EC - Energy and Reserve Dispatch with Distributionally Robust Joint Chance Constraints

Christos Ordoudis, Viet Anh Nguyen, Daniel Kuhn, Pierre Pinson

This document serves as an electronic companion (EC) for the paper "Energy and Reserve Dispatch with Distributionally Robust Joint Chance Constraints". It contains the proofs of the propositions in the original manuscript, a list of nomenclature, the formulation of the collective optimization model described in Section V.B. of the paper, the data for the IEEE 24-bus Reliability Test System, the procedure to generate the wind power data, the parameters related to the simulations for the numerical results and additional results for the Bonferroni and CVaR approximation approach.

1. Proofs of Propositions 1 and 2

Proof of Proposition 1. Using standard duality techniques, each worst-case CVaR in

$$\Omega_{\mathrm{BC}} \triangleq \left\{(x,Y) : \max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-}\operatorname{CVaR}_{\epsilon_k}\big[a_k(Y)^\top \xi - b_k(x)\big] \le 0 \quad \forall k \le K\right\} \tag{1}$$

can be rewritten as

$$\begin{aligned}
\max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-}\operatorname{CVaR}_{\epsilon_k}\big[a_k(Y)^\top \xi - b_k(x)\big] &= \max_{\mathbb{P}\in\mathcal{P}} \min_{\tau\in\mathbb{R}} \left\{\tau + \frac{1}{\epsilon_k}\mathbb{E}^{\mathbb{P}}\left[\left(a_k(Y)^\top\xi - b_k(x) - \tau\right)^+\right]\right\} && \text{(2a)}\\
&= \max_{\mathbb{P}\in\mathcal{P}} \min_{\tau\in\mathbb{R}} \left\{\mathbb{E}^{\mathbb{P}}\left[\max\left\{\tau, \frac{1}{\epsilon_k}\left(a_k(Y)^\top\xi - b_k(x)\right) + \left(1 - \frac{1}{\epsilon_k}\right)\tau\right\}\right]\right\} && \text{(2b)}\\
&\le \min_{\tau\in\mathbb{R}} \left\{\max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[\max\left\{\tau, \frac{1}{\epsilon_k}\left(a_k(Y)^\top\xi - b_k(x)\right) + \left(1 - \frac{1}{\epsilon_k}\right)\tau\right\}\right]\right\}, && \text{(2c)}
\end{aligned}$$

where (2a) uses the definition of CVaR from [1, Theorem 1], and the inequality in (2c) follows from the minimax inequality. Notice that the objective function of the max-min problem (2b) is convex in τ and linear in $\mathbb{P}$. Furthermore, by using [2, Theorem 7.12, (ii)], one can show that the ambiguity set $\mathcal{P}$ is weakly compact. As a result, [3, Theorem 4.2] implies that the inequality (2c) holds in fact as an equality. Because the integrand inside the expectation constitutes a pointwise maximum of two affine functions in ξ, [4, Corollary 5.1, (i)] allows us to reformulate the worst-case expectation in (2c) as a finite convex minimization problem, and thus we can rewrite the worst-case CVaR as

$$\max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-}\operatorname{CVaR}_{\epsilon_k}\big[a_k(Y)^\top \xi - b_k(x)\big] = \left\{\begin{array}{lll}
\min\limits_{\tau_k,\lambda_k,s_k,\gamma_k} & \lambda_k\rho + N^{-1}\sum_{i=1}^N s_{ik} & \\
\text{s. t.} & \tau_k \le s_{ik} & \forall i \le N\\
& a_k(Y)^\top\widehat{\xi}_i - b_k(x) + (\epsilon_k - 1)\tau_k & \\
& \quad +\epsilon_k\gamma_{ik}^\top(h - H\widehat{\xi}_i) \le \epsilon_k s_{ik} & \forall i \le N\\
& \|\epsilon_k H^\top\gamma_{ik} - a_k(Y)\|_* \le \epsilon_k\lambda_k & \forall i \le N\\
& \gamma_{ik} \in \mathbb{R}_+^{2W} & \forall i \le N\\
& \tau \in \mathbb{R}^K,\ \lambda \in \mathbb{R}^K,\ s \in \mathbb{R}^{N\times K}. &
\end{array}\right. \tag{3}$$

Consequently, Ω_{BC} admits an explicit conic representation, which is obtained by substituting the reformulation (3) into the feasible set (1). The claim then follows. $\square$

Proof of Proposition 2. By [1, Theorem 1], the worst-case CVaR appearing in the feasible set

$$\Omega_{\mathrm{C}}(\delta) \triangleq \left\{(x,Y) : \max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{- CVaR}_{\epsilon}\left[\max_{k\le K}\left\{\delta_k\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right] \le 0\right\}, \tag{4}$$

can be expressed as

$$\begin{aligned}
&\max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{- CVaR}_{\epsilon}\left[\max_{k\le K}\left\{\delta_k\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right] \\
&= \max_{\mathbb{P}\in\mathcal{P}} \min_{\tau\in\mathbb{R}} \left\{\tau + \frac{1}{\epsilon}\mathbb{E}^{\mathbb{P}}\left[\left(\max_{k\le K}\left\{\delta_k\left[a_k(Y)^\top\xi - b_k(x)\right]\right\} - \tau\right)^+\right]\right\} \\
&= \max_{\mathbb{P}\in\mathcal{P}} \min_{\tau\in\mathbb{R}} \left\{\mathbb{E}^{\mathbb{P}}\left[\max\left\{\tau, \left(\max_{k\le K}\left\{\frac{\delta_k}{\epsilon_k}\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right) + \left(1-\frac{1}{\epsilon}\right)\tau\right\}\right]\right\} \\
&\le \min_{\tau\in\mathbb{R}} \left\{\max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[\max\left\{\tau, \left(\max_{k\le K}\left\{\frac{\delta_k}{\epsilon_k}\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right) + \left(1-\frac{1}{\epsilon}\right)\tau\right\}\right]\right\}.
\end{aligned} \tag{5a}$$

Using the same reasoning as in the proof of Proposition 1, one can show that the inequality in (5a) holds in fact as an equality. The integrand inside the expectation operator constitutes a pointwise maximum of $K+1$ affine functions. As such, one can use [4, Corollary 5.1, (i)] to reformulate the worst-case CVaR over the Wasserstein ambiguity set $\mathcal{P}$ as

$$\begin{aligned}
&\max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{- CVaR}_{\epsilon}\left[\max_{k\le K}\left\{\delta_k\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right] = \\
&\left\{\begin{array}{lll}
\min\limits_{\tau,\lambda,s,\gamma} & \lambda\rho + \frac{1}{N}\sum_{i=1}^N s_i & \\
\text{s. t.} & \tau\in\mathbb{R},\ \lambda\in\mathbb{R},\ s\in\mathbb{R}^N,\ \gamma_{ik}\in\mathbb{R}_+^{2W} & \forall i\le N, k\le K \\
& \tau\le s_i & \forall i\le N \\
& \delta_k\left[a_k(Y)^\top\widehat{\xi}_i - b_k(x)\right] + (\epsilon-1)\tau & \\
& \qquad +\epsilon\gamma_{ik}^\top(h - H\widehat{\xi}_i) \le \epsilon s_i & \forall i\le N, k\le K \\
& \|\epsilon H^\top\gamma_{ik} - \delta_k a_k(Y)\|_* \le \epsilon\lambda & \forall i\le N, k\le K.
\end{array}\right.
\end{aligned} \tag{6}$$

Substituting (6) into the feasible set (4) completes the proof. □

2. Nomenclature

In Table 1 we list the symbols used in the original paper and describe each of them briefly.

Table 1. Nomenclature

Symbol	Description
y_1	First-stage power dispatch of conventional power plants
Y	Coefficient matrix of the affine policy of the conventional power plants' real time adjustments
r_-	Downward reserve capacity of conventional power plants
r_+	Upward reserve capacity of conventional power plants
μ	Predicted power production of the wind farms
ξ	Random variable with zero mean
C	Diagonal matrix of wind farm capacities
Q	Matrix of power transfer distribution factors
$\overline{r}$	Maximum reserve capacity offered by conventional power plants
$\overline{y}$	Maximum power production of conventional power plants
$\underline{y}$	Minimum power production of conventional power plants
$\overline{f}$	Capacity limits of transmission lines
$\overline{q}$	Capacity limits of pipelines
Φ	Matrix of gas transfer distribution factors
d	Electricity demands
c	Variable costs of conventional power plants
c_-	Cost of reserving downward capacity
c_+	Cost of reserving upward capacity
ϵ	Violation probabilities of joint chance constraints
$\mathbb{P}$	True probability distribution of the random vector ξ
$\widehat{\mathbb{P}}_N$	Empirical distribution on the training samples
N	Number of training samples
M	Number of test samples
$\mathcal{P}$	Wasserstein ambiguity set
ρ	Wasserstein radius
δ	Vector of scaling parameters for the optimized CVaR approximation
η	Minimum relative improvement per iteration in the sequential optimization algorithm
v	Auxiliary slack variables for the joint chance constraints
M	Big-M constant

3. Collective optimization model

Consider the Wasserstein ambiguity set $\mathcal{P}$ with a radius ρ defined as

$$\mathcal{P} \triangleq \left\{\mathbb{P} \in \mathcal{M}(\Xi) : \mathbb{W}(\mathbb{P}, \widehat{\mathbb{P}}_N) \leq \rho\right\}.$$

In the remainder of this section we provide a formal description of the collective optimization model and how it can be approximated by a tractable convex program for $\rho > 0$ and $\rho = 0$.

3.1. Strictly positive radius $\rho > 0$

We consider the following hybrid optimization where the objective is to minimize the worst-case expected cost subject to all probability measures in the Wasserstein ambiguity set, while the constraints are robust constraints. We assume that the uncertain factors ξ varies in the box uncertainty set $\Xi = \{\xi \in \mathbb{R}^W : -\mu \leq \xi \leq e - \mu\}$ where e denotes the vector of ones.

The collective optimization model becomes

$$\begin{array}{lll}
\min\limits_{\substack{y_1, r_+, r_- \\ y_2(\cdot), r(\cdot), l(\cdot), w(\cdot)}} & c^\top y_1 + c_+^\top r_+ + c_-^\top r_- + \max\limits_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[c^\top y_2(\xi) + c_l e^\top l(\xi) + c_w e^\top w(\xi) + c_r e^\top r(\xi)] & \\
\text{s.t.} & 0 \leq r_+ \leq \overline{r}, \quad 0 \leq r_- \leq \overline{r} & \\
& \underline{y} \leq y_1 - r_-, \quad y_1 + r_+ \leq \overline{y} & \\
& e^\top C(\mu + \xi) - e^\top w(\xi) + e^\top (y_1 + y_2(\xi)) = e^\top (d - l(\xi)) & \mathbb{P}\text{-a.s.} \\
& -(r_- + r(\xi)) \leq y_2(\xi) \leq r_+ & \mathbb{P}\text{-a.s.} \\
& -\overline{f} \leq Q^{\mathrm{g}}(y_1 + y_2(\xi)) + Q^{\mathrm{w}} C(\mu + \xi - w(\xi)) - Q^{\mathrm{d}}(d - l(\xi)) \leq \overline{f} & \mathbb{P}\text{-a.s.} \\
& 0 \leq \Phi(y_1 + y_2(\xi)) \leq \overline{q} & \mathbb{P}\text{-a.s.} \\
& 0 \leq r(\xi), \quad 0 \leq w(\xi) \leq C(\mu + \xi), \quad 0 \leq l(\xi) \leq d & \mathbb{P}\text{-a.s.,}
\end{array}$$

where we emphasize that the constraints are now satisfied for any $\xi \in \Xi$. By applying the linear decision rule approximation, i.e., by constraining $y_2(\xi) = Y\xi, r(\xi) = Y_2\xi, w(\xi) = Y_3\xi, l(\xi) = Y_4\xi$ for some $Y \in \mathbb{R}^{G\times W}, Y_2 \in \mathbb{R}^{G\times W}, Y_3 \in \mathbb{R}^{W\times W}$ and $Y_4 \in \mathbb{R}^{D\times W}$, we solve the following linear decision rule approximation problem

$$\begin{array}{lll}
\min\limits_{\substack{y_1, r_+, r_- \\ Y, Y_2, Y_3, Y_4}} & c^\top y_1 + c_+^\top r_+ + c_-^\top r_- + \max\limits_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[(c^\top Y + c_l e^\top Y_4 + c_w e^\top Y_3 + c_r e^\top Y_2)\xi] & \\
\text{s.t.} & 0 \leq r_+ \leq \overline{r}, \quad 0 \leq r_- \leq \overline{r} & \\
& \underline{y} \leq y_1 - r_-, \quad y_1 + r_+ \leq \overline{y} & \\
& e^\top C(\mu + \xi) - e^\top Y_3\xi + e^\top (y_1 + Y\xi) = e^\top (d - Y_4\xi) & \forall \xi \in \Xi \\
& -(r_- + Y_2\xi) \leq Y\xi \leq r_+ & \forall \xi \in \Xi \\
& -\overline{f} \leq Q^{\mathrm{g}}(y_1 + Y\xi) + Q^{\mathrm{w}} C(\mu + \xi - Y_3\xi) - Q^{\mathrm{d}}(d - Y_4\xi) \leq \overline{f} & \forall \xi \in \Xi \\
& 0 \leq \Phi(y_1 + Y\xi) \leq \overline{q} & \forall \xi \in \Xi \\
& 0 \leq Y_2\xi, \quad 0 \leq Y_3\xi \leq C(\mu + \xi), \quad 0 \leq Y_4\xi \leq d & \forall \xi \in \Xi.
\end{array}$$

The above program can be further re-expressed in the compact form as

$$\begin{array}{lll}
\min\limits_{\substack{y_1, r_+, r_- \\ Y, Y_2, Y_3, Y_4}} & c^\top y_1 + c_+^\top r_+ + c_-^\top r_- + \max\limits_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}[(c^\top Y + c_l e^\top Y_4 + c_w e^\top Y_3 + c_r e^\top Y_2)\xi] & \\
\text{s.t.} & 0 \leq r_+ \leq \overline{r}, \quad 0 \leq r_- \leq \overline{r}, \quad \underline{y} \leq y_1 - r_-, \quad y_1 + r_+ \leq \overline{y} & \\
& e^\top y_1 + e^\top C\mu = e^\top d, \quad e^\top Y + e^\top C - e^\top Y_3 + e^\top Y_4 = 0 & \\
& [A^j(Y) + A_2^j(Y_2) + A_3^j(Y_3) + A_4^j(Y_4)]\xi \leq b^j(x) & \forall \xi \in \Xi, \forall j \in \mathcal{J} \\
& 0 \leq Y_2\xi, \quad 0 \leq Y_3\xi \leq C(\mu + \xi), \quad 0 \leq Y_4\xi \leq d & \forall \xi \in \Xi,
\end{array}$$

where $\mathcal{J}$ captures the indices $\mathcal{J} \triangleq \{\text{gen}, \text{grid}, \text{gas}\}$. In addition, the definition of $A^j(Y)$ and $b^j(x)$ are similar to the main paper, i.e., for the generator constraints, we have

$$A^{\text{gen}}(Y) \triangleq \begin{bmatrix} Y \\ -Y \end{bmatrix}, \quad b^{\text{gen}}(x) \triangleq \begin{bmatrix} r_+ \\ -r_- \end{bmatrix}.$$

For the line capacity constraints, we have

$$A^{\text{grid}}(Y) \triangleq \begin{bmatrix} Q^{\text{g}}Y + Q^{\text{w}}C \\ -Q^{\text{g}}Y - Q^{\text{w}}C \end{bmatrix}, \quad b^{\text{grid}}(x) \triangleq \begin{bmatrix} \overline{f} - Q^{\text{w}}C\mu + Q^{\text{d}}d - Q^{\text{g}}y_1 \\ \overline{f} + Q^{\text{w}}C\mu - Q^{\text{d}}d + Q^{\text{g}}y_1 \end{bmatrix},$$

and for the pipeline capacity constraints, we have

$$A^{\text{gas}}(Y) \triangleq \begin{bmatrix} \Phi Y \\ -\Phi Y \end{bmatrix}, \quad b^{\text{gas}}(x) \triangleq \begin{bmatrix} \overline{q} - \Phi y_1 \\ \Phi y_1 \end{bmatrix}.$$

For the reserve increase variables, we have the corresponding mappings

$$A_2^{\text{gen}} = \begin{bmatrix} \mathbf{0} \\ -Y_2 \end{bmatrix}, \quad A_2^{\text{grid}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad A_2^{\text{gas}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

In a similar manner, the mappings for the wind spillage variables and load shedding variables are

$$\begin{aligned} A_3^{\text{gen}} &= \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, A_3^{\text{grid}} = \begin{bmatrix} -Q^{\text{w}}Y_3 \\ Q^{\text{w}}Y_3 \end{bmatrix}, A_3^{\text{gas}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \\ A_4^{\text{gen}} &= \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, A_4^{\text{grid}} = \begin{bmatrix} Q^{\text{d}}Y_4 \\ -Q^{\text{d}}Y_4 \end{bmatrix}, A_4^{\text{gas}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \end{aligned}$$

By definition, the uncertainty set Ξ has 2^W vertices, we replace each robust constraint by 2^W individual constraints for each vertex ξ_v of Ξ, $v = 1, \ldots, 2^W$. The objective function is also reformulated accordingly. The robust problem can be re-expressed as a linear distributionally robust optimization program

$$\begin{array}{lll} \min\limits_{\substack{y_1, r_+, r_- \\ Y, Y_2, Y_3, Y_4}} & c^\top y_1 + c_+^\top r_+ + c_-^\top r_- + \max\limits_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}}[(c^\top Y + c_l e^\top Y_4 + c_w e^\top Y_3 + c_r e^\top Y_2)\xi] & \\ \text{s. t.} & 0 \leq r_+ \leq \overline{r}, \quad 0 \leq r_- \leq \overline{r}, \quad \underline{y} \leq y_1 - r_-, \quad y_1 + r_+ \leq \overline{y} & \\ & e^\top y_1 + e^\top C\mu = e^\top d, \quad e^\top Y + e^\top C - e^\top Y_3 + e^\top Y_4 = 0 & \\ & [A^j(Y) + A_2^j(Y_2) + A_3^j(Y_3) + A_4^j(Y_4)]\xi_{\text{v}} \leq b^j(x) & \forall \text{v} = 1, \ldots, 2^W, \forall j \in \mathcal{J} \\ & 0 \leq Y_2\xi_{\text{v}}, \quad 0 \leq Y_3\xi_{\text{v}} \leq C(\mu + \xi_{\text{v}}), \quad 0 \leq Y_4\xi_{\text{v}} \leq d & \forall \text{v} = 1, \ldots, 2^W. \end{array}$$

The worst-case expected value in the objective function can be reformulated using the result of the Section IV.A in the main paper. The resulting problem is a tractable conic program.

3.2. Zero radius $\rho = 0$

In case the radius of the Wasserstein ball is set to zero $\rho = 0$, we solve the following sample average approximation problem

$$\begin{array}{lll} \min\limits_{\substack{y_1, r_+, r_- \\ y_2, r, l, w}} & c^\top y_1 + c_+^\top r_+ + c_-^\top r_- + N^{-1} \sum_{i=1}^N \left(c^\top y_{2i} + c_r e^\top r_i + c_l e^\top l_i + c_w e^\top w_i\right) & \\ \text{s. t.} & 0 \leq r_+ \leq \overline{r}, \quad 0 \leq r_- \leq \overline{r}, \quad \underline{y} \leq y_1 - r_-, \quad y_1 + r_+ \leq \overline{y} & \\ & 0 \leq y_1 + y_{2i} \leq \overline{y}, \ -(\widehat{r}_- + r_i) \leq y_{2i} \leq \widehat{r}_+ & \forall i = 1, \ldots, N \\ & e^\top y_1 + e^\top C\mu = e^\top d & \\ & e^\top y_{2i} + e^\top (C\widehat{\xi}_i - w_i) + e^\top l_i = 0 & \forall i = 1, \ldots, N \\ & r_- \leq y_{2i} \leq r_+ & \forall i = 1, \ldots, N \\ & -\overline{f} \leq Q^{\text{g}}(y_1 + y_{2i}) + Q^{\text{w}}(C\mu + C\widehat{\xi}_i - w_i) - Q^{\text{d}}(d - l_i) \leq \overline{f} & \forall i = 1, \ldots, N \\ & 0 \leq \Phi(y_1 + y_{2i}) \leq \overline{q} & \forall i = 1, \ldots, N \\ & 0 \leq r_i \leq y_1 - r_-, 0 \leq l_i \leq d, \ 0 \leq w_i \leq C(\mu + \widehat{\xi}_i) & \forall i = 1, \ldots, N. \end{array}$$

We emphasize that the second-stage decisions (y_2, r, l, w) are separated for each scenario $\widehat{\xi}_i, i = 1, \ldots, N$. The above program is a linear program, and thus is implementable and solvable by standard package.

The out-of-sample performance of the collective optimization model is evaluated on the test sample $\widehat{\xi}_i, i = N+1, \ldots, N+M$ in a similar procedure as in Section V of the main manuscript.

4. Data for the IEEE 24-bus RTS

The 24-bus power system is illustrated in Figure 1. The slack bus of the system is bus 13.

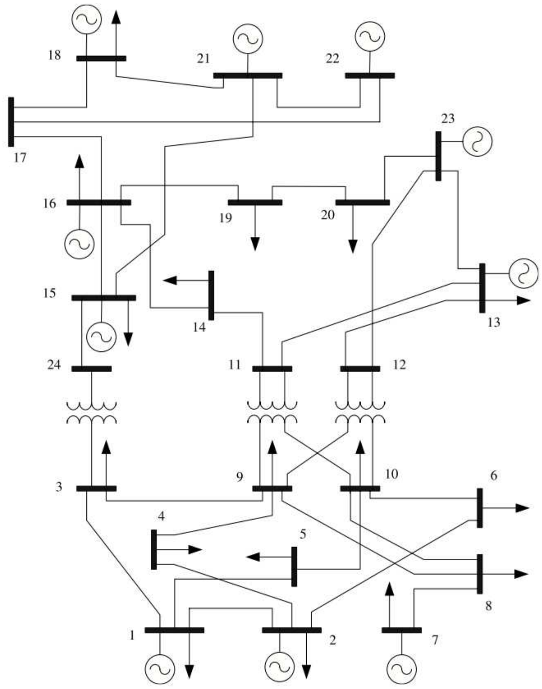

Figure 1. 24-bus power system – Single area RTS-96

The generating units of the power system offer a single block of energy, up and down reserve capacity. Table 2 provides the technical data of generating units, including the location on the power system, the relevant costs, as well as the conversion factor and the connection with the corresponding pipelines for the gas-fired power plants. There are three pipelines that the gas-fired power plants (GFPPs) are connected to, and the capacity of the gas pipelines is adopted from capacity of pipelines are based on the test case used in [5]: Pipeline 1 has a capacity of 10,000 kcf, Pipeline 2 has a capacity of 5,500 kcf and Pipeline 3 has a capacity of 7,000 kcf (kcf: 1,000 cubic feet).

Table 2. Technical Data of Generating Units

Unit #	Bus	$\overline{y}$ (MW)	$\underline{y}$ (MW)	$\overline{r}$ (MW)	c (\$/MWh)	c_+ (\$/MW)	c_- (\$/MW)	Φ (kcf/MWh)	Pipeline
1	1	152	0	60.8	17.5	3.5	3.5	12.65	1
2	2	152	0	60.8	20	4	4	13.45	3
3	7	300	0	120	15	3	3	-	-
4	13	591	0	236.4	27.5	5.5	5.5	-	-
5	15	60	0	24	30	6	6	11.12	2
6	15	155	0	62	22.5	4.5	4.5	-	-
7	16	155	0	62	25	5	5	14.88	1
8	18	400	0	160	5	1	1	-	-
9	21	400	0	160	7.5	1.5	1.5	-	-
10	22	300	0	120	32.5	6.5	6.5	-	-
11	23	310	0	124	10	2	2	16.8	2
12	23	350	0	140	12.5	2.5	2.5	15.6	3

Table 3 presents the bus location of the loads, as well as the load at each bus as a percentage of the total system demand. The total electricity demand over the whole grid is 2,650 MWh. The data concerning the electricity transmission lines is given in Table 4. These lines are characterized by the buses that it connects, as well as the reactance and the transmission capacity.

TABLE 3. Bus Location and Distribution of the Total System Demand

Load #	Bus	% of system load	Load #	Bus	% of system load
1	1	3.8	10	10	6.8
2	2	3.4	11	13	9.3
3	3	6.3	12	14	6.8
4	4	2.6	13	15	11.1
5	5	2.5	14	16	3.5
6	6	4.8	15	18	11.7
7	7	4.4	16	19	6.4
8	8	6	17	20	4.5
9	9	6.1			

TABLE 4. Reactance and Capacity of Transmission Lines

From	To	Reactance (p.u.)	Capacity (MW)	From	To	Reactance (p.u.)	Capacity (MW)
1	2	0.0146	175	11	13	0.0488	500
1	3	0.2253	175	11	14	0.0426	500
1	5	0.0907	400	12	13	0.0488	500
2	4	0.1356	175	12	23	0.0985	500
2	6	0.205	175	13	23	0.0884	500
3	9	0.1271	400	14	16	0.0594	1000
3	24	0.084	200	15	16	0.0172	500
4	9	0.111	175	15	21	0.0249	1000
5	10	0.094	400	15	24	0.0529	500
6	10	0.0642	400	16	17	0.0263	500
7	8	0.0652	600	16	19	0.0234	500
8	9	0.1762	175	17	18	0.0143	500
8	10	0.1762	175	17	22	0.1069	500
9	11	0.084	200	18	21	0.0132	1000
9	12	0.084	200	19	20	0.0203	1000
10	11	0.084	200	20	23	0.0112	1000
10	12	0.084	200	21	22	0.0692	500

There are 6 wind farms of nominal capacity 250 MW connected to the electricity grid at bus 1, 2, 11, 12, 12, 16 respectively. The data in Tables 2-4 are adopted from [5, 6]. The marginal cost of power production c in Table 2 are modified such that each power plant has a different cost to prioritize the order of dispatch.

5. Wind power data generation for the simulation of numerical results

The data consists of wind power forecast errors that is generated with the same method and the historical data given in [7]. The wind power data are provided by the Australian System Operator [8], which consists of wind power generation recordings from 22 wind farms in the southeastern Australia. The complete dataset can be found in [9]. Data from 2012 and 2013 are normalized by the nominal power of the corresponding wind farm which results in data being in the range of $[0, 1]$. From the csv file that is publicly available in [9], we have chosen the 6 wind farms presented in Table 5.

Table 5. Wind Farm Location and Name tag in csv file

WF	Bus	Name tag in csv
1	1	CAPTL_WF
2	2	CATHROCK
3	11	CULLRGWF
4	12	LKBONNY1
5	12	MTMILLAR
6	16	STARHLWF

The wind power data are generated with the following procedure:

(1) Load the 2 years of raw data for the 6 chosen wind farms.
(2) The data is projected to the range $[0.01, 0.99]$ and then a logit-normal transformation is performed (see [7, Equation (1)]).
(3) The sample mean $\widehat{\mu}$ and sample covariance matrix $\widehat{\Sigma}$ are calculated from the transformed data.
(4) We generate the required number of independent and identically distributed random samples by assuming that the transformed variables are normally distributed with distribution $\mathcal{N}(\widehat{\mu}, \widehat{\Sigma})$.
(5) Apply the inverse of logit-normal transformation (see [7, Equation (2)]) on the samples generated in the previous step.

We refer the reader to [7] for a detailed presentation of the method and additional insights in very-short-term probabilistic wind power forecasting. Note that the procedure can be found also in the source code provided online.

6. Data for the simulation of numerical results

The discretized search space of ρ for the Bonferroni and CVaR approximation is confined to the vector $\{0,\ \text{linspace}(10^{-4}, 24 \cdot 10^{-4}, 23)\}$, while the discretized search space for the optimized CVaR approximation is $\{0,\ 10^{-4},\ \text{linspace}(10^{-3}, 10^{-2}, 23)\}$. The MATLAB command linspace(a, b, n) generates n points evenly distributed on the interval $[a, b]$.

Table 6 presents the values of the parameters related to the algorithm utilized in the optimized CVaR approximation of distributionally robust joint chance constraints.

Table 6. Parameter for the sequential algorithm to solve the optimized CVaR approximation problem

Parameter	Value
$\bar{t}$	40
η	0.1
M	10^6

7. Additional Results for Bonferroni and CVaR approximation approach

Figures 2 and 3 illustrate the impact of Wasserstein radius ρ on the expected value and interquantile range between the 10$^{\text{th}}$ and 90$^{\text{th}}$ quantile of the empirical cost $\widehat{\mathcal{C}}$ using reoptimization. These two figures are the corresponding counterparts of Figure 2 and 3 in the original manuscript for the Bonferroni and CVaR approximation approach.

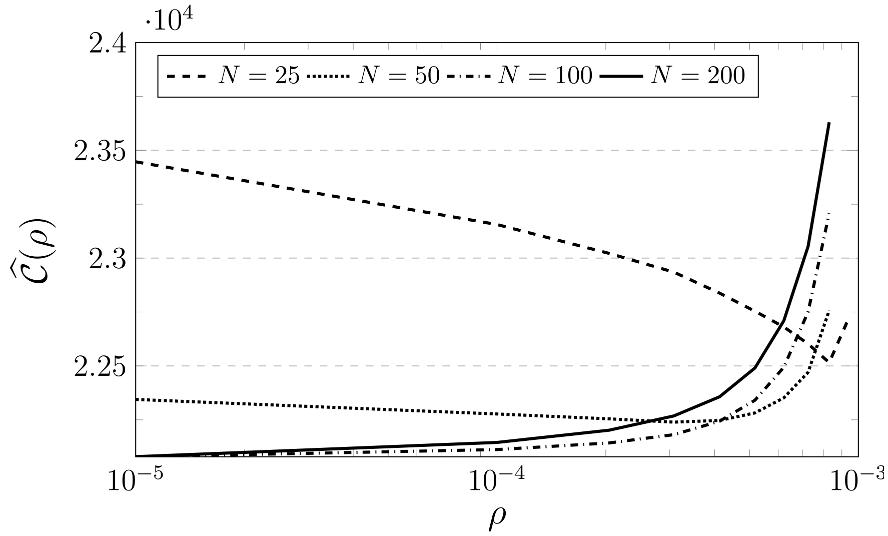

Figure 2. Average out-of-sample cost $\widehat{\mathcal{C}}(\rho)$ using the Bonferroni and CVaR approximation and reoptimization.

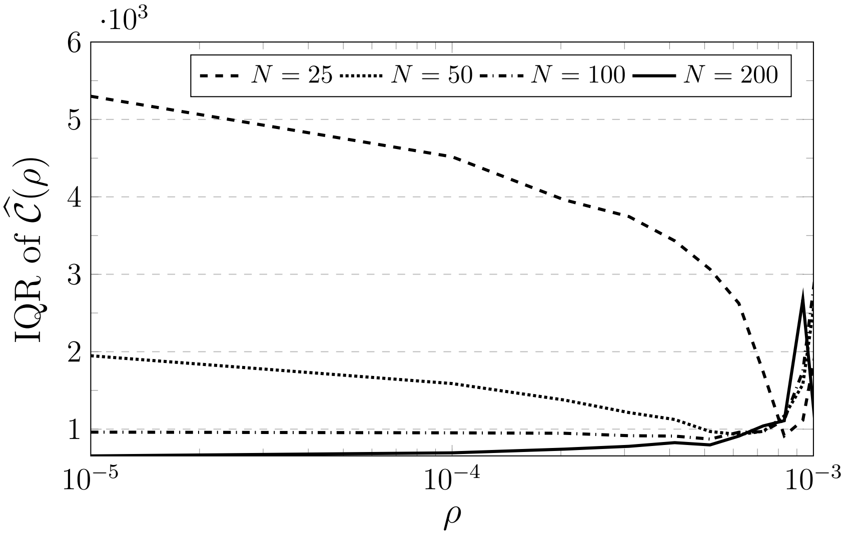

FIGURE 3. Interquantile range between the empirical 10% and 90% quantiles of $\widehat{\mathcal{C}}(\rho)$ using the Bonferroni and CVaR approximation and reoptimization.

REFERENCES

[1] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[2] C. Villani, *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[3] M. Sion, "On general minimax theorems," *Pacific J. Math.*, vol. 8, no. 1, pp. 171–176, 1958.

[4] P. Mohajerin Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, no. 1, pp. 115–166, 2018.

[5] C. Ordoudis, P. Pinson, and J. M. Morales, "An integrated market for electricity and natural gas systems with stochastic power producers," *Eur. J. Oper. Res.*, vol. 272, no. 2, pp. 642 – 654, 2019.

[6] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Problems*. International Series in Operations Research & Management Science, Springer US, 2014.

[7] J. Dowell and P. Pinson, "Very-short-term probabilistic wind power forecasts by sparse vector autoregression," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 763–770, 2016.

[8] Australian Energy Market Operator, "AEMO wind power data." http://www.aemo.com.au/, 2018. Online; Accessed: 26.11.2018.

[9] J. Browell, "AEMO wind power data." http://www.jethrobrowell.com/data-and-code.html, 2018. Online; Accessed: 26.11.2018.