Appendix for:
Statistical performance of subgradient step-size update rules in Lagrangian relaxations of chance-constrained optimization models

Charlotte Ritter[a], Bismark Singh[b,*]

[a] *Department of Mathematics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, 91058*
[b] *School of Mathematical Sciences, University of Southampton, Southampton, S017 1BJ, UK*

There are two components to this appendix. First, we provide a motivating example and algorithms we omitted in the main text to preserve space. Next, we provide three tables reporting additional computational results.

A. Algorithms

Here, we provide pseudocodes for the Lagrangian procedure of Section 2 and the three lower bounding procedures of Section 3. We first provide a motivating example that we refer to in Section 2.

Example S1 demonstrates that, unfortunately, in practice we cannot provide a strong optimality certificate as that given by the theoretical bound in the `BBP`.

Example S1. *Consider the following MIP:*

$$\begin{aligned} \mathcal{Z} = \max_{x,y} \quad & 2x + 3y \\ s.t. \quad & x + y \leq 1 \\ & 3x + y \leq 3 \\ & x, y \in \{0, 1\}. \end{aligned}$$

The optimal solution is $x = 0, y = 1$ with $\mathcal{Z} = 3$. Consider the following Lagrangian dual model:

$$\begin{aligned} \mathcal{Z}_D(\lambda) = \max_{x,y} \quad & 2x + 3y + \lambda(1 - x - y) \\ s.t. \quad & 3x + y \leq 3 \\ & x, y \in \{0, 1\}. \end{aligned}$$

Assume that the branch-and-bound solver terminates with the following feasible solution for the Lagrangian dual model: $x = 1, y = 0$ that provides $\underline{\mathcal{Z}_D}(\lambda) = 2, \forall \lambda \geq 0$. We note that this solution is feasible for the

*Corresponding author
Email address: `b.singh@southampton.ac.uk` (Bismark Singh)

true model. A corresponding upper bound is available from $x = 0, y = 1$ with $\overline{\mathcal{Z}_D}(\lambda) = 3, \forall \lambda \geq 0$. Since $\underline{\mathcal{Z}_D}(\lambda) < \mathcal{Z}$, the feasible solution reported by the solver cannot be used for the BBP. *Further, since the precise value of $\mathcal{Z}_D(\lambda) \in [\underline{\mathcal{Z}_D}(\lambda), \overline{\mathcal{Z}_D}(\lambda)]$ is unknown, the* BBP *is of no value.* □

Next, we provide the four algorithms.

Algorithm S1 Lagrangian relaxation procedure of model (1)

Input: *iter*; $\underline{\mathcal{Z}}$; *time*; θ; ψ; ι; step-size rule r.

Output: $\mathcal{Z}_D$; δ.

Solve LP relaxation of model (1); $\mathcal{Z}_D \leftarrow$ optimal objective function value; $\lambda \leftarrow$ optimal dual of constraint (2b).
$k \leftarrow 1$.
while $k \leq iter$ **do**
  Solve model (3); $\mathcal{Z}_D \leftarrow \min\{\mathcal{Z}_D, \overline{\mathcal{Z}_D}(\lambda)\}$; $\gamma \leftarrow \varepsilon - \sum_{\omega \in \Omega} p^\omega z^\omega$ with optimal z^ω.
  If no change in $\mathcal{Z}_D$ between current and previous two iterations, $\theta \leftarrow \frac{\theta}{2}$.
  Update Δ according to step-size rule r.
  $\lambda_{\text{new}} \leftarrow \max\{0, \lambda - \Delta\gamma\}$; $\zeta \leftarrow |\lambda_{\text{new}} - \lambda|$; $\lambda \leftarrow \lambda_{\text{new}}$.
  $\delta \leftarrow \frac{\mathcal{Z}_D - \underline{\mathcal{Z}}}{\mathcal{Z}_D}$.
  If $\delta \leq \psi$ or time $\geq$ *time* or $\zeta < \iota$, *STOP*.
  $k \leftarrow k + 1$; update time to the cumulative wall-clock time.

Algorithm S2 Iterative regularization bound of [3] for model (1)

Input: m; κ; *time*; *iter*; ρ; Oracle to generate independent scenarios for model (1); instance of model (1) with $|\Omega|$ scenarios.

Output: $\underline{\mathcal{Z}}$.

$\underline{\mathcal{Z}} \leftarrow 0$; $k \leftarrow 1$.
while time $\leq$ *time* and $k \leq iter$ **do**
  if $k = 1$ **then**
    Generate m independent scenarios from Oracle; solve SAA of model (1) with these m scenarios; $\hat{x} \leftarrow$ optimal x.
  else
    Generate m independent scenarios from Oracle; solve SAA of model (7) with these m scenarios; $\hat{x} \leftarrow$ optimal x.
  Solve input instance of model (1) with x fixed to $\hat{x}$; $z_m \leftarrow$ objective function value.
  $\underline{\mathcal{Z}} \leftarrow \max\{z_m, \underline{\mathcal{Z}}\}$.
  $m \leftarrow \lceil (1 + \kappa)m \rceil$.
  $k \leftarrow k + 1$; update time to wall-clock time.

Algorithm S3 Aggregation bound for model (1)

Input: m; κ; $time$; $iter$; ρ; p; Oracle to generate independent scenarios for model (1); procedure `cluster`(a, b) that aggregates set of a scenarios into set of b scenarios and their respective weights, $a > b$; instance of model (1) with $|\Omega|$ scenarios.

Output: $\underline{\mathcal{Z}}$.

1: Generate $p >> |\Omega|$ independent scenarios from Oracle.
2: $\underline{\mathcal{Z}} \leftarrow 0$; $k \leftarrow 1$.
3: **while** time $\leq time$ and $k \leq iter$ **do**
4:    **if** k=1 **then**
5:       Use `cluster`(p, m); solve SAA of model (1) with these m scenarios and their respective weights; $\hat{x} \leftarrow$ optimal x.
6:    **else**
7:       Use `cluster`(p, m); solve SAA of model (7) with these m scenarios and their respective weights; $\hat{x} \leftarrow$ optimal x.
8:    Solve input instance of model (1) with x fixed to $\hat{x}$; let z_m denote the objective function value.
9:    $\underline{\mathcal{Z}} \leftarrow \max\{z_m, \underline{\mathcal{Z}}\}$.
10:    $m \leftarrow \lceil (1+\kappa)m \rceil$.
11:    $k \leftarrow k + 1$; update time to wall-clock time.

Algorithm S4 Quantile bound of [2, 1] for equally likely scenarios for model (1)

Input: instance of model (1) with $|\Omega|$ scenarios.

Output: $\underline{\mathcal{Z}}$.

1: Solve model (1) separately for each ω in the input instance with $z^\omega \leftarrow 0$.
2: Sort corresponding objective function values in ascending order; $z^\omega \leftarrow 1$ for first $\lfloor |\Omega| \varepsilon \rfloor$ scenarios, $z^\omega \leftarrow 0$ for rest.
3: Solve input instance of model (1) with z fixed from Step 2.
4: $\underline{\mathcal{Z}} \leftarrow$ objective value.

B. Supplemental computational results

Here, we provide additional computational experiments accompanying those in the main text.

Table S1 presents the 95% confidence intervals for the objective function value, the time, and the optimality gap on solving model (1) naively.

Table S2 presents the 95% confidence intervals for the objective function value, the time, and the improvement in relation to the naive solution of the three lower bounding techniques. Here, the trend we report in Table 2 is further validated. For Model I, the `QP` bound is computed very quickly; however, compared to the `IR` bound over the larger instances, the improvements for `QP` are much smaller. For example, consider the $|\Omega| = 1500$ regime for Model I. The `QP` bound has an improvement of at most 3.3%, while the `IR` bound has an improvement that is at least double (6.7%). We further observe that the `AP` bound is practically useless for Model I. For all instances, in each of the 20 batches with $|\Omega| \geq 600$, we

$\lvert\Omega\rvert$	Objective CI	Time CI	Gap CI [%]
100	(280.94, 282.29)	(387, 453)	(0.01, 0.04)
600	(239.28, 244.64)	**T**	(14.77, 16.49)
900	(234.08, 240.19)	**T**	(49.19, 86.23)
1500	(233.42, 237.98)	**T**	(14.16, 37.89)

(a) Model I

$\lvert\Omega\rvert$	Objective CI	Time CI	Gap CI [%]
100	(6,385.63, 6,529.70)	**T**	(42.69, 43,53)
600	(4,588.53, 4,755.71)	**T**	(62.13, 63.38)
900	(4,296.65, 4,417.90)	**T**	(65.62, 66.55)
1500	(2,879.22, 3,308.35)	**T**	(74.92, 78.17)

(b) Model II

Table S1: Computational results for a naive solution method on 20 instances of model (1) for $\varepsilon = 0.05$. CI denotes the 95% confidence interval. For details, see Section 5.2.

	`IR`			`AP`			`QP`		
$\lvert\Omega\rvert$	Objective CI	Time CI	Improvement CI [%]	Objective CI	Time CI	Improvement CI [%]	Objective CI	Time CI	Improvement CI [%]
100	(261.70, 263.43)	(244, 245)	(-7.08, -6.35)	(232.76, 232.97)	(430, 431)	(-17.45, -17.08)	(279.93, 281.63)	(4, 4)	(-0.44, -0.05)
600	(253.65, 254.17)	(822, 827)	(3.80, 6.20)	✗	✗	✗	(246.37, 249.76)	(44, 48)	(1.43, 3.72)
900	(253.61, 254.10)	(1194, 1198)	(5.39, 8.37)	✗	✗	✗	(244.88, 247.96)	(73, 81)	(2.13, 5.37)
1500	(253.72, 254.29)	(2087, 2096)	(6.73, 8.87)	✗	✗	✗	(239.06, 243.83)	(165, 184)	(1.60, 3.28)

(a) Model I

	`IR`			`AP`			`QP`		
$\lvert\Omega\rvert$	Objective CI	Time CI	Improvement CI [%]	Objective CI	Time CI	Improvement CI [%]	Objective CI	Time CI	Improvement CI [%]
100	(5,650.76, 5,871.67)	**T**	(-12.45, -9.07)	(5,110.87, 5,342.39)	**T**	(-20.98, -17.07)	(6,729.79, 6,852.93)	**T**	(4.60, 5.77)
600	(5,494.22, 5,612.70)	**T**	(17.37, 20.57)	(5,061.44, 5,184.64)	**T**	(7.66, 11.92)	(4,700.14, 4,896.32)	**T**	(-0.21, 5.96)
900	(5,545.61, 5,621.69)	**T**	(26.29, 30.22)	(5,144.17, 5,241.68)	**T**	(17.57, 20.95)	(4,344.65, 4,563.58)	**T**	(-0.60, 5.23)
1500	(5,574.85, 5,638.68)	**T**	(69.50, 103.09)	(5,148.49, 5,215.68)	**T**	(56.35, 88.31)	(3,184.35, 3,477.76)	**T**	(-0.92, 23.07)

(b) Model II

Table S2: Computational results for three lower bounding heuristics on 20 instances of model (1) for $\varepsilon = 0.05$. CI denotes the 95% confidence interval. For details, see Section 5.3.

fail to obtain a feasible solution with the `AP` bound.

As in Table 2b, we reach the time limit in Table S2b for all methods for the computationally more challenging Model II. However, in contrast to Table S2a we observe that the `QP` bound lags behind not only the `IR` bound but also the `AP` bound. For larger instances of Model II, the upper limit of the CI for the `QP` bound is lower than the lower limit of the CI for both the `IR` and `AP` bounds. The `IR` bound performs the best here, consistently delivering the highest improvement CIs for all instances except the smallest instance with $|\Omega| = 100$. For the smallest instance, the `QP` performs better for both Model I and Model II, further validating our empirical claim that the `IR` and `AP` bounds are most suitable for the larger instances. Finally, we also observe that the improvement by the `IR` bound is significantly higher for Model II than for Model I. This is especially evident from the last rows of Table S2a and Table S2b where the improvements in Model II increase by an order of magnitude, e.g., from 6.7% to 69.5%.

Next, we run Algorithm S1 with a time limit of $time = 41000$ seconds. Compared to Table 3, we now have seven additional instances in Table S3b for Model I where we can reject the null hypothesis. For Model II, there are 16 such additional instances in Table S3a. This supports our premise that the previous time limit was not sufficiently high enough to allow for rejections. For Model I in Table 3a, Rule VI is statistically the best performer, however only for the smallest scenario regime of $|\Omega| = 100$. Now—given additional time as in Table S3b—Rule VI has rejections in its favor compared to all other rules for two scenario regimes $|\Omega| = \{100, 1500\}$. For Model II, given this markedly larger amount of additional time, Rule II has statistically significant evidence to perform better than all other rules for all regimes except the smallest one of $|\Omega| = 100$. We examine this particular scenario regime in greater detail. Here, Algorithm S1 manages to complete between five and eight iterations for the different instances.

(r,s)	$\lvert\Omega\rvert$			
	100	600	900	1500
(I,II)	✗	✗	✗	✗
(I,III)	✗	✗	✗	✗
(I,IV)	✗	✗	✗	✗
(I,V)	✗	✗	✗	✗
(I,VI)	✗	✗	✗	✗
(II,I)	✗	✗	✗	✗
(II,III)	✗	✗	✗	✗
(II,IV)	✗	✗	✗	✗
(II,V)	✗	✗	✗	✗
(II,VI)	✗	✗	✗	✗

(r,s)	$\lvert\Omega\rvert$			
	100	600	900	1500
(III,I)	✗	✗	✗	✗
(III,II)	✗	✗	✗	✗
(III,IV)	✗	✗	✗	✗
(III,V)	✗	✗	✗	✗
(III,VI)	✗	✗	✗	✗
(IV,I)	✗	✗	✗	✗
(IV,II)	✗	✗	✗	✗
(IV,III)	✗	✗	✗	✗
(IV,V)	✗	✗	✗	✗
(IV,VI)	✗	✗	✗	✗

(r,s)	$\lvert\Omega\rvert$			
	100	600	900	1500
(V,I)	✓	✗	✗	✗
(V,II)	✓	✗	✗	✗
(V,III)	✓	✗	✗	✗
(V,IV)	✓	✗	✗	✗
(V,VI)	✗	✗	✗	✗
(VI,I)	✓	✗	✗	✓
(VI,II)	✓	✓	✗	✓
(VI,III)	✓	✗	✗	✓
(VI,IV)	✓	✗	✗	✓
(VI,V)	✓	✓	✗	✓

(a) Model II

(r,s)	$\lvert\Omega\rvert$			
	100	600	900	1500
(I,II)	✓	✗	✗	✗
(I,III)	✗	✗	✗	✗
(I,IV)	✗	✗	✗	✗
(I,V)	✗	✗	✗	✗
(I,VI)	✗	✗	✗	✗
(II,I)	✗	✓	✓	✓
(II,III)	✗	✓	✓	✓
(II,IV)	✗	✓	✓	✓
(II,V)	✗	✓	✓	✓
(II,VI)	✗	✓	✓	✓

(r,s)	$\lvert\Omega\rvert$			
	100	600	900	1500
(III,I)	✓	✗	✗	✗
(III,II)	✓	✗	✗	✗
(III,IV)	✗	✗	✗	✗
(III,V)	✓	✗	✗	✗
(III,VI)	✗	✗	✗	✗
(IV,I)	✓	✗	✗	✗
(IV,II)	✓	✗	✗	✗
(IV,III)	✗	✗	✗	✗
(IV,V)	✓	✗	✗	✗
(IV,VI)	✗	✗	✗	✗

(r,s)	$\lvert\Omega\rvert$			
	100	600	900	1500
(V,I)	✗	✗	✗	✗
(V,II)	✓	✗	✗	✗
(V,III)	✗	✗	✗	✗
(V,IV)	✗	✗	✗	✗
(V,VI)	✗	✗	✗	✗
(VI,I)	✓	✗	✗	✗
(VI,II)	✓	✗	✗	✗
(VI,III)	✗	✗	✗	✗
(VI,IV)	✗	✗	✗	✗
(VI,V)	✓	✗	✗	✗

(b) Model I

Table S3: Analogous results to Table 3 for $\varepsilon = 0.05$ but with a time limit of $time = 41000$ in Algorithm S1.

References

[1] Ahmed, S., Luedtke, J., Song, Y., Xie, W., 2017. Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs. Mathematical Programming 162, 51–81. doi:10.1007/s10107-016-1029-z.

[2] Singh, B., Knueven, B., 2021. Lagrangian relaxation based heuristics for a chance-constrained optimization model of a hybrid solar-battery storage system. Journal of Global Optimization 80, 965–989. doi:10.1007/s10898-021-01041-y.

[3] Singh, B., Morton, D.P., Santoso, S., 2018. An adaptive model with joint chance constraints for a hybrid wind-conventional generator system. Computational Management Science 15, 563—-582. doi:10.1007/s10287-018-0309-x.