k} \sum_{i=1}^k \{(\pi_{T_i}^k)^\top (-T_T(\omega^i)) + \phi_{T_i}^k c^\top\}.$$

$$\gamma_{T_k}^k \leftarrow \frac{1}{k} \sum_{i=1}^k -\phi_{T_i}^k.$$

d. Determine the cut coefficients of the k -th cutting plane for all nodes on sample path

ω^k , starting from stage $t = T - 2, T - 3, \dots, 1$ by solving dual of problem

(6.34a)-(6.34c) to obtain α_t^k , $(\beta_t^k)^\top$ and γ_t^k using equation (6.35) and (6.36),

respectively.

- e. Update coefficients of all previously generated cuts associated with each node n on sample path ω^k .

$$\alpha_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)\alpha_{ni}^{k-1} + \left(\frac{1}{k}\right)L_n, \quad \forall i = 1, \dots, k-1.$$

$$\beta_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)\beta_{ni}^{k-1}, \quad \forall i = 1, \dots, k-1.$$

$$\gamma_{ni}^k \leftarrow \left(\frac{k-1}{k}\right)(\beta_{ni}^{k-1})_0, \quad \forall i = 1, \dots, k-1.$$

- f. Update $f_t^k(x_t, \omega^k)$ for each node n on sample path ω^k .

$$\eta_n^k - (\beta_{ni}^k)^\top x_{na} - \gamma_{ni}^k \psi_{na} \geq \alpha_{ni}^k, \quad \forall i = 1, \dots, k.$$

- g. Update approximation for each node n not on the sample path ω^k .

$$f_n^k(x_t, \omega^k) \leftarrow f_{n-1}^k(x_t, \omega^k)$$

Step 4. Update Incumbent Solution.

$$\text{if } f_1^k(x_1^k) - f_1^k(\bar{x}_1^{k-1}) < \delta[f_1^{k-1}(x_1^k) - f_1^{k-1}(\bar{x}_1^{k-1})],$$

$$\bar{x}_1^k \leftarrow x_1^k$$

else

$$\bar{x}_1^k \leftarrow \bar{x}_1^{k-1}.$$

Step 5. Solve Master Problem.

$$\text{Min}_{x, \eta^k} \quad c_1^\top x_1 + \lambda_2 \psi_1 + \eta_2^k$$

$$\text{s.t. } \quad x_1 \in X_1,$$

$$\eta_2^k - (\beta_{2i}^k)^\top x_1 - \gamma_{2i}^k \psi_1 \geq \alpha_{2i}^k, \quad \forall i = \{1, \dots, k\},$$

to get the new candidate solution x_1^{k+1} .

Step 6. Termination Criterion.

If the following criterion is not satisfied, return to step 1.

For a large enough k , terminate the algorithm if

$$f_1^{k-1}(\bar{x}_1^{k-1}) - f_1^{k-1}(x_1^k) \leq \epsilon, \text{ where } \epsilon > 0 \text{ is a given tolerance level.}$$

In the next section we apply the mean-risk MSD algorithm with risk measures EE, QDEV and CVaR to an instance of LTHS.

7. LONG-TERM HYDROTHERMAL SCHEDULING

We implemented the mean-risk MSD algorithms and applied it to the long-term hydrothermal scheduling (LTHS). The aim of LTHS is to find an optimal policy for scheduling the operation of hydro and thermal power systems over a multi-period planning horizon so that the expected cost of power generation is minimized. LTHS is one of the widely studied application of MSLP. This problem was first introduced by Sherkat et al. [10], who applied stochastic dynamic programming algorithm to LTHS. Two sampling strategies namely Latin hypercube sampling and randomized quasi-Monte Carlo sampling were applied for the generation of scenario trees for LTHS problem in [4] and a stopping criteria based on statistical hypothesis test was applied. A tutorial on stochastic programming using LTHS as an example along with the description of the stochastic optimization model, scenario tree generation and application of a nested decomposition algorithm for optimizing LTHS problem was provided in [11]. A risk-averse MSLP formulation with coherent risk measure *conditional value-at-risk* was used for optimizing LTHS problem in [12]. In this paper, we use the hydrothermal model described in [4, 11] and we use data for LTHS from [51].

7.1 LTHS Problem Formulation

The risk-neutral formulation of LTHS model at stage (month) t for scenario $\omega \in \Omega$ and decision vector x_{t-1} from stage $t - 1$, is given as follows:

$$f_t(x_{t-1}, \omega_{[t]}) := \text{Min} \quad \sum_{i \in I} c_i^\top p_{ti} + c^\top l_t + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})] \quad (7.1a)$$

$$\text{s.t.} \quad \sum_{i \in I} p_{ti} + l_t + \gamma q_t = d_t \quad (7.1b)$$

$$v_t + a(q_t + s_t) = v_{t-1} + ay_t(\omega_t) \quad (7.1c)$$

$$v_t \leq v_{\max}, q_t \leq q_{\max}, p_{ti} \leq p_{i\max}, \quad \forall i \in I \quad (7.1d)$$

$$v_t, q_t, s_t, p_{ti}, d_t \geq 0, \quad \forall i \in I,$$

where, I is the set of thermal power plants, c_i is incremental unit cost of thermal power generation in dollars from thermal power plant $i \in I$ and p_{ti} is the power generated from thermal power plant $i \in I$ at stage t . The cost associated with load shedding is denoted by c and decision variable l_t denotes amount of load shedding at stage t . The objective function (7.1a) gives the cost associated with power generation at stage t along with the expected future cost denoted by $\mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})]$.

The first constraint (7.1b), satisfies the power demand d_t at stage t with power generation from thermal power plants p_{ti} , hydro power plant γq_t (where γ is coefficient of hydro plant productivity and q_t is the volume of turbine outflow during stage t) and the load shedding l_t at stage t . The second constraint (7.1c) ensures the continuity of amount of water available in hydro power plant. The amount of water available at stage t and $t - 1$ is denoted by v_t and v_{t-1} , respectively and v_{\max} denotes the maximum storage capability of the reservoir. The constant a is used to convert water flow into an equivalent volume in a month, s_t denotes hydro plant spillage at stage t and $y_t(\omega)$ denotes random incremental flow realized in month t . The third set of constraints (7.1d) limits the amount of water available, volume of turbine outflow and power generation from thermal power plants based on their maximum capacity. Next we present risk-averse formulations of the LTHS problem with risk measures EE, QDEV and CVaR, respectively.

7.1.1 LTHS Formulation with EE

Given $\lambda_t \geq 0$ and a target $\psi_t \in \mathbb{R}$, the LTHS formulation at stage t for realization $\omega \in \Omega$ and decision vector x_{t-1} at month $t - 1$ with $\rho = \phi_{EE, \psi_t}$ is equivalent to the following formulation:

$$f_t(x_{t-1}, \omega_{[t]}) := \text{Min} \quad \sum_{i \in I} c_i^\top p_{ti} + c^\top l_t + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})] + \lambda_t \nu_t \quad (7.2)$$

$$\text{s.t.} \quad \sum_{i \in I} p_{ti} + l_t + \gamma q_t = d_t$$

$$v_t + a(q_t + s_t) = v_{t-1} + ay_t(\omega_t)$$

$$- \sum_{i \in I} c_i^\top p_{ti} - c^\top l_t - \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})] + \nu_t$$

$$\geq -\psi_t + \sum_{i \in I} c_i^\top p_{t-1,i} + c^\top l_{t-1} \quad (7.3)$$

$$v_t \leq v_{\max}, q_t \leq q_{\max}, p_{ti} \leq p_{i\max}, \quad \forall i \in I$$

$$v_t, q_t, s_t, p_{ti}, d_t \geq 0, \quad \forall i \in I,$$

where constraint (7.3) computes the excess cost of power generation for realization ω over target ψ_t .

7.1.2 LTHS Formulation with QDEV

Given $\lambda_t \geq 0$ and a quantile $\alpha \in (0, 1)$, the LTHS formulation at stage t for realization $\omega \in \Omega$ and decision vector x_{t-1} and quantile ψ_{t-1} at month $t-1$ with $\rho = \phi_{\text{QDEV}_{\psi_t}}$ is equivalent to the following formulation:

$$f_t(x_{t-1}, \omega_{[t]}) := \text{Min} \quad (1 - \lambda_t \epsilon_1) \left[(1 - \lambda_{t+1} \epsilon_1) \left(\sum_{i \in I} c_i^\top p_{ti} + c^\top l_t \right) + \lambda_{t+1} \epsilon_1 \psi_t \right. \\ \left. + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})] \right] + \lambda_t (\epsilon_1 + \epsilon_2) \nu_t \quad (7.4)$$

$$\text{s.t.} \quad \sum_{i \in I} p_{ti} + l_t + \gamma q_t = d_t \\ v_t + a(q_t + s_t) = v_{t-1} + ay_t(\omega_t) \\ - (1 - \lambda_{t+1} \epsilon_1) \left(\sum_{i \in I} c_i^\top p_{ti} + c^\top l_t \right) - \lambda_{t+1} \epsilon_1 \psi_t - \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})] \\ + \nu_t \geq -\psi_{t+1} + \sum_{i \in I} c_i^\top p_{t-1,i} + c^\top l_{t-1} \quad (7.5)$$

$$v_t \leq v_{\max}, q_t \leq q_{\max}, p_{ti} \leq p_{i\max}, \quad \forall i \in I$$

$$v_t, q_t, s_t, p_{ti}, d_t \geq 0, \quad \forall i \in I,$$

where constraint (7.5) computes the excess cost of power generation for realization ω over quantile ψ_{t-1} .

7.1.3 LTHS Formulation with EECVaR

Given $\lambda \geq 0$ and $\alpha \in (0, 1)$, the LTHS formulation at stage t for realization $\omega \in \Omega$, and decision vector x_{t-1} and quantile ψ_{t-1} at month $t-1$ with $\rho = \phi_{\text{CVaR}_\alpha}$ is equivalent to the following formulation:

$$f_t(x_{t-1}, \psi_{t-1}, \omega) := \text{Min} \quad \sum_{i \in I} c_i^\top p_{ti} + c^\top l_t + \lambda_{t+1} \psi_t + \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})] + \frac{\lambda_t}{1-\alpha} \nu_t \quad (7.6)$$

$$\text{s.t.} \quad \sum_{i \in I} p_{ti} + l_t + \gamma q_t = d_t,$$

$$v_t + a(q_t + s_t) = v_{t-1} + ay_t(\omega_t),$$

$$- \sum_{i \in I} c_i^\top p_{ti} - c^\top l_t - \lambda_{t+1} \psi_t - \mathbb{E}[f_{t+1}(x_t, \tilde{\omega}_{[t]})] + \nu_t$$

$$\geq -\psi_{t-1} + \sum_{i \in I} c_i^\top p_{t-1,i} + c^\top l_{t-1}, \quad (7.7)$$

$$v_t \leq v_{\max}, q_t \leq q_{\max}, p_{ti} \leq p_{i\max}, \quad \forall i \in I$$

$$v_t, q_t, s_t, p_{ti}, d_t \geq 0, \quad \forall i \in I,$$

where constraint (7.7) computes the excess cost of power generation for realization ω over quantile ψ_{t-1} .

7.2 Computational Results

We considered LTHS involving two thermal power generators p_1 and p_2 , one hydro power plant q and 60 stages, where each stage represents a month. The maximum power generation capacity of the thermal power plants p_1 and p_2 is 65 MWh and 50 MWh, respectively. The cost of power generation from each thermal power plant is 20 \$/MWh and 100 \$/MWh, respectively. The maximum power generation capacity γq of hydro power plant is 100MWh, while the maximum energy storage capacity of the reservoir is 150MWh. The power generation from the hydro power plant is essentially free. Load shedding occurs when there is not enough energy generation from thermal and hydro power plants to satisfy the demand. This incurs a high cost of \$1000/MWh to satisfy the unmet demand.

The amount of water inflow per month shown in Figure 7.1, is a random variable with three observations representing a high, medium and low water inflow. This results in total of 3^{60} possible scenarios. Note that the scenario tree is stage-wise independent, that is, realization of the random

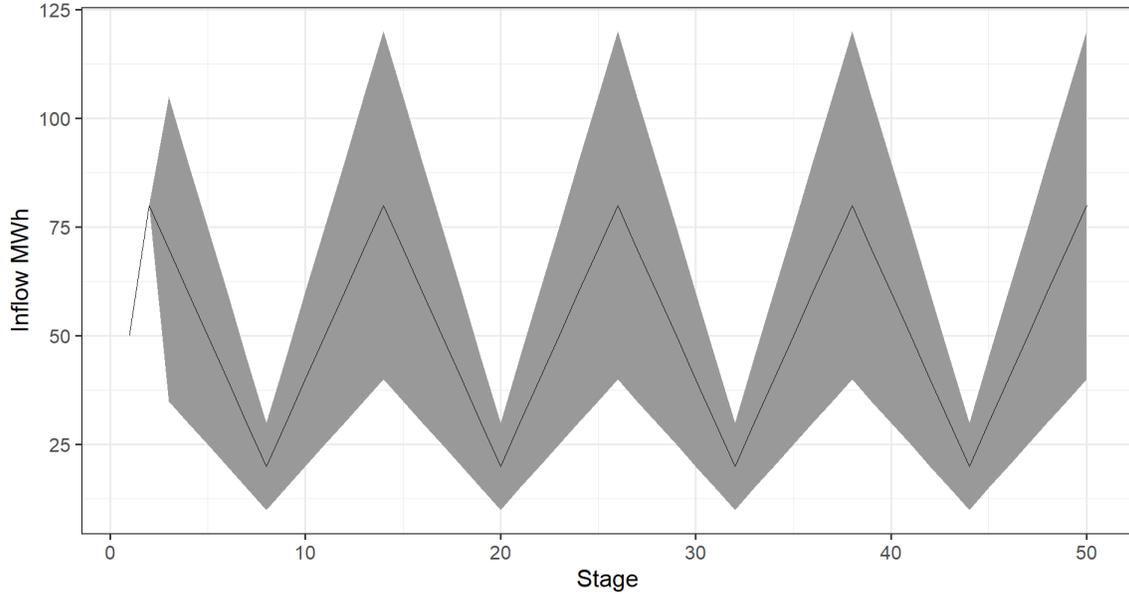


Figure 7.1: Inflow per Stage

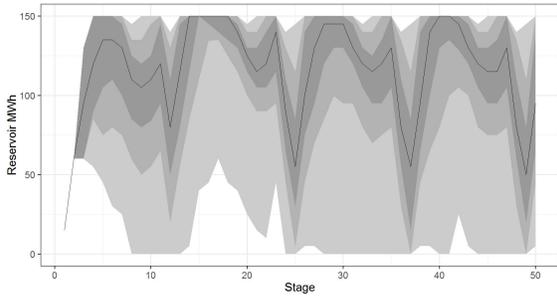
variable at stage t is independent of realization at stage $t - 1$. The demand is constant for each stage and is equal to 100MWh.

7.2.1 Risk-Neutral MSLP

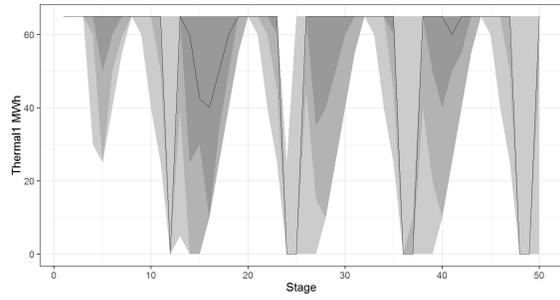
Now we present the results of risk-neutral MSLP for the LTHS instance solved using the above specified data. As mentioned, the instance is solved for 60 stages, but the results are reported for 50 stages. This is done to mitigate the impact of the end effects.

In Figure 7.2, the results for risk-neutral case are presented using a variety of plots. Each plot represents the results of 5000 randomly selected scenarios. The solid line in each graph represents the median. To demonstrate the spread of results, we have used different shades to distinguish between quantiles. The three shades from lightest to the darkest represent 0-1, 0.1-0.9 and 0.25-0.75 quantiles, respectively.

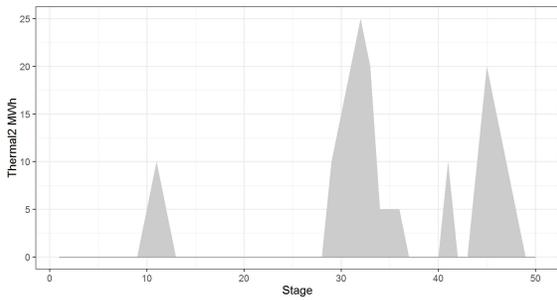
Figure 7.2a displays amount of energy storage (MWh) in the reservoir after the power generation at each stage, for the risk-neutral case. From the figure it is evident that the median reservoir level stays fairly high throughout all the stages and only in few instances the reservoir level touches the nadir. Figures 7.2b, 7.2c and 7.2d show amount of power generated in MWh from power plant



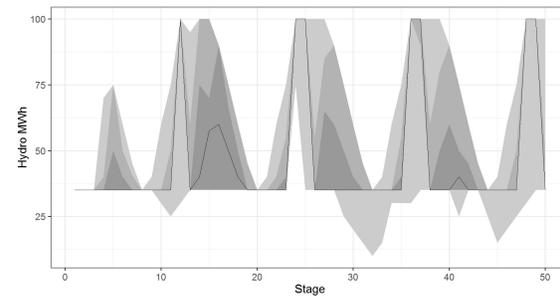
(a) Reservoir per stage



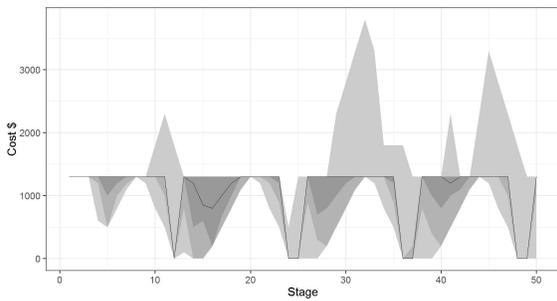
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage



(d) Power generation γq per stage



(e) Cost of operation per stage

Figure 7.2: Results for Risk-Neutral MSLP for LTHS

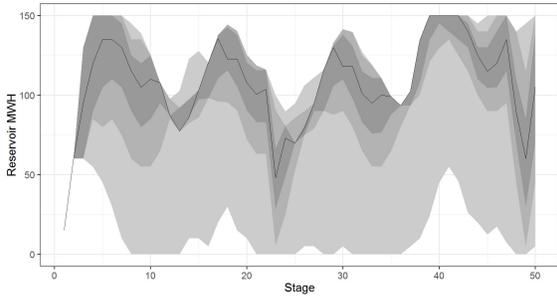
p_1 , p_2 and q , respectively. As can be observed from figure 7.2b and 7.2d, the power generation from plants p_1 and q complement each other and they seem to follow the seasonal variation in the water availability. Power generation from the plant q stays above 35MWh for most of the stages and reaches 100MWh for the stages where reservoir is at the full capacity. Power generation from the plant p_2 occurs only during few rare occasions and it corresponds to the stages where the reservoir is empty. Finally, 7.2e represents the cost incurred in each stage and the cost pattern mostly mirrors the power generation pattern of plant p_1 . The median cost stays at or below \$1300 for all the stages.

7.2.2 MR-MSLP with EE

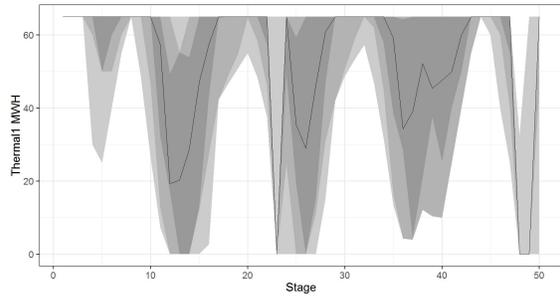
Figure 7.3 displays the results for the LHTS instance with EE and $\lambda = 0.1$. The selection of target in case of MSLP with nested formulation is not straight forward, as it is in the two-stage case. Since, the target at a given stage has to account for the total future cost and also for the risk associated with all the future stages. Hence, to set the target at each stage we computed the 90th percentile of cost C_t at each stage for the risk neutral case. Then we computed the average difference D between the maximum cost and the 90th percentile of cost for all stages. Using these two values the target ψ_t at each stage was determined as follows:

$$\psi_t = C_{t-1} + (D * (61 - t)/2) \quad (7.8)$$

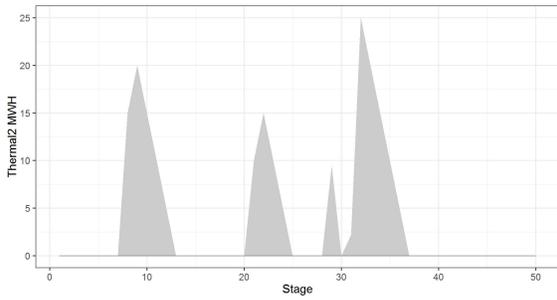
Figure 7.3a displays amount of water available in the reservoir at each stage after power generation with EE as risk measure. It can be observed that the median of energy stored in the reservoir stays fairly high, but the reservoir level is rarely full or empty. As can be observed from figure 7.3b and 7.3d, the power generation from plants p_1 and q complement each other at a given stage and they have a inverse relationship. In case of plant q the power generation deeps below 35MWh for some stages signifying unavailability of water. The power generation from plant p_2 shown in figure 7.3c corresponds with the low inflow of water. Finally, figure 7.3e represents the cost incurred in each stage, and it is evident that for certain stages the cost is higher than \$1300 on account of



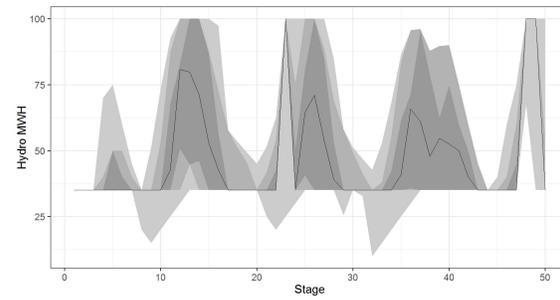
(a) Reservoir per stage



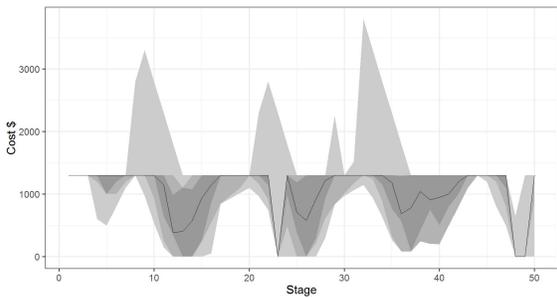
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage



(d) Power generation γq per stage



(e) Cost of operation per stage

Figure 7.3: Results for MR-MSLP with EE for LTHS ($\lambda = 0.1$)

power generation from the plant p_2 . Figures C.1, C.2 and C.3 represent results for MSD-EE with λ value of 0.25, 0.5 and 1 respectively and are included in Appendix C.

7.2.3 MR-MSLP with QDEV

The results for the LHTS instance with QDEV, are displayed in Figure 7.4. The risk trade-off parameter used for all the stages was $\lambda_t = 0.1$ and the α -quantile was set to 0.95. Figure 7.4a displays amount of water available in the reservoir at each stage after power generation and in contrast to the risk-neutral case the median water level stays close to zero for most of the initial stages. In case of QDEV, it appears that the model stresses on using as much water as available in a given stage. Figures C.4, C.5 and C.6 from Appendix C display the results for MSD-QDEV with trade-off factor equal to 0.25, 0.5 and 1, respectively.

From figure 7.4b, 7.4c and 7.4d, one can observe that the power generation from plant p_1 , p_2 and q follow a cyclic pattern which reflects the cyclic availability of water. Also, it is quite evident that the expensive thermal powerplant p_2 is regularly used for power generation. This is likely the result of policy of using all the available water at a given stage and hence during the times of water scarcity the demand can be met only through power plant p_2 . Finally, 7.4e represents the cost incurred in each stage, and it is evident that there is lot more variation from stage to stage in terms of cost incurred. The cost at stages associated with water scarcity often crosses the threshold of \$1300, whereas for the stages with plenty of water availability the cost touches zero mark. Although, the median cost at any stage stays below \$1300. The results reflect the nature of QDEV, that is QDEV captures the two-sided deviation from the α -quantile and therefore it tries to reduce overall variance at each stage at the expense of expected cost.

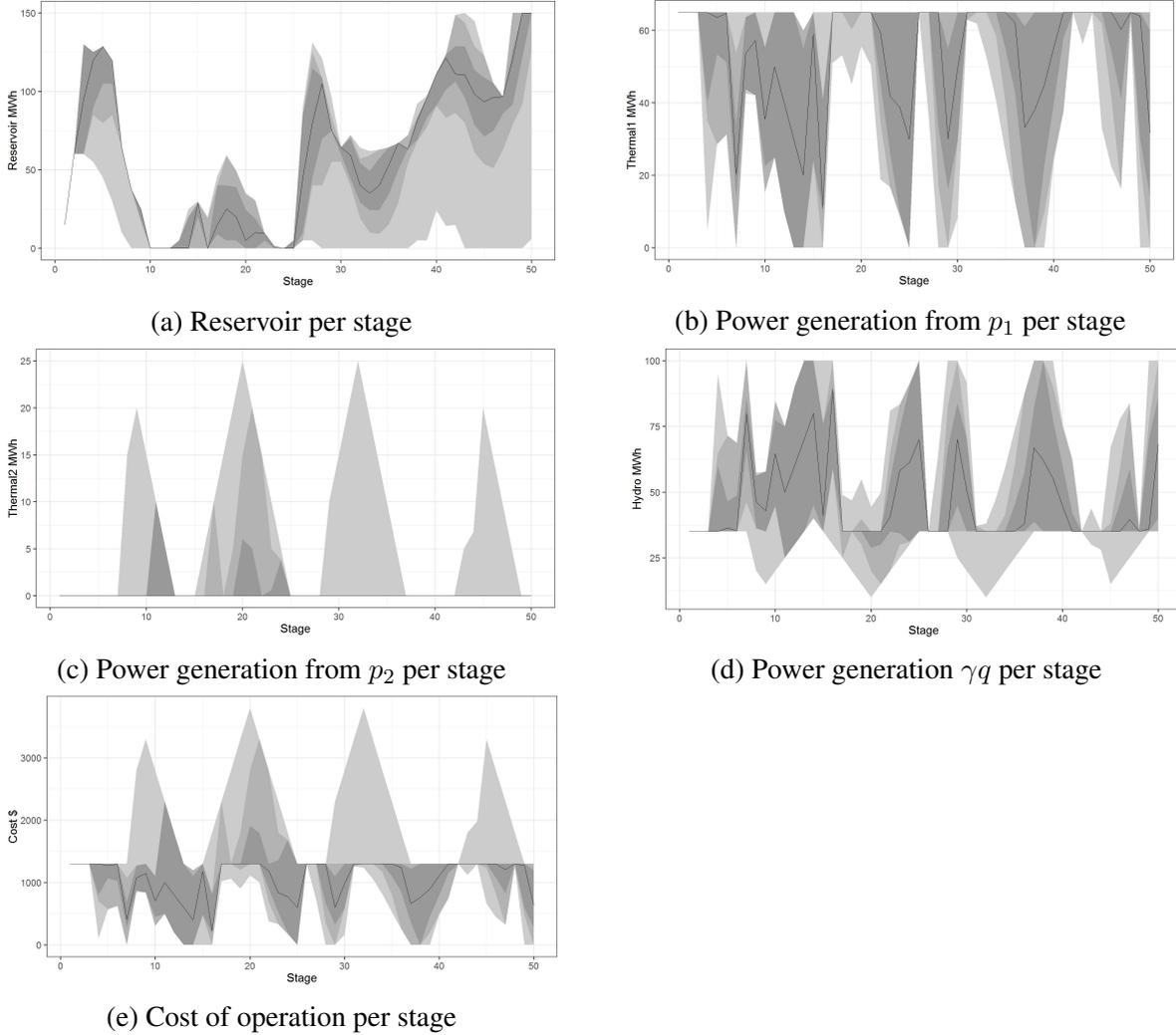


Figure 7.4: Results for MR-MSLP with QDEV for LTHS ($\lambda = 0.1$)

7.2.4 MR-MSLP with CVaR

Figure 7.5 displays the results for the LTHS instance with CVaR. For all the stages the risk trade-off parameter $\lambda = 0.1$ was used and the α -quantile was set to 0.95. The water availability at each stage after the power generation is shown in 7.5a. In case of CVaR, the model appears to conserve water at each stage and keep the level of stored energy as close to the maximum reservoir level as possible.

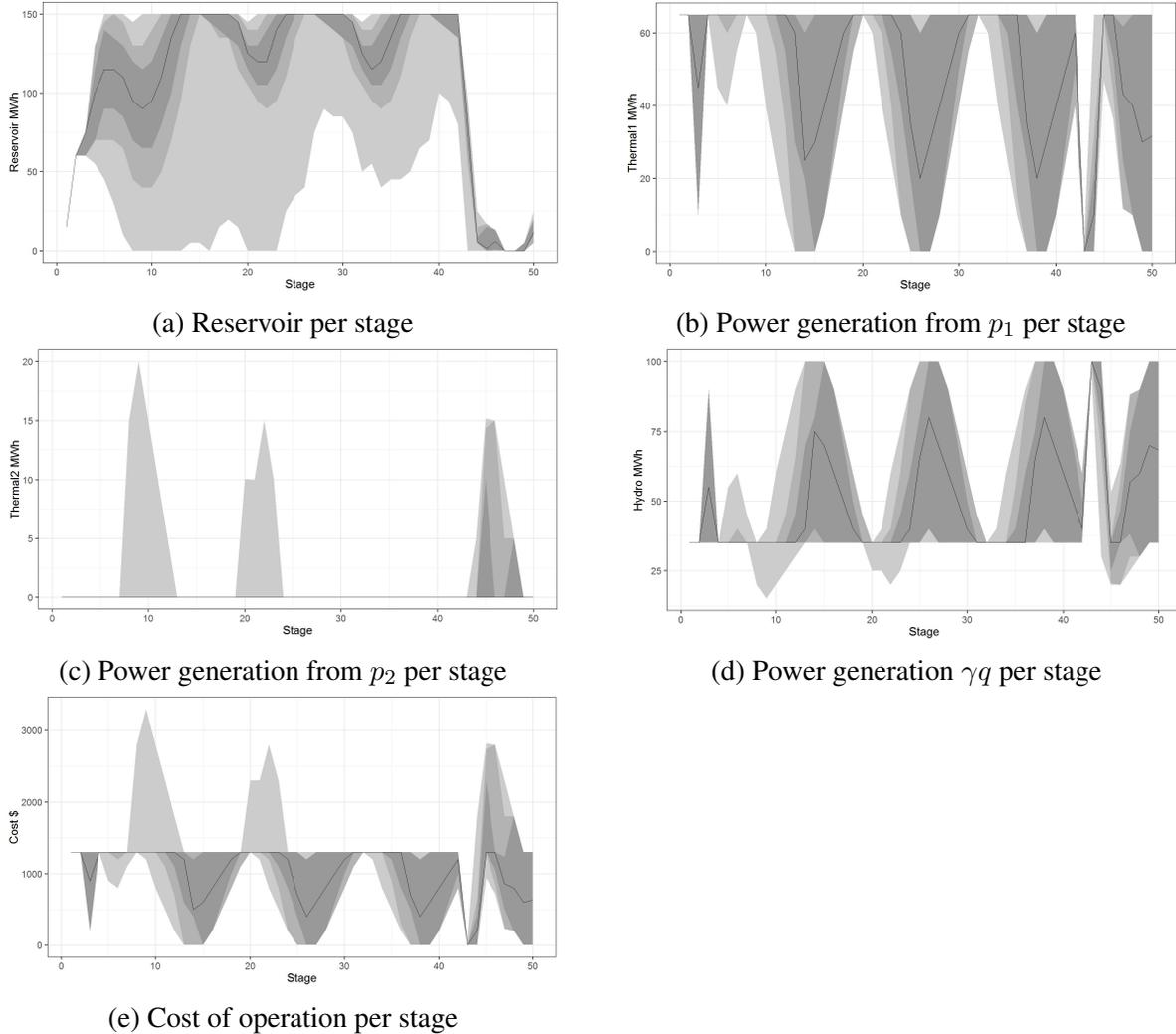


Figure 7.5: Results for MR-MSLP with CVaR for LTHS ($\lambda = 0.1$)

From figure 7.5b and 7.5d, it is evident that the power generation from plants p_1 and q complement each other and they reflect the cyclic availability of water. The power generation from the hydro power plant q rarely dips below 35 MWh, which is an indication of enough availability of water in the reservoir. The use of thermal powerplant p_2 for power generation as shown in figure 7.5c, is very rare in initial stages and is often engaged after 45th stage. Figure 7.5e represents the cost incurred in each stage and it can be inferred that the cost very rarely exceeds \$1300 except in case of the terminal stages. Figures C.7, C.8 and C.9 represent results for MSD-CVaR with risk trade-off factor equal to 0.25, 0.5 and 1 respectively and are included in Appendix C.

7.2.5 Discussion

Figures 7.2a, 7.3a, 7.4a, and 7.5a, display energy storage (MWh) of the reservoir at each stage after power generation for the risk-neutral, EE, QDEV and CVaR case respectively. In case of risk-neutral and CVaR, the energy storage level in the reservoir is maintained as high as possible, that is the stress of the model is on water conservation for future use. Especially the storage level for CVaR is highest compared to all other risk measures and even compared to the risk-neutral case. In contrast to the risk-neutral approach, the EE and QDEV cases stress on consumption of water for power generation. The QDEV case has the worst energy storage level compared to any other risk measures and its median storage level stays close to zero for most of the stages.

The power generation from plant p_2 for risk-neutral, EE, QDEV and CVaR is shown in figures 7.2c, 7.3c, 7.4c, and 7.5c, respectively. Since, the cost of power generation from p_2 is \$100/MW, the use of this power plant significantly increases the cost associated with satisfying the power demand. In the risk-neutral and EE cases, power plant p_2 is rarely used for power generation, since energy storage in the reservoir is enough to satisfy the demand that cannot be met through power plant p_1 alone. Whereas in case of QDEV and CVaR, power plant p_2 is used very often and coincides with the stages associated with water shortages. The policy of water consumption rather than conservation in case of QDEV often leads to scenarios with empty storage and hence the unmet power demand has to be met through power plant p_2 .

Table 7.1: Average Cost of Operation per Stage

λ	EE	QDEV	CVaR
0.1	\$1002.52	\$1036.16	\$1062.06
0.25	\$1007.32	\$1155.72	\$1268.95
0.5	\$1027.72	\$1207.34	\$1313.47
1	\$1024.67	\$1264.81	\$1361.18

Finally, figures 7.2e, 7.3e, 7.4e and 7.5e represent the cost incurred in each stage for risk-neutral, EE, QDEV and CVaR, respectively. The median cost at any given stage never crosses threshold of \$1300 for risk measure EE, QDEV and CVaR. Though the instances of cost exceeding \$1300 is more common in QDEV and CVaR. The average cost per stage for risk-neutral stage was found to be \$1001.58, which was smaller than all other three cases. For $\lambda = 0.1$ the average cost per stage increased slightly to \$1002.52 in case of EE and significantly increased to \$1036.16 and \$1062.06 in case of QDEV and CVaR respectively. Table 7.1, summarizes the average cost of operation for various combination of risk measures and risk trade-off factors.

7.2.6 Impact of Risk Measures on Extreme Scenarios

The use of MR-MSLP for modeling sequential decision making problems is justified by the fact that risk measures provide a suitable mechanism for handling losses, controlling variability and hedging against extreme scenarios. Therefore, it was necessary to gauge the impact of risk measures on such extreme scenarios. Hence, for each risk measure and the risk neutral case, we collected the top 1% most expensive scenarios (that is the scenarios with average cost greater than 0.99 percentile).

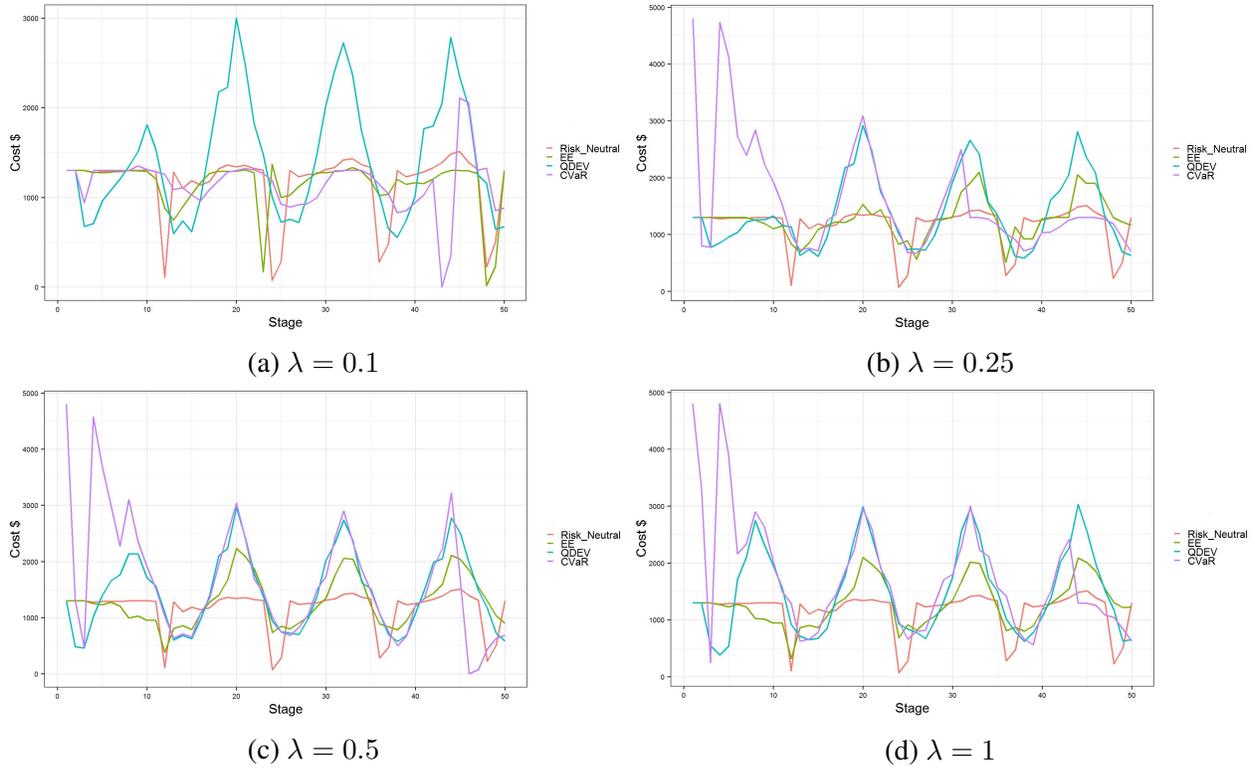


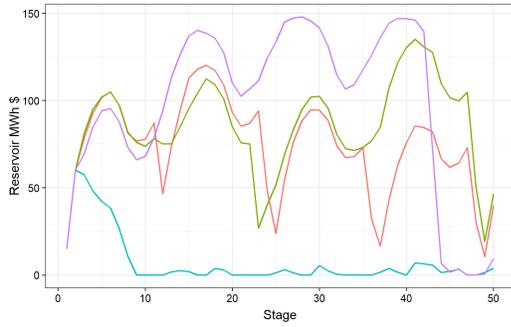
Figure 7.6: Cost of Operation per Stage

Figure 7.6, displays the average cost at each stage for the extreme scenarios associated with risk neutral, EE, QDEV and CVaR case. As it can be observed, at higher values of trade-off factor risk measure QDEV and CVaR offer no relief from the cost fluctuations. It is at the lower value of $\lambda = 0.1$ that CVaR exhibits much more control over the cost fluctuation. The average cost of this extreme scenarios for risk neutral case was found to be \$1161.23 and only MSD-EE at $\lambda = 0.1$ outperformed the risk neutral case. For various values of λ the average cost of extreme scenarios per stage for each risk measure is shown in Table 7.2.

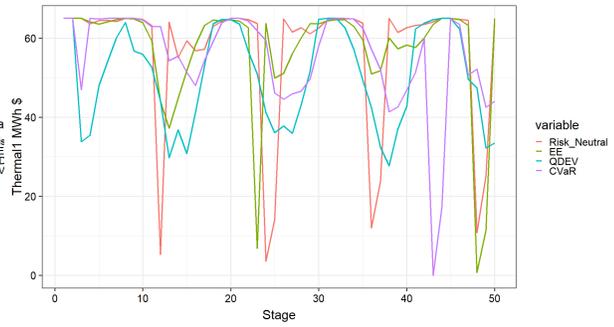
Table 7.2: Average Cost per Stage for Extreme Scenarios

λ	EE	QDEV	CVaR
0.1	\$1158.21	\$1234.02	\$1236.81
0.25	\$1178.30	\$1420.27	\$1501.17
0.5	\$1236.46	\$1478.41	\$1582.25
1	\$1280.01	\$1523.58	\$1589.98

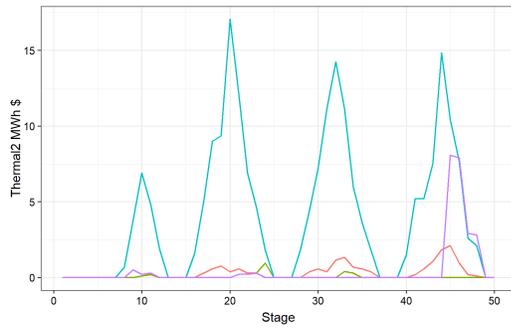
Figure 7.7 displays the impact different risk measures have on the individual decision variable associated with the extreme scenarios. Figure 7.7a represents the storage level of reservoir at different stages for each risk measure ($\lambda = 0.1$) and the risk neutral case. It is evident from the figure that in case of CVaR the emphasis is on conserving the water. In contrast, use of risk measure QDEV results in occurrence of empty reservoir.



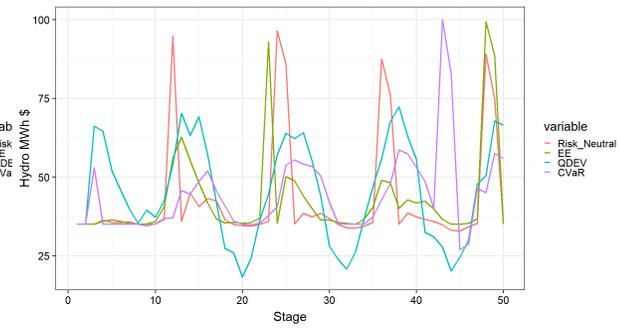
(a) Reservoir per stage



(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage



(d) Power generation γq per stage

Figure 7.7: Impact of Risk Measures on Extreme Scenarios ($\lambda = 0.1$)

8. CONCLUSIONS AND FUTURE RESEARCH

8.1 Conclusions

Risk-averse stochastic programming goes beyond the prevalent expected value framework and aims at controlling variability of cost associated with different outcomes. We began this dissertation by introduction to risk-averse stochastic programming and followed it by presenting an overview of literature on stochastic programming, risk-aversion and computational difficulties associated with solving large scale SLP problems. We then introduced the readers to the definition of risk measures and their formulations. Next we developed and implemented stochastic decomposition algorithms for solving MR-SLP and MR-MSLP with three different risk measures: expected excess (EE), quantile deviation (QDEV) and conditional value-at-risk (CVaR). We have also provided proofs of convergence for these algorithms and results of computational experiments.

Our numerical results for MR-SLP show that even for the large-scale instances such as *Lands* and *storm*, the stochastic decomposition (SD) algorithm generally requires a relatively small number of scenarios to converge to an optimal solution. This is possible due to the interior sampling approach of generating one scenario per iteration used in SD algorithm; if we were to consider all possible scenarios it is usually very computationally demanding to solve large-scale instances to optimality. Moreover, unlike SAA, in SD we do not need to specify the number of scenarios to sample for a given instance. We can set a desired tolerance level for the termination criteria and the algorithm will only sample required number of scenarios to satisfy the criteria. This feature guarantees a desired level of accuracy in solution and at the same time offers a competitive computational time, which is an advantage over the SAA approach.

In case of MR-SLP, the computational study also provides several insights useful for understanding the effect of the risk measure and the risk trade-off factor on expected cost and variability in two-stage setting. We show that the QDEV has more impact on expected cost and the cost associated with extreme scenarios compared to the impact of CVaR and EE. We also observed that for

higher target values, the risk measure EE becomes effective only for a relatively small number of scenarios and has little to no-effect on the solution for small values of the trade-off factor.

Like its two-stage counterpart, risk-averse MSD offers several advantages over the traditional algorithms. First and foremost, the combinatorial explosion of MSLP resulting from large number of stages and scenarios is avoided by only solving the LPs associated with the sample path. Also for implementing MSD, the entire scenario tree need not be revealed at the beginning of the algorithm because at every iteration the subgradients are not computed using every potential outcome.

The application of mean-risk MSD to an instance of long-term hydrothermal scheduling helped us in better understanding of risk-aversion in multistage settings. We observed that the trade-off actor λ has an outsize impact on the optimal solution and is the key in achieving desirable results. At $\lambda = 0.1$, CVaR exhibited better control over the extreme scenarios compared to EE and QDEV. Also the experiment revealed the policy adopted by each risk measure in controlling the variability. In case of CVaR, the emphasis was on conserving water as much as possible for the future use and hence as the value of risk-trade of factor increased so did the reservoir level. In contrast, for risk measure QDEV the optimal policy involved maximum use of available water at each stage and at higher values of λ the reservoir level stayed close to zero. Among the three risk measures EE offered the best performance in terms of average cost and hedging against the extreme scenarios, but this is contingent upon selection of appropriate target values.

8.2 Future Research

Unlike the two-stage case where risk aversion is well understood and widely applied, there is no obvious way of formulating risk measures in the multistage settings. Hence, different formulations of risk measures could be tested for various applications to determine their appropriateness. Scope also exist to define new risk measures more suitable for the multistage settings.

In future multistage stochastic decomposition algorithm can be extended to expected conditional risk measures (ECRM), and then the results can be compared to the nested formulations. Modifications to MSD can be made to further improve its convergence rate and to develop a more sophisticated stopping criterion. Also, a detailed computational study using various instances could

be conducted to assess the performance of MSD algorithm with respect to SDDP.

Future work could also include deriving and implementing stochastic decomposition algorithms for other risk measures such as *excess probability* and *absolute semi-deviation*. In case of *excess probability* the problem becomes a mix-integer program and in case of *absolute semi-deviation* the use of expected mean to compute excess makes the problem averse to benders decomposition.

REFERENCES

- [1] D. Carino and W. Ziemba, “Formulation of the Russell-Yasuda Kasai financial planning mode.,” *Operations Research*, vol. 46.4, pp. 443–449, 1998.
- [2] M. Pereira and L. Pinto, “Multi-stage stochastic optimization applied to energy planning.,” *Mathematical Programming*, vol. 52.1-3, pp. 359–375, 1991.
- [3] R. Rockafellar and J. Wets, “Scenarios and policy aggregation in optimization under uncertainty.,” *Mathematics of Operations Research*, vol. 16.1, pp. 119–147, 1991.
- [4] T. H. de Mello, V. D. Matos, and E. Finardi, “Sampling strategies and stopping criteria for stochastic dual dynamic programming: a case study in long-term hydrothermal scheduling.,” *Energy Systems*, vol. 2.1, pp. 1–31, 2011.
- [5] S. Sen and J. Higle, “Finite master programs in regularized stochastic decomposition.,” *Mathematical Programming*, vol. 67.1-3, pp. 143–168, 1994.
- [6] J. Higle and S. Sen, “Stochastic decomposition: An algorithm for two-stage linear programs with recourse,” *Mathematics of Operations Research*, vol. 16.3, pp. 650–669, 1991.
- [7] A. Märkert and R. Schultz, “On deviation measures in stochastic integer programming,” *Operations Research Letters*, vol. 33, no. 5, pp. 441 – 449, 2005.
- [8] W. Ogryczak and A. Ruszczyński, “Dual stochastic dominance and related mean-risk models,” *SIAM Journal on Optimization*, vol. 13, no. 1, pp. 60–78, 2002.
- [9] T. Rockafellar and S. Uryasev, “Optimization of conditional value-at-risk,” *Journal of Risk*, vol. 2, pp. 21–42, 2000.
- [10] V. Sherkat, R. Campo, K. Moleshi, and E. Lo, “Stochastic long-term hydrothermal optimization for a multireservoir system.,” *IEEE Transactions on Power Apparatus and Systems*, vol. 8, pp. 2040–2050, 1985.

- [11] B. Decker, V. D. Matos, and E. Finardi, “An introductory tutorial on stochastic programming using a long-term hydrothermal scheduling problem.,” *Journal of Control, Automation and Electrical Systems*, vol. 24.3, pp. 361–376, 2013.
- [12] A. Philpott and V. D. Matos, “Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion.,” *European Journal of Operational Research*, vol. 218.2, pp. 470–483, 2012.
- [13] R. Schultz, “Risk aversion in two-stage stochastic integer programming,” in *International Series in Operations Research & Management Science* (G. Infanger, ed.), vol. 150, pp. 165–187, New York: Springer Science+Business Media, 2011.
- [14] W. Ogryczak and A. Ruszczyński, “Dual stochastic dominance and related mean-risk models,” *SIAM Journal on Optimization*, vol. 13, pp. 60–78, 2002.
- [15] P. Jorion, “Risk2: Measuring the risk in value at risk,” *Financial Analysts Journal*, vol. 52.6, pp. 47–56, 1996.
- [16] M. Pritsker, “Evaluating value at risk methodologies: accuracy versus computational time,” *Journal of Financial Services Research*, vol. 12.2, pp. 201–242, 1997.
- [17] P. Artzner, F. Delbean, J. Eber, and D. Heath, “Coherent measures of risk,” *Mathematical Finance*, vol. 9, pp. 203–228, 1999.
- [18] S. Saryklain, G. Serraino, and S. Uryasev, “Value-at-risk vs. conditional value-at-risk in risk management and optimization,” *INFORMS*, pp. 270–294, 2008.
- [19] R. Schultz and S. Tiedemann, “Risk aversion via excess probabilities in stochastic programs with mixed-integer recourse,” *SIAM Journal on Optimization*, vol. 14, no. 1, pp. 115–138, 2003.
- [20] T. Kristoffersen, “Deviation measures in linear two-stage stochastic programming,” *Mathematical Methods of Operations Research*, vol. 62, no. 2, pp. 255–274, 2005.

- [21] M. Riis and R. Schultz, “Applying the minimum risk criterion in stochastic recourse programs,” *Computational Optimization and Applications*, vol. 24, no. 2-3, pp. 267–287, 2003.
- [22] A. Shapiro and T. H. de Mello, “A simulation-based approach to two-stage stochastic programming with recourse,” *Mathematical Programming*, vol. 81.3, pp. 301–325, 1998.
- [23] E. Plambeck, B. Fu, S. Robinson, and R. Suri, “Sample-path optimization of convex stochastic performance functions,” *Mathematical Programming*, vol. 75.2, pp. 137–176, 1996.
- [24] S. Ahmed, “Convexity and decomposition of mean-risk stochastic programs,” *Mathematical Programming*, vol. 106, no. 3, pp. 433–446, 2006.
- [25] T. Cotton and L. Ntaimo, “Computational study of decomposition algorithms for mean-risk stochastic linear programs,” *Mathematical Programming Computation*, vol. 7.4, pp. 471–499, 2015.
- [26] T. Homem-de Mello and G. Bayraksan, “Monte Carlo sampling-based methods for stochastic optimization,” *Surveys in Operations Research and Management Science*, vol. 19, no. 1, pp. 56–85, 2014.
- [27] J. Higle and S. Sen, *Stochastic Decomposition*. 101 Phillip Drive, Norwell, MA 02061: Kluwer Academic Publishers, 1996.
- [28] P. Krokmal and M. Z. nad S. Uryasev, “Modeling and optimization of risk.,” *Handbook of the Fundamentals of Financial Decision Making: Part II*, pp. 555–600, 2013.
- [29] J. Birge, “Decomposition and partitioning methods for multistage stochastic linear programs.,” *Operations Research*, vol. 33.5, pp. 989–1007, 1985.
- [30] S. Sen and Z. Zhou, “Multistage stochastic decomposition: a bridge between stochastic programming and approximate dynamic programming.,” *SIAM Journal on Optimization*, vol. 24.1, pp. 127–153, 2014.
- [31] V. Kozmík and D. Morton, “Evaluating policies in risk-averse multi-stage stochastic programming.,” *Mathematical Programming*, vol. 152.1-2, pp. 275–300, 2015.

- [32] C. Carøe and J. Tind, “A cutting-plane approach to mixed 0-1 stochastic integer programs,” *European Journal of Operational Research*, vol. 101, pp. 306–316, 1997.
- [33] J. Bion-Nadal, “Dynamic risk measures: Time consistency and risk measures from bmo martingales,” *Finance and Stochastics*, vol. 12, pp. 219–244, 2008.
- [34] A. Ruszczyński and A. Shapiro, “Optimization of convex risk functions,” *Mathematics of Operations Research*, vol. 31.3, pp. 433–452, 2006.
- [35] S. Weber, “Distribution-invariant dynamic risk measures, information and dynamic consistency,” *Mathematical Finance*, vol. 16, pp. 419–441, 2006.
- [36] A. Shapiro, “On a time consistency concept in risk averse multistage stochastic programming,” *Operations Research*, vol. 37.3, pp. 143–147, 2009.
- [37] A. Shapiro, “Minimax and risk averse multistage stochastic programming,” *European Journal of Operational Research*, vol. 219.3, pp. 719–726, 2012.
- [38] R. Collado, D. Papp, and A. Ruszczyński, “Scenario decomposition of risk-averse multistage stochastic programming problems,” *Annals of Operations Research*, vol. 200.1, pp. 147–170, 2012.
- [39] S. LeRoy, “Risk aversion and the martingale property of stock prices,” *International Economic Review*, pp. 436–446, 1973.
- [40] R. Chou, R. Engle, and A. Kane, “Measuring risk aversion from excess returns on a stock index,” *Journal of Econometrics*, vol. 52.1-2, pp. 201–224, 1992.
- [41] A. Ruszczyński and A. Shapiro, “Optimization of convex risk functions,” *Mathematics of Operations Research*, vol. 31, no. 3, pp. 433–452, 2006.
- [42] P. Krokmal, J. Palmquist, and S. Uryasev, “Portfolio optimization with conditional value-at-risk objective and constraints,” *Journal of Risk*, vol. 4, pp. 43–68, 2002.
- [43] S. Zhu and M. Fukushima, “Worst-case conditional value-at-risk with application to robust portfolio management,” *Operations Research*, vol. 57.5, pp. 1155–1168, 2009.

- [44] J. Benders, “Partitioning procedures for solving mixed-variable programming problems,” *Numerische Mathematik*, vol. 4, pp. 238–252, 1962.
- [45] I. I. CPLEX, *User’s Manual for CPLEX*, vol. V12.0. 2012.
- [46] J. Linderoth, A. Shapiro, and S. Wright, “The empirical behavior of sampling methods for stochastic programming,” *Annals of Operation Research*, vol. 142, pp. 215–241, 2006.
- [47] G. Dantzig, *Linear Programming Extensions*. New Jersey: Princeton University Press, 1963.
- [48] F. Louveaux and Y. Smeers, *Optimal Investments for Electricity Generation: A Stochastic Model and a Test Problem*, ch. Numerical techniques for stochastic optimization problems. Berlin: Springer-Verlag, 1998.
- [49] J. Mulvey and A. Ruszcynski, “A new scenario decomposition method for large scale stochastic optimization,” *Operations Research*, vol. 43, pp. 477–490, 1995.
- [50] A. Shapiro and T. H. de Mello, “A simulation-based approach to two-stage stochastic programming with recourse.,” *Mathematical Programming*, vol. 81(3), pp. 301–325, 1998.
- [51] A. Brigatto, A. Street, and D. Valladão, “Assessing the cost of timeinconsistent operation policies in hydrothermal power systems.,” *IEEE Transactions on Power Systems*, vol. 32, pp. 4540–4550, 2017.

APPENDIX A

COMPUTATIONAL RESULTS TABLES

A.1 EE

Table A.1: Results for SD-EE for *pgp2* ($\psi = 450.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	444.39	450.86	28.75	450.86	5.29	830.20
	3.01	3.93	2.62	3.93	3.53	233.36
0.10	447.76	450.72	27.94	453.52	8.49	945.20
	3.26	3.13	2.48	3.30	7.83	300.13
0.20	450.73	452.35	29.54	458.25	4.82	783.20
	4.24	6.18	4.37	7.00	3.61	213.24
0.30	453.45	450.40	27.69	458.71	5.43	848.20
	3.90	2.78	2.26	3.26	3.59	215.06
0.40	455.91	450.91	27.66	461.97	3.50	707.70
	5.08	3.45	2.92	4.43	2.26	179.05
0.50	459.19	450.21	27.20	463.81	6.01	811.40
	4.80	2.21	2.09	2.70	6.77	287.81
0.60	461.90	450.95	27.14	467.24	5.22	814.90
	4.94	3.87	2.33	5.16	2.63	172.65
0.70	463.98	450.82	27.02	469.73	6.98	842.10
	4.98	3.02	2.38	4.53	7.87	323.80
0.80	465.50	449.56	27.14	471.27	9.18	854.90
	5.04	2.60	2.05	3.86	15.33	409.84
0.90	468.89	450.59	26.73	474.65	3.43	687.80
	5.92	3.58	1.44	4.56	1.83	138.79
1.00	471.25	451.47	26.19	475.66	6.21	867.10
	5.32	2.04	1.75	3.61	4.68	215.03

Table A.2: Results for SAA-EE for *pgp2* ($\psi = 450.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	445.86	448.32	29.85	448.32	4.21	31.43
	2.79	1.80	1.53	1.80	0.55	3.96
0.10	448.39	448.32	29.85	451.30	4.09	30.63
	2.95	1.80	1.53	1.95	0.38	2.77
0.20	450.91	448.32	29.85	454.29	4.23	30.33
	3.11	1.80	1.53	2.10	1.02	3.25
0.30	453.43	448.32	29.85	457.27	4.10	30.47
	3.27	1.80	1.53	2.25	0.45	3.33
0.40	455.95	448.32	29.85	460.26	4.15	30.90
	3.44	1.80	1.53	2.40	0.50	3.57
0.50	458.47	448.32	29.85	463.25	4.19	31.07
	3.60	1.80	1.53	2.55	0.51	3.70
0.60	460.97	448.60	29.41	466.25	4.09	30.53
	3.76	1.79	1.52	2.68	0.56	4.03
0.70	463.46	448.60	29.41	469.19	4.09	30.43
	3.92	1.79	1.52	2.83	0.62	4.51
0.80	465.95	448.60	29.41	472.13	4.20	31.30
	4.08	1.79	1.52	2.98	0.62	4.44
0.90	468.42	448.60	29.41	475.07	4.53	32.70
	4.25	1.79	1.52	3.14	0.89	4.31
1.00	470.90	448.60	29.41	478.01	4.38	32.57
	4.42	1.79	1.52	3.29	0.61	3.57

Table A.3: Results for SD-EE for *pgp2e* ($\psi = 420.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	411.61	417.81	50.49	418.81	2.98	690.50
	5.31	4.36	3.67	4.36	1.09	105.57
0.10	416.40	417.69	49.12	422.60	4.92	785.30
	5.48	5.28	3.32	5.53	4.16	239.25
0.20	420.94	419.08	49.21	428.93	5.19	805.70
	5.87	4.66	5.33	5.31	3.56	217.07
0.30	425.20	419.18	49.33	433.98	3.47	705.10
	6.90	5.49	5.15	6.84	1.71	120.94
0.40	430.06	420.30	49.24	439.99	3.99	710.20
	9.36	3.64	3.15	4.66	3.06	208.47
0.50	433.48	421.74	52.76	448.12	3.92	743.10
	7.75	3.74	3.02	4.73	1.50	121.40
0.60	437.06	420.22	50.58	450.57	5.60	849.90
	7.87	3.40	3.82	4.48	3.14	173.63
0.70	441.58	421.31	50.49	456.65	4.37	762.50
	8.24	2.81	3.96	4.95	2.57	180.67
0.80	450.01	418.45	47.92	456.79	4.71	781.90
	6.87	3.93	3.61	5.72	3.13	184.92
0.90	452.62	420.03	48.62	463.79	3.76	732.20
	9.84	4.89	5.08	8.80	1.51	108.39
1.00	457.30	420.40	48.85	469.25	5.37	797.30
	11.49	5.05	4.77	8.94	4.20	241.81

Table A.4: Results for SAA-EE for *pgp2e* ($\psi = 420.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	412.40	414.28	50.25	414.28	5.44	39.90
	4.78	2.81	2.47	2.81	0.78	5.84
0.10	417.10	413.55	47.60	418.31	5.10	37.27
	5.06	2.33	1.93	2.52	0.87	6.50
0.20	421.78	413.54	47.60	423.06	5.04	37.13
	5.34	2.33	1.93	2.70	0.65	4.93
0.30	426.46	413.61	47.40	427.83	5.10	37.47
	5.62	2.32	1.93	2.88	0.84	6.41
0.40	431.13	413.61	47.40	432.57	5.14	37.40
	5.91	2.32	1.93	3.06	0.75	5.55
0.50	435.74	413.61	47.40	437.31	5.38	39.23
	6.19	2.32	1.93	3.25	0.65	4.94
0.60	440.33	413.61	47.40	442.05	5.62	41.03
	6.45	2.32	1.93	3.44	0.85	6.49
0.70	444.92	413.89	46.96	446.77	5.74	41.37
	6.71	2.31	1.92	3.61	0.81	6.05
0.80	449.48	413.89	46.96	451.46	5.92	42.50
	6.96	2.31	1.92	3.80	0.80	5.89
0.90	454.04	413.89	46.96	456.16	5.80	41.57
	7.22	2.31	1.92	3.98	0.85	6.32
1.00	458.58	414.16	46.71	460.87	5.70	41.13
	7.48	2.30	1.92	4.16	0.84	6.35

Table A.5: Results for SD-EE for gbd ($\psi = 1672.33$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	1645.55	1660.63	245.39	1662.63	3.58	694.50
	23.84	15.79	12.84	15.79	2.33	182.25
0.10	1668.77	1664.93	246.83	1689.61	3.68	732.00
	19.55	17.17	14.57	18.47	1.37	105.00
0.20	1687.81	1664.95	248.15	1714.58	4.25	739.90
	21.98	17.97	14.62	20.78	3.22	201.96
0.30	1718.53	1661.99	244.82	1735.44	3.24	686.40
	23.82	18.23	13.37	21.94	1.11	103.14
0.40	1740.51	1665.98	248.78	1765.49	4.02	716.90
	25.59	21.71	16.35	28.01	2.63	198.10
0.50	1770.21	1660.68	244.24	1782.80	3.94	710.10
	34.78	15.39	12.43	20.67	3.11	216.83
0.60	1790.68	1663.31	246.36	1811.13	4.36	757.50
	28.68	20.24	15.44	29.26	2.98	206.33
0.70	1816.52	1660.66	243.72	1831.27	3.69	711.20
	28.95	20.98	13.35	29.96	2.65	167.99
0.80	1840.28	1663.87	245.45	1860.24	3.75	723.10
	28.57	19.73	13.79	29.98	1.14	104.63
0.90	1863.66	1665.35	244.85	1885.72	2.70	626.20
	38.12	21.37	14.21	33.69	1.32	110.21
1.00	1894.18	1665.13	244.73	1904.86	5.43	818.90
	32.54	23.57	16.68	39.92	4.74	223.11

Table A.6: Results for SAA-EE for gbd ($\psi = 162.73$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	1648.44	1653.95	243.99	1653.95	3.83	30.40
	20.84	10.15	7.51	10.15	0.37	2.81
0.10	1672.58	1653.95	243.99	1678.35	4.04	31.73
	22.45	10.15	7.51	10.87	0.42	3.13
0.20	1696.70	1653.94	243.81	1702.70	4.04	31.80
	24.07	10.14	7.51	11.58	0.37	2.67
0.30	1720.82	1653.94	243.81	1727.08	4.04	31.97
	25.71	10.14	7.51	12.30	0.35	2.51
0.40	1744.93	1653.94	243.81	1751.46	4.16	32.40
	27.35	10.14	7.51	13.03	0.37	2.76
0.50	1769.03	1653.94	243.81	1775.84	4.40	34.20
	28.99	10.14	7.51	13.76	0.44	3.29
0.60	1793.14	1653.94	243.81	1800.22	4.34	34.13
	30.64	10.14	7.51	14.49	0.36	2.76
0.70	1817.23	1653.94	243.81	1824.61	4.36	34.27
	32.30	10.14	7.51	15.22	0.37	2.77
0.80	1841.32	1653.94	243.81	1848.99	4.53	35.10
	33.96	10.14	7.51	15.96	0.44	3.18
0.90	1865.40	1654.03	243.67	1873.33	4.54	34.87
	35.62	10.14	7.51	16.69	0.37	2.78
1.00	1889.47	1654.03	243.67	1897.70	4.72	37.30
	37.28	10.14	7.51	17.43	0.62	4.68

Table A.7: Results for SD-EE for *LandS* ($\psi = 227.40$).

λ	Obj Val	$\mathbb{E}[f(x, \hat{\omega})]$	$\mathbb{E}[\nu(\hat{\omega})]$	UB	CPU	Iteration
0.00	224.97	226.98	23.94	226.98	3.46	724.10
	0.87	0.87	1.30	2.48	2.72	186.03
0.10	227.24	227.04	23.89	229.43	2.81	699.90
	1.78	1.84	1.14	1.94	1.21	109.15
0.20	229.69	227.22	24.59	232.13	2.91	679.70
	1.58	2.51	1.78	2.83	2.15	158.00
0.30	232.10	227.38	23.78	234.52	2.82	682.90
	1.29	2.24	1.38	2.61	1.22	103.20
0.40	234.42	227.58	23.57	237.01	3.02	682.50
	2.06	2.46	1.32	2.86	1.97	164.87
0.50	236.26	227.50	23.67	239.34	3.62	745.90
	1.62	1.87	1.40	2.50	1.92	153.79
0.60	238.65	226.81	23.59	240.96	3.75	741.90
	1.92	2.00	1.10	2.56	2.23	168.22
0.70	240.99	226.91	23.27	243.19	3.53	736.70
	2.07	2.27	1.36	3.17	1.60	144.39
0.80	243.71	228.08	23.44	246.83	3.43	712.60
	1.91	1.89	1.20	2.66	2.40	168.16
0.90	245.24	227.08	23.66	248.37	3.05	677.20
	2.42	1.61	1.11	2.53	2.08	152.39
1.00	248.31	227.94	22.89	250.83	3.75	723.00
	1.89	1.81	0.92	2.38	3.19	187.01

Table A.8: Results for SAA-EE for *LandS* ($\psi = 227.40$).

λ	Obj Val	$\mathbb{E}[f(x, \hat{\omega})]$	$\mathbb{E}[\nu(\hat{\omega})]$	UB	CPU	Iteration
0.00	224.78	225.52	23.14	225.52	4.43	31.63
	1.93	0.92	0.57	0.92	0.38	2.58
0.10	227.06	225.53	23.02	227.83	4.38	31.53
	2.03	0.91	0.57	0.96	0.38	2.58
0.20	229.33	225.54	22.95	230.13	4.38	31.47
	2.12	0.91	0.57	1.01	0.54	3.58
0.30	231.60	225.56	22.87	232.42	4.34	31.30
	2.22	0.90	0.57	1.06	0.38	2.59
0.40	233.85	225.58	22.81	234.71	4.28	30.97
	2.31	0.90	0.56	1.11	0.53	3.61
0.50	236.10	225.61	22.75	236.99	4.25	30.67
	2.41	0.90	0.56	1.16	0.32	2.17
0.60	238.35	225.64	22.69	239.26	4.26	30.70
	2.51	0.90	0.56	1.22	0.34	2.35
0.70	240.59	225.64	22.69	241.53	4.30	31.00
	2.60	0.90	0.56	1.27	0.35	2.38
0.80	242.82	225.70	22.61	243.79	4.29	30.90
	2.70	0.89	0.56	1.32	0.26	1.79
0.90	245.05	225.74	22.57	246.05	4.31	30.90
	2.80	0.89	0.56	1.37	0.36	2.43
1.00	247.28	225.76	22.54	248.31	4.40	31.67
	2.90	0.89	0.56	1.42	0.39	2.60

Table A.9: Results for SD-EE for *storm* ($\psi = 15580000.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	15420970.00	15563800.00	130966.60	15573800.00	129.92	702.20
	21470.59	37756.88	18172.03	37756.88	36.52	79.83
0.10	15440670.00	15566840.00	126987.30	15579520.00	190.38	763.00
	16553.85	28230.31	12767.68	29496.30	152.69	188.53
0.20	15439600.00	15582160.00	134747.00	15609090.00	184.99	739.70
	20582.42	39807.94	20669.83	43934.10	178.24	210.31
0.30	15440860.00	15576710.00	132268.00	15616390.00	117.32	654.70
	26092.92	38906.93	20729.33	45113.71	61.78	113.36
0.40	15443650.00	15567070.00	126857.40	15617830.00	158.28	705.60
	24236.97	25998.73	13492.27	31338.19	115.95	181.34
0.50	15451580.00	15605800.00	148318.10	15679960.00	101.58	623.90
	34227.32	59799.50	34489.04	76985.17	57.61	114.34
0.60	15468620.00	15601820.00	146132.70	15689490.00	202.20	759.00
	17552.60	63855.29	34311.69	84417.55	170.71	226.51
0.70	15479620.00	15566860.00	126751.90	15655570.00	138.54	671.80
	16796.89	25258.31	12798.47	34145.37	126.13	179.41
0.80	15486770.00	15566290.00	126411.20	15667410.00	218.29	800.30
	19361.51	26702.94	13019.36	37055.19	161.75	192.24
0.90	15496190.00	15572580.00	130198.80	15689750.00	163.54	727.40
	26900.17	34394.99	16059.95	48725.54	105.13	162.01
1.00	15500930.00	15560320.00	123821.20	15684140.00	134.85	696.20
	23415.94	34395.20	17843.76	52205.75	60.22	104.37

Table A.10: Results for SAA-EE for *storm* ($\psi = 15580000.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0.00	15501523.33	15498813.33	96695.31	15498813.33	62.63	51.17
	10118.70	4215.06	2128.78	4215.06	6.52	5.02
0.10	15511223.33	15498786.67	96682.92	15508466.67	63.81	51.93
	10404.86	4211.70	2128.55	4380.23	5.89	4.72
0.20	15520910.00	15498780.00	96681.90	15518133.33	59.94	50.73
	10526.33	4134.52	2092.76	4477.97	7.44	6.21
0.30	15527575.00	15497293.75	96396.99	15526206.25	63.45	54.00
	10574.11	4220.11	2122.67	4782.06	6.22	5.04
0.40	15540276.67	15498813.33	96688.78	15537486.67	64.48	53.27
	11147.61	4144.21	2092.90	4820.56	6.12	4.66
0.50	15549970.00	15498803.33	96689.52	15547143.33	65.92	54.77
	11477.87	4148.13	2092.75	4989.97	6.81	5.49
0.60	15559650.00	15498813.33	96686.37	15556820.00	62.23	53.33
	11811.15	4144.21	2092.64	5172.77	6.42	5.40
0.70	15569340.00	15498813.33	96680.17	15566486.67	63.36	53.03
	12367.21	4218.57	2127.78	5451.40	6.10	4.94
0.80	15579026.67	15498783.33	96681.40	15576136.67	64.70	54.30
	12717.49	4209.93	2128.44	5633.37	7.78	6.38
0.90	15588716.67	15498776.67	96681.07	15585803.33	66.71	54.30
	13085.77	4205.88	2128.93	5826.43	6.12	5.31
1.00	15598413.33	15498780.00	96682.28	15595470.00	63.23	52.67
	13456.05	4205.20	2128.74	6010.57	7.38	5.98

A.2 QDEV

Table A.11: Results for SD-QDEV for *pgp2*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	443.87	450.71	-	-	450.71	5.43	818.47
	3.14	2.36	-	-	2.36	3.71	202.37
0.1	457.09	450.88	2.97	557.76	467.50	4.04	720.13
	4.45	3.52	3.39	21.12	6.70	3.16	183.12
0.2	470.49	452.08	1.89	557.13	480.63	3.86	710.40
	5.24	4.55	0.96	11.14	4.08	2.16	143.46
0.3	484.64	452.93	1.54	557.32	493.51	4.11	705.47
	5.35	6.03	0.82	9.04	4.85	2.96	173.70
0.4	496.12	452.89	1.84	554.03	508.08	3.82	700.60
	5.27	5.07	0.93	9.60	5.16	2.45	161.49
0.5	508.02	452.90	1.62	554.56	519.89	3.69	673.40
	5.41	4.85	0.73	6.24	5.62	2.42	149.85
0.6	520.22	453.38	1.70	554.56	534.52	2.59	617.43
	5.50	5.89	1.05	7.52	9.73	1.16	94.42
0.7	532.46	453.17	1.58	555.23	546.70	2.42	588.27
	6.46	5.37	0.97	6.02	11.10	1.65	105.31
0.8	545.05	454.13	1.57	555.01	559.97	1.91	559.13
	6.37	5.28	0.98	5.94	12.58	0.48	47.00
0.9	556.85	453.93	1.64	553.35	573.02	1.77	548.63
	6.33	4.98	0.94	5.40	14.57	0.42	48.09
1	568.82	454.83	1.55	555.00	585.89	1.77	547.47
	6.71	6.19	1.00	5.90	16.89	0.49	53.07

Table A.12: Results for SAA-QDEV for *pgp2*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	446.15	449.05	-	-	449.05	4.47	29.67
	2.87	1.64	-	-	1.64	0.55	3.06
0.10	461.05	447.42	2.05	552.05	461.94	5.98	43.20
	3.84	1.08	0.62	2.67	1.93	0.86	5.99
0.20	474.43	447.69	2.06	551.75	476.28	5.79	42.07
	4.51	1.07	0.62	4.28	3.01	0.98	7.03
0.30	487.19	447.99	2.06	553.88	490.48	6.92	50.47
	5.03	1.05	0.62	5.11	4.16	1.02	7.53
0.40	499.05	448.23	2.03	553.88	504.58	5.78	44.23
	5.18	1.05	0.62	5.11	5.32	1.39	10.61
0.50	510.88	448.23	2.03	553.82	518.67	5.64	41.87
	5.36	1.05	0.62	4.98	6.49	1.25	9.05
0.60	522.71	448.78	1.96	553.82	532.13	5.46	39.43
	5.56	1.04	0.62	4.98	7.67	1.88	9.77
0.70	534.53	448.78	1.96	553.82	546.02	5.29	38.77
	5.76	1.04	0.62	4.98	8.85	1.35	9.86
0.80	546.34	448.78	1.96	553.82	559.91	5.41	39.20
	5.97	1.04	0.62	4.98	10.04	1.31	9.29
0.90	558.10	448.78	1.96	553.82	573.80	5.83	41.70
	6.16	1.04	0.62	4.98	11.22	1.31	9.31
1.00	569.64	448.78	1.96	553.77	587.55	5.93	43.57
	6.35	1.04	0.62	4.96	12.39	1.53	11.48

Table A.13: Results for SD-QDEV for *pgp2e*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	410.03	417.75	-	-	417.75	4.78	773.57
	6.26	5.43	-	-	5.43	3.44	193.43
0.1	439.60	419.03	5.37	637.17	451.59	3.48	666.30
	5.88	4.51	3.73	38.79	4.70	3.08	180.98
0.2	469.06	421.77	3.34	643.06	479.41	3.60	684.53
	6.10	4.16	1.73	24.05	4.31	2.44	163.39
0.3	493.79	421.66	3.14	638.81	505.63	3.11	664.80
	6.58	4.85	1.47	12.59	5.74	1.47	113.84
0.4	521.46	424.29	2.94	633.50	531.46	2.69	629.93
	5.88	5.04	1.08	15.44	4.09	1.24	105.61
0.5	546.24	425.20	2.70	637.62	558.39	2.76	631.33
	6.81	5.06	1.12	16.69	5.75	1.01	84.23
0.6	572.01	429.14	2.61	633.15	582.89	2.76	622.53
	6.64	4.75	1.10	12.01	9.69	1.65	120.29
0.7	595.36	430.85	2.93	629.34	610.83	2.02	567.50
	7.73	3.54	1.50	12.52	15.76	0.49	52.11
0.8	618.82	431.46	3.02	627.75	636.84	1.80	542.20
	8.89	2.32	1.79	15.56	19.72	0.32	31.62
0.9	642.51	431.04	2.99	628.94	662.94	1.85	544.47
	9.20	2.82	1.49	11.46	21.34	0.37	38.32
1	665.67	431.55	2.93	628.20	686.55	2.42	585.33
	10.24	2.32	1.55	12.71	22.60	0.94	75.22

Table A.14: Results for SAA-QDEV for *pgp2e*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	413.08	415.47	-	-	415.47	5.60	40.40
	4.64	2.11	-	-	2.11	0.64	4.69
0.10	444.45	419.06	2.46	640.93	446.46	8.21	58.33
	5.40	1.78	0.31	8.23	1.89	1.67	11.77
0.20	472.23	419.64	2.39	640.48	473.61	9.13	63.10
	5.89	1.75	0.31	7.54	2.14	1.54	11.13
0.30	498.88	421.47	2.26	640.78	500.05	8.83	59.83
	6.37	1.68	0.30	7.94	2.46	1.72	12.04
0.40	524.45	422.52	2.26	638.79	526.08	7.56	54.00
	6.69	1.66	0.30	8.11	2.89	1.49	11.07
0.50	549.51	426.58	2.25	634.51	551.68	7.82	55.80
	7.03	1.58	0.30	8.32	3.35	1.61	12.21
0.60	573.82	427.55	2.25	632.91	576.48	6.79	48.93
	7.49	1.56	0.30	8.81	3.86	1.60	11.97
0.70	597.72	431.32	2.26	632.37	600.77	5.70	41.10
	8.09	1.51	0.30	9.25	4.38	1.44	10.34
0.80	621.51	431.32	2.26	632.13	624.98	5.41	38.70
	8.83	1.51	0.30	9.23	4.91	1.03	7.09
0.90	645.29	431.32	2.26	632.11	649.18	5.30	38.10
	9.59	1.51	0.30	9.31	5.45	1.12	7.64
1.00	668.84	431.32	2.26	632.19	673.15	5.50	37.83
	10.35	1.51	0.30	9.37	5.99	1.20	7.80

Table A.15: Results for SD-QDEV for *gbd*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	1638.72	1666.58	-	-	1666.58	3.55	675.73
	28.50	10.35	-	-	10.35	2.07	149.71
0.1	1832.27	1669.77	25.73	3140.75	1868.33	3.49	636.90
	31.06	11.65	12.19	220.11	13.60	1.54	117.45
0.2	2028.16	1672.13	23.18	3169.36	2064.28	3.12	610.03
	37.79	15.30	10.23	190.43	13.15	1.56	110.09
0.3	2220.46	1671.89	23.04	3159.05	2256.26	3.46	626.90
	41.98	13.29	7.89	145.51	17.96	1.33	100.01
0.4	2416.60	1687.36	23.13	3147.37	2456.37	3.24	604.80
	54.29	20.34	13.60	221.77	33.22	1.08	86.09
0.5	2597.60	1701.25	23.49	3119.72	2645.40	3.02	578.13
	68.47	30.11	9.90	176.76	29.24	0.88	72.59
0.6	2762.74	1962.86	23.42	2915.06	2815.27	3.43	595.20
	73.64	125.37	8.36	105.18	26.15	1.48	96.67
0.7	2896.51	2125.57	17.72	2914.37	2925.80	3.30	588.17
	70.58	62.29	5.47	84.93	24.71	1.22	84.13
0.8	3010.04	2152.21	16.84	2928.23	3042.45	3.09	571.63
	65.54	47.37	5.55	100.89	26.77	0.78	58.69
0.9	3111.26	2174.20	16.01	2941.87	3153.31	2.96	559.73
	82.52	59.16	6.34	106.05	32.84	0.84	63.01
1	3221.10	2187.47	15.69	2942.46	3255.19	3.43	591.97
	78.29	42.57	4.92	81.97	34.65	1.44	92.93

Table A.16: Results for SAA-QDEV for *gbd*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	1649.34	1658.49	-	-	1658.49	4.13	30.13
	19.25	11.98	-	-	11.98	0.37	2.54
0.10	1843.75	1658.49	20.71	3179.05	1854.27	7.64	52.23
	22.09	11.98	1.58	78.83	13.24	0.84	5.82
0.20	2037.44	1661.56	21.24	3165.40	2049.75	8.11	55.07
	25.80	11.94	1.60	80.68	14.90	0.91	6.03
0.30	2230.27	1661.84	21.59	3164.87	2243.93	8.42	55.10
	29.99	11.91	1.62	80.03	16.81	1.19	7.59
0.40	2421.95	1670.38	21.42	3138.16	2440.23	7.89	56.20
	34.62	11.82	1.63	87.51	18.85	1.12	7.73
0.50	2608.38	1690.46	22.96	3109.56	2627.75	7.52	52.93
	39.49	11.79	1.73	85.02	22.00	1.47	9.97
0.60	2779.16	2030.04	20.83	2903.43	2799.83	7.66	50.63
	42.25	9.54	1.59	54.24	22.12	1.48	9.33
0.70	2898.64	2148.71	17.83	2897.88	2923.01	7.30	49.83
	40.56	7.85	1.36	45.96	20.90	1.44	9.33
0.80	3007.37	2176.52	15.56	2918.44	3031.11	7.20	48.67
	41.41	7.93	1.23	45.88	20.95	1.35	8.86
0.90	3112.63	2187.72	14.96	2939.02	3137.34	7.00	47.13
	43.00	7.97	1.20	50.72	22.18	1.53	9.90
1.00	3214.50	2207.03	13.42	2952.87	3239.45	6.55	45.53
	44.64	8.04	1.11	53.56	22.26	1.63	10.96

Table A.17: Results for SD-QDEV for *LandS*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	224.32	226.51	-	-	226.51	3.28	686.77
	2.06	1.38	-	-	1.38	2.95	173.47
0.1	235.32	226.51	1.30	318.98	238.36	3.51	688.33
	1.84	1.21	0.76	11.90	1.23	2.25	148.15
0.2	247.00	226.55	0.87	324.34	249.58	2.57	628.07
	2.01	1.02	0.45	10.55	0.97	1.01	91.35
0.3	258.11	226.96	0.88	322.31	260.82	2.93	651.63
	1.96	1.55	0.43	8.10	1.33	1.89	133.08
0.4	269.20	227.46	0.98	318.49	271.68	3.05	630.87
	2.18	1.50	0.35	6.18	1.16	2.15	139.61
0.5	280.29	228.09	0.81	320.18	282.20	2.68	629.10
	2.36	1.37	0.31	6.52	1.05	1.65	124.45
0.6	291.03	228.63	0.86	319.13	293.19	2.40	598.23
	2.80	1.89	0.33	5.65	1.61	0.91	82.93
0.7	301.82	228.84	0.80	319.12	303.27	2.22	596.13
	2.91	1.71	0.25	4.43	1.30	0.82	74.23
0.8	312.50	229.60	0.84	318.31	314.08	1.93	568.83
	3.25	1.58	0.35	5.85	1.58	0.55	56.60
0.9	323.29	229.63	0.81	318.43	324.15	1.68	538.87
	3.64	1.32	0.24	4.35	1.29	0.24	27.91
1	333.88	230.79	0.83	318.55	335.05	1.91	562.63
	3.66	2.04	0.34	5.53	2.16	0.49	52.81

Table A.18: Results for SAA-QDEV for *LandS*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	225.29	225.72	-	-	225.72	4.45	30.83
	1.83	0.84	-	-	0.84	0.42	2.83
0.10	236.86	225.82	1.03	322.19	237.31	7.10	48.80
	1.70	0.82	0.06	2.75	0.79	0.86	5.74
0.20	248.12	226.05	1.01	321.08	248.60	7.12	48.80
	1.62	0.81	0.06	2.65	0.75	1.01	6.79
0.30	259.21	226.41	1.05	320.25	259.75	7.12	48.60
	1.60	0.79	0.06	2.76	0.75	1.04	7.03
0.40	270.13	226.83	1.02	319.59	270.71	6.67	47.57
	1.62	0.78	0.06	2.73	0.75	0.87	6.05
0.50	280.90	227.59	0.99	318.76	281.47	6.62	47.57
	1.70	0.76	0.05	2.56	0.76	0.83	5.91
0.60	291.53	228.29	0.98	318.51	292.13	6.20	44.47
	1.80	0.74	0.05	2.53	0.80	0.93	6.61
0.70	302.08	228.77	0.98	318.57	302.74	6.35	45.30
	1.94	0.74	0.05	2.61	0.85	1.23	8.49
0.80	312.56	229.15	0.96	318.57	313.25	6.03	43.20
	2.10	0.73	0.05	2.61	0.91	1.01	7.10
0.90	322.99	229.46	0.96	318.53	323.71	6.08	41.93
	2.28	0.73	0.05	2.64	0.97	1.05	7.11
1.00	333.28	229.78	0.95	318.58	334.04	6.15	42.60
	2.48	0.73	0.05	2.52	1.04	1.17	7.88

Table A.19: Results for SD-QDEV for *storm*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	15422680.00	15591983.33	-	-	15591983.33	207.32	759.83
	20372.57	57245.94	-	-	57245.94	147.90	182.17
0.1	15454273.33	15591373.33	23665.86	16124053.33	15691973.33	186.08	708.60
	29652.17	63630.46	33749.04	325617.64	74244.40	143.66	157.62
0.2	15521170.00	15607940.00	18874.62	16065373.33	15774926.67	187.75	690.40
	24363.43	55467.60	14333.34	173342.40	74569.29	102.48	137.28
0.3	15534932.67	15664860.00	16662.85	16199363.33	15871326.67	218.97	720.23
	22970.40	91761.89	20575.04	145574.35	68234.16	203.41	190.77
0.4	15674830.00	15579053.33	12055.10	16071606.67	15872516.67	186.74	683.30
	27533.55	39294.14	8467.83	106310.51	56309.56	109.32	140.21
0.5	15751153.33	15566736.67	12879.48	16042400.00	15933363.33	360.41	677.73
	26867.57	39336.99	8493.82	94654.26	62823.54	892.89	158.27
0.6	15827613.33	15557926.67	12641.39	16041616.67	15999830.00	172.91	613.87
	26673.27	30313.47	9353.36	108010.72	67340.00	252.42	125.71
0.7	15896616.67	15556690.00	11933.28	16033393.33	16057450.00	107.13	586.80
	28676.67	27048.96	6306.74	77121.45	48598.24	49.23	93.30
0.8	15989576.67	15537776.67	10159.71	16037323.33	16099973.33	109.38	589.47
	31775.61	17129.73	5524.26	70178.49	42593.03	45.62	79.37
0.9	16060676.67	15528690.00	8944.29	16038656.67	16148666.67	95.62	533.07
	26052.15	15790.36	3845.86	52747.77	31349.80	52.07	29.90
1	16155620.00	15508736.67	7856.80	16030400.00	16186853.33	80.41	504.20
	30759.23	8354.38	2064.63	37108.35	14097.18	57.77	2.95

Table A.20: Results for SAA-QDEV for *storm*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.10	15566925.00	15498032.14	-	-	15565571.43	87.01	70.39
	9561.37	4238.19	-	-	4227.24	7.69	6.04
0.20	15632553.33	15498093.33	7176.79	16043990.00	15633230.00	101.62	75.93
	13403.49	4420.40	545.51	30387.89	4844.27	13.43	9.85
0.30	15700537.93	15498141.38	7182.94	16044527.59	15700696.55	101.36	75.45
	15283.40	4500.12	542.72	30798.10	5385.20	10.88	7.77
0.40	15767700.00	15498136.67	7244.27	16044283.33	15768440.00	103.34	79.93
	17163.17	4428.47	547.88	29416.59	5949.71	14.76	11.28
0.50	15835253.33	15498133.33	7221.02	16043733.33	15835980.00	102.47	80.63
	19282.67	4425.81	546.99	29255.61	6668.27	16.28	12.61
0.60	15902790.00	15498163.33	7235.07	16044000.00	15903613.33	98.71	80.07
	21498.99	4413.65	547.42	30863.51	7439.72	14.18	11.23
0.70	15970316.67	15498140.00	7218.85	16043140.00	15971133.33	104.78	79.07
	23797.71	4433.07	546.90	30211.49	8273.11	15.97	11.74
0.80	16036666.67	15500000.00	7224.87	16043333.33	16006666.67	98.09	79.13
	49013.25	0.00	547.01	50400.69	25370.81	15.17	12.09
0.90	16105300.00	15498173.33	7211.91	16043213.33	16106270.00	105.91	80.43
	28568.94	4421.26	546.52	29483.72	10019.16	17.70	13.13
1.00	16172080.00	15498173.33	7211.66	16044086.67	16173176.67	100.43	80.90
	30979.14	4421.26	546.52	30093.17	10933.65	18.95	14.84

A.3 CVaR

Table A.21: Results for SD-CVaR for *pgp2*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	443.34	450.53	-	-	450.53	5.16	805.13
	3.69	2.40	-	-	2.40	4.31	238.52
0.1	502.18	450.40	2.48	557.06	511.06	5.01	776.17
	5.39	3.44	1.90	17.85	4.15	4.28	228.52
0.2	560.36	451.72	2.12	557.08	571.60	4.00	734.80
	5.81	4.95	1.92	16.71	7.60	1.97	149.21
0.3	617.14	452.44	1.76	554.69	629.40	3.62	701.00
	6.91	4.15	0.69	8.30	5.11	2.16	138.43
0.4	675.00	452.34	1.80	556.73	689.46	4.12	726.70
	7.42	4.42	1.45	9.36	9.88	3.04	173.25
0.5	731.30	452.26	1.73	556.20	747.66	3.02	651.27
	8.46	4.78	0.92	7.77	7.37	1.61	117.92
0.6	788.04	453.25	1.63	555.82	806.34	3.05	660.20
	8.99	5.51	0.91	7.59	8.54	1.42	115.24
0.7	844.94	452.58	1.65	555.94	864.83	3.06	654.93
	9.75	4.59	0.91	9.12	9.81	1.69	119.93
0.8	903.16	453.20	1.62	556.12	924.08	2.77	624.30
	9.78	4.53	0.97	6.09	12.12	2.29	135.57
0.9	958.84	452.61	1.64	555.98	982.48	2.67	625.53
	10.49	4.85	0.92	6.92	12.72	1.11	99.32
1	1016.07	452.57	1.60	554.81	1039.34	3.35	667.33
	10.82	4.74	0.68	5.94	10.64	2.20	148.85

Table A.22: Results for SAA-CVaR for *pgp2*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	446.62	448.05	-	-	448.05	4.09	30.40
	3.19	1.78	-	-	1.78	0.49	3.64
0.10	506.50	447.24	2.00	554.03	506.40	5.91	43.13
	4.94	0.86	0.59	4.56	1.78	1.00	7.08
0.20	565.02	449.90	1.62	553.78	566.84	5.71	42.20
	6.00	0.80	0.43	5.34	2.17	0.92	6.82
0.30	623.23	450.10	1.59	553.39	625.39	5.94	43.60
	7.03	0.80	0.42	5.18	2.91	1.01	7.36
0.40	681.07	458.77	0.96	555.93	690.65	6.90	51.07
	7.96	0.75	0.11	5.79	1.27	1.24	8.59
0.50	738.33	459.32	0.89	555.86	748.49	6.26	46.70
	8.59	0.75	0.10	5.74	1.40	1.40	10.18
0.60	795.55	459.32	0.89	555.86	806.33	6.12	45.57
	9.23	0.75	0.10	5.74	1.58	1.37	9.86
0.70	852.76	459.56	0.87	555.86	864.04	5.78	43.20
	9.87	0.74	0.10	5.74	1.74	1.47	10.52
0.80	909.97	459.56	0.87	555.86	921.82	5.53	41.13
	10.52	0.74	0.10	5.74	1.93	1.43	10.33
0.90	967.18	459.56	0.87	555.86	979.60	5.44	40.63
	11.17	0.74	0.10	5.74	2.12	1.38	10.12
1.00	1024.38	459.56	0.87	555.86	1037.39	5.60	41.87
	11.83	0.74	0.10	5.74	2.31	1.46	10.57

Table A.23: Results for SD-CVaR for *pgp2e*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	409.68	417.73	-	-	417.73	4.29	774.80
	6.16	5.46	-	-	5.46	2.97	190.21
0.1	481.80	419.09	5.27	644.24	494.05	5.49	824.97
	6.44	3.34	4.16	41.99	5.26	3.57	209.79
0.2	551.28	420.06	3.17	647.44	562.21	4.64	766.63
	7.11	3.95	1.47	20.17	4.06	3.76	198.68
0.3	618.25	421.14	3.09	645.91	633.48	3.94	706.43
	8.15	4.18	1.79	19.73	6.34	3.71	203.86
0.4	687.29	422.14	3.40	636.06	703.78	3.59	712.27
	9.05	4.71	1.43	16.14	8.63	1.80	134.16
0.5	754.27	421.79	2.84	645.63	772.97	2.66	631.07
	10.00	4.49	1.28	18.08	9.30	1.54	120.46
0.6	822.88	423.35	2.82	637.23	839.55	2.88	660.03
	9.00	3.86	0.96	15.45	6.28	1.07	94.64
0.7	889.48	423.48	3.02	634.89	910.19	3.25	684.07
	11.26	4.51	1.38	12.91	12.60	1.68	126.31
0.8	956.31	424.43	3.01	635.82	981.21	2.57	630.33
	12.29	4.56	1.60	17.48	17.57	1.09	92.85
0.9	1023.32	425.27	2.94	635.45	1050.08	2.57	623.67
	12.73	5.69	1.54	14.41	20.25	1.30	115.64
1	1090.77	425.61	2.99	633.99	1119.35	2.49	619.63
	14.81	4.46	1.57	16.88	20.38	0.98	85.62

Table A.24: Results for SAA-CVaR for *pgp2e*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	413.37	414.31	-	-	414.31	5.43	38.83
	5.09	2.45	-	-	2.45	0.93	6.83
0.10	486.13	414.05	6.90	639.91	491.32	8.42	57.60
	5.54	2.08	1.30	8.91	4.30	1.60	11.01
0.20	555.81	414.33	6.90	639.44	568.48	8.89	61.03
	6.06	2.06	1.30	8.32	6.78	1.63	11.31
0.30	624.89	415.14	6.83	637.30	645.33	9.11	63.07
	6.66	2.04	1.30	7.93	9.31	1.97	14.01
0.40	693.42	429.70	1.91	638.58	703.92	9.21	63.97
	7.34	1.54	0.16	7.63	2.50	2.16	15.10
0.50	761.38	429.70	1.91	638.78	772.48	8.32	57.90
	7.99	1.54	0.16	8.50	2.78	1.65	11.90
0.60	829.21	430.76	1.92	636.53	840.89	7.97	56.33
	8.62	1.52	0.16	7.78	3.06	1.57	11.59
0.70	896.97	430.76	1.92	637.07	909.25	7.95	55.90
	9.31	1.52	0.16	8.61	3.36	1.71	12.37
0.80	964.69	431.50	1.89	635.65	977.61	8.11	56.40
	9.99	1.51	0.16	7.44	3.63	1.71	12.14
0.90	1032.24	433.15	1.89	634.20	1045.73	8.16	57.57
	10.63	1.48	0.16	8.25	3.91	1.79	13.26
1.00	1099.57	435.33	1.91	632.28	1113.74	7.56	53.43
	11.29	1.45	0.16	7.69	4.20	1.61	11.81

Table A.25: Results for SD-CVaR for *gbd*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	1640.30	1667.43	-	-	1667.43	3.00	646.17
	27.04	11.86	-	-	11.86	1.68	131.35
0.1	1995.71	1670.49	28.54	3097.69	2037.34	3.73	679.17
	30.53	14.95	13.68	208.39	19.03	2.10	155.49
0.2	2351.91	1668.10	24.10	3166.71	2397.85	3.15	621.70
	44.19	14.68	12.44	219.02	24.36	1.35	106.14
0.3	2714.99	1674.57	23.47	3164.79	2764.82	3.01	602.60
	57.42	16.09	8.57	175.47	28.70	1.23	97.40
0.4	3071.76	1675.68	23.78	3153.01	3127.12	3.85	660.07
	56.69	14.71	11.07	192.02	28.76	1.87	122.80
0.5	3427.88	1678.45	23.49	3154.64	3490.65	3.49	635.73
	75.74	24.51	8.11	149.41	38.67	2.29	143.20
0.6	3778.60	1686.93	23.33	3135.17	3847.96	3.21	619.57
	73.23	23.80	7.96	140.05	40.01	1.05	85.93
0.7	4145.13	1690.85	21.78	3149.21	4200.17	3.78	657.87
	90.75	24.53	6.88	136.60	32.92	2.10	135.02
0.8	4484.25	1699.55	25.75	3078.46	4574.27	3.32	628.03
	104.74	32.09	8.60	131.76	51.13	1.13	87.33
0.9	4842.37	1707.90	23.99	3113.89	4942.17	3.11	598.30
	121.71	58.43	11.41	207.09	101.46	1.21	96.46
1	5197.87	1701.97	23.07	3120.71	5284.17	2.83	583.67
	136.94	30.11	7.52	153.80	50.67	0.95	77.95

Table A.26: Results for SAA-CVaR for *gbd*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	1654.56	1656.12	-	-	1656.12	4.15	30.57
	21.35	8.23	-	-	8.23	0.37	2.65
0.10	2015.53	1656.11	23.14	3195.16	2017.30	8.29	54.30
	27.37	8.23	1.84	103.73	10.67	1.05	6.57
0.20	2376.04	1659.40	22.69	3186.64	2378.79	8.54	56.37
	34.28	8.14	1.82	103.18	13.40	1.01	6.41
0.30	2735.87	1659.40	22.50	3179.08	2738.41	8.54	55.87
	41.91	8.14	1.80	109.92	16.55	1.11	6.93
0.40	3095.38	1659.48	22.54	3179.83	3098.02	8.37	54.63
	49.76	8.14	1.81	102.56	19.88	1.25	7.93
0.50	3454.78	1659.48	22.49	3181.16	3457.62	8.56	54.97
	57.74	8.14	1.80	105.05	23.27	1.18	7.37
0.60	3813.84	1659.48	22.54	3167.78	3817.29	8.85	57.50
	65.95	8.14	1.81	110.66	26.76	1.38	8.76
0.70	4170.97	1682.86	22.54	3142.78	4174.39	8.70	56.60
	74.45	7.95	1.82	112.54	30.07	1.70	10.72
0.80	4526.70	1682.86	22.54	3135.40	4530.32	8.38	54.53
	82.81	7.95	1.82	109.22	33.62	1.72	10.81
0.90	4882.25	1684.09	22.31	3133.13	4885.83	8.07	53.73
	91.14	7.92	1.80	108.02	36.90	1.62	10.50
1.00	5237.35	1691.74	22.61	3127.28	5241.46	7.80	52.47
	99.50	7.91	1.82	109.48	40.80	1.62	10.30

Table A.27: Results for SD-CVaR for *LandS*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	224.28	226.29	-	-	226.29	3.79	746.43
	1.95	1.38	-	-	1.38	3.42	224.00
0.1	258.15	226.41	1.16	322.35	260.97	4.31	774.07
	2.28	1.04	0.85	13.54	1.29	3.08	200.04
0.2	291.96	226.72	1.01	321.03	294.97	3.22	699.07
	1.95	1.40	0.55	9.37	1.45	2.00	153.71
0.3	325.34	226.57	0.97	321.37	328.79	2.87	662.70
	2.60	1.28	0.47	7.90	1.69	1.84	145.96
0.4	358.99	226.87	1.09	318.37	362.97	2.17	611.47
	2.86	1.62	0.51	8.46	2.14	0.94	96.17
0.5	393.16	227.16	0.90	320.20	396.29	3.20	683.40
	3.02	1.45	0.35	5.76	1.92	2.07	151.16
0.6	426.02	227.14	0.97	319.59	430.49	2.36	620.00
	3.20	1.33	0.42	7.17	2.04	1.14	99.92
0.7	459.97	227.61	0.84	321.11	464.19	2.72	650.37
	3.87	1.71	0.39	6.48	1.83	1.56	130.50
0.8	493.46	228.08	0.84	319.89	497.40	2.62	648.13
	4.24	1.63	0.37	6.57	2.56	1.44	106.84
0.9	527.21	228.00	0.88	319.32	531.22	2.42	634.77
	4.52	1.49	0.34	6.15	2.46	0.89	82.63
1	560.21	227.87	0.83	319.59	564.15	2.32	613.03
	4.71	1.41	0.30	5.77	2.15	1.48	122.56

Table A.28: Results for SAA-CVaR for *LandS*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	225.18	225.59	-	-	225.59	4.30	31.13
	1.86	0.76	-	-	0.76	0.39	2.57
0.10	259.37	225.68	1.08	322.76	259.77	6.95	49.23
	1.94	0.74	0.07	3.47	0.81	0.89	5.96
0.20	293.29	225.91	1.01	321.90	293.67	6.86	48.67
	2.06	0.73	0.06	3.39	0.87	0.84	5.77
0.30	327.10	226.14	1.02	321.24	327.49	6.96	49.50
	2.20	0.72	0.06	3.34	0.94	0.79	5.60
0.40	360.83	226.32	1.01	320.90	361.23	6.99	49.67
	2.37	0.72	0.06	3.37	1.03	0.77	5.43
0.50	394.51	226.57	0.99	320.44	394.89	6.74	48.17
	2.56	0.71	0.06	3.44	1.11	0.89	6.24
0.60	428.14	226.85	0.95	320.18	428.45	6.94	49.43
	2.77	0.71	0.06	3.31	1.18	0.81	5.66
0.70	461.73	227.06	0.95	319.99	462.02	6.84	48.97
	2.98	0.70	0.06	3.39	1.28	0.86	5.95
0.80	495.28	227.33	0.94	319.68	495.56	6.70	48.00
	3.21	0.70	0.06	3.21	1.37	0.82	5.72
0.90	528.81	227.57	0.95	319.57	529.10	6.56	47.23
	3.44	0.69	0.06	3.29	1.47	0.82	5.82
1.00	562.32	227.64	0.96	319.48	562.64	6.56	47.33
	3.67	0.69	0.06	3.27	1.59	0.94	6.67

Table A.29: Results for SD-CVaR for *storm*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0	15422153.33	15570036.67	-	-	15570036.67	191.40	739.53
	21608.42	34373.48	-	-	34373.48	227.03	220.22
0.1	16987016.67	15607136.67	23322.37	16095873.33	17263370.00	272.22	760.00
	35869.36	73111.65	29027.16	266676.28	110094.17	375.31	265.84
0.2	18594783.33	15593096.67	20395.16	16053623.33	18885396.67	319.22	666.37
	38707.59	53172.65	23136.93	189259.23	102311.86	841.17	135.70
0.3	20211133.33	15606653.33	13726.66	16137503.33	20530256.67	231.80	707.40
	38037.66	82402.14	18702.58	176675.56	171916.47	203.88	163.98
0.4	21824453.33	15572880.00	19165.61	16001770.00	22126903.33	172.63	683.23
	38020.36	41067.18	16961.07	130221.76	116502.30	95.08	128.08
0.5	23438120.00	15585206.67	13697.15	16063480.00	23753910.00	178.56	676.60
	38422.29	60973.58	11721.88	126651.06	113212.19	76.93	116.86
0.6	25019463.33	15560723.33	13475.37	16035830.00	25343920.00	196.94	690.17
	58900.53	30896.44	8714.03	131448.93	62531.33	101.88	146.96
0.7	26638170.00	15597443.33	20496.93	16016420.00	27115896.67	224.50	711.97
	114688.66	230764.69	12350.35	121430.34	3017201.12	113.15	142.99
0.8	28275906.67	15573540.00	13736.90	16041686.67	28626693.33	481.73	760.70
	48369.73	47694.26	9996.43	104554.91	124117.80	202.21	378.96
0.9	29874676.67	15561676.67	13553.30	16032080.00	30034510.00	204.17	654.77
	64696.27	23976.76	8753.25	102739.16	91098.11	145.02	144.37
1	31444123.33	15666460.00	11785.30	16077246.67	31781916.67	228.69	652.73
	67541.26	50906.39	9978.03	160756.74	163565.04	118.57	107.35

Table A.30: Results for SAA-CVaR for *storm*

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	ψ	UB	CPU	Iteration
0.00	15497596.67	15498453.33	-	-	15498453.33	69.01	52.13
	10945.40	3908.32	-	-	3908.32	6.34	4.35
0.10	17114070.00	15498463.33	5016.45	16037800.00	17115960.00	101.27	73.03
	11636.51	3908.65	411.22	22545.86	4409.75	11.39	7.88
0.20	18730530.00	15498470.00	4907.32	16037680.00	18733553.33	93.26	75.70
	12784.80	3907.21	407.71	22934.97	4984.72	10.90	8.72
0.30	20346980.00	15498500.00	5009.00	16038146.67	20350970.00	95.21	76.57
	14263.68	3913.90	411.13	22831.51	5635.18	8.24	6.44
0.40	21963416.67	15498463.33	4989.46	16037786.67	21968466.67	99.27	77.23
	16008.10	3906.01	410.35	22999.97	6319.56	11.15	8.38
0.50	23579850.00	15498473.33	4983.58	16038343.33	23585993.33	101.66	79.10
	17920.72	3912.93	410.20	22581.92	7020.02	13.68	10.28
0.60	25196283.33	15498470.00	4981.02	16038160.00	25203513.33	94.16	76.70
	19968.42	3913.12	409.98	22817.49	7748.22	12.77	10.11
0.70	26812706.67	15498463.33	4996.37	16038330.00	26820960.00	102.71	78.30
	22090.39	3908.65	410.54	22851.73	8512.29	12.34	9.27
0.80	28429133.33	15498510.00	4986.67	16038066.67	28438523.33	102.65	77.53
	24301.37	3909.52	409.99	22674.04	9280.04	15.25	11.29
0.90	30045556.67	15498493.33	4993.40	16037963.33	30055996.67	98.21	78.77
	26566.68	3910.55	410.31	22651.25	10044.85	14.98	11.93
1.00	31661976.67	15498510.00	4988.75	16038090.00	31673533.33	101.01	76.17
	28862.18	3909.52	410.06	23181.12	10820.20	15.87	11.54

Table A.31: Results for SD-EE for *pgp2* ($\psi = 557.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0	443.31	451.77	5.57	451.77	3.62	717.23
	3.71	3.30	3.36	3.30	2.12	152.35
0.1	443.63	450.79	3.81	451.17	6.49	843.23
	2.92	2.86	2.40	2.90	5.94	266.91
0.2	444.02	451.06	4.66	452.00	4.69	751.20
	3.52	4.77	5.44	5.67	3.92	220.74
0.3	444.47	450.27	3.78	451.40	4.62	748.40
	3.96	2.37	2.85	2.76	2.50	150.75
0.4	444.52	450.45	3.61	451.89	6.29	773.77
	3.62	2.78	2.50	3.06	8.11	277.79
0.5	444.80	450.33	3.22	451.94	5.02	743.77
	4.00	2.10	2.51	2.46	4.25	219.31
0.6	445.24	451.11	3.81	453.40	7.13	843.77
	3.40	5.68	2.72	6.48	10.25	297.59
0.7	445.68	449.99	2.79	451.94	7.56	862.97
	3.97	3.11	1.94	3.59	10.95	316.61
0.8	445.76	450.64	2.97	453.01	5.14	817.03
	4.13	3.16	2.21	4.28	3.47	195.66
0.9	445.77	450.15	2.63	452.52	5.00	775.97
	4.02	2.99	2.06	3.80	3.70	205.30
1	446.26	450.09	2.72	452.81	5.97	805.37
	3.85	2.93	2.09	4.16	5.63	261.50

Table A.32: Results for SD-EE for $pgp2e$ ($\psi = 645.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0	410.11	417.83	9.41	417.83	4.48	768.57
	5.77	3.84	5.31	3.84	3.55	208.90
0.1	410.51	418.46	9.74	419.44	4.65	767.50
	5.56	4.53	6.49	4.95	4.18	235.55
0.2	410.67	417.02	6.63	418.34	4.78	786.53
	5.65	3.39	2.28	3.28	3.50	220.77
0.3	411.07	417.69	7.97	420.08	3.43	713.87
	5.71	3.89	4.45	4.31	2.06	143.50
0.4	412.43	417.89	7.05	420.71	3.50	721.60
	5.44	4.57	4.70	5.52	1.85	142.34
0.5	412.40	417.43	6.80	420.83	4.82	773.10
	5.98	3.04	2.85	3.35	4.90	249.88
0.6	412.49	417.63	5.60	421.00	3.75	720.43
	5.99	3.13	2.18	2.95	3.07	188.06
0.7	413.66	417.77	6.40	422.25	4.15	739.87
	5.93	3.86	3.76	5.22	4.13	212.05
0.8	413.79	418.00	6.38	423.10	4.38	762.10
	6.10	3.79	2.97	5.38	3.79	214.18
0.9	413.14	417.84	5.93	423.18	4.61	768.23
	6.11	3.39	2.74	3.15	4.18	227.98
1	415.19	417.22	5.10	422.31	3.88	736.30
	6.19	2.94	2.06	2.94	2.62	185.92

Table A.33: Results for SD-EE for gbd ($\psi = 3100.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0	1641.51	1666.42	26.62	1666.42	3.86	721.60
	26.84	10.96	2.12	10.96	2.04	160.30
0.1	1642.55	1670.21	26.99	1672.90	2.95	642.30
	26.52	13.67	2.39	13.81	1.91	139.35
0.2	1647.57	1666.22	26.40	1671.50	4.03	723.93
	23.80	13.79	2.01	13.99	2.47	170.13
0.3	1647.96	1666.98	26.70	1674.99	4.02	718.77
	28.94	11.79	2.37	12.23	2.01	150.08
0.4	1655.10	1665.09	26.36	1675.64	3.73	688.53
	28.86	9.74	1.77	10.12	2.40	162.07
0.5	1655.98	1665.76	26.30	1678.91	3.94	707.57
	27.86	10.60	1.86	11.01	3.22	203.82
0.6	1658.22	1668.36	26.49	1684.25	3.68	705.43
	28.27	13.81	2.13	14.54	2.61	176.66
0.7	1658.07	1667.47	26.28	1685.86	2.94	657.83
	28.28	9.62	2.05	10.42	1.05	97.08
0.8	1662.08	1668.89	26.68	1690.24	3.51	683.90
	30.31	13.40	2.30	14.64	2.01	157.16
0.9	1666.54	1667.30	26.21	1690.89	3.87	692.70
	30.36	11.38	2.06	12.30	2.52	172.72
1	1667.42	1668.63	26.25	1694.88	3.29	681.07
	30.54	14.05	2.02	14.93	1.94	155.07

Table A.34: Results for SD-EE for *Lands* ($\psi = 320.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0	224.49	226.49	1.15	226.49	3.25	733.00
	2.07	1.39	0.27	1.39	1.96	149.16
0.1	224.44	226.35	1.11	226.46	3.05	711.33
	2.17	1.67	0.22	1.66	1.68	143.28
0.2	224.62	226.86	1.32	227.13	3.69	699.00
	2.16	1.88	0.46	1.95	1.90	146.94
0.3	224.55	226.40	1.17	226.75	4.12	701.60
	2.08	1.34	0.25	1.36	3.05	185.09
0.4	224.77	226.25	1.26	226.75	4.31	736.73
	1.96	1.28	0.23	1.28	3.61	186.11
0.5	224.84	226.25	1.25	226.87	4.86	768.83
	2.07	1.18	0.26	1.24	3.12	200.19
0.6	224.97	226.65	1.12	227.32	3.31	674.20
	1.93	1.29	0.22	1.29	1.69	132.37
0.7	225.01	226.46	1.23	227.32	4.07	726.10
	1.98	1.33	0.29	1.35	2.80	169.29
0.8	225.41	226.13	1.16	227.06	3.89	720.43
	1.86	1.02	0.20	1.02	2.07	158.44
0.9	225.43	226.40	1.21	227.49	3.90	711.07
	2.25	1.03	0.31	1.06	2.84	186.82
1	225.69	226.13	1.12	227.25	5.18	774.03
	2.32	1.04	0.18	1.08	5.05	246.79

Table A.35: Results for SD-EE for *storm* ($\psi = 16095000.00$).

λ	Obj Val	$\mathbb{E}[f(x, \tilde{\omega})]$	$\mathbb{E}[\nu(\tilde{\omega})]$	UB	CPU	Iteration
0	15419746.67	15578190.00	8498.31	15578190.00	156.94	720.67
	20682.84	50816.22	4393.71	50816.22	83.05	139.59
0.1	15415116.67	15600636.67	10325.17	15601680.00	197.04	694.27
	28036.60	62845.43	6212.93	63440.48	344.69	238.83
0.2	15417050.00	15568406.67	7789.96	15569896.67	191.33	715.87
	17536.68	41499.32	3350.86	42098.40	236.56	212.55
0.3	15420573.33	15569030.00	7700.01	15571330.00	174.22	711.97
	21650.57	38121.63	2686.25	38896.23	131.97	172.23
0.4	15417566.67	15596040.00	9642.59	15599896.67	157.07	719.37
	24212.44	50237.69	4323.41	51928.97	85.27	135.94
0.5	15416063.33	15587026.67	8966.41	15591510.00	161.50	697.50
	24732.05	48558.38	4018.83	50493.08	158.39	196.18
0.6	15421076.67	15595036.67	9422.34	15600683.33	217.11	767.90
	19930.26	48701.06	3940.82	51025.18	210.67	232.20
0.7	15419156.67	15576753.33	7941.04	15582310.00	170.24	730.00
	27802.06	30135.77	1898.27	31390.07	129.20	168.36
0.8	15421436.67	15578403.33	8271.27	15585016.67	202.01	748.13
	21919.82	46861.12	3348.14	49468.58	222.35	217.70
0.9	15416430.00	15582686.67	8664.87	15590480.00	126.34	657.20
	23726.46	47875.48	3657.28	51098.56	67.29	114.59
1	15421143.33	15575190.00	8102.88	15583293.33	243.57	783.47
	24703.14	42513.64	3122.72	45554.49	225.88	239.31

APPENDIX B

CONVERGENCE PROOFS FOR SD

B.1 Convergence Proof for SD-QDEV

For SD-QDEV, we follow similar approach as that of SD-EE by proving that the approximation of the recourse function generated during the execution of the algorithm uniformly converges and the accumulation point of the candidate solutions is an optimal solution. We begin by proving that the variable $\nu(\omega)$ for realization ω used for computing the excess over and below the α -quantile ψ is always finite for $x \in X$. Therefore, the assumption **A2** holds for two-stage SLP with QDEV. For the convergence results of SD-QDEV and SD-CVaR, iterate x^k will encompass both the decision variables at first stage and the quantile variable ψ generated at iteration k .

COROLLARY B.1.1. *Suppose that assumptions A1-A2 hold, then for any $x \in X$ and $\omega \in \Omega$, the dispersion statistic for QDEV $\mathbb{E}[\nu(\tilde{\omega})] < \infty$.*

Proof. By assumption (A1), $X \subseteq \mathbb{R}_+^{n_1}$ is a compact set and from Equation (3.1) we have cost vector $c \in \mathbb{R}^{n_1}$. Therefore the elements of set $\{c^\top x^k\}_{k=1}^\infty$ will always be finite.

From assumption (A2), for any given $x \in X$ we have $\mathbb{E}[f(x, \tilde{\omega})] < \infty$. Hence from Equation (3.2), for any $\omega \in \Omega$, we have $q^\top y(\omega) < \infty$.

Since α -quantile $\psi \in \mathbb{R}$, from Proposition 3.4.2 we have

$$-c^\top x^k - q^\top y(\omega) + \psi + \nu(\omega) \geq 0 \quad \forall \omega \in \Omega. \quad (\text{B.1})$$

Hence, we have $\mathbb{E}[\nu(\tilde{\omega})] < \infty$. □

In the following statements for denoting limits we use \lim , for denoting upper limits and lower limits we use $\overline{\lim}$ and $\underline{\lim}$ respectively. The functions $h_k(x, \omega)$ and $h(x, \omega)$ are defined as follows:

$$h_k(x, \omega) = \max \{(\pi)^\top [r(\omega^k) - T(\omega^k)x^k] + \phi[c^\top x - \psi] \mid \pi \in V^k, \phi \in U^k\},$$

$h(x, \omega) = \max \{(\pi)^\top [r(\omega^k) - T(\omega^k)x^k] + \phi[c^\top x - \psi] \mid \pi \in V, \phi \in U\}$, where set V and U are collections of all dual vertices of the subproblem.

LEMMA B.1.2. *Suppose that assumptions A1-A2 hold, then the sequence $\{h_k\}_{k=1}^\infty$ of functions $h_k(x, \omega)$, converges uniformly on $X \times \Omega$.*

Proof. Note that the set $V^k \subseteq V^{k+1} \subseteq V$ and set $U^k \subseteq U^{k+1} \subseteq U$, This implies that $h_k(x, \omega) \leq h_{k+1}(x, \omega) \leq h(x, \omega)$ for all k and for all $(x, \omega) \in X \times \Omega$. Since $\{h_k\}_{k=1}^\infty$ increases monotonically and is bounded from above by the function $h(x, \omega)$, it follows that $\{h_k\}_{k=1}^\infty$ converges pointwise to some function $g(x, \omega) \leq h(x, \omega)$. Since set $V^k \subseteq V^{k+1} \subseteq V$ and set $U^k \subseteq U^{k+1} \subseteq U$ for all k ,

$$\bar{V} = \lim_{k \rightarrow \infty} V_k \subseteq V, \quad (\text{B.2})$$

and

$$\bar{U} = \lim_{k \rightarrow \infty} U_k \subseteq U. \quad (\text{B.3})$$

By assumption (A2) and Corollary B.2.1, V and U are sets of finite elements and so are \bar{V} and \bar{U} , hence

$$\begin{aligned} g(x, \omega) &= \lim_{k \rightarrow \infty} h_k(x, \omega) \\ &= \lim_{k \rightarrow \infty} \{ \text{Max} \quad \{ \pi^\top [r(\omega) - T(\omega)x] + \phi[c^\top x - \psi] \mid \pi \in V_k, \phi \in U_k \} \} \\ &= \text{Max} \quad \{ \pi^\top [r(\omega) - T(\omega)x] + \phi[c^\top x - \psi] \mid \pi \in \bar{V}, \phi \in \bar{U} \}. \end{aligned} \quad (\text{B.4})$$

Therefore, from the statements (B.2), (B.3) and (B.4), we can conclude that $\{h_k\}_{k=1}^\infty$ converges uniformly to the function $g(x, \omega)$, since $\{h_k\}_{k=1}^\infty$ is a monotone sequence of continuous functions and $X \times \Omega$ is a compact set. \square

Let Ψ be the set of iterates $\psi^k \in \mathbb{R}$ used for estimating α -quantile. Since, Ψ is a closed and bounded set by Heine-Borel theorem Ψ is also a compact set. Therefore every sequence that belongs to Ψ will have a converging subsequence.

THEOREM B.1.3. *Let $\{x^{k_n}\}_{n=1}^\infty$ be an infinite subsequence of $\{x^k\}_{k=1}^\infty$. Suppose that assumptions A1-A4 hold and if $x^{k_n} \rightarrow \hat{x}$, then with probability one*

$$\frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n}) \rightarrow \mathbb{E}[h(\hat{x}, \tilde{\omega})].$$

Proof. From the equation (4.11) and step 2 of the SD-QDEV algorithm, we know that

$$h_{k_n}(x^{k_n}, \omega^t) = \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n})$$

and

$$\frac{1}{k_n} \sum_{t=1}^{k_n} h_{k_n}(x^{k_n}, \omega^t) = \frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n}).$$

By Lemma B.1.2, there exists a function $g(x, \omega) \leq h(x, \omega)$ such that $\{h_{k_n}\}_{n=1}^\infty$ converges uniformly to $g(x, \omega)$. Thus, since we have

$$\frac{1}{k_n} \sum_{t=1}^{k_n} [h_{k_n}(x^{k_n}, \omega^t) - g(\hat{x}, \omega^t)] \rightarrow 0 \quad \text{and} \quad \frac{1}{k_n} \sum_{t=1}^{k_n} h(x, \omega^t) \rightarrow \mathbb{E}[h(x, \tilde{\omega})],$$

it is sufficient to show that $g(\hat{x}, \omega^t) = h(\hat{x}, \omega^t)$ with probability one. Since $h(x, \omega)$ is a continuous function and $\{h_{k_n}\}_{n=1}^\infty$ is a uniformly convergent sequence of continuous function, for every $\epsilon > 0$ there exist $\delta > 0$ and $N < \infty$ such that

$$|(\hat{x}, \omega^t) - (x, \omega)| < \delta \Rightarrow |h(\hat{x}, \omega^t) - h(x, \omega)| < \frac{\epsilon}{3} \quad \forall n \geq N$$

and

$$|h_{k_n}(\hat{x}, \omega^t) - h_{k_n}(x, \omega)| < \frac{\epsilon}{3} \quad \forall n \geq N.$$

Thus, since $x^{k_n} \rightarrow \hat{x}$, for every $\epsilon > 0$ there exist a further subsequence $\{(x^{k'_n}, \omega^{k'_n})\}_{n=1}^\infty$ and $K < \infty$ such that

$$|h(\hat{x}, \omega^t) - h(\hat{x}, \omega^{k'_n})| < \epsilon/3,$$

$$|h(\hat{x}, \omega^{k'_n}) - h(x^{k'_n}, \omega^{k'_n})| < \epsilon/3$$

and

$$|h_{k'_n}(x^{k'_n}, \omega^{k'_n}) - h_{k'_n}(x^{k'_n}, \omega^t)| < \epsilon/3,$$

for all $k'_n \geq K$. By construction we have $h_{k'_n}(x^{k'_n}, \omega^{k'_n}) = h(x^{k'_n}, \omega^{k'_n})$. Thus, for every $\epsilon > 0$ there exist a subsequence $\{x^{k'_n}\}_{n=1}^\infty$ and $K < \infty$ such that

$$\begin{aligned} |h(\hat{x}, \omega^t) - h_{k'_n}(x^{k'_n}, \omega^t)| &\leq |h(\hat{x}, \omega^t) - h(\hat{x}, \omega^{k'_n})| \\ &\quad + |h(\hat{x}, \omega^{k'_n}) - h(x^{k'_n}, \omega^{k'_n})| \\ &\quad + |h(x^{k'_n}, \omega^{k'_n}) - h_{k'_n}(x^{k'_n}, \omega^t)| < \epsilon, \end{aligned}$$

for all $k'_n \geq K$. Hence, by the uniqueness of the sequential limit, it follows that $g(\hat{x}, \omega^t) = h(\hat{x}, \omega^t)$. Therefore by probability one, we have

$$\frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n}) \rightarrow \mathbb{E}[h(\hat{x}, \tilde{\omega})].$$

Also since $h(x, \omega^t) = \operatorname{argmax}\{\pi(r(\omega^t) - T(\omega^t)x) + \phi(c^\top x - \psi) | \pi \in V, \phi \in U\}$, $V^k \subseteq V$, and $U^k \subseteq U \forall k$, it follows that

$$\begin{aligned} c^\top x + \lambda\psi + \frac{1}{k_n} \sum_{t=1}^{k_n} h(x, \omega^t) &\geq c^\top x + \lambda\psi + \frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n}) \\ &= c^\top x + \lambda\psi + \alpha_{k_n}^{k_n} + \beta_{k_n}^{k_n} x \quad x \in X. \end{aligned}$$

□

THEOREM B.1.4. *Suppose that assumptions A1-A4 hold, then there exists a subsequence $\{x^{k_n}\}_{n=1}^\infty$ of $\{x^k\}_{k=1}^\infty$, such that $\lim_{n \rightarrow \infty} [F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n})] = 0$, with probability one.*

Proof. If assumption (A2) is satisfied, then for every $\omega \in \Omega$ there exist $M(\omega) \in \mathbb{R}_+$, such that $|h(x^1, \omega) - h(x^2, \omega)| \leq M(\omega) \|x^1 - x^2\|$ for all $x^1, x^2 \in X$. Let $\epsilon > 0$ be given, let $M = \mathbb{E}[M(\omega)]$, let $r = \frac{\epsilon}{2M}$, and let $B_r(x)$ denote an open ball of radius r centered at x . Then $\cup_{x \in X} B_r(x)$ is an open cover of X . Since X is a compact set, there exist $N_\epsilon \leq \infty$ and $\{x_i\}_{i=1}^{N_\epsilon} \subset X$ such that $X \subset \cup_{i=1}^{N_\epsilon} B_r(x_i)$. Moreover, since $\{x^k\} \subset X$, it follows that each iterate is contained in one or more of the open balls $\{B_r(x_i)\}_{i=1}^{N_\epsilon}$. Thus, there exist two sequence of indices, $\{k_n\}$ and $\{t_n\}$ such that

$$0 < k_n - t_n \leq N_\epsilon + 1 \quad \text{and} \quad \|x^{k_n} - x^{t_n}\| < r.$$

By assumption (A1), we know that X is a compact set. Thus, without loss of generality we may assume that

$$\lim_{n \rightarrow \infty} x^{k_n} = \hat{x}_k \quad \text{and} \quad \lim_{n \rightarrow \infty} x^{t_n} = \hat{x}_t,$$

where \hat{x}_k and \hat{x}_t are accumulation points of sequences x^{k_n} and x^{t_n} respectively.

Now, in iteration k the cutting plane generated during iteration t appears as (please refer to step 2(d) of the SD-QDEV algorithm)

$$\alpha_t^k + \beta_t^k x + \gamma_t^k \psi = \frac{t}{k} (\alpha_t^t + \beta_t^t x + \gamma_t^t \psi) + \frac{k-t}{k} L. \quad (\text{B.5})$$

As per the step 2 of algorithm we have

$$\eta_{k-1}(x^k) = \text{Max}\{\alpha_t^{k-1} + \beta_t^{k-1}x^k + \gamma_t^k\psi^k | t = 1, \dots, k-1\}. \quad (\text{B.6})$$

Therefore, using equation (B.5) we can rewrite equation (B.6) as follows

$$\eta_{k-1}(x^k) \geq \frac{t}{k-1}(\alpha_t^t + \beta_t^t x^k + \gamma_t^k \psi^k) + (1 - \frac{t}{k-1})L \quad \forall t = 1, \dots, k-1.$$

Thus,

$$\begin{aligned} \eta_{k_n-1}(x^{k_n}) &\geq \frac{t_n}{k_n-1}(\alpha_{t_n}^{t_n} + \beta_{t_n}^{t_n} x^{k_n} + \gamma_{t_n}^{k_n} \psi^{k_n}) + (1 - \frac{t_n}{k_n-1})L \\ &= \frac{t_n}{k_n-1}(\alpha_{t_n}^{t_n} + \beta_{t_n}^{t_n} x^{t_n} + \gamma_{t_n}^{t_n} \psi^{t_n}) + \frac{t_n}{k_n-1} \beta_{t_n}^{t_n} (x^{k_n} - x^{t_n}) \\ &\quad + \frac{t_n}{k_n-1} \gamma_{t_n}^{t_n} (\psi^{k_n} - \psi^{t_n}) + (1 - \frac{t_n}{k_n-1})L \\ &= \frac{t_n}{k_n-1} \eta_{t_n}(x^{t_n}) + \frac{t_n}{k_n-1} \beta_{t_n}^{t_n} (x^{k_n} - x^{t_n}) + \frac{t_n}{k_n-1} \gamma_{t_n}^{t_n} (\psi^{k_n} - \psi^{t_n}) + (1 - \frac{t_n}{k_n-1})L, \end{aligned} \quad (\text{B.7})$$

where the last equality follows the fact that $\eta_k(x^k) = \alpha_k^k + \beta_k^k x^k + \gamma_k^k \psi^k$ for all k . Furthermore by definition,

$$\begin{aligned} F_k(x^k) - F_{k-1}(x^k) &= c^\top x^k + \eta_k(x^k) - c^\top x^k - \eta_{k-1}(x^k) \\ &= \eta_k(x^k) - \eta_{k-1}(x^k). \end{aligned} \quad (\text{B.8})$$

Using equations (B.7) and (B.8) we have

$$\begin{aligned}
F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) &\leq \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n - 1}(\eta_{t_n}(x^{t_n}) + \beta_{t_n}^{t_n}(x^{k_n} - x^{t_n})) \\
&\quad + \gamma_{t_n}^{t_n}(\psi^{k_n} - \psi^{t_n}) - \left(1 - \frac{t_n}{k_n - 1}\right)L.
\end{aligned}$$

By construction,

$$0 < k_n - t_n \leq N_\epsilon + 1 < \infty,$$

we have

$$\begin{aligned}
\|\beta_{t_n}^{t_n}\| &\leq \frac{1}{t_n} \sum_{t=1}^{t_n} M(\omega^t), \\
\lim_{n \rightarrow \infty} \frac{t_n}{k_n - 1} &= 1 \quad \text{and} \\
\lim_{n \rightarrow \infty} \left(1 - \frac{t_n}{k_n - 1}\right)L &= 0.
\end{aligned}$$

Moreover, Theorem B.1.3 ensures that $\eta_{k_n}(x^{k_n}) \rightarrow \mathbb{E}[h(\hat{x}_k, \tilde{\omega})]$ and $\eta_{t_n}(x^{t_n}) \rightarrow \mathbb{E}[h(\hat{x}_t, \tilde{\omega})]$ with probability one. It follows that

$$\begin{aligned}
0 &\leq \underline{\lim}_{k \rightarrow \infty} F_k(x^k) - F_{k-1}(x^k) \\
&\leq \underline{\lim}_{n \rightarrow \infty} F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \\
&\leq \underline{\lim}_{n \rightarrow \infty} \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n - 1} (\eta_{t_n}(x^{t_n}) + \|\beta_{t_n}^{t_n}\| \|x^{k_n} - x^{t_n}\|) - (1 - \frac{t_n}{k_n - 1})L. \\
&\leq \underline{\lim}_{n \rightarrow \infty} \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n - 1} (\eta_{t_n}(x^{t_n}) + \frac{1}{t_n} \sum_{t=1}^{t_n} M(\omega^t) \|x^{k_n} - x^{t_n}\|) - (1 - \frac{t_n}{k_n - 1})L. \\
&= \mathbb{E}[h(\hat{x}_k, \tilde{\omega})] - \mathbb{E}[h(\hat{x}_t, \tilde{\omega})] + M \|\hat{x}_k - \hat{x}_t\| \\
&\leq |\mathbb{E}[h(\hat{x}_k, \tilde{\omega})] - \mathbb{E}[h(\hat{x}_t, \tilde{\omega})]| + M \|\hat{x}_k - \hat{x}_t\| \\
&\leq 2M \|\hat{x}_k - \hat{x}_t\| \\
&\leq 2M \left(\frac{\epsilon}{2M} \right) \\
&= \epsilon.
\end{aligned}$$

Thus, for every $\epsilon > 0$, $0 \leq \lim_{n \rightarrow \infty} F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \leq \epsilon$ and hence the result. □

THEOREM B.1.5. *There exist a subsequence $\{x^{k_n}\}_{n=1}^{\infty}$ of $\{x^k\}_{k=1}^{\infty}$, such that every accumulation point of $\{x^{k_n}\}_{n=1}^{\infty}$ is an optimal solution x^* with probability one.*

Proof. From Theorem 4.4.3, we know that there exist subsequence $\{x^{k_n}\}_{n=1}^{\infty}$ such that $\lim_{n \rightarrow \infty} h_{k_n}(x^{k_n}) - h_{k_n-1}(x^{k_n}) = 0$. Let $\{x^{k_n}\}_{n \in N}$ be a subsequence such that $\lim_{n \in N} x^{k_n} = \hat{x}$. By assumption (A1) we will always have accumulation point $\hat{x} \in \mathcal{X}$ thus for an optimal solution x^* we have

$$F(x^*) \leq F(\hat{x}), \tag{B.9}$$

where $F(x) = (1 - \lambda\epsilon_1)c^\top x + \lambda\epsilon_1\psi + \mathbb{E}[h(x, \omega)]$, and also by construction we have

$$\lim_{k \in K} F_k(x^*) \leq (1 - \lambda\epsilon_1)c^\top x^* + \lambda\epsilon_1\psi^* + \mathbb{E}[h(x^*, \omega)] = F(x^*). \quad (\text{B.10})$$

As per the step 4 of algorithm we know that x^k minimizes F_{k-1} , therefore

$$F_{k-1}(x^k) \leq F_{k-1}(x^*). \quad (\text{B.11})$$

Now using Theorem B.1.3 and result $\lim_{n \in N} F_{k_n}(x^{k_n}) = F(\hat{x})$ we get $\lim_{n \in N} F_{k_{n-1}}(x^{k_n}) = F(\hat{x})$, with probability one. Combining equations (B.9), (B.10) and (B.11) we have

$$F(x^*) \leq F(\hat{x}) = \lim_{n \in N} F_{k_{n-1}}(x^{k_n}) \leq \lim_{k \in K} F_k(x^*) \leq F(x^*).$$

Hence accumulation point of subsequence $\{x^{k_n}\}_{n=1}^\infty$ is an optimal solution x^* with probability one. \square

We have now proved that the SD-QDEV algorithm will generate an optimal solution with probability one. Next we need to show that with probability one, there exists at least one optimal accumulation point of the sequence of incumbents $\{\bar{x}^k\}$. We will use Theorems B.1.3 and B.1.4 along with the incumbent test used in step 3 of SD-QDEV algorithm to prove this result. We begin by noting that for all k ,

$$(1 - \lambda\epsilon_1)c^\top \bar{x}^k + \lambda\epsilon_1\bar{\psi}^k + \alpha_{i_k}^k + \beta_{i_k}^k x \leq F_k(\bar{x}^k) \leq (1 - \lambda\epsilon_1)c^\top \bar{x}^k + \lambda\epsilon_1\bar{\psi}^k + \frac{1}{k} \sum_{t=1}^k h(\bar{x}^k, \omega^t).$$

Since $h(x, \omega)$ is continuous in x for all $\omega \in \Omega$, if $\{\bar{x}^{k_n}\}_{n=1}^\infty$ is a subsequence such that $\bar{x}^{k_n} \rightarrow \bar{x}$, $h(\bar{x}^{k_n}, \omega^t) \rightarrow h(\bar{x}, \omega^t)$ for all t . Thus,

$$\begin{aligned}
\underline{\lim}_{n \rightarrow \infty} (1 - \lambda \epsilon_1) c^\top \bar{x}^{k_n} + \lambda \epsilon_1 \bar{\psi}^{k_n} + \alpha_{i_k}^{k_n} + \beta_{i_{k_n}}^{k_n} x &\leq \underline{\lim}_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) \\
&\leq \overline{\lim}_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) \\
&\leq \overline{\lim}_{n \rightarrow \infty} c^\top \bar{x}^{k_n} + \lambda \epsilon_1 \bar{\psi}^{k_n} + \frac{1}{k_n} \sum_{t=1}^{k_n} h(\bar{x}^{k_n}, \omega^t).
\end{aligned}$$

With probability one, both the upper and lower limits described above are $F(\bar{x})$ (please see Theorem B.1.3). Thus, it follows that

$$\lim_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) = F(\bar{x}).$$

Similarly, if $\{\bar{x}^{k_n}\}_{n=1}^\infty \rightarrow \bar{x}$, then the nature of the update mechanism described in step 2 of algorithm and Theorem B.1.4 ensures that

$$\lim_{n \rightarrow \infty} F_{k_n+1}(\bar{x}^{k_n}) = F(\bar{x}).$$

The above results can be summarized and formerly stated in the following corollary.

COROLLARY B.1.6. *Let $\{\bar{x}^k\}$ denote the sequence of incumbent solutions, and let $\{\bar{x}^{k_n}\}_{n=1}^\infty$ be an infinite subsequence such that $\{\bar{x}^{k_n}\} \rightarrow \bar{x}$. If the assumptions (A1)-(A4) hold, then with probability one*

$$\lim_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) = \lim_{n \rightarrow \infty} F_{k_n+1}(\bar{x}^{k_n}) = F(\bar{x}).$$

To establish that an optimal accumulation point of the incumbent sequence exists, next we explore the implication of the incumbent test described in step 3 of the algorithm.

LEMMA B.1.7. *Suppose that assumption A1-A4 hold. Let $\theta^k = F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1})$, and let $\{k_n\}_{n \in \mathbb{N}}$ represent the sequence of iterations at which the incumbent is changed. If N is finite, then $\overline{\lim}_{k \rightarrow \infty} \theta^k = 0$, with probability one. Otherwise, $\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} = 0$ with probability one.*

Proof. By definition, $\theta^k = F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1}) \leq 0$ for all k . If N is a finite set, there exist \bar{x} and $K < \infty$ such that $\bar{x}^k = \bar{x}$ for all $k \geq K$ and thus

$$F_k(x^k) - F_k(\bar{x}) \geq \delta[F_{k-1}(x^k) - F_{k-1}(\bar{x})] = \delta\theta \quad \forall k \geq K.$$

By Theorem B.1.4 and Corollary B.1.6, there exist a subsequence indexed by set \mathcal{K} such that

$$\begin{aligned} \lim_{k \in \mathcal{K}} x^k &= \hat{x} \\ \lim_{k \in \mathcal{K}} F_k(x^k) &= f(\hat{x}), \quad \lim_{k \in \mathcal{K}} F_k(\bar{x}) = f(\bar{x}), \\ \lim_{k \in \mathcal{K}} F_{k-1}(x^k) &= f(\hat{x}), \quad \lim_{k \in \mathcal{K}} F_{k-1}(\bar{x}) = f(\bar{x}), \end{aligned}$$

with probability one. Thus,

$$\begin{aligned} \lim_{k \in \mathcal{K}} \{F_k(x^k) - F_k(\bar{x})\} &\geq \delta[\lim_{k \in \mathcal{K}} \{F_{k-1}(x^k) - F_{k-1}(\bar{x})\}] \\ &\Rightarrow F(\hat{x}) - F(\bar{x}) \geq \delta[F(\hat{x}) - F(\bar{x})] \\ &= \lim_{k \in \mathcal{K}} \delta\theta^k, \end{aligned}$$

with probability one. Now since $\delta \in (0, 1)$ and $\theta^k \leq 0$ for all k , it follows that $F(\hat{x}) - F(\bar{x}) = 0$

and thus $\lim_{k \in \mathcal{K}} \delta \theta^k \leq 0$, with probability one. Now suppose N is not a finite set. By hypothesis,

$$F_{k_n}(x^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) < \delta[F_{k_n-1}(x^{k_n}) - F_{k_n-1}(\bar{x}^{k_n-1})] = \delta \theta^{k_n} \leq 0 \quad \forall n.$$

By definition of the subsequence $\{k_n\}$, we note that $\bar{x}^{k_n-1} = \bar{x}^{k_n-1}$. Therefore

$$F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) \leq \delta \theta^{k_n} \leq 0 \quad \forall n.$$

Thus,

$$\begin{aligned} \frac{1}{m} \sum_{n=1}^m \{F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1})\} &\leq \frac{\delta}{m} \sum_{n=1}^m \theta^{k_n} \leq 0 \quad \forall m \\ &\Rightarrow \frac{1}{m} \left\{ \left(\sum_{n=1}^{m-1} F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) \right) + (F_{k_m}(\bar{x}^{k_m}) - F_{k_1}(\bar{x}^{k_0})) \right\} \\ &\leq \frac{\delta}{m} \sum_{n=1}^m \theta^{k_n} \leq 0 \quad \forall m. \end{aligned}$$

Assumptions (A1)-(A4) ensure that there exist $M < \infty$, such that

$$|F_{k_m}(\bar{x}^{k_m}) - F_{k_1}(\bar{x}^{k_0})| < M \quad \forall m.$$

Thus, since $\bar{x}^{k_n} = \bar{x}^{k_{n+1}-1}$, the left hand side converges to zero with probability one as m approaches ∞ . Thus,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} = 0.$$

□

THEOREM B.1.8. *Suppose that assumptions A1-A4 are satisfied. Let $\{\bar{x}^k\}_{k=1}^{\infty}$ represent the sequence of incumbents and let X^* represent set of optimal solutions. Then there exist a subsequence*

$\{\bar{x}^k\}_{k \in K}$ for which every accumulation point is contained in X^* , with probability one.

Proof. Let $\{k_n\}_{n \in N}$ represent the sequence of iterations at which the incumbent is changed. Note that if N is infinite set,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} \leq \overline{\lim}_{n \rightarrow \infty} \theta^{k_n} \leq 0.$$

Thus, as a result of lemma B.1.7, whether N is finite or infinite, there exist a subsequence indexed by set \mathcal{K} such that

$$\lim_{k \in \mathcal{K}} \theta^{k+1} = 0.$$

Note that

$$\theta^{k+1} = F_k(x^{k+1}) - F_k(\bar{x}^k) \leq F_k(x^*) - F_k(\bar{x}^k) \quad \forall k \in K.$$

Thus as a result of Corollary 6.5, It follows that if \bar{x} is an accumulation point of $\{\bar{x}^k\}_{k \in \mathcal{K}}$, then

$$\begin{aligned} F(\bar{x}) &\leq \overline{\lim}_{k \in \mathcal{K}} F_k(x^*) \\ &\leq c^\top x^* + \lim_{k \in \mathcal{K}} \frac{1}{k} \sum_{t=1}^k h(x^*, \omega^t) \\ &\leq F(x^*), \end{aligned}$$

and thus, $\bar{x} \in X^*$, with probability one. □

B.2 Convergence Proof for SD-CVaR

For convergence proof of SD-CVaR, We begin by proving that the variable $\nu(\omega)$ for realization ω used for computing the excess over the α -quantile ψ is always finite for $x \in X$ and the iterate $\psi^k \in \mathbb{R}$ used for estimating α -quantile belong to a compact set Ψ . Therefore, the assumption **A2** holds for two-stage SLP with CVaR.

COROLLARY B.2.1. *Suppose that assumptions A1-A2 hold, then for any $x \in X$ and $\omega \in \Omega$, the dispersion statistic for CVaR $\mathbb{E}[\nu(\tilde{\omega})] < \infty$.*

Proof. By assumption (A1), $X \subseteq \mathbb{R}_+^{n_1}$ is a compact set and from Equation (3.1) we have cost vector $c \in \mathbb{R}^{n_1}$. Therefore the elements of set $\{c^\top x^k\}_{k=1}^\infty$ will always be finite.

From assumption (A2), for any given $x \in X$ we have $\mathbb{E}[f(x, \tilde{\omega})] < \infty$. Hence from Equation (3.2), for any $\omega \in \Omega$, we have $q^\top y(\omega) < \infty$.

Since α -quantile $\psi \in \mathbb{R}$, from Proposition 3.4.3 we have

$$-c^\top x^k - q^\top y(\omega) + \nu(\omega) \geq \psi \quad \forall \omega \in \Omega. \quad (\text{B.12})$$

Hence, we have $\mathbb{E}[\nu(\tilde{\omega})] < \infty$. □

In the following statements for denoting limits we use \lim , for denoting upper limits and lower limits we use $\overline{\lim}$ and $\underline{\lim}$ respectively. The functions $h_k(x, \omega)$ and $h(x, \omega)$ are defined as follows:

$$h_k(x, \omega) = \max \{(\pi)^\top [r(\omega^k) - T(\omega^k)x^k] + \phi[c^\top x - \psi] \mid \pi \in V^k, \phi \in U^k\},$$

$$h(x, \omega) = \max \{(\pi)^\top [r(\omega^k) - T(\omega^k)x^k] + \phi[c^\top x - \psi] \mid \pi \in V, \phi \in U\}, \text{ where set } V \text{ and } U$$

are collections of all dual vertices of the subproblem.

LEMMA B.2.2. *Suppose that assumptions A1-A2 hold, then the sequence $\{h_k\}_{k=1}^\infty$ of functions $h_k(x, \omega)$, converges uniformly on $X \times \Omega$.*

Proof. Note that the set $V^k \subseteq V^{k+1} \subseteq V$ and set $U^k \subseteq U^{k+1} \subseteq U$, This implies that $h_k(x, \omega) \leq h_{k+1}(x, \omega) \leq h(x, \omega)$ for all k and for all $(x, \omega) \in X \times \Omega$. Since $\{h_k\}_{k=1}^\infty$ increases monotonically

and is bounded from above by the function $h(x, \omega)$, it follows that $\{h_k\}_{k=1}^{\infty}$ converges pointwise to some function $g(x, \omega) \leq h(x, \omega)$. Since set $V^k \subseteq V^{k+1} \subseteq V$ and set $U^k \subseteq U^{k+1} \subseteq U$ for all k ,

$$\bar{V} = \lim_{k \rightarrow \infty} V_k \subseteq V, \quad (\text{B.13})$$

and

$$\bar{U} = \lim_{k \rightarrow \infty} U_k \subseteq U. \quad (\text{B.14})$$

By assumption (A2) and Corollary B.2.1, V and U are sets of finite elements and so are \bar{V} and \bar{U} , hence

$$\begin{aligned} g(x, \omega) &= \lim_{k \rightarrow \infty} h_k(x, \omega) \\ &= \lim_{k \rightarrow \infty} \{ \text{Max} \quad \{ \pi^\top [r(\omega) - T(\omega)x] + \phi [c^\top x - \psi] \mid \pi \in V_k, \phi \in U_k \} \} \\ &= \text{Max} \quad \{ \pi^\top [r(\omega) - T(\omega)x] + \phi [c^\top x - \psi] \mid \pi \in \bar{V}, \phi \in \bar{U} \}. \end{aligned} \quad (\text{B.15})$$

Therefore, from the statements (B.13), (B.14) and (B.15), we can conclude that $\{h_k\}_{k=1}^{\infty}$ converges uniformly to the function $g(x, \omega)$, since $\{h_k\}_{k=1}^{\infty}$ is a monotone sequence of continuous functions and $X \times \Omega$ is a compact set. \square

Let Ψ be the set of iterates $\psi^k \in \mathbb{R}$ used for estimating α -quantile. Since, Ψ is a closed and bounded set by Heine-Borel theorem Ψ is also a compact set. Therefore every sequence that belongs to Ψ will have a converging subsequence.

THEOREM B.2.3. *Let $\{x^{k_n}\}_{n=1}^{\infty}$ be an infinite subsequence of $\{x^k\}_{k=1}^{\infty}$. Suppose that assumptions A1-A4 hold and if $x^{k_n} \rightarrow \hat{x}$, then with probability one*

$$\frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n}) \rightarrow \mathbb{E}[h(\hat{x}, \tilde{\omega})].$$

Proof. From the equation (4.18) and step 2 of the SD-CvaR algorithm, we know that

$$h_{k_n}(x^{k_n}, \omega^t) = \pi_t^{k_n}(r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n}(c^\top x^{k_n} - \psi^{k_n})$$

and

$$\frac{1}{k_n} \sum_{t=1}^{k_n} h_{k_n}(x^{k_n}, \omega^t) = \frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n}(r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n}(c^\top x^{k_n} - \psi^{k_n}).$$

By Lemma B.2.2, there exists a function $g(x, \omega) \leq h(x, \omega)$ such that $\{h_{k_n}\}_{n=1}^\infty$ converges uniformly to $g(x, \omega)$. Thus, since we have

$$\frac{1}{k_n} \sum_{t=1}^{k_n} [h_{k_n}(x^{k_n}, \omega^t) - g(\hat{x}, \omega^t)] \rightarrow 0 \quad \text{and} \quad \frac{1}{k_n} \sum_{t=1}^{k_n} h(x, \omega^t) \rightarrow \mathbb{E}[h(x, \tilde{\omega})],$$

it is sufficient to show that $g(\hat{x}, \omega^t) = h(\hat{x}, \omega^t)$ with probability one. Since $h(x, \omega)$ is a continuous function and $\{h_{k_n}\}_{n=1}^\infty$ is a uniformly convergent sequence of continuous function, for every $\epsilon > 0$ there exist $\delta > 0$ and $N < \infty$ such that

$$|(\hat{x}, \omega^t) - (x, \omega)| < \delta \Rightarrow |h(\hat{x}, \omega^t) - h(x, \omega)| < \frac{\epsilon}{3} \quad \forall n \geq N$$

and

$$|h_{k_n}(\hat{x}, \omega^t) - h_{k_n}(x, \omega)| < \frac{\epsilon}{3} \quad \forall n \geq N.$$

Thus, since $x^{k_n} \rightarrow \hat{x}$, for every $\epsilon > 0$ there exist a further subsequence $\{(x^{k'_n}, \omega^{k'_n})\}_{n=1}^\infty$ and $K < \infty$ such that

$$|h(\hat{x}, \omega^t) - h(\hat{x}, \omega^{k'_n})| < \epsilon/3,$$

$$|h(\hat{x}, \omega^{k'_n}) - h(x^{k'_n}, \omega^{k'_n})| < \epsilon/3$$

and

$$|h_{k'_n}(x^{k'_n}, \omega^{k'_n}) - h_{k'_n}(x^{k'_n}, \omega^t)| < \epsilon/3,$$

for all $k'_n \geq K$. By construction we have $h_{k'_n}(x^{k'_n}, \omega^{k'_n}) = h(x^{k'_n}, \omega^{k'_n})$. Thus, for every $\epsilon > 0$ there exist a subsequence $\{x^{k'_n}\}_{n=1}^\infty$ and $K < \infty$ such that

$$\begin{aligned} |h(\hat{x}, \omega^t) - h_{k'_n}(x^{k'_n}, \omega^t)| &\leq |h(\hat{x}, \omega^t) - h(\hat{x}, \omega^{k'_n})| \\ &\quad + |h(\hat{x}, \omega^{k'_n}) - h(x^{k'_n}, \omega^{k'_n})| \\ &\quad + |h(x^{k'_n}, \omega^{k'_n}) - h_{k'_n}(x^{k'_n}, \omega^t)| < \epsilon, \end{aligned}$$

for all $k'_n \geq K$. Hence, by the uniqueness of the sequential limit, it follows that $g(\hat{x}, \omega^t) = h(\hat{x}, \omega^t)$. Therefore by probability one, we have

$$\frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n}) \rightarrow \mathbb{E}[h(\hat{x}, \tilde{\omega})].$$

Also since $h(x, \omega^t) = \operatorname{argmax}\{\pi(r(\omega^t) - T(\omega^t)x) + \phi(c^\top x - \psi) | \pi \in V, \phi \in U\}$, $V^k \subseteq V$, and $U^k \subseteq U \forall k$, it follows that

$$\begin{aligned} c^\top x + \lambda\psi + \frac{1}{k_n} \sum_{t=1}^{k_n} h(x, \omega^t) &\geq c^\top x + \lambda\psi + \frac{1}{k_n} \sum_{t=1}^{k_n} \pi_t^{k_n} (r(\omega^t) - T(\omega^t)x^{k_n}) + \phi^{k_n} (c^\top x^{k_n} - \psi^{k_n}) \\ &= c^\top x + \lambda\psi + \alpha_{k_n}^{k_n} + \beta_{k_n}^{k_n} x \quad x \in X. \end{aligned}$$

□

THEOREM B.2.4. *Suppose that assumptions A1-A4 hold, then there exists a subsequence $\{x^{k_n}\}_{n=1}^\infty$ of $\{x^k\}_{k=1}^\infty$, such that $\lim_{n \rightarrow \infty} [F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n})] = 0$, with probability one.*

Proof. If assumption (A2) is satisfied, then for every $\omega \in \Omega$ there exist $M(\omega) \in \mathbb{R}_+$, such that $|h(x^1, \omega) - h(x^2, \omega)| \leq M(\omega) \|x^1 - x^2\|$ for all $x^1, x^2 \in X$. Let $\epsilon > 0$ be given, let $M = \mathbb{E}[M(\omega)]$, let $r = \frac{\epsilon}{2M}$, and let $B_r(x)$ denote an open ball of radius r centered at x . Then $\cup_{x \in X} B_r(x)$ is an

open cover of X . Since X is a compact set, there exist $N_\epsilon \leq \infty$ and $\{x_i\}_{i=1}^{N_\epsilon} \subset X$ such that $X \subset \cup_{i=1}^{N_\epsilon} B_r(x_i)$. Moreover, since $\{x^k\} \subset X$, it follows that each iterate is contained in one or more of the open balls $\{B_r(x_i)\}_{i=1}^{N_\epsilon}$. Thus, there exist two sequence of indices, $\{k_n\}$ and $\{t_n\}$ such that

$$0 < k_n - t_n \leq N_\epsilon + 1 \quad \text{and} \quad \|x^{k_n} - x^{t_n}\| < r.$$

By assumption (A1), we know that X is a compact set. Thus, without loss of generality we may assume that

$$\lim_{n \rightarrow \infty} x^{k_n} = \hat{x}_k \quad \text{and} \quad \lim_{n \rightarrow \infty} x^{t_n} = \hat{x}_t,$$

where \hat{x}_k and \hat{x}_t are accumulation points of sequences x^{k_n} and x^{t_n} respectively.

Now, in iteration k the cutting plane generated during iteration t appears as (please refer to step 2(d) of the SD-CVaR algorithm)

$$\alpha_t^k + \beta_t^k x = \frac{t}{k}(\alpha_t^t + \beta_t^t x) + \frac{k-t}{k}L. \quad (\text{B.16})$$

As per the step 2 of algorithm we have

$$\eta_{k-1}(x^k) = \text{Max}\{\alpha_t^{k-1} + \beta_t^{k-1}x^k | t = 1, \dots, k-1\}. \quad (\text{B.17})$$

Therefore, using equation (B.16) we can rewrite equation (B.17) as follows

$$\eta_{k-1}(x^k) \geq \frac{t}{k-1}(\alpha_t^t + \beta_t^t x^k) + (1 - \frac{t}{k-1})L \quad \forall t = 1, \dots, k-1.$$

Thus,

$$\begin{aligned}
\eta_{k_n-1}(x^{k_n}) &\geq \frac{t_n}{k_n-1}(\alpha_{t_n}^{t_n} + \beta_{t_n}^{t_n}x^{k_n}) + (1 - \frac{t_n}{k_n-1})L \\
&= \frac{t_n}{k_n-1}(\alpha_{t_n}^{t_n} + \beta_{t_n}^{t_n}x^{t_n}) + \frac{t_n}{k_n-1}\beta_{t_n}^{t_n}(x^{k_n} - x^{t_n}) + (1 - \frac{t_n}{k_n-1})L \\
&= \frac{t_n}{k_n-1}\eta_{t_n}(x^{t_n}) + \frac{t_n}{k_n-1}\beta_{t_n}^{t_n}(x^{k_n} - x^{t_n}) + (1 - \frac{t_n}{k_n-1})L, \tag{B.18}
\end{aligned}$$

where the last equality follows the fact that $\eta_k(x^k) = \alpha_k^k + \beta_k^k x^k$ for all k . Furthermore by definition,

$$\begin{aligned}
F_k(x^k) - F_{k-1}(x^k) &= c^\top x^k + \lambda\psi^k + \eta_k(x^k) - c^\top x^k - \lambda\psi^k + -\eta_{k-1}(x^k) \\
&= \eta_k(x^k) - \eta_{k-1}(x^k). \tag{B.19}
\end{aligned}$$

Using equations (B.18) and (B.19) we have

$$F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \leq \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n-1}(\eta_{t_n}(x^{t_n}) + \beta_{t_n}^{t_n}(x^{k_n} - x^{t_n})) - (1 - \frac{t_n}{k_n-1})L.$$

By construction,

$$0 < k_n - t_n \leq N_\epsilon + 1 < \infty,$$

we have

$$\begin{aligned}\|\beta_{t_n}^{t_n}\| &\leq \frac{1}{t_n} \sum_{t=1}^{t_n} M(\omega^t), \\ \lim_{n \rightarrow \infty} \frac{t_n}{k_n - 1} &= 1 \quad \text{and} \\ \lim_{n \rightarrow \infty} \left(1 - \frac{t_n}{k_n - 1}\right)L &= 0.\end{aligned}$$

Moreover, Theorem B.2.3 ensures that $\eta_{k_n}(x^{k_n}) \rightarrow \mathbb{E}[h(\hat{x}_k, \tilde{\omega})]$ and $\eta_{t_n}(x^{t_n}) \rightarrow \mathbb{E}[h(\hat{x}_t, \tilde{\omega})]$ with probability one. It follows that

$$\begin{aligned}0 &\leq \underline{\lim}_{k \rightarrow \infty} F_k(x^k) - F_{k-1}(x^k) \\ &\leq \underline{\lim}_{n \rightarrow \infty} F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \\ &\leq \underline{\lim}_{n \rightarrow \infty} \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n - 1} (\eta_{t_n}(x^{t_n}) + \|\beta_{t_n}^{t_n}\| \|x^{k_n} - x^{t_n}\|) - \left(1 - \frac{t_n}{k_n - 1}\right)L. \\ &\leq \underline{\lim}_{n \rightarrow \infty} \eta_{k_n}(x^{k_n}) - \frac{t_n}{k_n - 1} (\eta_{t_n}(x^{t_n}) + \frac{1}{t_n} \sum_{t=1}^{t_n} M(\omega^t) \|x^{k_n} - x^{t_n}\|) - \left(1 - \frac{t_n}{k_n - 1}\right)L. \\ &= \mathbb{E}[h(\hat{x}_k, \tilde{\omega})] - \mathbb{E}[h(\hat{x}_t, \tilde{\omega})] + M \|\hat{x}_k - \hat{x}_t\| \\ &\leq |\mathbb{E}[h(\hat{x}_k, \tilde{\omega})] - \mathbb{E}[h(\hat{x}_t, \tilde{\omega})]| + M \|\hat{x}_k - \hat{x}_t\| \\ &\leq 2M \|\hat{x}_k - \hat{x}_t\| \\ &\leq 2M \left(\frac{\epsilon}{2M}\right) \\ &= \epsilon.\end{aligned}$$

Thus, for every $\epsilon > 0$, $0 \leq \lim_{n \rightarrow \infty} F_{k_n}(x^{k_n}) - F_{k_n-1}(x^{k_n}) \leq \epsilon$ and hence the result. □

THEOREM B.2.5. *There exist a subsequence $\{x^{k_n}\}_{n=1}^{\infty}$ of $\{x^k\}_{k=1}^{\infty}$, such that every accumulation point of $\{x^{k_n}\}_{n=1}^{\infty}$ is an optimal solution x^* with probability one.*

Proof. From Theorem B.2.3, we know that there exist subsequence $\{x^{k_n}\}_{n=1}^\infty$ such that $\lim_{n \rightarrow \infty} h_{k_n}(x^{k_n}) - h_{k_{n-1}}(x^{k_n}) = 0$. Let $\{x^{k_n}\}_{n \in N}$ be a subsequence such that $\lim_{n \in N} x^{k_n} = \hat{x}$. By assumption (A1) we will always have accumulation point $\hat{x} \in \mathcal{X}$ thus for an optimal solution x^* we have

$$F(x^*) \leq F(\hat{x}), \quad (\text{B.20})$$

where $F(x) = c^\top x + \lambda\psi + \mathbb{E}[h(x, \omega)]$, and also by construction we have

$$\lim_{k \in K} F_k(x^*) \leq c^\top x^* + \lambda\psi^* + \mathbb{E}[h(x^*, \omega)] = F(x^*). \quad (\text{B.21})$$

As per the step 4 of SD-CVaR algorithm we know that x^k minimizes F_{k-1} , therefore

$$F_{k-1}(x^k) \leq F_{k-1}(x^*). \quad (\text{B.22})$$

Now using Theorem B.2.3 and result $\lim_{n \in N} F_{k_n}(x^{k_n}) = F(\hat{x})$ we get $\lim_{n \in N} F_{k_{n-1}}(x^{k_n}) = F(\hat{x})$, with probability one. Combining equations (B.20), (B.21) and (B.22) we have

$$F(x^*) \leq F(\hat{x}) = \lim_{n \in N} F_{k_{n-1}}(x^{k_n}) \leq \lim_{k \in K} F_k(x^*) \leq F(x^*).$$

Hence accumulation point of subsequence $\{x^{k_n}\}_{n=1}^\infty$ is an optimal solution x^* with probability one. □

We have now proved that the SD-CVaR algorithm will generate an optimal solution with probability one. Next we need to show that with probability one, there exists at least one optimal accumulation point of the sequence of incumbents $\{\bar{x}^k\}$. We will use Theorems B.2.3 and B.2.4 along with the incumbent test used in step 3 of algorithm to prove this result. We begin by noting

that for all k ,

$$c^\top \bar{x}^k + \lambda \bar{\psi}^k + \alpha_{i_k}^k + \beta_{i_k}^k x \leq F_k(\bar{x}^k) \leq c^\top \bar{x}^k + \lambda \bar{\psi}^k + \frac{1}{k} \sum_{t=1}^k h(\bar{x}^k, \omega^t).$$

Since $h(x, \omega)$ is continuous in x for all $\omega \in \Omega$, if $\{\bar{x}^{k_n}\}_{n=1}^\infty$ is a subsequence such that $\bar{x}^{k_n} \rightarrow \bar{x}$, $h(\bar{x}^{k_n}, \omega^t) \rightarrow h(\bar{x}, \omega^t)$ for all t . Thus,

$$\begin{aligned} \underline{\lim}_{n \rightarrow \infty} c^\top \bar{x}^{k_n} + \lambda \bar{\psi}^{k_n} + \alpha_{i_{k_n}}^{k_n} + \beta_{i_{k_n}}^{k_n} x &\leq \underline{\lim}_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) \\ &\leq \overline{\lim}_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) \\ &\leq \overline{\lim}_{n \rightarrow \infty} c^\top \bar{x}^{k_n} + \lambda \bar{\psi}^{k_n} + \frac{1}{k_n} \sum_{t=1}^{k_n} h(\bar{x}^{k_n}, \omega^t). \end{aligned}$$

With probability one, both the upper and lower limits described above are $F(\bar{x})$ (please see Theorem B.2.3). Thus, it follows that

$$\lim_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) = F(\bar{x}).$$

Similarly, if $\{\bar{x}^{k_n}\}_{n=1}^\infty \rightarrow \bar{x}$, then the nature of the update mechanism described in step 2 of SD-CvaR algorithm and Theorem B.2.4 ensures that

$$\lim_{n \rightarrow \infty} F_{k_n+1}(\bar{x}^{k_n}) = F(\bar{x}).$$

The above results can be summarized and formerly stated in the following corollary.

COROLLARY B.2.6. *Let $\{\bar{x}^k\}$ denote the sequence of incumbent solutions, and let $\{\bar{x}^{k_n}\}_{n=1}^\infty$ be an infinite subsequence such that $\{\bar{x}^{k_n}\} \rightarrow \bar{x}$. If the assumptions (A1)-(A4) hold, then with probability one*

$$\lim_{n \rightarrow \infty} F_{k_n}(\bar{x}^{k_n}) = \lim_{n \rightarrow \infty} F_{k_n+1}(\bar{x}^{k_n}) = F(\bar{x}).$$

To establish that an optimal accumulation point of the incumbent sequence exists, next we explore the implication of the incumbent test described in step 3 of the algorithm.

LEMMA B.2.7. *Suppose that assumption A1-A4 hold. Let $\theta^k = F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1})$, and let $\{k_n\}_{n \in \mathbb{N}}$ represent the sequence of iterations at which the incumbent is changed. If N is finite, then $\overline{\lim}_{k \rightarrow \infty} \theta^k = 0$, with probability one. Otherwise, $\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} = 0$ with probability one.*

Proof. By definition, $\theta^k = F_{k-1}(x^k) - F_{k-1}(\bar{x}^{k-1}) \leq 0$ for all k . If N is a finite set, there exist \bar{x} and $K < \infty$ such that $\bar{x}^k = \bar{x}$ for all $k \geq K$ and thus

$$F_k(x^k) - F_k(\bar{x}) \geq \delta[F_{k-1}(x^k) - F_{k-1}(\bar{x})] = \delta\theta \quad \forall k \geq K.$$

By Theorem ?? and Corollary B.2.6, there exist a subsequence indexed by set \mathcal{K} such that

$$\begin{aligned} \lim_{k \in \mathcal{K}} x^k &= \hat{x} \\ \lim_{k \in \mathcal{K}} F_k(x^k) &= f(\hat{x}), \quad \lim_{k \in \mathcal{K}} F_k(\bar{x}) = f(\bar{x}), \\ \lim_{k \in \mathcal{K}} F_{k-1}(x^k) &= f(\hat{x}), \quad \lim_{k \in \mathcal{K}} F_{k-1}(\bar{x}) = f(\bar{x}), \end{aligned}$$

with probability one. Thus,

$$\begin{aligned}
\lim_{k \in \mathcal{K}} \{F_k(x^k) - F_k(\bar{x})\} &\geq \delta [\lim_{k \in \mathcal{K}} \{F_{k-1}(x^k) - F_{k-1}(\bar{x})\}] \\
&\Rightarrow F(\hat{x}) - F(\bar{x}) \geq \delta [F(\hat{x}) - F(\bar{x})] \\
&= \lim_{k \in \mathcal{K}} \delta \theta^k,
\end{aligned}$$

with probability one. Now since $\delta \in (0, 1)$ and $\theta^k \leq 0$ for all k , it follows that $F(\hat{x}) - F(\bar{x}) = 0$ and thus $\lim_{k \in \mathcal{K}} \delta \theta^k \leq 0$, with probability one. Now suppose N is not a finite set. By hypothesis,

$$F_{k_n}(x^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) < \delta [F_{k_n-1}(x^{k_n}) - F_{k_n-1}(\bar{x}^{k_n-1})] = \delta \theta^{k_n} \leq 0 \quad \forall n.$$

By definition of the subsequence $\{k_n\}$, we note that $\bar{x}^{k_n-1} = \bar{x}^{k_n-1}$. Therefore

$$F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) \leq \delta \theta^{k_n} \leq 0 \quad \forall n.$$

Thus,

$$\begin{aligned}
\frac{1}{m} \sum_{n=1}^m \{F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1})\} &\leq \frac{\delta}{m} \sum_{n=1}^m \theta^{k_n} \leq 0 \quad \forall m \\
&\Rightarrow \frac{1}{m} \left\{ \left(\sum_{n=1}^{m-1} F_{k_n}(\bar{x}^{k_n}) - F_{k_n}(\bar{x}^{k_n-1}) \right) + (F_{k_m}(\bar{x}^{k_m}) - F_{k_1}(\bar{x}^{k_0})) \right\} \\
&\leq \frac{\delta}{m} \sum_{n=1}^m \theta^{k_n} \leq 0 \quad \forall m.
\end{aligned}$$

Assumptions (A1)-(A4) ensure that there exist $M < \infty$, such that

$$|F_{k_m}(\bar{x}^{k_m}) - F_{k_1}(\bar{x}^{k_0})| < M \quad \forall m.$$

Thus, since $\bar{x}^{k_n} = \bar{x}^{k_{n+1}-1}$, the left hand side converges to zero with probability one as m approaches ∞ . Thus,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} = 0.$$

□

THEOREM B.2.8. *Suppose that assumptions A1-A4 are satisfied. Let $\{\bar{x}^k\}_{k=1}^{\infty}$ represent the sequence of incumbents and let X^* represent set of optimal solutions. Then there exist a subsequence $\{\bar{x}^k\}_{k \in K}$ for which every accumulation point is contained in X^* , with probability one.*

Proof. Let $\{k_n\}_{n \in \mathbb{N}}$ represent the sequence of iterations at which the incumbent is changed. Note that if N is infinite set,

$$\lim_{m \rightarrow \infty} \frac{1}{m} \sum_{n=1}^m \theta^{k_n} \leq \overline{\lim}_{n \rightarrow \infty} \theta^{k_n} \leq 0.$$

Thus, as a result of lemma B.2.7, whether N is finite or infinite, there exist a subsequence indexed by set \mathcal{K} such that

$$\lim_{k \in \mathcal{K}} \theta^{k+1} = 0.$$

Note that

$$\theta^{k+1} = F_k(x^{k+1}) - F_k(\bar{x}^k) \leq F_k(x^*) - F_k(\bar{x}^k) \quad \forall k \in K.$$

Thus as a result of Corollary B.2.6, It follows that if \bar{x} is an accumulation point of $\{\bar{x}^k\}_{k \in \mathcal{K}}$, then

$$\begin{aligned} F(\bar{x}) &\leq \overline{\lim}_{k \in \mathcal{K}} F_k(x^*) \\ &\leq c^\top x^* + \lim_{k \in \mathcal{K}} \frac{1}{k} \sum_{t=1}^k h(x^*, \omega^t) \\ &\leq F(x^*), \end{aligned}$$

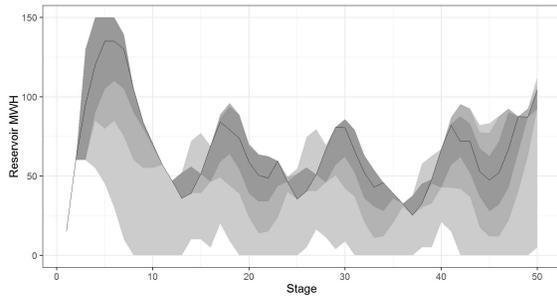
and thus, $\bar{x} \in X^*$, with probability one.

□

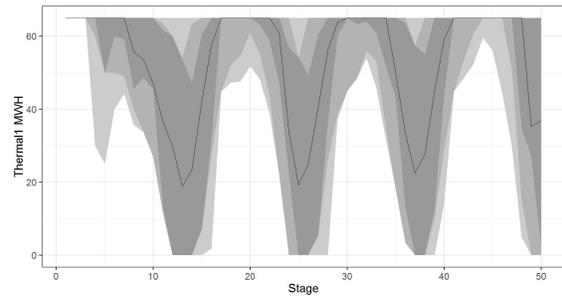
APPENDIX C

LTHS: GRAPHS

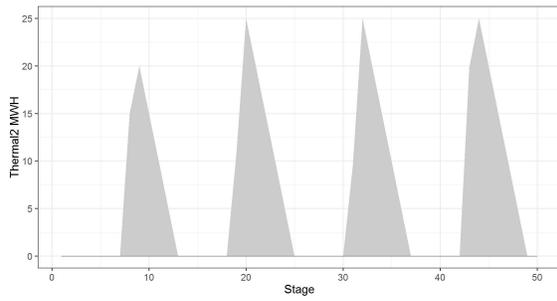
C.1 MR-MSLP with EE



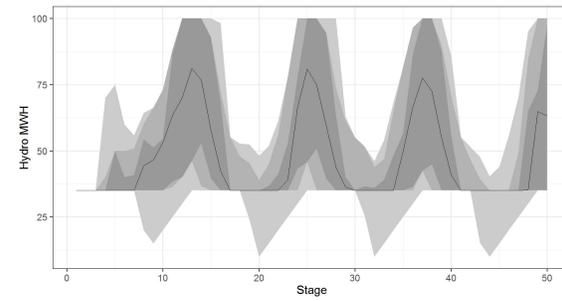
(a) Reservoir per stage



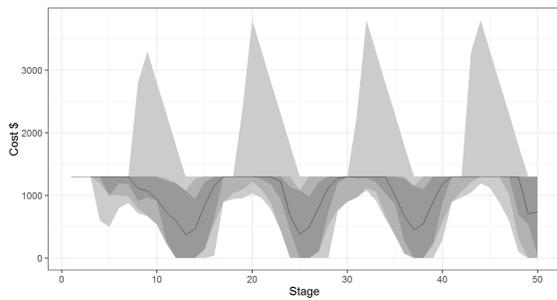
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage

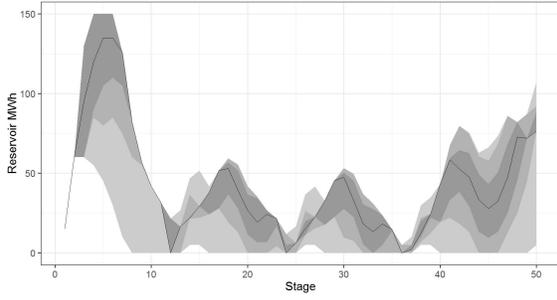


(d) Power generation γq per stage

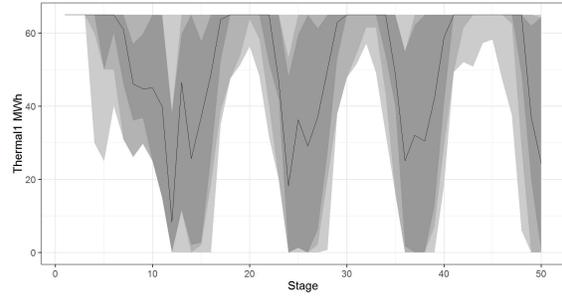


(e) Cost of operation per stage

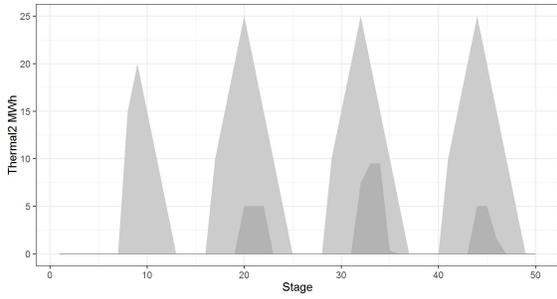
Figure C.1: Results for MR-MSLP with EE for LTHS ($\lambda = 0.25$)



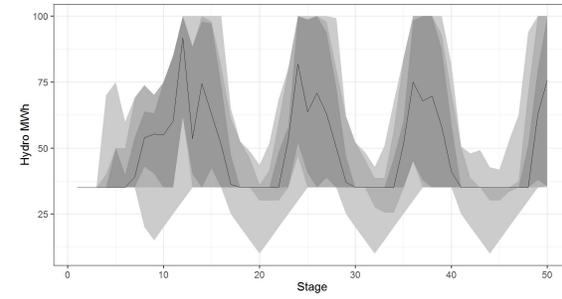
(a) Reservoir per stage



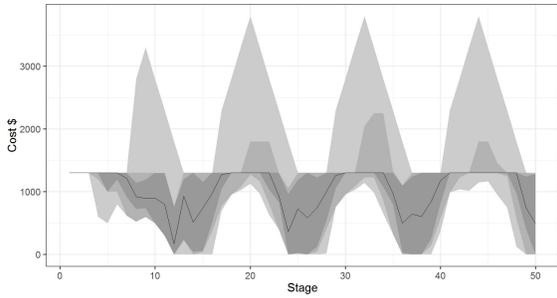
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage

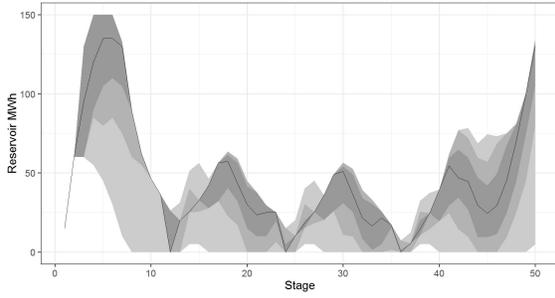


(d) Power generation γq per stage

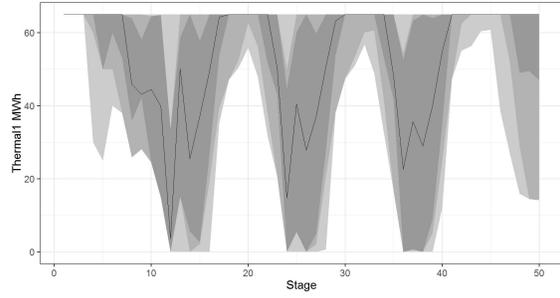


(e) Cost of operation per stage

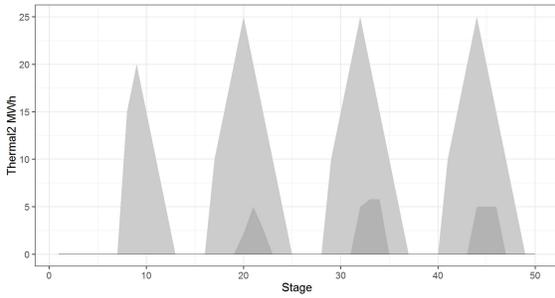
Figure C.2: Results for MR-MSLP with EE for LTHS ($\lambda = 0.5$)



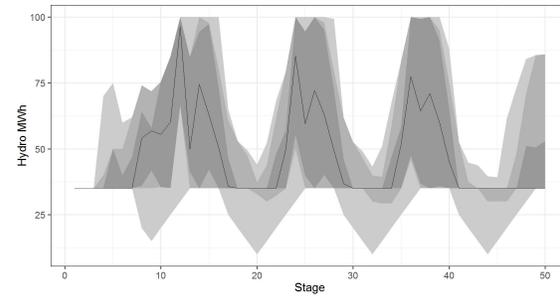
(a) Reservoir per stage



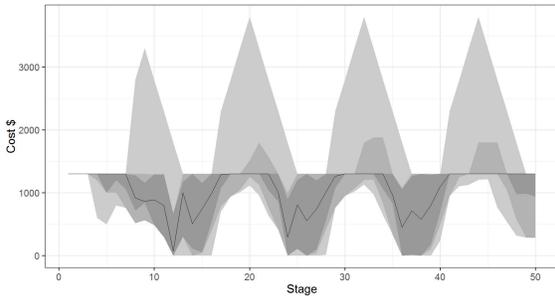
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage



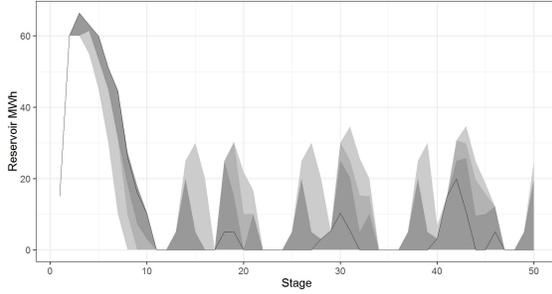
(d) Power generation γq per stage



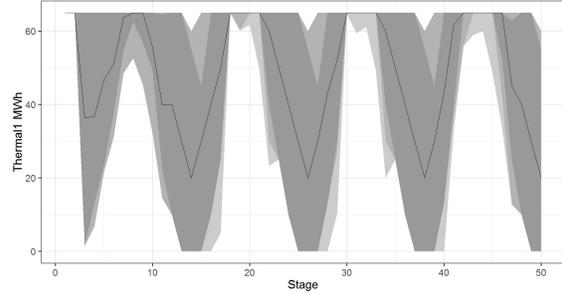
(e) Cost of operation per stage

Figure C.3: Results for MR-MSLP with EE for LTHS ($\lambda = 1$)

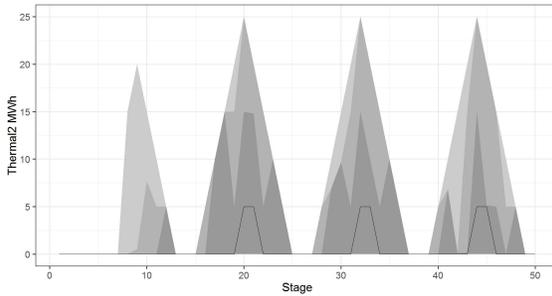
C.2 MR-MSLP with QDEV



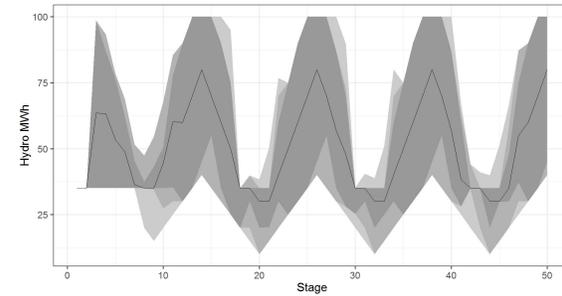
(a) Reservoir per stage



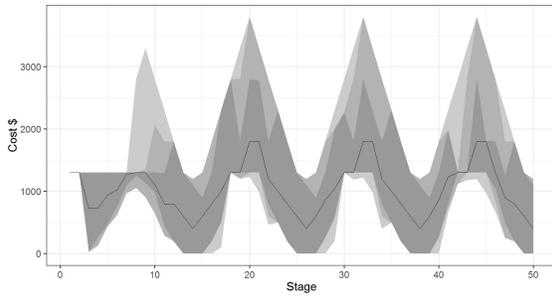
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage

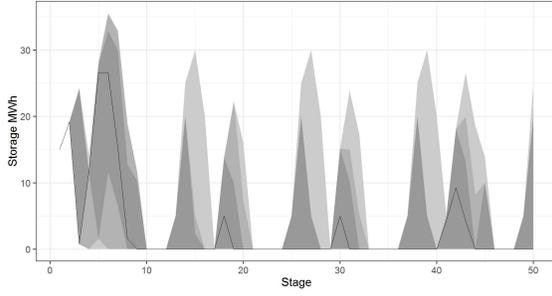


(d) Power generation γq per stage

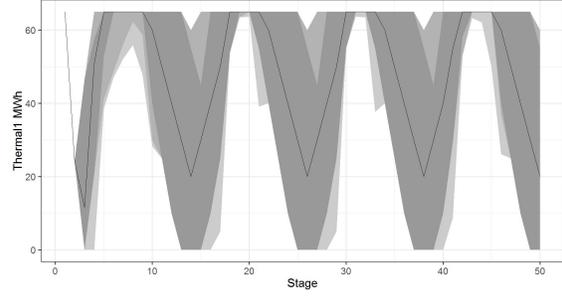


(e) Cost of operation per stage

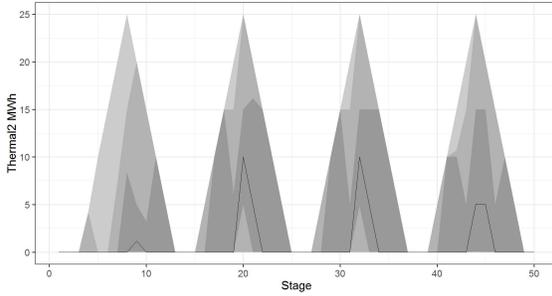
Figure C.4: Results for MR-MSLP with QDEV for LTHS ($\lambda = 0.25$)



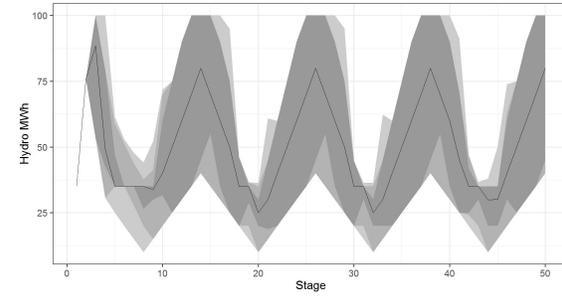
(a) Reservoir per stage



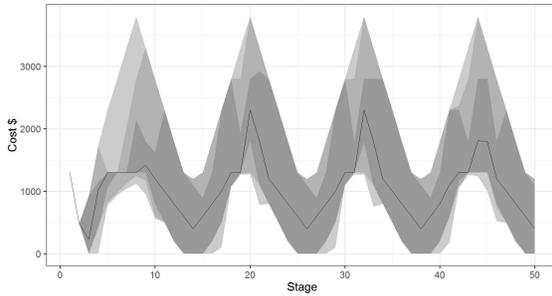
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage

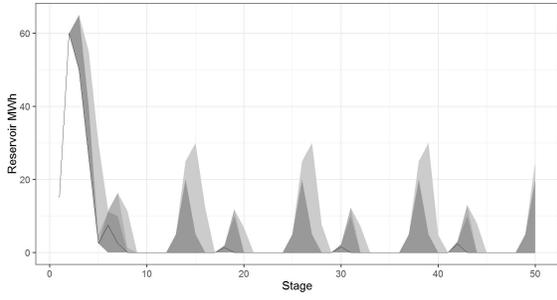


(d) Power generation γq per stage

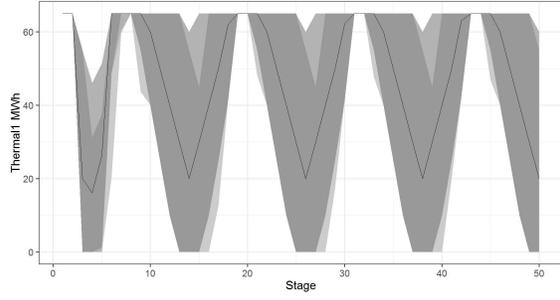


(e) Cost of operation per stage

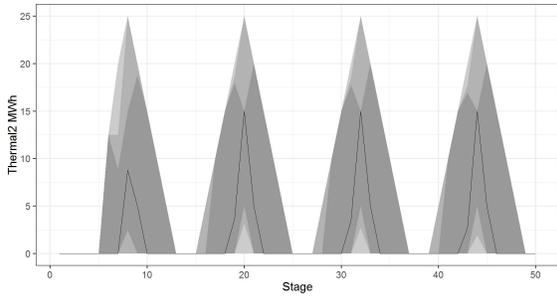
Figure C.5: Results for MR-MSLP with QDEV for LTHS ($\lambda = 0.5$)



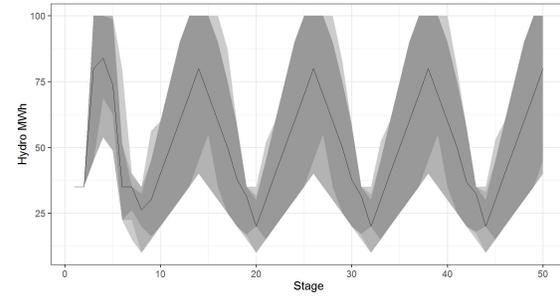
(a) Reservoir per stage



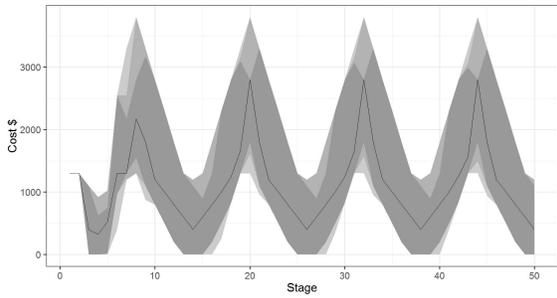
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage



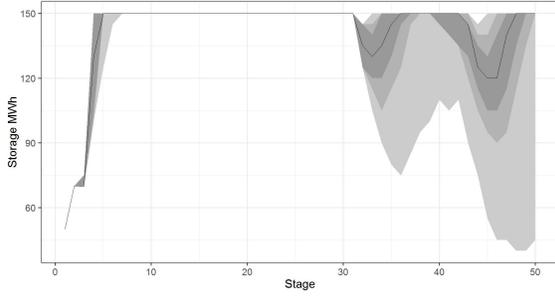
(d) Power generation γq per stage



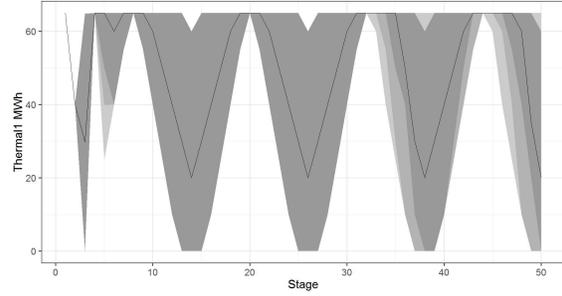
(e) Cost of operation per stage

Figure C.6: Results for MR-MSLP with QDEV for LTHS ($\lambda = 1$)

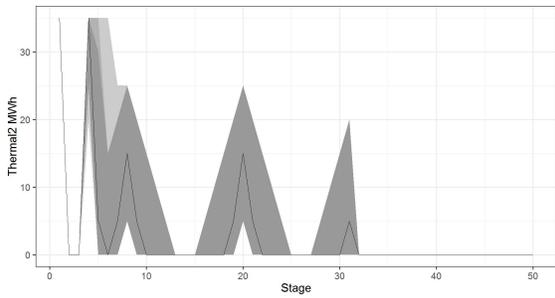
C.3 MR-MSLP with CVaR



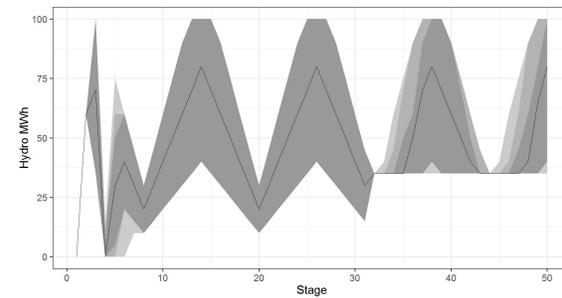
(a) Reservoir per stage



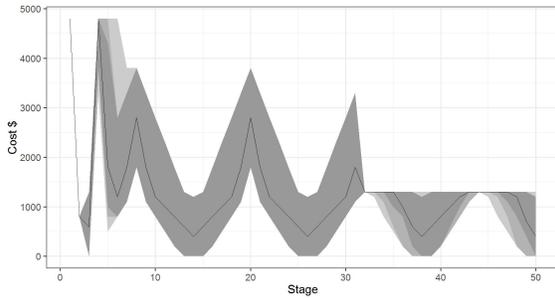
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage

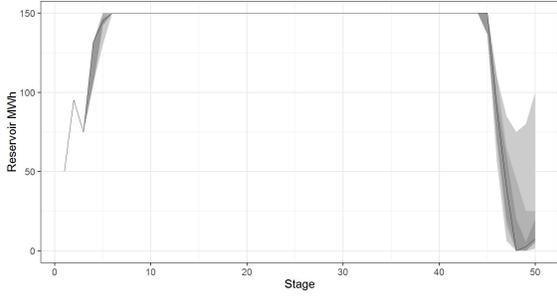


(d) Power generation γq per stage

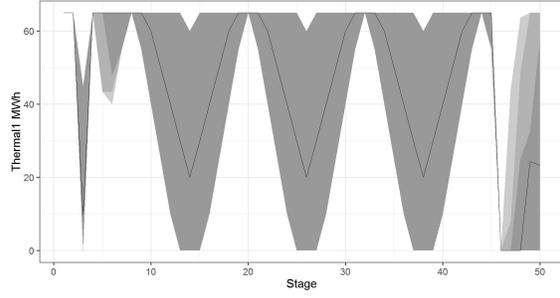


(e) Cost of operation per stage

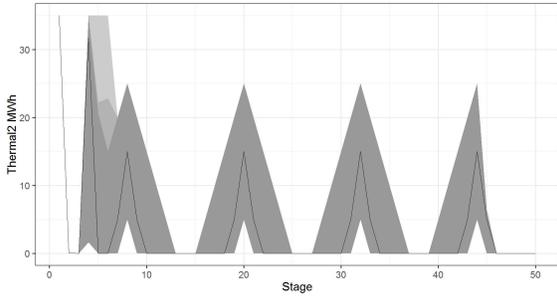
Figure C.7: Results for MR-MSLP with CVaR for LTHS ($\lambda = 0.25$)



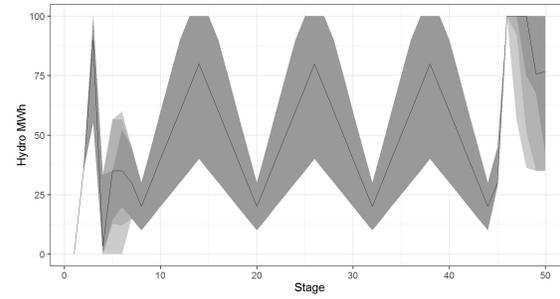
(a) Reservoir per stage



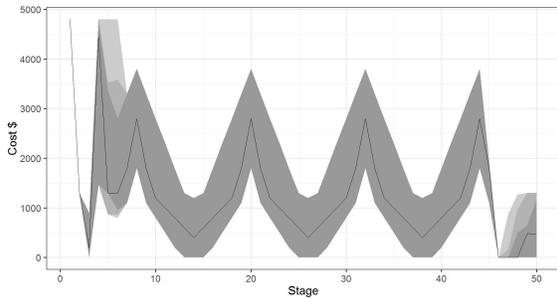
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage

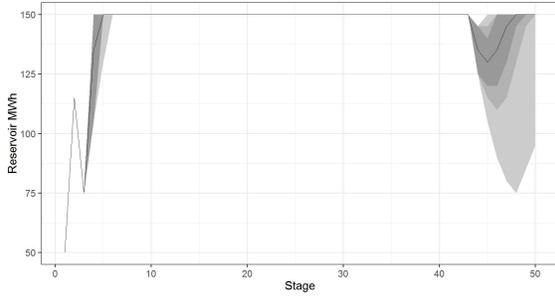


(d) Power generation γq per stage

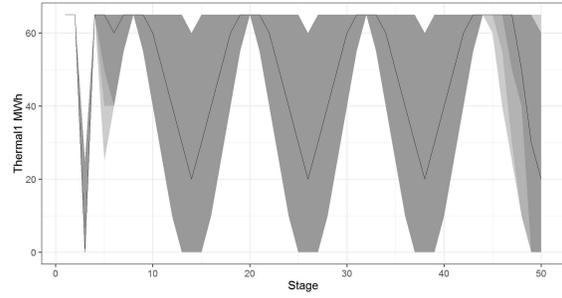


(e) Cost of operation per stage

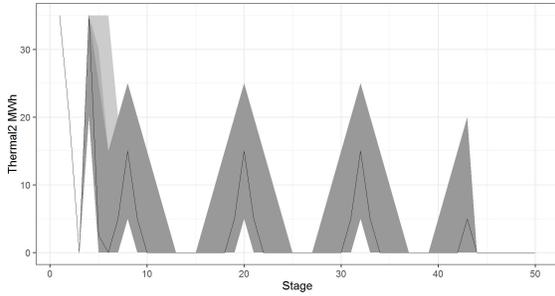
Figure C.8: Results for MR-MSLP with CVaR for LTHS ($\lambda = 0.5$)



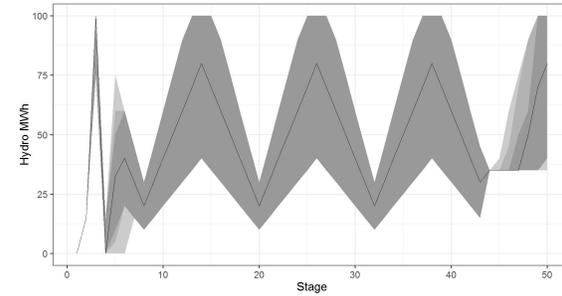
(a) Reservoir per stage



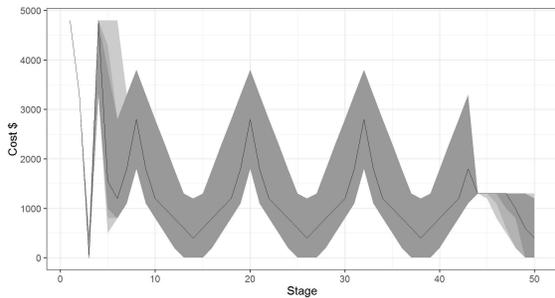
(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage



(d) Power generation γq per stage



(e) Cost of operation per stage

Figure C.9: Results for MR-MSLP with CVaR for LTHS ($\lambda = 1$)

C.4 Impact of Risk Measures

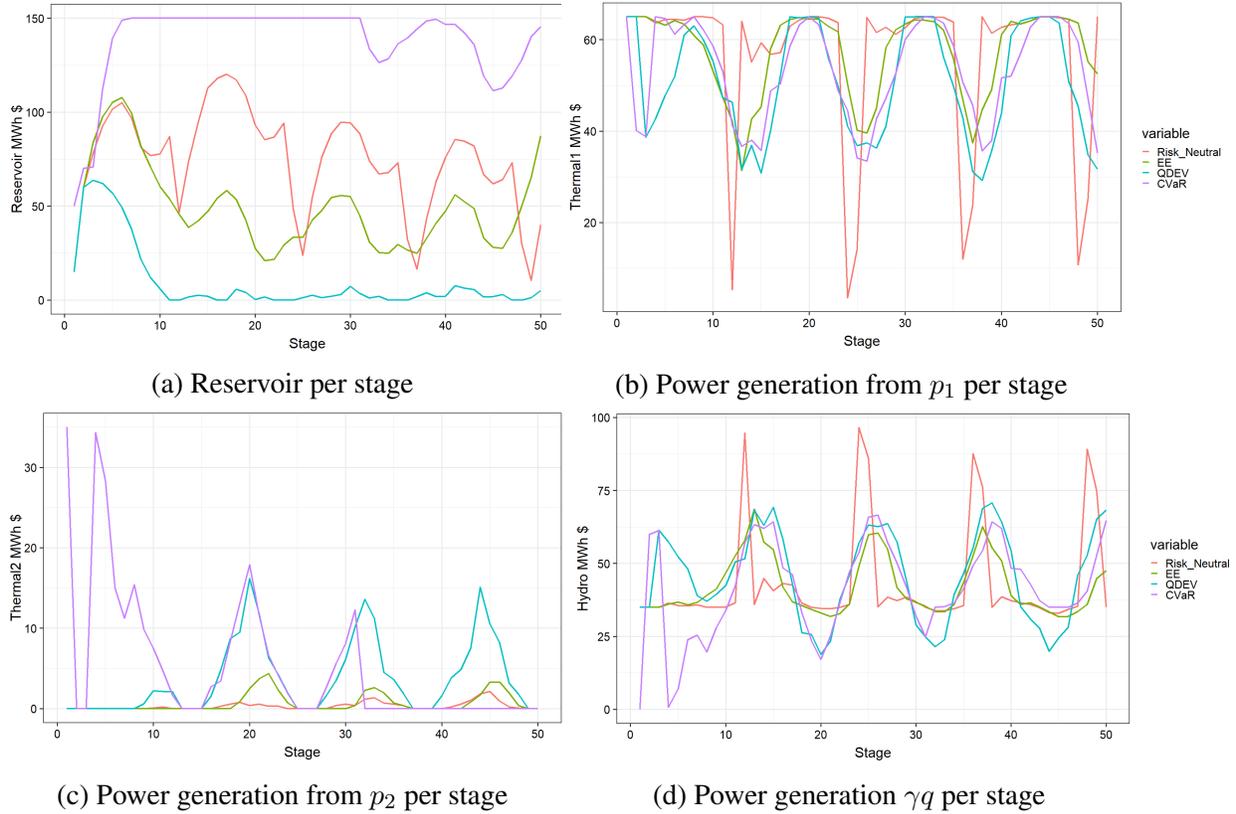
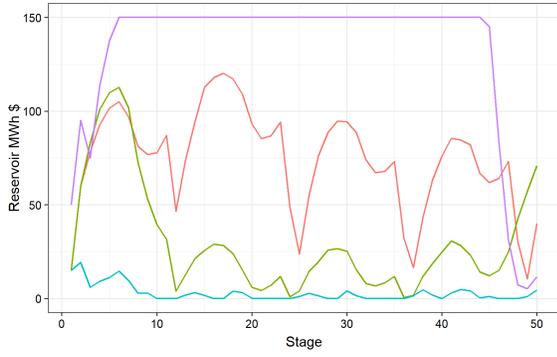
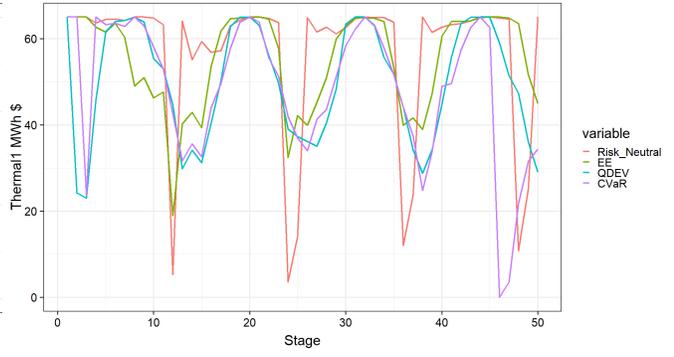


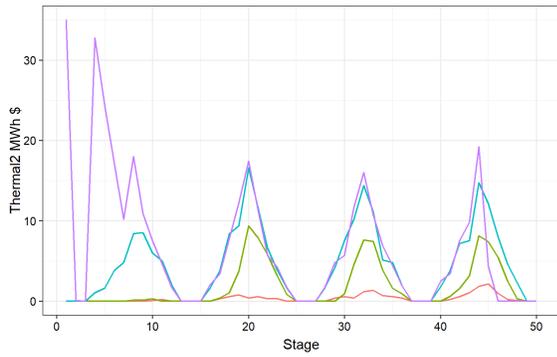
Figure C.10: Impact of Risk Measures on Extreme Scenarios ($\lambda = 0.25$)



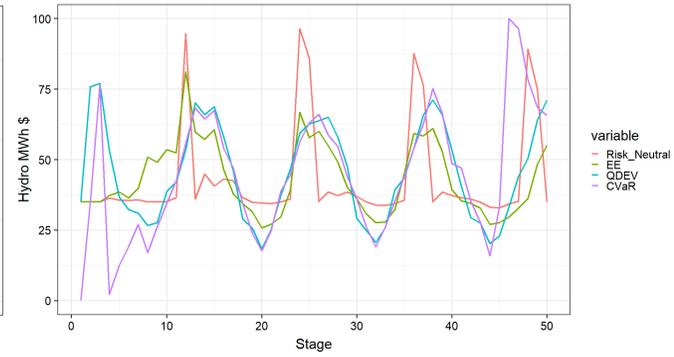
(a) Reservoir per stage



(b) Power generation from p_1 per stage

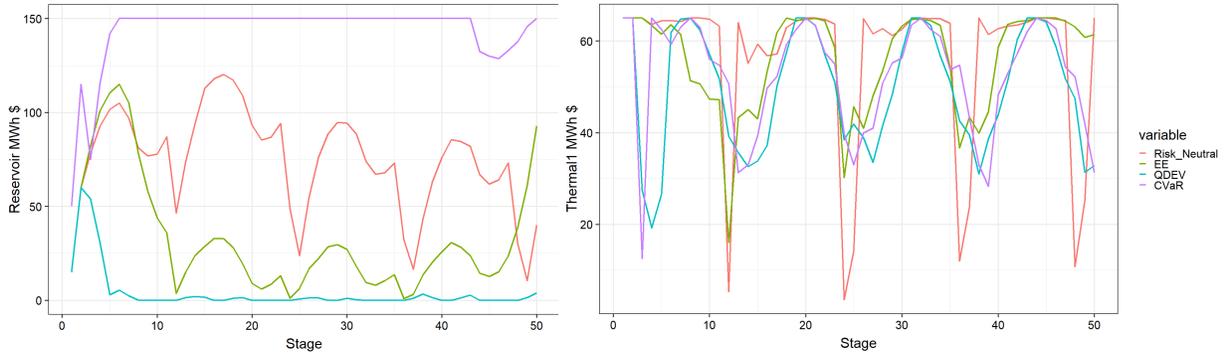


(c) Power generation from p_2 per stage



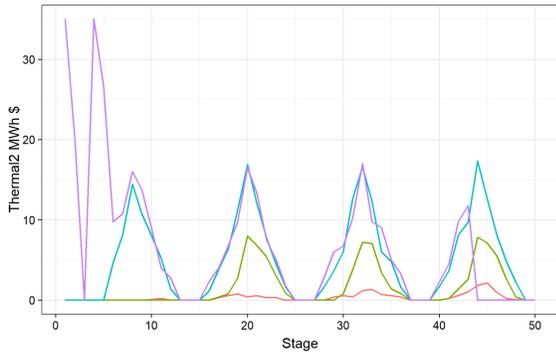
(d) Power generation γq per stage

Figure C.11: Impact of Risk Measures on Extreme Scenarios ($\lambda = 0.5$)

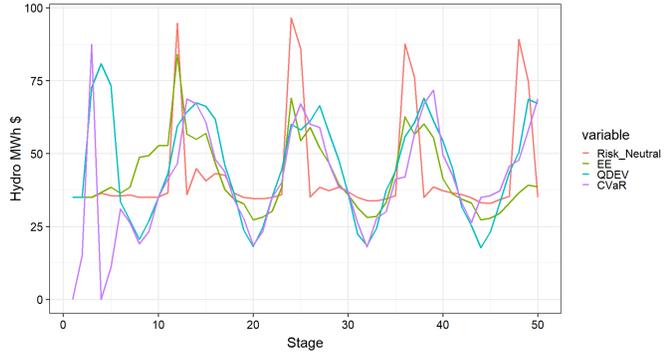


(a) Reservoir per stage

(b) Power generation from p_1 per stage



(c) Power generation from p_2 per stage



(d) Power generation γq per stage

Figure C.12: Impact of Risk Measures on Extreme Scenarios

APPENDIX D

PROOF FOR PROPERTIES OF RISK MEASURES

D.1 Expected Excess

LEMMA D.1.1. *Risk measure expected excess does not satisfy the property of translation invariance.*

Proof. For $a \in \mathbb{R}$, random variable $S \in \mathcal{F}$ and target ψ . The risk measure expected excess ρ is defined as follows:

$$\begin{aligned}\rho(a + S) &= \mathbb{E}[\max\{a + S - \psi, 0\}] \\ &= a + \mathbb{E}[\max\{S - \psi, -a\}] \\ &\neq a + \mathbb{E}[\max\{S - \psi, 0\}], \quad \forall a \neq 0.\end{aligned}$$

□

LEMMA D.1.2. *Risk measure expected excess does not satisfy the property of Positive Homogeneity.*

Proof. For $c > 0$, random variable $S \in \mathcal{F}$ and target ψ , risk measure expected excess ρ is defined as follows:

$$\begin{aligned}\rho(cS) &= \mathbb{E}[\max\{cS - \psi, 0\}] \\ &\neq c\mathbb{E}[\max\{S - \psi, 0\}], \quad \forall c \neq 1.\end{aligned}$$

□

LEMMA D.1.3. *Risk measure expected excess satisfies the property of monotonicity.*

Proof. Let the target be ψ and random variables $S_1 \in \mathcal{F}$ and $S_2 \in \mathcal{F}$, be such that $S_1 \leq S_2 \Rightarrow \mathbb{E}[S_1 - \psi] \leq \mathbb{E}[S_2 - \psi]$. Then risk measure expected excess ρ is defined as:

$$\begin{aligned}\rho(S_1) &= \mathbb{E}[\max\{S_1 - \psi, 0\}] \\ &\leq \mathbb{E}[\max\{S_2 - \psi, 0\}] = \rho(S_2).\end{aligned}$$

□

LEMMA D.1.4. *Risk measure expected excess satisfies the property of convexity.*

Proof. For random variables $S_1 \in \mathcal{F}$ and $S_2 \in \mathcal{F}$, $\lambda \in [0, 1]$ and target ψ , the risk measure expected excess ρ is defined as:

$$\begin{aligned}\rho(\lambda S_1 + (1 - \lambda)S_2) &= \mathbb{E}[\max\{\lambda S_1 + (1 - \lambda)S_2 - \psi, 0\}] \\ &= \mathbb{E}[\max\{\lambda S_1 - \lambda\psi + (1 - \lambda)S_2 - (1 - \lambda)\psi, 0\}] \\ &= \mathbb{E}[\max\{\lambda(S_1 - \psi) + (1 - \lambda)(S_2 - \psi), 0\}] \\ &\leq \mathbb{E}[\lambda \max\{S_1 - \psi, 0\}] + (1 - \lambda) \max\{S_2 - \psi, 0\}. \\ &\leq \lambda \mathbb{E}[\max\{S_1 - \psi, 0\}] + (1 - \lambda) \mathbb{E}[\max\{S_2 - \psi, 0\}].\end{aligned}$$

□

D.2 Quantile Deviation

LEMMA D.2.1. *Risk measure quantile deviation does not satisfy the property of translation invariance.*

Proof. For $a \in \mathbb{R}$, $S \in \mathcal{F}$ as random variable and $\psi \in \mathbb{R}$ as quantile, the risk measure quantile deviation ρ is defined as:

$$\begin{aligned}
\rho(a + S) &= \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - (S + a), 0\} + \epsilon_2 \max\{S + a - \psi, 0\}]\} \\
&= \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{(\psi - a) - S, 0\} + \epsilon_2 \max\{S - (\psi - a), 0\}]\} \\
\text{Let } \psi - a = \psi' &\Rightarrow \\
&= \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi' - S, 0\} + \epsilon_2 \max\{S - \psi', 0\}]\} \\
&= \rho(S) \\
&\neq a + \rho(S), \quad \forall \quad a \neq 0.
\end{aligned}$$

□

LEMMA D.2.2. *Risk measure quantile deviation satisfies the property of positive homogeneity.*

Proof. For $c > 0$, $S \in \mathcal{F}$ as random variable and $\psi \in \mathbb{R}$ as quantile, the risk measure quantile deviation ρ is defined as:

$$\begin{aligned}
\rho(cS) &= \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - cS, 0\} + \epsilon_2 \max\{cS - \psi, 0\}]\} \\
&= c\mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\frac{\psi}{c} - S, 0\} + \epsilon_2 \max\{S - \frac{\psi}{c}, 0\}]\} \\
\text{Let } \frac{\psi}{c} = \psi' &\Rightarrow \\
&= c\mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi' - S, 0\} + \epsilon_2 \max\{S - \psi', 0\}]\} \\
&= c\rho(S).
\end{aligned}$$

□

LEMMA D.2.3. *Risk measure quantile deviation does not satisfy the property of monotonicity.*

Proof. Let $S_1, S_2 \in \mathcal{F}$ be uniform random variables, such that $S_1 \leq S_2$. let $\psi \in \mathbb{R}$ be the quantile.

Then the risk measure quantile deviation ρ is defined as:

$$\begin{aligned}\rho(S_1) - \rho(S_2) &= \text{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - S_1, 0\} + \epsilon_2 \max\{S_1 - \psi, 0\}]\} \\ &\quad - \text{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - S_2, 0\} + \epsilon_2 \max\{S_2 - \psi, 0\}]\}.\end{aligned}$$

Let S_1 and S_2 be uniformly distributed random variables, such that $S_1 \in \{1, 2\}$ and $S_2 \in \{2, 3\}$. Let $\alpha = 0.1$, therefore we have $\epsilon_1 = 0.9$ and $\epsilon_2 = 0.1$. Then for risk measure ρ as quantile deviation,

$$\begin{aligned}\rho(S_1) - \rho(S_2) &= \text{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - S_1, 0\} + \epsilon_2 \max\{S_1 - \psi, 0\}]\} \\ &\quad - \text{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - S_2, 0\} + \epsilon_2 \max\{S_2 - \psi, 0\}]\} \\ &= 0.5 - 0.25 \\ &= 0.25 \geq 0.\end{aligned}$$

□

LEMMA D.2.4. *Risk measure quantile deviation satisfies the property of convexity.*

Proof. Let $S_1, S_2 \in \mathcal{F}$ be random variables, $\lambda \in [0, 1]$ and $\psi \in \mathbb{R}$ be the quantile, then the risk measure quantile deviation ρ is defined as:

$$\begin{aligned}
\rho(\lambda S_1 + (1 - \lambda)S_2) &= \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - (\lambda S_1 + (1 - \lambda)S_2), 0\}] \\
&\quad + \epsilon_2 \max\{\lambda S_1 + (1 - \lambda)S_2 - \psi, 0\}]\} \\
&= \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\lambda\psi + (1 - \lambda)\psi - (\lambda S_1 + (1 - \lambda)S_2), 0\}] \\
&\quad + \epsilon_2 \max\{\lambda S_1 + (1 - \lambda)S_2 - \lambda\psi - (1 - \lambda)\psi, 0\}]\} \\
&= \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\lambda\psi - \lambda S_1 + (1 - \lambda)\psi - (1 - \lambda)S_2, 0\}] \\
&\quad + \epsilon_2 \max\{\lambda S_1 - \lambda\psi + (1 - \lambda)S_2 - (1 - \lambda)\psi, 0\}]\} \\
&\leq \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\lambda\psi - \lambda S_1, 0\}] + \epsilon_1 \max\{(1 - \lambda)\psi - (1 - \lambda)S_2, 0\}] \\
&\quad + \epsilon_2 \max\{\lambda S_1 - \lambda\psi, 0\} + \epsilon_2 \max\{(1 - \lambda)S_2 - (1 - \lambda)\psi, 0\}]\} \\
&\leq \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\lambda\psi - \lambda S_1, 0\}] + \epsilon_2 \max\{\lambda S_1 - \lambda\psi, 0\}]\} \\
&\quad + \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{(1 - \lambda)\psi - (1 - \lambda)S_2, 0\}] \\
&\quad + \epsilon_2 \max\{(1 - \lambda)S_2 - (1 - \lambda)\psi, 0\}]\} \\
&\leq \lambda \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - S_1, 0\}] + \epsilon_2 \max\{S_1 - \psi, 0\}]\} \\
&\quad + (1 - \lambda) \mathbf{Min}\{\mathbb{E}[\epsilon_1 \max\{\psi - S_2, 0\}] + \epsilon_2 \max\{S_2 - \psi, 0\}]\} \\
&\leq \lambda \rho(S_1) + (1 - \lambda) \rho(S_2).
\end{aligned}$$

□

D.3 Conditional Value-at-Risk

LEMMA D.3.1. *Risk measure CVaR satisfies the property of translation invariance.*

Proof. For $a \in \mathbb{R}$, $S \in \mathcal{F}$ as random variable and $\psi \in \mathbb{R}$ as quantile, the risk measure conditional value-at-risk ρ is defined as:

$$\begin{aligned}
\rho(a + S) &= \text{Min}\left\{\psi + \frac{1}{1 - \alpha} \mathbb{E}[\max\{S + a - \psi, 0\}]\right\} \\
&= \text{Min}\left\{\psi + \frac{1}{1 - \alpha} \mathbb{E}[\max\{S - (\psi - a), 0\}]\right\} \\
\text{Let } \psi' &= \psi - a \Rightarrow \\
&= \text{Min}\left\{\psi' + a + \frac{1}{1 - \alpha} \mathbb{E}[\max\{S - \psi', 0\}]\right\} \\
&= a + \text{Min}\left\{\psi' + \frac{1}{1 - \alpha} \mathbb{E}[\max\{S - \psi', 0\}]\right\} \\
&= a + \rho(S).
\end{aligned}$$

□

LEMMA D.3.2. *Risk measure CVaR satisfies the property of positive homogeneity.*

Proof. For $c > 0$, $S \in \mathcal{F}$ as random variable and ψ as quantile, the risk measure conditional value-at-risk ρ is defined as:

$$\begin{aligned}
\rho(cS) &= \text{Min}\left\{\psi + \frac{1}{1 - \alpha} \mathbb{E}[\max\{cS - \psi, 0\}]\right\} \\
&= c \left(\text{Min}\left\{\frac{\psi}{c} + \frac{1}{1 - \alpha} \mathbb{E}[\max\{S - \frac{\psi}{c}, 0\}]\right\} \right) \\
\text{Let } \psi' &= \frac{\psi}{c} \Rightarrow \\
&= \text{Min}\left\{\psi' + \frac{1}{1 - \alpha} \mathbb{E}[\max\{S - \psi', 0\}]\right\} \\
&= c\rho(S).
\end{aligned}$$

□

LEMMA D.3.3. *Risk measure CVaR satisfies the property of monotonicity.*

Proof. Let $S_1 \in \mathcal{F}$ and $S_2 \in \mathcal{F}$ be random variables, such that $S_1 \leq S_2$. Therefore for a quantile $\psi \in \mathbb{R}$, we have $\mathbb{E}[\max\{S_1 - \psi, 0\}] \leq \mathbb{E}[\max\{S_2 - \psi, 0\}]$. Then the risk measure conditional

value-at-risk ρ is defined as:

$$\begin{aligned}
\rho(S_1) &= \mathbf{Min}\left\{\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{S_1 - \psi, 0\}]\right\} \\
&\leq \mathbf{Min}\left\{\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{S_2 - \psi, 0\}]\right\} \\
&\leq \rho(S_2).
\end{aligned}$$

□

LEMMA D.3.4. *Risk measure CVaR satisfies the property of convexity.*

Proof. Let $S_1, S_2 \in \mathcal{F}$ be random variables, $\lambda \in [0, 1]$ and $\psi \in \mathbb{R}$ be the quantile, then the risk measure conditional value-at-risk ρ is defined as:

$$\begin{aligned}
\rho(\lambda S_1 + (1-\lambda)S_2) &= \mathbf{Min}\left\{\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{\lambda S_1 + (1-\lambda)S_2 - \psi, 0\}]\right\} \\
&= \mathbf{Min}\left\{\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{\lambda S_1 - \lambda\psi + (1-\lambda)S_2 - (1-\lambda)\psi, 0\}]\right\} \\
&\leq \mathbf{Min}\left\{\lambda\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{\lambda S_1 - \lambda\psi, 0\}]\right. \\
&\quad \left. + (1-\lambda)\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{(1-\lambda)S_2 - (1-\lambda)\psi, 0\}]\right\} \\
&\leq \lambda\left[\mathbf{Min}\left\{\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{S_1 - \psi, 0\}]\right\}\right] \\
&\quad + (1-\lambda)\left[\mathbf{Min}\left\{\psi + \frac{1}{1-\alpha}\mathbb{E}[\max\{S_2 - \psi, 0\}]\right\}\right] \\
&\leq \lambda\rho(S_1) + (1-\lambda)\rho(S_2).
\end{aligned}$$

□