Computational Management Science manuscript No.
(will be inserted by the editor)

Distributionally Robust SDDP

A.B. Philpott[1], V.L. de Matos[2], L. Kapelevich[1]

[1] The University of Auckland e-mail: a.philpott@auckland.ac.nz
[2] Plan4 Engenharia e-mail: vitor@plan4.com.br
[3] The University of Auckland e-mail: lkap311@aucklanduni.ac.nz

The date of receipt and acceptance will be inserted by the editor

This paper is dedicated to the memory of our friend and colleague Maarten van der Vlerk.

Abstract We study a version of stochastic dual dynamic programming (SDDP) with a distributionally robust objective. The modifications to SDDP are described, and the algorithm is illustrated by applying it to the New Zealand hydrothermal electricity system.

Key words SDDP, Distributionally robust, Hydroelectric reservoir optimization

1 Introduction

Stochastic Dual Dynamic Programming (SDDP) has been widely used to build policies for multistage stochastic problems in many practical problems, with a historical focus on problems related to energy and hydrothermal scheduling. When SDDP was first introduced by [16] the objective was to build a policy to optimize the expected value of a multistage linear stochastic problem. In recent years several contributions have led to the discussion of non-convex and risk averse cases. In this paper we are particularly interested in models of risk aversion proposed within the SDDP framework.

The relevance of analyzing such modeling in the SDDP algorithm comes from the fact that there are several conditions that need to be met in order to be able to build a valid outer approximation for the Bellman function defining an optimal policy [13]. In recent years the incopration into SDDP of coherent risk measures [1] such as Conditional Value-at-Risk (CVaR) has received lot of attention (see [21, 17, 18, 15]). The Average Value-at-Risk of a random cost can be viewed (in a minimization problem) as an expectation that assigns positive probabilities to only the most expensive outcomes. To be able to compute a policy, in practice these outcomes are the finite sample space of an approximating problem (such as a sample average approximation [22]) endowed with a finite probability distribution.

In this paper we consider the problem of distributionally robust optimization (DRO) in SDDP. This seeks a policy that minimizes expected cost over the worst-case probability distribution in some family of distributions. The appeal of such a model comes from the fact that in practical applications the real probability distribution is not known and in most cases we consider the historical data to build an approximation which will then be used to sample possible realizations that are going to be used in the SDDP valuation process. We might expect a robust approach to avoid overfitting the policy to a single distribution that has been estimated from data, which is particularly important when there are only a small number of outcomes per stage in the model.

Robust optimization has received a lot of attention in recent years. A good summary is provided in the review article [3]. Our interest in this paper is confined to distributionally robust models in probability spaces with a finite number of random outcomes. We assume that these outcomes are fixed and model distributional ambiguity by allowing changes in the probability measure on these outcomes, as long as the new probability measure is close to a given nominal measure. A broad class of such distance measures are the so-called φ-divergence distances [2]. A different strand of research uses uncertainty sets based on probability metrics such as the Wasserstein distance [11],[10]. As observed by [11], these have advantages over φ-divergence distances in that the probability measure is not confined to a given set of points, but computational methods are more complicated.

Our purpose in this paper is to investigate the effect of distributional robustness on policies computed by SDDP. Since standard SDDP implementations on real models take several hours to converge, our choice of uncertainty set is dictated to some extent by computational convenience, so we seek a method that is distributionally robust without requiring a big increase in computational effort. Our choice of distance is a φ-divergence based on the Euclidean norm on the difference in probabilities which gives rise to a so-called *modified* χ^2 distance (see [2]). A version of this (the χ^2 distance) was used in an inventory lot-sizing setting by [14], and has been discussed in general stochastic optimization by [4] and [5].

The χ^2 distance has the following goodness-of-fit interpretation. Given a sample of historical data of size N, represented for example frequency n_i in bin i of a histogram, and a probability distribution p_i for the same histogram bins, one can test the hypothesis that the sample was obtained from the probability distribution using the statistic $\sum_i \frac{(n_i - Np_i)^2}{Np_i}$ that has a χ^2 distribution. Thus given a sample, an uncertainty set $\mathcal{P}$ of probabilities p that would not be rejected under a goodness of fit at some confidence level takes the form $\mathcal{P} = \{p : \sum_i p_i = 1, p_i \geq 0, \sum_i \frac{(n_i - Np_i)^2}{Np_i} \leq r^2\}$. As shown by [14] this leads to a problem with second-order cone constraints defining $\mathcal{P}$.

In the current paper we adopt a slightly different approach to that of [14]. We assume that given a set of m historical data points, the nominal distribution assigns equal probability $\frac{1}{m}$ to these. For a sample of N observations we would expect $\frac{N}{m}$ observations for each point. We then seek a set of possible sample frequencies n_i, $i = 1, \ldots, m$ for these outcomes that would be such that we would not reject the null hypothesis that each outcome had probability $\frac{1}{m}$. Expressing these frequencies in terms of ratios $p_i = \frac{n_i}{N}$ gives an uncertainty set

$$\mathcal{P} = \{p : \sum_i p_i = 1, p_i \geq 0, \sum_i (p_i - \frac{1}{m})^2 \leq \frac{r^2}{mN}\}.$$

This is a modified χ^2 distance. As we show below this choice of $\mathcal{P}$ leads to a solution with a closed form that enables fast optimization of stage problems.

In this paper we make the following contributions:

1. We derive a distributionally robust SDDP algorithm, and show that it converges almost surely to an optimal policy;
2. We implement the SDDP algorithm in the Julia language[6] and the modeling package JuMP[9], and demonstrate its out-of-sample performance on a hydrothermal planning problem in New Zealand;

The paper is laid out as follows. We first describe a multistage distributionally robust optimization model in a finite probability space. We then look at a specific example where the uncertainty set takes the form of a unit simplex intersected with a ball centred on the probability measure with equal weights. By varying the radius of the ball we can construct increasingly conservative robust optimization problems. The inner maximization for these problems can be computed by a simple algorithm. In section 4 we show how this can be imbedded in SDDP to give a distributionally robust version of this algorithm. Section 5 applies this algorithm to some hydrothermal scheduling models from New Zealand to illustrate the effect of increasing conservatism on water release policies. The paper concludes with a discussion of the results of these experiments.

2 Multistage distributionally robust optimization

The type of problem we consider has T stages, denoted $t = 1, 2, \ldots, T$, in each of which a random right-hand-side vector $b_t(\omega_t) \in \mathbb{R}^m$ has a finite number of realizations defined by $\omega_t \in \Omega_t$. We assume that the outcomes ω_t are stagewise independent, and that Ω_1 is a singleton (i.e. we know ω_1). For $t > 1$, the probability of each outcome ω_t is not known exactly, but lies in some convex set $\mathcal{P}_{t-1}$ of probability distributions.

The assumption of finite probability spaces greatly simplifies the analysis, whereby we can dispense with most measurability assumptions, such as, for example, specifying constraints that hold almost surely. If we let $\Omega =$

$\times_{t=1}^{T} \Omega_t$ then the evolution of $b_t(\omega_t)$ defines a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\{\emptyset, \Omega\} = \mathcal{F}_1 \subset \mathcal{F}_2 \ldots \subset \mathcal{F}_T \subset \mathcal{F}$ of σ-fields where b_1 is assumed to be deterministic. The decision variables x_t, $t = 1, 2, \ldots, T$, are constrained to be non-negative $\mathcal{F}_t$-measurable random variables that obey the linear dynamics

$$A_t x_t = b_t(\omega_t) - H_t x_{t-1},$$

where for simplicity we assume that A_t and H_t are deterministic $m \times n$ matrices. The objective function to be minimized is

$$c_1^\top x_1 + \max_{\mathbb{P} \in \mathcal{P}_1} \mathbb{E}_{\mathbb{P}}[c_2^\top x_2 + \max_{\mathbb{P} \in \mathcal{P}_2} \mathbb{E}_{\mathbb{P}}[c_3^\top x_3 + \ldots + \max_{\mathbb{P} \in \mathcal{P}_{T-1}} \mathbb{E}_{\mathbb{P}}[c_T^\top x_T] \ldots]]$$

where $c_t \in \mathbb{R}^n$ is a cost vector. In our setting, this construction leads us to a recursive form for the dynamic programming problem to be solved. The first-stage problem is

$$\begin{aligned} z = \min\ & c_1^\top x_1 + \max\nolimits_{\mathbb{P} \in \mathcal{P}_1} \mathbb{E}_{\mathbb{P}}[Q_2(x_1, \omega_2)] \\ \text{s.t. } & A_1 x_1 = b_1, \\ & x_1 \geq 0, \end{aligned} \tag{1}$$

where for $t = 2, 3, \ldots, T$,

$$\begin{aligned} Q_t(x_{t-1}, \omega_t) = \min\ & c_t^\top x_t + \max\nolimits_{\mathbb{P} \in \mathcal{P}_t} \mathbb{E}_{\mathbb{P}}[Q_{t+1}(x_t, \omega_{t+1})] \\ \text{s.t. } & A_t x_t = b_t(\omega_t) - H_t x_{t-1}, \\ & x_t \geq 0, \end{aligned} \tag{2}$$

and in the last stage we assume for simplicity that $Q_{T+1}(x_T, \omega_{T+1}) = 0$. (The approach can be easily modified to make use of a known (convex) polyhedral function that defines $Q_{T+1}(x_T, \omega_{T+1})$.)

Observe that the convexity of $\mathcal{P}_t$ implies that $\max_{\mathbb{P} \in \mathcal{P}_t} \mathbb{E}_{\mathbb{P}}[Q_{t+1}(x_t, \omega_{t+1})]$ is a coherent risk measure, so it is monotonic and convex. It follows that $\max_{\mathbb{P} \in \mathcal{P}_t} \mathbb{E}_{\mathbb{P}}[Q_{t+1}(x_t, \omega_{t+1})]$ is a convex function of x_t whenever $Q_{t+1}(x_t, \omega_{t+1})$ is convex in x_t for every ω_{t+1}. This means that $Q_t(x_{t-1}, \omega_t)$ is convex in x_{t-1} for every ω_t whenever $Q_{t+1}(x_t, \omega_{t+1})$ is convex in x_t for every ω_{t+1}, and so it follows by induction that for every $t = 2, 3, \ldots, T$, $Q_t(x_{t-1}, \omega_t)$ is convex in x_{t-1} for every ω_t.

Our goal is to construct an approximately optimal solution for the multistage problem defined by (1) and (2). We define

$$\mathcal{X}_1(\omega_1) = \{x_1 \geq 0 : A_1 x_1 = b_1\}$$

and for $t = 2, 3, \ldots T$, recursively we let

$$\mathcal{X}_t(\omega_t) = \{x_t \geq 0 : A_t x_t = b_t(\omega_t) - H_t x_{t-1}, \quad x_{t-1} \in \mathcal{X}_{t-1}(\omega_{t-1})\}.$$

Under the assumption that the random disturbances ω_t are stagewise independent, the solution has the form of a *policy* defined for each stage t by a mapping $\boldsymbol{\pi}$ from $\mathcal{X}_{t-1}(\omega_{t-1}) \times \Omega_t$ to $\mathcal{X}_t(\omega_t)$, specifying the decision $x_t(x_{t-1}, \omega_t)$ made by the policy at time t.

3 Some preliminary results

In our distributionally robust version of SDDP we need to solve a subproblem of the form

$$\max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[Z(x,\omega)]$$

where $Z(x,\omega)$ is a cost. The choice of $\mathcal{P}$ can be made in different ways. We assume that ω takes a finite number of values ω_i, $i = 1, 2, \ldots, m$, each with a nominal probability q_i. In the special case where the outcomes are obtained by sampling, for example in a sample average approximation, we have $q_i = \frac{1}{m}$. We now define

$$\mathcal{P} = \{p \in \mathbb{R}^m \mid \sum_{i=1}^{m} p_i = 1,\ p \geq 0,\ \ \|p - q\|_2 \leq r\}.$$

This results in a subproblem of the form

$$\begin{aligned} \text{P:}\ \max\ & \textstyle\sum_{i=1}^{m} z_i p_i \\ \text{s.t.}\ & \textstyle\sum_{i=1}^{m} p_i = 1, \\ & \|p - q\|_2 \leq r, \\ & p \geq 0. \end{aligned}$$

where $z_i = Z(x,\omega_i)$ and we assume without loss of generality that $z_1 \leq z_2 \leq \ldots \leq z_m$. We use the notation

$$\bar{z} = \frac{\sum_{i=1}^{m} z_i}{m}$$

and

$$s = \sqrt{\frac{\sum_{i=1}^{m} (z_i - \bar{z})^2}{m}},$$

and given z_i, $i = 1, 2, \ldots, m$, and r, we denote the optimal value of P by $\rho_r(z)$.

Lemma 1 *Given z_i, $i = 1, 2, \ldots, m$, $\rho_r(z)$ is a continuous and nondecreasing function of $r \geq 0$.*

Proof See Appendix 1.

We will make use of a more general formulation of problem P. This is

$$\begin{aligned} \text{P}(n)\text{:} \min\ & -\textstyle\sum_{i=1}^{n} z_i y_i \\ \text{s.t.}\ & \textstyle\sum_{i=1}^{n} y_i = a, \\ & \textstyle\sum_{i=1}^{n} y_i^2 \leq b^2. \end{aligned}$$

Observe that P is an instance of P(n) in which we set $n = m$, $a = 0$, and $b = r$, and change variables by setting $p = q + y$.

Lemma 2 *P(n) has an optimal solution if and only if $a^2 \leq nb^2$.*

Proof See Appendix 1.

Lemma 3 *Suppose $a^2 \le nb^2$. The optimal solution to P(n) is*

$$y_i = \frac{a}{n} + \sqrt{nb^2 - a^2}\frac{z_i - \bar{z}}{ns}.$$

Proof See Appendix 1.

Lemma 4 *If $r \le \sqrt{\frac{m}{(m-1)}}\min\{q_i\}$ then P has optimal solution*

$$p_i = q_i + \frac{z_i - \bar{z}}{\sqrt{m}}\frac{r}{s}$$

and solution value

$$\sum_{i=1}^{m} q_i z_i + \left(\sqrt{m}s\right) r.$$

Proof See Appendix 1.

We now consider the case where $r > \sqrt{\frac{m}{(m-1)}}\min\{q_i\}$. Suppose $z_1 < z_2 < \ldots < z_m$. We are no longer guaranteed that $p_i = q_i + \frac{z_i - \bar{z}}{\sqrt{m}}\frac{r}{s} \ge 0$ for every i. Consider a candidate solution that identifies an index set K for which we set $p_i = 0$, $i \notin K$, and solve P for the remaining p_i, which will solve

$$\begin{aligned} \min\ & -\textstyle\sum_{i\in K} z_i p_i \\ \text{s.t.}\ & \textstyle\sum_{i\in K} p_i = 1, \\ & \textstyle\sum_{i\notin K} \left(-q_i\right)^2 + \sum_{i\in K} \left(p_i - q_i\right)^2 \le r^2 \end{aligned}$$

or

$$\begin{aligned} \min\ & -\textstyle\sum_{i\in K} z_i p_i \\ \text{s.t.}\ & \textstyle\sum_{i\in K} p_i = 1, \\ & \textstyle\sum_{i\in K} \left(p_i - q_i\right)^2 \le r^2 - \sum_{i\notin K} q_i^2. \end{aligned}$$

For $i \in K$ we define $y_i = p_i - q_i$, which yields

$$\begin{aligned} \min\ & -\textstyle\sum_{i\in K}(q_i + y_i) z_i \\ \text{s.t.}\ & \textstyle\sum_{i\in K}(q_i + y_i) = 1, \\ & \textstyle\sum_{i\in K} y_i^2 \le r^2 - \sum_{i\notin K} q_i^2. \end{aligned}$$

To compute y we solve

$$\begin{aligned} \text{Q:} \min\ & -\textstyle\sum_{i\in K} z_i y_i \\ \text{s.t.}\ & \textstyle\sum_{i\in K} y_i = \sum_{i\notin K} q_i, \\ & \textstyle\sum_{i\in K} y_i^2 \le r^2 - \sum_{i\notin K} q_i^2 \end{aligned}$$

and then add y_i to q_i, for $i \in K$.

Observe that problem Q is of the form of P(n) where $n = |K|$, $a = \sum_{i\notin K} q_i$, and $b^2 = r^2 - \sum_{i\notin K} q_i^2$. By Lemma 2 P($n$) has a solution if and only if $nb^2 - a^2 \ge 0$. Thus Q has a solution if and only if

$$r^2 \geq \sum_{i \notin K} q_i^2 + \frac{1}{|K|} \left(\sum_{i \notin K} q_i \right)^2 .$$

The optimal solution to Q when $K \subset \{1, 2, \ldots, m\}$ is

$$\begin{aligned} y_i &= \frac{a}{n} + \sqrt{nb^2 - a^2} \frac{z_i - \bar{z}}{ns} \\ &= \frac{\sum_{j \notin K} q_j}{|K|} + \sqrt{|K| \, (r^2 - \sum_{j \notin K} q_j^2) - \left(\sum_{j \notin K} q_j \right)^2} \frac{z_i - \bar{z}}{|K| \, s} \end{aligned}$$

whence

$$p_i = q_i + \frac{1}{|K|} \left(\sum_{j \notin K} q_j + \sqrt{|K| \, (r^2 - \sum_{j \notin K} q_j^2) - \left(\sum_{j \notin K} q_j \right)^2} \frac{z_i - \bar{z}}{s} \right) .$$

Observe that $\bar{z}$ and s must now be computed using a smaller set of data. In other words

$$\bar{z} = \frac{1}{|K|} \sum_{i \in K} z_i$$

and

$$s = \sqrt{\frac{1}{|K|} \sum_{i \in K} z_i^2 - \bar{z}^2}.$$

In the special case where r is large enough so that we must set $K = \{m\}$, then we obtain

$$\begin{aligned} \text{Q:} \; & \min \; -z_m y_m \\ \text{s.t.} \; & \; y_m = 1 - q_m, \\ & \; y_m^2 \leq r^2 - \textstyle\sum_{i \notin K} q_i^2 \end{aligned}$$

so $y_m = 1 - q_m$, and $p_m = y_m + q_m = 1$. In this case we choose probabilities equal to the worst-case measure. So by varying r, the solution to P can vary from expectation when $r = 0$ to worst case when r is large enough so that the unit simplex is a subset of the ball

$$\{p : \|p - q\|_2 \leq r\}.$$

Now recall

$$\begin{aligned} \text{P:} \; & \max \; \textstyle\sum_{i=1}^m z_i p_i \\ \text{s.t.} \; & \textstyle\sum_{i=1}^m p_i = 1, \\ & \|p - q\|_2 \leq r, \\ & p \geq 0. \end{aligned}$$

We propose the following algorithm for computing the solution to P.

Algorithm 1: Solving P

1. Set $K = \{1, 2, \ldots, m\}$
2. While $|K| > 1$
   (a)
$$\bar{z} = \frac{1}{|K|} \sum_{i \in K} z_i$$
   and
$$s = \sqrt{\frac{1}{|K|} \sum_{i \in K} z_i^2 - \bar{z}^2}.$$
   (b) If $k = m$ then let
$$p_i = q_i + \frac{z_i - \bar{z}}{\sqrt{m} s} r$$
   else let
$$p_i = \begin{cases} 0 & i \notin K, \\ q_i + \frac{1}{|K|}\left(\sum_{j \notin K} q_j + \sqrt{|K|\,(r^2 - \sum_{j \notin K} q_j^2) - \left(\sum_{j \notin K} q_j\right)^2} \frac{z_i - \bar{z}}{s} \right) & i \in K. \end{cases}$$
   (c) If $p_i \geq 0$, $i \in K$, then STOP and return p as the optimal solution.
   (d) Find critical $j \in K$. This is the last index of $p_i < 0$ to become positive as we decrease r. This can be found by a line search or by analysis of the formula for p_i. Set $K = K \setminus \{j\}$.
3. Return
$$p_i = \begin{cases} 0 \; i \notin K, \\ 1 \; i \in K. \end{cases}.$$

3.1 Equal nominal probabilities

The algorithm for computing the solution to P takes a simpler form when $q_i = \frac{1}{m}$. Then the set K can be shown to take the form $\{k+1, k+2, \ldots, m\}$. This gives

Algorithm 2: Solving P (equal probabilities)

1. For $k = 0$ to $m - 2$ do
   (a) Compute
$$\bar{z} = \frac{1}{(m-k)} \sum_{i=k+1}^{m} z_i$$
   and
$$s = \sqrt{\frac{1}{(m-k)} \sum_{i=k+1}^{m} z_i^2 - \bar{z}^2}.$$

(b) Compute

$$p_i = \begin{cases} 0 & i = 1, \ldots, k, \\ \frac{1}{(m-k)} + \frac{\sqrt{(m-k)r^2 - \frac{k}{m}} \frac{z_i - \bar{z}}{s}}{(m-k)} & i = k+1, \ldots, m. \end{cases} \tag{3}$$

(c) If $p_{k+1} \geq 0$ then STOP and return p as the optimal solution.

2. Return

$$p_i = \begin{cases} 0 \; i = 1, \ldots, m-1, \\ 1 \; i = m. \end{cases}.$$

The value of k that is computed by Algorithm 2 is called the *threshold* value of k for r, denoted $k(r)$. If $k(r) = 0$ then all probabilities in $\mathcal{P}$ are positive. If $k(r) = m - 1$ then $\mathcal{P}$ corresponds to a worst-case risk measure.

Recall $\rho_r(\cdot)$ to be the risk measure computed using a distributional uncertainty set with radius r. Depending on the value of $k(r)$, we get

$$\rho_r(Z(x)) = \begin{cases} \bar{z} + s\sqrt{(m-k(r))r^2 - \frac{k(r)}{m}} & k(r) < m-1, \\ z_m & k(r) = m-1. \end{cases} \tag{4}$$

Lemma 5 *If $r_1 < r_2$ then $k(r_1) \leq k(r_2)$.*

Proof See Appendix 1.

4 Using P in SDDP

We now show how the problem P can be incorporated into SDDP to give a distributionally robust version of this algorithm. The SDDP algorithm performs a sequence of major iterations known as the *forward pass* and the *backward pass* to build an outer approximation of the Bellman function at each stage. This approximation defines a policy in which the action at each stage solves a problem of the form (2) with $\max_{\mathbb{P}\in\mathcal{P}_t} \mathbb{E}_{\mathbb{P}}[Q_{t+1}(x_t, \omega_{t+1})]$ replaced by its approximation. In each forward pass, a single scenario is sampled from the scenario tree and decisions are taken according to the approximate policy, starting in the first stage and moving forward up to the last stage. In each stage, the observed values of the decision variables x_t, and the costs of each node are saved. The backward pass improves the outer approximation of the Bellman function at each stage by adding a single cutting plane computed using information from the optimal decision variables.

To obtain the cut coefficients, we use the following proposition which is a special case of [22, Theorem 6.11].

Proposition 1 *Suppose that $Z(x, \omega)$ is a convex function of x for each $\omega \in \Omega$, and that $g(\tilde{x}, \omega)$ is a subgradient of $Z(x, \omega)$ at $\tilde{x}$. Then $\mathbb{E}_{\mathbb{P}^*}[g(\tilde{x}, \omega)]$ is a subgradient of* $\max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[Z(x, \omega)]$ *at $\tilde{x}$, where* $\mathbb{P}^* \in \arg\max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}[Z(\tilde{x}, \omega_m)]$.

Proof See Appendix 1.

The approximation at stage t replaces $\max_{\mathbb{P}\in\mathcal{P}_t} \mathbb{E}_{\mathbb{P}}[Q_{t+1}(x_t, \omega_{t+1})]$ by the variable θ_{t+1} which is constrained by the set of linear inequalities

$$\theta_{t+1} + \bar{\pi}_{t+1,k}^\top H_{t+1} x_t \geq h_{t+1,k} \qquad \text{for } k = 1, 2, ..., \nu, \tag{5}$$

where ν is the number of cuts. Given $\mathbb{P}_t^* \in \arg\max_{\mathbb{P}\in\mathcal{P}_t} \mathbb{E}_{\mathbb{P}}[Q_{t+1}(x_t, \omega_{t+1})]$, we set $\bar{\pi}_{t+1,k} = \mathbb{E}_{\mathbb{P}_t^*}[\pi_{t+1}(\omega_{t+1})]$, which defines the gradient $-\bar{\pi}_{t+1,k}^\top H_{t+1}$ and the intercept $h_{t+1,k}$ for cut k in stage t, where

$$h_{t+1,k} = \mathbb{E}_{\mathbb{P}_t^*}[\tilde{Q}_{t+1}(x_t^k, \omega_{t+1})] + \bar{\pi}_{t+1,k}^\top H_{t+1} x_t^k,$$

and we define $\tilde{Q}_t$ and $\pi_t(\omega_t)$ (the Lagrange multipliers of the constraints) by the approximate stage problem

$$\begin{aligned} \tilde{Q}_t(x_{t-1}, \omega_t) = \min\ & c_t^\top x_t + \theta_{t+1} \\ \text{s.t. } & A_t x_t = b_t(\omega_t) - H_t x_{t-1}, \qquad [\pi_t(\omega_t)] \\ & \theta_{t+1} + \bar{\pi}_{t+1,k}^\top H_{t+1} x_t \geq h_{t+1,k},\ k = 1, 2, \ldots, \nu, \\ & x_t \geq 0. \end{aligned} \tag{6}$$

Thus if we denote $\Omega_{t+1} = \{\omega_{t+1}^1, \omega_{t+1}^2, \ldots, \omega_{t+1}^m\}$ and $\mathbb{P}_t^*(\omega_{t+1}^i) = p_i$, then the cut parameters are defined by

$$\begin{aligned} \bar{\pi}_{t+1,k} &= \sum_{i=1}^m p_i \pi_{t+1,k}(\omega_{t+1,i}), \\ h_{t+1,k} &= \sum_{i=1}^m p_i \tilde{Q}_{t+1}(x_t^k, \omega_{t+1,i}) + \bar{\pi}_{t+1,k}^\top H_{t+1} x_t^k. \end{aligned} \tag{7}$$

The Distributionally Robust SDDP algorithm can now be defined as follows.

Algorithm 3: Distributionally robust SDDP

1. Set $\nu = 0$.
2. Sample a scenario $\omega_t, t = 2, ..., T$;
3. Forward Pass
   For $t = 1$, solve (6) where and save $x_1(\nu)$ and z;
   For $t = 2, ..., T$,
      Solve (6), and save $x_t(\nu)$ and $\tilde{Q}_t(x_{t-1}(\nu), \omega_t)$.
4. Backward Pass
   For $t = T, ..., 2$,
      For $\omega_{t,i} \in \Omega_t$, solve (6) using $x_{t-1}(\nu)$ and save $\pi_t(\omega_{t,i})$ and $z_i = \tilde{Q}_t(x_{t-1}(\nu), \omega_{t,i})$, $i = 1, 2, \ldots, m$.
      Apply Algorithm 2 to compute p_i, $i = 1, 2, \ldots, m$.
      Calculate a cut using 7 with probabilities p for iteration ν, and add it to all nodes in stage
   Set $\nu = \nu + 1$.

5. If $\nu < \nu^{\max}$, go to step 2. Otherwise, stop.

In order to compute a cut it is necessary to add a step in the backward pass which calculates the worst-case probabilities p when the outcomes for the stage problem are z_i, $i = 1, 2, \ldots, m$. This is done using Algorithm 2. The cut computation then proceeds using the probabilities p. Since the probabilities p change from iteration to iteration, we need to verfify that the cuts computed do not violate the outer-approximation property. Since Algorithm 3 is essentially identical to the SDDP algorithm with coherent risk measures described in [18] we can reiterate the argument from [18]. This proceeds as follows.

In any backwards pass, we begin with an exact Bellman function $Q_{T+1}(x_T, \omega_{T+1})$ that is a convex (trivial) outer approximation to the future cost. If we denote

$$\mathcal{X}_t(\omega_t) = \{x_t \geq 0 : A_t x_t = b_t(\omega_t) - H_t x_{t-1}, \quad x_{t-1} \in \mathcal{X}_{t-1}(\omega_{t-1})\}.$$

then by construction for any $x_T \in \mathcal{X}_T(\omega_T)$, and every $k = 1, 2, \ldots, \nu$,

$$h_{T+1,k} - \bar{\pi}_{T+1,k}^\top H_{T+1} x_T \leq \max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[Q_{T+1}(x_T, \omega_{T+1})]$$

so $\max_{k=1,2,\ldots,\nu}\{h_{T+1,k} - \bar{\pi}_{T+1,k}^\top H_{T+1} x_T\}$ is a convex outer approximation to $\max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[Q_{T+1}(x_T, \omega_{T+1})]$. It follows that $\tilde{Q}_T(x_{T-1}, \omega_T)$ defined by (6) is a convex outer approximation to $Q_T(x_{T-1}, \omega_T)$ defined by (2). Extending this argument to every t we have

Proposition 2 *If for any* $x_t \in \mathcal{X}_t(\omega_t)$, $h_{t+1,k} - \bar{\pi}_{t+1,k}^\top H_{t+1} x_t \leq \mathbb{E}_{\mathbb{P}_t^*}[Q_{t+1}(x_t, \omega_{t+1})]$ *for every* $k = 1, 2, \ldots, \nu$, *then*

$$\tilde{Q}_t(x_{t-1}, \omega_t) \leq Q_t(x_{t-1}, \omega_t).$$

Proof See Appendix 1.

By Proposition 2, the convex outer approximation property of $\tilde{Q}_t$ is inherited in every step of the backwards pass, and so it is maintained throughout every iteration of the algorithm.

5 Computational results

We now present some computational results of applying SDDP to hydrothermal scheduling problems in New Zealand, with and without distributional robustness.

The aim in a hydrothermal scheduling problem is to construct a policy for managing water levels in hydroelectric reservoirs, subject to uncertain inflows. We consider a model of the New Zealand national electricity system, described in [19]. In this model, the state of the system is defined by the reservoir levels of seven lakes, which supply water to 25 hydro generators in various locations. The releases and spills from each lake can be viewed as decision variables. Any electricity demand that is not satisfied due to hydro releases must be met by thermal generation, or considered as lost load. Our objective is then to minimize the costs of thermal generation and lost load over a finite planning horizon. In our experiments, we construct policies for a planning period of one year, which is divided into 52 weekly stages.

5.1 Implementation

The SDDP algorithm is implemented in Julia [6], using the Julia library SDDP.jl [8] and a Julia model of HTSP formulated in JuMP [9]. All linear subproblems are solved in Gurobi [12].

During policy generation, we sample historical realizations of weekly inflows over the years 1970 — 1999, giving $m = 30$ stagewise independent random outcomes per week. The policies we compute using these samples aim to minimize the (risk-adjusted) expected cost of meeting electricity demand over the calendar year 2008. We apply a cut selection heuristic [7] every 50 iterations of SDDP to reduce computational effort.

Policies were generated using Algorithm 3 with radii $\frac{1}{m}, \frac{2}{m}, \frac{4}{m}$, and also without using a distributionally robust approach. We also created policies using a nested risk measure based on the one-stage measure

$$\rho(Z) = (1 - \lambda)\mathbb{E}[Z] + \lambda CVaR_{1-\beta}[Z]$$

as described in [18]. We chose $\lambda = 0.5$ and tested values of $\beta = 0.1$, 0.2, and 0.3.

Each policy was created with 10,000 cuts and simulated using the ten historical inflow sequences observed in the years 2000 — 2009. Since none of the inflow data in these ten years is used to generate the policy, the simulations can be viewed as out-of-sample tests of the policy. It is also important to note that Algorithm 3 samples in its forward pass, and so even after a large number of SDDP iterations, policies generated from different samples of forward passes will not be identical. Policies that are not identical will not necessarily give rise to the same sequence of actions when simulated. In practice, we set a random seed at the start of policy generation, and pseudo-random numbers dictate the inflows that are sampled while a policy is computed. In order to explore the effect of different uncertainty sets in SDDP, it is necessary to generate samples of policies for each uncertainty set.

Policies were created using ten different random seeds (ten different sets of 10,000 scenarios), for each uncertainty set. Observe that this is a form of out-of-sample testing that uses only one hold-out observation (the year to be studied). In practice the year in question would be dealt with by solving SDDP once using historical (in-sample) data and then applying the policy. Since SDDP creates (ten) random policies for each uncertainty set, we can view each policy simulated on the year in question to be an out-of-sample test. The results of these are shown in Table 1 and Table 2.

Table 1 Mean simulated cost (NZ$(M)) for policies created with different uncertainty sets.

	No DRO	DRO			CVaR with $\lambda = 0.5$		
Years	$r = 0$	$r = 1/m$	$r = 2/m$	$r = 4/m$	β= 0.1	β= 0.2	β= 0.3
2000	286.226	285.742	286.809	301.685	286.119	288.513	288.651
2001	475.933	476.991	477.778	480.112	477.221	478.571	479.516
2002	342.190	342.633	343.483	348.345	343.794	344.504	344.292
2003	387.975	385.473	385.503	386.780	386.560	385.913	385.681
2004	256.212	254.958	254.189	253.308	255.121	254.153	253.958
2005	487.581	484.349	482.792	483.685	483.240	481.887	481.532
2006	349.597	350.264	350.573	353.512	350.150	351.004	350.329
2007	449.999	449.803	449.988	450.603	450.974	449.803	449.858
2008	501.069	471.477	461.730	481.015	454.249	466.564	471.199
2009	345.046	344.281	343.431	347.191	344.973	343.130	342.986

Table 2 Standard deviation of simulated cost (NZ$(M)) for policies created with different uncertainty sets.

	No DRO	DRO			CVaR with $\lambda = 0.5$		
Years	$r = 0$	$r = 1/m$	$r = 2/m$	$r = 4/m$	β= 0.1	β= 0.2	β= 0.3
2000	0.082	0.138	0.282	3.485	1.678	2.483	0.146
2001	0.088	0.040	0.033	0.245	0.042	0.261	0.268
2002	0.100	0.111	0.332	2.090	0.155	0.214	0.277
2003	5.159	0.560	0.368	1.296	0.267	0.124	0.122
2004	0.348	0.140	0.258	0.597	0.442	0.467	0.555
2005	0.529	0.191	0.188	2.721	0.351	0.275	0.198
2006	0.064	0.203	0.438	1.408	0.164	0.425	0.077
2007	0.008	0.021	0.617	1.606	0.031	0.032	0.188
2008	42.335	18.743	18.424	60.108	12.657	19.668	27.152
2009	0.019	0.025	0.026	3.964	0.209	0.157	0.031

The years (2001, 2005, 2008) with high average costs are those with dry winters in which a lot of thermal fuel is consumed as well as some load shedding in extreme cases. The years with low average costs (2000, 2004) are those with winters with high reservoir inflows. In some of the dry years increasing the uncertainty-set radius (to $1/m$ or $2/m$) appears to give lower average cost out of sample. This is particularly noticeable in 2008. In high inflow years this outcome is more ambiguous. Observe also that distributional robustness decreases out-of-sample variation in dry years, but this is not always the case for the other years.

6 Conclusions

We have shown how SDDP can be extended to solve a distributionally robust model. This generalizes the capability of SDDP beyond models with nested coherent risk measures as discussed in [18]. We have also indicated how uncertainty sets might be constructed that vary with stored energy levels, so that decision makers can adapt their levels of conservatism with observed hydrological conditions.

Our research has been aimed at understanding the effect on policies of a distributionally robust approach. From our experiments the outcomes of the policies become less variable and less costly for sample years where there are substantial risks of high costs.

Acknowledgements

We are very grateful for discussions with Oscar Dowson, Eddie Anderson, Karen Willcox, and Michael Kapteyn, and for comments from participants at the Workshop on Stochastic Programming Honoring Maarten van der Vlerk held at the University of Groningen on August 10-11, 2017. Andy Philpott acknowledges the financial support of the New Zealand Marsden Fund under contract UOA1520.

References

1. P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, *Coherent measures of risk*, Mathematical Finance **9** (1999), no. 3, 203–228.
2. G. Bayraksan and D.K. Love, *Data-driven stochastic programming using phi-divergences*, The Operations Research Revolution, INFORMS, 2015, pp. 1–19.
3. D. Bertsimas, D.B. Brown, and C. Caramanis, *Theory and applications of robust optimization*, SIAM review **53** (2011), no. 3, 464–501.
4. D. Bertsimas, V. Gupta, and N. Kallus, *Data-driven robust optimization*, arXiv preprint arXiv:1401.0212 (2013).
5. ———, *Robust SAA*, arXiv preprint arXiv:1408.4445 (2014), 617.
6. J. Bezanson, A. Edelman, S. Karpinski, and V.B. Shah, *Julia: A fresh approach to numerical computing*, SIAM Review **59** (2017), no. 1, 65–98.
7. V.L. de Matos, A.B. Philpott, and E.C. Finardi, *Improving the performance of stochastic dual dynamic programming*, Journal of Computational and Applied Mathematics **290** (2015), 196–208.
8. O. Dowson, *SDDP in Julia*, Tech. report, University of Auckland, 2017.
9. I. Dunning, J. Huchette, and M. Lubin, *JuMP: A modeling language for mathematical optimization*, SIAM Review **59** (2017), no. 2, 295–320.
10. P.M. Esfahani and D. Kuhn, *Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations*, arXiv preprint arXiv:1505.05116 (2015).
11. R. Gao and A.J. Kleywegt, *Distributionally robust stochastic optimization with Wasserstein distance*, arXiv preprint arXiv:1604.02199 (2016).
12. Gurobi Optimization Inc., *Gurobi Optimizer Reference Manual*, 2016.
13. G. Infanger and D.P. Morton, *Cut sharing for multistage stochastic linear programs with interstage dependency*, Mathematical Programming **75** (1996), no. 2, 241–256.
14. D. Klabjan, D. Simchi-Levi, and M. Song, *Robust stochastic lot-sizing by means of histograms*, Production and Operations Management **22** (2013), no. 3, 691–710.

15. V. Kozmík and D.P. Morton, *Evaluating policies in risk-averse multi-stage stochastic programming*, Mathematical Programming **152** (2015), no. 1, 275–300.
16. M.V.F. Pereira and L.M.V.G. Pinto, *Multi-stage stochastic optimization applied to energy planning*, Mathematical Programming **52** (1991), no. 1-3, 359–375.
17. A.B. Philpott and V.L. de Matos, *Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion*, European Journal of Operational Research **218** (2012), no. 2, 470–483.
18. A.B. Philpott, V.L. de Matos, and E. Finardi, *On solving multistage stochastic programs with coherent risk measures*, Operations Research **61** (2013), no. 4, 957–970.
19. A.B. Philpott and Pritchard G., *EMI-DOASA, downloadble from https://www.emi.ea.govt.nz/Content/Tools/Doasa*, 2013.
20. R.T. Rockafellar, *Convex analysis*, Princeton University Press, 1972.
21. A. Shapiro, *Analysis of stochastic dual dynamic programming method*, European Journal of Operational Research **209** (2011), no. 1, 63 – 72.
22. A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming: modeling and theory*, SIAM, 2009.

A Appendix

This appendix contains proofs of all the results in the paper. We begin by proving some technical lemmas. Recall the problem

$$\begin{aligned} \rho_r(z) = \max\ & \textstyle\sum_{i=1}^m z_i p_i \\ \text{s.t. } & \textstyle\sum_{i=1}^m p_i = 1, \\ & \|p - q\|_2 \le r, \\ & p \ge 0. \end{aligned}$$

Lemma 1 *Given z_i, $i = 1, 2, \ldots, m$, $\rho_r(z)$ is a continuous and nondecreasing function of $r \ge 0$.*

Proof Observe that given $r \ge 0$ and z_i, $i = 1, 2, \ldots, m$, P is a convex optimization problem with a compact feasible region, so it has an optimal solution, with optimal value denoted $\rho_r(z)$. Moreover for fixed z, the optimal value function is a concave increasing function of r. By [20, Theorem 10.1] $\rho_r(z)$ is therefore continuous on $\{r : r > 0\}$. It is also easy to show that $\lim_{r \to 0} \rho_r(z) = \sum_{i=1}^m q_i z_i$, so $\rho_r(z)$ is also continuous at $r = 0$.

Consider the problem

$$\begin{aligned} \mathrm{P}(n)\colon \min\ & -\textstyle\sum_{i=1}^n z_i y_i \\ \text{s.t. } & \textstyle\sum_{i=1}^n y_i = a, \\ & \textstyle\sum_{i=1}^n y_i^2 \le b^2. \end{aligned}$$

Lemma 2 *P(n) has an optimal solution if and only if $a^2 \le nb^2$.*

Proof The objective function is continuous and the feasible region is compact, and so P0 has an optimal solution if and only if the feasible region is nonempty. If $a^2 \leq nb^2$ then $y_i = \frac{a}{n}$ is feasible as it satisfies

$$\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n} \left(\frac{a}{n}\right)^2 = \frac{a^2}{n} \leq b^2.$$

If $a^2 > nb^2$ then the feasible region is empty, for any feasible y satisfies $\sum_{i=1}^{n} y_i = a$ so

$$\left(\sum_{i=1}^{n} y_i\right)^2 > nb^2 \geq n \sum_{i=1}^{n} y_i^2$$

But this contradicts

$$n \sum_{i=1}^{n} y_i^2 \geq \left(\sum_{i=1}^{n} y_i\right)^2$$

which is true because

$$n \sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2 = n \sum_{i=1}^{n} (y_i - \frac{\sum_{i=1}^{n} y_i}{n})^2$$
$$\geq 0.$$

Lemma 3 *Suppose $a^2 \leq nb^2$. The optimal solution to P(n) is*

$$y_i = \frac{a}{n} + \sqrt{nb^2 - a^2}\,\frac{z_i - \bar{z}}{ns}.$$

Proof Since P(n) is a convex program we can solve it by minimizing the Lagrangian, giving

$$\min_y -\textstyle\sum_{i=1}^{n} z_i y_i + \mu(\sum_{i=1}^{n} y_i - a) + \lambda(\sum_{i=1}^{n} y_i^2 - b^2).$$

Differentiating gives

$$-z_i + \mu + 2\lambda y_i = 0$$

so

$$y_i = \frac{z_i - \mu}{2\lambda}$$

Since

$$\sum_{i=1}^{n} y_i = a$$

we have

$$\sum_{i=1}^{n} z_i - n\mu = 2\lambda a.$$

If $a = 0$ then

$$\mu = \bar{z}$$

so

$$y_i = \frac{z_i - \bar{z}}{2\lambda}$$

This gives

$$\sum_{i=1}^{n} y_i^2 = \frac{ns^2}{4\lambda^2} = b^2$$

so

$$2\lambda = \sqrt{n}\frac{s}{b}$$

and

$$y_i = \frac{b\left(z_i - \bar{z}\right)}{s\sqrt{n}}$$

as required.

If $a \neq 0$ then substituting for λ gives

$$y_i = \frac{-z_i + \mu}{-\sum_{i=1}^{n} z_i + n\mu} a$$

Now

$$\sum_{i=1}^{n} y_i^2 = b^2$$

so

$$\sum_{i=1}^{n} \left(\frac{-z_i + \mu}{-\sum_{i=1}^{n} z_i + n\mu}\right)^2 a^2 = b^2$$

$$a^2 \sum_{i=1}^{n} \left(-z_i + \mu\right)^2 = b^2 (-\sum_{i=1}^{n} z_i + n\mu)^2$$

$$a^2 \left(\sum_{i=1}^{n} z_i^2 - 2\mu \sum_{i=1}^{n} z_i + n\mu^2\right) = b^2 \left((\sum_{i=1}^{n} z_i)^2 - 2n\mu \sum_{i=1}^{n} z_i + n^2\mu^2\right) \quad (8)$$

Let $z = \sum_{i=1}^{n} z_i$, and $d = \sum_{i=1}^{n} z_i^2$, so

$$\begin{aligned} dn - z^2 &= n \sum_{i=1}^{n} z_i^2 - \left(\sum_{i=1}^{n} z_i\right)^2 \\ &= n^2 s^2. \end{aligned}$$

Equation (8) is a quadratic in μ,

$$\left(a^2 n - n^2 b^2\right) \mu^2 - \left(2a^2 z - 2znb^2\right) \mu + \left(a^2 d - z^2 b^2\right) = 0,$$

that has roots

$$\mu \in \{\bar{z} - \frac{a}{\sqrt{nb^2 - a^2}} s, \bar{z} + \frac{a}{\sqrt{nb^2 - a^2}} s\}.$$

The first root is

$$\mu = \bar{z} - \frac{a}{\sqrt{nb^2 - a^2}} s.$$

This gives

$$\begin{aligned} y_i &= \frac{-z_i + \mu}{-\sum_{i=1}^n z_i + n\mu} a \\ &= \frac{-z_i + \bar{z} - \frac{a}{\sqrt{nb^2 - a^2}} s}{-\sum_{i=1}^n z_i + n\bar{z} - \frac{na}{\sqrt{nb^2 - a^2}} s} a \\ &= \frac{a}{n} + \sqrt{nb^2 - a^2} \frac{z_i - \bar{z}}{ns} \end{aligned}$$

with objective

$$\begin{aligned} -\sum_{i=1}^n z_i y_i &= -\sum_{i=1}^n z_i \frac{a}{n} - \sum_{i=1}^n \sqrt{nb^2 - a^2} \frac{z_i^2 - \bar{z} z_i}{ns} \\ &= -a\bar{z} - s\sqrt{nb^2 - a^2} \end{aligned}$$

(The other root of (8) gives

$$y_i = \frac{a}{n} - \sqrt{nb^2 - a^2} \frac{c_i - \bar{c}}{ns}$$

which has value $-a\bar{z} + s\sqrt{nb^2 - a^2}$, a local maximum of P($n$).)

Lemma 6 *Let* $\bar{z} = \frac{\sum_{i=1}^n z_i}{n}$. *For all* i,

$$|z_i - \bar{z}| \leq \sqrt{\frac{(n-1)}{n} \sum_{j=1}^n (z_j - \bar{z})^2}.$$

Proof Without loss of generality choose $i = 1$. Then

$$
\begin{aligned}
&(n-1)\sum_{j=2}^{n}\left(z_j-\frac{1}{n}\sum_{i=1}^{n}z_i\right)^2-\left(z_1-\frac{1}{n}\sum_{i=1}^{n}z_i\right)^2\\
&=(n-1)\sum_{j=2}^{n}(z_j-\bar{z})^2-(z_1-\bar{z})^2\\
&=(n-1)\sum_{j=2}^{n}\left(z_j^2-2\bar{z}z_j+\bar{z}^2\right)-(z_1-\bar{z})^2\\
&=(n-1)\sum_{j=2}^{n}z_j^2-2(n-1)\bar{z}\sum_{j=2}^{n}z_j+(n-1)^2\bar{z}^2-z_1^2+2\bar{z}z_1-\bar{z}^2\\
&=(n-1)\sum_{j=2}^{n}z_j^2-2(n-1)\bar{z}(n\bar{z}-z_1)+(n-1)^2\bar{z}^2-z_1^2+2\bar{z}z_1-\bar{z}^2\\
&=(n-1)\sum_{j=2}^{n}z_j^2-(z_1-\bar{z}n)^2\\
&=(n-1)\sum_{j=2}^{n}z_j^2-\left(\sum_{j=2}^{n}z_j\right)^2.
\end{aligned}
$$

This expression is $(n-1)^2$ times the variance of the quantities $z_2, z_3, \ldots, z_n$ which is nonnegative, so it follows that

$$
n\left(z_1-\frac{1}{n}\sum_{i=1}^{n}z_i\right)^2\leq(n-1)\left(\sum_{j=2}^{n}\left(z_j-\frac{1}{n}\sum_{i=1}^{n}z_i\right)^2+\left(z_1-\frac{1}{n}\sum_{i=1}^{n}z_i\right)^2\right)
$$

from which the result follows.

Lemma 4 *If* $r \leq \sqrt{\frac{m}{m-1}}\min_i\{q_i\}$ *then* P *has optimal solution*

$$
p_i = q_i + \frac{z_i-\bar{z}}{\sqrt{m}}\frac{r}{s}.
$$

with optimal value $\sum_{i=1}^{m} q_i z_i + (\sqrt{m}s)\, r$.

Proof We consider a solution in which we drop the constraint $p \geq 0$. This gives $p_i = q_i + y_i$ where y solves

$$
\begin{array}{ll}
\min & -\sum_{i=1}^{m} z_i y_i\\
\text{s.t.} & \sum_{i=1}^{m} y_i = 0,\\
& \sum_{i=1}^{m} y_i^2 \leq r^2.
\end{array}
$$

Thus Lemma 3 with $a = 0$ gives

$$
y_i = \sqrt{mr^2}\frac{z_i-\bar{z}}{ms}
$$

whence

$$p_i = q_i + \frac{z_i - \bar{z}}{\sqrt{m}} \frac{r}{s}.$$

Now by Lemma 6

$$\begin{aligned} \bar{z} - z_i &\leq \sqrt{\frac{(m-1)}{m} \sum_i (z_i - \bar{z})^2} \\ &= \sqrt{m-1}s \end{aligned}$$

so

$$\begin{aligned} p_i &= q_i + \frac{\bar{z} - z_i}{\sqrt{m}} \frac{r}{s} \\ &\geq q_i + \frac{\sqrt{m-1}r}{\sqrt{m}} \\ &\geq 0 \end{aligned}$$

as $r \leq \sqrt{\frac{m}{m-1}} \min_i\{q_i\}$ by assumption.

We also have

$$\begin{aligned} \rho(Z(x)) &= \sum_{i=1}^{m} p_i z_i \\ &= \sum_{i=1}^{m} q_i z_i + \sum_{i=1}^{m} \frac{z_i - \bar{z}}{\sqrt{m}} \frac{r}{s} z_i \\ &= \sum_{i=1}^{m} q_i z_i + \frac{r}{\sqrt{m}s} \sum_{i=1}^{m} \left(z_i^2 - \bar{z} z_i\right) \\ &= \sum_{i=1}^{m} q_i z_i + \frac{r}{\sqrt{m}s} \left(\sum_{i=1}^{m} z_i^2 - m\bar{z}^2\right) \\ &= \sum_{i=1}^{m} q_i z_i + \frac{r}{\sqrt{m}s} \left(ms^2\right) \\ &= \sum_{i=1}^{m} q_i z_i + \left(\sqrt{m}s\right) r. \end{aligned}$$

Lemma 5 *If* $r_1 < r_2$ *then* $k(r_1) \leq k(r_2)$.

Proof Recall for any fixed k and r that if $r \leq \sqrt{\frac{1}{m(m-1)}}$,

$$p_i(r) = \frac{1}{m} + \frac{r}{s\sqrt{m}}(z_i - \bar{z}),\ i = 1, \ldots, m,$$

and otherwise

$$p_i(r) = \begin{cases} 0 & i = 1, \ldots, k(r), \\ \frac{1}{(m-k(r))} + \frac{\sqrt{(m-k(r))r^2 - \frac{k(r)}{m}} \frac{z_i - \bar{z}}{s}}{(m-k(r))} & i = k(r)+1, \ldots, m. \end{cases}$$

If we assume that $k(r)$ is constant at value k as r varies then

$$\frac{dp_i(r)}{dr} = \begin{cases} 0 & i = 1, \dots, k, \\ \frac{z_i - \bar{z}}{s} \frac{r}{\sqrt{(m-k)r^2 - \frac{k}{m}}} & i = k+1, \dots, m. \end{cases} \tag{9}$$

Since $z_i \leq z_{i+1}$, for a fixed k and r we have

$$p_i(r) \leq p_{i+1}(r),\ i = k+1, \dots, m-1.$$

Also $\frac{dp_{k+1}(r)}{dr} \leq \frac{dp_i(r)}{dr}$ for $i = k+2, \dots, m$ and $\frac{dp_{k+1}(r)}{dr} \leq 0$. This means that as r increases $p_{k+1}(r)$ becomes negative before (or at that same r as) any $p_i(r)$, $i = k+2, \dots, m$. In other words as r increases $k(r)$ remains constant and then increases at some value of r. If the z_i are distinct then $k(r)$ increases by 1 at each step. We therefore have $k(0) = 0$, and $k(r)$ is piecewise constant and nondecreasing with r.

Proposition 1 *Suppose that $Z(x, \omega)$ is a convex function of x for each $\omega \in \Omega$, and that $g(\tilde{x}, \omega)$ is a subgradient of $Z(x, \omega)$ at $\tilde{x}$. Then $\mathbb{E}_{\mathbb{P}^*}[g(\tilde{x}, \omega)]$ is a subgradient of $\max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[Z(x, \omega)]$ at $\tilde{x}$, where $\mathbb{P}^* \in \arg\max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[Z(\tilde{x}, \omega_m)]$.*

Proof For any x,

$$\begin{aligned} \max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}[Z(x, \omega)] &= \mathbb{E}_{\mathbb{P}^*}[Z(x, \omega)] \\ &\geq \mathbb{E}_{\mathbb{P}^*}[Z(\tilde{x}, \omega) + g(\tilde{x}, \omega)^\top (x - \tilde{x})] \\ &= \mathbb{E}_{\mathbb{P}^*}[Z(\tilde{x}, \omega)] + \left(\mathbb{E}_{\mathbb{P}^*}[g(\tilde{x}, \omega)]\right)^\top (x - \tilde{x}) \end{aligned}$$

which demonstrates that $\mathbb{E}_{\mathbb{P}^*}[g(\tilde{x}, \omega)]$ is a subgradient at $\tilde{x}$.

Proposition 2 *If for any $x_t \in \mathcal{X}_t(\omega_t)$, $h_{t+1,k} - \bar{\pi}_{t+1,k}^\top H_{t+1} x_t \leq \mathbb{E}_{\mathbb{P}_t^*}[Q_{t+1}(x_t, \omega_{t+1})]$ for every $k = 1, 2, \dots, \nu$, then*

$$\tilde{Q}_t(x_{t-1}, \omega_t) \leq Q_t(x_{t-1}, \omega_t).$$

Proof For any $x_t \in \mathcal{X}_t(\omega_t)$ the optimal choice of θ_{t+1} satisfies

$$\begin{aligned} c_t^\top x_t + \theta_{t+1} &= c_t^\top x_t + \max_k \{h_{t+1,k} - \bar{\pi}_{t+1,k}^\top H_{t+1} x_t\} \\ &\leq c_t^\top x_t + \mathbb{E}_{\mathbb{P}_t^*}[Q_{t+1}(x_t, \omega_{t+1})] \end{aligned}$$

by hypothesis. It follows that

$$\begin{aligned} \tilde{Q}_t(x_{t-1}, \omega_t) &= \min_{x_t \in \mathcal{X}_t(\omega_t)} \{c_t^\top x_t + \max_k \{h_{t+1,k} - \bar{\pi}_{t+1,k}^\top H_{t+1} x_t\}\} \\ &\leq \min_{x_t \in \mathcal{X}_t(\omega_t)} \{c_t^\top x_t + \mathbb{E}_{\mathbb{P}_t^*}[Q_{t+1}(x_t, \omega_{t+1})]\} \\ &= Q_t(x_{t-1}, \omega_t) \end{aligned}$$

giving the desired result.