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, the data for the IEEE 24-bus RTS, the procedure to generate the wind power data, parameters related to the simulations for the numerical results and additional results.

1. Proofs of Proposition 1 and Proposition 2

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

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

can be rewritten as

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

where in (1a) we use the definition of CVaR in [1, Theorem 1], and the inequality in (1c) comes from weak duality. Notice that the objective function of (1b) is convex in τ and linear in $\mathbb{P}$. Furthermore, by the result of [2, Theorem 7.12, part (ii)], one can show that the ambiguity set $\mathcal{P}$ is weakly compact. As a result, [3, Theorem 4.2] implies that sup-inf equals inf-sup, and (1c) holds with equality. Because the expression inside the expectation is a pointwise maximum of two affine functions in terms of ξ, [4, Corollary 5.1, part (i)] applies and the worst-case expectation over the probability measure $\mathbb{P}$ in (1c) admits a reformulation in the dual form, and we can rewrite each worst-case CVaR value as

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

and Ω_{BC} can be reformulated as the intersection of K feasible sets by substituting the value of (2) into (1). This completes the proof. □

Proof of Proposition 2. For an individual chance constraint from the feasible set

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

the worst case CVaR can be expressed based on definition in [1, Theorem 1] as

$$\max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-}\operatorname{CVaR}_{\epsilon}\left[\max_{k\leq K}\left\{\delta_k\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right] \tag{3a}$$

$$= \max_{\mathbb{P}\in\mathcal{P}} \min_{\tau\in\mathbb{R}} \left\{\tau + \frac{1}{\epsilon}\mathbb{E}^{\mathbb{P}}\left[\left(\max_{k\leq K}\left\{\delta_k\left[a_k(Y)^\top\xi - b_k(x)\right]\right\} - \tau\right)^+\right]\right\} \tag{3b}$$

$$= \max_{\mathbb{P}\in\mathcal{P}} \min_{\tau\in\mathbb{R}} \left\{\mathbb{E}^{\mathbb{P}}\left[\max\left\{\tau, \left(\max_{k\leq K}\left\{\frac{\delta_k}{\epsilon_k}\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right) + \left(1 - \frac{1}{\epsilon}\right)\tau\right\}\right]\right\} \tag{3c}$$

$$\leq \min_{\tau\in\mathbb{R}} \left\{\max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[\max\left\{\tau, \left(\max_{k\leq K}\left\{\frac{\delta_k}{\epsilon_k}\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right) + \left(1 - \frac{1}{\epsilon}\right)\tau\right\}\right]\right\}. \tag{3d}$$

Using the same reasoning as in the proof of Proposition 1, the inequality in (3d) holds with equality. The expression inside the expectation is the pointwise maximum of $K+1$ affine functions. Thus, we can reformulate the worst-case CVaR based on [4, Corollary 5.1, part (i)] as

$$\max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-}\operatorname{CVaR}_{\epsilon}\left[\max_{k\leq K}\left\{\delta_k\left[a_k(Y)^\top\xi - b_k(x)\right]\right\}\right] =$$
$$\left\{\begin{array}{lll} \displaystyle\min_{\tau,\lambda,s,\gamma} & \lambda\rho + \frac{1}{N}\sum_{i=1}^N s_i & \\ \text{s. t.} & \tau\in\mathbb{R},\ \lambda\in\mathbb{R},\ s\in\mathbb{R}^N,\ \gamma_{ik}\in\mathbb{R}_+^{2W} & \forall i\leq N, k\leq K \\ & \tau \leq s_i & \forall i \leq N \\ & \delta_k\left[a_k(Y)^\top\widehat{\xi}_i - b_k(x)\right] + (\epsilon-1)\tau & \\ & \quad +\epsilon\gamma_{ik}^\top(h - H\widehat{\xi}_i) \leq \epsilon s_i & \forall i\leq N, k\leq K \\ & \|\epsilon H^\top\gamma_{ik} - \delta_k a_k(Y)\|_* \leq \epsilon\lambda & \forall i\leq N, k\leq K. \end{array}\right. \tag{4}$$

Replacing (4) into (3) completes the proof. □

2. Nomenclature

In Table 1, we present the symbols used in the original paper and the description for each one of them.

Table 1. Nomenclature

Symbol	Description
y_1	First-stage power dispatch of conventional power plants
Y	Afine policy to approximate the recourse decisions of conventional power plants
r_-	First-stage downward reserve capacity of conventional power plants
r_+	First-stage upward reserve capacity of conventional power plants
μ	Predicted power production of wind farm
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 limit of transmission line
$\overline{q}$	Capacity limit of pipeline
Φ	Matrix of gas distribution factors (containing power conversion factor)
d	Electricity demand
c	Variable cost 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 random variable ξ
$\widehat{\mathbb{P}}_N$	Empirical distribution: uniform distribution on the training samples
N	Number of training samples drawn from $\mathbb{P}$
N'	Number of realizations in the testing dataset
$\mathcal{P}$	Ambiguity set
ρ	Wasserstein radius
δ	Vector of scaling parameters
η	Minimum relative improvement per iteration in iterative algorithm of Optimized CVaR approximation
v	Auxiliary slack variable in iterative algorithm of Optimized CVaR approximation
M	Big-M constant in iterative algorithm of Optimized CVaR approximation

3. Collective optimization model when $\rho > 0$

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. In the robust optimization approach, 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.

Consider the ambiguity set $\mathcal{P}$ defined as

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

with the radius $\rho > 0$. The collective optimization model becomes

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

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 approximation problem

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

which can be re-expressed in the compact form as

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

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, we have

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

For the wind spillage, we have

$$A_3^{\text{gen}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad A_3^{\text{grid}} = \begin{bmatrix} -Q^{\text{w}} Y_3 \\ Q^{\text{w}} Y_3 \end{bmatrix}, \quad A_3^{\text{gas}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

For the load shedding, we have

$$A_4^{\text{gen}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \quad A_4^{\text{grid}} = \begin{bmatrix} Q^{\text{d}} Y_4 \\ -Q^{\text{d}} Y_4 \end{bmatrix}, \quad A_4^{\text{gas}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

Because 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

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

where 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 out-of-sample performance evaluated on the test sample $\widehat{\xi}_i, i = N+1, \ldots, N+10^3$ is calculated in a similar procedure as in Section V of the main paper.

4. Collective optimization model when $\rho = 0$

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

$$\begin{array}{lll}
\min\limits_{y_1, r_+, r_-, y_2, r, l, w} & c^\top y_1 + c_+^\top r_+ + c_-^\top r_- + N^{-1} \sum_{i=1}^N \left(c^\top y_{1i} + c_r e^\top r_i + c_l e^\top l_i + c_w e^\top w_i \right) & \\
\text{s.t.} & 0 \le r_+ \le \overline{r}, \quad 0 \le r_- \le \overline{r}, \quad \underline{y} \le y_1 - r_-, \quad y_1 + r_+ \le \overline{y} & \\
& 0 \le y_1 + y_{1i} \le \overline{y}, \; -(\widehat{r}_- + r_i) \le y_{1i} \le \widehat{r}_+ & \forall i = 1, \ldots, N \\
& e^\top y_1 + e^\top C\mu = e^\top d & \\
& e^\top y_{1i} + e^\top (C\widehat{\xi}_i - w_i) + e^\top l_i = 0 & \forall i = 1, \ldots, N \\
& r_- \le y_{1i} \le r_+ & \forall i = 1, \ldots, N \\
& -\overline{f} \le Q^{\text{g}}(y_1 + y_{1i}) + Q^{\text{w}}(C\mu + C\widehat{\xi}_i - w_i) - Q^{\text{d}}(d - l_i) \le \overline{f} & \forall i = 1, \ldots, N \\
& 0 \le \Phi(y_1 + y_{1i}) \le \overline{q} & \forall i = 1, \ldots, N \\
& 0 \le r_i \le y_1 - r_-, 0 \le l_i \le d, \; 0 \le w_i \le C(\mu + \widehat{\xi}_i), & \forall i = 1, \ldots, N.
\end{array} \tag{5}$$

We emphasize that the second-stage decision is 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.

5. 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.

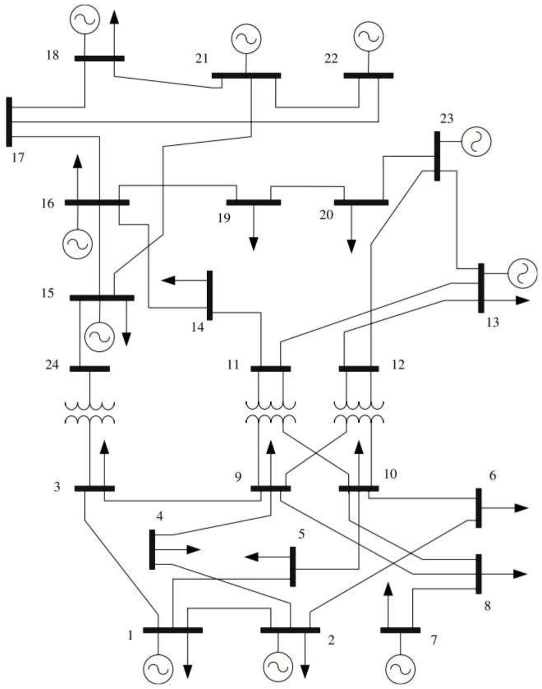

Figure 1. 24-bus power system – Single area RTS-96

Table 2 presents the generating units' data of the power system. The generating units offer a single block of energy, up and down reserve capacity. Table 2 provides the technical data of generating units, the costs, the location on the power system, 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 are connected. 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 is 2,650 MWh and the cost of load shedding is $1,000 /MWh.

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			

The transmission lines data is given in Table 4. The lines are characterized by the bus that are connected, as well as the reactance and the capacity of each line.

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 250 MW with different locations throughout the grid. The wind farms are connected at the 1, 2, 11, 12, 12, 16 bus. The data in Tables 2-4 are based on [5]. We have modified the marginal cost of power production to have different costs for each power plant. The value of power conversion factor and capacity of pipelines are based on the test case used in [6].

6. 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 can be found in [9], we have picked the 6 wind farms presented in Table 5.

Table 5. Wind Farm Location and Name tag in csv

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) The 2 years of data for the 6 wind farms is loaded.
(2) The data range is adjusted to [0.01,0.99] to permit the logit-normal transformation in the next step.
(3) A logit-normal transformation is performed based on [7, Equation (1)].
(4) The mean and covariance are calculated.
(5) We generate the required number of independent and identically distributed random samples by assuming the transformed variable to be normally distributed.
(6) The inverse of logit-normal transformation is applied based on [7, Equation (2)].

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.

7. Data for the simulation of numerical results

The discretized search space of ρ for the Combined Bonferroni and CVaR approximation is confined to the vector $\{0,\ \text{linspace}(10^{-4}, 24 \cdot 10^{-4}, 23)\}$, while the 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

8. Additional Results for Bonferroni and CVaR approximation approach

Figs. 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 reoptimized cost $\widehat{\mathcal{C}}$. For the Bonferroni and CVaR approximation, we have similar observations as the ones for Optimized CVaR approximation presented in the original manuscript.

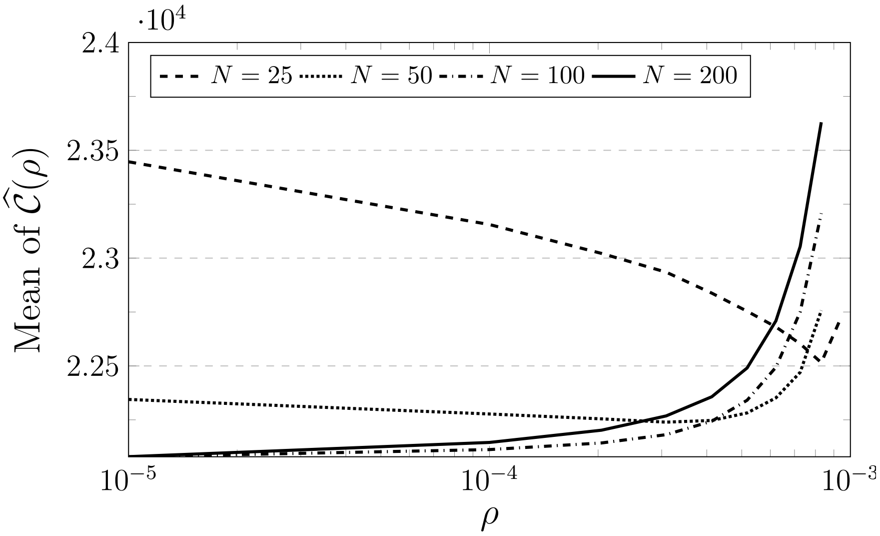

Figure 2. Average out-of-sample cost $\widehat{\mathcal{C}}(\rho)$ for the reoptimization approach using the Bonferroni and CVaR approximation.

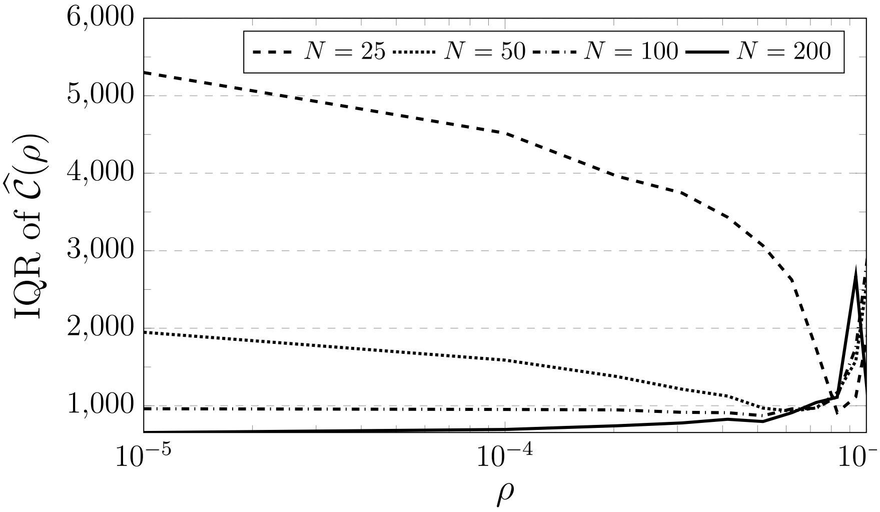

FIGURE 3. Interquantile range between the empirical 10% and 90% quantiles of $\widehat{\mathcal{C}}(\rho)$ for the reoptimization approach using the Bonferroni and CVaR approximation.

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.*, Jul 2017.

[5] 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.

[6] 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.*, 2018.

[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: 06.09.2018.

[9] J. Browell, "AEMO wind power data." http://www.jethrobrowell.com/data-and-code.html, 2018. Online; Accessed: 06.09.2018.