Distributionally Robust Optimization and its Tractable Approximations

Joel Goh * Melvyn Sim †

submitted April 2009, revised July 2009

Abstract

In this paper, we focus on a linear optimization problem with uncertainties, having expectations in the objective and in the set of constraints. We present a modular framework to obtain an approximate solution to the problem that is distributionally robust, and more flexible than the standard technique of using linear rules. Our framework begins by firstly affinely-extending the set of primitive uncertainties to generate new linear decision rules of larger dimensions, and are therefore more flexible. Next, we develop new piecewise-linear decision rules which allow a more flexible re-formulation of the original problem. The reformulated problem will generally contain terms with expectations on the positive parts of the recourse variables. Finally, we convert the uncertain linear program into a deterministic convex program by constructing distributionally robust bounds on these expectations. These bounds are constructed by first using different pieces of information on the distribution of the underlying uncertainties to develop separate bounds, and next integrating them into a combined bound that is better than each of the individual bounds.

*NUS Business School, National University of Singapore. Email: joelgoh@nus.edu.sg

†NUS Business School and NUS Risk Management Institute, National University of Singapore. Email: dscsimm@nus.edu.sg. The research of the author is supported by Singapore-MIT Alliance and NUS academic research grant R-314-000-068-122.

1 Introduction

Traditionally, robust optimization has been used to immunize deterministic optimization problems against infeasibility caused by perturbations in model parameters, while simultaneously preserving computational tractability. The general approach involves reformulation of the original uncertain optimization problem into a deterministic convex program, such that each feasible solution of the new program is feasible for all allowable realizations of the model uncertainties. The deterministic program is therefore "robust" against perturbations in the model parameters. This approach dates back to Soyster [36], who considered a deterministic linear optimization model that is feasible for all data lying in a convex set. Recent works using this general approach include Ben-Tal and Nemirovski [3, 4, 5], Bertsimas and Sim [8], El Ghaoui and Lebret [20], and El Ghaoui et al. [22].

Ben-Tal et al. [2] noted that the traditional robust optimization approach was limited in the sense that it only allowed for all decisions to be made before the realization of the underlying uncertainties. They noted that in the modeling of real-world problems with multiple stages, it might be permissable for a subset of the decisions to be made after the realization of all or part of the underlying uncertainties. To overcome this limitation, the authors introduced the Affinely Adjustable Robust Counterpart (AARC), which allowed for delayed decisions that are affinely dependent upon the primitive uncertainties. Chen and Zhang [15] also introduced the Extended Affinely Adjustable Robust Counterpart (EAARC) as an extension of the AARC by an affine re-parameterization the primitive uncertainties. In a related work, Chen et al. [14] introduced several piecewise-linear decision rules which are more flexible than (and improve upon) regular LDRs, and they show that under their new rules, computational tractability is preserved.

Typically, robust optimization problems do not require specifications of the exact distribution of the model uncertainties. This is the general distinction between the approaches of robust optimization and stochastic programming towards modeling problems with uncertainties. In the latter, uncertainties are typically modeled as random variables with known distributions, and has been used to obtain analytic solutions to important classes of problems (see, e.g. Birge and Louveaux [10], Ruszczynski and Shapiro [31]). In the framework of robust optimization, however, uncertainties are usually modeled as random variables with true distributions that are unknown to the modeler, but are constrained to lie within a known support. Each approach has its advantages: if the exact distribution of uncertainties is precisely known, optimal solutions to the robust problem would be overly and unnecessarily conservative. Conversely, if the assumed distribution of uncertainties is in fact different from the actual distribution, the optimal solution using a stochastic programming approach may perform poorly. Bertsimas and Thiele [9] reported computational results for an inventory model showing that even in the case when the assumed demand distribution had identical first and second moments to the actual demand distribution, an inventory policy which is heavily-tuned to the assumed distribution might perform poorly when used against the true distribution.

A body which aims to bridge the gap between the conservatism of robust optimization and the specificity of stochastic programming is the minimax stochastic programming approach, where optimal decisions are sought for the *worst-case* probability distributions within a family of possible distributions, defined by certain properties such as their support and moments. This approach was pioneered by by Záčková [38] and studied in many other works (e.g. Dupačová [18], Breton and El Hachem [11], Shapiro and Kleywegt [34]). This approach has seen numerous applications, dating back to Scarf's [32] study of an optimal single-product newsvendor problem under an unknown distribution with known mean and variance, as well as the subsequent simplification of Scarf's proof by Gallego and Moon [24] and their extensions to include recourse and fixed cost. El Ghaoui et al. [21] developed worst-case Value-at-Risk bounds for a robust portfolio selection problem, when only the bounds on the means and covariance matrix of the assets are known. Chen et al. [13] introduced directional deviations as an additional means to characterize a family of distributions and were applied by Chen and Sim [12] to a goal-driven optimization problem. In a recent work, Delage and Ye [17] study distributionally robust stochastic programs when the mean and covariance of the primitive uncertainties are themselves subject to uncertainty.

Our paper aims to extend this body of work in a similar direction: for a linear optimization problem with partially-characterized uncertainties, we seek a solution that is *distributionally robust* i.e. feasible for the worst-case probability distribution within the family of distributions. The model that we study is different from most minimax stochastic programs in that we allow for expectations of recourse variables in the constraint specifications. In addition, our model allows for non-anticipativity requirements, which occurs in many practical problems (e.g. multi-stage problems). Such problems are known to be difficult to solve exactly (see Shapiro and Nemirovski [35]), but are important in practice. We approach the problem by first using a simple LDR model of recourse to tractably approximate the problem, and subsequently build more complex piecewise-linear decision rules to improve the quality of the approximation. The overarching motivation for our work is to design a general framework for modeling and solving linear robust optimization problems, which can then be automated. In an ongoing parallel work [25], we are concurrently designing software to model robust problems within this framework. The key contributions of this paper are summarized below:

1. We present a new flexible non-anticipative decision rule, which we term the bi-deflected LDR, which generalizes both the previous deflected LDR in [14] and the truncated LDR in See and Sim [33]. Furthermore, being non-anticipative, our new decision rule is also suitable for multi-stage modeling. We show that our new decision rule is an improvement over the original deflected LDR, as well as the standard LDR.

2. We discuss a technique of segregating the primitive uncertainties to obtain a new set of uncertainties. We show that by applying LDRs on the new segregated uncertainties, we obtain decision rules (which we term Segregated LDRs, or SLDRs in short) that are more flexible than the original LDRs, which preserve the non-anticipativity requirements of the original LDRs. We study how these SLDRs can used in conjunction with other partially-known characteristics of the original uncertainty distribution (such as its mean and covariance) to construct distributionally ambiguous bounds on the expected positive part of an SLDR.

This paper is structured as follows: In Section 2, we present the general optimization problem that we attempt to solve, discussing our motivation and some applications. In Section 3, we discuss the model of uncertainty which we will use for the rest of the paper, and highlight the distributional properties of the model uncertainties which we assume we have knowledge of. In Section 4, we discuss a tractable linear approximation to the general problem, and how segregated uncertainties can be used to improve the quality of the approximation. In Section 5, we present and extend existing known bounds on the expected value of the positive part of a random variable, which is used in Section 6, where we introduce the two-stage and non-anticipative bi-deflected LDR and discuss its properties. Section 7 concludes. The mathematical proofs in this paper are relegated to appendices A, B, and C.

Notations We denote a random variable by the tilde sign, i.e., $\tilde{x}$. Bold lower case letters such as $\boldsymbol{x}$ represent vectors and the upper case letters such as $\boldsymbol{A}$ denote matrices. In addition, $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$. The same notation can be used on vectors, such as $\boldsymbol{y}^+$ and $\boldsymbol{z}^-$ which denotes that the corresponding operations are performed component-wise. For any set S, we will denote by $\mathbb{1}_S$ the indicator function on the set. Also, we will denote by $[N]$ the set of positive running indices to N, i.e. $[N] = \{1, 2, \dots, N\}$, for some positive integer N. For completeness, we assume $[0] = \emptyset$. We also denote with a superscripted letter "c" the complement of a set, e.g. I^c. We denote by $\boldsymbol{e}$ the vector of all ones, and by $\boldsymbol{e}^i$ the i^{th} standard basis vector. In addition, we denote the identity matrix using the upper case boldface symbol $\boldsymbol{I}$, and for brevity, we will omit specifying its dimension where it is contextually clear. We will use superscripted indices on vectors to index members of a collection of vectors, while subscripted indexes on a vector denotes its components, i.e. $x_i = \boldsymbol{e}^{i\prime}\boldsymbol{x}$. Finally, we distinguish between models which are intractable against those that are computationally tractable, by denoting the optimal objectives of intractable functions with a superscripted asterisk, i.e. Z^*.

2 Linear Optimization with Expectation Constraints

2.1 General Model

Let $\tilde{\boldsymbol{z}} = (\tilde{z}_1, \dots \tilde{z}_N)$ be a vector of N random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $\tilde{\boldsymbol{z}}$ represents the primitive uncertainties of our model. We do not presume knowledge of the actual joint distribution of $\tilde{\boldsymbol{z}}$. Instead, we shall assume that the true joint distribution $\mathbb{P}$ lies in some family of distributions $\mathbb{F}$. We shall denote by $\boldsymbol{x} \in \Re^n$ the vector of decision variables, representing the *here-and-now* decisions which are unaffected by realizations of the primitive uncertainties. We also optimize over a set of K *wait-and-see* decision rules (also known as recourse variables), denoted by $\boldsymbol{y}^k(\cdot) \in \Re^{m_k}$, which are functions of the primitive uncertainties. In general, each decision rule may only depend on a subset of the primitive uncertainties. For each $k \in [K]$, we denote by $I_k \subseteq [N]$ the index set of dependent uncertainties for $\boldsymbol{y}^k(\cdot)$. Furthermore, for any index set $I \subseteq [N]$, we denote by $\mathcal{Y}(m, N, I)$ the space of allowable recourse decisions, which are measurable functions, defined as

$$\mathcal{Y}(m, N, I) \triangleq \left\{ \boldsymbol{f} : \Re^N \to \Re^m \ : \ \boldsymbol{f}\left(\boldsymbol{z} + \sum_{i \notin I} \lambda_i \boldsymbol{e}^i \right) = \boldsymbol{f}(\boldsymbol{z}), \forall \boldsymbol{\lambda} \in \Re^N \right\}, \tag{2.1}$$

and $\boldsymbol{y}^k \in \mathcal{Y}(m_k, N, I_k), \forall k \in [K]$. For example, if $I = \{1, 2\}$, and $\boldsymbol{y} \in \mathcal{Y}(m, N, I)$, then $\boldsymbol{y}$ only depends on the first two components of the primitive uncertainty vector $\tilde{\boldsymbol{z}}$. From a practical perspective, the specific structure of $\{I_k\}_{k=1}^K$ often translates naturally into practical modeling phenomena. For example, the condition $I_1 \subseteq I_2 \subseteq \ldots \subseteq I_K$ reflects the non-anticipativity requirement in information-dependent modeling (of which multi-stage problems are a special case) where we have successive revelation of information at each stage. We consider the ambiguity-averse minimization of a linear expected cost, with a finite set of M linear expectation constraints. The general problem can be expressed as:

$$
\begin{array}{rcll}
Z_{GEN}{}^* = & \min\limits_{\boldsymbol{x},\{\boldsymbol{y}^k(\cdot)\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{y}^k(\tilde{\boldsymbol{z}})\right) & \\
& \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{y}^k(\tilde{\boldsymbol{z}})\right) \le b_l & \forall l \in [M] \\
& & \boldsymbol{T}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\boldsymbol{y}^k(\tilde{\boldsymbol{z}}) = \boldsymbol{v}(\tilde{\boldsymbol{z}}) & \\
& & \underline{\boldsymbol{y}}^k \le \boldsymbol{y}^k(\tilde{\boldsymbol{z}}) \le \overline{\boldsymbol{y}}^k & \forall k \in [K] \\
& & \boldsymbol{x} \ge \boldsymbol{0} & \\
& & \boldsymbol{y}^k \in \mathcal{Y}(m_k, N, I_k) & \forall k \in [K],
\end{array}
\tag{2.2}
$$

where the model data $(\boldsymbol{c}^l, \boldsymbol{d}^l, \boldsymbol{T}(\tilde{\boldsymbol{z}}), \boldsymbol{v}(\tilde{\boldsymbol{z}}), \boldsymbol{U}^k, \underline{\boldsymbol{y}}^k, \overline{\boldsymbol{y}}^k, I_k)$ is deterministic, and we assume that $\boldsymbol{T}(\tilde{\boldsymbol{z}}), \boldsymbol{v}(\tilde{\boldsymbol{z}})$ are affinely dependent on $\tilde{\boldsymbol{z}}$, given by $\boldsymbol{T}(\tilde{\boldsymbol{z}}) = \boldsymbol{T}^0 + \sum_{j=1}^N \tilde{z}_j \boldsymbol{T}^j$ and $\boldsymbol{v}(\tilde{\boldsymbol{z}}) = \boldsymbol{v}^0 + \sum_{j=1}^N \tilde{z}_j \boldsymbol{v}^j$. In our model, the matrices $\boldsymbol{U}^k$ are unaffected by the uncertainties, corresponding to the case of *fixed recourse* in the stochastic programming literature. Uncertainty in the values of $\boldsymbol{c}^l$ or b_l can be handled by reformulating the problem and adding slack variables, however, our model does not handle the cases when $\boldsymbol{U}^k, \boldsymbol{d}^{l,k}$ are uncertain. In such cases, Ben-Tal et al. [2, Section 4] showed that even using LDRs for recourse decisions can result in intractability.

The bounds on the recourse variables, $\underline{\boldsymbol{y}}^k$ and $\overline{\boldsymbol{y}}^k$, for each k, are specified constants which can be infinite. Explicitly, $\underline{\boldsymbol{y}}^k \in (\Re \cup \{-\infty\})^{m_k}$ and $\overline{\boldsymbol{y}}^k \in (\Re \cup \{+\infty\})^{m_k}$. For ease of exposition later, we will denote the index sets for the non-infinite bounds as follows:

$$
\begin{array}{rcl}
\underline{J}^k & = & \left\{ i \in [m_k] : \underline{y}_i^k > -\infty \right\}, \\
\overline{J}^k & = & \left\{ i \in [m_k] : \overline{y}_i^k < +\infty \right\}.
\end{array}
\tag{2.3}
$$

For brevity, we adopt the convention here and throughout this paper that (in)equalities involving recourse variables hold almost surely for all probability distributions $\mathbb{P}$ in the family $\mathbb{F}$, i.e. $\boldsymbol{y}(\tilde{\boldsymbol{z}}) \le \boldsymbol{u} \Leftrightarrow \mathbb{P}\left(\boldsymbol{y}(\tilde{\boldsymbol{z}}) \le \boldsymbol{u}\right) = 1, \forall \mathbb{P} \in \mathbb{F}$.

2.2 Motivation

The general model (2.2) that we consider has a linear structure, which may appear overly restrictive. In this section, however, we will proceed to show how (2.2) can be used to model important classes of problems with piecewise-linear structures.

2.2.1 Piecewise-linear Utility Functions

In the modeling of certain problems, such as newsvendor-type models, it is common to encounter constraints of the form $\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\boldsymbol{y}(\tilde{\boldsymbol{z}})^+\right) \leq \boldsymbol{b}$, which can be modeled in the form of (2.2) using a slack decision rule $\boldsymbol{s}(\tilde{\boldsymbol{z}})$ as follows

$$\begin{array}{rcl} \sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\boldsymbol{s}(\tilde{\boldsymbol{z}})\right) & \leq & \boldsymbol{b} \\ \boldsymbol{s}(\tilde{\boldsymbol{z}}) & \geq & \boldsymbol{0} \\ \boldsymbol{s}(\tilde{\boldsymbol{z}}) & \geq & \boldsymbol{y}(\tilde{\boldsymbol{z}}) \\ \boldsymbol{s},\boldsymbol{y} & \in & \mathcal{Y}\left(m,N,I\right). \end{array} \tag{2.4}$$

2.2.2 CVaR Constraints

The Conditional Value-at-Risk (CVaR) risk metric was popularized by Rockafellar and Uryasev [30], and is the smallest law-invariant convex risk measure which is continuous from above that dominates Value-at-Risk (VaR) (Föllmer and Schied [23, Theorem 4.61]). CVaR is a coherent measure of risk (as axiomatized by Artzner et al. [1]) and is typically parameterized by a level $\beta \in (0,1)$. Furthermore, the β-CVaR can be derived as a special case of the negative optimized certainty equivalent (OCE) introduced by Ben-Tal and Teboulle [6, 7]. The worst-case β-CVaR when the actual uncertainty distribution $\mathbb{P}$ lies in a family of distributions $\mathbb{F}$ can be expressed as

$$\beta\text{-}\operatorname{CVaR}_{\mathbb{F}}(\tilde{x}) \triangleq \inf_{v\in\Re}\left\{v + \frac{1}{1-\beta}\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left((\tilde{x}-v)^+\right)\right\}. \tag{2.5}$$

The β-CVaR can be used to approximate chance-constraints, by using the relation $\beta\text{-}\operatorname{CVaR}_{\mathbb{F}}(\tilde{x}) \leq b \Rightarrow \mathbb{P}(\tilde{x} \geq b) \leq 1-\beta$, which holds for any distribution $\mathbb{P}$ in the family of distributions $\mathbb{F}$. Using a similar argument as (2.4), the constraint $\beta\text{-CVaR}_{\mathbb{F}}(y(\tilde{\boldsymbol{z}})) \leq b$, for a scalar-valued decision rule $y(\tilde{\boldsymbol{z}})$, can therefore be expressed as:

$$\begin{array}{rcl} v + \dfrac{1}{1-\beta}\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(s(\tilde{\boldsymbol{z}})\right) & \leq & b \\ s(\tilde{\boldsymbol{z}}) & \geq & 0 \\ s(\tilde{\boldsymbol{z}}) & \geq & y(\tilde{\boldsymbol{z}}) - v \\ s,y & \in & \mathcal{Y}(1,N,I). \end{array} \tag{2.6}$$

3 Model of Uncertainty, U

In the modeling of most problems, even though the problem data contains elements of uncertainty, the modeler may have access to some crude or partial information about the data. We assume that we may have knowledge of certain descriptive statistics of the primitive uncertainty vector $\tilde{\boldsymbol{z}}$, as follows:

Support. We denote by $\mathcal{W} \subseteq \Re^N$ the smallest convex set *containing* the support of $\tilde{\boldsymbol{z}}$, which can also be unbounded. For example, if the actual support of $\tilde{\boldsymbol{z}}$ is non-convex, we can take $\mathcal{W}$ as its convex hull. We further assume that $\mathcal{W}$ is a full-dimensional tractable conic representable set, which we take to mean a set that can be represented (exactly or approximately) by a polynomial number of linear and/or second order conic constraints[1].

Mean. We denote by $\hat{\boldsymbol{z}}$ the mean of $\tilde{\boldsymbol{z}}$. Instead of modeling the mean as a precisely-known quantity, we consider a generalization in which the mean $\hat{\boldsymbol{z}}$ is itself uncertain, with corresponding (possibly unbounded) support contained in a set $\hat{\mathcal{W}}$. We again assume that $\hat{\mathcal{W}}$ is a tractable conic representable set. This includes the case of a known mean, which corresponds to $\hat{\mathcal{W}}$ being a singleton set.

Covariance. We denote by $\boldsymbol{\Sigma}$ the covariance of $\tilde{\boldsymbol{z}}$. Unlike the mean, which we assume to be known to lie within a set, we assume that the covariance is precisely known[2].

Directional Deviations. While $\tilde{\boldsymbol{z}}$ may not have stochastically independent components, we may be able to find a linear transformation of $\tilde{\boldsymbol{z}}$, parameterized by a matrix $\boldsymbol{H}_\sigma \in \Re^{N_\sigma \times N}$, generally with $N_\sigma \le N$, that yields a vector $\boldsymbol{H}_\sigma \tilde{\boldsymbol{z}} = \tilde{\boldsymbol{z}}_\sigma$, which has stochastically independent components. We denote by $\hat{\boldsymbol{z}}_\sigma$ the mean of $\tilde{\boldsymbol{z}}_\sigma$, which lies in a set $\hat{\mathcal{W}}_\sigma \triangleq \left\{ \boldsymbol{H}_\sigma \hat{\boldsymbol{z}} : \hat{\boldsymbol{z}} \in \hat{\mathcal{W}} \right\}$. We denote by $\boldsymbol{\sigma_f}$ and $\boldsymbol{\sigma_b}$ the upper bounds of the forward and backward deviations of $\tilde{\boldsymbol{z}}_\sigma$, i.e. $\boldsymbol{\sigma_f} \ge \boldsymbol{\sigma_f}_{\mathbb{P}}(\tilde{\boldsymbol{z}}_\sigma)$ and $\boldsymbol{\sigma_b} \ge \boldsymbol{\sigma_b}_{\mathbb{P}}(\tilde{\boldsymbol{z}}_\sigma)$, where $\boldsymbol{\sigma}_{f,\mathbb{P}}(\cdot)$, $\boldsymbol{\sigma}_{b,\mathbb{P}}(\cdot)$ are defined component-wise by Chen, Sim, and Sun [13] as:

$$\begin{aligned} \boldsymbol{\sigma_f}_{\mathbb{P}}(\tilde{\boldsymbol{z}}_\sigma)'\boldsymbol{e}^j &= \sigma_{f\mathbb{P}}(\tilde{z}_{\sigma,j}) &\triangleq \sup_{\theta>0} \left\{ \sqrt{2 \ln \mathrm{E}_{\mathbb{P}} \left(\exp\left(\theta(\tilde{z}_{\sigma,j} - \hat{z}_{\sigma,j}) \right) \right) / \theta^2} \right\}, \\ \boldsymbol{\sigma_b}_{\mathbb{P}}(\tilde{\boldsymbol{z}}_\sigma)'\boldsymbol{e}^j &= \sigma_{b\mathbb{P}}(\tilde{z}_{\sigma,j}) &\triangleq \sup_{\theta>0} \left\{ \sqrt{2 \ln \mathrm{E}_{\mathbb{P}} \left(\exp\left(-\theta(\tilde{z}_{\sigma,j} - \hat{z}_{\sigma,j}) \right) \right) / \theta^2} \right\}, \end{aligned}$$

for $j \in [N_\sigma]$. We consider upper bounds of directional deviations in order to characterize a family of distributions. We note that numerical values of these bounds can be estimated from empirical data, and we refer interested readers to Chen et al. [13], Natarajan et al. [27], or See and Sim [33] for examples of how directional deviations can be estimated and used.

These distributional properties characterize the family of distributions $\mathbb{F}$. We will see in Section 5 how each property can be used to construct bounds to approximately solve the general problem (2.2), which will, in turn, be used in Section 6 where we introduce the deflected linear decision rules.

4 Linear Approximations of the General Model

Solving model (2.2) exactly is generally a computationally intractable endeavor. For instance, when the family $\mathbb{F}$ contains a single distribution, Dyer and Stougie [19] formally showed that a two-stage problem

[1]While the results in this paper will still hold even if the definition encompasses semidefinite cones, we focus on second-order conic programs (SOCP) because the study of semidefinite programming (SDP) is still an active area of research, while SOCPs can already be solved with high efficiency, and SOCP solvers are even commercially available.

[2]Although our analysis can be extended to the case of unknown covariance (using, e.g., techniques developed by Delage and Ye [17]), we do not consider this generalization for two reasons. Firstly, this generally increases the computational complexity of the problem from an SOCP to an SDP, which we prefer to avoid (see footnote 1). Secondly, there are important applications which motivate this assumption. For example, in the study of portfolio management, Chopra and Ziemba [16] showed empirically that the impact of estimation errors in mean asset returns is about an order of magnitude more severe than the corresponding impact of estimation error in asset variances and covariances.

is #P-hard to solve. In the robust case, i.e. when $\mathbb{F}$ is solely defined by a support set, a two-stage problem can be NP-hard (see the Adjustable Robust Counterpart of Ben Tal et al. [2]). However, by applying a suitable restriction to the space of allowable decision rules, we can obtain a tractable approximation to the problem. Instead of considering all possible choices of $\boldsymbol{y}^k(\cdot)$ from $\mathcal{Y}(m,N,I)$, we restrict ourselves to Linear Decision Rules (LDRs), where each $\boldsymbol{y}^k(\cdot)$ is instead chosen from the space of *affine* functions of $\tilde{\boldsymbol{z}}$, denoted by $\mathcal{L}(m,N,I) \subset \mathcal{Y}(m,N,I)$, and defined as follows:

$$\mathcal{L}(m,N,I) = \left\{ \boldsymbol{f} : \Re^N \to \Re^m : \exists (\boldsymbol{y}^0, \boldsymbol{Y}) \in \Re^m \times \Re^{m \times N} : \begin{array}{c} \boldsymbol{f}(\boldsymbol{z}) = \boldsymbol{y}^0 + \boldsymbol{Y}\boldsymbol{z} \\ \boldsymbol{Y}\boldsymbol{e}^i = \boldsymbol{0}, \forall i \notin I \end{array} \right\}. \tag{4.1}$$

We notice that the final condition, $\boldsymbol{Y}\boldsymbol{e}^i = \boldsymbol{0}$ enforces the information dependency upon the index set I. Therefore, using a linear model of recourse, the recourse decision can be explicitly written as $\boldsymbol{y}^k(\tilde{\boldsymbol{z}}) = \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k \tilde{\boldsymbol{z}}$, for each $k \in [K]$.

Using LDRs as our model of recourse, and denoting by $\mathbb{F}$ the family of distributions $\mathbb{P}$ with distributional properties as specified in the Model of Uncertainty, Problem (2.2) is approximated as:

$$\begin{array}{llll} Z_{LDR} = & \min\limits_{\boldsymbol{x}, \left\{\boldsymbol{y}^{0,k}, \boldsymbol{Y}^k\right\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{y}^{0,k} + \sup\limits_{\hat{\boldsymbol{z}} \in \hat{\mathcal{W}}} \left(\sum\limits_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{Y}^k \hat{\boldsymbol{z}} \right) & \\ & \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{y}^{0,k} + \sup\limits_{\hat{\boldsymbol{z}} \in \hat{\mathcal{W}}} \left(\sum\limits_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{Y}^k \hat{\boldsymbol{z}} \right) \le b_l & \forall l \in [M] \\ & & \boldsymbol{T}^0\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{U}^k \boldsymbol{y}^{0,k} = \boldsymbol{v}^0 & \\ & & \boldsymbol{T}^j\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{U}^k \boldsymbol{Y}^k \boldsymbol{e}^j = \boldsymbol{v}^j & \forall j \in [N] \\ & & \underline{\boldsymbol{y}}^k \le \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k \boldsymbol{z} \le \overline{\boldsymbol{y}}^k \qquad \forall \boldsymbol{z} \in \mathcal{W} & \forall k \in [K] \\ & & \boldsymbol{Y}^k \boldsymbol{e}^j = \boldsymbol{0} & \forall j \notin I_k, \forall k \in [K] \\ & & \boldsymbol{x} \ge \boldsymbol{0}. & \end{array} \tag{4.2}$$

We formalize this in the following proposition:

Proposition 1 *If $\mathcal{L}(m_k, N, I_k)$ is used to approximate $\mathcal{Y}(m_k, N, I_k)$, then under the approximation, Problem (2.2) is equivalent to Problem (4.2).*

Proof : Please see Appendix A.1.

In the transformed problem, we notice that the constraints $\underline{\boldsymbol{y}}^k \le \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k \boldsymbol{z} \le \overline{\boldsymbol{y}}^k$ over all $\boldsymbol{z} \in \mathcal{W}$, as well as the suprema over $\hat{\boldsymbol{z}} \in \hat{\mathcal{W}}$ in the objective and first M constraints, can be converted into their robust counterparts. This will render Problem (4.2) as a tractable conic optimization problem, since $\mathcal{W}$ and $\hat{\mathcal{W}}$ are tractable conic representable sets (see Ben-Tal and Nemirovski [3], Bertsimas and Sim [8]). In particular, if $\mathcal{W}$ and $\hat{\mathcal{W}}$ are polyhedral, then Problem (4.2) becomes a linear program.

While the LDR approximation is computationally tractable, the quality of the approximation can be very poor, even being infeasible for very simple constraints (see e.g., the discussion by Chen et al. [14]). In the remainder of this section, we discuss how we may segregate the primitive uncertainty vector to obtain a more flexible model of recourse. In Section 6, we further discuss how to exploit the structure of the constraints to construct even more flexible piecewise-linear decision rules.

4.1 Remapping the Primitive Uncertainty Vector

In the positive-and-negative Segregated LDR introduced by Chen et al. [14], the authors split the original uncertainty vector into positive and negative half spaces, and applied LDRs to the split uncertainties in order to increase the flexibility of the LDR. We extend their result, and discuss here how we may segregate primitive uncertainties into intervals, and more importantly, how we can use the segregated uncertainties in our modeling framework. To begin, we consider a functional mapping $\boldsymbol{M} : \Re^N \rightarrow \Re^{N_E}$, where $N_E \geq N$, that satisfies the following relationship for any $\boldsymbol{z} \in \Re^N$, for some given matrix $\boldsymbol{F} \in \Re^{N \times N_E}$ and vector $\boldsymbol{g} \in \Re^N$, such that

$$\boldsymbol{z} = \boldsymbol{F}\boldsymbol{M}(\boldsymbol{z}) + \boldsymbol{g}. \tag{4.3}$$

We notice that $\boldsymbol{F}$ has to be full rank since $\mathcal{W}$ is assumed to have a non-empty interior. We denote by $\mathcal{V}^*$ the image of $\boldsymbol{M}(\cdot)$ corresponding to the domain $\mathcal{W}$, i.e.

$$\mathcal{V}^* \triangleq \{\boldsymbol{M}(\boldsymbol{z}) : \boldsymbol{z} \in \mathcal{W}\}, \tag{4.4}$$

and similarly by $\hat{\mathcal{V}}^*$ the image of $\boldsymbol{M}(\cdot)$ corresponding to the domain $\hat{\mathcal{W}}$, i.e. $\hat{\mathcal{V}}^* \triangleq \left\{\boldsymbol{M}(\hat{\boldsymbol{z}}) : \hat{\boldsymbol{z}} \in \hat{\mathcal{W}}\right\}$.

Again, we require $\boldsymbol{M}(\cdot)$ to be such that the convex hull of $\mathcal{V}^*$ is full-dimensional. We notice that while $\boldsymbol{M}(\cdot)$ is invertible by an affine mapping, $\boldsymbol{M}(\cdot)$ itself is not required to be affine in its argument. Indeed, in the example which follows, we present a piecewise-affine $\boldsymbol{M}(\cdot)$, and show how it can be used to segregate a scalar primitive uncertainty into different regions of interest.

4.2 Example: Segregating a Scalar Uncertainty

Suppose we have a scalar primitive uncertainty, $\tilde{z}$ with support $\Re$, which wish to segregate into three regions, $(-\infty, -1]$, $[-1, 1]$, and $[1, +\infty)$. Denoting the points, $(p_1, p_2, p_3, p_4) = (-\infty, -1, 1, +\infty)$, we can construct the segregated uncertainty by applying the following nonlinear mapping: $\tilde{\boldsymbol{\zeta}} = \boldsymbol{M}(\tilde{z})$, where

$$\tilde{\zeta}_i = \begin{cases} \tilde{z} & \text{if } p_i \leq \tilde{z} \leq p_{i+1}, \\ p_i & \text{if } \tilde{z} \leq p_i, \\ p_{i+1} & \text{if } \tilde{z} \geq p_{i+1}, \end{cases}$$

for $i \in \{1, 2, 3\}$. We notice that

$$\tilde{\zeta}_1 + \tilde{\zeta}_2 + \tilde{\zeta}_3 = \begin{cases} \tilde{z} - 1 + 1 = \tilde{z} & \text{if } \tilde{z} \leq -1, \\ -1 + \tilde{z} + 1 = \tilde{z} & \text{if } -1 < \tilde{z} < 1, \\ -1 + 1 + \tilde{z} = \tilde{z} & \text{if } \tilde{z} \geq 1, \end{cases}$$

and therefore Equation (4.3) holds with $\boldsymbol{F} = [1, 1, 1]$ and $g = 0$.

4.3 Mappings which Represent Segregations

In the preceding example, we notice that the segregation resulted in a new uncertainty vector $\tilde{\boldsymbol{\zeta}}$, the components of which provide local information of the original scalar uncertainty $\tilde{z}$. In general, the purpose of segregating uncertainties into intervals is to obtain a finer resolution of the original uncertainty.

In Section 4.4, we will use segregated uncertainties to define more flexible decision rules. To better understand how we can construct such segregations, we proceed to characterize the mapping functions $\boldsymbol{M}(\cdot)$ which represents a segregation of a primitive uncertainty vector.

For some positive integer L, we consider a collection of $N \times (L+1)$ distinct points on the extended real line, denoted by ξ_{ij} for some $i \in [N], j \in [L+1]$, with the following properties $\forall i \in [N]$:

$$\begin{aligned} \xi_{i,1} &= -\infty, \\ \xi_{i,L+1} &= +\infty, \\ \xi_{i,j_1} &< \xi_{i,j_2} \quad \text{iff } j_1 < j_2. \end{aligned} \tag{4.5}$$

Furthermore, we denote by $\boldsymbol{\xi}$ the $N \times (L+1)$ matrix which collects these points. We call $\boldsymbol{M}(\cdot)$ a segregation if $\forall \boldsymbol{\zeta} = \boldsymbol{M}(\boldsymbol{z})$, its components $\forall j \in [N_E]$, are given by

$$\zeta_j = \begin{cases} z_i & \text{if } \xi_{i,k} \le z_i \le \xi_{i,k+1}, \\ \xi_{i,k} & \text{if } z_i \le \xi_{i,k}, \\ \xi_{i,k+1} & \text{if } z_i \ge \xi_{i,k+1}, \end{cases} \tag{4.6}$$

where $j = i + (k-1)N$ for some $i \in [N], k \in [L]$. We notice that for a particular $j \in [N_E]$, i and k can be uniquely obtained as:

$$\begin{aligned} i &= ((j-1) \bmod N) + 1, \\ k &= \lceil j/N \rceil . \end{aligned}$$

In the following proposition, we prove that if $\boldsymbol{M}(\cdot)$ is a segregation, it is also affinely-invertible by proper choice of $\boldsymbol{F}$ and $\boldsymbol{g}$.

Proposition 2 *If $\boldsymbol{M}(\cdot)$ is a segregation, and if we choose $\boldsymbol{F}$ and $\boldsymbol{g}$ to be*

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{I} & \ldots & \boldsymbol{I} \end{bmatrix} \quad , \quad \boldsymbol{g} = -\sum_{i=2}^{L} \boldsymbol{\xi}\boldsymbol{e}^i \ ,$$

then $\boldsymbol{z} = \boldsymbol{F}\boldsymbol{M}(\boldsymbol{z}) + \boldsymbol{g}\ \forall \boldsymbol{z} \in \Re^N$, where $\boldsymbol{F} \in \Re^{N \times LN}, \boldsymbol{g} \in \Re^N$, and $\boldsymbol{I}$ is the $N \times N$ identity matrix.

Proof : Please see Appendix A.2.

Remark : Although we consider a segregation $\boldsymbol{M}(\cdot)$ which segments each component of $\tilde{\boldsymbol{z}}$ uniformly into L parts, it is clear from the proof that Proposition 2 still holds even if each component of $\tilde{\boldsymbol{z}}$ was segmented into a different number of positive integer parts, $\{L_i\}_{i=1}^N$, albeit with a different choice of $\boldsymbol{F}$ and $\boldsymbol{g}$. This would come at the expense of more notation and bookkeeping. Hence, for simplicity, here and for the rest of this paper, we will only discuss only case of uniform segregation.

4.4 Segregated Linear Decision Rules

As seen above, the segregation $\boldsymbol{M}(\cdot)$ can be used to define an new uncertainty vector $\tilde{\boldsymbol{\zeta}} \in \Re^{N_E}$, which we will term the segregated uncertainty vector. By considering LDRs on the segregated uncertainty vector $\tilde{\boldsymbol{\zeta}}$, we obtain a new set of decision rules, which we term segregated LDRs or SLDRs for short. By using the SLDRs, the recourse decsisions become

$$\boldsymbol{y}^k(\tilde{\boldsymbol{z}}) = \boldsymbol{r}^k(\boldsymbol{M}(\tilde{\boldsymbol{z}})) = \boldsymbol{r}^{0,k} + \boldsymbol{R}^k\boldsymbol{M}(\tilde{\boldsymbol{z}}),\ \forall k \in [K]\,,$$

which is effectively the composition of an affine functional with the segregating mapping $\boldsymbol{M}(\cdot)$. Under these SLDRs, Problem (2.2) becomes:

$$\begin{array}{rcll}
Z_{SLDR}{}^* = & \displaystyle\min_{\boldsymbol{x},\left\{\boldsymbol{r}^k(\cdot)\right\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \displaystyle\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{r}^k(\boldsymbol{M}(\tilde{\boldsymbol{z}}))\right) & \\
& \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \displaystyle\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{r}^k(\boldsymbol{M}(\tilde{\boldsymbol{z}}))\right) \le b_l & \forall l \in [M] \\
& & \boldsymbol{T}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \displaystyle\sum_{k=1}^K \boldsymbol{U}^k\boldsymbol{r}^k(\boldsymbol{M}(\tilde{\boldsymbol{z}})) = \boldsymbol{v}(\tilde{\boldsymbol{z}}) & \\
& & \underline{\boldsymbol{y}}^k \le \boldsymbol{r}^k(\boldsymbol{M}(\tilde{\boldsymbol{z}})) \le \overline{\boldsymbol{y}}^k & \forall k \in [K] \\
& & \boldsymbol{x} \ge \boldsymbol{0} & \\
& & \boldsymbol{r}^k(\boldsymbol{M}(\tilde{\boldsymbol{z}})) = \boldsymbol{r}^{0,k} + \boldsymbol{R}^k\boldsymbol{M}(\tilde{\boldsymbol{z}}) & \forall k \in [K] \\
& & \boldsymbol{r}^k \circ \boldsymbol{M} \in \mathcal{Y}(m_k, N, I_k) & \forall k \in [K]\,.
\end{array} \tag{4.7}$$

Problem (A.4) of Appendix A presents an equivalent, and more explicit formulation of Problem (4.7). We notice that Problem (A.4), and, equivalently (4.7), is generally intractable, since $\mathcal{V}^*$ and $\hat{\mathcal{V}}^*$ are generally non-convex. Furthermore, it is not obvious how to handle the non-anticipativity constraints. We therefore aim to construct an approximation of Problem (4.7) that would still improve upon the standard LDR.

4.5 Approximating $\mathcal{V}^*$ and $\hat{\mathcal{V}}^*$

We begin by constructing sets $\mathcal{V} \subseteq \mathcal{V}^*$ and $\hat{\mathcal{V}} \subseteq \hat{\mathcal{V}}^*$ which approximate the sets $\mathcal{V}^*$ and $\hat{\mathcal{V}}^*$. Both $\mathcal{V}$ and $\hat{\mathcal{V}}$ should be tractable conic representable sets, and should satisfy the implicit set relations:

$$\begin{array}{rcl}
\mathcal{W} &=& \left\{\boldsymbol{F}\boldsymbol{\zeta} + \boldsymbol{g} : \boldsymbol{\zeta} \in \mathcal{V}\right\}, \\
\hat{\mathcal{W}} &=& \left\{\boldsymbol{F}\hat{\boldsymbol{\zeta}} + \boldsymbol{g} : \hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}\right\},
\end{array} \tag{4.8}$$

To begin explicit construction of $\mathcal{V}$, we note that the best convex approximant of $\mathcal{V}^*$ is its convex hull, $\mathrm{conv}(\mathcal{V}^*)$, and ideally, we would choose $\mathcal{V} = \mathrm{conv}(\mathcal{V}^*)$. Indeed, Wang et al [37] show how to describe $\mathrm{conv}(\mathcal{V}^*)$ when $\mathcal{W}$ has a special structure, which they term an *absolute set*. However, in general, for $\mathcal{W}$ of arbitrary structures, it may not be easy to describe $\mathrm{conv}(\mathcal{V}^*)$. Instead, we construct $\mathcal{V}$ using a proxy set $\mathcal{H}$ by defining

$$\mathcal{V} \triangleq \left\{\boldsymbol{\zeta} \in \mathcal{H} : \boldsymbol{F}\boldsymbol{\zeta} + \boldsymbol{g} \in \mathcal{W}\right\}, \tag{4.9}$$

where $\mathcal{H}$ is explicitly defined as:

$$\mathcal{H} \triangleq \left\{\boldsymbol{\zeta} \in \Re^{N_E} : \xi_{i,k} \le \zeta_j \le \xi_{i,k+1}\ \forall j \in [N_E] : i \in [N]\,, k \in [L] : j = i + (k-1)N\right\}. \tag{4.10}$$

We observe that from the definition of a segregation (4.6), $\mathcal{H}$ satisfies the property that $\boldsymbol{z} \in \mathcal{W} \Rightarrow \boldsymbol{M}(\boldsymbol{z}) \in \mathcal{H}$, and clearly, $\mathcal{V}$ defined in this manner will satisfy (4.8). The motivation for this seemingly extraneous construction of $\mathcal{V}$ is that $\mathcal{H}$ depends only on the segregation $\boldsymbol{M}(\cdot)$, and is decoupled from $\mathcal{W}$, making it easy to specify in practice. In particular, for the example presented in Section 4.2, $\mathcal{H}$ can be represented by $\mathcal{H} = \left\{\boldsymbol{\zeta} \in \Re^3 : \zeta_1 \leq -1, -1 \leq \zeta_2 \leq 1, \zeta_3 \geq 1\right\}$. Using a similar argument, we define $\hat{\mathcal{V}}$ by the same set $\mathcal{H}$ as $\hat{\mathcal{V}} \triangleq \left\{\hat{\boldsymbol{\zeta}} \in \mathcal{H} : \boldsymbol{F}\hat{\boldsymbol{\zeta}} + \boldsymbol{g} \in \hat{\mathcal{W}}\right\}$.

4.6 Approximating Problem (4.7)

After constructing $\mathcal{V}$ and $\hat{\mathcal{V}}$, we now discuss how we can approximate Problem (4.7). For convenience, we define the collection of index sets for each $k \in [K]$ as

$$\Phi_k = \{j \in [N_E] : \exists i \in I_k : (i-1) \equiv (j-1) \mod N\}. \tag{4.11}$$

We then define the tractable SLDR approximation to Problem (4.7) as the following:

$$\begin{array}{rlll}
Z_{SLDR} = & \min\limits_{\boldsymbol{x}, \left\{\boldsymbol{r}^{0,k}, \boldsymbol{R}^k\right\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{r}^{0,k} + \sup\limits_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} \left(\sum\limits_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{R}^k\hat{\boldsymbol{\zeta}}\right) & \\
& \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{r}^{0,k} + \sup\limits_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} \left(\sum\limits_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{R}^k\hat{\boldsymbol{\zeta}}\right) \leq b_l & \forall l \in [M] \\
& & \boldsymbol{\mathcal{T}}^0\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{U}^k\boldsymbol{r}^{0,k} = \boldsymbol{\nu}^0 & \\
& & \boldsymbol{\mathcal{T}}^j\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{U}^k\boldsymbol{R}^k\boldsymbol{e}^j = \boldsymbol{\nu}^j & \forall j \in [N_E] \\
& & \underline{\boldsymbol{y}}^k \leq \boldsymbol{r}^{0,k} + \boldsymbol{R}^k\boldsymbol{\zeta} \leq \overline{\boldsymbol{y}}^k \qquad \forall \boldsymbol{\zeta} \in \mathcal{V} & \forall k \in [K] \\
& & \boldsymbol{R}^k\boldsymbol{e}^j = \boldsymbol{0} & \forall j \notin \Phi_k, \forall k \in [K] \\
& & \boldsymbol{x} \geq \boldsymbol{0}, &
\end{array} \tag{4.12}$$

where the transformed model data is defined as:

$$\begin{array}{rclcrcll}
\boldsymbol{\nu}^0 & = & \boldsymbol{v}^0 + \sum\limits_{i=1}^N g_i\boldsymbol{v}^i & & \boldsymbol{\mathcal{T}}^0 & = & \boldsymbol{T}^0 + \sum\limits_{i=1}^N g_i\boldsymbol{T}^i & \\
\boldsymbol{\nu}^j & = & \sum\limits_{i=1}^N F_{ij}\boldsymbol{v}^i & & \boldsymbol{\mathcal{T}}^j & = & \sum\limits_{i=1}^N F_{ij}\boldsymbol{T}^i & \forall j \in [N_E].
\end{array} \tag{4.13}$$

The following proposition relates the objectives under the exact (intractable) SLDR, the approximate SLDR, and the LDR models of recourse:

Proposition 3 *The following inequality holds:* $Z_{SLDR^*} \leq Z_{SLDR} \leq Z_{LDR}$.

Proof : Please see Appendix A.3.

Remark 1: Proposition 3 shows that irrespective of how crudely $(\mathcal{V}, \hat{\mathcal{V}})$ approximates $(\mathcal{V}^*, \hat{\mathcal{V}}^*)$, using the approximate SLDR will nonetheless not be worse than using the original LDR. Furthermore, using the SLDR retains the linear structure of the problem. Specifically, if $\mathcal{V}$ and $\hat{\mathcal{V}}$ are polyhedral, the SLDR approximation (Problem (4.12)) reduces to a linear program.

Remark 2: A key difference in the SLDR which we describe here and the SLDR of Chen et. al. [14] is that they assume precise knowledge of the mean and covariance of the segregated uncertainty vector. In our tractable SLDR model, we only exploit the support information of the segregated uncertainty vector, which is captured by the set $\mathcal{H}$. Our SLDR however, does include their model as a special case, since, if we did have knowledge of the segregated moments, we can simply reformulate the problem, expressing what they term as the segregated uncertainty vector as our primitive uncertainty vector.

4.7 Interpreting the SLDR Approximation (4.12)

We notice that the structure of the SLDR approximation (4.12) above closely resembles the form of the LDR approximation (4.2). Indeed, we can interpret the SLDR as a linear approximation of the following uncertain optimization problem, defined over a different uncertainty vector $\tilde{\boldsymbol{\zeta}}$:

$$
\begin{array}{rcll}
Z^*_{GEN,2} = & \displaystyle\min_{\boldsymbol{x},\{\boldsymbol{r}^k(\cdot)\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \displaystyle\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}})\right) & \\
& \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \displaystyle\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}})\right) \leq b_l & \forall l \in [M] \\
& & \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \displaystyle\sum_{k=1}^K \boldsymbol{U}^k\boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) & \\
& & \underline{\boldsymbol{y}}^k \leq \boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}}) \leq \overline{\boldsymbol{y}}^k & \forall k \in [K] \\
& & \boldsymbol{x} \geq \boldsymbol{0} & \\
& & \boldsymbol{r}^k \in \mathcal{Y}(m_k, N_E, \Phi_k) & \forall k \in [K]\,,
\end{array}
\tag{4.14}
$$

where

$$
\boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}}) \;=\; \boldsymbol{\mathcal{T}}^0 + \sum_{j=1}^N \tilde{\zeta}_j \boldsymbol{\mathcal{T}}^j \quad , \quad \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) \;=\; \boldsymbol{\nu}^0 + \sum_{j=1}^N \tilde{\zeta}_j \boldsymbol{\nu}^j.
$$

After approximating $\mathcal{Y}(m_k, N_E, \Phi_k)$ with $\mathcal{L}(m_k, N_E, \Phi_k)$, and following the same steps as the LDR approximation (see Appendix A.1), (4.14) will reduce to (4.12). When we interpret (4.12) as an LDR approximation of Problem (4.14) above, the approximate sets $\mathcal{V}$ and $\hat{\mathcal{V}}$ should therefore be respectively interpreted as supersets of the support and mean support of the new uncertainty vector $\tilde{\boldsymbol{\zeta}}$. Furthermore, the collection of index sets $\{\Phi_k\}_{k=1}^K$, which was somewhat arbitrarily defined before, now has the natural interpretation as the information index sets of the new decision rules, $\boldsymbol{r}^k(\cdot)$.

4.8 Example: Specifying Distributional Properties for Segregated Uncertainties

We provide a concrete example of how various distributional properties can be specified for segregated uncertainties to illustrate how segregations might work in practice. Consider a primitive uncertainty vector $\tilde{\boldsymbol{z}} \in \Re^4$, where only $\tilde{z}_2$ and $\tilde{z}_3$ are stochastically independent. The distributional properties

$\mathcal{W}$, $\hat{\mathcal{W}}$, and $\boldsymbol{\Sigma}$ of the primitive uncertainty vector $\tilde{\boldsymbol{z}}$ can be specified directly. Furthermore, we wish to segregate each component of $\tilde{\boldsymbol{z}}$ into three regions, $(-\infty,-1]$, $[-1,1]$, and $[1,\infty)$ as in the earlier example of Section 4.2. This results in a segregated uncertainty vector $\tilde{\boldsymbol{\zeta}} \in \Re^{12}$, which obeys the relation $\tilde{\boldsymbol{z}} = \boldsymbol{F}\tilde{\boldsymbol{\zeta}} + \boldsymbol{g}$ for parameters $\boldsymbol{F} \in \Re^{4\times 12}$, $\boldsymbol{g} \in \Re^4$, where

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{I} & \boldsymbol{I} \end{bmatrix} \quad \text{and} \quad \boldsymbol{g} = \boldsymbol{0},$$

and $\boldsymbol{I}$ represents the 4-by-4 identity matrix. Based on the segregation, we can choose $\mathcal{H}$ as

$$\mathcal{H} = \left\{ \boldsymbol{\zeta} \in \Re^{12} : \begin{array}{rcl} \zeta_i \le -1 & \text{if} & i = 1,2,3,4 \\ -1 \le \zeta_i \le 1 & \text{if} & i = 5,6,7,8 \\ \zeta_i \ge 1 & \text{if} & i = 9,10,11,12 \end{array} \right\}.$$

Finally, we notice that expressing $\tilde{\boldsymbol{z}}_\sigma = \boldsymbol{F}_\sigma\tilde{\boldsymbol{\zeta}} + \boldsymbol{g}_\sigma$, where

$$\boldsymbol{F}_\sigma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \boldsymbol{F} \quad \text{and} \quad \boldsymbol{g}_\sigma = \boldsymbol{0},$$

we can construct an uncertainty vector $\tilde{\boldsymbol{z}}_\sigma$ with independent components. We can therefore specify the directional deviations $(\boldsymbol{\sigma_f}, \boldsymbol{\sigma_b})$ of $\tilde{\boldsymbol{z}}_\sigma$.

5 Distributionally Ambiguous Bounds for $\mathrm{E}_{\mathbb{P}}\left((\cdot)^+\right)$

When we specify partial distributional information of the model uncertainties, we effectively characterize a family of distributions $\mathbb{F}$, which contains the true uncertainties distribution $\mathbb{P}$. In this section, we discuss how we may evaluate the supremum of the expected positive part of an SLDR (recall that according to the discussion in Section 4.7, the SLDR can be interpreted as an LDR on the segregated uncertainties), i.e.

$$\begin{array}{l} \sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\left(\boldsymbol{r}(\tilde{\boldsymbol{\zeta}})\right)^+\right) \\ \boldsymbol{r} \in \mathcal{L}(m, N_E, \Phi), \end{array}$$

where the family $\mathbb{F}$ is partially characterized by the distributional properties as laid out in the Model of Uncertainty in Section 3. This is in anticipation of our discussion of deflected linear decision rules (DLDRs) in Section 6, where we will use these bounds. We will show how each pair of distributional properties:

- Mean and (segregated) Support
- Mean and Covariance
- Mean and Directional Deviation

establishes a distinct bound, and we conclude this section by showing how to combine these separate bounds, when we have access to a combination of distributional information from these three categories.

We show that each bound can be constructed by solving a deterministic optimization problem and we adopt the standard convention in convex programming that an infeasible minimization (maximization) problem has an optimal value of $+\infty$ $(-\infty)$. Also, we will only present results for $\mathrm{E}_{\mathbb{P}}\left((\cdot)^+\right)$, since the results for the negative case can be easily derived by applying the identity $x^- = (-x)^+$. For generality, we will discuss the bound of $\mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$, applied to the segregated LDR as defined in Section 4, since bounding the expectation of the positive part of a standard LDR, i.e. $\mathrm{E}_{\mathbb{P}}\left(\left(y^0 + \boldsymbol{y}'\tilde{\boldsymbol{z}}\right)^+\right)$ can be derived as a special case when $\boldsymbol{M}(\boldsymbol{z}) = \boldsymbol{z}$.

5.1 Mean and Support Information

Theorem 1 *Let $\mathbb{F}_1$ be the family of all distributions $\mathbb{P}$ such that the random variable $\tilde{\boldsymbol{\zeta}}$ has support in $\mathcal{V}$, and its mean, $\hat{\boldsymbol{\zeta}}$ has support in $\hat{\mathcal{V}}$, i.e.*

$$\mathbb{F}_1 = \left\{\mathbb{P} : \hat{\boldsymbol{\zeta}} = \mathrm{E}_{\mathbb{P}}\left(\tilde{\boldsymbol{\zeta}}\right) \in \hat{\mathcal{V}}, \mathbb{P}\left(\tilde{\boldsymbol{\zeta}} \in \mathcal{V}\right) = 1\right\}.$$

Then $\pi^1\left(r^0, \boldsymbol{r}\right)$ is a tight upper bound for $\mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ over all distributions $\mathbb{P} \in \mathbb{F}_1$, i.e.

$$\sup_{\mathbb{P}\in\mathbb{F}_1} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = \pi^1\left(r^0, \boldsymbol{r}\right),$$

where

$$\pi^1\left(r^0, \boldsymbol{r}\right) \triangleq \inf_{\boldsymbol{s}\in\Re^{N_E}} \left(\sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}}\left\{\boldsymbol{s}'\hat{\boldsymbol{\zeta}}\right\} + \sup_{\boldsymbol{\zeta}\in\mathcal{V}}\left(\max\left\{r^0 + \boldsymbol{r}'\boldsymbol{\zeta} - \boldsymbol{s}'\boldsymbol{\zeta}, -\boldsymbol{s}'\boldsymbol{\zeta}\right\}\right)\right). \tag{5.1}$$

Proof : Please see Appendix B.1.

Remark : Firstly, we notice that for given $(r^0, \boldsymbol{r})$,

$$r^0 + \boldsymbol{r}'\boldsymbol{\zeta} \geq 0,\ \forall \boldsymbol{\zeta} \in \mathcal{V} \Rightarrow \sup_{\mathbb{P}\in\mathbb{F}_1} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = r^0 + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \boldsymbol{r}'\hat{\boldsymbol{\zeta}},$$

which is attained by $\pi^1\left(r^0, \boldsymbol{r}\right)$ by choosing $\boldsymbol{s} = \boldsymbol{r}$. Also,

$$r^0 + \boldsymbol{r}'\boldsymbol{\zeta} \leq 0,\ \forall \boldsymbol{\zeta} \in \mathcal{V} \Rightarrow \sup_{\mathbb{P}\in\mathbb{F}_1} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = 0,$$

which is attained by $\pi^1\left(r^0, \boldsymbol{r}\right)$ by choosing $\boldsymbol{s} = \boldsymbol{0}$. Secondly, we note that the epigraph of $\pi^1\left(r^0, \boldsymbol{r}\right)$ can be reformulated as a set of robust constraints over the support sets $\mathcal{V}$ and $\hat{\mathcal{V}}$. Since $\mathcal{V}$ and $\hat{\mathcal{V}}$ are tractable conic representable sets by assumption, the bound $\pi^1\left(r^0, \boldsymbol{r}\right)$ is consequently computationally tractable.

5.2 Mean and Covariance Information

Theorem 2 *Let $\mathbb{F}_2$ be the family of all distributions $\mathbb{P}$ such that the mean of the segregated uncertainties, $\hat{\boldsymbol{\zeta}}$, has support in $\hat{\mathcal{V}}$, and the primitive uncertainty vector has covariance matrix $\boldsymbol{\Sigma}$, i.e.*

$$\mathbb{F}_2 = \left\{ \mathbb{P} : \hat{\boldsymbol{\zeta}} = \mathrm{E}_{\mathbb{P}}\left(\tilde{\boldsymbol{\zeta}}\right) \in \hat{\mathcal{V}}, \mathrm{E}_{\mathbb{P}}\left(\boldsymbol{F}\left(\tilde{\boldsymbol{\zeta}} - \hat{\boldsymbol{\zeta}}\right)\left(\tilde{\boldsymbol{\zeta}} - \hat{\boldsymbol{\zeta}}\right)' \boldsymbol{F}' \right) = \boldsymbol{\Sigma} \right\}.$$

Then $\pi^2\left(r^0, \boldsymbol{r}\right)$ is a tight upper bound for $\mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ over all distributions $\mathbb{P} \in \mathbb{F}_2$, i.e.

$$\sup_{\mathbb{P}\in\mathbb{F}_2} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = \pi^2\left(r^0, \boldsymbol{r}\right),$$

where

$$\pi^2\left(r^0, \boldsymbol{r}\right) \triangleq \inf_{\boldsymbol{y}\in\{\boldsymbol{y}:\boldsymbol{F}'\boldsymbol{y}=\boldsymbol{r}\}} \left\{ \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \left\{ \frac{1}{2}\left(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}}\right) + \frac{1}{2}\sqrt{(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}})^2 + \boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}} \right\} \right\}. \tag{5.2}$$

Proof : Please see Appendix B.2.

Remark : We observe that the function $f(u) = \frac{1}{2}u + \frac{1}{2}\sqrt{u^2 + \boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}}$ is everywhere non-decreasing in u, which allows us to express $\pi^2\left(r^0, \boldsymbol{r}\right)$ as the following tractable conic optimization problem:

$$\begin{aligned} \pi^2\left(r^0, \boldsymbol{r}\right) = \inf_{u,\boldsymbol{y}} \quad & \frac{1}{2}u + \frac{1}{2}\sqrt{u^2 + \boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}} \\ & r^0 + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \boldsymbol{r}'\hat{\boldsymbol{\zeta}} \le u \\ & \boldsymbol{F}'\boldsymbol{y} = \boldsymbol{r}. \end{aligned} \tag{5.3}$$

5.3 Mean and Directional Deviation Information

Theorem 3 *Let $\mathbb{F}_3$ be the family of all distributions $\mathbb{P}$ such that the mean of the segregated uncertainties, $\hat{\boldsymbol{\zeta}}$ has support in $\hat{\mathcal{V}}$, and the projected uncertainty vector $\tilde{\boldsymbol{z}}_\sigma = \boldsymbol{F}_\sigma\tilde{\boldsymbol{\zeta}} + \boldsymbol{g}_\sigma$ has independent components with directional deviations bounded above by $\boldsymbol{\sigma_f}$ and $\boldsymbol{\sigma_b}$, for known parameters $(\boldsymbol{F}_\sigma, \boldsymbol{g}_\sigma)$, i.e.*

$$\mathbb{F}_3 = \left\{ \mathbb{P} : \hat{\boldsymbol{\zeta}} = \mathrm{E}_{\mathbb{P}}\left(\tilde{\boldsymbol{\zeta}}\right) \in \hat{\mathcal{V}}, \boldsymbol{\sigma}_{\boldsymbol{f}\mathbb{P}}\left(\tilde{\boldsymbol{z}}_\sigma\right) \le \boldsymbol{\sigma_f}, \boldsymbol{\sigma}_{\boldsymbol{b}\mathbb{P}}\left(\tilde{\boldsymbol{z}}_\sigma\right) \le \boldsymbol{\sigma_b} \right\}.$$

Then $\pi^3\left(r^0, \boldsymbol{r}\right)$ is an upper bound for $\mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ over all distributions $\mathbb{P} \in \mathbb{F}_3$, i.e.

$$\sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) \le \pi^3\left(r^0, \boldsymbol{r}\right),$$

where

$$\pi^3\left(r^0, \boldsymbol{r}\right) \triangleq \inf_{\substack{s^0,\boldsymbol{s},x^0,\boldsymbol{x} \\ x^0+\boldsymbol{x}'\boldsymbol{g}_\sigma=r^0 \\ \boldsymbol{F}_\sigma'\boldsymbol{x}=\boldsymbol{r}}} \left\{ \begin{array}{lll} (r^0 - s^0 - \boldsymbol{s}'\boldsymbol{g}_\sigma) & + & \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}}(\boldsymbol{r}' - \boldsymbol{s}'\boldsymbol{F}_\sigma)\hat{\boldsymbol{\zeta}} \\ & + & \psi(s^0 - x^0, \boldsymbol{s} - \boldsymbol{x}) + \psi(s^0, \boldsymbol{s}) \end{array} \right\}, \tag{5.4}$$

and

$$\psi(x^0, \boldsymbol{x}) = \inf_{\lambda>0} \left\{ \frac{\lambda}{e} \exp\left(\frac{1}{\lambda} \sup_{\hat{\boldsymbol{z}}_\sigma \in \hat{\mathcal{W}}_\sigma} \left\{ x^0 + \boldsymbol{x}'\hat{\boldsymbol{z}}_\sigma \right\} + \frac{\|\boldsymbol{u}\|_2^2}{2\lambda^2} \right) \right\},$$

and $u_j = \max\{x_j\sigma_{f,j}, -x_j\sigma_{b,j}\}$.

Proof : Please see Appendix B.3.

Remark : Firstly, we notice that from the Model of Uncertainty in Section 3, we specify the linear transform parameter $\boldsymbol{H}_\sigma$ mapping the primitive uncertainty vector $\tilde{\boldsymbol{z}}$ to $\tilde{\boldsymbol{z}}_\sigma$, instead of specifying the affine transform parameters mapping the segregated uncertainty vector $\tilde{\boldsymbol{\zeta}}$ to $\tilde{\boldsymbol{z}}_\sigma$ directly. Nevertheless, $(\boldsymbol{F}_\sigma, \boldsymbol{g}_\sigma)$ is obtained as: $\boldsymbol{F}_\sigma = \boldsymbol{H}_\sigma\boldsymbol{F}$ and $\boldsymbol{g}_\sigma = \boldsymbol{H}_\sigma\boldsymbol{g}$. Secondly, unlike $\pi^1\left(r^0, \boldsymbol{r}\right)$ and $\pi^2\left(r^0, \boldsymbol{r}\right)$, this bound, while valid, is not tight. However, it remains useful because it any *component-wise independence* of the uncertainties to form a bound, and it has been shown computationally by See and Sim [33] to significantly improve the quality of the solution for a robust inventory problem. Finally, we note that evaluating $\pi^3\left(r^0, \boldsymbol{r}\right)$ involves exponential functions with quadratic arguments. Chen and Sim [12, Appendix B] showed that a small number of second-order conic constraints can be used to approximate such functions with good accuracy, and perform numerical studies showing the usefulness of the bound.

5.4 Unified Bounds

Each of the above functions separately bound $\mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ from above for $\mathbb{P}$ belonging to a given family of distributions $\mathbb{F} \in \{\mathbb{F}_1, \mathbb{F}_2, \mathbb{F}_3\}$. We now consider whether we are able to construct a better bound if we know that the actual distribution $\mathbb{P}$ lies in the intersection of these families. This can be done via a well-known technique in convex analysis known as infimal convolution (see Chen and Sim [12]). Due to its importance in our discussion, we reproduce it here in the following theorem:

Theorem 4 *Let $S \subseteq \{1,2,3\}$ be an index set of the bounds to be combined. Then the bound $\pi\left(r^0, \boldsymbol{r}\right)$, defined as*

$$\begin{array}{rll} \pi\left(r^0, \boldsymbol{r}\right) = & \min & \sum_{s\in S} \pi^s\left(r^{0,s}, \boldsymbol{r}^s\right) \\ & \text{s.t.} & r^0 = \sum_{s\in S} r^{0,s} \\ & & \boldsymbol{r} = \sum_{s\in S} \boldsymbol{r}^s, \end{array} \tag{5.5}$$

is a better bound for $\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ *than* $\pi^s\left(r^0, \boldsymbol{r}\right), \forall s \in S$, *where* $\mathbb{F} = \bigcap_{s\in S} \mathbb{F}_s$, *such that*

$$\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) \le \pi\left(r^0, \boldsymbol{r}\right) \le \pi^s\left(r^0, \boldsymbol{r}\right) \quad \forall s \in S. \tag{5.6}$$

Proof : Please see Appendix B.4.2.

Remark : On the theoretical front, we notice that the unified bound $\pi\left(r^0, \boldsymbol{r}\right)$ has the mathematically desirable properties of convexity and positive homogeneity. From a practical perspective, $\pi\left(r^0, \boldsymbol{r}\right)$ is able to selectively synthesize disparate pieces of information about the distribution of primitive uncertainties, and present a combined bound which takes into account all pieces of information.

6 Deflected Linear Decision Rules

We earlier showed that SLDRs improve over LDRs, while retaining a linear model of recourse. We aim to investigate if we can do even better. The deflected linear decision rule (DLDR) proposed by Chen et al. [14] exploited the *structure* of the model constraints (by solving a series of sub-problems based on the model parameters) to obtain an even more flexible decision rule. We adapt this idea here, similarly solving a series of sub-problems to exploit structural information within the model to generate a better decision rule, which we term the bi-deflected linear decision rule (BDLDR).

Although the original DLDR has been shown to be more flexible in comparison to LDRs, we will present an example in Section 6.1.1 where the DLDR can be further improved upon by an alternate piecewise-linear decision rule, which suggests that there is room for further improvement. In addition, the original DLDR of Chen et al. [14] does not explicitly handle expectation constraints or non-anticipativity requirements. We seek to address these limitations in the BDLDR which we present here.

Since the basic LDR can be obtained from the SLDR by choosing $\boldsymbol{M}(\boldsymbol{z}) = \boldsymbol{z}$, our results in this section also hold if we choose to omit constructing the SLDR as an intermediate step. However, we choose to present the techniques in this section as an additional layer of improvement over the SLDR for greater generality.

6.1 DLDR of Chen et. al. [14]

We review the two-stage optimization problem as in [14] under linear recourse, with non-negative constraints for a subset $J \subseteq [m]$ of indices as follows:

$$\begin{array}{rll} \min\limits_{\boldsymbol{x},\boldsymbol{r}(\cdot)} & \boldsymbol{c}'\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\boldsymbol{d}'\boldsymbol{r}(\tilde{\boldsymbol{\zeta}})\right) & \\ \text{s.t.} & \boldsymbol{T}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) & \\ & r_j(\tilde{\boldsymbol{\zeta}}) \geq 0 & \forall j \in J \\ & \boldsymbol{r} \in \mathcal{L}\left(m, N_E, \Phi\right). & \end{array} \tag{6.1}$$

We consider a series of sub-problems for each $i \in J$:

$$\begin{array}{rll} \min\limits_{\boldsymbol{p}} & \boldsymbol{d}'\boldsymbol{p} & \\ \text{s.t.} & \boldsymbol{U}\boldsymbol{p} = \boldsymbol{0} & \\ & p_i = 1 & \\ & p_j \geq 0 & \forall j \in J. \end{array} \tag{6.2}$$

Denoting by $J^\circ \subseteq J$ the set of indices where the problem (6.2) has a feasible solution, and by $\bar{\boldsymbol{p}}^i$, the optimal solution to the sub-problem (6.2) for each $i \in J^\circ$, the DLDR is then defined from the SLDR, by the following relation

$$\hat{\boldsymbol{r}}_{\boldsymbol{D}}(\tilde{\boldsymbol{\zeta}}) \triangleq \boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) + \sum_{i\in J^\circ} \left(r_i(\tilde{\boldsymbol{\zeta}})\right)^- \bar{\boldsymbol{p}}^i, \tag{6.3}$$

where the SLDR satisfies:

$$\begin{aligned} \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) &= \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) \\ r_j(\tilde{\boldsymbol{\zeta}}) &\geq 0 \qquad \forall j \in J \setminus J^\circ \\ \boldsymbol{r} &\in \mathcal{L}\left(m, N_E, \Phi\right). \end{aligned} \tag{6.4}$$

Problem (6.1) under the DLDR can be approximated by

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{r}(\cdot)} \quad & \boldsymbol{c}'\boldsymbol{x} + \sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\boldsymbol{d}'\boldsymbol{r}(\tilde{\boldsymbol{\zeta}})\right) + \sum_{i\in J_R^\circ} \sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\left(r_i(\tilde{\boldsymbol{\zeta}})\right)^-\right)\boldsymbol{d}'\bar{\boldsymbol{p}}^i \\ \text{s.t.} \quad & \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) \\ & r_j(\tilde{\boldsymbol{\zeta}}) \geq 0 && \forall j \in J \setminus J^\circ \\ & \boldsymbol{r} \in \mathcal{L}\left(m, N_E, \Phi\right). \end{aligned}$$

Where we define the reduced index set as $J_R^\circ = \left\{i \in J^\circ : \boldsymbol{d}'\bar{\boldsymbol{p}}^i > 0\right\}$ to avoid non-convexity in the objective. The objective is bounded from above by summing over only indices in J_R°, since the respective summation terms for $i \notin J_R^\circ$ are non-positive. We notice that the objective involves summing over terms of the form $\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left((\cdot)^-\right)$, which we can bound from above using the unified bound $\pi(\cdot)$ in Section 5. The two-stage DLDR model can therefore be expressed explicitly as:

$$\begin{aligned} Z_{DLDR} = \min_{\boldsymbol{x},\boldsymbol{r}^0,\boldsymbol{R}} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{r}^0 + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}}\left\{\boldsymbol{d}'\boldsymbol{R}\hat{\boldsymbol{\zeta}}\right\} + \sum_{i\in J_R^\circ} \pi\left(-r_i^0, -\boldsymbol{R}'\boldsymbol{e}^i\right)\boldsymbol{d}'\bar{\boldsymbol{p}}^i \\ \text{s.t.} \quad & \boldsymbol{\mathcal{T}}^0\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}^0 = \boldsymbol{\nu}^0 \\ & \boldsymbol{\mathcal{T}}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{R}\boldsymbol{e}^j = \boldsymbol{\nu}^j && \forall j \in [N_E] \\ & r_j^0 + \boldsymbol{e}^{j\prime}\boldsymbol{R}\boldsymbol{\zeta} \geq 0 \qquad \forall \boldsymbol{\zeta} \in \mathcal{V} && \forall j \in J \setminus J^\circ. \end{aligned} \tag{6.5}$$

6.1.1 Example: Limitation of DLDR and LDR

We will now consider an example of the two-stage problem (6.1), which will illustrate the limitation of the DLDR and motivate our subsequent exposition. For simplicity, we will discuss applying the DLDR to the LDR instead of the SLDR (i.e. using $\boldsymbol{M}(\boldsymbol{z}) = \boldsymbol{z}$). We consider the family $\mathbb{F}$ of scalar ($N = 1$) uncertainty distributions with infinite support ($\mathcal{W} = \Re$), zero mean $\left(\hat{\mathcal{W}} = \{0\}\right)$, and unit variance $\left(\sigma^2 = 1\right)$. We consider the following uncertain optimization problem with scalar recourse variables $y(\tilde{z})$, $u(\tilde{z})$, and $v(\tilde{z})$:

$$\begin{aligned} \min_{y(\cdot),u(\cdot),v(\cdot)} \quad & \sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(u(\tilde{z}) + v(\tilde{z})\right) \\ \text{s.t.} \quad & u(\tilde{z}) - v(\tilde{z}) = y(\tilde{z}) - \tilde{z} \\ & 0 \leq y(\tilde{z}) \leq 1 \\ & u(\tilde{z}), v(\tilde{z}) \geq 0 \\ & y, u, v \in \mathcal{Y}\left(1, 1, \{1\}\right). \end{aligned} \tag{6.6}$$

Using LDRs as our model of recourse, (i.e. using $\mathcal{L}$ to approximate $\mathcal{Y}$), we note that Problem (6.6) is infeasible (i.e. $Z_{LDR} = +\infty$), since the inequalities over the infinite support of $\tilde{z}$ cannot simultaneously fulfill the equality constraint. Now suppose we attempt to apply the DLDR to improve the solution, we will need a slack variable to convert the problem into the form of (6.1), with has model parameters:

$$\boldsymbol{U} = \begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{d} = \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix}, \quad J = \{1,2,3,4\}$$

Solving sub-problem (6.2) leads to the piecewise-linear decision rules:

$$\begin{aligned} \hat{u}_D(z) &= \left(u^0 + uz\right)^+ + \left(v^0 + vz\right)^- \\ \hat{v}_D(z) &= \left(v^0 + vz\right)^+ + \left(u^0 + uz\right)^- \\ y(z) &= y^0 + yz. \end{aligned}$$

Applying these decision rules, we obtain the following reformulation,

$$\begin{aligned} \min_{y^0, y, u^0, u, v^0, v} \quad & \sup_{\mathbb{P} \in \mathbb{F}} \mathrm{E}_{\mathbb{P}} \left(\left|u^0 + u\tilde{z}\right| + \left|v^0 + v\tilde{z}\right| \right) \\ \text{s.t.} \quad & u^0 - v^0 = y^0 \\ & u - v = y - 1 \\ & 0 \leq y^0 + yz \leq 1 \qquad\qquad \forall z \in \Re. \end{aligned}$$

After applying the bounds in Section 5, and noticing that the last inequality over all of $\Re$ implies $y = 0$, we get the final deterministic formulation which determines Z_{DLDR},

$$\begin{aligned} Z_{DLDR} \; = \; \min_{u^0, u, v^0, v} \quad & \left\| \begin{pmatrix} u^0 \\ u \end{pmatrix} \right\|_2 + \left\| \begin{pmatrix} v^0 \\ v \end{pmatrix} \right\|_2 \\ \text{s.t.} \quad & u - v = -1 \\ & 0 \leq u^0 - v^0 \leq 1. \end{aligned} \tag{6.7}$$

Solving, we get $Z_{DLDR} = 1$, which is a significant improvement over the LDR solution. We notice that even after applying the the DLDR, the decision rule $y(\cdot)$ remains as an LDR, and we would like to investigate whether we can further improve on this. Now, suppose we consider the following hypothetical piecewise-linear decision rule:

$$\begin{aligned} \hat{u}(z) &= \left(u^0 + uz\right)^+ + \left(v^0 + vz\right)^- + \left(y^0 + yz\right)^- \\ \hat{v}(z) &= \left(v^0 + vz\right)^+ + \left(u^0 + uz\right)^- + \left(y^0 - 1 + yz\right)^+ \\ \hat{y}(z) &= \left(y^0 + yz\right)^+ - \left(y^0 - 1 + yz\right)^+. \end{aligned}$$

We notice that under these decision rules, Problem (6.6), after applying the bounds, can be reduced to

$$\begin{aligned} Z_0 \; = \; \min_{y^0, y, u^0, u, v^0, v} \quad & \left\| \begin{pmatrix} u^0 \\ u \end{pmatrix} \right\|_2 + \left\| \begin{pmatrix} v^0 \\ v \end{pmatrix} \right\|_2 + \frac{1}{2} \left\| \begin{pmatrix} y^0 \\ y \end{pmatrix} \right\|_2 + \frac{1}{2} \left\| \begin{pmatrix} y^0 - 1 \\ y \end{pmatrix} \right\|_2 - \frac{1}{2} \\ \text{s.t.} \quad & u^0 - v^0 = y^0 \\ & u - v = y - 1. \end{aligned} \tag{6.8}$$

Solving the problem above, the optimal value is given by $Z_0 = \dfrac{1}{\sqrt{2}} < Z_{DLDR}$, a further improvement over the DLDR. We therefore seek a decision rule that would encompass our hypothetical piecewise-linear model of recourse.

6.2 Two-stage Bi-Deflected Linear Decision Rule

In this subsection, we will first introduce the BDLDR for a two stage problem, and later generalize it to a non-anticipative BDLDR (including the multi-stage as a special case) in the following subsection. We consider a two-stage problem similar to (6.1), as follows:

$$\begin{array}{rl} \min\limits_{\boldsymbol{x},\boldsymbol{r}(\cdot)} & \boldsymbol{c}'\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\boldsymbol{d}'\boldsymbol{r}(\tilde{\boldsymbol{\zeta}})\right) \\ \text{s.t.} & \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) \\ & \underline{\boldsymbol{y}} \le \boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) \le \overline{\boldsymbol{y}} \\ & \boldsymbol{r} \in \mathcal{L}(m, N_E, \Phi). \end{array} \tag{6.9}$$

Notice that for a two-stage problem we should have $\Phi = [N_E]$. Similar to the definition (2.3) of non-infinite bounds in our general model, we denote the index sets of non-infinite bounds in our two-stage model:

$$\begin{array}{l} \underline{J} = \left\{ i \in [m] : \underline{y}_i > -\infty \right\}, \\ \overline{J} = \left\{ i \in [m] : \overline{y}_i < +\infty \right\}. \end{array} \tag{6.10}$$

Notice that if we were to choose

$$\begin{array}{rcl} \underline{y}_j & = & \begin{cases} 0 & \forall j \in J \\ -\infty & \forall j \in [m] \setminus J \end{cases} \\ \overline{y}_j & = & +\infty \qquad \forall j \in [m], \end{array} \tag{6.11}$$

we obtain model (6.1) exactly. To construct the BDLDR, we consider the following pairs of optimization problems. Firstly, for each $i \in \underline{J}$,

$$\begin{array}{rll} \min\limits_{\boldsymbol{p}} & \boldsymbol{d}'\boldsymbol{p} & \\ \text{s.t.} & \boldsymbol{U}\boldsymbol{p} = \boldsymbol{0} & \\ & p_i = 1 & \\ & p_j \ge 0 & \forall j \in \underline{J} \\ & p_j \le 0 & \forall j \in \overline{J} \setminus \{i\}. \end{array} \tag{6.12}$$

Notice that for $j \in \left(\underline{J} \cap \overline{J}\right) \setminus \{i\}$, the constraints imply that $p_j = 0$. Similarly, for $i \in \overline{J}$,

$$\begin{array}{rll} \min\limits_{\boldsymbol{q}} & \boldsymbol{d}'\boldsymbol{q} & \\ \text{s.t.} & \boldsymbol{U}\boldsymbol{q} = \boldsymbol{0} & \\ & q_i = -1 & \\ & q_j \le 0 & \forall j \in \overline{J} \\ & q_j \ge 0 & \forall j \in \underline{J} \setminus \{i\}. \end{array} \tag{6.13}$$

By defining $\underline{J}^{\circ} \subseteq \underline{J}$ as the index set of i such that problem (6.12) is feasible, $\overline{J}^{\circ} \subseteq \overline{J}$ as the index set of i such that problem (6.13) is feasible, and $\bar{\boldsymbol{p}}^i, \bar{\boldsymbol{q}}^i$ as the respective optimal solutions for each i in $\underline{J}^{\circ}$ and $\overline{J}^{\circ}$ respectively, we consider an SLDR satisfying

$$\begin{array}{rll} \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) & & \\ r_j(\tilde{\boldsymbol{\zeta}}) \ge \underline{y}_j & & \forall j \in \underline{J} \setminus \underline{J}^{\circ} \\ r_j(\tilde{\boldsymbol{\zeta}}) \le \overline{y}_j & & \forall j \in \overline{J} \setminus \overline{J}^{\circ}. \end{array} \tag{6.14}$$

The associated BDLDR is then defined from the SLDR by:

$$\hat{\boldsymbol{r}}(\tilde{\boldsymbol{\zeta}}) \triangleq \boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) + \sum_{i\in \underline{J}^\circ} \left(r_i(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i\right)^- \overline{\boldsymbol{p}}^i + \sum_{i\in \overline{J}^\circ} \left(r_i(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i\right)^+ \overline{\boldsymbol{q}}^i. \tag{6.15}$$

Some properties of the BDLDR are stated in Proposition 4 which follows.

Proposition 4 *The Bi-Deflected Linear Decision Rule, $\hat{\boldsymbol{r}}(\tilde{\boldsymbol{\zeta}})$, satisfies the following properties*

1. $\boldsymbol{U}\hat{\boldsymbol{r}}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}})$,
2. $\underline{\boldsymbol{y}} \le \hat{\boldsymbol{r}}(\tilde{\boldsymbol{\zeta}}) \le \overline{\boldsymbol{y}}$.

Proof : Please see Appendix C.1.2.
This implies that as long as we have an SLDR that satisfies (6.14), we can find a feasible BDLDR. Under the BDLDR, problem (6.9) becomes:

$$\begin{array}{lll}
\min\limits_{\boldsymbol{x},\boldsymbol{r}(\cdot)} & \boldsymbol{c}'\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}} \left(\boldsymbol{d}'\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) + \sum\limits_{i\in \underline{J}^\circ} \left(r_i(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i\right)^- \boldsymbol{d}'\overline{\boldsymbol{p}}^i + \sum\limits_{i\in \overline{J}^\circ} \left(r_i(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i\right)^+ \boldsymbol{d}'\overline{\boldsymbol{q}}^i \right) & \\
\text{s.t.} & & \\
& \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) & \\
& r_j(\tilde{\boldsymbol{\zeta}}) \ge \underline{y}_j & \forall j \in \underline{J} \setminus \underline{J}^\circ \\
& r_j(\tilde{\boldsymbol{\zeta}}) \le \overline{y}_j & \forall j \in \overline{J} \setminus \overline{J}^\circ \\
& \boldsymbol{r} \in \mathcal{L}(m, N_E, \Phi). &
\end{array} \tag{6.16}$$

When $\boldsymbol{d}'\overline{\boldsymbol{p}}^i$ or $\boldsymbol{d}'\overline{\boldsymbol{p}}^i$ is negative, the objective becomes non-convex. Thus, we consider an approximation of Problem (6.16) from above, by defining the reduced index sets:

$$\begin{array}{rcl}
\underline{J}_R^\circ & \triangleq & \left\{ i \in \underline{J}^\circ : \boldsymbol{d}'\overline{\boldsymbol{p}}^i > 0 \right\}, \\
\overline{J}_R^\circ & \triangleq & \left\{ i \in \overline{J}^\circ : \boldsymbol{d}'\overline{\boldsymbol{q}}^i > 0 \right\}.
\end{array} \tag{6.17}$$

We then use the convexity of the supremum to obtain the formulation of the BDLDR problem:

$$\begin{array}{rll}
Z_{BDLDR}{}^* = \min\limits_{\boldsymbol{x},\boldsymbol{r}(\cdot)} & \boldsymbol{c}'\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}} \left(\boldsymbol{d}'\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) \right) & \\
& + \sum\limits_{i\in \underline{J}_R^\circ} \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}} \left(\left(r_i(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i\right)^- \right) \boldsymbol{d}'\overline{\boldsymbol{p}}^i + \sum\limits_{i\in \overline{J}_R^\circ} \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}} \left(\left(r_i(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i\right)^+ \right) \boldsymbol{d}'\overline{\boldsymbol{q}}^i & \\
\text{s.t.} & & \\
& \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) & \\
& r_j(\tilde{\boldsymbol{\zeta}}) \ge \underline{y}_j & \forall j \in \underline{J} \setminus \underline{J}^\circ \\
& r_j(\tilde{\boldsymbol{\zeta}}) \le \overline{y}_j & \forall j \in \overline{J} \setminus \overline{J}^\circ \\
& \boldsymbol{r} \in \mathcal{L}(m, N_E, \Phi). &
\end{array}$$

Using the bounds developed in Section 5 to approximate $\sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left((\cdot)^{\pm}\right)$, we obtain the explicit final form of the BDLDR model:

$$
\begin{aligned}
Z_{BDLDR} = \min_{\boldsymbol{x},\boldsymbol{r}^0,\boldsymbol{R}} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{r}^0 + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \left\{\boldsymbol{d}'\boldsymbol{R}\hat{\boldsymbol{\zeta}}\right\} \\
& + \sum_{i\in\underline{J}_R^\circ} \pi\left(-r_i^0 + \underline{y}_i, -\boldsymbol{R}'\boldsymbol{e}^i\right)\boldsymbol{d}'\overline{\boldsymbol{p}}^i + \sum_{i\in\overline{J}_R^\circ} \pi\left(r_i^0 - \overline{y}_i, \boldsymbol{R}'\boldsymbol{e}^i\right)\boldsymbol{d}'\overline{\boldsymbol{q}}^i \\
\text{s.t.} \quad & \\
& \boldsymbol{\mathcal{T}}^0\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}^0 = \boldsymbol{\nu}^0 \\
& \boldsymbol{\mathcal{T}}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{R}\boldsymbol{e}^j = \boldsymbol{\nu}^j \qquad \forall j \in [N_E] \\
& r_j^0 + \boldsymbol{e}^{j\prime}\boldsymbol{R}\boldsymbol{\zeta} \geq \underline{y}_j \qquad \forall\boldsymbol{\zeta}\in\mathcal{V} \quad \forall j\in\underline{J}\setminus\underline{J}^\circ \\
& r_j^0 + \boldsymbol{e}^{j\prime}\boldsymbol{R}\boldsymbol{\zeta} \leq \overline{y}_j \qquad \forall\boldsymbol{\zeta}\in\mathcal{V} \quad \forall j\in\overline{J}\setminus\overline{J}^\circ.
\end{aligned}
\tag{6.18}
$$

6.3 Comparison of BDLDR with DLDR and LDR

In this subsection, we will proceed to show that the BDLDR improves upon the DLDR and SLDR for Problem (6.9). Without loss of generality, we can consider a simplified version of the problem, such that the lower recourse constraint has the structure:

$$\underline{y}_j = 0, \ \forall j \in \underline{J},$$

since we can simply apply a change of variables if the above was not true. For reference, we begin by writing Problem (6.9) under the SLDR explicitly as:

$$
\begin{aligned}
Z_{SLDR} = \min_{\boldsymbol{x},\boldsymbol{r}^0,\boldsymbol{R}} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{r}^0 + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \left\{\boldsymbol{d}'\boldsymbol{R}\hat{\boldsymbol{\zeta}}\right\} \\
\text{s.t.} \quad & \boldsymbol{\mathcal{T}}^0\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}^0 = \boldsymbol{\nu}^0 \\
& \boldsymbol{\mathcal{T}}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{R}\boldsymbol{e}^j = \boldsymbol{\nu}^j \qquad \forall j \in [N_E] \\
& r_j^0 + \boldsymbol{e}^{j\prime}\boldsymbol{R}\boldsymbol{\zeta} \geq \underline{y}_j \qquad \forall\boldsymbol{\zeta}\in\mathcal{V} \quad \forall j\in\underline{J} \\
& r_j^0 + \boldsymbol{e}^{j\prime}\boldsymbol{R}\boldsymbol{\zeta} \leq \overline{y}_j \qquad \forall\boldsymbol{\zeta}\in\mathcal{V} \quad \forall j\in\overline{J}.
\end{aligned}
\tag{6.19}
$$

Next, we consider the DLDR. In order to apply the DLDR we need to introduce slack linear recourse variables $s_j(\tilde{\boldsymbol{\zeta}})$, $\forall j \in \overline{J}$ to convert (6.9) into the form of (6.1). The constraint set becomes:

$$
\begin{aligned}
& \boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) \\
& r_j(\tilde{\boldsymbol{\zeta}}) + s_j(\tilde{\boldsymbol{\zeta}}) = \overline{y}_j \qquad \forall j \in \overline{J} \\
& r_j(\tilde{\boldsymbol{\zeta}}) \geq 0 \qquad \forall j \in \underline{J} \\
& s_j(\tilde{\boldsymbol{\zeta}}) \geq 0 \qquad \forall j \in \overline{J} \\
& \boldsymbol{r}, \boldsymbol{s} \in \mathcal{L}(m, N_E, \Phi).
\end{aligned}
$$

Under the DLDR model of recourse, we need to solve the sub-problem (6.2) for each $i \in \underline{J}$, which corresponds to the inequalities $r_i(\tilde{\boldsymbol{\zeta}}) \geq 0$. After eliminating the slack variables, the problem takes the form:

$$\begin{aligned} \min_{\boldsymbol{p}} \quad & \boldsymbol{d}'\boldsymbol{p} \\ \text{s.t.} \quad & \boldsymbol{U}\boldsymbol{p} = \boldsymbol{0} \\ & p_i = 1 \\ & p_j \geq 0 \qquad \forall j \in \underline{J} \\ & p_j \leq 0 \qquad \forall j \in \overline{J}. \end{aligned} \tag{6.20}$$

Similarly, the inequalities $s_i(\tilde{\boldsymbol{\zeta}}) \geq 0$ requires us to solve the following sub-problem (again after eliminating the slacks) for each $i \in \overline{J}$,

$$\begin{aligned} \min_{\boldsymbol{q}} \quad & \boldsymbol{d}'\boldsymbol{q} \\ \text{s.t.} \quad & \boldsymbol{U}\boldsymbol{q} = \boldsymbol{0} \\ & q_i = -1 \\ & q_j \geq 0 \qquad \forall j \in \underline{J} \\ & q_j \leq 0 \qquad \forall j \in \overline{J}. \end{aligned} \tag{6.21}$$

We denote by $\underline{J}_D^\circ$ the set of indices $i \in \underline{J}$ such that (6.20) has a feasible solution, with corresponding optimal solution $\overline{\boldsymbol{p}}_{\boldsymbol{D}}^i$, and by $\overline{J}_D^\circ$ the set of indices $i \in \overline{J}$ such that (6.21) is feasible, with corresponding optimal solution $\overline{\boldsymbol{q}}_{\boldsymbol{D}}^i$. From the DLDR formulation (6.5), after rearrangement of terms, Problem (6.9) under the DLDR reduces to:

$$\begin{aligned} Z_{DLDR} = \min_{\boldsymbol{x}, \boldsymbol{r}^0, \boldsymbol{R}} \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{d}'\boldsymbol{r}^0 + \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} \left\{ \boldsymbol{d}'\boldsymbol{R}\hat{\boldsymbol{\zeta}} \right\} \\ & + \sum_{i \in \underline{J}_{D,R}^\circ} \pi\Big(-r_i^0 + \underline{y}_i, -\boldsymbol{R}'\boldsymbol{e}^i\Big)\boldsymbol{d}'\overline{\boldsymbol{p}}_{\boldsymbol{D}}^i + \sum_{i \in \overline{J}_{D,R}^\circ} \pi\Big(r_i^0 - \overline{y}_i, \boldsymbol{R}'\boldsymbol{e}^i\Big)\boldsymbol{d}'\overline{\boldsymbol{q}}_{\boldsymbol{D}}^i \\ \text{s.t.} \quad & \\ & \boldsymbol{\mathcal{T}}^0\boldsymbol{x} + \boldsymbol{U}\boldsymbol{r}^0 = \boldsymbol{\nu}^0 \\ & \boldsymbol{\mathcal{T}}^j\boldsymbol{x} + \boldsymbol{U}\boldsymbol{R}\boldsymbol{e}^j = \boldsymbol{\nu}^j \qquad\qquad \forall j \in [N_E] \\ & r_j^0 + \boldsymbol{e}^{j\prime}\boldsymbol{R}\boldsymbol{\zeta} \geq \underline{y}_j \qquad \forall \boldsymbol{\zeta} \in \mathcal{V} \quad \forall j \in \underline{J} \setminus \underline{J}_D^\circ \\ & r_j^0 + \boldsymbol{e}^{j\prime}\boldsymbol{R}\boldsymbol{\zeta} \leq \overline{y}_j \qquad \forall \boldsymbol{\zeta} \in \mathcal{V} \quad \forall j \in \overline{J} \setminus \underline{J}_D^\circ, \end{aligned} \tag{6.22}$$

where we define the reduced index sets: $\underline{J}_{D,R}^\circ \triangleq \left\{ i \in \underline{J}_D^\circ : \boldsymbol{d}'\overline{\boldsymbol{p}}_{\boldsymbol{D}}^i > 0 \right\}$ and $\overline{J}_{D,R}^\circ \triangleq \left\{ i \in \overline{J}_D^\circ : \boldsymbol{d}'\overline{\boldsymbol{q}}_{\boldsymbol{D}}^i > 0 \right\}$. We now summarize the result relating the optimal objectives to Problem (6.9) under the SLDR, DLDR, and BDLDR in the following proposition:

Proposition 5 *The optimal objective to Problem (6.9) under the BDLDR, DLDR, and the SLDR, and are related by the inequality:* $Z_{BDLDR} \leq Z_{DLDR} \leq Z_{SLDR}$.

Proof : Please see Appendix C.2.

Remark : The discussion above underscores an important distinction between deterministic linear optimization and robust linear optimization: In the deterministic case, it is trivial to include slack variables to convert the feasible set of a linear program into the standard form $\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$. The decision to convert or not is typically a result of the trade-off between storage space and solver performance, and would not affect the final optimal solution. However, in the robust case, using the LP "standard form" obscures the information of the set $\overline{J}$, which is in turn required to use the BDLDR. As we have seen, this could potentially deteriorate the optimal solution of the robust linear program.

6.4 Non-anticipative Bi-Deflected Linear Decision Rule

In the previous section, we introduced the two-stage BDLDR and showed that it generalizes and improves upon the DLDR and SLDR. We notice that each component of the two-stage BDLDR is a sum of its original underlying SLDR and piecewise-linear functions of other SLDRs, disregarding any information dependencies between the original SLDRs. Using the two-stage BDLDR could therefore violate non-anticipativity constraints present in the model, and result in an optimal solution which nonetheless improves upon the SLDR, but would otherwise be practically meaningless. In this section, we will further adapt the BDLDR for the case of non-anticipative recourse. This includes, but is not restricted to, multi-stage models. We will introduce the following notation:

$$\begin{array}{rcl} N^+(k) & = & \{j \in [K] : \Phi_k \subseteq \Phi_j\}\,, \\ N^-(k) & = & \{j \in [K] : \Phi_j \subseteq \Phi_k\}\,. \end{array} \tag{6.23}$$

Notice that $k \in N^{\pm}(k)\ \forall k \in [K]$, guaranteeing that $N^{\pm}(k)$ cannot be empty. Furthermore, the following property follows directly from the definitions above:

$$j \in N^+(k) \Leftrightarrow k \in N^-(j). \tag{6.24}$$

Now, for each $k \in [K]$, $i \in \underline{J}_k$, we consider the polyhedron $P(i,k)$ defined by the constraints:

$$\begin{array}{rcllll} \sum_{j\in N^+(k)} \boldsymbol{U}^j \boldsymbol{p}^{i,k,j} & = & \boldsymbol{0} & & \\ p_l^{i,k,j} & \geq & 0 & \forall l \in \underline{J}_j & \forall j \in N^+(k) \\ p_l^{i,k,j} & \leq & 0 & \forall l \in \overline{J}_j & \forall j \in N^+(k) \setminus \{k\} \\ p_l^{i,k,k} & \leq & 0 & \forall l \in \overline{J}_k \setminus \{i\} & \\ p_i^{i,k,k} & = & 1. & & \end{array} \tag{6.25}$$

Similarly, for each $k \in [K]$, $i \in \overline{J}_k$, we consider the polyhedron $Q(i,k)$ defined by the constraints:

$$\begin{array}{rcllll} \sum_{j\in N^+(k)} \boldsymbol{U}^j \boldsymbol{q}^{i,k,j} & = & \boldsymbol{0} & & \\ q_l^{i,k,j} & \leq & 0 & \forall l \in \overline{J}_j & \forall j \in N^+(k) \\ q_l^{i,k,j} & \geq & 0 & \forall l \in \underline{J}_j & \forall j \in N^+(k) \setminus \{k\} \\ q_l^{i,k,k} & \geq & 0 & \forall l \in \underline{J}_k \setminus \{i\} & \\ q_i^{i,k,k} & = & -1. & & \end{array} \tag{6.26}$$

For convenience, we collect the indices i which yield feasible instances of (6.25) and (6.26) in the following index sets $\forall k \in [K]$:

$$\begin{array}{l} \underline{J}_k^\circ = \{i \subseteq \underline{J}_k : P(i,k) \neq \emptyset\}\,, \\ \overline{J}_k^\circ = \left\{i \subseteq \overline{J}_k : Q(i,k) \neq \emptyset\right\}. \end{array} \tag{6.27}$$

Now suppose that we have a set of SLDRs, $\boldsymbol{r}^k \in \mathcal{L}(m_k, N_E, \Phi_k)$, $\forall k \in [K]$, which satisfies

$$\begin{aligned}
&\boldsymbol{\mathcal{T}}(\tilde{\boldsymbol{\zeta}})\boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{\nu}(\tilde{\boldsymbol{\zeta}}) \\
&\quad r_j^k(\tilde{\boldsymbol{\zeta}}) \geq \underline{y}_j^k \qquad \forall k \in [K], \forall j \in \underline{J}_k \setminus \underline{J}_k^\circ \\
&\quad r_j^k(\tilde{\boldsymbol{\zeta}}) \leq \overline{y}_j^k \qquad \forall k \in [K], \forall j \in \overline{J}_k \setminus \overline{J}_k^\circ \\
&\quad \boldsymbol{r}^k \in \mathcal{L}(m_k, N_E, \Phi_k) \qquad \forall k \in [K].
\end{aligned} \tag{6.28}$$

Based on the SLDR, we define the non-anticipative BDLDR, denoted by $\hat{\boldsymbol{r}}^k\left(\tilde{\boldsymbol{\zeta}}\right), \forall k \in [K]$, as

$$\hat{\boldsymbol{r}}^k(\tilde{\boldsymbol{\zeta}}) \triangleq \boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}}) + \sum_{j \in N^-(k)} \left(\sum_{i \in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j \right)^- \boldsymbol{p}^{i,j,k} + \sum_{i \in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j \right)^+ \boldsymbol{q}^{i,j,k} \right). \tag{6.29}$$

6.5 Properties of BDLDR

Proposition 6 *The information index set of the k^{th} BDLDR, denoted by $\hat{\Phi}_k$, is contained in the information index set of its underlying SLDR, Φ_k, i.e. $\hat{\Phi}_k \subseteq \Phi_k$. Furthermore, if problem (4.12) has a feasible solution, then $\hat{\Phi}_k = \Phi_k$.*

Proof : Please see Appendix C.3

Remark : This proposition implies that the non-anticipative BDLDR only uses information that is available to the the SLDR in forming the recourse decision. For example, if we are modeling a problem with a temporal revelation of information, the non-anticipative BDLDR uses information that has *already* been revealed.

Proposition 7 *Each non-anticipative BDLDR, $\hat{\boldsymbol{r}}^k(\tilde{\boldsymbol{\zeta}})$ satisfies the following properties:*

1. $\sum_{k=1}^K \boldsymbol{U}^k \hat{\boldsymbol{r}}^k(\tilde{\boldsymbol{\zeta}}) = \sum_{k=1}^K \boldsymbol{U}^k \boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}})$,
2. $\underline{\boldsymbol{y}}^k \leq \hat{\boldsymbol{r}}^k(\tilde{\boldsymbol{\zeta}}) \leq \overline{\boldsymbol{y}}^k, \quad \forall k \in [K]$.

Proof : Please see Appendix C.4.

6.6 Comparison of non-anticipative BDLDR with SLDR

We define the reduced index sets $\forall l \in \{0\} \cup [M], \forall k \in [K], \forall j \in N^-(k)$ as:

$$\begin{aligned}
\underline{J}_{l,j,k}^\circ &\triangleq \left\{ i \in \underline{J}_j^\circ : \boldsymbol{d}^{l,k\prime} \boldsymbol{p}^{i,j,k} > 0 \right\}, \\
\overline{J}_{l,j,k}^\circ &\triangleq \left\{ i \in \overline{J}_j^\circ : \boldsymbol{d}^{l,k\prime} \boldsymbol{q}^{i,j,k} > 0 \right\}.
\end{aligned}$$

Using Proposition 7, Problem (2.2) under the BDLDR then is approximated as:

$$
\begin{aligned}
& Z_{BDLDR} \\
= & \min_{\boldsymbol{x},\left\{\boldsymbol{r}^{0,k},\boldsymbol{R}^k\right\}_{k=1}^K} && \boldsymbol{c}^{0\prime}\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{r}^{0,k} + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}}\left\{\sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{R}^k\hat{\boldsymbol{\zeta}}\right\} \\
& && + \sum_{k=1}^K \sum_{j\in N^-(k)} \sum_{i\in \underline{J}^\circ_{0,j,k}} \pi\left(-r_i^{0,j} + \underline{y}_i^j, -\boldsymbol{R}^{j\prime}\boldsymbol{e}^i\right)\boldsymbol{d}^{0,k\prime}\boldsymbol{p}^{i,j,k} \\
& && + \sum_{k=1}^K \sum_{j\in N^-(k)} \sum_{i\in \overline{J}^\circ_{0,j,k}} \pi\left(r_i^{0,j} - \overline{y}_i^j, \boldsymbol{R}^{j\prime}\boldsymbol{e}^i\right)\boldsymbol{d}^{0,k\prime}\boldsymbol{q}^{i,j,k} \\
& \text{s.t.} && \boldsymbol{c}^{l\prime}\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{r}^{0,k} + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}}\left\{\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{R}^k\hat{\boldsymbol{\zeta}}\right\} \\
& && + \sum_{k=1}^K \sum_{j\in N^-(k)} \sum_{i\in \underline{J}^\circ_{l,j,k}} \pi\left(-r_i^{0,j} + \underline{y}_i^j, -\boldsymbol{R}^{j\prime}\boldsymbol{e}^i\right)\boldsymbol{d}^{l,k\prime}\boldsymbol{p}^{i,j,k} \\
& && + \sum_{k=1}^K \sum_{j\in N^-(k)} \sum_{i\in \overline{J}^\circ_{l,j,k}} \pi\left(r_i^{0,j} - \overline{y}_i^j, \boldsymbol{R}^{j\prime}\boldsymbol{e}^i\right)\boldsymbol{d}^{l,k\prime}\boldsymbol{q}^{i,j,k} \le b_l \quad \forall l\in[M] \\
& && \boldsymbol{\mathcal{T}}^0\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\boldsymbol{r}^{0,k} = \boldsymbol{\nu}^0 \\
& && \boldsymbol{\mathcal{T}}^j\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\boldsymbol{R}^k\boldsymbol{e}^j = \boldsymbol{\nu}^j \qquad \forall k\in[K], \forall j\in[N_E] \\
& && r_j^{0,k} + \boldsymbol{e}^{j\prime}\boldsymbol{R}^k\boldsymbol{\zeta} \ge \underline{y}_j^k \quad \forall\boldsymbol{\zeta}\in\mathcal{V} \qquad \forall k\in[K], j\in \underline{J}_k\setminus\underline{J}_k^\circ \\
& && r_j^{0,k} + \boldsymbol{e}^{j\prime}\boldsymbol{R}^k\boldsymbol{\zeta} \le \overline{y}_j^k \quad \forall\boldsymbol{\zeta}\in\mathcal{V} \qquad \forall k\in[K], j\in \overline{J}_k\setminus\overline{J}_k^\circ \\
& && \boldsymbol{x}\ge\boldsymbol{0}.
\end{aligned}
\tag{6.30}
$$

Proposition 8 *Problem (6.30) has an optimal objective not worse than Problem (4.12), i.e.* $Z_{BDLDR} \le Z_{SLDR}$.

Proof : Please see Appendix C.5.

Remark : Although we have shown that the non-anticipative BDLDR improves upon the SLDR solution, we notice that the BDLDR requires a choice of feasible points $\left\{\boldsymbol{p}^{i,k,j}\right\}_{j\in N^+(k)}$ and $\left\{\boldsymbol{q}^{i,k,j}\right\}_{j\in N^+(k)}$ which satisfy the polyhedral constraint sets (6.25) and (6.26) respectively. We leave it as an open question how to optimally choose points within these feasible polyhedra in the general case. If the model (2.2) can be reformulated such that there are no expectation constraints, we have a similar situation to the simpler two-stage BDLDR, and the optimal choices $\left\{\bar{\boldsymbol{p}}^{i,k,j}\right\}_{j\in N^+(k)}$ and $\left\{\bar{\boldsymbol{q}}^{i,k,j}\right\}_{j\in N^+(k)}$ can be found from the solutions of the following pairs of optimization problems:

$$
\begin{aligned}
\min \quad & \sum_{j\in N^+(k)} \boldsymbol{d}^{0,j}\boldsymbol{p}^{i,k,j} & \qquad \min \quad & \sum_{j\in N^+(k)} \boldsymbol{d}^{0,j}\boldsymbol{q}^{i,k,j} \\
\text{s.t.} \quad & \left\{\boldsymbol{p}^{i,k,j}\right\}_{j\in N^+(k)} \in P(i,k), & \qquad \text{s.t.} \quad & \left\{\boldsymbol{q}^{i,k,j}\right\}_{j\in N^+(k)} \in Q(i,k).
\end{aligned}
\tag{6.31}
$$

In the general case, due to the coupling between the expectation constraints and the objective, solving the pair of sub-problems (6.31) will no longer guarantee optimal $\left\{\bar{\boldsymbol{p}}^{i,k,j}\right\}_{j\in N^+(k)}$ and $\left\{\bar{\boldsymbol{q}}^{i,k,j}\right\}_{j\in N^+(k)}$ for use in the BDLDR. However, since (6.31) explicitly decreases the objective of the original problem, we feel that it remains as a viable heuristic for choosing a feasible points for use in the BDLDR, which will nonetheless be an improvement over the original SLDR.

7 Conclusions

We have presented a framework for the robust optimization of linear programs under uncertainty, by using linear-based decision rules to model the recourse variables. We have introduced SLDRs and BDLDRs, which are more flexible models of recourse decisions than LDRs, and have shown how they can be used in a non-anticipative modeling context. In a parallel work [25], we are developing a modeling language to model and solve the class of problems described in this paper. There we present modeling examples and comprehensive numerical studies for a service-constrained inventory management problem and a portfolio optimization problem. In particular, we demonstrate numerically that the non-anticipative BDLDR improves significantly over the LDR, verifying some of the theoretical results developed in this paper.

8 Acknowledgements

The authors would like to thank the associate editor and two anonymous referees for their comments and critique of our paper. Their helpful suggestions have helped us to improve the quality of this work.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[2] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99:351–376, 2004.

[3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[4] A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–13, August 1999.

[5] A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming Series A*, 88(3):411–424, 2000.

[6] A. Ben-Tal and M. Teboulle. Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Science*, 32(11):1445–1466, 1986.

[7] A. Ben-Tal and M. Teboulle. An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance*, 17(3):449–476, 2007.

[8] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

[9] D. Bertsimas and A. Thiele. A robust optimization approach to inventory theory. *Operations Research*, 54(1):150–168, 2006.

[10] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer, New York, 1997.

[11] M. Breton and S. El Hachem. Algorithms for the solution of stochastic dynamic minimax problems. *Computational Optimization and Applications*, 4:317–345, 1995.

[12] W. Chen and M. Sim. Goal-driven optimization. *Operations Research*, 57(2)(2):342–357, 2009.

[13] X. Chen, M. Sim, and P. Sun. A robust optimization perspective on stochastic programming. *Operations Research*, 55(6):1058–1071, 2007.

[14] X. Chen, M. Sim, P. Sun, and J. Zhang. A linear decision-based approximation approach to stochastic programming. *Operations Research*, 56(2):344–357, 2008.

[15] X. Chen and Y. Zhang. Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research*, 2009.

[16] V. K. Chopra and W. T. Ziemba. The effects of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19(2):6–11, 1993.

[17] E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, forthcoming.

[18] J. Dupačová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20(1):73–88, January 1987.

[19] M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. *Math. Programming Ser. A*, 106:423–432, 2006.

[20] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, October 1997.

[21] L. El Ghaoui, M. Oks, and F. Oustry. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, 2003.

[22] L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal of Optimization*, 9:33–52, 1998.

[23] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

[24] G. Gallego and I. Moon. The distribution free newsboy problem: Review and extensions. *The Journal of the Operational Research Society*, 44(8):825–834, 1993.

[25] J. Goh and M. Sim. Distributionally robust optimization with roam: Robust optimization via algebraic modeling. *Working Paper*, 2009.

[26] K. Isii. On sharpness of tchebycheff-type inequalities. *Annals of the Institute of Statistical Mathematics*, 14(1):185–197, December 1962.

[27] K. Natarajan, D. Pachamanova, and M. Sim. Incorporating asymmetric distributional information in robust value-at-risk optimization. *Management Science*, 54(3):573–585, 2008.

[28] K. Natarajan, M. Sim, and J. Uichanco. Tractable robust expected utility and risk models for portfolio optimization. *Mathematical Finance, forthcoming.*

[29] I. Popescu. Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–122, 2007.

[30] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:493–517, 2000.

[31] A. Ruszczynski and A. Shapiro, editors. *Stochastic Programming: Handbook in Operations Research and Management Science*. Elsevier Science, Amsterdam, 2003.

[32] H. Scarf. A min-max solution of an inventory problem. *Studies in The Mathematical Theory of Inventory and Production*, pages 201–209, 1958.

[33] C.-T. See and M. Sim. Robust approximation to multi-period inventory management. *Operations Research, forthcoming*, 2009.

[34] A. Shapiro and A. Kleywegt. Minimax analysis of stochastic programs. *Optimization Methods and Software*, 17(3):523–542, 2002.

[35] A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In V. Jeyakumar and A. Rubinov, editors, *Continuous Optimization*, pages 111–146. Springer USA, 2005.

[36] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157, 1973.

[37] C. Wang, C.-J. Ong, and M. Sim. Model predictive control using segregated disturbance feedback. In *Proceedings of the American Control Conference*, 2008.

[38] J. Žáčková. On minimax solution of stochastic linear programming problems. *Časopis pro Pěstování Matematiky*, 91:423–430, 1966.

Appendix A Proofs of LDR and SLDR reductions

A.1 Proof of Proposition 1

We use $\mathcal{L}(m_k, N, I_k)$ to approximate $\mathcal{Y}(m_k, N, I_k)$ in Problem (2.2). Applying the definition of $\mathcal{L}(m_k, N, I_k)$ in (4.1) for each $k \in [K]$, the problem above equivalently becomes:

$$\begin{array}{lll} \min\limits_{\boldsymbol{x}, \left\{\boldsymbol{y}^{0,k}, \boldsymbol{Y}^k\right\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\left(\boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\tilde{\boldsymbol{z}}\right)\right) & \\ \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \sup\limits_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\left(\boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\tilde{\boldsymbol{z}}\right)\right) \le b_l & \forall l \in [M] \\ & \boldsymbol{T}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\left(\boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\tilde{\boldsymbol{z}}\right) = \boldsymbol{v}(\tilde{\boldsymbol{z}}) & \\ & \underline{\boldsymbol{y}}^k \le \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\tilde{\boldsymbol{z}} \le \overline{\boldsymbol{y}}^k & \forall k \in [K] \\ & \boldsymbol{Y}^k\boldsymbol{e}^j = \boldsymbol{0} & \forall j \notin I_k, \forall k \in [K] \\ & \boldsymbol{x} \ge \boldsymbol{0}. & \end{array} \tag{A.1}$$

We now proceed to show that Problems (A.1) and (4.2) are equivalent. We first notice that due to linearity, the expectation terms in the objective first M constraints can be expressed as:

$$\begin{aligned} \sup_{\mathbb{P}\in\mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\left(\boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\tilde{\boldsymbol{z}}\right)\right) &= \sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{y}^{0,k} + \sup_{\mathbb{P}\in\mathbb{F}}\left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{Y}^k\mathrm{E}_{\mathbb{P}}\left(\tilde{\boldsymbol{z}}\right)\right) \\ &= \sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{y}^{0,k} + \sup_{\hat{\boldsymbol{z}}\in\hat{\mathcal{W}}}\left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{Y}^k\hat{\boldsymbol{z}}\right), \end{aligned}$$

for any $l \in \{0\} \cup [M]$. Next, for the following constraints to hold for the random variable $\tilde{\boldsymbol{z}}$,

$$\begin{array}{ll} \boldsymbol{T}(\tilde{\boldsymbol{z}})\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\left(\boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\tilde{\boldsymbol{z}}\right) = \boldsymbol{v}(\tilde{\boldsymbol{z}}) & \\ \underline{\boldsymbol{y}}^k \le \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\tilde{\boldsymbol{z}} \le \overline{\boldsymbol{y}}^k & \forall k \in [K], \end{array}$$

it is necessary and sufficient for the constraints to hold within the support, i.e.

$$\begin{array}{ll} \boldsymbol{T}(\boldsymbol{z})\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\left(\boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\boldsymbol{z}\right) = \boldsymbol{v}(\boldsymbol{z}) & \forall \boldsymbol{z} \in \mathcal{W} \\ \underline{\boldsymbol{y}}^k \le \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k\boldsymbol{z} \le \overline{\boldsymbol{y}}^k & \forall \boldsymbol{z} \in \mathcal{W}, \forall k \in [K]. \end{array}$$

Since the model data $\boldsymbol{T}(\cdot), \boldsymbol{v}(\cdot)$ are assumed to be affine in their respective arguments, we can equivalently re-write the equality constraint as a sum of the components of $\boldsymbol{z}$, as:

$$\left(\boldsymbol{T}^0\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\boldsymbol{y}^{0,k} - \boldsymbol{v}^0\right) + \sum_{i=1}^N z_i\left(\boldsymbol{T}^i\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k\boldsymbol{Y}^k\boldsymbol{e}^i - \boldsymbol{v}^i\right) = \boldsymbol{0} \quad \forall \boldsymbol{z} \in \mathcal{W}.$$

Finally, since $\mathcal{W}$ is assumed to be full-dimensional, the constraint holds iff the individual coefficients vanish, i.e.

$$\begin{aligned}\boldsymbol{T}^0\boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{y}^{0,k} &= \boldsymbol{v}^0 \\ \boldsymbol{T}^i\boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{Y}^k \boldsymbol{e}^i &= \boldsymbol{v}^i \quad \forall i \in [N] \,.\end{aligned}$$

Putting these all together, Problems (A.1) and (4.2) are equivalent as desired. ■

A.2 Proof of Proposition 2

For an arbitrary $\boldsymbol{z} \in \Re^N$, we consider $\boldsymbol{\zeta} = \boldsymbol{M}(\boldsymbol{z})$, and the components of $\boldsymbol{\gamma} = \boldsymbol{F}\boldsymbol{\zeta}$, $\forall i \in [N]$,

$$\begin{aligned}\gamma_i &= \boldsymbol{e}^{i\prime}\boldsymbol{F}\boldsymbol{\zeta} \\ &= \sum_{l=0}^{L-1} \boldsymbol{e}^{i+lN\prime}\boldsymbol{\zeta} \quad \text{(by structure of } \boldsymbol{F}\text{)} \\ &= \sum_{j\in\Phi(i)} \zeta_j .\end{aligned} \tag{A.2}$$

Where the set $\Phi(i)$ is defined for brevity as

$$\Phi(i) \triangleq \{ j \in [N_E] : i = ((j-1) \bmod N) + 1 \} \,. \tag{A.3}$$

Firstly, for each $i \in [N]$, we first consider the case where $\exists j \in \Phi(i), \zeta_j \notin \{\xi_{i,k}\}_{k=2}^{L}$. We can omit $k = 1$ and $k = L+1$ in the consideration of the set above, since $\zeta_j \neq \pm\infty$. In this case, since $\{\xi_{i,k}\}_{k=1}^{L+1}$ segments the extended real line, $\exists k^* \in [L]$ such that $\xi_{i,k^*} < \zeta_{j^*} < \xi_{i,k^*+1}$ for some $j^* \in \Phi(i)$. Furthermore, using the definition of a segregation (4.6), $\zeta_{j^*} = z_i$, and $\xi_{i,k^*} < z_i < \xi_{i,k^*+1}$. Using property (4.5), $\forall k \in [L+1]$, we get

$$\begin{aligned} z_i &> \xi_{i,k} \quad \text{if} \quad k < k^*, \\ z_i &< \xi_{i,k} \quad \text{if} \quad k > k^*. \end{aligned}$$

Hence, using the definition of a segregation (4.6) again, and recalling that $k = \lceil j/N \rceil$, this implies that $\forall j \in \Phi(i)$,

$$\begin{aligned} \zeta_j &= \xi_{i,k+1} \quad \text{if} \quad j < j^*, \\ \zeta_j &= \xi_{i,k} \quad \text{if} \quad j > j^*, \\ \zeta_j &= z_i \quad \text{if} \quad j = j^*. \end{aligned}$$

Hence, substituting into (A.2), we get

$$\begin{aligned}\gamma_i &= \sum_{j\in\Phi(i)} \zeta_j \\ &= \zeta_{j^*} + \sum_{\substack{j\in\Phi(i)\\ j<j^*}} \zeta_j + \sum_{\substack{j\in\Phi(i)\\ j>j^*}} \zeta_j \\ &= z_i + \sum_{k=2}^{L} \xi_{i,k} \\ &= z_i - g_i .\end{aligned}$$

Since this holds for each $i \in [N]$, we obtain $\boldsymbol{F}\boldsymbol{M}(\boldsymbol{z}) = \boldsymbol{z} - \boldsymbol{g}$, $\forall \boldsymbol{z} \in \Re^N$ such that $\exists j \in \Phi(i), \boldsymbol{e}^{j\prime}\boldsymbol{M}(\boldsymbol{z}) \notin \{\xi_{i,k}\}_{k=2}^{L}$.

Next, we consider the case that $\forall j \in \Phi(i), \zeta_j \in \{\xi_{i,k}\}_{k=2}^{L}$. There are L elements in $\Phi(i)$, and ζ_j can take on $L-1$ distinct values. Hence, we can apply the pigeonhole principle, which implies that $\exists k^* \in \{2, \ldots, L\}, \exists j_1, j_2 \in \Phi(i)$, such that $\zeta_{j_1} = \zeta_{j_2} = \xi_{i,k^*}$. From the definition of a segregation, (4.6), we can establish that $|j_1 - j_2| = N$, and $\zeta_{j_1} = \zeta_{j_2} = z_i$. We can express this alternatively as $\exists j^* \in \Phi(i)$, such that $z_i = \zeta_{j^*} = \zeta_{j^*+N} = \xi_{i,k^*}$. Finally, recalling that $k = \lceil j/N \rceil$, we again have

$$\begin{array}{rcll} \zeta_j & = & \xi_{i,k+1} & \text{if} \quad j < j^*, \\ \zeta_j & = & \xi_{i,k} & \text{if} \quad j > j^*, \\ \zeta_j & = & \xi_{i,k^*} = z_i & \text{if} \quad j = j^*. \end{array}$$

By the same argument as the previous case, we obtain $\boldsymbol{z} = \boldsymbol{F}\boldsymbol{M}(\boldsymbol{z}) + \boldsymbol{g}$, $\forall \boldsymbol{z} \in \Re^N$ such that $\forall j \in \Phi(i), \boldsymbol{e}^{j\prime}\boldsymbol{M}(\boldsymbol{z}) \in \{\xi_{i,k}\}_{k=2}^{L}$. Combining these two cases allows us to conclude that $\forall \boldsymbol{z} \in \Re^N$, $\boldsymbol{z}{=}\boldsymbol{F}\boldsymbol{M}(\boldsymbol{z}){+}\boldsymbol{g}$. ■

A.3 Proof of Proposition 3

To prove the first inequality, we express $\boldsymbol{\zeta} = \boldsymbol{M}(\boldsymbol{z})$, and equivalently re-express Problem (4.7) as

$$\begin{array}{llll} Z_{SLDR}{}^* = & \displaystyle\min_{\boldsymbol{x}, \{\boldsymbol{r}^k(\cdot)\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \displaystyle\sum_{k=1}^{K} \boldsymbol{d}^{0,k\prime}\boldsymbol{r}^{0,k} + \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}^*} \left(\sum_{k=1}^{K} \boldsymbol{d}^{0,k\prime}\boldsymbol{R}^k\hat{\boldsymbol{\zeta}} \right) & \\ & \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \displaystyle\sum_{k=1}^{K} \boldsymbol{d}^{l,k\prime}\boldsymbol{r}^{0,k} + \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}^*} \left(\sum_{k=1}^{K} \boldsymbol{d}^{l,k\prime}\boldsymbol{R}^k\hat{\boldsymbol{\zeta}} \right) \leq b_l & \forall l \in [M] \\ & & \boldsymbol{T}(\boldsymbol{F}\boldsymbol{\zeta} + \boldsymbol{g})\boldsymbol{x} + \left(\displaystyle\sum_{k=1}^{K} \boldsymbol{U}^k\boldsymbol{r}^{0,k} + \boldsymbol{U}^k\boldsymbol{R}^k\boldsymbol{\zeta} \right) = \boldsymbol{v}(\boldsymbol{F}\boldsymbol{\zeta} + \boldsymbol{g}) \quad \forall \boldsymbol{\zeta} \in \mathcal{V}^* & \\ & & \underline{\boldsymbol{y}}^k \leq \boldsymbol{r}^{0,k} + \boldsymbol{R}^k\boldsymbol{\zeta} \leq \overline{\boldsymbol{y}}^k \quad \forall \boldsymbol{\zeta} \in \mathcal{V}^* & \forall k \in [K] \\ & & \boldsymbol{x} \geq \boldsymbol{0} & \\ & & \boldsymbol{r}^k \circ \boldsymbol{M} \in \mathcal{Y}(m_k, N, I_k) & \forall k \in [K]. \end{array} \tag{A.4}$$

We assume that we have some $\left(\boldsymbol{x}, \left\{\boldsymbol{r}^{0,k}, \boldsymbol{R}^k\right\}_{k=1}^K\right)$ that is feasible in the approximated SLDR problem (4.12). The inclusion $\hat{\mathcal{V}}^* \subseteq \hat{\mathcal{V}}$, implies that the first M inequalities in Problem (A.4) are satisfied. Furthermore, the inclusion $\mathcal{V}^* \subseteq \mathcal{V}$ implies that the upper and lower bounds in Problem (A.4) are also satisfied. To show that the equality constraint in Problem (A.4) is satisfied, we consider the following expression for an arbitrary $\boldsymbol{\zeta} \in \mathcal{V}^*$:

$$
\begin{aligned}
& \boldsymbol{T}(\boldsymbol{F}\boldsymbol{\zeta}+\boldsymbol{g})\boldsymbol{x} + \left(\sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^{0,k} + \boldsymbol{U}^k \boldsymbol{R}^k \boldsymbol{\zeta}\right) - \boldsymbol{v}(\boldsymbol{F}\boldsymbol{\zeta}+\boldsymbol{g}) \\
= & \left(\boldsymbol{T}^0 \boldsymbol{x} + \sum_{i=1}^{N} (\boldsymbol{e}^{i\prime} \boldsymbol{F}\boldsymbol{\zeta} + g_i) \boldsymbol{T}^i \boldsymbol{x}\right) + \left(\sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^{0,k} + \boldsymbol{U}^k \boldsymbol{R}^k \boldsymbol{\zeta}\right) - \left(\boldsymbol{v}^0 + \sum_{i=1}^{N} (\boldsymbol{e}^{i\prime} \boldsymbol{F}\boldsymbol{\zeta} + g_i) \boldsymbol{v}^i\right) \\
= & \left(\boldsymbol{T}^0 \boldsymbol{x} + \sum_{i=1}^{N} g_i \boldsymbol{T}^i \boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^{0,k} - \boldsymbol{v}^0 - \sum_{i=1}^{N} g_i \boldsymbol{v}^i\right) + \sum_{i=1}^{N} (\boldsymbol{e}^{i\prime} \boldsymbol{F}\boldsymbol{\zeta}) \left(\boldsymbol{T}^i \boldsymbol{x} - \boldsymbol{v}^i\right) + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{R}^k \boldsymbol{\zeta} \\
= & \left(\boldsymbol{\mathcal{T}}^0 \boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^{0,k} - \boldsymbol{\nu}^0\right) + \sum_{i=1}^{N} (\boldsymbol{e}^{i\prime} \boldsymbol{F}\boldsymbol{\zeta}) \left(\boldsymbol{T}^i \boldsymbol{x} - \boldsymbol{v}^i\right) + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{R}^k \boldsymbol{\zeta} \\
= & \left(\boldsymbol{\mathcal{T}}^0 \boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^{0,k} - \boldsymbol{\nu}^0\right) + \sum_{i=1}^{N} \sum_{j=1}^{N_E} F_{ij} \zeta_j \left(\boldsymbol{T}^i \boldsymbol{x} - \boldsymbol{v}^i\right) + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{R}^k \sum_{i=j}^{N_E} \zeta_j \boldsymbol{e}^j \\
= & \left(\boldsymbol{\mathcal{T}}^0 \boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^{0,k} - \boldsymbol{\nu}^0\right) + \sum_{j=1}^{N_E} \zeta_j \sum_{i=1}^{N} F_{ij} \left(\boldsymbol{T}^i \boldsymbol{x} - \boldsymbol{v}^i\right) + \sum_{j=1}^{N_E} \zeta_j \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{R}^k \boldsymbol{e}^j \\
= & \left(\boldsymbol{\mathcal{T}}^0 \boldsymbol{x} + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{r}^{0,k} - \boldsymbol{\nu}^0\right) + \sum_{j=1}^{N_E} \zeta_j \left(\boldsymbol{\mathcal{T}}^j \boldsymbol{x} - \boldsymbol{\nu}^j + \sum_{k=1}^{K} \boldsymbol{U}^k \boldsymbol{R}^k \boldsymbol{e}^j\right).
\end{aligned}
$$

Hence, the system of equality constraints in Problem (4.12) implies that the above expression vanishes component-wise for any $\boldsymbol{\zeta} \in \Re^{N_E}$, which in turn implies that it vanishes component-wise for any $\boldsymbol{\zeta} \in \mathcal{V}^*$. Finally, we consider the non-anticipativity requirement. Denoting by $(\boldsymbol{M}(\boldsymbol{z}))_j$ the j^{th} component of $\boldsymbol{M}(\boldsymbol{z})$ for some $j \in [N_E]$,

$$
j \in \Phi_k \Leftrightarrow \left(\boldsymbol{M}\left(\boldsymbol{z} + \sum_{i \notin I_k} \lambda_i \boldsymbol{e}^i\right)\right)_j = \left(\boldsymbol{M}\left(\boldsymbol{z}\right)\right)_j \quad \forall \boldsymbol{\lambda} \in \Re^N \tag{A.5}
$$

for each $k \in [K]$. The forward direction follows directly from the definition of Φ_k in (4.11). The reverse direction results from requiring the equality to hold for all $\boldsymbol{\lambda} \in \Re^N$. The only components $(\boldsymbol{M}(\boldsymbol{z}))_j$ that are invariant to all $\boldsymbol{\lambda}$ are those with indices j in the set $\{j \in [N_E] : (i-1) \equiv (j-1) \bmod N\} = \Phi_k$.

Next, we expand the composite function:

$$
\begin{aligned}
\boldsymbol{r}^k \circ \boldsymbol{M}\left(\boldsymbol{z}\right) &= \boldsymbol{r}^{0,k} + \boldsymbol{R}^k \boldsymbol{M}\left(\boldsymbol{z}\right) \\
&= \boldsymbol{r}^{0,k} + \boldsymbol{R}^k \sum_{j=1}^{N_E} \left(\boldsymbol{M}\left(\boldsymbol{z}\right)\right)_j \boldsymbol{e}^j \\
&= \boldsymbol{r}^{0,k} + \sum_{j=1}^{N_E} \left(\boldsymbol{M}\left(\boldsymbol{z}\right)\right)_j \boldsymbol{R}^k \boldsymbol{e}^j \\
&= \boldsymbol{r}^{0,k} + \sum_{j \in \Phi_k} \left(\boldsymbol{M}\left(\boldsymbol{z}\right)\right)_j \boldsymbol{R}^k \boldsymbol{e}^j,
\end{aligned} \tag{A.6}
$$

where the last equality is due to the assumption of feasibility in (4.12), which gives $\boldsymbol{R}^k \boldsymbol{e}^j = \boldsymbol{0}, \forall j \notin \Phi_k$.

Hence, for an arbitrary $\boldsymbol{\lambda} \in \Re^N$,

$$\begin{aligned}
\boldsymbol{r}^k \circ \boldsymbol{M}\left(\boldsymbol{z} + \sum_{i \notin I_k} \lambda_i \boldsymbol{e}^i\right) &= \boldsymbol{r}^{0,k} + \sum_{j \in \Phi_k} \left(\boldsymbol{M}\left(\boldsymbol{z} + \sum_{i \notin I_k} \lambda_i \boldsymbol{e}^i\right)\right)_j \boldsymbol{R}^k \boldsymbol{e}^j && \text{(by (A.6))} \\
&= \boldsymbol{r}^{0,k} + \sum_{j \in \Phi_k} \left(\boldsymbol{M}(\boldsymbol{z})\right)_j \boldsymbol{R}^k \boldsymbol{e}^j && \text{(by (A.5))} \\
&= \boldsymbol{r}^k \circ \boldsymbol{M}(\boldsymbol{z}) && \text{(by (A.6))},
\end{aligned}$$

implying that $\boldsymbol{r}^k \circ \boldsymbol{M} \in \mathcal{Y}(m_k, N, I_k)$ as required. Therefore, we have established that any feasible solution to (4.12) is always feasible in (A.4), with an objective that is not smaller. Hence we have $Z_{SLDR}{}^* \leq Z_{SLDR}$.

To prove the second inequality, we consider Problem (4.12), and choose $\forall k \in [K]$,

$$\begin{aligned}
\boldsymbol{R}^k &= \boldsymbol{Y}^k \boldsymbol{F}, \\
\boldsymbol{r}^{0,k} &= \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k \boldsymbol{g}.
\end{aligned}$$

Problem (4.12) becomes:

$$\begin{aligned}
Z_{SLDR} = \min_{\boldsymbol{x}, \left\{\boldsymbol{y}^{0,k}, \boldsymbol{Y}^k\right\}_{k=1}^K} \quad & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{y}^{0,k} + \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} \left(\sum_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{Y}^k(\boldsymbol{F}\hat{\boldsymbol{\zeta}} + \boldsymbol{g})\right) \\
\text{s.t.} \quad & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{y}^{0,k} + \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} \left(\sum_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{Y}^k(\boldsymbol{F}\hat{\boldsymbol{\zeta}} + \boldsymbol{g})\right) \leq b_l \quad \forall l \in [M] \\
& \boldsymbol{\mathcal{T}}^0 \boldsymbol{x} + \left(\sum_{k=1}^K \boldsymbol{U}^k \boldsymbol{y}^{0,k} + \boldsymbol{U}^k \boldsymbol{Y}^k \boldsymbol{g}\right) = \boldsymbol{\nu}^0 \\
& \boldsymbol{\mathcal{T}}^j \boldsymbol{x} + \sum_{k=1}^K \boldsymbol{U}^k \boldsymbol{Y}^k \boldsymbol{F} \boldsymbol{e}^j = \boldsymbol{\nu}^j \qquad \forall j \in [N_E] \\
& \underline{\boldsymbol{y}}^k \leq \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k(\boldsymbol{F}\boldsymbol{\zeta} + \boldsymbol{g}) \leq \overline{\boldsymbol{y}}^k \quad \forall \boldsymbol{\zeta} \in \mathcal{V} \qquad \forall k \in [K] \\
& \boldsymbol{Y}^k \boldsymbol{F} \boldsymbol{e}^j = \boldsymbol{0} \qquad \forall j \notin \Phi_k, \forall k \in [K] \\
& \boldsymbol{x} \geq \boldsymbol{0}.
\end{aligned}$$

Expanding the terms $\left\{\boldsymbol{\mathcal{T}}^j, \boldsymbol{\nu}^j\right\}_{j=0}^{N_E}$, we obtain

$$\begin{array}{rlll}
Z_{SLDR} = & \min\limits_{\boldsymbol{x},\left\{\boldsymbol{y}^{0,k},\boldsymbol{Y}^k\right\}_{k=1}^K} & \boldsymbol{c}^{0\prime}\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{y}^{0,k} + \sup\limits_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \left(\sum\limits_{k=1}^K \boldsymbol{d}^{0,k\prime}\boldsymbol{Y}^k(\boldsymbol{F}\hat{\boldsymbol{\zeta}} + \boldsymbol{g})\right) & \\
& \text{s.t.} & \boldsymbol{c}^{l\prime}\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{y}^{0,k} + \sup\limits_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \left(\sum\limits_{k=1}^K \boldsymbol{d}^{l,k\prime}\boldsymbol{Y}^k(\boldsymbol{F}\hat{\boldsymbol{\zeta}} + \boldsymbol{g})\right) \leq b_l & \forall l \in [M] \\
& & \left(\boldsymbol{T}^0\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{U}^k\boldsymbol{y}^{0,k} - \boldsymbol{v}^0\right) & \\
& & \quad + \sum\limits_{i=1}^N g_i \left(\boldsymbol{T}^i\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{U}^k\boldsymbol{Y}^k\boldsymbol{e}^i - \boldsymbol{v}^i\right) = \boldsymbol{0} & \\
& & \sum\limits_{i=1}^N F_{ij} \left(\boldsymbol{T}^i\boldsymbol{x} + \sum\limits_{k=1}^K \boldsymbol{U}^k\boldsymbol{Y}^k\boldsymbol{e}^i - \boldsymbol{v}^i\right) = \boldsymbol{0} & \forall j \in [N_E] \\
& & \underline{\boldsymbol{y}}^k \leq \boldsymbol{y}^{0,k} + \boldsymbol{Y}^k(\boldsymbol{F}\boldsymbol{\zeta} + \boldsymbol{g}) \leq \overline{\boldsymbol{y}}^k \qquad \forall \boldsymbol{\zeta} \in \mathcal{V} & \forall k \in [K] \\
& & \boldsymbol{Y}^k\boldsymbol{F}\boldsymbol{e}^j = \boldsymbol{0} & \forall j \notin \Phi_k, \forall k \in [K] \\
& & \boldsymbol{x} \geq \boldsymbol{0}. &
\end{array}$$

Since $\boldsymbol{M}(\cdot)$ represents a segregation, $\boldsymbol{F}$ is the horizontal concatenation of L identity matrices. Hence, $\forall j \in [N_E]$, $\boldsymbol{F}\boldsymbol{e}^j = \boldsymbol{e}^i$ where $(i-1) \equiv (j-1) \bmod N$. Notice that $\boldsymbol{e}^j \in \Re^{N_E}$ while $\boldsymbol{e}^i \in \Re^N$. In particular, for any index set $I \subseteq [N]$,

$$\boldsymbol{Y}\boldsymbol{e}^i = \boldsymbol{0} \quad \forall i \in I \quad \Leftrightarrow \quad \boldsymbol{Y}\boldsymbol{F}\boldsymbol{e}^j = \boldsymbol{0} \quad \forall j \in \{j \in [N_E] : \exists i \in I : (i-1) \equiv (j-1) \bmod N\}.$$

In particular, if we choose $I = I_k^c$, then using the definition of Φ_k (4.11), we can express the above as

$$\boldsymbol{Y}\boldsymbol{e}^i = \boldsymbol{0} \quad \forall i \notin I_k \quad \Leftrightarrow \quad \boldsymbol{Y}\boldsymbol{F}\boldsymbol{e}^j = \boldsymbol{0} \quad \forall j \notin \Phi_k,$$

by applying the definition of Φ_k (4.11), Hence, using (4.8), any feasible point $\left(\boldsymbol{x}, \left\{\boldsymbol{y}^{0,k}, \boldsymbol{Y}^k\right\}_{k=1}^K\right)$ in Problem (4.2) is also feasible in Problem (4.12), and since their objectives coincide, we have $Z_{SLDR} \leq Z_{LDR}$. ■

Appendix B Proofs of Bounds on $\mathrm{E}_{\mathbb{P}}\left((\cdot)^+\right)$

B.1 Proof of Theorem 1

For the case of a fixed mean $\hat{\boldsymbol{\zeta}} = \boldsymbol{\mu}$, Natarajan et al. [28, Theorem 2.2], provided a tight bound in for the expectation of a general piecewise-linear utility function applied to an LDR. We specialize their result for the case of the utility function $u(x) = x^+$, to obtain

$$\sup_{supp(\tilde{\boldsymbol{\zeta}})\subseteq\mathcal{V},\hat{\boldsymbol{\zeta}}=\boldsymbol{\mu}} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = \inf_{\boldsymbol{s}\in\Re^{N_E}} \left(\boldsymbol{s}'\boldsymbol{\mu} + \sup_{\boldsymbol{\zeta}\in\mathcal{V}}\left(\max\left\{r^0 + \boldsymbol{r}'\boldsymbol{\zeta} - \boldsymbol{s}'\boldsymbol{\zeta}, -\boldsymbol{s}'\boldsymbol{\zeta}\right\}\right)\right),$$

and equality is obtained because of the strong duality result of Isii [26]. In general, if the mean is not fixed, the ambiguity-averse bound on $\mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ is simply obtained by taking the supremum over the allowed values of $\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}$, which yields

$$\sup_{\mathbb{P}\in\mathbb{F}_1} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = \pi^1\left(r^0, \boldsymbol{r}\right),$$

as required. ■

B.2 Proof of Theorem 2

In Natarajan et al [28, Theorem 2.1], the authors use a projection method by Popescu [29, Theorem 1] to show that if $\tilde{\boldsymbol{\zeta}}$ has a known mean $\hat{\boldsymbol{\zeta}} = \boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}_{\mathcal{V}}$, the following equality holds:

$$\sup_{\tilde{\boldsymbol{\zeta}}\sim(\boldsymbol{\mu},\boldsymbol{\Sigma}_{\mathcal{V}})} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = \frac{1}{2}\left(r^0 + \boldsymbol{r}'\boldsymbol{\mu}\right) + \frac{1}{2}\sqrt{(r^0 + \boldsymbol{r}'\boldsymbol{\mu})^2 + \boldsymbol{r}'\boldsymbol{\Sigma}_{\mathcal{V}}\boldsymbol{r}}.$$

To construct the worst-case bounds of $\mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ over all $\mathbb{P} \in \mathbb{F}_2$, we simply need to find the supremum over all allowable $\left(\hat{\boldsymbol{\zeta}}, \boldsymbol{\Sigma}_{\mathcal{V}}\right)$ in $\mathbb{F}_2$. We obtain the bound by solving the following optimization problem:

$$\begin{aligned}\sup_{\mathbb{P}\in\mathbb{F}_2} \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = \eta^2(r^0, \boldsymbol{r}) \triangleq \quad & \sup_{\hat{\boldsymbol{\zeta}},\boldsymbol{\Sigma}_{\mathcal{V}}} \left\{\frac{1}{2}\left(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}}\right) + \frac{1}{2}\sqrt{(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}})^2 + \boldsymbol{r}'\boldsymbol{\Sigma}_{\mathcal{V}}\boldsymbol{r}}\right\} \\ & \text{s.t.} \quad \boldsymbol{F}\boldsymbol{\Sigma}_{\mathcal{V}}\boldsymbol{F}' = \boldsymbol{\Sigma} \\ & \qquad \boldsymbol{\Sigma}_{\mathcal{V}} \in \mathbb{S}_+^{N_E} \\ & \qquad \hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}},\end{aligned}$$

where $\mathbb{S}_+^N$ denotes the positive semidefinite cone of symmetric $N \times N$ matrices. We complete the proof by showing that the bounds $\eta^2(r^0, \boldsymbol{r})$ and $\pi^2\left(r^0, \boldsymbol{r}\right)$ are equivalent. Suppose $\exists \boldsymbol{y} \in \Re^N$ such that $\boldsymbol{F}'\boldsymbol{y} = \boldsymbol{r}$, then $\boldsymbol{r}'\boldsymbol{\Sigma}_{\mathcal{V}}\boldsymbol{r} = \boldsymbol{y}'\boldsymbol{F}\boldsymbol{\Sigma}_{\mathcal{V}}\boldsymbol{F}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{\Sigma}\boldsymbol{y}$, and the bounds are easily seen to be equivalent. Now suppose $\nexists \boldsymbol{y} \in \Re^N$ such that $\boldsymbol{F}'\boldsymbol{y} = \boldsymbol{r}$. This causes the outer optimization problem defining $\pi^2\left(r^0, \boldsymbol{r}\right)$ to be infeasible, and $\pi^2\left(r^0, \boldsymbol{r}\right) = +\infty$. We proceed to establish that $\eta^2(r^0, \boldsymbol{r}) = +\infty$ as well. To begin the proof, we choose

$$\boldsymbol{\Sigma}_{\mathcal{V}}^0 = \boldsymbol{F}'\left(\boldsymbol{F}\boldsymbol{F}'\right)^{-1}\boldsymbol{\Sigma}\left(\boldsymbol{F}\boldsymbol{F}'\right)^{-1}\boldsymbol{F},$$

which satisfies $\boldsymbol{F}\boldsymbol{\Sigma}_{\mathcal{V}}^{0}\boldsymbol{F}' = \boldsymbol{\Sigma}$. We notice that $\boldsymbol{F}\boldsymbol{F}'$ is invertible since $N_E \geq N$, and $\boldsymbol{F}$ is assumed to be full rank. Next, we choose $\boldsymbol{y} = (\boldsymbol{F}\boldsymbol{F}')^{-1}\boldsymbol{F}\boldsymbol{r}$, and express $\boldsymbol{r} = \boldsymbol{F}'\boldsymbol{y} + \boldsymbol{r}_{\perp}$. By assumption, $\boldsymbol{F}'\boldsymbol{y} \neq \boldsymbol{r}$, which implies $\boldsymbol{r}_{\perp} \neq \boldsymbol{0}$. Furthermore, we have

$$\begin{aligned} \boldsymbol{F}\boldsymbol{r}_{\perp} &= \boldsymbol{F}\boldsymbol{r} - \boldsymbol{F}\boldsymbol{F}'\boldsymbol{y} \\ &= \boldsymbol{F}\boldsymbol{r} - \boldsymbol{F}\boldsymbol{F}'(\boldsymbol{F}\boldsymbol{F}')^{-1}\boldsymbol{F}\boldsymbol{r} \\ &= \boldsymbol{0}. \end{aligned}$$

Now, for some $\lambda \in \Re_+$, consider

$$\boldsymbol{\Sigma}_{\mathcal{V}}(\lambda) = \boldsymbol{\Sigma}_{\mathcal{V}}^{0} + \lambda \boldsymbol{r}_{\perp}\boldsymbol{r}_{\perp}'.$$

We notice that $\boldsymbol{\Sigma}_{\mathcal{V}}(\lambda) \in \mathbb{S}_{+}^{N_E}$. Furthermore, we have

$$\begin{aligned} \boldsymbol{F}\Big(\boldsymbol{\Sigma}_{\mathcal{V}}(\lambda)\Big)\boldsymbol{F}' &= \boldsymbol{F}\boldsymbol{\Sigma}_{\mathcal{V}}^{0}\boldsymbol{F}' + \boldsymbol{0} \\ &= \boldsymbol{\Sigma}\ . \end{aligned}$$

Hence, for any $\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}$, $\lambda \in \Re_+$, $\eta^2\left(r^0, \boldsymbol{r}\right)$ is bounded from below by

$$\begin{aligned} \eta^2\left(r^0, \boldsymbol{r}\right) &\geq \frac{1}{2}\left(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}}\right) + \frac{1}{2}\sqrt{(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}})^2 + \boldsymbol{r}'\boldsymbol{\Sigma}_{\mathcal{V}}(\lambda)\boldsymbol{r}} \\ &= \frac{1}{2}\left(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}}\right) + \frac{1}{2}\sqrt{(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}})^2 + \boldsymbol{r}'\boldsymbol{\Sigma}_{\mathcal{V}}^{0}\boldsymbol{r} + \lambda\left(\boldsymbol{r}'\boldsymbol{r}_{\perp}\right)^2} \\ &= \frac{1}{2}\left(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}}\right) + \frac{1}{2}\sqrt{(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}})^2 + \boldsymbol{r}'\boldsymbol{\Sigma}_{\mathcal{V}}^{0}\boldsymbol{r} + \lambda\left(\boldsymbol{y}'\boldsymbol{F}\boldsymbol{r}_{\perp} + \boldsymbol{r}_{\perp}'\boldsymbol{r}_{\perp}\right)^2} \\ &= \frac{1}{2}\left(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}}\right) + \frac{1}{2}\sqrt{(r^0 + \boldsymbol{r}'\hat{\boldsymbol{\zeta}})^2 + \boldsymbol{r}'\boldsymbol{\Sigma}_{\mathcal{V}}^{0}\boldsymbol{r} + \lambda\left\|\boldsymbol{r}_{\perp}\right\|_2^4}. \end{aligned}$$

Taking the limit as $\lambda \to \infty$, the lower bound (i.e. right-hand side) approaches $+\infty$. Thus, if $\nexists \boldsymbol{y}$ such that $\boldsymbol{F}'\boldsymbol{y} = \boldsymbol{r}$, then $\eta^2(r^0, \boldsymbol{r}) = +\infty$ as desired. ■

B.3 Proof of Theorem 3

We only have to prove the bound in the non-infinite case. We begin by noticing that we can express $x^0 + \boldsymbol{x}'\tilde{\boldsymbol{z}}_\sigma$ as:

$$x^0 + \boldsymbol{x}'\tilde{\boldsymbol{z}}_\sigma \equiv x^0 + \boldsymbol{x}'\hat{\boldsymbol{z}}_\sigma + \boldsymbol{x}'\left(\tilde{\boldsymbol{z}}_\sigma - \hat{\boldsymbol{z}}_\sigma\right).$$

Now, we use the property $\forall \lambda > 0$ that

$$w^+ \leq \frac{\lambda}{e}\exp\left(\frac{w}{\lambda}\right), \forall w \in \Re,$$

and the independence of each component of $\tilde{\boldsymbol{z}}_\sigma$ to obtain the general bound

$$\mathrm{E}_{\mathbb{P}}\left(\left(x^0 + \boldsymbol{x}'\tilde{\boldsymbol{z}}_\sigma\right)^+\right) \leq \frac{\lambda}{e}\exp\left(\frac{1}{\lambda}\sup_{\hat{\boldsymbol{z}}_\sigma \in \hat{\mathcal{W}}_\sigma}\left\{x^0 + \boldsymbol{x}'\hat{\boldsymbol{z}}_\sigma\right\}\right)\prod_{j=1}^{N_\sigma}\mathrm{E}_{\mathbb{P}}\left(\exp\left(\frac{x_j\left(\tilde{z}_{\sigma,j} - \hat{z}_{\sigma,j}\right)}{\lambda}\right)\right).$$

Now, using the definition of the forward and backward deviations in [13, Equations (17) and (18)], $\forall \mathbb{P} \in \mathbb{F}_3$,

$$\ln\left(\mathrm{E}_{\mathbb{P}}\left(\exp\left(\frac{x_j\left(\tilde{z}_{\sigma,j} - \hat{z}_{\sigma,j}\right)}{\lambda}\right)\right)\right) \leq \begin{cases} x_j^2\sigma_{f,j}^2/2\lambda^2 & \text{if } x_j \geq 0, \\ x_j^2\sigma_{b,j}^2/2\lambda^2 & \text{otherwise.} \end{cases}$$

Combining these results, when $(r^0, \boldsymbol{r}) = (x^0 + \boldsymbol{x}'\boldsymbol{g}_\sigma, \boldsymbol{F}'_\sigma \boldsymbol{x})$, we get

$$\sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) = \sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left((x^0 + \boldsymbol{x}'\tilde{\boldsymbol{z}}_\sigma)^+\right) \le \psi(x^0, \boldsymbol{x}). \tag{B.1}$$

Next, using the identity $x^+ \equiv x + x^- \ \forall x \in \Re$, by the same argument,

$$\sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) \le \left(r^0 + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \boldsymbol{r}'\hat{\boldsymbol{\zeta}}\right) + \psi(-x^0, -\boldsymbol{x}). \tag{B.2}$$

Since choosing $(s^0, \boldsymbol{s}) = (x^0, \boldsymbol{x})$ in (5.4) reduces to (B.1) and choosing $(s^0, \boldsymbol{s}) = (0, \boldsymbol{0})$ in (5.4) reduces to (B.2), we have shown that $\pi^3\left(r^0, \boldsymbol{r}\right)$ is not larger than either (B.1) or (B.2). Finally, we establish that $\pi^3\left(r^0, \boldsymbol{r}\right)$ indeed bounds $\mathrm{E}_\mathbb{P}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ from above. For any $(s^0, \boldsymbol{s}, x^0, \boldsymbol{x})$ such that $(x^0 + \boldsymbol{x}'\boldsymbol{g}_\sigma, \boldsymbol{F}'_\sigma \boldsymbol{x}) = (r^0, \boldsymbol{r})$, we have

$$\begin{aligned}
&(r^0 - s^0) - \boldsymbol{s}'\boldsymbol{g}_\sigma + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \left(\boldsymbol{r} - \boldsymbol{s}'\boldsymbol{F}_\sigma\right)\hat{\boldsymbol{\zeta}} + \psi(s^0 - x^0, \boldsymbol{s} - \boldsymbol{x}) + \psi(s^0, \boldsymbol{s}) \\
\ge\ & (x^0 - s^0) + (\boldsymbol{x} - \boldsymbol{s})'\boldsymbol{g}_\sigma + \sup_{\hat{\boldsymbol{\zeta}}\in\hat{\mathcal{V}}} \left((\boldsymbol{x} - \boldsymbol{s})'\boldsymbol{F}_\sigma\hat{\boldsymbol{\zeta}}\right) + \psi(s^0 - x^0, \boldsymbol{s} - \boldsymbol{x}) + \psi(s^0, \boldsymbol{s}) \\
\ge\ & \sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left(\left((x^0 - s^0) + (\boldsymbol{x} - \boldsymbol{s})'\tilde{\boldsymbol{z}}_\sigma\right)^+\right) + \sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left(\left(s^0 + \boldsymbol{s}'\tilde{\boldsymbol{z}}_\sigma\right)^+\right) \quad \text{(by (B.1) and (B.2))} \\
\ge\ & \sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left(\left(x^0 + \boldsymbol{x}'\tilde{\boldsymbol{z}}_\sigma\right)^+\right) \quad \text{(by subadditivity)} \\
=\ & \sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right).
\end{aligned}$$

Since the above inequality holds for *any* choice of $(s^0, \boldsymbol{s}, x^0, \boldsymbol{x})$ which satisfies $(x^0 + \boldsymbol{x}'\boldsymbol{g}_\sigma, \boldsymbol{F}'_\sigma \boldsymbol{x}) = (r^0, \boldsymbol{r})$, it also holds when we take the infimum, and hence $\pi^3\left(r^0, \boldsymbol{r}\right)$ bounds $\sup_{\mathbb{P}\in\mathbb{F}_3} \mathrm{E}_\mathbb{P}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ from above as required. ■

B.4 Proof of Theorem 4

B.4.1 Lemma: Positive Homogeneity of $\pi^s\left(r^0, \boldsymbol{r}\right)$

Lemma 5 *The bounding functions $\pi^s\left(r^0, \boldsymbol{r}\right)$ are positively homogeneous for each $s \in \{1, 2, 3\}$.*

Proof : The bounding functions $\pi^1\left(r^0, \boldsymbol{r}\right)$ and $\pi^2\left(r^0, \boldsymbol{r}\right)$ are easily seen to be positive homogeneous. We shall only explicitly prove the positive homogeneity of $\pi^3\left(r^0, \boldsymbol{r}\right)$. For any $\mu > 0$, we notice that

$$\psi(\mu x^0, \mu\boldsymbol{x}) = \inf_{\lambda>0}\left\{\frac{\lambda}{e}\exp\left(\frac{\mu}{\lambda}\sup_{\hat{\boldsymbol{z}}_\sigma\in\hat{\mathcal{W}}_\sigma}\left\{x^0 + \boldsymbol{x}'\hat{\boldsymbol{z}}_\sigma\right\} + \frac{\mu^2\left\|\boldsymbol{u}\right\|_2^2}{2\lambda^2}\right)\right\},$$

from the positive homogeneity of the supremum and norm operators. Re-expressing the minimization problem in terms of a new variable, $\nu = \frac{\lambda}{\mu}$,

$$\begin{aligned}
\psi(\mu x^0, \mu\boldsymbol{x}) &= \inf_{\nu>0}\left\{\frac{\mu\nu}{e}\exp\left(\frac{1}{\nu}\sup_{\hat{\boldsymbol{z}}_\sigma\in\hat{\mathcal{W}}_\sigma}\left\{x^0 + \boldsymbol{x}'\hat{\boldsymbol{z}}_\sigma\right\} + \frac{\left\|\boldsymbol{u}\right\|_2^2}{2\nu^2}\right)\right\} \\
&= \mu\psi(x^0, \boldsymbol{x}),
\end{aligned}$$

where the final equality comes from the positive homogeneity of the infimum operator. Now we consider

$$\pi^3\left(\mu r^0, \mu \boldsymbol{r}\right) = \inf_{\substack{s^0, \boldsymbol{s}, x^0, \boldsymbol{x} \\ x^0 + \boldsymbol{x}'\boldsymbol{g}_\sigma = \mu r^0 \\ \boldsymbol{F}_\sigma' \boldsymbol{x} = \mu \boldsymbol{r}}} \left\{ \begin{array}{lcl} (\mu r^0 - s^0 - \boldsymbol{s}'\boldsymbol{g}_\sigma) & + & \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} (\mu \boldsymbol{r}' - \boldsymbol{s}'\boldsymbol{F}_\sigma)\hat{\boldsymbol{\zeta}} \\ & + & \psi(s^0 - x^0, \boldsymbol{s} - \boldsymbol{x}) + \psi(s^0, \boldsymbol{s}) \end{array} \right\}$$

and, using the same idea as before, express the minimization problem in terms of new variables $(q^0, \boldsymbol{q}) = \left(\frac{s^0}{\mu}, \frac{\boldsymbol{s}}{\mu}\right)$, and $(w^0, \boldsymbol{w}) = \left(\frac{x^0}{\mu}, \frac{\boldsymbol{x}}{\mu}\right)$, we get

$$\pi^3\left(\mu r^0, \mu \boldsymbol{r}\right) = \inf_{\substack{q^0, \boldsymbol{q}, w^0, \boldsymbol{w} \\ w^0 + \boldsymbol{w}'\boldsymbol{g}_\sigma = r^0 \\ \boldsymbol{F}_\sigma' \boldsymbol{w} = \boldsymbol{r}}} \left\{ \begin{array}{lcl} \mu(r^0 - q^0 - \boldsymbol{q}'\boldsymbol{g}_\sigma) & + & \mu \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} (\boldsymbol{r}' - \boldsymbol{q}'\boldsymbol{F}_\sigma)\hat{\boldsymbol{\zeta}} \\ & + & \psi(\mu q^0 - \mu w^0, \mu \boldsymbol{q} - \mu \boldsymbol{w}) + \psi(\mu q^0, \mu \boldsymbol{q}) \end{array} \right\}.$$

Using the positive homogeneity of the infimum operator and $\psi(x^0, \boldsymbol{x})$ established earlier, we obtain $\pi^3\left(\mu r^0, \mu \boldsymbol{r}\right) = \mu \pi^3\left(r^0, \boldsymbol{r}\right) \ \forall \mu > 0$.

We consider the case of $\mu = 0$ separately. We first notice that the $\boldsymbol{H}_\sigma$ is full rank by definition, since it represents a mapping to an uncertainty vector $\boldsymbol{z}_\sigma$ with stochastically independent components. Hence $\boldsymbol{F}_\sigma = \boldsymbol{H}_\sigma \boldsymbol{F}$ is also full rank, and from the constraints, $(r^0, \boldsymbol{r}) = (0, \boldsymbol{0})$ implies $(x^0, \boldsymbol{x}) = (0, \boldsymbol{0})$. Simplifying, we have

$$\pi^3(0, \boldsymbol{0}) = \inf_{s^0, \boldsymbol{s}} \left\{ (s^0 - \boldsymbol{s}'\boldsymbol{g}_\sigma) + \sup_{\hat{\boldsymbol{\zeta}} \in \hat{\mathcal{V}}} (-\boldsymbol{s}'\boldsymbol{F}_\sigma)\hat{\boldsymbol{\zeta}} + 2\psi(s^0, \boldsymbol{s}) \right\}.$$

We know that $\pi^3(0, \boldsymbol{0}) \geq 0$ due to the upper bound property (Theorem 3). Furthermore, substituting the feasible $(s^0, \boldsymbol{s}) = (0, \boldsymbol{0})$ in the inner expression, and noticing that $\psi(0, \boldsymbol{0}) = 0$, we get $\pi^3(0, \boldsymbol{0}) \leq 0$. Thus $\pi^3(0, \boldsymbol{0}) = 0$, and $\pi^3\left(r^0, \boldsymbol{r}\right)$ is positive homogeneous. ■

B.4.2 Proof of Theorem 4

We begin by noticing that each $\pi^s\left(r^0, \boldsymbol{r}\right)$ is convex and positive homogeneous (Lemma 5) in its arguments. From positive homogeneity of each $\pi^s\left(r^0, \boldsymbol{r}\right)$, we have $\pi^s(0, \boldsymbol{0}) = 0$, which gives us the second inequality of (5.6). To establish that $\pi\left(r^0, \boldsymbol{r}\right)$ does indeed bound $\sup_{\mathbb{P} \in \mathbb{F}} \mathrm{E}_{\mathbb{P}}\left(\left(r_0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right)$ from above, we have for each $\mathbb{P} \in \bigcap_{s \in S} \mathbb{F}_s$

$$\begin{array}{rcll} \sum_{s \in S} \pi^s\left(r^{0,s}, \boldsymbol{r}^s\right) & \geq & \sum_{s \in S} \mathrm{E}_{\mathbb{P}}\left(\left(r^{0,s} + \boldsymbol{r}^{s\prime}\tilde{\boldsymbol{\zeta}}\right)^+\right) & (\text{using } \mathbb{P} \in \mathbb{F}_s) \\ & = & \mathrm{E}_{\mathbb{P}}\left(\sum_{s \in S}\left(r^{0,s} + \boldsymbol{r}^{s\prime}\tilde{\boldsymbol{\zeta}}\right)^+\right) & (\text{by linearity}) \\ & \geq & \mathrm{E}_{\mathbb{P}}\left(\left(\sum_{s \in S} r^{0,s} + \boldsymbol{r}^{s\prime}\tilde{\boldsymbol{\zeta}}\right)^+\right) & (\text{by subadditivity of } (\cdot)^+) \\ & = & \mathrm{E}_{\mathbb{P}}\left(\left(r^0 + \boldsymbol{r}'\tilde{\boldsymbol{\zeta}}\right)^+\right) & (\text{from (5.5)}). \end{array}$$

■

Appendix C Proofs of BDLDR Properties

C.1 Proof of Proposition 4

C.1.1 Lemma: Bounding a portion of the BDLDR

For clarity of exposition, we begin the proof with the following lemma:

Lemma 6 *For each $j \in [m]$, the following inequality holds:*

$$\underline{y}_j \le r_j(\tilde{\boldsymbol{\zeta}}) + \left(r_j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_j\right)^- \mathbb{1}_{\{j \in \underline{J}^\circ\}} - \left(r_j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_j\right)^+ \mathbb{1}_{\{j \in \overline{J}^\circ\}} \le \overline{y}_j \ .$$

Proof : We divide the proof into 4 cases:

Case 1: When $j \in [m] \setminus \left(\underline{J}^\circ \cup \overline{J}^\circ\right)$, the result holds directly from Equation (6.14).

Case 2: When $j \in \underline{J}^\circ \setminus \overline{J}^\circ$: To prove the upper bound, it suffices to consider the case when $j \in \overline{J}$. Furthermore, by assumption, $j \notin \overline{J}^\circ$, we can apply the linear constraints of Equation (6.14), $r_j(\tilde{\boldsymbol{\zeta}}) \le \overline{y}_j$. Together with the obvious $\underline{y}_j \le \overline{y}_j$, we get

$$r_j(\tilde{\boldsymbol{\zeta}}) + \left(r_j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_j\right)^- = \max\left\{r_j(\tilde{\boldsymbol{\zeta}}), \underline{y}_j\right\} \le \overline{y}_j \ .$$

Thus proving the upper bound. The lower bound follows directly from

$$\underline{y}_j \le \max\left\{r_j(\tilde{\boldsymbol{\zeta}}), \underline{y}_j\right\}.$$

Case 3: When $j \in \overline{J}^\circ \setminus \underline{J}^\circ$, the proof follows an identical argument to **Case 2**.

Case 4: When $j \in \overline{J}^\circ \cap \underline{J}^\circ$, we have

$$r_j(\tilde{\boldsymbol{\zeta}}) + \left(r_j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_j\right)^- - \left(r_j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_j\right)^+ = \max\left\{\min\left\{r_j(\tilde{\boldsymbol{\zeta}}), \overline{y}_j\right\}, \underline{y}_j\right\},$$

which directly satisfies both upper and lower bounds.

■

C.1.2 Proof of Proposition 4

We notice that statement *1* of the proposition follows directly from the feasibility of each $\bar{\boldsymbol{p}}^i$ and $\bar{\boldsymbol{q}}^i$ in (6.12) and (6.13). To prove statement *2* of the proposition, we consider the j^{th} component of the BDLDR, by considering the statement $\underline{y}_j \le \hat{r}_j(\tilde{\boldsymbol{\zeta}}) \le \overline{y}_j, \ \forall j \in [m]$. We notice that the BDLDR can be written in the following verbose form:

$$\begin{aligned}\hat{r}_j(\tilde{\boldsymbol{\zeta}}) = \Bigg(r_j(\tilde{\boldsymbol{\zeta}}) \quad &+ \quad \left(r_j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_j\right)^- \mathbb{1}_{\{j \in \underline{J}^\circ\}} \quad - \quad \left(r_j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_j\right)^+ \mathbb{1}_{\{j \in \overline{J}^\circ\}} \Bigg) \\ &+ \sum_{i \in \underline{J}^\circ \setminus \{j\}} \left(r_i(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i\right)^- \bar{p}^i_j \quad + \sum_{i \in \overline{J}^\circ \setminus \{j\}} \left(r_i(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i\right)^+ \bar{q}^i_j \ .\end{aligned}$$

To prove the upper bound, it suffices to consider $j \in \overline{J}$. We notice that we have explicitly removed j from both summation terms, so in both sums, $i \neq j$. Furthermore, since we only sum over indices i such that $\bar{\boldsymbol{p}}^i$ and $\bar{\boldsymbol{q}}^i$ are feasible in (6.12) and (6.13) respectively and we have established that $j \in \overline{J} \setminus \{i\}$, we have $\bar{p}^i_j \leq 0$ and $\bar{q}^i_j \leq 0$. Finally, the upper bound of Lemma 6 establishes the upper bound in the proposition statement. The lower bound can be proven with an identical argument. ■

C.2 Proof of Proposition 5

To show the second inequality, we begin by noting that each feasible solution of (6.19) is feasible in (6.22). We further note that $\forall i \in \underline{J},\ \pi^1(-r^0_i + \underline{y}_i, -\boldsymbol{R'e}^i) = 0$, (see Remark in Section 5.1). Similarly, $\forall i \in \overline{J},\ \pi^1(r^0_i - \overline{y}_i, \boldsymbol{R'e}^i) = 0$. Since $\underline{J}^\circ_{D,R} \subseteq \underline{J}$ and $\overline{J}^\circ_{D,R} \subseteq \overline{J}$, and using the property that $\pi(\cdot) \leq \pi^1(\cdot)$, we obtain $Z_{DLDR} \leq Z_{SLDR}$.

To prove the first inequality, we consider the sub-problems for the BDLDR, (6.12) and (6.13) against the corresponding sub-problems for the DLDR, (6.20) and (6.21). We notice that they are identical, with the sole exception that the BDLDR sub-problems (6.12, 6.13) have one less inequality constraint compared with the DLDR counterparts (6.20, 6.21). In particular, whenever $i \in \overline{J}$, the first DLDR sub-problem (6.20) is always infeasible. Conversely, if $i \notin \overline{J}$, both the BDLDR sub-problem (6.12) and DLDR sub-problem (6.20) are identical. A similar relation holds for the second sub-problem. This leads to the following set relations:

$$\begin{aligned} \underline{J}^\circ_D &= \underline{J}^\circ \setminus \overline{J}, \\ \overline{J}^\circ_D &= \overline{J}^\circ \setminus \underline{J}, \end{aligned} \tag{C.1}$$

and relations for the optimal solutions to the sub-problems:

$$\begin{aligned} \bar{\boldsymbol{p}}^i_{\boldsymbol{D}} &= \bar{\boldsymbol{p}}^i \quad \forall i \in \underline{J}^\circ_D, \\ \bar{\boldsymbol{q}}^i_{\boldsymbol{D}} &= \bar{\boldsymbol{q}}^i \quad \forall i \in \overline{J}^\circ_D. \end{aligned} \tag{C.2}$$

Together, these imply the set relations for the reduced index sets:

$$\begin{aligned} \underline{J}^\circ_{D,R} &= \underline{J}^\circ_R \setminus \overline{J}, \\ \overline{J}^\circ_{D,R} &= \overline{J}^\circ_R \setminus \underline{J}. \end{aligned} \tag{C.3}$$

Hence, using these relations, any feasible solution of (6.22) is feasible in (6.18). Using a similar argument to the DLDR vs SLDR above, we can relate the objectives by $Z_{BDLDR} \leq Z_{DLDR}$. ■

C.3 Proof of Proposition 6

From the BDLDR definition (6.29), it is obvious that the k^{th} BDLDR, $\boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}})$ has no dependency for any $\zeta_j,\ \forall j \in \Phi^c_k$. Hence, $\hat{\Phi}_k \subseteq \Phi_k$ follows directly. Now, assuming problem (4.12) is feasible, we denote a feasible set of SLDRs to the problem as $\{\boldsymbol{r}^k\}^K_{k=1}$. We notice that $\{\boldsymbol{r}^k\}^K_{k=1}$ lies within the feasible region of problem (6.28), and is an valid candidate to construct our DLDR. However, feasibility in problem (4.12) implies that the nonlinear terms in (6.29) vanish, giving us $\hat{\boldsymbol{r}}^k(\tilde{\boldsymbol{\zeta}}) = \boldsymbol{r}^k(\tilde{\boldsymbol{\zeta}})$. Hence it is necessary that their information index sets agree, i.e. $\hat{\Phi}_k = \Phi_k$. ■

C.4 Proof of Proposition 7

We consider:

$$\begin{aligned}
&\sum_{k=1}^{K} \sum_{j\in N^-(k)} \left(\sum_{i\in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^- \boldsymbol{U}^k \boldsymbol{p}^{i,j,k} + \sum_{i\in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^+ \boldsymbol{U}^k \boldsymbol{q}^{i,j,k} \right) \\
=\ & \sum_{k=1}^{K} \sum_{j=1}^{K} \left(\sum_{i\in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^- \boldsymbol{U}^k \boldsymbol{p}^{i,j,k} + \sum_{i\in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^+ \boldsymbol{U}^k \boldsymbol{q}^{i,j,k} \right) \mathbb{1}_{\{j\in N^-(k)\}} \\
=\ & \sum_{j=1}^{K} \sum_{k=1}^{K} \left(\sum_{i\in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^- \boldsymbol{U}^k \boldsymbol{p}^{i,j,k} + \sum_{i\in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^+ \boldsymbol{U}^k \boldsymbol{q}^{i,j,k} \right) \mathbb{1}_{\{k\in N^+(j)\}} \\
=\ & \sum_{j=1}^{K} \sum_{k\in N^+(j)} \left(\sum_{i\in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^- \boldsymbol{U}^k \boldsymbol{p}^{i,j,k} + \sum_{k\in N^+(j)} \sum_{i\in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^+ \boldsymbol{U}^k \boldsymbol{q}^{i,j,k} \right) \\
=\ & \sum_{j=1}^{K} \left(\sum_{k\in N^+(j)} \sum_{i\in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^- \boldsymbol{U}^k \boldsymbol{p}^{i,j,k} + \sum_{k\in N^+(j)} \sum_{i\in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^+ \boldsymbol{U}^k \boldsymbol{q}^{i,j,k} \right),
\end{aligned}$$

where we reverse the order of summation in the second equality, and use property (6.24). Considering the final expression, we note that the polyhedral regions $P(i,j)$ and $Q(i,j)$ are non-empty for $i \in \underline{J}_j^\circ$ and $i \in \overline{J}_j^\circ$ respectively, allowing us to apply the set of constraints (6.25) and (6.26). Applying the first constraint in each constraint set, we notice that the first summation vanishes for each $i \in \underline{J}_j^\circ$, and the second summation vanishes for each $i \in \overline{J}_j^\circ$. This causes the entire expression above to vanish, implying the result in statement *1* of the proposition. We prove statement *2* of the proposition by establishing the upper and lower bounds component-wise. We consider the n^{th} component of the k^{th} BDLDR, and rewrite it in the more verbose form:

$$\begin{aligned}
\hat{r}_n^k(\tilde{\boldsymbol{\zeta}}) &= r_n^k(\tilde{\boldsymbol{\zeta}}) + \sum_{j\in N^-(k)} \sum_{i\in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^- p_n^{i,j,k} + \sum_{j\in N^-(k)} \sum_{i\in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^+ q_n^{i,j,k} \\
&= r_n^k(\tilde{\boldsymbol{\zeta}}) + \sum_{i\in \underline{J}_k^\circ} \left(r_i^k(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^k\right)^- p_n^{i,k,k} + \sum_{j\in N^-(k)\setminus\{k\}} \sum_{i\in \underline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^- p_n^{i,j,k} \\
&\quad + \sum_{i\in \overline{J}_k^\circ} \left(r_i^k(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^k\right)^+ q_n^{i,k,k} + \sum_{j\in N^-(k)\setminus\{k\}} \sum_{i\in \overline{J}_j^\circ} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^+ q_n^{i,j,k}.
\end{aligned}$$

And extracting the $i = n$ term from the first and third sums, we get the final expression:

$$\begin{aligned}
\hat{r}_n^k(\tilde{\boldsymbol{\zeta}}) \;=\; r_n^k(\tilde{\boldsymbol{\zeta}}) \;&+\; \left(r_n^k(\tilde{\boldsymbol{\zeta}}) - \underline{y}_n^k\right)^{-} \mathbb{1}_{\{n \in \underline{J}_k^{\circ}\}} \;-\; \left(r_n^k(\tilde{\boldsymbol{\zeta}}) - \overline{y}_n^k\right)^{+} \mathbb{1}_{\{n \in \overline{J}_k^{\circ}\}} \\
&+ \underbrace{\sum_{i \in \underline{J}_k^{\circ} \setminus \{n\}} \left(r_i^k(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^k\right)^{-} p_n^{i,k,k}}_{(A)} + \underbrace{\sum_{j \in N^-(k) \setminus \{k\}} \sum_{i \in \underline{J}_j^{\circ}} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \underline{y}_i^j\right)^{-} p_n^{i,j,k}}_{(B)} \\
&+ \underbrace{\sum_{i \in \overline{J}_k^{\circ} \setminus \{n\}} \left(r_i^k(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^k\right)^{+} q_n^{i,k,k}}_{(C)} + \underbrace{\sum_{j \in N^-(k) \setminus \{k\}} \sum_{i \in \overline{J}_j^{\circ}} \left(r_i^j(\tilde{\boldsymbol{\zeta}}) - \overline{y}_i^j\right)^{+} q_n^{i,j,k}}_{(D)} .
\end{aligned}$$

To prove the upper bound of statement *2* of the proposition, it suffices to consider $n \in \overline{J}_k$. Again, in each of the four sums in the expression above, we sum over indices i which correspond to feasible instances of constraint sets (6.25) and (6.26), and hence we can apply these constraints. We consider the sums (A) – (D) in turn. For (A), we notice that since $i \neq n$, using the third inequality of (6.25), $p_n^{i,k,k} \leq 0$. For (B), we notice that similar to (6.24), we have

$$j \in N^-(k) \setminus \{k\} \Leftrightarrow k \in N^+(j) \setminus \{j\} .$$

Hence, using the second inequality of (6.25), $p_n^{i,j,k} \leq 0$ in (B). Also, using the first inequality of (6.26), $q_n^{i,k,k} \leq 0$ in (C) and $q_n^{i,j,k} \leq 0$ in (D). The upper bound follows directly using Lemma 6. The lower bound can be proven using an identical argument. ∎

C.5 Proof of Proposition 8

Any SLDR solution to Problem (2.2) will take the form of Problem (4.7), using the support sets $\mathcal{V}$ and $\hat{\mathcal{V}}$ to approximate the exact supports $\mathcal{V}^*$ and $\hat{\mathcal{V}}^*$. We begin by noting that $\forall j \in [K]$, $i \in \underline{J}_j$ implies $\pi^1\left(-r_i^{0,j} + \underline{y}_i, -\boldsymbol{R}^{j\prime}\boldsymbol{e}^i\right) = 0$, and similarly $i \in \overline{J}_j$ implies $\pi^1\left(r_i^{0,j} - \overline{y}_i, \boldsymbol{R}^{j\prime}\boldsymbol{e}^i\right) = 0$. Now, in each of the two summation terms (in the objective and constraints), $j \in N^-(k) \subseteq [K]$, and $i \in \underline{J}_{l,j,k}^{\circ} \subseteq \underline{J}_j$ (in the first sum) or $i \in \overline{J}_{l,j,k}^{\circ} \subseteq \overline{J}_j$ (in the second sum), for some $l \in \{0\} \cup [M]$. We further note that $\pi(\cdot) \leq \pi^1(\cdot)$. Hence, any feasible solution of the approximated Problem (4.7) is feasible in (6.30), and since their objectives coincide, the objectives are related by $Z_{BDLDR} \leq Z_{SLDR}$ as desired. ∎