Online Companion - Enhanced Wasserstein Distributionally Robust OPF With Dependence Structure and Support Information

Adriano Arrigo[1], Jalal Kazempour[2], Zacharie De Grève[1],
Jean-François Toubeau[1], François Vallée[1]
[1] Electrical Power Engineering Unit, University of Mons, Mons, Belgium
{adriano.arrigo, zacharie.degreve, jean-francois.toubeau, francois.vallee}@umons.ac.be
[2] Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark
seykaz@dtu.dk

This paper serves as an electronic companion to the paper [1]. Section 1 presents the nomenclature used along the document. Sections 2, 3 and 4 introduces mathematical definitions regarding, respectively, the CVaR approximation for Distributionally Robust Chance Constraints (DRCCs), the Wasserstein ambiguity set, the worst-case expectation under the metric-based ambiguity set $\mathcal{A}_1$. The complete final model formulation under ambiguity sets $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ are given in Sections 4, 5 and 6. The real-time stage optimization problem is given in Section 7. Finally, Section 8 introduces the economical and technical network parameters of the considered IEEE 24-node reliability test system.

1 Nomenclature

		Sets
$g \in \mathcal{G}$	:	Set of generators.
$l \in \mathcal{L}$	:	Set of transmission lines.
$d \in \mathcal{D}$	:	Set of demands.
$w \in \mathcal{W}$	:	Set of renewable generators.
$\mathbb{Q} \in \mathcal{A}$	:	Ambiguity set.
$\mathcal{U}$	:	Support of uncertainty.
$i \in \{1, ..., N\}$	:	Set of historical observations of wind power deviation from day-ahead forecast.
		Parameters
$\mathbf{c} \in \mathbb{R}^{\|\mathcal{G}\|}$	:	Production cost of conventional generators [\$/MWh].
$\overline{\mathbf{c}}, \underline{\mathbf{c}} \in \mathbb{R}^{\|\mathcal{G}\|}$	:	Procurement cost of upward, downward reserve from conventional generators [\$/MW].
$\underline{\mathbf{c}} \in \mathbb{R}^{\|\mathcal{G}\|}$	:	Vector of procurement cost of downward reserve from conventional generators [\$/MW].
$\mathbf{d} \in \mathbb{R}^{\|\mathcal{D}\|}$	:	Consumption level of demands [MW].
$\mathbf{f}^{\max} \in \mathbb{R}^{\|\mathcal{L}\|}$	:	Capacity of transmission lines [MW].
$\mathbf{r}^{\max} \in \mathbb{R}^{\|\mathcal{G}\|}$	:	Maximum reserve provision capability of conventional generators [MW].
$\mathbf{p}^{\max} \in \mathbb{R}^{\|\mathcal{G}\|}$	:	Capacity of conventional generators [MW].
$\mathbf{W} \in \mathbb{R}^{\|\mathcal{W}\|\times\|\mathcal{W}\|}$	:	Diagonal matrix of installed capacity of renewable generators [MW].
$\mathbf{Z}^{\mathcal{G}} \in \mathbb{R}^{\|\mathcal{L}\|\times\|\mathcal{G}\|}$	:	Matrix of power transfer distribution factor for conventional generators.
$\mathbf{Z}^{\mathcal{W}} \in \mathbb{R}^{\|\mathcal{L}\|\times\|\mathcal{W}\|}$	:	Matrix of power transfer distribution factor for renewable generators.
$\mathbf{Z}^{\mathcal{D}} \in \mathbb{R}^{\|\mathcal{L}\|\times\|\mathcal{D}\|}$	:	Matrix of power transfer distribution factor for demands.
$\overline{\epsilon}_g$, $\underline{\epsilon}_g$, $\epsilon_l \in \mathbb{R}$	:	Violation probabilities.
$\boldsymbol{\mu} \in \mathbb{R}^{\|\mathcal{W}\|}$	:	Day-ahead power production forecast of renewable generators.
$\rho \in \mathbb{R}$	:	Wasserstein ball radius.
$\mathbf{M}$	:	Matrix for the support definition.
$\boldsymbol{\Sigma}_0 \in \mathbb{R}^{\|\mathcal{W}\|\times\|\mathcal{W}\|}$	:	Matrix defining the ellipsoidal support.
$\boldsymbol{\xi}_0 \in \mathbb{R}^{2\|\mathcal{W}\|}$	:	Center of the ellipsoidal support.

$\boldsymbol{\mu}_0 \in \mathbb{R}^{\vert\mathcal{W}\vert}$	:	Empirical mean of $\tilde{\boldsymbol{\xi}}$.
$\boldsymbol{\Sigma} \in \mathbb{R}^{\vert\mathcal{W}\vert\times\vert\mathcal{W}\vert}$	:	Empirical covariance matrix of $\tilde{\boldsymbol{\xi}}$.
$\tilde{\boldsymbol{\xi}} \in \mathbb{R}^{\vert\mathcal{W}\vert}$	:	Renewable power deviations from day-ahead forecast (random variables).
$\hat{\boldsymbol{\xi}}_i \in \mathbb{R}^{\vert\mathcal{W}\vert}$	:	Historical observations of random variables $\tilde{\boldsymbol{\xi}}$.
		Decision variables
$\mathbf{p} \in \mathbb{R}^{\vert\mathcal{G}\vert}$	:	Power dispatch of conventional generators [MW].
$\overline{\mathbf{r}}, \underline{\mathbf{r}} \in \mathbb{R}^{\vert\mathcal{G}\vert}$	:	Upward, downward reserve dispatch of conventional generators [MW].
$\mathbf{B} \in \mathbb{R}^{\vert\mathcal{G}\vert\times\vert\mathcal{W}\vert}$	:	Matrix of participation factor of conventional generators.
$\lambda, \sigma_i, \boldsymbol{\Lambda}, \boldsymbol{\zeta}_i, \beta_i$	:	Auxiliary variables emanating from the proposed reformulations.

2 CVaR Approximation of DRCCs

As thoroughly discussed in [7], [4], a generic DRCC may be conservatively approximated by a CVaR constraint, as suggested by the following implication

$$\max_{\mathbb{A}\in\mathcal{A}} \ \mathbb{Q}\text{-CVaR}_\epsilon(.) \leq 0 \Rightarrow \min_{\mathbb{Q}\in\mathcal{A}} \ \mathbb{Q}\left(. \leq 0\right) \geq 1-\epsilon. \tag{1}$$

The approximation is conservative because CVaR inherently accounts for the severity (i.e., amplitude) of the constraint violation, which reduces the constraint violation probability pursued by the decision maker. CVaR approximation allows to get rid of the non-convex probability operator $\mathbb{Q}\left(.\right)$ while preserving linearity. The CVaR constraint can eventually be recast as

$$\max_{\mathbb{Q}\in\mathcal{A}} \ \mathbb{Q}\text{-CVaR}_\epsilon(.) = \min_{\tau\in\mathbb{R}} \ \tau + \frac{1}{\epsilon} \ \max_{\mathbb{Q}\in\mathcal{A}} \ \mathbb{E}^{\mathbb{Q}}\left[\lceil . - \tau\rceil^+\right] \tag{2}$$

where $\tau \in \mathbb{R}$ represents an auxiliary variable. Using this technique to reformulate the DRCCs yields

$$\min_{\mathbf{p},\overline{\mathbf{r}},\underline{\mathbf{r}},\mathbf{B}} \ \mathbf{c}^\top\mathbf{p} + \overline{\mathbf{c}}^\top\overline{\mathbf{r}} + \underline{\mathbf{c}}^\top\underline{\mathbf{r}} + \max_{\mathbb{Q}\in\mathcal{A}} \ \mathbb{E}^{\mathbb{Q}}\left[\mathbf{c}^\top\mathbf{B}\tilde{\boldsymbol{\xi}}\right] \tag{3a}$$

$$\text{s.t.} \ \ \mathbf{p} + \overline{\mathbf{r}} \leq \mathbf{p}^{\max} \tag{3b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \geq \mathbf{0} \tag{3c}$$

$$\mathbf{0} \leq \underline{\mathbf{r}} \leq \mathbf{r}^{\max}; \mathbf{0} \leq \overline{\mathbf{r}} \leq \mathbf{r}^{\max} \tag{3d}$$

$$\mathbb{1}^\top\mathbf{p} + \mathbb{1}^\top\mathbf{W}\boldsymbol{\mu} - \mathbb{1}^\top\mathbf{d} = 0 \tag{3e}$$

$$\mathbb{1}^\top\mathbf{B} + \mathbb{1}^\top\mathbf{W} = \mathbf{0} \tag{3f}$$

$$\min_{\underline{\tau}_g} \ \underline{\tau}_g + \frac{1}{\underline{\epsilon}_g}\max_{\mathbb{Q}\in\mathcal{A}} \ \mathbb{E}^{\mathbb{Q}}\left[\left\lceil -\underline{r}_g - \mathbf{B}_g\tilde{\boldsymbol{\xi}} - \underline{\tau}_g\right\rceil^+\right] \leq 0 \quad \forall g\in\mathcal{G} \tag{3g}$$

$$\min_{\overline{\tau}_g} \ \overline{\tau}_g + \frac{1}{\overline{\epsilon}_g}\max_{\mathbb{Q}\in\mathcal{A}} \ \mathbb{E}^{\mathbb{Q}}\left[\left\lceil -\overline{r}_g + \mathbf{B}_g\tilde{\boldsymbol{\xi}} - \overline{\tau}_g\right\rceil^+\right] \leq 0 \quad \forall g\in\mathcal{G} \tag{3h}$$

$$\begin{aligned}\min_{\tau_l} \ \tau_l + \frac{1}{\epsilon_l}\max_{\mathbb{Q}\in\mathcal{A}} \ \mathbb{E}^{\mathbb{Q}}\Big[\Big\lceil &\mathbf{Z}_l^{\mathcal{G}}\left(\mathbf{p} + \mathbf{B}\tilde{\boldsymbol{\xi}}\right) + \mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu} + \tilde{\boldsymbol{\xi}}\right) \\ &-\mathbf{Z}_l^{\mathcal{D}}\mathbf{d} - f_l^{\max} - \tau_l\Big\rceil^+\Big] \leq 0 \quad \forall l\in\mathcal{L}\end{aligned} \tag{3i}$$

where $\lceil . \rceil^+ = \max\left(., 0\right)$. The problem (3) now includes worst-case expectations (i.e., in the objective function and in the constraints) which may be reformulated following the same technique. In the next sections, we provide the reformulation of this worst-case expectation under ambiguity sets $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$.

3 Mathematical definition of the Wasserstein distance

The closeness between two distributions, namely $\mathbb{Q}$ and $\hat{\mathbb{Q}}_N$, may be assessed via the Wasserstein metric, whose mathematical formulation is given by equation (4).

$$d_W(\mathbb{Q}, \hat{\mathbb{Q}}_N) = \left\{ \begin{array}{l} \min_{\Pi} \int_{\Xi^2} ||\tilde{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}_N||\Pi(d\tilde{\boldsymbol{\xi}}, d\hat{\boldsymbol{\xi}}_N) \\ \text{s.t.} \begin{array}{l} \Pi \text{ is a joint distribution of } \tilde{\boldsymbol{\xi}} \text{ and } \hat{\boldsymbol{\xi}}_N \\ \text{with marginals } \mathbb{Q} \text{ and } \hat{\mathbb{Q}}_N \text{, respectively} \end{array} \end{array} \right\}. \tag{4}$$

Problem (4) refers to a transportation problem that displace the samples from $\hat{\mathbb{Q}}_N$ to $\mathbb{Q}$, in the cheapest way in view of the transportation cost $||\tilde{\boldsymbol{\xi}} - \hat{\boldsymbol{\xi}}_N||$. The Wasserstein distance corresponds therefore to the optimal transportation cost which is achieved by the optimal transportation plan Π.

4 Reformulation of the worst-case expectation under ambiguity set $\mathcal{A}_1$

The worst-case expectation problem under metric-based ambiguity set $\mathcal{A}_1$, as proposed in [3], writes as:

$$\max_{\mathbb{Q} \in \mathcal{A}_1} \mathbb{E}^{\mathbb{Q}} \left[\mathbf{a}(\mathbf{x})^\top \boldsymbol{\xi} + b(\mathbf{x}) \right] = \tag{5a}$$

$$\left\{ \begin{array}{lll} \min_{\lambda, \sigma_i} & \lambda\rho + \frac{1}{N} \sum_{i=1}^{N} \sigma_i & \\ \text{s.t.} & \mathbf{a}(\mathbf{x})^\top \hat{\boldsymbol{\xi}}_i + b(\mathbf{x}) \leq \sigma_i & \forall i \in \{1, ..., N\} \\ & ||\, \mathbf{a}(\mathbf{x}) \,||_* \leq \lambda & \forall i \in \{1, ..., N\} \end{array} \right. \tag{5b}$$

where $\lambda \in \mathbb{R}$ and $\sigma \in \mathbb{R}^N$ are additional auxiliary variables. The complete day-ahead DR-OPF problem formulation based on $\mathcal{P}_1$ is provided hereinafter

$$\min_{\substack{\mathbf{p},\overline{\mathbf{r}},\underline{\mathbf{r}},\mathbf{B},\lambda,\sigma_i,\boldsymbol{\gamma}_i,\\ \tau_l,\lambda_l,\sigma_{l,i},\gamma_{l,i,1},\gamma_{l,i,2},\\ \overline{\tau}_g,\overline{\lambda}_g,\overline{\sigma}_{g,i},\overline{\gamma}_{g,i,1},\overline{\gamma}_{g,i,2},\underline{\tau}_g,\underline{\lambda}_g,\underline{\sigma}_{g,i},\underline{\gamma}_{g,i,1},\underline{\gamma}_{g,i,2}}} \quad \mathbf{c}^\top \mathbf{p} + \overline{\mathbf{c}}^\top \overline{\mathbf{r}} + \underline{\mathbf{c}}^\top \underline{\mathbf{r}} + \lambda\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i \tag{6a}$$

$$\text{s.t.} \quad \mathbf{p} + \overline{\mathbf{r}} \le \mathbf{p}^{\max} \tag{6b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \ge \mathbf{0} \tag{6c}$$

$$\mathbf{0} \le \underline{\mathbf{r}} \le \mathbf{r}^{\max}; \mathbf{0} \le \overline{\mathbf{r}} \le \mathbf{r}^{\max} \tag{6d}$$

$$\mathbb{1}^\top \mathbf{p} + \mathbb{1}^\top \mathbf{W}\boldsymbol{\mu} - \mathbb{1}^\top \mathbf{d} = 0 \tag{6e}$$

$$\mathbb{1}^\top \mathbf{B} + \mathbb{1}^\top \mathbf{W} = \mathbf{0} \tag{6f}$$

$$\begin{cases} \mathbf{c}^\top \mathbf{B}\hat{\boldsymbol{\xi}}_i + \boldsymbol{\gamma}_{\boldsymbol{i}}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \le \sigma_i & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \boldsymbol{\gamma}_{\boldsymbol{i}} - \mathbf{c}^\top \mathbf{B}||_* \le \lambda & \forall i \in \{1, ..., N\} \\ \boldsymbol{\gamma}_{\boldsymbol{i}} \ge 0 & \forall i \in \{1, ..., N\} \end{cases} \tag{6g}$$

$$\left.\begin{cases} \overline{\tau}_g + \frac{1}{\epsilon}\left(\overline{\lambda}_g\rho + \frac{1}{N}\sum_{i=1}^{N}\overline{\sigma}_{g,i}\right) \le 0 \\ \mathbf{B}_g\hat{\boldsymbol{\xi}}_i - \overline{r}_g - \overline{\tau}_g + \overline{\boldsymbol{\gamma}}_{g,i,1}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \le \overline{\sigma}_{g,i} & \forall i \in \{1, ..., N\} \\ \overline{\boldsymbol{\gamma}}_{g,i,2}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \le \overline{\sigma}_{g,i} & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \overline{\boldsymbol{\gamma}}_{g,i,1} - \mathbf{B}_g||_* \le \overline{\lambda}_g & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \overline{\boldsymbol{\gamma}}_{g,i,2}||_* \le \overline{\lambda}_g & \forall i \in \{1, ..., N\} \\ \overline{\boldsymbol{\gamma}}_{g,i,1} \ge 0; \overline{\boldsymbol{\gamma}}_{g,i,2} \ge 0 & \forall i \in \{1, ..., N\} \end{cases}\right\} \forall g \in \mathcal{G} \tag{6h}$$

$$\left.\begin{cases} \underline{\tau}_g + \frac{1}{\epsilon}\left(\underline{\lambda}_g\rho + \frac{1}{N}\sum_{i=1}^{N}\underline{\sigma}_{g,i}\right) \le 0 \\ -\mathbf{B}_g\hat{\boldsymbol{\xi}}_i - \underline{r}_g - \underline{\tau}_g + \underline{\boldsymbol{\gamma}}_{g,i,1}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \le \underline{\sigma}_{g,i} & \forall i \in \{1, ..., N\} \\ \underline{\boldsymbol{\gamma}}_{g,i,2}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \le \underline{\sigma}_{g,i} & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \underline{\boldsymbol{\gamma}}_{g,i,1} + \mathbf{B}_g||_* \le \underline{\lambda}_g & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \underline{\boldsymbol{\gamma}}_{g,i,2}||_* \le \underline{\lambda}_g & \forall i \in \{1, ..., N\} \\ \underline{\boldsymbol{\gamma}}_{g,i,1} \ge 0; \underline{\boldsymbol{\gamma}}_{g,i,2} \ge 0 & \forall i \in \{1, ..., N\} \end{cases}\right\} \forall g \in \mathcal{G} \tag{6i}$$

$$\left.\begin{cases} \tau_l + \frac{1}{\epsilon}\left(\lambda_l\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_{l,i}\right) \le 0 \\ \left(\mathbf{Z}_l^{\mathcal{G}}\left(\mathbf{p} + \mathbf{B}\boldsymbol{\xi}\right) + \mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu} + \boldsymbol{\xi}\right) - \mathbf{Z}_l^{\mathcal{D}}\mathbf{d} - f_l^{\max}\right) - \tau_l + \boldsymbol{\gamma}_{l,i,1}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \le \sigma_{l,i} & \forall i \in \{1, ..., N\} \\ \boldsymbol{\gamma}_{l,i,2}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \le \sigma_{l,i} & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \boldsymbol{\gamma}_{l,i,1} - \left(\mathbf{Z}_l^{\mathcal{G}}\mathbf{B} + \mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\right)||_* \le \lambda_l & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \boldsymbol{\gamma}_{l,i,2}||_* \le \lambda_l & \forall i \in \{1, ..., N\} \\ \boldsymbol{\gamma}_i \ge 0 & \forall i \in \{1, ..., N\} \end{cases}\right\} \forall l \in \mathcal{L}. \tag{6j}$$

5 Model reformulation under ambiguity set $\mathcal{A}_2$

The complete day-ahead DR-OPF problem formulation based on $\mathcal{A}_2$ reads as

$$\min_{\substack{\mathbf{p},\overline{\mathbf{r}},\underline{\mathbf{r}},\mathbf{B}\\ y_i,\boldsymbol{\zeta}_i,\lambda\geq 0,\boldsymbol{\Lambda}\succeq 0}} \quad \mathbf{c}^\top\mathbf{p} + \overline{\mathbf{c}}^\top\overline{\mathbf{r}} + \underline{\mathbf{c}}^\top\underline{\mathbf{r}} + \lambda\rho + \langle\boldsymbol{\Lambda},\boldsymbol{\Sigma}\rangle_{\mathrm{F}} + \frac{1}{N}\sum_{i=1}^{N} y_i \tag{7a}$$

$$\text{s.t.} \quad \mathbf{p} + \overline{\mathbf{r}} \leq \mathbf{p}^{\max} \tag{7b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \geq \mathbf{0} \tag{7c}$$

$$\mathbf{0} \leq \underline{\mathbf{r}} \leq \mathbf{r}^{\max}; \mathbf{0} \leq \overline{\mathbf{r}} \leq \mathbf{r}^{\max} \tag{7d}$$

$$\mathbb{1}^\top\mathbf{p} + \mathbb{1}^\top\mathbf{W}\boldsymbol{\mu} - \mathbb{1}^\top\mathbf{d} = 0 \tag{7e}$$

$$\mathbb{1}^\top\mathbf{B} + \mathbb{1}^\top\mathbf{W} = \mathbf{0} \tag{7f}$$

$$\begin{bmatrix} \boldsymbol{\Lambda} & -\frac{1}{2}\mathbf{B}^\top\mathbf{c} + \frac{1}{2}\boldsymbol{\zeta}_i \\ \left(-\frac{1}{2}\mathbf{B}^\top\mathbf{c} + \frac{1}{2}\boldsymbol{\zeta}_i\right)^\top & -\boldsymbol{\zeta}_i^\top\hat{\boldsymbol{\xi}}_i + y_i \end{bmatrix} \succeq 0 \ \ \forall i \in \{1,\ldots,N\} \tag{7g}$$

$$||\boldsymbol{\zeta}_i||_* \leq \lambda \ \ \forall i \in \{1,\ldots,N\} \tag{7h}$$

$$\underline{\tau}_g + \frac{1}{\underline{\epsilon}_g}\left(\underline{\lambda}_g\rho + \left\langle\underline{\boldsymbol{\Lambda}}_g,\boldsymbol{\Sigma}\right\rangle_{\mathrm{F}} + \frac{1}{N}\sum_{i=1}^{N}\underline{y}_{g,i}\right) \leq 0 \quad \forall g \in \mathcal{G} \tag{7i}$$

$$\begin{bmatrix} \underline{\boldsymbol{\Lambda}}_g & \frac{1}{2}B_g + \frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,1} \\ \left(\frac{1}{2}B_g + \frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,1}\right)^\top & \underline{\tau}_g + \underline{r}_g - \underline{\boldsymbol{\zeta}}_{g,i,1}^\top\hat{\boldsymbol{\xi}}_i + \underline{y}_{g,i} \end{bmatrix} \succeq 0 \ \ \forall i \in \{1,\ldots,N\} \ \ \forall g \in \mathcal{G} \tag{7j}$$

$$\begin{bmatrix} \underline{\boldsymbol{\Lambda}}_g & \frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,2} \\ \left(\frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,2}\right)^\top & -\underline{\boldsymbol{\zeta}}_{g,i,2}^\top\hat{\boldsymbol{\xi}}_i + \underline{y}_{g,i} \end{bmatrix} \succeq 0 \ \ \forall i \in \{1,\ldots,N\} \ \ \forall g \in \mathcal{G} \tag{7k}$$

$$||\underline{\boldsymbol{\zeta}}_{g,i,j}||_* \leq \underline{\lambda}_g \ \ \forall j \in \mathcal{J} \ \ \forall i \in \{1,\ldots,N\} \ \ \forall g \in \mathcal{G} \tag{7l}$$

$$\overline{\tau}_g + \frac{1}{\overline{\epsilon}_g}\left(\overline{\lambda}_g\rho + \left\langle\overline{\boldsymbol{\Lambda}}_g,\boldsymbol{\Sigma}\right\rangle_{\mathrm{F}} + \frac{1}{N}\sum_{i=1}^{N}\overline{y}_{g,i}\right) \leq 0 \quad \forall g \in \mathcal{G} \tag{7m}$$

$$\begin{bmatrix} \overline{\boldsymbol{\Lambda}}_g & -\frac{1}{2}B_g + \frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,1} \\ \left(-\frac{1}{2}B_g + \frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,1}\right)^\top & \overline{\tau}_g + \overline{r}_g - \overline{\boldsymbol{\zeta}}_{g,i,1}^\top\hat{\boldsymbol{\xi}}_i + \overline{y}_{g,i} \end{bmatrix} \succeq 0 \ \ \forall i \in \{1,\ldots,N\} \ \ \forall g \in \mathcal{G} \tag{7n}$$

$$\begin{bmatrix} \overline{\boldsymbol{\Lambda}}_g & \frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,2} \\ \left(\frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,2}\right)^\top & -\overline{\boldsymbol{\zeta}}_{g,i,2}^\top\hat{\boldsymbol{\xi}}_i + \overline{y}_{g,i} \end{bmatrix} \succeq 0 \ \ \forall i \in \{1,\ldots,N\} \ \ \forall g \in \mathcal{G} \tag{7o}$$

$$||\overline{\boldsymbol{\zeta}}_{g,i,j}||_* \leq \overline{\lambda}_g \ \ \forall j \in \mathcal{J} \ \ \forall i \in \{1,\ldots,N\} \ \ \forall g \in \mathcal{G} \tag{7p}$$

$$\tau_l + \frac{1}{\epsilon_l}\left(\lambda_l\rho + \langle\boldsymbol{\Lambda}_l,\boldsymbol{\Sigma}\rangle_{\mathrm{F}} + \frac{1}{N}\sum_{i=1}^{N} y_{l,i}\right) \leq 0 \quad \forall l \in \mathcal{L} \tag{7q}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_l & -\frac{1}{2}\left(\mathbf{Z}_l^{\mathcal{G}}\mathbf{B} + \mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\right) + \frac{1}{2}\boldsymbol{\zeta}_{l,i,1} \\ \left(-\frac{1}{2}\left(\mathbf{Z}_l^{\mathcal{G}}\mathbf{B} + \mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\right) + \frac{1}{2}\boldsymbol{\zeta}_{l,i,1}\right)^\top & \tau_l + \left(f_l^{\max} - \mathbf{Z}_l^{\mathcal{G}}\mathbf{p} + \mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\boldsymbol{\mu} - \mathbf{Z}_l^{\mathcal{D}}\mathbf{D}\right) - \boldsymbol{\zeta}_{l,i,1}^\top\hat{\boldsymbol{\xi}}_i + y_{l,i} \end{bmatrix} \succeq 0 \ \ \forall i \in \{1,\ldots,N\} \ \ \forall l \in \mathcal{L} \tag{7r}$$

$$\begin{bmatrix} \boldsymbol{\Lambda}_l & \frac{1}{2}\boldsymbol{\zeta}_{l,i,2} \\ \left(\frac{1}{2}\boldsymbol{\zeta}_{l,i,2}\right)^\top & -\boldsymbol{\zeta}_{l,i,2}^\top\hat{\boldsymbol{\xi}}_i + y_{l,i} \end{bmatrix} \succeq 0 \ \ \forall i \in \{1,\ldots,N\} \ \ \forall l \in \mathcal{L} \tag{7s}$$

$$||\boldsymbol{\zeta}_{l,i,j}||_* \leq \lambda_l \ \ \forall j \in \mathcal{J} \ \ \forall i \in \{1,\ldots,N\} \ \ \forall l \in \mathcal{L} \tag{7t}$$

6 Model reformulation under ambiguity set $\mathcal{A}_3$

The complete day-ahead DR-OPF problem formulation based on $\mathcal{A}_3$ reads as

$$\min_{\substack{\mathbf{p},\overline{\mathbf{r}},\underline{\mathbf{r}},\mathbf{B}\\ y_i,\boldsymbol{\zeta}_i,\lambda\geq 0,\boldsymbol{\Lambda}\succeq 0}} \quad \mathbf{c}^\top\mathbf{p}+\overline{\mathbf{c}}^\top\overline{\mathbf{r}}+\underline{\mathbf{c}}^\top\underline{\mathbf{r}}+\lambda\rho+\langle\boldsymbol{\Lambda},\boldsymbol{\Sigma}\rangle_{\mathrm{F}}+\frac{1}{N}\sum_{i=1}^{N}y_i \tag{8a}$$

$$\text{s.t.}\quad \mathbf{p}+\overline{\mathbf{r}}\leq\mathbf{p}^{\max} \tag{8b}$$

$$\mathbf{p}-\underline{\mathbf{r}}\geq\mathbf{0} \tag{8c}$$

$$\mathbf{0}\leq\underline{\mathbf{r}}\leq\mathbf{r}^{\max};\mathbf{0}\leq\overline{\mathbf{r}}\leq\mathbf{r}^{\max} \tag{8d}$$

$$\mathbb{1}^\top\mathbf{p}+\mathbb{1}^\top\mathbf{W}\boldsymbol{\mu}-\mathbb{1}^\top\mathbf{d}=0 \tag{8e}$$

$$\mathbb{1}^\top\mathbf{B}+\mathbb{1}^\top\mathbf{W}=\mathbf{0} \tag{8f}$$

$$\begin{bmatrix}\boldsymbol{\Lambda} & -\frac{1}{2}\mathbf{B}^\top\mathbf{c}+\frac{1}{2}\boldsymbol{\zeta}_i\\ \left(-\frac{1}{2}\mathbf{B}^\top\mathbf{c}+\frac{1}{2}\boldsymbol{\zeta}_i\right)^\top & -\boldsymbol{\zeta}_i^\top\hat{\boldsymbol{\xi}}_i+y_i\end{bmatrix}\succeq-\sigma_i M\ \ \forall i\in\{1,\dots,N\} \tag{8g}$$

$$||\boldsymbol{\zeta}_i||_*\leq\lambda\ \ \forall i\in\{1,\dots,N\} \tag{8h}$$

$$\underline{\tau}_g+\frac{1}{\underline{\epsilon}_g}\left(\underline{\lambda}_g\rho+\left\langle\underline{\boldsymbol{\Lambda}}_g,\boldsymbol{\Sigma}\right\rangle_{\mathrm{F}}+\frac{1}{N}\sum_{i=1}^{N}\underline{y}_{g,i}\right)\leq 0\quad\forall g\in\mathcal{G} \tag{8i}$$

$$\begin{bmatrix}\underline{\boldsymbol{\Lambda}}_g & \frac{1}{2}B_g+\frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,1}\\ \left(\frac{1}{2}B_g+\frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,1}\right)^\top & \underline{\tau}_g+\underline{r}_g-\underline{\boldsymbol{\zeta}}_{g,i,1}^\top\hat{\boldsymbol{\xi}}_i+\underline{y}_{g,i}\end{bmatrix}\succeq-\underline{\sigma}_{g,i,1}*M\ \ \forall i\in\{1,\dots,N\}\ \ \forall g\in\mathcal{G} \tag{8j}$$

$$\begin{bmatrix}\underline{\boldsymbol{\Lambda}}_g & \frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,2}\\ \left(\frac{1}{2}\underline{\boldsymbol{\zeta}}_{g,i,2}\right)^\top & -\underline{\boldsymbol{\zeta}}_{g,i,2}^\top\hat{\boldsymbol{\xi}}_i+\underline{y}_{g,i}\end{bmatrix}\succeq-\underline{\sigma}_{g,i,2}*M\ \ \forall i\in\{1,\dots,N\}\ \ \forall g\in\mathcal{G} \tag{8k}$$

$$||\underline{\boldsymbol{\zeta}}_{g,i,j}||_*\leq\underline{\lambda}_g\ \ \forall j\in\mathcal{J}\ \ \forall i\in\{1,\dots,N\}\ \ \forall g\in\mathcal{G} \tag{8l}$$

$$\overline{\tau}_g+\frac{1}{\overline{\epsilon}_g}\left(\overline{\lambda}_g\rho+\left\langle\overline{\boldsymbol{\Lambda}}_g,\boldsymbol{\Sigma}\right\rangle_{\mathrm{F}}+\frac{1}{N}\sum_{i=1}^{N}\overline{y}_{g,i}\right)\leq 0\quad\forall g\in\mathcal{G} \tag{8m}$$

$$\begin{bmatrix}\overline{\boldsymbol{\Lambda}}_g & -\frac{1}{2}B_g+\frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,1}\\ \left(-\frac{1}{2}B_g+\frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,1}\right)^\top & \overline{\tau}_g+\overline{r}_g-\overline{\boldsymbol{\zeta}}_{g,i,1}^\top\hat{\boldsymbol{\xi}}_i+\overline{y}_{g,i}\end{bmatrix}\succeq-\overline{\sigma}_{g,i,1}*M\ \ \forall i\in\{1,\dots,N\}\ \ \forall g\in\mathcal{G} \tag{8n}$$

$$\begin{bmatrix}\overline{\boldsymbol{\Lambda}}_g & \frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,2}\\ \left(\frac{1}{2}\overline{\boldsymbol{\zeta}}_{g,i,2}\right)^\top & -\overline{\boldsymbol{\zeta}}_{g,i,2}^\top\hat{\boldsymbol{\xi}}_i+\overline{y}_{g,i}\end{bmatrix}\succeq-\overline{\sigma}_{g,i,2}*M\ \ \forall i\in\{1,\dots,N\}\ \ \forall g\in\mathcal{G} \tag{8o}$$

$$||\overline{\boldsymbol{\zeta}}_{g,i,j}||_*\leq\overline{\lambda}_g\ \ \forall j\in\mathcal{J}\ \ \forall i\in\{1,\dots,N\}\ \ \forall g\in\mathcal{G} \tag{8p}$$

$$\tau_l+\frac{1}{\epsilon_l}\left(\lambda_l\rho+\langle\boldsymbol{\Lambda}_l,\boldsymbol{\Sigma}\rangle_{\mathrm{F}}+\frac{1}{N}\sum_{i=1}^{N}y_{l,i}\right)\leq 0\quad\forall l\in\mathcal{L} \tag{8q}$$

$$\begin{bmatrix}\boldsymbol{\Lambda}_l & -\frac{1}{2}\left(\mathbf{Z}_l^{\mathcal{G}}\mathbf{Y}+\mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\right)+\frac{1}{2}\boldsymbol{\zeta}_{l,i,1}\\ \left(-\frac{1}{2}\left(\mathbf{Z}_l^{\mathcal{G}}\mathbf{B}+\mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\right)+\frac{1}{2}\boldsymbol{\zeta}_{l,i,1}\right)^\top & \tau_l+\left(f_l^{\max}-\mathbf{Z}_l^{\mathcal{G}}\mathbf{p}+\mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\boldsymbol{\mu}-\mathbf{Z}_l^{\mathcal{D}}\mathbf{D}\right)-\boldsymbol{\zeta}_{l,i,1}^\top\hat{\boldsymbol{\xi}}_i+y_{l,i}\end{bmatrix}\succeq-\sigma_{l,i,1}*M\ \ \forall i\in\{1,\dots,N\}\ \ \forall l\in\mathcal{L} \tag{8r}$$

$$\begin{bmatrix}\boldsymbol{\Lambda}_l & \frac{1}{2}\boldsymbol{\zeta}_{l,i,2}\\ \left(\frac{1}{2}\boldsymbol{\zeta}_{l,i,2}\right)^\top & -\boldsymbol{\zeta}_{l,i,2}^\top\hat{\boldsymbol{\xi}}_i+y_{l,i}\end{bmatrix}\succeq-\sigma_{l,i,2}*M\ \ \forall i\in\{1,\dots,N\}\ \ \forall l\in\mathcal{L} \tag{8s}$$

$$||\boldsymbol{\zeta}_{l,i,j}||_*\leq\lambda_l\ \ \forall j\in\mathcal{J}\ \ \forall i\in\{1,\dots,N\}\ \ \forall l\in\mathcal{L} \tag{8t}$$

7 Real-time stage program

This section presents the real-time stage optimization problem (9) that is used to assess the ex-post performance of the different sets of decisions.

$$\min_{\mathbf{B},\boldsymbol{\Delta}\mathbf{d},\boldsymbol{\Delta}\mathbf{w}}\mathbf{c}^\top\mathbf{B}\tilde{\boldsymbol{\xi}}_j+\mathbf{v}_{\mathrm{Shed}}^\top\boldsymbol{\Delta}\mathbf{d} \tag{9a}$$

$$\text{s.t.}\ \ 0\leq\boldsymbol{\Delta}\mathbf{d}\leq\mathbf{d} \tag{9b}$$

$$0\leq\boldsymbol{\Delta}\mathbf{w}\leq\mathbf{W}(\boldsymbol{\mu}+\tilde{\boldsymbol{\xi}}_j) \tag{9c}$$

$$-\underline{\mathbf{r}}\leq\mathbf{B}\tilde{\boldsymbol{\xi}}_j\leq\overline{\mathbf{r}} \tag{9d}$$

$$\mathbf{e}^\top\mathbf{B}\tilde{\boldsymbol{\xi}}_j+\mathbf{e}^\top\mathbf{W}\tilde{\boldsymbol{\xi}}_j+\mathbf{e}^\top\boldsymbol{\Delta}\mathbf{d}-\mathbf{e}^\top\boldsymbol{\Delta}\mathbf{w}=0 \tag{9e}$$

$$\mathbf{Z}_l^{\mathcal{G}}\left(\mathbf{p}+\mathbf{B}\tilde{\boldsymbol{\xi}}_{\mathbf{j}}\right)+\mathbf{Z}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu}+\tilde{\boldsymbol{\xi}}_j\right)-\mathbf{Z}_l^{\mathcal{D}}\mathbf{d}\leq f_l^{\max}\qquad\forall l\in\mathcal{L}. \tag{9f}$$

The objective function (9a) models the real-time operational costs incurred by the energy activation costs and the value of load shedding $\mathbf{v}_{\mathrm{Shed}}\in\mathbb{R}^{|\mathcal{D}|}$ (the wind spillage cost is assumed to be equal to zero). Equations (9b) and (9c) imposes physical limitations on wind power spillage

$\mathbf{\Delta w} \in \mathbb{R}^{|\mathcal{W}|}$ and load shedding $\mathbf{\Delta d} \in \mathbb{R}^{|\mathcal{D}|}$. The actual activation of reserves $\mathbf{B}\tilde{\boldsymbol{\xi}}_j$ is limited by the capacity bound $\underline{\mathbf{r}}$ and $\overline{\mathbf{r}}$ determined in day-ahead (9d). The real-time power balance and line limit capacity are ensured respectively by (9e) and (9f).

8 Network Parameters

We build our model upon the IEEE 24-node Reliability Test System [5] and the economic data available in [6]. The system is represented in Fig. 1.

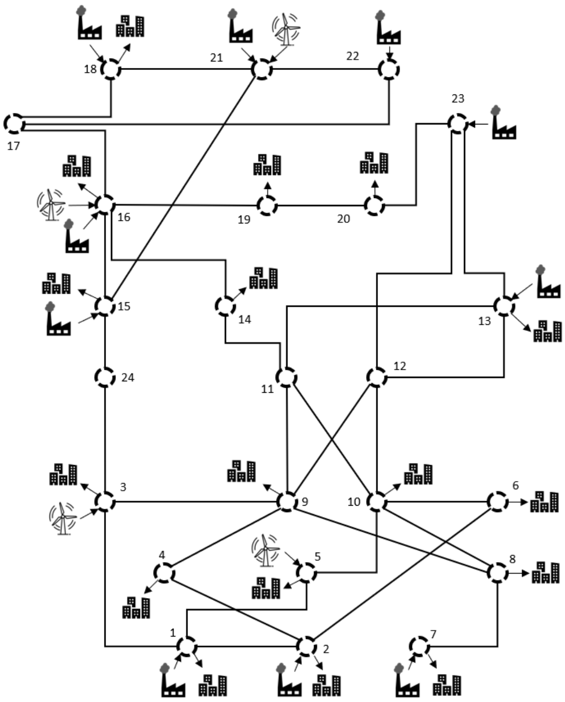

Fig. 1: IEEE RTS 24-node network case study

The network data is given by Table 3. It includes generator parameters such as location node, production cost c_g in \$/MWh, upward reserve capacity procurement cost $\overline{c}_g$ in €/MW, downward reserve capacity procurement cost $\underline{c}_g$ in \$/MW, maximum capacity $p_g^{\max}$ in MW, maximum upward regulation capability $\overline{r}_g^{\max}$ in MW and maximum downward regulation capability $\underline{r}_g^{\max}$ in MW. The 12 generators total capacity is 2,362.5 MW, including 798 MW of upward or downward total flexibility.

Wind farms are also connected to network on nodes 3, 5, 16 and 21 enabling power system studies with high share of renewable generation. The corresponding day-ahead wind forecast $\boldsymbol{\mu}$ in MW, maximum wind farm capacity $W_{(w,w)}$ in MW and expected value in MW are also given in Table 3.

The 17 loads gather 2,207 MW of power demand. Their respective location node, consumption $\mathbf{d}$ in MW and value of curtailed load $\mathbf{v}_{\text{Shed}}$ in \$/MWh are referred in Table 3. The lines are characterized by the nodes they connect, their per-unit inverse susceptance $1/B$ as well as their maximum line capacity $F_{mn}^{\max}$ in MW.

Table 3: Network parameters

Generators						1	2	3	4	5	6	7	8	9	10	11	12
Node						1	2	7	13	15	15	16	18	21	22	23	23
c_g [$ /MWh]						13.32	13.32	20.7	20.93	26.11	10.52	10.52	6.02	5.47	7	10.52	10.89
$\overline{c}_g$ [$ /MW]						1.68	1.68	3.30	4.07	1.89	5.48	5.48	4.98	5.53	8.00	3.45	5.11
$\underline{c}_g$ [$ /MW]						2.32	2.32	4.67	3.93	3.11	3.52	3.52	5.02	4.97	6.00	2.52	2.89
$P_g^{\max}$ [MW]						106.4	106.4	245	413.7	42	108.5	108.5	280	280	210	217	245
$\overline{r}_g$ [MW]						48	48	84	216	42	36	36	60	60	48	72	48
$\underline{r}_g$ [MW]						48	48	84	216	42	36	36	60	60	48	72	48
Wind farms																1	2
Node																3	5
$P_q^{\max}$ [MW]																500	500
Loads	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Node	1	2	3	4	5	6	7	8	9	10	13	14	15	16	18	19	20
d [MW]	84	75	139	58	55	106	97	132	135	150	205	150	245	77	258	141	100
V_{Shed} [$ /MWh]	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
Lines: From node	1	1	1	2	2	3	3	4	5	6	7	8	8	9	9	10	10
To node	2	3	5	4	6	9	24	9	10	10	8	9	10	11	12	11	12
$1/B$ [pu]	0.0146	0.2253	0.0907	0.1356	0.205	0.1271	0.084	0.111	0.094	0.0642	0.0652	0.1762	0.1762	0.084	0.084	0.084	0.084
F_{max} [MW]	175	175	350	175	175	175	400	175	350	175	350	175	175	400	400	400	400
Lines: From node	11	11	12	12	13	14	15	15	15	16	16	17	17	18	19	20	21
To node	13	14	13	23	23	16	16	21	24	17	19	18	22	21	20	23	22
$1/B$ [pu]	0.0488	0.0426	0.0488	0.0985	0.0884	0.0594	0.0172	0.0249	0.0529	0.0263	0.0234	0.0143	0.1069	0.0132	0.0203	0.0112	0.0692
F_{max} [MW]	500	500	500	500	250	250	500	400	500	500	500	500	500	1000	1000	1000	500

References

[1] A. Arrigo, J. Kazempour, Z. De Grève, J.-F. Toubeau and F. Vallée, “Enhanced Wasserstein distributionally robust OPF with dependence structure and support information,” submitted for 14^{th} IEEE PowerTech Conference, Madrid, Spain, 2021.

[2] L. V. Kantorovich and G. S. Rubinshtein, On a space of totally additive functions, Vestnik Leningradskogo Universiteta, 13 (1958), pp. 52-59.

[3] P. Mohajerin Esfahani and D. Kuhn, “Data-Driven Distributionally Robust Optimization Using the Wasserstein Metric: Performance Guarantees and Tractable Reformulations”, in *Mathematical Programming*, vol. 171(1-2), pp. 115-166, 2017.

[4] S. Zymler, D. Kuhn, and B. Rustem, “Distributionally robust joint chance constraints with second-order moment information”, in *Math. Program.*, vol. 137(1-2), pp. 167-198, 2013.

[5] C. Grigg et al., ”The IEEE Reliability Test System 1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee”, IEEE Trans. Power Syst., vol. 14(3), pp. 1010-1020, 1999.

[6] C. Ordoudis, P. Pinson, J. M. Morales and M. Zugno, ”An Updated Version of the IEEE RTS 24-Bus System for Electricity Market and Power System Operation Studies”, Technical University of Denmark (DTU), 2016.

[7] W. Wiesemann and D. Kuhn and M. Sim, “Distributionally robust convex optimization,” in *Oper. Res.*, vol. 62(6), pp. 1358-1376, 2014.