DATA-DRIVEN DISTRIBUTIONALLY ROBUST CHANCE-CONSTRAINED LINEAR MATRIX INEQUALITIES

FENGJIE LIANG

MPhil

The Hong Kong Polytechnic University

2024

The Hong Kong Polytechnic University
Department of Logistics and Maritime Studies

Data-Driven Distributionally Robust Chance-Constrained Linear Matrix Inequalities

Fengjie Liang

A thesis submitted in partial fulfillment of the requirements for
the degree of Master of Philosophy
May 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

Signature:____________________

Name of Student: Fengjie Liang

Abstract

In this thesis, we present approximation and reformulation techniques for problem with distributionally robust chance-constrained linear matrix inequality (DRCCLMI), aiming at overcoming the computational challenges posed by multidimensional integration and nonconvexity of feasible sets. DRCCLMI seek a robust solution which guarantees that the chance constraints are fulfilled for a wide range of possible distribution within an ambiguity set. Specifically, we consider a data-driven ambiguity set which includes all the potential distributions sharing the same moment information. We first propose an inner approximation for DRCCLMI in a general form to deal with common constraint structures encountered in real-world scenarios. The key method we use for approximation is the Conditional Value-at-Risk (CVaR) approach, which enables us to approximate DRCCLMI in a way that ensures a certain level of solution quality while maintaining the robustness of the original constraint. Second, we derive an inner approximation and exact reformulations for a special case of DRCCLMI with and without support information, respectively. Specifically, this special case refers to the situation where a block matrix structure is inherent in the linear matrix inequality. Notably, our approximation and reformulation techniques facilitate the transformation of the original DRCCLMI into a more tractable semidefinite programming (SDP) problem, simplifying the computational process and improving the accuracy and efficiency of the solution. The practicality and effectiveness of these techniques are demonstrated through numerical studies on two real-world applications: truss topology design problem and calibration problem.

Acknowledgements

I am extremely grateful to my chief supervisor, Dr. Kai Pan, for his dedicated guidance and support. I am honoured to have benefited from his supervision and advice. I would also like to express my heartfelt thanks to Dr. Jianqiang Cheng, who has given me great help and insights from different perspectives in my research. In addition, I would like to thank all the teachers in our department who provided me with a good research environment. Of course, I would also like to thank my parents and boyfriend. Without their support and encouragement, I could not have done all the work. I hereby give you my genuine blessing.

Contents

List of Tables

Chapter 1

Introduction

A Linear Matrix Inequality (LMI) is a mathematical constraint in the form of an inequality that involves a linear or affine function, maps from a finite-dimensional vector space to a set of Hermitian or symmetric matrices (Scherer and Weiland 2000). LMI has been emerging as a powerful tool for control, systems, and other research fields in engineering and applied mathematics (S. Boyd et al. 1994). For example, Masubuchi et al. (1998) propose a unique approach for linear controller synthesis, where multiple LMI constraints represent control specifications. Zhong et al. (2003) study the uncertain linear time-invariant systems by using robust fault detection filters in LMI format.

There are several reasons why LMI plays such an important and widespread role. First, LMI provides a unified language into which different types of constraints appearing in optimization problems can easily be expressed (El Ghaoui and Niculescu 2000). These constraints encompass linear, quadratic, semidefinite, and spectral constraints. Second, LMI copes well with the high-dimensional problem, which is typical of many modern engineering problems (Wang et al. 2017). Third, convex constraints in LMI can attract full benefits from the convex optimization theory (Scherer and Weiland 2000). Finally, with the invention of interior point approach (Nesterov and Nemirovskii 1994), the Ellipsoid algorithm, and other sound numerical algorithms, solving LMI efficaciously even in large-scale cases is feasible.

Dealing with uncertainties in real-world systems prompts an important question, while LMI is suitable for deterministic systems. An effective answer in this context arises as chance-constrained programming. It first comes into being in Charnes and Cooper (1959), and further development occurs in Miller and Wagner (1965). It can be used to guarantee constraint satisfaction in a probabilistic sense when the exact values of random variables are unknown or may change in the future. Chance-constrained programming is important especially in areas like robust control and risk management, where model establishment needs stability under different situations (Calafiore and Campi 2006; Jiang and Guan 2016). Applying chance constraints, not only uncertainties can be effectively managed but also systems can be designed to fulfill certain operational requirements. Thus, by incorporating chance constraints, we can design a system of constraints to ensure stability and reliability in unpredictable environments.

Although the integration of chance constraints with LMI is a major advance, it assumes a crucial assumption. Under the assumption, the probability distributions are known (Shapiro, Dentcheva, et al. 2021). Nevertheless, collecting all of the information is not practicable, particularly in practical senarios, where we have only partial information due to lack of historical data. Hence, we need a data driven solution to manage the uncertainty. In this case, the idea of Distributionally Robust Optimization (DRO) plays an important role. DRO doesn't only consider one probability distribution, it takes multiple distributions inside the ambiguity set into consideration. This is especially beneficial for system models when we don't know the exact probability distribution in advance while constructing models. In this scenario, historical information is particularly valuable because it gives us empirical data from which we could construct the ambiguity set. Once we put these kinds of information in our problem formulation, we can benefit a lot in terms of capturing uncertainty and improving robustness.

In recent years, numerous disciplines, including transportation (e.g., Yang et al. 2023), supply chain management (e.g., Zhong et al. 2023), and energy systems (e.g., Zhao et al. 2019), have embraced DRO techniques. The mathematical structure of uncertain variables in these applications constitutes an inherent property, and has tremendous influence on the outcomes of optimization problems. Admitting the point of view of DRO can yield more intelligent and efficient decisions, with less conservatism and thereby narrowing the robustness-applicability gap.

In spite of its attractive qualities, DRO does not come without challenges. One of the main challenges is in the computational complexity to solve DRO problems, especially when it comes to large-scale systems, or when ambiguity sets are complex (Shapiro, Dentcheva, et al. 2021). It comes from the requirement to optimize over a number of possible distributions, which can each lead to very different optimization outcomes. Furthermore, even the selection and construction of an acceptable ambiguity set alone poses a challenge (Mohajerin Esfahani and Kuhn 2018). It requires a fine balance that avoids being too conservative (and hence, providing overly cautious solutions), while in the meantime being too optimistic (and hence, ignoring robustness requirements).

Crossing DRO, Chance-Constrained Programming, and LMIs, our research forges a path in the pursuit of a unified system that intermingles these three powerful mathematical constructs, referred to as Distributionally Robust Chance-constrained Linear Matrix Inequality (DRCCLMI) herein. To join these three mathematical tools together, we face the daunting task of dealing with complicated, practical optimization challenges plagued by distributional uncertainties and probabilistic constraints simultaneously, without sacrificing computational tractability.

Several benefits accrue from such integration. First, DRO is protective, guarding against unfavorable outcomes by preconditioning on a panorama of plausible distributions that are contained within the specified ambiguity set. This robustness is especially critical when one knows but a little about their underlying data distribution. Second, the chance-constrained optimization framework can be directly incorporated into probabilistic constraints, making the optimization model considerably more robust and, hence, realistic. Lastly, LMIs furnish a mathematical framework that is highly expressive, thus permitting the modelling of a broad collection of applied problems.

Our contributions are as follows:

1. By utilizing the Conditional Value-at-Risk (CVaR) constraints and piecewise linear decision rule, we are able to develop an inner approximation for general DRCCLMI.

2. By using CVaR constraints, we also develop an inner approximation and exact reformulations for DRCCLMI with a block matrix structure in the linear matrix inequalities for both ambiguity sets with and without support information. This exact reformulations allows for more accurate and efficient solutions.

3. To verify the performance of our proposed reformulations and approximations, we conducted experiments on two industrial application problems: truss topology design problem and calibration problem. The results of the numerical study demonstrate that our proposed method performs well.

The remaining chapters are organized as follows. Chapter 2 introduces the original DRCCLMI of interest and an overview of CVaR constraints to approximate robust chance constraints. By using CVaR, we derive a conservative approximation for the general DRCCLMI problems in Chapter 3, and also an approximation and exact reformulations for a special block matrix case of DRCCLMI in Chapter 4. Chapter 5 presents numerical experiments based on applications to truss topology design problem and calibration problem. Finally, Chapter 6 summarizes our contributions.

Chapter 2

Problem Formulation

2.1 DRCCLMI

Given a realization of a random variable $\boldsymbol{\xi} \in \mathbb{R}^m$, the following semi-definite program seeks an $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^n$ such that the total cost ($\mathbf{c}^\top \mathbf{x}$) is minimized and a LMI is satisfied:

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \mathbf{c}^\top \mathbf{x} \tag{2.1a}$$

$$\text{s.t.} \quad \mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0. \tag{2.1b}$$

Here, $\mathbf{c} \in \mathbb{R}^n$, $\mathcal{X}$ is a convex set, and $\mathbf{A}_i(\mathbf{x}) \in \mathbb{S}_{d \times d}$ ($\forall i = 1, \ldots, m$) are affine in $\mathbf{x}$. Note that Problem (2.1) is a convex program, and the above LMI (2.1b) generalizes linear and second-order cone (SOC) constraints. To hedge against potential risks due to the uncertainty characterized by a joint probability distribution $\mathbb{P}$, we consider a risk tolerance parameter $\epsilon \in (0, 1)$ and solve the following chance-constrained program (CCP):

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \mathbf{c}^\top \mathbf{x} \tag{2.2a}$$

$$\text{s.t.} \quad \mathbb{P}\left\{ \mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0 \right\} \geq 1 - \epsilon, \tag{2.2b}$$

which ensures that the LMI (2.1b) can be satisfied at a high probability $1 - \epsilon$.

The distribution $\mathbb{P}$ is often unknown due to limited information to accurately estimate it, though some historical data can be available. Thus, we use the data to statistically estimate the moment information of $\mathbb{P}$ and construct the following data-driven distributional ambiguity set $\mathcal{D}$ that includes all the potential probability distributions sharing the same moment information:

$$\mathcal{D}(\mathcal{S}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left\{ \mathbb{P} \left| \begin{array}{l} \mathbb{P}(\boldsymbol{\xi} \in \mathcal{S}) = 1 \\ \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}] = \boldsymbol{\mu} \\ \mathbb{E}_{\mathbb{P}}\left[(\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})^{\top}\right] \preceq \boldsymbol{\Sigma} \end{array} \right. \right\}. \tag{2.3}$$

Specifically, any distribution in $\mathcal{D}$ specifies the support $\mathcal{S}$, mean $\boldsymbol{\mu}$, and upper bound $\boldsymbol{\Sigma} \succ 0$ for the covariance matrix of random variable $\boldsymbol{\xi}$. Throughout this thesis, we consider the support set $\mathcal{S}$ to be $\{\boldsymbol{\xi} \in \mathbb{R}^m : \|\boldsymbol{\xi}\| \leq t\}$. By hedging against the worst-case probability distribution in $\mathcal{D}$, we consider a distributionally robust counterpart of the chance constraint (2.2b), leading to the following problem with a DRCCLMI:

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \mathbf{c}^{\top}\mathbf{x} \tag{2.4a}$$

$$\text{s.t.} \quad \inf_{\mathbb{P} \in \mathcal{D}} \mathbb{P}\left\{ \mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0 \right\} \geq 1 - \epsilon, \tag{2.4b}$$

Solving Problem (2.4) leads to a robust solution ensuring that the chance constraint (2.2b) is fulfilled for a wide range of possible distribution $\mathbb{P}$ within the ambiguity set $\mathcal{D}$.

Nevertheless, solving Problem (2.4) can be computationally challenging. First, the feasible region of Problem (2.4) is nonconvex in general, even if $\mathcal{X}$ is a convex set and $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x})$ is an affine mapping in $\mathbf{x}$. Second, computing $\mathbb{P}\{\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0\}$ is difficult due to the computational complexity of the high-dimensional integration (Nemirovski and Shapiro 2007; Shapiro, Dentcheva, et al. 2021). Third, the stochastic dependence among the random variables and the presence of the ambiguity set further complicate the solution process (Cheung et al. 2012). Therefore, we aim to develop reformulations and approximations to help solve Problem (2.4) efficiently with high-quality solutions.

2.2 CVaR Risk Measure

To address the above challenges, we approximate constraint (2.4b) using a CVaR constraint. Specifically, given a measurable loss function $\boldsymbol{L} : \mathbb{R}^m \to \mathbb{R}$, the corresponding CVaR (Rockafellar and Uryasev 2000) at level ϵ associated with $\mathbb{P}$ is specified as

$$\mathbb{P}\text{−CVaR}_\epsilon \left(\boldsymbol{L}\left(\boldsymbol{\xi}\right)\right) = \inf_{\beta \in \mathbb{R}} \left\{ \beta + \frac{1}{\epsilon} \mathbb{E}_{\mathbb{P}} \left[\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right) - \beta \right)^+ \right] \right\},$$

which is a risk assessment measure calculates the conditional expectation of the loss beyond the $(1-\epsilon)$-quantile, i.e., expected tail loss. We further consider the following distributionally robust chance constraint (DRCC)

$$\inf_{\mathbb{P} \in \mathcal{D}} \mathbb{P}\left\{ \boldsymbol{L}\left(\boldsymbol{\xi}\right) \leq 0 \right\} \geq 1 - \epsilon.$$

Here, DRCC ensures the probability of $\boldsymbol{L}(\boldsymbol{\xi}) \leq 0$ in the worst case to be at least $1-\epsilon$. It is well known that for any loss function $\boldsymbol{L}(\boldsymbol{\xi})$,

$$\mathbb{P}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right) \leq \mathbb{P}\text{−CVaR}_\epsilon \left(\boldsymbol{L}\left(\boldsymbol{\xi}\right)\right)\right) \geq 1 - \epsilon.$$

This implies that $\mathbb{P}\text{−CVaR}_\epsilon(\boldsymbol{L}(\boldsymbol{\xi})) \leq 0$ is already sufficient to indicate that $\mathbb{P}(\boldsymbol{L}(\boldsymbol{\xi}) \leq 0) \geq 1-\epsilon$. Since this holds for any distribution $\mathbb{P}$, we can derive that the following distributionally robust CVaR constraint (2.5)

$$\sup_{\mathbb{P} \in \mathcal{D}} \mathbb{P}\text{−CVaR}_\epsilon \left(\boldsymbol{L}\left(\boldsymbol{\xi}\right)\right) \leq 0 \tag{2.5}$$

constitutes a conservative approximation for DRCC.

Moreover, a notable refinement arises when considering the nature of the loss function $\boldsymbol{L}(\boldsymbol{\xi})$, as shown in the following proposition. Here, we emphasize that Proposition 1 holds under the assumption that no support information $\mathcal{S}$ is defined in the ambiguity set $\mathcal{D}$.

PROPOSITION 1. Assume that no support information $\mathcal{S}$ is defined within the ambiguity set $\mathcal{D}$. Let $\boldsymbol{L} : \mathbb{R}^m \to \mathbb{R}$ be a continuous function that is either

1. concave in $\boldsymbol{\xi}$, or

2. quadratic in $\boldsymbol{\xi}$

Then, we have

$$\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\text{−CVaR}_{\epsilon}\left(\boldsymbol{L}(\boldsymbol{\xi})\right) \leq 0 \Longleftrightarrow \inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right) \leq 0\right) \geq 1-\epsilon. \tag{2.6}$$

Proof. Firstly, we aim to prove:

$$\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right) \leq 0\right) \geq 1-\epsilon \Longleftrightarrow \mathsf{WC}-\mathsf{VaR}_{\epsilon}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right)\right) \leq 0, \tag{2.7}$$

where the $\mathsf{WC}-\mathsf{VaR}_{\epsilon}$ is the worst-case Value-at-Risk (VaR) of $\boldsymbol{L}(\boldsymbol{\xi})$. According to the definition, the $\mathsf{WC}-\mathsf{VaR}_{\epsilon}$ is equivalent to

$$\mathsf{WC}-\mathsf{VaR}_{\epsilon}\left(\boldsymbol{L}(\boldsymbol{\xi})\right) = \inf_{\gamma\in\mathbb{R}} \left\{\gamma : \inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right) \leq \gamma\right) \geq 1-\epsilon\right\}. \tag{2.8}$$

Observe that if the left hand side of the equivalence (2.7) holds, then we can conclude that $\gamma = 0$ is a feasible solution to (2.8). This means $\mathsf{WC}-\mathsf{VaR}_{\epsilon}(\boldsymbol{L}(\boldsymbol{\xi})) \leq 0$ holds. Next, we prove that the converse implication holds. In the previous work of Pagnoncelli et al. (2009), they conclude that the mapping function from γ to $\mathbb{P}(\boldsymbol{L}(\boldsymbol{\xi}) \leq \gamma)$ is upper semicontinuous for any fixed distribution $\mathbb{P} \in \mathcal{D}$. Here, we can also derive that the mapping function from γ to $\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}(\boldsymbol{L}(\boldsymbol{\xi}) \leq \gamma)$ is upper semicontinuous. Observe that if $\mathsf{WC}-\mathsf{VaR}_{\epsilon}(\boldsymbol{L}(\boldsymbol{\xi})) \leq 0$ holds, there exists a sequence $\{\gamma_n\}_{n\in\mathbb{N}}$ which converges to 0 and is a feasible solution to (2.8) as well. This means

$$\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right) \leq 0\right) \geq \lim_{n\to\infty} \sup \inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right) \leq \gamma_n\right) \geq 1-\epsilon.$$

Thus, the left hand side of (2.7) holds as well.

Then, we need to prove that

$$\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\text{−CVaR}_{\epsilon}\left(\boldsymbol{L}(\boldsymbol{\xi})\right) = \mathsf{WC}-\mathsf{VaR}_{\epsilon}\left(\boldsymbol{L}(\boldsymbol{\xi})\right).$$

Observe that (2.8) can be equivalently expressed as

$$\mathrm{WC}-\mathrm{VaR}_{\epsilon}\left(\boldsymbol{L}(\boldsymbol{\xi})\right)=\inf_{\gamma\in\mathbb{R}}\left\{\gamma:\sup_{\mathbb{P}\in\mathcal{D}}\mathbb{P}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right)>\gamma\right)\leq\epsilon\right\}.$$

Next, we simplify $\sup\limits_{\mathbb{P}\in\mathcal{D}}\mathbb{P}(\boldsymbol{L}(\boldsymbol{\xi})>\gamma)\leq\epsilon$ according to Lemma 2 in Appendix II. Then, equation (2.8) can be reformulated as

$$\begin{aligned}
\mathrm{WC}-\mathrm{VaR}_{\epsilon}\left(\boldsymbol{L}(\boldsymbol{\xi})\right)=\quad \inf_{\gamma,\tau,\mathbf{s},\mathbf{q},\mathbf{Q}}\quad & \gamma && (2.9a)\\
\text{s.t.}\quad & \frac{1}{\epsilon}\left(\mathbf{s}+2\boldsymbol{\mu}^{\top}\mathbf{q}+\langle\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^{\top},\mathbf{Q}\rangle\right)\leq\tau, && (2.9b)\\
& \mathbf{s}+2\boldsymbol{\xi}^{\top}\mathbf{q}+\boldsymbol{\xi}^{\top}\mathbf{Q}\boldsymbol{\xi}\geq 0,\ \forall\boldsymbol{\xi}\in\mathbb{R}^{m}, && (2.9c)\\
& -\tau+\mathbf{s}+2\boldsymbol{\xi}^{\top}\mathbf{q}+\boldsymbol{\xi}^{\top}\mathbf{Q}\boldsymbol{\xi}+\gamma-\boldsymbol{L}\left(\boldsymbol{\xi}\right)\geq 0,\\
& \qquad\forall\boldsymbol{\xi}\in\mathbb{R}^{m}, && (2.9d)\\
& \gamma\in\mathbb{R},\tau\in\mathbb{R},\tau\geq 0,\mathbf{s}\in\mathbb{R}, && (2.9e)\\
& \mathbf{q}\in\mathbb{R}^{m},\mathbf{Q}\in\mathbb{R}^{m\times m},\mathbf{Q}\succeq 0. && (2.9f)
\end{aligned}$$

Here, we define $\beta=\gamma-\tau$, which enables us to eliminate variable γ. Problem (2.9) is equivalent to

$$\begin{aligned}
\mathrm{WC}-\mathrm{VaR}_{\epsilon}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right)\right)=\quad \inf_{\beta,\tau,\mathbf{s},\mathbf{q},\mathbf{Q}}\quad & \beta+\tau\\
\text{s.t.}\quad & \frac{1}{\epsilon}\left(\mathbf{s}+2\boldsymbol{\mu}^{\top}\mathbf{q}+\langle\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^{\top},\mathbf{Q}\rangle\right)\leq\tau,\\
& \mathbf{s}+2\boldsymbol{\xi}^{\top}\mathbf{q}+\boldsymbol{\xi}^{\top}\mathbf{Q}\boldsymbol{\xi}\geq 0,\ \forall\boldsymbol{\xi}\in\mathbb{R}^{m},\\
& \beta+\mathbf{s}+2\boldsymbol{\xi}^{\top}\mathbf{q}+\boldsymbol{\xi}^{\top}\mathbf{Q}\boldsymbol{\xi}-\boldsymbol{L}\left(\boldsymbol{\xi}\right)\geq 0,\ \forall\boldsymbol{\xi}\in\mathbb{R}^{m},\\
& \beta\in\mathbb{R},\tau\in\mathbb{R},\mathbf{s}\in\mathbb{R},\\
& \mathbf{q}\in\mathbb{R}^{m},\mathbf{Q}\in\mathbb{R}^{m\times m},\mathbf{Q}\succeq 0.
\end{aligned}$$

Note that at optimality $\tau=\frac{1}{\epsilon}(\mathbf{s}+2\boldsymbol{\mu}^{\top}\mathbf{q}+\langle\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^{\top},\mathbf{Q}\rangle)$, which gives us the final expression of $\mathrm{WC}-\mathrm{VaR}_{\epsilon}(\boldsymbol{L}(\boldsymbol{\xi}))$

$$\begin{aligned}
\mathrm{WC}-\mathrm{VaR}_{\epsilon}\left(\boldsymbol{L}\left(\boldsymbol{\xi}\right)\right)=\quad \inf_{\beta,\mathbf{s},\mathbf{q},\mathbf{Q}}\quad & \beta+\frac{1}{\epsilon}\left(\mathbf{s}+2\boldsymbol{\mu}^{\top}\mathbf{q}+\langle\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^{\top},\mathbf{Q}\rangle\right) && (2.11a)\\
\text{s.t.}\quad & \mathbf{s}+2\boldsymbol{\xi}^{\top}\mathbf{q}+\boldsymbol{\xi}^{\top}\mathbf{Q}\boldsymbol{\xi}\geq 0,\ \forall\boldsymbol{\xi}\in\mathbb{R}^{m}, && (2.11b)
\end{aligned}$$

$$\boldsymbol{\beta} + \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} - \boldsymbol{L}(\boldsymbol{\xi}) \geq 0, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^m, \tag{2.11c}$$

$$\boldsymbol{\beta} \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \tag{2.11d}$$

$$\mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0. \tag{2.11e}$$

Note that by Lemma 1 we have

$$\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\text{−CVaR}_\epsilon\left(\boldsymbol{L}(\boldsymbol{\xi})\right) = \inf_{\boldsymbol{\beta}\in\mathbb{R}} \left\{\boldsymbol{\beta} + \frac{1}{\epsilon}\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}}\left[\left(\boldsymbol{L}(\boldsymbol{\xi}) - \boldsymbol{\beta}\right)^+\right]\right\} \tag{2.12a}$$

$$= \inf \quad \boldsymbol{\beta} + \frac{1}{\epsilon}\left(\mathbf{s} + 2\boldsymbol{\mu}^\top\mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}\rangle\right) \tag{2.12b}$$

$$\text{s.t.} \quad \mathbf{s} + 2\boldsymbol{\xi}^\top\mathbf{q} + \boldsymbol{\xi}^\top\mathbf{Q}\boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \tag{2.12c}$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top\mathbf{q} + \boldsymbol{\xi}^\top\mathbf{Q}\boldsymbol{\xi} \geq \boldsymbol{L}(\boldsymbol{\xi}) - \boldsymbol{\beta}, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^m, \tag{2.12d}$$

$$\boldsymbol{\beta} \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \tag{2.12e}$$

$$\mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0, \tag{2.12f}$$

which is just the same as (2.11). This completes the proof. □

The above techniques enable us to derive conservative approximations or exact reformulations for DRCC based on the property of the loss function $\boldsymbol{L}(\boldsymbol{\xi})$. Note that other alternative approaches can also be used for approximating DRCC. For example, Jiang and Xie (2023) approximate DRCC using ALSO-X# method, where a data-driven q−Wasserstein ambiguity set and an uncertain constraint system specified by multiple linear constraints are considered. Building upon the above theoretical framework, our research aims at developing computationally efficient approximations and reformulations for DRCCLMI. This is achieved by transforming Problem (2.4) into tractable SDP Problems by using the above techniques.

Chapter 3

General Case

3.1 Inner Approximation Using CVaR

In this chapter, we consider a general formulation of Linear Matrix Inequality $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0$. We apply the CVaR approximation method introduced in Chapter 2 to derive an inner approximation for DRCCLMI (2.4b).

PROPOSITION 2. The following formulation provides an inner approximation for Problem (2.4)

$$
\begin{aligned}
\min_{\mathbf{x},\lambda,\boldsymbol{\beta},\mathbf{s},\mathbf{q},\mathbf{Q}} \quad & \mathbf{c}^\top \mathbf{x} && \text{(3.1a)}\\
\text{s.t.} \quad & \boldsymbol{\beta} + \frac{1}{\epsilon}\left(\mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}\rangle\right) \leq 0, && \text{(3.1b)}\\
& \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} \geq 0,\ \forall \boldsymbol{\xi} \in \mathcal{S}, && \text{(3.1c)}\\
& \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \boldsymbol{\beta} + \lambda(\boldsymbol{\xi}) \geq 0, \forall \boldsymbol{\xi} \in \mathcal{S}, && \text{(3.1d)}\\
& \mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}) - \lambda(\boldsymbol{\xi})\mathbf{I}_d \succeq 0, \forall \boldsymbol{\xi} \in \mathcal{S}, && \text{(3.1e)}\\
& \boldsymbol{\beta} \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0, && \text{(3.1f)}
\end{aligned}
$$

where $\lambda(\boldsymbol{\xi})$ is a function mapping from $\mathbb{R}^m$ to $\mathbb{R}$.

Proof. Given the definition of positive semidefinite (PSD), the linear matrix inequality $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0$ can equivalently be expressed as

$$\begin{aligned}
&\bar{\mathbf{z}}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \right) \bar{\mathbf{z}} \geq 0, \ \forall \bar{\mathbf{z}} \in \mathbb{R}^d, \\
\iff \quad &\left(\frac{\bar{\mathbf{z}}}{\|\bar{\mathbf{z}}\|} \right)^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \right) \left(\frac{\bar{\mathbf{z}}}{\|\bar{\mathbf{z}}\|} \right) \geq 0, \ \forall \bar{\mathbf{z}} \in \mathbb{R}^d. \quad (3.2)
\end{aligned}$$

If we let $\mathbf{z} = \frac{\bar{\mathbf{z}}}{\|\bar{\mathbf{z}}\|}$, then we have $\mathbf{z} \in \mathbb{R}^d$ and $\|\mathbf{z}\| = 1$. Constraint (3.2) can be represented by:

$$\begin{aligned}
&\mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} \geq 0, \ \forall \mathbf{z} \in \mathbb{R}^d, \|\mathbf{z}\| = 1, \\
\iff \quad &\min_{\mathbf{z} \in \mathbb{R}^d, \|\mathbf{z}\|=1} \ \mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} \geq 0.
\end{aligned}$$

Starting from constraint (2.4b), if we replace $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0$ by $\min\limits_{\|\mathbf{z}\|=1} \ \mathbf{z}^\top (\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}))\mathbf{z} \geq 0$, then we have:

$$\begin{aligned}
&\inf_{\mathbb{P} \in \mathcal{D}} \ \mathbb{P} \left\{ \mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0 \right\} \geq 1 - \epsilon, \\
\iff \quad &\inf_{\mathbb{P} \in \mathcal{D}} \ \mathbb{P} \left\{ \min_{\|\mathbf{z}\|=1} \ \mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} \geq 0 \right\} \geq 1 - \epsilon, \\
\iff \quad &\inf_{\mathbb{P} \in \mathcal{D}} \ \mathbb{P} \left\{ \max_{\|\mathbf{z}\|=1} \ -\mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} \leq 0 \right\} \geq 1 - \epsilon. \quad (3.3)
\end{aligned}$$

The second equivalence is reasonable because minimizing a function and maximizing its negative yield the same result under identical constraints. Observe that $\max\limits_{\|\mathbf{z}\|=1} \ -\mathbf{z}^\top (\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}))\mathbf{z}$ in constraint (3.3) is neither concave nor quadratic in $\boldsymbol{\xi}$. If we consider it as the loss function $\boldsymbol{L}(\boldsymbol{\xi})$, then the following distributionally robust CVaR constraint (3.4) provides a conservative approximation for constraint (3.3):

$$\sup_{\mathbb{P} \in \mathcal{D}} \mathbb{P}\text{-CVaR}_\epsilon \left(\max_{\|\mathbf{z}\|=1} \ -\mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} \right) \leq 0. \quad (3.4)$$

According to the definition of CVaR, constraint (3.4) can be further expressed as:

$$\sup_{\mathbb{P}\in\mathcal{D}} \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}} \left(\left(\max_{\|\mathbf{z}\|=1} -\mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} - \beta \right)^+ \right) \right\} \leq 0, \quad (3.5)$$

$$\Longleftarrow \quad \inf_{\beta\in\mathbb{R}} \left\{ \beta + \frac{1}{\epsilon} \sup_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}} \left(\left(\max_{\|\mathbf{z}\|=1} -\mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} - \beta \right)^+ \right) \right\} \leq 0. \quad (3.6)$$

Note that $\Longleftarrow$ implies an inner approximation after the interchange of the supremum and infimum. This is based on the max–min inequality (Boyd and Vandenberghe 2004). By applying Lemma 1, the subordinate problem $\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}}((\max_{\|\mathbf{z}\|=1} -\mathbf{z}^\top(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}))\mathbf{z} - \beta)^+)$ in constraint (3.6) can be expressed as:

$$\begin{aligned}
\inf_{\beta,\mathbf{s},\mathbf{q},\mathbf{Q}} \quad & \mathbf{s} + 2\mu^\top \mathbf{q} + \langle \Sigma + \mu\mu^\top, \mathbf{Q} \rangle \\
\text{s.t.} \quad & \mathbf{s} + 2\xi^\top \mathbf{q} + \xi^\top \mathbf{Q} \xi \geq \max_{\|\mathbf{z}\|=1} -\mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} - \beta, \\
& \qquad \forall \xi \in \mathcal{S}, \\
& \mathbf{s} + 2\xi^\top \mathbf{q} + \xi^\top \mathbf{Q} \xi \geq 0, \ \forall \xi \in \mathcal{S}, \\
& \beta \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0.
\end{aligned}$$

Combining the above problem into constraint (3.6), the following constraints provide a conservative approximation for constraint (2.4b):

$$\beta + \frac{1}{\epsilon}\left(\mathbf{s} + 2\mu^\top \mathbf{q} + \langle \Sigma + \mu\mu^\top, \mathbf{Q} \rangle \right) \leq 0, \quad (3.7a)$$

$$\mathbf{s} + 2\xi^\top \mathbf{q} + \xi^\top \mathbf{Q} \xi + \beta + \min_{\|\mathbf{z}\|=1} \mathbf{z}^\top \left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}) \right) \mathbf{z} \geq 0, \quad \forall \xi \in \mathcal{S}, \quad (3.7b)$$

$$\mathbf{s} + 2\xi^\top \mathbf{q} + \xi^\top \mathbf{Q} \xi \geq 0, \ \forall \xi \in \mathcal{S}, \quad (3.7c)$$

$$\beta \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0. \quad (3.7d)$$

Since $\min_{\|\mathbf{z}\|=1} \mathbf{z}^\top(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x}))\mathbf{z}$ is nothing but the minimum eigenvalue of matrix $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x})$, constraint (3.7b) can be equivalently rewritten as:

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \boldsymbol{\beta} + \lambda_{\min}\left(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x})\right) \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{S}. \quad (3.8)$$

Here, $\lambda_{\min}(\cdot)$ represents the minimum eigenvalue of a matrix. Note that $\lambda_{\min}(\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}))$ is related to the realisation of $\boldsymbol{\xi}$. Next, we define a function $\lambda(\boldsymbol{\xi}) : \mathbb{R}^m \rightarrow \mathbb{R}$ to characterise the relationship between $\boldsymbol{\xi}$ and the minimum eigenvalue. Specifically, constraint (3.8) can be equivalently reformulated as:

$$\begin{aligned}
&\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \boldsymbol{\beta} + \lambda(\boldsymbol{\xi}) \geq 0, \forall \boldsymbol{\xi} \in \mathcal{S}, \\
&\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) - \lambda(\boldsymbol{\xi}) \mathbf{I}_d \succeq 0, \forall \boldsymbol{\xi} \in \mathcal{S}.
\end{aligned}$$

This completes the proof. □

3.2 Inner Approximation Using Recourse Functions

A prevalent method to address the intractability of $\lambda(\boldsymbol{\xi})$ in Proposition 2 is to specifically define this wait-and-see decisions in (3.1d) and (3.1e) to some specific function with respect to the random variable $\boldsymbol{\xi}$. For example, Linear Decision Rules (LDR) and Quadratic Decision Rules (QDR) restrict the function to be linear and quadratic in $\boldsymbol{\xi}$, respectively, see, e.g., Ben-Tal et al. (2004) and Zhen et al. (2022). In this subsection, we derive a conservative approximation for problem (3.1) through Piecewise Linear Decision Rule (PLDR).

Here, we partition the original support set $\mathcal{S}$ in problem (3.1) into K subsets $\mathcal{S}_1, \ldots, \mathcal{S}_K$. Specifically, the way we construct the subset partitions is corresponding to Voronoi regions (Fan and Hanasusanto 2024). We set $\boldsymbol{\xi}'_k, \ldots, \boldsymbol{\xi}'_K$ to be K constructor points, which are randomly chosen from the sample points $\{\hat{\boldsymbol{\xi}}_i\}_{k\in[N]}$. If we start from the k-th constructor point $\{\boldsymbol{\xi}'_k\}_{k\in[K]}$, the Voronoi region of subset $\mathcal{S}_k$ is expressed as all the points in support set $\mathcal{S}$ who has the closest Euclidean distance to constructor $\boldsymbol{\xi}'_k$ compared to the other constructor points. We define the k-th subset as follows:

$$\begin{aligned} \mathcal{S}_k &= \left\{ \boldsymbol{\xi} \in \mathcal{S} : \|\boldsymbol{\xi} - \boldsymbol{\xi}'_k\| \leq \|\boldsymbol{\xi} - \boldsymbol{\xi}'_i\|, \forall i \in [K] : i \neq k \right\} \\ &= \left\{ \boldsymbol{\xi} \in \mathcal{S} : 2\left(\boldsymbol{\xi}'_i - \boldsymbol{\xi}'_k\right)^\top \boldsymbol{\xi} \leq \boldsymbol{\xi}'^\top_i \boldsymbol{\xi}'_i - \boldsymbol{\xi}'^\top_k \boldsymbol{\xi}'_k, \forall i \in [K] : i \neq k \right\}. \end{aligned} \tag{3.9}$$

Furthermore, we adopt PLDR in each partition separately. Specifically, the PLDR for partition k is given by $\boldsymbol{a}_k^\top \boldsymbol{\xi} + \boldsymbol{b}_k$, where $\boldsymbol{a}_k \in \mathbb{R}^m$ and $\boldsymbol{b}_k \in \mathbb{R}$. Thus, we express $\lambda(\boldsymbol{\xi})$ for the k-th partition as

$$\lambda_k(\boldsymbol{\xi}) = \boldsymbol{a}_k^\top \boldsymbol{\xi} + \boldsymbol{b}_k, \forall \boldsymbol{\xi} \in \mathcal{S}_k, k \in [K].$$

Next, we derive an inner approximation for Problem (3.1) through PLDR.

PROPOSITION 3. By adopting PLDR within each partition, the following formulation provides an inner approximation for Problem (3.1):

$$\min_{\mathbf{x},\boldsymbol{\beta},\mathbf{s},\mathbf{q},\mathbf{Q},\boldsymbol{a},\boldsymbol{b}} \quad \mathbf{c}^\top \mathbf{x} \tag{3.10a}$$

$$\text{s.t.} \quad \boldsymbol{\beta} + \frac{1}{\epsilon}\left(\mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle\right) \leq 0, \tag{3.10b}$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{S}, \tag{3.10c}$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} + \boldsymbol{\beta} + \boldsymbol{a}_j^\top \boldsymbol{\xi} + \boldsymbol{b}_j \geq 0, \quad \forall \boldsymbol{\xi} \in \mathcal{S}_j, j \in [K], \tag{3.10d}$$

$$\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) - \left(\boldsymbol{a}_j^\top \boldsymbol{\xi} + \boldsymbol{b}_j\right) \mathbf{I}_d \succeq 0, \quad \forall \boldsymbol{\xi} \in \mathcal{S}_j, j \in [K], \tag{3.10e}$$

$$\boldsymbol{\beta} \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m \times m}, \mathbf{Q} \succeq 0, \tag{3.10f}$$

$$\boldsymbol{a}_j \in \mathbb{R}^m, \boldsymbol{b}_j \in \mathbb{R}, j \in [K], \tag{3.10g}$$

where the set $\mathcal{S}_j$ of partition j is defined in (3.9).

3.3 Inner Approximation for LMI with Ellipsoidal Support

According to the definition of subset $\mathcal{S}_k$ in (3.9), all the constraints are linear except $\boldsymbol{\xi} \in \mathcal{S}$, i.e., $\|\boldsymbol{\xi}\| \leq t$. We further need to derive approximations for constraint (3.10e) with this kind of nonlinear constraints embedded, i.e., we consider the following LMI with ellipsoidal support

$$\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m : \|\boldsymbol{\xi}\| \leq t. \tag{3.11}$$

Here, we introduce two types of conservative approximation method.

3.3.1 Inner Approximation for Ellipsoidal Support

The first approximation method is directly derived from Theorem 2.1 proposed by Ben-Tal et al. (1998).

PROPOSITION 4. The following semidefinite program

$$\max \quad \mathbf{c}^\top \mathbf{x} \tag{3.12}$$

$$\text{s.t.} \quad \mathbf{S} + \mathbf{Q} \preceq 2\mathbf{A}_0(\mathbf{x}), \tag{3.13}$$

$$\begin{bmatrix} \mathbf{S} & t\mathbf{A}_1(\mathbf{x}) & t\mathbf{A}_2(\mathbf{x}) & \dots & t\mathbf{A}_m(\mathbf{x}) \\ t\mathbf{A}_1(\mathbf{x}) & \mathbf{Q} & & & \\ t\mathbf{A}_2(\mathbf{x}) & & \mathbf{Q} & & \\ \vdots & & & \ddots & \vdots \\ t\mathbf{A}_m(\mathbf{x}) & & & & \mathbf{Q} \end{bmatrix} \succeq 0, \tag{3.14}$$

in variable $\mathbf{x}$, $\mathbf{S}$, $\mathbf{Q}$ is an approximation for constraint (3.11).

3.3.2 Polyhedral Approximation for Ellipsoidal Uncertainty

In this subsection, we consider the polyhedral approximation for the ellipsoidal support $\|\boldsymbol{\xi}\| \le t$. Our goal is to approximate this conic quadratic constraint

$$\sqrt{\xi_1^2 + \xi_2^2 + \ldots + \xi_m^2} \le t$$

of dimension m by a system of conic quadratic constraints of dimension 3 each. Firstly, we provide a polyhedral approximation for

$$L^2 = \left\{ (y_1, y_2, y_3) \middle| \sqrt{y_1^2 + y_2^2} \le y_3 \right\}. \tag{3.15}$$

We consider a polyhedral approximation for the set (3.15) using the following constraints:

$$\begin{aligned} &\theta_0 \ge |y_1|, \\ &\eta_0 \ge |y_2|, \\ &\cos\left(\frac{\tau\pi}{2^v}\right)\theta_0 + \sin\left(\frac{\tau\pi}{2^v}\right)\eta_0 \le y_3, \forall \tau \in \{0\} \cap \left\{2^{v-1}\right\}. \end{aligned}$$

Then, we can approximate constraint (3.11) by

$$\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0, \ \forall \boldsymbol{\xi} \in \mathcal{U}. \tag{3.16}$$

Here, $\mathcal{U}$ is defined as a set consisting of the intersection of N non-empty polyhedral uncertainty sets. Specifically,

$$\mathcal{U} = \{\boldsymbol{\xi} : \mathbf{C}_1\boldsymbol{\xi} \le \mathbf{c}_1\} \cap \{\boldsymbol{\xi} : \mathbf{C}_2\boldsymbol{\xi} \le \mathbf{c}_2\} \cap \ldots \cap \{\boldsymbol{\xi} : \mathbf{C}_N\boldsymbol{\xi} \le \mathbf{c}_N\}, \tag{3.17}$$

where $\mathbf{C}_1, \ldots, \mathbf{C}_N \in \mathbb{R}^{p\times m}$ and $\mathbf{c}_1, \ldots, \mathbf{c}_N \in \mathbb{R}^p$ are given parameters. Specifically, the way we define $\mathcal{U}$ is based on the polyhedral approximation for a system of L^2 cone. Next, the following proposition gives an inner approximation for (3.16).

PROPOSITION 5. The following constraints provide a conservative approximation for constraint (3.16)

$$\begin{aligned}
&\sum_{k=1}^{N} \mathbf{c}_k^\top \mathbf{v}_k \leq 0, \\
&\mathbf{A}_0(\mathbf{x}) - \sum_{k=1}^{N}\sum_{j=1}^{p} \mathbf{c}_{kj} \mathbf{V}_k^{(j)} \succeq 0, \\
&\sum_{k=1}^{N} \mathbf{C}_{ki}^\top \mathbf{v}_k = 0, i = 1, \ldots, m, \\
&\mathbf{A}_i(\mathbf{x}) + \sum_{k=1}^{N}\sum_{j=1}^{p} \mathbf{C}_{kij} \mathbf{V}_k^{(j)} = 0, i = 1, \ldots, m, \\
&\mathbf{v}_k \geq 0, k \in [N], \\
&\mathbf{V}_k^{(j)} \succeq 0, k \in [N], j = 1, \ldots, p.
\end{aligned}$$

where $\mathbf{v}_k \in \mathbb{R}^p, k \in [N]$ and $\mathbf{V}_k^{(j)} \in \mathbb{R}^{d\times d}, k \in [N], j = 1, \ldots, p$. Here the support $\mathcal{U}$ is defined in (3.17).

Proof. Given the definition of PSD, if $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^m \xi_i \mathbf{A}_i(\mathbf{x})$ is positive semidefinite, then the trace of its product with any PSD matrix is positive. Thus, we can derive:

$$\begin{aligned}
& \forall \boldsymbol{\xi} \in \mathcal{U} : \mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) \succeq 0, \\
\iff\quad & \forall \mathbf{G} \succeq 0, \forall \boldsymbol{\xi} \in \mathcal{U} : \langle \mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}), \mathbf{G} \rangle \geq 0, \\
\iff\quad & \forall \mathbf{G} \succeq 0, \forall \boldsymbol{\xi} \in \mathcal{U} : \langle \mathbf{A}_0(\mathbf{x}), \mathbf{G} \rangle + \sum_{i=1}^{m} \xi_i \langle \mathbf{A}_i(\mathbf{x}), \mathbf{G} \rangle \geq 0, \\
\iff\quad & \forall \mathbf{G} \succeq 0 : \langle \mathbf{A}_0(\mathbf{x}), \mathbf{G} \rangle + \min_{\boldsymbol{\xi} \in \mathcal{U}} \left\{ \sum_{i=1}^{m} \xi_i \langle \mathbf{A}_i(\mathbf{x}), \mathbf{G} \rangle \right\} \geq 0.
\end{aligned}$$

We now assign dual variables $\boldsymbol{\lambda}_k \in \mathbb{R}^p, \boldsymbol{\lambda}_k \geq 0, k \in [N]$ to the k^{th} polyhedral set in $\mathcal{U}$, respectively. If we dualize over $\boldsymbol{\xi}$, then we have

$$\begin{aligned}
& \forall \mathbf{G} \succeq 0 : \\
& \langle \mathbf{A}_0(\mathbf{x}), \mathbf{G} \rangle + \max_{\lambda \geq 0} \left\{ -\sum_{k=1}^{N} \mathbf{c}_k^\top \boldsymbol{\lambda}_k \;\middle|\; \sum_{k=1}^{N} \mathbf{C}_{ki}^\top \boldsymbol{\lambda}_k = -\langle \mathbf{A}_i(\mathbf{x}), \mathbf{G} \rangle, i = 1, \ldots, m \right\} \geq 0,
\end{aligned}$$

$$\iff \quad \exists \boldsymbol{\lambda}_k \geq 0, k \in [N], \forall \mathbf{G} \succeq 0 : \begin{cases} \langle \mathbf{A}_0(\mathbf{x}), \mathbf{G} \rangle \geq \sum_{k=1}^{N} \mathbf{c}_k^\top \boldsymbol{\lambda}_k, \\ \sum_{k=1}^{N} \mathbf{C}_{ki}^\top \boldsymbol{\lambda}_k = -\langle \mathbf{A}_i(\mathbf{x}), \mathbf{G} \rangle, i = 1, \ldots, m. \end{cases} \tag{3.18}$$

Here, we restrict $\boldsymbol{\lambda}$ using linear decision rule:

$$\lambda_{kj}(\mathbf{G}) = \mathbf{v}_{kj} + \langle \mathbf{V}_k^{(j)}, \mathbf{G} \rangle, k \in [N], j = 1, \ldots, p,$$

where $\mathbf{v}_k \in \mathbb{R}^p, k \in [N]$ and $\mathbf{V}_k^{(j)} \in \mathbb{R}^{d \times d}, k \in [N], j = 1, \ldots, p$. Thus, constraint (3.18) can be equivalently expressed as:

$$\forall \mathbf{G} \succeq 0 : \begin{cases} \langle \mathbf{A}_0(\mathbf{x}), \mathbf{G} \rangle \geq \sum_{k=1}^{N} \mathbf{c}_k^\top \mathbf{v}_k + \sum_{k=1}^{N} \sum_{j=1}^{p} \mathbf{c}_{kj} \langle \mathbf{V}_k^{(j)}, \mathbf{G} \rangle, \\ \sum_{k=1}^{N} \mathbf{C}_{ki}^\top \mathbf{v}_k + \sum_{k=1}^{N} \sum_{j=1}^{p} \mathbf{C}_{kij} \langle \mathbf{V}_k^{(j)}, \mathbf{G} \rangle = -\langle \mathbf{A}_i(\mathbf{x}), \mathbf{G} \rangle, i = 1, \ldots, m, \\ \mathbf{v}_{kj} + \langle \mathbf{V}_k^{(j)}, \mathbf{G} \rangle \geq 0, k \in [N], j = 1, \ldots, p, \end{cases}$$

$$\iff \begin{cases} \min_{\mathbf{G} \succeq 0} \langle \mathbf{A}_0(\mathbf{x}) - \sum_{k=1}^{N} \sum_{j=1}^{p} \mathbf{c}_{kj} \mathbf{V}_k^{(j)}, \mathbf{G} \rangle \geq \sum_{k=1}^{N} \mathbf{c}_k^\top \mathbf{v}_k, \\ \sum_{k=1}^{N} \mathbf{C}_{ki}^\top \mathbf{v}_k + \langle \mathbf{A}_i(\mathbf{x}), \mathbf{G} \rangle + \sum_{k=1}^{N} \sum_{j=1}^{p} \mathbf{C}_{kij} \langle \mathbf{V}_k^{(j)}, \mathbf{G} \rangle = 0, \forall \mathbf{G} \succeq 0, i = 1, \ldots, m, \\ \mathbf{v}_{kj} + \min_{\mathbf{G} \succeq 0} \langle \mathbf{V}_k^{(j)}, \mathbf{G} \rangle \geq 0, k \in [N], j = 1, \ldots, p, \end{cases}$$

$$\iff \begin{cases} \sum_{k=1}^{N} \mathbf{c}_k^\top \mathbf{v}_k \leq 0, \\ \mathbf{A}_0(\mathbf{x}) - \sum_{k=1}^{N} \sum_{j=1}^{p} \mathbf{c}_{kj} \mathbf{V}_k^{(j)} \succeq 0, \\ \sum_{k=1}^{N} \mathbf{C}_{ki}^\top \mathbf{v}_k = 0, i = 1, \ldots, m, \\ \mathbf{A}_i(\mathbf{x}) + \sum_{k=1}^{N} \sum_{j=1}^{p} \mathbf{C}_{kij} \mathbf{V}_k^{(j)} = 0, i = 1, \ldots, m, \\ \mathbf{v}_k \geq 0, k \in [N], \\ \mathbf{V}_k^{(j)} \succeq 0, k \in [N], j = 1, \ldots, p. \end{cases}$$

□

Chapter 4

Special Case

4.1 Special Case without Support Information

In the previous chapter, we conduct a detailed exploration of the problem in its general DRCCLMI form. Next, we can gain further insights by studying specific cases where the DRCCLMI is more structured.

We turn our attention to a special case where the inner $d \times d$ symmetric linear matrix, $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x})$, in equation (2.4b) exhibits a specific block structure. In particular, we decompose this matrix into a form : $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x}) = \begin{bmatrix} a_{11}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \\ \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix}$. Here, a_{11} is the element located at the first row and the first column, $\mathbf{a_{12}}$ is a $(d-1)$-dimensional column vector, and $\mathbf{A_{22}}(\mathbf{x})$ is a $(d-1) \times (d-1)$ symmetric submatrix. Further, a_{11} and $\mathbf{a_{12}}$ are affine with respect to $\boldsymbol{\xi}$ and $\mathbf{x}$. In particular, we consider the following special case with DRCCLMI (4.1b):

$$\min_{\mathbf{x} \in \mathcal{X}} \quad \mathbf{c}^\top \mathbf{x} \tag{4.1a}$$

$$\text{s.t.} \quad \inf_{\mathbb{P} \in \mathcal{D}} \mathbb{P} \left\{ \begin{bmatrix} a_{11}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \\ \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \succeq 0 \right\} \geq 1 - \epsilon, \tag{4.1b}$$

where the uncertain set $\mathcal{D}$ is of the form:

$$\mathcal{D}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left\{ \mathbb{P} \middle| \begin{array}{l} \mathbb{E}_{\mathbb{P}}[\boldsymbol{\xi}] = \boldsymbol{\mu} \\ \mathbb{E}_{\mathbb{P}}\left[(\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})^\top\right] \preceq \boldsymbol{\Sigma} \end{array} \right\}. \tag{4.2}$$

By using Proposition 6, we derive an exact reformulation for this special case.

PROPOSITION 6. If the $(d-1) \times (d-1)$ symmetric submatrix $\mathbf{A_{22}}(\mathbf{x})$ is positive definite (PD), then the following problem provides an exact reformulation of Problem (4.1)

$$\min_{\mathbf{x}, \boldsymbol{\beta}, \mathbf{s}, \mathbf{q}, \mathbf{Q}} \quad \mathbf{c}^\top \mathbf{x} \tag{4.3a}$$

$$\text{s.t.} \quad \boldsymbol{\beta} + \frac{1}{\epsilon}\left(\mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle\right) \leq 0, \tag{4.3b}$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} + \boldsymbol{\beta} - \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) + a_{11}(\boldsymbol{\xi}, \mathbf{x}) \geq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \tag{4.3c}$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \tag{4.3d}$$

$$\boldsymbol{\beta} \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m \times m}, \mathbf{Q} \succeq 0. \tag{4.3e}$$

Proof. Given the definition of PSD, we have the following equivalence holds:

$$\begin{aligned} & \begin{bmatrix} a_{11}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \\ \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \succeq 0, \\ \Leftrightarrow \quad & \begin{bmatrix} 1 & \mathbf{z}^\top \end{bmatrix} \begin{bmatrix} a_{11}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \\ \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} 1 & \mathbf{z}^\top \end{bmatrix}^\top \geq 0, \ \forall \mathbf{z} \in \mathbb{R}^{d-1} \\ \Leftrightarrow \quad & \min_{\mathbf{z} \in \mathbb{R}^{d-1}} \begin{bmatrix} 1 & \mathbf{z}^\top \end{bmatrix} \begin{bmatrix} a_{11}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \\ \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} 1 & \mathbf{z}^\top \end{bmatrix}^\top \geq 0. \end{aligned}$$

To prove the first equivalence, we study two possible scenarios for any $(z_0 \in \mathbb{R}, \mathbf{z}^\top \in \mathbb{R}^{d-1})^\top \in \mathbb{R}^d$: (1) if $z_0 = 0$, then $(z_0, \mathbf{z}^\top)\, \mathbf{A}(\boldsymbol{\xi}, \mathbf{x})\, (z_0, \mathbf{z}^\top)^\top = \mathbf{z}^\top \mathbf{A_{22}}(\mathbf{x})\, \mathbf{z} > 0$, because $\mathbf{A_{22}}(\mathbf{x})$ is positive definite; (2) if $z_0 \neq 0$, then we have $(z_0, \mathbf{z}^\top)\, \mathbf{A}(\boldsymbol{\xi}, \mathbf{x})\, (z_0, \mathbf{z}^\top)^\top = z_0^2 \left(1, \frac{\mathbf{z}^\top}{z_0}\right) \mathbf{A}(\boldsymbol{\xi}, \mathbf{x}) \left(1, \frac{\mathbf{z}^\top}{z_0}\right)^\top > 0$. Therefore, $\Rightarrow$ holds and vice versa. Using the above equivalence, we can reformulate constraint (4.1b) as follows:

$$\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left\{\min_{\mathbf{z}\in\mathbb{R}^{d-1}} \begin{bmatrix}1 & \mathbf{z}^\top\end{bmatrix} \begin{bmatrix} a_{11}(\boldsymbol{\xi},\mathbf{x}) & \mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \\ \mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \begin{bmatrix}1 & \mathbf{z}^\top\end{bmatrix}^\top \geq 0\right\} \geq 1-\epsilon, \tag{4.4}$$

$$\iff \inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left\{\min_{\mathbf{z}\in\mathbb{R}^{d-1}} a_{11}(\boldsymbol{\xi},\mathbf{x}) + 2\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{z} + \mathbf{z}^\top \mathbf{A_{22}}(\mathbf{x})\,\mathbf{z} \geq 0\right\} \geq 1-\epsilon. \tag{4.5}$$

Observe that $\min_{\mathbf{z}\in\mathbb{R}^{d-1}} a_{11}(\boldsymbol{\xi},\mathbf{x}) + 2\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{z} + \mathbf{z}^\top \mathbf{A_{22}}(\mathbf{x})\,\mathbf{z}$ is quadratic in $\mathbf{z}$. By taking the first derivative of $\mathbf{z}$, we can get $\mathbf{z}^* = -\mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})$. Then the following constraint can be obtained by substituting $\mathbf{z}^*$ into constraint (4.5):

$$\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left\{\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) \leq 0\right\} \geq 1-\epsilon. \tag{4.6}$$

Note that $\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x})$ in constraint (4.6) is quadratic in $\boldsymbol{\xi}$. If we consider it as the loss function $\boldsymbol{L}(\boldsymbol{\xi})$, then the following distributionally robust CVaR constraint provides an exact reformulation for constraint (4.6) accroding to Proposition 1:

$$\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\text{–CVaR}_\epsilon\left(\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x})\right) \leq 0. \tag{4.7}$$

According to the definition of CVaR, constraint (4.7) can be further expressed as:

$$\sup_{\mathbb{P}\in\mathcal{D}} \inf_{\beta\in\mathbb{R}} \left\{\beta + \frac{1}{\epsilon}\mathbb{E}_{\mathbb{P}}\left(\left(\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) - \beta\right)^+\right)\right\} \leq 0,$$

$$\iff \inf_{\beta\in\mathbb{R}} \left\{\beta + \frac{1}{\epsilon}\sup_{\mathbb{P}\in\mathcal{D}}\mathbb{E}_{\mathbb{P}}\left(\left(\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) - \beta\right)^+\right)\right\} \leq 0. \tag{4.8}$$

Here, the equivalence is due to the saddle point theorem (Shapiro and Kleywegt 2002). By applying Corollary 1, $\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}}((\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) -$

$a_{11}(\boldsymbol{\xi},\mathbf{x})-\boldsymbol{\beta})^{+})$ in constraint (4.8) can equivalently be transformed into the following problem:

$$
\begin{aligned}
\inf_{\mathbf{s},\mathbf{q},\mathbf{Q}} \quad & \mathbf{s}+2\boldsymbol{\mu}^{\top}\mathbf{q}+\langle\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^{\top},\mathbf{Q}\rangle \\
\text{s.t.} \quad & \mathbf{s}+2\boldsymbol{\xi}^{\top}\mathbf{q}+\boldsymbol{\xi}^{\top}\mathbf{Q}\boldsymbol{\xi} \geq \mathbf{a_{12}}\left(\boldsymbol{\xi},\mathbf{x}\right)^{\top}\mathbf{A_{22}}\left(\mathbf{x}\right)^{-1}\mathbf{a_{12}}\left(\boldsymbol{\xi},\mathbf{x}\right)-a_{11}\left(\boldsymbol{\xi},\mathbf{x}\right)-\boldsymbol{\beta}, \\
& \qquad\qquad \forall\boldsymbol{\xi}\in\mathbb{R}^{m}, \\
& \mathbf{s}+2\boldsymbol{\xi}^{\top}\mathbf{q}+\boldsymbol{\xi}^{\top}\mathbf{Q}\boldsymbol{\xi}\geq 0,\ \forall\boldsymbol{\xi}\in\mathbb{R}^{m}, \\
& \mathbf{s}\in\mathbb{R},\mathbf{q}\in\mathbb{R}^{m},\mathbf{Q}\in\mathbb{R}^{m\times m},\mathbf{Q}\succeq 0.
\end{aligned}
$$

Combining the above problem into constraint (4.8), we can get the desired result. This completes the proof. □

EXAMPLE 1. We consider the case where the column vector $\mathbf{a_{12}}$ is equivalent to $\boldsymbol{\xi}$, and a_{11} is affine in $\mathbf{x}$, i.e., $\mathbf{A}_0\left(\mathbf{x}\right)+\sum_{i=1}^{m}\xi_i\mathbf{A}_i\left(\mathbf{x}\right)=\begin{bmatrix} a_{11}\left(\mathbf{x}\right) & \boldsymbol{\xi}^{\top} \\ \boldsymbol{\xi} & \mathbf{A_{22}}\left(\mathbf{x}\right)\end{bmatrix}$. Specifically, we consider the following problem with a special block matrix case of DRCCLMI:

$$
\min_{\mathbf{x}\in\mathcal{X}} \quad \mathbf{c}^{\top}\mathbf{x} \tag{4.9a}
$$

$$
\text{s.t.} \quad \inf_{\mathbb{P}\in\mathcal{D}}\mathbb{P}\left\{\begin{bmatrix} a_{11}\left(\mathbf{x}\right) & \boldsymbol{\xi}^{\top} \\ \boldsymbol{\xi} & \mathbf{A_{22}}\left(\mathbf{x}\right)\end{bmatrix}\succeq 0\right\}\geq 1-\epsilon, \tag{4.9b}
$$

where $\mathcal{D}$ taking the form of (4.2). We can reformulate the above Problem (4.9) into a tractable SDP problem as shown in Proposition 7.

PROPOSITION 7. If $\mathbf{A_{22}}\left(\mathbf{x}\right)$ is PD, then Problem (4.9) can be exactly reformulated by:

$$
\min_{\mathbf{x},\boldsymbol{\beta},\mathbf{s},\mathbf{q},\mathbf{Q},\mathbf{B}} \quad \mathbf{c}^{\top}\mathbf{x} \tag{4.10a}
$$

$$
\text{s.t.} \quad \boldsymbol{\beta}+\frac{1}{\epsilon}\left(\mathbf{s}+2\boldsymbol{\mu}^{\top}\mathbf{q}+\langle\boldsymbol{\Sigma}+\boldsymbol{\mu}\boldsymbol{\mu}^{\top},\mathbf{Q}\rangle\right)\leq 0, \tag{4.10b}
$$

$$
\begin{bmatrix} \mathbf{s}+\boldsymbol{\beta}+a_{11}\left(\mathbf{x}\right) & \mathbf{q}^{\top} \\ \mathbf{q} & \mathbf{Q}-\mathbf{B}\end{bmatrix}\succeq 0, \tag{4.10c}
$$

$$\begin{bmatrix} \mathbf{s} & \mathbf{q}^\top \\ \mathbf{q} & \mathbf{Q} \end{bmatrix} \succeq 0, \tag{4.10d}$$

$$\begin{bmatrix} \mathbf{B} & \mathbf{I} \\ \mathbf{I} & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \succeq 0, \tag{4.10e}$$

$$\boldsymbol{\beta} \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{B} \in \mathbb{S}^{d-1}, \mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0 \tag{4.10f}$$

Proof. Observe that Problem (4.9) is a special case of Problem (4.1). If we plug in $\mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) = \boldsymbol{\xi}$ and $a_{11}(\boldsymbol{\xi}, \mathbf{x}) = a_{11}(\mathbf{x})$ in Proposition 6, then Problem (4.9) is equivalent to

$$\begin{aligned}
\min_{\mathbf{x},\boldsymbol{\beta},\mathbf{s},\mathbf{q},\mathbf{Q}} \quad & \mathbf{c}^\top \mathbf{x} && (4.11a)\\
\text{s.t.} \quad & \boldsymbol{\beta} + \frac{1}{\epsilon}\left(\mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}\rangle\right) \le 0, && (4.11b)\\
& \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \boldsymbol{\beta} - \boldsymbol{\xi}^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\boldsymbol{\xi} + a_{11}(\mathbf{x}) \ge 0, \\
& \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^m, && (4.11c)\\
& \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} \ge 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, && (4.11d)\\
& \boldsymbol{\beta} \in \mathbb{R}, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0. && (4.11e)
\end{aligned}$$

Observe that constraint (4.11d) can be equivalently represented as:

$$\begin{aligned}
& \begin{bmatrix} 1 & \boldsymbol{\xi}^\top \end{bmatrix} \begin{bmatrix} \mathbf{s} & \mathbf{q}^\top \\ \mathbf{q} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} 1 & \boldsymbol{\xi}^\top \end{bmatrix}^\top \ge 0, \forall \boldsymbol{\xi} \in \mathbb{R}^m, \\
\iff \quad & \begin{bmatrix} \mathbf{s} & \mathbf{q}^\top \\ \mathbf{q} & \mathbf{Q} \end{bmatrix} \succeq 0.
\end{aligned}$$

Moreover, constraint (4.11c) can be equivalently reformulated as:

$$\begin{aligned}
& \begin{bmatrix} 1 & \boldsymbol{\xi}^\top \end{bmatrix} \begin{bmatrix} \mathbf{s} + \boldsymbol{\beta} + a_{11}(\mathbf{x}) & \mathbf{q}^\top \\ \mathbf{q} & \mathbf{Q} - \mathbf{A_{22}}(\mathbf{x})^{-1} \end{bmatrix} \begin{bmatrix} 1 & \boldsymbol{\xi}^\top \end{bmatrix}^\top \ge 0, \forall \boldsymbol{\xi} \in \mathbb{R}^m, \\
\iff \quad & \begin{bmatrix} \mathbf{s} + \boldsymbol{\beta} + a_{11}(\mathbf{x}) & \mathbf{q}^\top \\ \mathbf{q} & \mathbf{Q} - \mathbf{A_{22}}(\mathbf{x})^{-1} \end{bmatrix} \succeq 0. && (4.12)
\end{aligned}$$

Now we introduce a new variable $\mathbf{B} \in \mathbb{S}^{d-1}$, and let $\mathbf{B} \succeq \mathbf{A_{22}}(\mathbf{x})^{-1}$. Thus constraint (4.12) can be represented by the following constraint:

$$\begin{bmatrix} \mathbf{s} + \boldsymbol{\beta} + a_{11}(\mathbf{x}) & \mathbf{q}^\top \\ \mathbf{q} & \mathbf{Q} - \mathbf{B} \end{bmatrix} \succeq 0. \tag{4.13}$$

Furthermore, according to Schur's complement, $\mathbf{B} \succeq \mathbf{A_{22}}(\mathbf{x})^{-1}$ can be represented as $\begin{bmatrix} \mathbf{B} & \mathbf{I} \\ \mathbf{I} & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \succeq 0$, which completes the proof. □

4.2 Special Case with Support Information

In this section, we still consider the following special case problem with DRC-CLMI:

$$\min_{\mathbf{x}\in\mathcal{X}} \quad \mathbf{c}^\top \mathbf{x} \tag{4.14a}$$

$$\text{s.t.} \quad \inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left\{ \begin{bmatrix} a_{11}(\boldsymbol{\xi},\mathbf{x}) & \mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \\ \mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \succeq 0 \right\} \geq 1-\epsilon, \tag{4.14b}$$

but the uncertainty set is taking the form of (2.3), i.e., a support set $\mathcal{S}$ is included within the ambiguity set $\mathcal{D}$.

By using Proposition 8, we derive an inner approximation for the above problem.

PROPOSITION 8. If the $(d-1)\times(d-1)$ symmetric submatrix $\mathbf{A_{22}}(\mathbf{x})$ is PD, then the following formulation provides an inner approximation for Problem (4.14)

$$\min_{\mathbf{x},\mathbf{m},\mathbf{s}',\mathbf{q}',\mathbf{Q}'} \quad \mathbf{c}^\top \mathbf{x} \tag{4.15a}$$

$$\text{s.t.} \quad \mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \leq \mathbf{m}\epsilon, \tag{4.15b}$$

$$\mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} \geq 0 \quad \forall \boldsymbol{\xi} \in \mathcal{S}, \tag{4.15c}$$

$$-\mathbf{m} + \mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} + a_{11}(\boldsymbol{\xi},\mathbf{x}) - \mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) \geq 0, \forall \boldsymbol{\xi} \in \mathcal{S} \tag{4.15d}$$

$$\mathbf{m} \in \mathbb{R}, \mathbf{m} \geq 0, \mathbf{s}' \in \mathbb{R}, \tag{4.15e}$$

$$\mathbf{q}' \in \mathbb{R}^m, \mathbf{Q}' \in \mathbb{R}^{m\times m}, \mathbf{Q}' \succeq 0. \tag{4.15f}$$

Proof. Starting from constraint (4.6), the key idea here is to replace this constraint with

$$\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left\{\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) < 0\right\} \geq 1-\epsilon. \tag{4.16}$$

This is stricter than constraint (4.6) and more relaxed than

$$\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left\{\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) \leq \eta\right\} \geq 1-\epsilon,$$

for any $\eta < 0$. Observe that (4.16) is equivalent to:

$$\sup_{\mathbb{P}\in\mathcal{D}} \mathbb{E}_{\mathbb{P}}\left[\mathbb{1}\left\{\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) \geq 0\right\}\right] \leq \epsilon, \tag{4.17}$$

where $\mathbb{1}(\cdot)$ is the indicator function. By duality theory, the left hand side problem of constraint (4.17) can be rewritten by

$$\begin{aligned}
\max_{\mathbb{P}\in\mathcal{D}} \quad & \int_{\mathcal{S}} \mathbb{1}\left\{\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) \geq 0\right\} \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) && (4.18a)\\
\text{s.t.} \quad & \int_{\mathcal{S}} \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) = 1, && (4.18b)\\
& \int_{\mathcal{S}} \boldsymbol{\xi}\mathbf{d}\mathbb{P}(\boldsymbol{\xi}) = \boldsymbol{\mu}, && (4.18c)\\
& \int_{\mathcal{S}} (\boldsymbol{\xi}-\boldsymbol{\mu})(\boldsymbol{\xi}-\boldsymbol{\mu})^\top \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) \preceq \boldsymbol{\Sigma}. && (4.18d)
\end{aligned}$$

Considering $\mathbf{s} \in \mathbb{R}$, $\mathbf{q} \in \mathbb{R}^m$, $\mathbf{Q} \in \mathbb{R}^{m\times m}$ and $\mathbf{Q} \succeq 0$ as the Lagrangian multipliers of constraints (4.18b), (4.18c), and (4.18d), respectively, we formulate the following problem as the Lagrangian dual problem of (4.18):

$$\begin{aligned}
\min_{\mathbf{s},\mathbf{q},\mathbf{Q}} \quad & \mathbf{s} + 2\boldsymbol{\mu}^\top\mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}\rangle \\
\text{s.t.} \quad & -\mathbb{1}\left\{\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1}\mathbf{a_{12}}(\boldsymbol{\xi},\mathbf{x}) - a_{11}(\boldsymbol{\xi},\mathbf{x}) \geq 0\right\} \\
& \qquad + \mathbf{s} + 2\boldsymbol{\xi}^\top\mathbf{q} + \boldsymbol{\xi}^\top\mathbf{Q}\boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{S}, \\
& \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{Q} \succeq 0.
\end{aligned}$$

Combining the above problem, constraint (4.17) can be equivalently represented as

$$\mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle \leq \epsilon, \tag{4.19a}$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{S}, \tag{4.19b}$$

$$\begin{aligned} &-1 + \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} \geq 0, \\ &\forall \boldsymbol{\xi} \in \mathcal{S} \cap \left\{ \boldsymbol{\xi} \in \mathbb{R}^m | \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) - a_{11}(\boldsymbol{\xi}, \mathbf{x}) \geq 0 \right\} \end{aligned} \tag{4.19c}$$

Studying more closely the third constraint (4.19c), we can reformulate it as

$$\begin{aligned} \min_{\boldsymbol{\xi} \in \mathcal{S}} \sup_{\lambda \geq 0} \ & -1 + \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \\ & \lambda \left(a_{11}(\boldsymbol{\xi}, \mathbf{x}) - \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) \right) \geq 0. \end{aligned}$$

Then the above constraint can be approximated by the following constraint

$$\begin{aligned} \sup_{\lambda \geq 0} \min_{\boldsymbol{\xi} \in \mathcal{S}} \ & -1 + \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \\ & \lambda \left(a_{11}(\boldsymbol{\xi}, \mathbf{x}) - \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) \right) \geq 0. \end{aligned}$$

Since $\mathcal{S}$ is compact, this can be further reformulated as

$$\begin{aligned} \sup_{\lambda > 0} \min_{\boldsymbol{\xi} \in \mathcal{S}} \ & -1 + \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \\ & \lambda \left(a_{11}(\boldsymbol{\xi}, \mathbf{x}) - \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) \right) \geq 0, \end{aligned}$$

If we let $\mathbf{m} := (1/\lambda)$, $\mathbf{s}' := (1/\lambda)\,\mathbf{s}$, $\mathbf{q}' := (1/\lambda)\,\mathbf{q}$, $\mathbf{Q}' := (1/\lambda)\,\mathbf{Q}$, then constraints (4.19a), (4.19b) and (4.19c) can therefore be approximated as

$$\mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \leq \mathbf{m}\epsilon, \tag{4.23a}$$

$$\mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}'\boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{S}, \tag{4.23b}$$

$$\begin{aligned} & -\mathbf{m} + \mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}'\boldsymbol{\xi} + a_{11}(\boldsymbol{\xi}, \mathbf{x}) - \\ & \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x})^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \mathbf{a_{12}}(\boldsymbol{\xi}, \mathbf{x}) \geq 0, \forall \boldsymbol{\xi} \in \mathcal{S}. \end{aligned} \tag{4.23c}$$

This completes the proof. □

EXAMPLE 2. We consider the case where $\mathbf{A}_0(\mathbf{x}) + \sum_{i=1}^{m} \xi_i \mathbf{A}_i(\mathbf{x})$ in constraint (2.4b) is equal to $\begin{bmatrix} a_{11}(\mathbf{x}) & \boldsymbol{\xi}^\top \\ \boldsymbol{\xi} & \mathbf{A}_{22}(\mathbf{x}) \end{bmatrix}$, where $\mathbf{A}_{22}(\mathbf{x})$ is symmetric matrix. Specifically, we consider the following problem with a special block matrix case of DRCCLMI:

$$\min_{\mathbf{x}\in\mathcal{X}} \quad \mathbf{c}^\top \mathbf{x} \tag{4.24a}$$

$$\text{s.t.} \quad \inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P}\left\{ \begin{bmatrix} a_{11}(\mathbf{x}) & \boldsymbol{\xi}^\top \\ \boldsymbol{\xi} & \mathbf{A}_{22}(\mathbf{x}) \end{bmatrix} \succeq 0 \right\} \geq 1-\epsilon, \tag{4.24b}$$

where $\mathcal{D}$ is taking the form of (2.3). The above Problem (4.24) can be approximated by a tractable SDP problem as shown in Proposition 9.

PROPOSITION 9. If $\mathbf{A}_{22}(\mathbf{x})$ is positive definite, and the support information is taking the form of $\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^m : \|\boldsymbol{\xi}\| \leq t\}$, then Problem (4.24) can be approximated by

$$\min_{\mathbf{x},\mathbf{m},\delta,\gamma,\mathbf{s}',\mathbf{q}',\mathbf{B},\mathbf{Q}'} \quad \mathbf{c}^\top \mathbf{x} \tag{4.25a}$$

$$\text{s.t.} \quad \mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \leq \mathbf{m}\epsilon, \tag{4.25b}$$

$$\begin{bmatrix} \mathbf{Q}' + \delta\mathbf{I} & \mathbf{q}' \\ \mathbf{q}'^\top & \mathbf{s}' - t^2\delta \end{bmatrix} \succeq 0, \tag{4.25c}$$

$$\begin{bmatrix} \mathbf{Q}' - \mathbf{B} + \gamma\mathbf{I} & \mathbf{q}' \\ \mathbf{q}'^\top & -\mathbf{m} + \mathbf{s}' + a_{11}(\mathbf{x}) - t^2\gamma \end{bmatrix} \succeq 0, \tag{4.25d}$$

$$\begin{bmatrix} \mathbf{B} & \mathbf{I} \\ \mathbf{I} & \mathbf{A}_{22}(\mathbf{x}) \end{bmatrix} \succeq 0, \tag{4.25e}$$

$$\mathbf{m}, \delta, \gamma \in \mathbb{R}, \ \mathbf{m}, \delta, \gamma \geq 0, \tag{4.25f}$$

$$\mathbf{s}' \in \mathbb{R}, \mathbf{q}' \in \mathbb{R}^m, \mathbf{B} \in \mathbb{S}^{d-1}, \mathbf{Q}' \in \mathbb{R}^{m\times m}, \mathbf{Q}' \succeq 0. \tag{4.25g}$$

Proof. According to Proposition 8, if we plug in $\mathbf{a}_{12}(\boldsymbol{\xi}, \mathbf{x}) = \boldsymbol{\xi}$ and $a_{11}(\boldsymbol{\xi}, \mathbf{x}) = a_{11}(\mathbf{x})$, then we have the following approximation

$$\min_{\mathbf{x},\mathbf{m},\mathbf{s}',\mathbf{q}',\mathbf{Q}'} \quad \mathbf{c}^\top \mathbf{x}$$

$$\text{s.t.} \quad \mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \leq \mathbf{m}\epsilon,$$

$$\mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} \geq 0 \ \ \forall \boldsymbol{\xi} \in \mathcal{S},$$
$$-\mathbf{m} + \mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} + a_{11}(\mathbf{x}) - \boldsymbol{\xi}^\top \mathbf{A_{22}}(\mathbf{x})^{-1} \boldsymbol{\xi} \geq 0 \ \ \forall \boldsymbol{\xi} \in \mathcal{S},$$
$$\mathbf{m} \in \mathbb{R}, \mathbf{s}' \in \mathbb{R}, \mathbf{q}' \in \mathbb{R}^m, \mathbf{Q}' \in \mathbb{R}^{m \times m} \mathbf{Q}' \succeq 0, \mathbf{m} \geq 0.$$

Since we consider the support information taking the form of $\mathcal{S} = \{\boldsymbol{\xi} \in \mathbb{R}^m : \|\boldsymbol{\xi}\| \leq t\}$, by duality theory, the above problem can be reformulated as

$$\begin{aligned}
\min_{\mathbf{x}, \mathbf{m}, \delta, \gamma, \mathbf{s}', \mathbf{q}', \mathbf{Q}'} \quad & \mathbf{c}^\top \mathbf{x} \\
\text{s.t.} \quad & \mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \leq \mathbf{m}\epsilon, \\
& \begin{bmatrix} \mathbf{Q}' + \delta \mathbf{I} & \mathbf{q}' \\ \mathbf{q}'^\top & \mathbf{s}' - t^2 \delta \end{bmatrix} \succeq 0, \\
& \begin{bmatrix} \mathbf{Q}' - \mathbf{A_{22}}(\mathbf{x})^{-1} + \gamma \mathbf{I} & \mathbf{q}' \\ \mathbf{q}'^\top & -\mathbf{m} + \mathbf{s}' + a_{11}(\mathbf{x}) - t^2 \gamma \end{bmatrix} \succeq 0, \\
& \mathbf{Q}' \succeq 0, \mathbf{m} \geq 0.
\end{aligned}$$

As shown in the proof of Proposition 7, we can introduce a new variable $\mathbf{B} \in \mathbb{S}^{d-1}$, and let $\mathbf{B} \succeq \mathbf{A_{22}}(\mathbf{x})^{-1}$ here. According to Schur's complement, $\mathbf{B} \succeq \mathbf{A_{22}}(\mathbf{x})^{-1}$ can be further represented as $\begin{bmatrix} \mathbf{B} & \mathbf{I} \\ \mathbf{I} & \mathbf{A_{22}}(\mathbf{x}) \end{bmatrix} \succeq 0$, which completes the proof.

□

Chapter 5

Numerical Experiments

We consider the application of our method in two real-world problems: truss topology design problem and calibration problem. Numerical experiments evaluate the efficiency and accuracy of the proposed approximation and reformulation method. Both of these problem formulations are modeled using the CVX 2.2 MATLAB package. We conduct experiments on a computer with an Intel Core i7-4600U 3.40GHz processor and 16.0GB of random access memory using MOSEK solver.

5.1 Truss Topology Design Problem

5.1.1 Formulation

We conduct experiments on a truss topology design (TTD) problem. A truss consists of bars connected at the nodes. The configuration of a truss structure has a large impact on its load carrying capacity, and a proper configuration design can achieve high stiffness with a simple structure while satisfying other requirements. However, as the application scenarios of truss structure become more and more complicated and the construction scale keeps expanding, the difficulty of truss configuration design increases gradually.

The purpose of this problem is to design a truss of a certain weight that can best carry a given load. This problem can be formulated by:

$$\min_{\tau,t} \left\{ \tau : \left[\begin{array}{c|c} 2\tau & \mathbf{h}^T \\ \hline \mathbf{h} & \sum_{i=1}^{m} \mathbf{t}_i \mathbf{b}_i \mathbf{b}_i^T \end{array} \right] \succeq 0, \quad \mathbf{t} \geq 0, \quad \sum_{i=1}^{m} \mathbf{t}_i = 1 \right\}, \tag{5.1}$$

where m is the number of bars; n is freedom degrees of all the nodes and τ is the rigidity of the truss. Variable $\mathbf{t}_i \in \mathbb{R}$ is the weight of the bar i; parameter $\mathbf{b}_i \in \mathbb{R}^n$ is the given geometry of all the nodes; and parameter $\mathbf{h} \in \mathbb{R}^n$ is a given load placed on the nodes.

In reality, a truss is required to withstand not only a given load $\mathbf{h}$, but also some random loads. This kind of random loads affect the nodes used when the truss structure is subjected to different external forces. This has to be taken into account when we design the truss; otherwise, the truss will probably be destroyed by the occasional loads. A way to design a truss that can withstand random loads is to reformulate the original problem. Specifically, we assume that the distribution of occasional loads is within an ambiguity set $\mathcal{D}$ with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. According to Karimi et al. (2021), an equivalent reformulation of Problem (5.1) is:

$$\min_{\rho,\mathbf{t}} \quad \rho \tag{5.2a}$$

$$\text{s.t.} \quad \overbrace{\left[\begin{array}{c|c} 2\hat{\tau} & \mathbf{h}^\top \\ \hline \mathbf{h} & \sum_{i=1}^{l} \mathbf{t}_i \mathbf{b}_i \mathbf{b}_i^\top \end{array} \right]}^{\mathbf{A}(\mathbf{t})} \succeq 0, \tag{5.2b}$$

$$\inf_{\mathbb{P}\in\mathcal{D}} \mathbb{P} \left\{ \underbrace{\left[\begin{array}{c|c} 2\hat{\tau}\rho & \boldsymbol{\xi}^\top \\ \hline \boldsymbol{\xi} & \sum_{i=1}^{l} \mathbf{t}_i \mathbf{b}_i \mathbf{b}_i^\top \end{array} \right]}_{\mathbf{A}_0(\mathbf{t},\rho)+\sum_{l=1}^{n} \xi_l \mathbf{A}_l(\mathbf{t})} \succeq 0 \right\} \geq 1-\epsilon, \tag{5.2c}$$

$$\sum_{i=1}^{m} \mathbf{t}_i = 1, \quad \mathbf{t} \geq 0. \tag{5.2d}$$

5.1.2 Numerical Setting and Results

We consider that the distribution of $\boldsymbol{\xi} = (\xi_1, \dots, \xi_n)$ can be captured by an ambiguity set $\mathcal{D}$ with mean $\boldsymbol{\mu} = \mathbf{1}$ and covariance $\boldsymbol{\Sigma} = \mathbf{I}_n$, which means that $\xi_1, \dots, \xi_n$ are iid distributed. The total number of bars m is set to be 54. The satisfactory parameter is set to $1 - \epsilon = 95\%$.

We report the numerical results in Table 5.1 and 5.2. "Obj" shows the average optimal value and "Time" shows the average computation time. "General" represents that we derive the conservative approximation for this TTD problem and use the proposed method in Chapter 3. "Special" represents that we derive exact reformulation and use SDP-solver to get the optimal solution. "Cheung" represents our benchmark approximation approach proposed by Cheung et al. (2012), which is under the setting that $\boldsymbol{\xi}$ follows a given distribution (Gaussian distribution in our experiment). In other words, they derive an inner approximation for CCLMI, instead of DRCCLMI.

5.1.2.1 Numerical Results without Support Information

We firstly conduct the following experiment to show the performance of our reformulation method in Section 4.1, which assumes that the random variable is without support information. We consider the approximation method of Cheung et al. (2012) as benchmark. The dimension of $\boldsymbol{\xi}$ is considered to be $n = 5, 10, 15, 20$.

Table 5.1: Numerial Results (without Support Information)

$\boldsymbol{\xi}$ Dimension	Special		Cheung		General	
	Obj	Time(sec)	Obj	Time(sec)	Obj	Time(sec)
5	2.45349	0.68	4.33643	0.75	-	7.48
10	2.0698	0.88	8.25792	1.37	-	9.76
15	1.65983	0.97	9.84256	2.21	-	12.33
20	0.99008	1.34	12.1399	2.76	-	14.67

Based on the numerical results provided in Table 5.1, we can draw the following conclusions:

1. Objective Value Performance: Our "Special" method, which uses exact reformulation, consistently yields the best objective values across all dimensions of ξ. This superiority is notable even when compared to Cheung's approximation, which is tailored for a specific distribution (namely, the normal distribution). When employing the "General" case's inner approximation for this problem, the objective value performance deteriorates significantly, reaching infinity across all dimensions.

2. Solution Time Performance: In terms of average CPU time ("Time(sec)"), our "Special" method's solution speed is fairly comparable to that of Cheung's approach. The differences in time are marginal, even as the dimensionality of ξ increases. On the other hand, the "General" case tends to take a substantially longer time, especially as we move to higher dimensions.

5.1.2.2 Numerical Results with Support Information

Next, we conduct experiments where the support information is defined. Here, we employ the technique mentioned in Section 4.2 to reformulate this problem. We consider the sample average (SA) to be the benchmark method. The dimension of ξ is considered to be $n = 5, 10, 15, 20$. We also consider support information $\|\xi\| \leq t$. Specifically, t is considered to be $3, 10, 20, 50, 100$.

Based on the numerical results provided in Table 5.2, we observe that:

1. Objective Value: The "Special" method consistently produces the lowest objective values across all dimensions and $\|\xi\|$ values, indicating superior optimization performance. The SA method yields higher objective values, which increase significantly as $\|\xi\|$ increases. The General method is not applicable (indicated by "-") for any of the dimensions and norms considered.

2. Computation Time: The "Special" method requires significantly more computation time, especially for higher dimensions and larger $\|\xi\|$. This suggests a trade-off between optimization quality and efficiency. The

computation time for the SA method increases with larger $\|\boldsymbol{\xi}\|$ but remains lower than that of the Special method for the same $\|\boldsymbol{\xi}\|$ values.

Table 5.2: Numerial Results (with Support Information)

$\boldsymbol{\xi}$ Dimension	$\|\boldsymbol{\xi}\|$	SA		Special		General	
		Obj	Time (sec)	Obj	Time (sec)	Obj	Time (sec)
5	3	0.2339	1.93	0.1764	5.14	-	3.23
	10	2.8575	4.03	0.5204	6.93	-	2.91
	20	12.0096	4.44	1.0827	9.14	-	8.05
	50	75.7125	16.11	1.0827	25.66	-	23.91
	100	309.7470	40.80	1.0827	47.18	-	50.55
10	3	0.3189	3.32	0.2568	6.27	-	5.37
	10	4.2019	6.87	0.6553	6.33	-	5.93
	20	17.7470	7.56	1.4239	15.04	-	6.13
	50	112.0273	27.28	1.4239	35.92	-	7.56
	100	458.3884	69.01	1.4239	56.73	-	9.86
15	3	0.4709	6.17	0.3424	11.07	-	8.93
	10	6.1642	12.85	0.9136	10.48	-	9.88
	20	26.0242	14.15	1.7202	20.16	-	12.27
	50	164.2595	51.26	1.7202	40.10	-	13.16
	100	672.1144	129.78	1.7202	66.39	-	15.28
20	3	0.8428	9.61	0.7185	13.21	-	9.86
	10	10.7077	20.02	1.6128	17.38	-	11.28
	20	45.1195	22.06	2.1384	30.42	-	15.77
	50	284.6424	79.94	2.1384	51.29	-	18.94
	100	1164.6121	202.40	2.1384	69.70	-	19.23

In summary, our exact reformulation (“Special”) stands out as an effective approach both with respect to solution quality and efficiency, even when benchmarked against Cheung’s approximation, which is designed for a specific distribution. The “General” case’s inner approximation seems to be less suitable for this problem, with poorer objective value outcomes.

5.2 Calibration Problem

5.2.1 Formulation

We conduct experiments on a calibration problem using the approximation method proposed in Chapter 3 for the general case. Mathematically, the problem takes the following form (Ben-Tal and Nemirovski 2009):

$$\begin{aligned} \rho^* = \quad & \max \sum_{l=1}^{n} \rho_l \\ & \text{s.t.} \quad \inf_{\mathbb{P} \in \mathcal{D}} \mathbb{P} \left\{ -\mathbf{A}_0 \leq \sum_{l=1}^{n} \rho_l \xi_l \mathbf{A}_l \leq \mathbf{A}_0 \right\} \geq p, \end{aligned}$$

where $\boldsymbol{\xi}$ is a random vector of n-dimension; $\boldsymbol{\rho} \in \mathbb{R}^n$ is a decision variable and $\mathbf{A}_l \in \mathbb{S}^{d \times d}, l = 0, \ldots, n$ are given parameters such that for a given $\vartheta^* > 0$,

$$\begin{aligned} & \text{Arrow}\left(\vartheta^* \mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_n\right) \\ := & \begin{bmatrix} \vartheta^* \mathbf{A}_0 & \mathbf{A}_1 & \ldots & \mathbf{A}_n \\ \mathbf{A}_1 & \vartheta^* \mathbf{A}_0 & & \\ \vdots & & \ddots & \\ \mathbf{A}_n & & & \vartheta^* \mathbf{A}_0 \end{bmatrix} \succeq 0. \end{aligned}$$

5.2.2 Numerical Setting and Results

We assume that the probability distribution of $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ is captured by an ambiguity set $\mathcal{D}$ with mean $\boldsymbol{\mu} = \mathbf{0}$, covariance $\boldsymbol{\Sigma} = \mathbf{I}_n$ and support information $\|\boldsymbol{\xi}\| \leq t$. Specifically, t is considered to be $3, 5, 10, 15, 20, \infty$. Here, we set the dimension of $\boldsymbol{\xi}$ to be 20, the partition number of the support set to be 20 and the satisfactory parameter to be $1 - \epsilon = 95\%$.

The experimental results presented in Table 5.3 offer insightful analysis into the performance of different approximation approaches for the specific calibration problem. The the sample average (SA) approach is one of our benchmark. This method provides an inner approximation, where $\boldsymbol{\xi}$ is generated based on the information from the ambiguity set $\mathcal{D}$ and incorporated into constraints. The "Ellipsoid" represents the results of the inner approximation method for

Table 5.3: Numerial Results of Calibration Problem

t	Ellipsoid		Polyhedral		SA	
	Obj	Time (min)	Obj	Time (min)	Obj	Time (sec)
3	3.3045	10.4	1.7256	54.8	9.4724	33.4
5	2.1169	9.6	1.3269	50.1	4.6356	56.6
10	1.5531	10.9	0.6302	52.4	2.9732	162.3
15	0.9709	11.2	0.3025	59.3	2.7984	301.9
20	0.9682	12.5	0.2977	60.6	2.7704	492.6
∞	0.2342	13.9	0.0216	72.9	\	\

ellipsoidal support (proposed in Section 3.3). The "Polyhedral" represents the experiment for the polyhedral approximation method (proposed in Section 3.3.2).

Based on Table 5.3, the experimental results for the calibration problem indicate varying performance among the three inner approximation methods. In terms of objective values, all of the three methods show a decreasing trend as t increases. The "SA" method serves as a useful benchmark for inner approximation. The "Ellipsoid" method does not perform as well as "SA" but is better than Polyhedral. Regarding computational time, the "Ellipsoid" method is the most efficient, maintaining low times ranging from 9.6 to 12.5 minutes. Therefore, for scenarios where computational time is crucial, the "Ellipsoid" method provides a faster alternative with moderate accuracy. The "SA" method, serving as an effective benchmark, offers balanced performance but becomes less practical for larger t due to its increasing computational time.

Chapter 6

Conclusion

This thesis presents the approximation and reformulation methods for DRCCLMI that can be used to overcome the limitations of the multi-dimension of integration and non-convex problem space. It mainly focuses on expressing chance constraints in terms of convex optimization problems which is important in a variety of engineering and optimization applications.

The essence of the proposed approach is deriving an inner approximation formulation for general DRCCLMI using the CVaR method. This method is especially suitable for addressing the uncertainty and risk related to the distributional robustness, thus greatly increasing the practicality and stability of the solution. However, with the help of the CVaR method it is much easier to tackle the problem because it is translated into a form that does not have the non-convex feasible set.

Furthermore, this thesis focuses on analyzing a particular type of DRCCLMI characterized by a block matrix in its linear matrix inequalities. In such cases, we obtain an inner approximation and exact reformulations of the problems to get closer solutions and be more focused on solving those particular problems. This part of our study is particularly relevant as it leads to efficient solutions to real problems having a block matrix structure.

Overall, our methods together help in converting the original DRCCLMI into a semidefinite programming (SDP) problem. This conversion is critical since it maps a hard problem to one that can be solved using current optimization methodologies and tools. The SDP framework can provide a clear approach to the computation process; It also helps improve the accuracy and efficiency of the solutions.

We apply our methods to two real-world applications: the truss topology design problem and the calibration problem. These applications are selected to cover the diversity and stability of applying the proposed methods to different real-world problems. Our methods help to enhance further the TTD problem, the key concern in the field of structural engineering, to find one of the best solutions. In the same manner, in the calibration problem, which is crucial in numerous analytical and quantitative fields, our methods prove efficient in obtaining accurate and stable solutions.

Chapter 7

References

Ben-Tal, Aharon, Alexander Goryashko, Elana Guslitzer, and Arkadi Nemirovski (2004). "Adjustable robust solutions of uncertain linear programs". In: Mathematical Programming 99.2, pp. 351–376.

Ben-Tal, Aharon, Tal Laurent, El Ghaoui, and Arkadi Nemirovski (Sept. 1998). "Robust semidefinite programming". In.

Ben-Tal, Aharon and Arkadi Nemirovski (2009). "On safe tractable approximations of chance-constrained linear matrix inequalities". In: Mathematics of Operations Research 34.1, pp. 1–25.

Boyd, Stephen, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan (1994). Linear matrix inequalities in system and control theory. SIAM.

Boyd, Stephen P and Lieven Vandenberghe (2004). Convex optimization. Cambridge university press.

Calafiore, Giuseppe Carlo and Marco C Campi (2006). "The scenario approach to robust control design". In: IEEE Transactions on Automatic Control 51.5, pp. 742–753.

Charnes, Abraham and William W Cooper (1959). "Chance-constrained programming". In: Management Science 6.1, pp. 73–79.

Cheung, Sin-Shuen, Anthony Man-Cho So, and Kuncheng Wang (2012). "Linear matrix inequalities with stochastically dependent perturbations and applications to chance-constrained semidefinite optimization". In: SIAM Journal on Optimization 22.4, pp. 1394–1430.

Delage, Erick and Yinyu Ye (2010). "Distributionally robust optimization under moment uncertainty with application to data-driven problems". In: Operations Research 58.3, pp. 595–612.

El Ghaoui, Laurent and Silviu-lulian Niculescu (2000). Advances in linear matrix inequality methods in control. SIAM.

Fan, Xiangyi and Grani A Hanasusanto (2024). "A decision rule approach for two-stage data-driven distributionally robust optimization problems with random recourse". In: INFORMS Journal on Computing 36.2, pp. 526–542.

Jiang, Nan and Weijun Xie (2023). "ALSO-X#: Better convex approximations for distributionally robust chance constrained programs". In: arXiv preprint arXiv:2302.01737.

Jiang, Ruiwei and Yongpei Guan (2016). "Data-driven chance constrained stochastic program". In: Mathematical Programming 158.1-2, pp. 291–327.

Karimi, Roya, Jianqiang Cheng, and Miguel A Lejeune (2021). "A framework for solving chance-constrained linear matrix inequality programs". In: INFORMS Journal on Computing 33.3, pp. 1015–1036.

Masubuchi, Izumi, Atsumi Ohara, and Nobuhide Suda (1998). "LMI-based controller synthesis: a unified formulation and solution". In: International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal 8.8, pp. 669–686.

Miller, Bruce L and Harvey M Wagner (1965). "Chance constrained programming with joint constraints". In: Operations Research 13.6, pp. 930–945.

Mohajerin Esfahani, Peyman and Daniel Kuhn (2018). "Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations". In: Mathematical Programming 171.1-2, pp. 115–166.

Nemirovski, Arkadi and Alexander Shapiro (2007). "Convex approximations of chance constrained programs". In: SIAM Journal on Optimization 17.4, pp. 969–996.

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

Pagnoncelli, Bernardo K, Shabbir Ahmed, and Alexander Shapiro (2009). "Sample average approximation method for chance constrained programming: theory and applications". In: Journal of Optimization Theory and Applications 142.2, pp. 399–416.

Rockafellar, R Tyrrell and Stanislav Uryasev (2000). "Optimization of conditional value-at-risk". In: Journal of Risk 2, pp. 21–42.

Scherer, Carsten and Siep Weiland (2000). "Linear matrix inequalities in control". In: Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands 3.2.

Shapiro, Alexander, Darinka Dentcheva, and Andrzej Ruszczynski (2021). Lectures on Stochastic Programming: Modeling and Theory. SIAM.

Shapiro, Alexander and Anton Kleywegt (2002). "Minimax analysis of stochastic problems". In: Optimization Methods and Software 17.3, pp. 523–542.

Wang, Zhong, Ming He, Tang Zheng, Zhiliang Fan, and Guangbin Liu (2017). "Guaranteed cost consensus for high-dimensional multi-agent systems with time-varying delays". In: IEEE/CAA Journal of Automatica Sinica 5.1, pp. 181–189.

Yang, Yongjian, Yunqiang Yin, Dujuan Wang, Joshua Ignatius, TCE Cheng, and Lalitha Dhamotharan (2023). "Distributionally robust multi-period location-allocation with multiple resources and capacity levels in humanitarian logistics". In: European Journal of Operational Research 305.3, pp. 1042–1062.

Zhao, Pengfei, Chenghong Gu, Da Huo, Yichen Shen, and Ignacio Hernando-Gil (2019). "Two-stage distributionally robust optimization for energy hub systems". In: IEEE Transactions on Industrial Informatics 16.5, pp. 3460–3469.

Zhen, Jianzhe, Frans JCT de Ruiter, Ernst Roos, and Dick den Hertog (2022). "Robust optimization for models with uncertain second-order cone and semidefinite programming constraints". In: INFORMS Journal on Computing 34.1, pp. 196–210.

Zhong, Maiying, Steven X Ding, James Lam, and Haibo Wang (2003). "An LMI approach to design robust fault detection filter for uncertain LTI systems". In: Automatica 39.3, pp. 543–550.

Zhong, Yuanguang, Ju Liu, Yong-Wu Zhou, Bin Cao, and Xueliang Zheng (2023). "The Role of Ambiguity Aversion in Contract-farming Supply Chains: A Distributionally Robust Approach". In: Omega 117, p. 102827.

Zymler, Steve, Daniel Kuhn, and Berç Rustem (2013). "Distributionally robust joint chance constraints with second-order moment information". In: Mathematical Programming 137, pp. 167–198.

Chapter 8

Appendices

8.1 Appendix I: Notation

Non-bold symbols	scalar values, e.g., s and π
Bold symbols	vectors, e.g., $\boldsymbol{x} = (x_1, \ldots, x_m)^\top$ and $\boldsymbol{q}$
Bold capital symbols	matrices, e.g., $\mathbf{A}$ and $\boldsymbol{\Sigma}$
$\mathbb{E}_{\mathbb{P}}[\cdot]$	expectation over distribution $\mathbb{P}$
$\|\cdot\|$	Euclidean norm
$\mathbb{1}(\cdot)$	indicator function
$(t)^+$	$\max\{t, 0\}$
$\mathbf{A} \bullet \mathbf{B} = \sum_{i,j} A_{ij} B_{ij}$	inner product of two conformal matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A} \succeq 0$	matrix $\mathbf{A}$ is positive semi-definite
$\mathbf{I}_m$	identity matrix of size m
$\mathbf{0}_m$	a zero vector of size m
$\mathbf{0}_{r \times c}$	a zero matrix of size $r \times c$
$\mathbb{S}_{d \times d}$	symmetric matrix of size $d \times d$

Table 8.1: Notation

8.2 Appendix II: Reformulations for Worst-Case Expectation and Probability Problems

LEMMA 1. Let $f : \mathbb{R}^m \to \mathbb{R}$ be a measurable function, and define the worst-case expectation θ_{wc} as

$$\theta_{wc} = \sup_{\mathbb{P} \in \mathcal{D}} \mathbb{E}_{\mathbb{P}} \left((f(\boldsymbol{\xi}))^+ \right),$$

where $\mathcal{D}$ is defined as (2.3). Then,

$$\begin{aligned} \theta_{wc} = \inf_{\mathbf{s},\mathbf{q},\mathbf{Q}} \quad & \mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle \\ \text{s.t.} \quad & \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq f(\boldsymbol{\xi}), \ \forall \boldsymbol{\xi} \in \mathcal{S}, \\ & \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{S}, \\ & \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m \times m}, \mathbf{Q} \succeq 0. \end{aligned}$$

Proof. This proof adheres to the same proof logic as Lemma A.1 in Zymler et al. (2013). First of all, the worst-case expectation $\sup_{\mathbb{P} \in \mathcal{D}} \mathbb{E}_{\mathbb{P}}((f(\boldsymbol{\xi}))^+)$ can equivalently be expressed as:

$$\sup_{\mathbb{P} \in \mathcal{D}} \quad \int_{\mathcal{S}} \max\{0, f(\boldsymbol{\xi})\} \, \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) \tag{8.1a}$$

$$\text{s.t.} \quad \int_{\mathcal{S}} \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) = 1, \tag{8.1b}$$

$$\int_{\mathcal{S}} \boldsymbol{\xi} \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) = \boldsymbol{\mu}, \tag{8.1c}$$

$$\int_{\mathcal{S}} (\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})^\top \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) \preceq \boldsymbol{\Sigma}. \tag{8.1d}$$

We now assign dual variables $\mathbf{s} \in \mathbb{R}$, $\mathbf{q} \in \mathbb{R}^m$, $\mathbf{Q} \in \mathbb{R}^{m \times m}$ and $\mathbf{Q} \succeq 0$ to constraints (8.1b), (8.1c), and (8.1d), respectively, and introduce the following dual problem (see, e.g., Delage and Ye 2010):

$$\inf_{\mathbf{s},\mathbf{q},\mathbf{Q}} \quad \mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle \tag{8.2a}$$

$$\text{s.t.} \quad \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq \max\{0, f(\boldsymbol{\xi})\}, \ \forall \boldsymbol{\xi} \in \mathcal{S}, \tag{8.2b}$$

$$\mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m \times m}, \mathbf{Q} \succeq 0. \tag{8.2c}$$

Furthermore, constraint (8.2b) can be expanded in terms of two constraints (8.3a) and (8.3b):

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq f(\boldsymbol{\xi}), \ \forall \boldsymbol{\xi} \in \mathcal{S}, \tag{8.3a}$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathcal{S}. \tag{8.3b}$$

This completes the proof. □

COROLLARY 1. If the support $\mathcal{S}$ defined in the ambiguity set $\mathcal{D}$ in (2.3) is unconstrained, then the worst-case expectation θ_{wc} in Lemma 1 further reduces to

$$\begin{aligned} \theta_{wc} = \inf_{\mathbf{s},\mathbf{q},\mathbf{Q}} \quad & \mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle \\ \text{s.t.} \quad & \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq f(\boldsymbol{\xi}), \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \\ & \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \\ & \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m \times m}, \mathbf{Q} \succeq 0. \end{aligned}$$

LEMMA 2. Assume that no support information $\mathcal{S}$ is defined in the ambiguity set $\mathcal{D}$. Let $\boldsymbol{L} : \mathbb{R}^m \to \mathbb{R}$ be a continuous loss function that is either concave or quadratic in $\boldsymbol{\xi}$, the following worst-case probability constraint

$$\sup_{\mathbb{P} \in \mathcal{D}} \mathbb{P}\left(\boldsymbol{L}(\boldsymbol{\xi}) > \gamma\right) \leq \epsilon,$$

can be equivalently expressed as:

$$\begin{aligned} & \frac{1}{\epsilon}\left(\mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle\right) \leq \boldsymbol{\tau}, \\ & \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \\ & -\boldsymbol{\tau} + \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} + \gamma - \boldsymbol{L}(\boldsymbol{\xi}) \geq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \\ & \boldsymbol{\tau} \in \mathbb{R}, \boldsymbol{\tau} \geq 0, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m \times m}, \mathbf{Q} \succeq 0. \end{aligned}$$

Proof. The worst-case probability problem $\sup_{\mathbb{P} \in \mathcal{D}} \mathbb{P}(\boldsymbol{L}(\boldsymbol{\xi}) > \gamma)$ can be equivalently expressed as:

$$\max_{\mathbb{P} \in \mathcal{D}} \quad \int \mathbb{1}\left\{\boldsymbol{L}(\boldsymbol{\xi}) > \gamma\right\} \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) \tag{8.4a}$$

$$\text{s.t.} \quad \int \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) = 1, \tag{8.4b}$$

$$\int \boldsymbol{\xi} \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) = \boldsymbol{\mu}, \tag{8.4c}$$

$$\int (\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})^\top \mathbf{d}\mathbb{P}(\boldsymbol{\xi}) \preceq \boldsymbol{\Sigma}, \tag{8.4d}$$

where $\mathbb{1}$ is the indicator function that gives one if the statement is verified and zero otherwise. Considering $\mathbf{s}' \in \mathbb{R}$, $\mathbf{q}' \in \mathbb{R}^m$, $\mathbf{Q}' \in \mathbb{R}^{m\times m}$ and $\mathbf{Q}' \succeq 0$ as the Lagrangian multipliers of constraints (8.4b), (8.4c) and (8.4d). We formulate its Lagrangian dual problem:

$$\min_{\mathbf{s}',\mathbf{q}',\mathbf{Q}'} \quad \mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \tag{8.5a}$$

$$\text{s.t.} \quad -\mathbb{1}\left\{ \boldsymbol{L}(\boldsymbol{\xi}) > \gamma \right\} + \mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} \geq 0,\ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \tag{8.5b}$$

$$\mathbf{s}' \in \mathbb{R}, \mathbf{q}' \in \mathbb{R}^m, \mathbf{Q}' \in \mathbb{R}^{m\times m}, \mathbf{Q}' \succeq 0. \tag{8.5c}$$

Thus, the worst-case probability constraint $\sup\limits_{\mathbb{P}\in\mathcal{D}} \mathbb{P}(\boldsymbol{L}(\boldsymbol{\xi}) > \gamma) \leq \epsilon$ can be equivalently represented as the following constraints:

$$\mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \leq \epsilon, \tag{8.6a}$$

$$\mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} \geq 0,\ \forall \boldsymbol{\xi} \in \mathbb{R}^m, \tag{8.6b}$$

$$-1 + \mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} \geq 0,\ \forall \boldsymbol{\xi} \in \mathbb{R}^m : \boldsymbol{L}(\boldsymbol{\xi}) > \gamma, \tag{8.6c}$$

$$\mathbf{s}' \in \mathbb{R}, \mathbf{q}' \in \mathbb{R}^m, \mathbf{Q}' \in \mathbb{R}^{m\times m}, \mathbf{Q}' \succeq 0. \tag{8.6d}$$

Studying more closely the third constraint (8.6c), we can reformulate it as:

$$\min_{\boldsymbol{\xi}\in\mathbb{R}^m} \sup_{\boldsymbol{\lambda}\geq 0} \ -1 + \mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} + \boldsymbol{\lambda}\left(\gamma - \boldsymbol{L}(\boldsymbol{\xi})\right) \geq 0.$$

By Sion's minimax theorem, this is further equivalent to

$$\sup_{\boldsymbol{\lambda}\geq 0} \min_{\boldsymbol{\xi}\in\mathbb{R}^m} \ -1 + \mathbf{s}' + 2\boldsymbol{\xi}^\top \mathbf{q}' + \boldsymbol{\xi}^\top \mathbf{Q}' \boldsymbol{\xi} + \boldsymbol{\lambda}\left(\gamma - \boldsymbol{L}(\boldsymbol{\xi})\right) \geq 0.$$

Here, if we consider $\boldsymbol{\lambda} = 0$, then we can derive that $\mathbf{s}' + 2\boldsymbol{\mu}^\top \mathbf{q}' + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q}' \rangle \geq 1$ for any feasible $\boldsymbol{\xi}$. However, since $\epsilon \leq 1$, this is in conflict with constraint (8.6a). Thus, $\boldsymbol{\lambda}$ should be greater than zero. By a simple replacement of variables $\boldsymbol{\tau} := (1/\boldsymbol{\lambda})$, $\mathbf{s} := (1/\boldsymbol{\lambda})\,\mathbf{s}'$, $\mathbf{q} := (1/\boldsymbol{\lambda})\,\mathbf{q}'$, $\mathbf{Q} := (1/\boldsymbol{\lambda})\,\mathbf{Q}'$, constraints (8.6a), (8.6b), (8.6c) and (8.6d) can therefore be restated as

$$\mathbf{s} + 2\boldsymbol{\mu}^\top \mathbf{q} + \langle \boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}^\top, \mathbf{Q} \rangle \leq \boldsymbol{\tau}\epsilon,$$

$$\mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q} \boldsymbol{\xi} \geq 0,\ \forall \boldsymbol{\xi} \in \mathbb{R}^m,$$

$$-\boldsymbol{\tau} + \mathbf{s} + 2\boldsymbol{\xi}^\top \mathbf{q} + \boldsymbol{\xi}^\top \mathbf{Q}\boldsymbol{\xi} + \gamma - \boldsymbol{L}(\boldsymbol{\xi}) \geq 0, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m,$$

$$\boldsymbol{\tau} \in \mathbb{R}, \boldsymbol{\tau} \geq 0, \mathbf{s} \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m, \mathbf{Q} \in \mathbb{R}^{m \times m}, \mathbf{Q} \succeq 0.$$

This completes the proof. □