Online Companion - Distributionally Robust Chance Constrained Energy and Reserve Dispatch: An Exact and Physically Bounded Approach

Adriano Arrigo[1], Christos Ordoudis[2], 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

{chror, 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 Wasserstein ambiguity set, the worst-case expectation inside the objective function and the CVaR approximation for Distributionally Robust Chance Constraints (DRCCs). Sections 5, 6 and 7 presents the final model following each technique used to reformulate the DRCCs, namely, the CVaR approximation, the exact MILP reformulation and the exact reformulation with support (or physical bounds), respectively. The real-time stage optimization problem is given in Section 8. Finally, Section 9 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 wind farms.
$i \in \{1, ..., N\}$	:	Set of in-sample database indices.
$j \in \{1, ..., Z\}$	:	Set of out-of-sample database indices.
$\mathbb{P} \in \mathcal{P}$	:	Ambiguity set collecting a family of distributions.
$\mathcal{M}(\Xi)$	:	Set of all distributions on the support Ξ.
$\mathcal{Y}$	:	Set of day-ahead decisions $\{\mathbf{p}, \overline{\mathbf{r}}, \underline{\mathbf{r}}, \mathbf{Y}\}$.
$\boldsymbol{S}(\mathcal{Y})$	:	Safe set.
$\overline{\boldsymbol{S}}(\mathcal{Y})$	:	Unsafe set.
		Parameters
$\mathbf{c} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of production costs of generators [$/MWh].
$\overline{\mathbf{c}}, \underline{\mathbf{c}} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of upward and downward reserve capacity procurement costs of generators [$/MW].
$\mathbf{d} \in \mathbb{R}^{\lvert\mathcal{D}\rvert}$	:	Vector of consumptions of demands [MW].
$\mathbf{f}^{\max} \in \mathbb{R}^{\lvert\mathcal{L}\rvert}$	:	Vector of transmission line capacities [MW].
$\mathbf{r}^{\max} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of maximum regulation capability of generators [MW].
$\mathbf{p}^{\max} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of generator capacities [MW].
$\mathbf{W} \in \mathbb{R}^{\lvert\mathcal{W}\rvert \times \lvert\mathcal{W}\rvert}$	:	Diagonal matrix of wind farm capacities [MW].
$\mathbf{T}^{\mathcal{G}} \in \mathbb{R}^{\lvert\mathcal{L}\rvert \times \lvert\mathcal{G}\rvert}$	:	Matrix of power transfer distribution factors for generators.
$\mathbf{T}^{\mathcal{W}} \in \mathbb{R}^{\lvert\mathcal{L}\rvert \times \lvert\mathcal{W}\rvert}$	:	Matrix of power transfer distribution factors for wind farms.
$\mathbf{T}^{\mathcal{D}} \in \mathbb{R}^{\lvert\mathcal{L}\rvert \times \lvert\mathcal{D}\rvert}$	:	Matrix of power transfer distribution factors for demands.
$\overline{\epsilon}_g$, $\underline{\epsilon}_g$, ϵ_l	:	Violation probability of upward, downward regulation capability constraints and line capacity constraints.
$\boldsymbol{\mu} \in \mathbb{R}^{\lvert\mathcal{W}\rvert}$	:	Forecast wind power production [MW].
$\mathbf{v}_{\text{Shed}} \in \mathbb{R}^{\lvert\mathcal{D}\rvert}$	:	Value of shed load of demands [$/MWh].
$\rho \in \mathbb{R}$	:	Wasserstein ball radius.
$\mathbf{Q} \in \mathbb{R}^{2\lvert\mathcal{W}\rvert \times \lvert\mathcal{W}\rvert}, \mathbf{h} \in \mathbb{R}^{2\lvert\mathcal{W}\rvert}$	:	Parameters for the support definition.

		Uncertain parameters
$\boldsymbol{\xi} \in \mathbb{R}^{\vert\mathcal{W}\vert}$	:	Random variable modeling the wind power deviation from its day-ahead forecast.
$\hat{\boldsymbol{\xi}}_i \in \mathbb{R}^{\vert\mathcal{W}\vert}$	:	Observed sample of the random variable $\boldsymbol{\xi}$.
$\tilde{\boldsymbol{\xi}}_j \in \mathbb{R}^{\vert\mathcal{W}\vert}$	:	Realization of the uncertainty $\boldsymbol{\xi}$ in real time.
		Variables
$\mathbf{p} \in \mathbb{R}^{\vert\mathcal{G}\vert}$	:	Power dispatch of generators [MW].
$\overline{\mathbf{r}}, \underline{\mathbf{r}} \in \mathbb{R}^{\vert\mathcal{G}\vert}$	:	Upward and downward reserve capacity of generators [MW].
$\mathbf{Y} \in \mathbb{R}^{\vert\mathcal{G}\vert \times \vert\mathcal{W}\vert}$	:	Participation factor matrix.

2 Wasserstein Ambiguity set

The closeness between two distributions, namely $\mathbb{P}_1$ and $\mathbb{P}_2$, may be assessed via the Wasserstein metric, whose mathematical formulation is given in Definition 2.1. This metric can be seen as an optimal transportation problem that aims at minimizing the cost of transporting the probability mass from $\mathbb{P}_1$ to $\mathbb{P}_2$.

Definition 2.1 *(Wasserstein metric [2]) The Wasserstein metric $d_W : \mathcal{M}(\Xi) \times \mathcal{M}(\Xi) \to \mathbb{R}$ is defined via*

$$d_W(\mathbb{P}_1, \mathbb{P}_2) = \left\{ \begin{array}{ll} \min & \displaystyle\int_{\Xi^2} ||\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2|| \Pi(d\boldsymbol{\xi}_1, d\boldsymbol{\xi}_2) \\ \text{s.t.} & \begin{array}{l} \Pi \text{ is a joint distribution of } \boldsymbol{\xi}_1 \text{ and } \boldsymbol{\xi}_2 \\ \text{with marginals } \mathbb{P}_1 \text{ and } \mathbb{P}_2, \text{ respectively} \end{array} \end{array} \right\}. \tag{1}$$

The objective function of (1) represents the transportation cost of moving the probability mass from $\mathbb{P}_1$ to $\mathbb{P}_2$. The chosen cost for moving each data sample is the norm $||\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2||$ and, is defined as the Wasserstein distance. The joint distribution $\Pi \in \mathcal{M}(\Xi)$, which is the optimization variable, reflects the optimal transportation plan. Based on this definition, the Wasserstein metric-based ambiguity set collects the closest distributions from an empirical one $\hat{\mathbb{P}}_N$, that is typically constituted of N observed samples, each assigned with probability $\frac{1}{N}$. We thereby mathematically define the Wasserstein ambiguity set as follows:

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{M}(\Xi) : d_W(\mathbb{P}, \hat{\mathbb{P}}_N) \leq \rho \right\}. \tag{2}$$

3 Reformulation of the objective function

The objective function (1a) in [1] optimizes the decisions for the worst-case distribution $\mathbb{P}$ within the ambiguity set $\mathcal{P}$, a Wasserstein ambiguity set with radius ρ. Using findings in [3], a reformulation of the worst-case expectation problem can be determined. In this way, the objective function can be written as:

$$\max_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{Q}} \left[\mathbf{c}^\top \mathbf{Y} \boldsymbol{\xi} \right] = \tag{3a}$$

$$\begin{cases} \displaystyle\min_{\lambda, \sigma_i, \boldsymbol{\gamma_i}} \ \lambda\rho + \frac{1}{N} \sum_{i=1}^{N} \sigma_i \\ \text{s.t. } \ \mathbf{c}^\top \mathbf{Y} \hat{\boldsymbol{\xi}}_i + \boldsymbol{\gamma_i}^\top \left(\mathbf{h} - \mathbf{Q} \hat{\boldsymbol{\xi}}_i \right) \leq \sigma_i & \forall i \in \{1, ..., N\} \\ \qquad ||\mathbf{Q}^\top \boldsymbol{\gamma_i} - \mathbf{c}^\top \mathbf{Y}||_* \leq \lambda & \forall i \in \{1, ..., N\} \\ \qquad \boldsymbol{\gamma_i} \geq 0 & \forall i \in \{1, ..., N\}. \end{cases} \tag{3b}$$

In formulation (3), $\lambda \in \mathbb{R}$, $\boldsymbol{\sigma} \in \mathbb{R}^N$ and $\boldsymbol{\gamma}_i \in \mathbb{R}^{2|\mathcal{W}|}$ are new additional variables. The support (e.g., defined by physical bounds) definition $\mathbf{Q}\boldsymbol{\xi} \leq \mathbf{h}$ restricts the worst distribution to take realistic values. The min operator can then be merged with the min operator over day-ahead decisions, arising in a single level optimization problem.

4 Reformulation of the DRCCs with CVaR Approximation

Following the definition in [1], a generic DRCC can be written as (4). Replacing (4) by a Conditional-Value-at-Risk (CVaR) formulation (5) allows to approximate the DRCC. Using the mathematical definition of CVaR [4], the constraint can eventually be cast into (6).

$$\min_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left(\mathbf{a}^\top \boldsymbol{\xi} \leq b\right) \geq 1-\epsilon \tag{4}$$

$$\max_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\text{-CVaR}_\epsilon(\mathbf{a}^\top \boldsymbol{\xi} - b) \leq 0 \tag{5}$$

$$\min_{\tau\in\mathbb{R}} \tau + \frac{1}{\epsilon} \max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[\lceil \mathbf{a}^\top \boldsymbol{\xi} - b\rceil^+\right] \leq 0. \tag{6}$$

We can next reformulate the worst-case expectation appearing in (6) following (3) as follows:

$$\tau + \frac{1}{\epsilon}\left(\lambda^{\text{CVaR}}\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i^{\text{CVaR}}\right) \leq 0 \tag{7a}$$

$$\mathbf{a}^\top\hat{\boldsymbol{\xi}}_i - b - \tau + \boldsymbol{\gamma}_{i,1}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i^{\text{CVaR}} \qquad \forall i\in\{1,...,N\} \tag{7b}$$

$$\boldsymbol{\gamma}_{i,2}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i^{\text{CVaR}} \qquad \forall i\in\{1,...,N\} \tag{7c}$$

$$||\mathbf{Q}^\top\boldsymbol{\gamma}_{i,1} - \mathbf{a}||_* \leq \lambda^{\text{CVaR}} \qquad \forall i\in\{1,...,N\} \tag{7d}$$

$$||\mathbf{Q}^\top\boldsymbol{\gamma}_{i,2}||_* \leq \lambda^{\text{CVaR}} \qquad \forall i\in\{1,...,N\} \tag{7e}$$

$$\boldsymbol{\gamma}_i \geq 0 \qquad \forall i\in\{1,...,N\}. \tag{7f}$$

The set of *linear* equations (7) involves dual variables λ^{CVaR}, $\boldsymbol{\sigma}^{\text{CVaR}} \in \mathbb{R}^N$ and $\boldsymbol{\gamma}_{i,1}, \boldsymbol{\gamma}_{i,2} \in \mathbb{R}^{2|\mathcal{W}|}$ as well as $\tau \in \mathbb{R}$ that comes from the definition of the CVaR. Those variables can be merged with the dispatch decision variables $\mathcal{Y}$ resulting in a single level optimization problem. This approximation leads to a *conservative* insight of uncertainty because the CVaR inherently accounts for the amplitude of the probability violation, resulting in a lower violation probability than the predefined one.

5 Complete model formulation based on CVaR approximation of DRCCs

The complete model formulation based on CVaR approximation reads as

$$\min_{\substack{\mathbf{p},\overline{\mathbf{r}},\underline{\mathbf{r}},\mathbf{Y},\lambda,\sigma_i,\boldsymbol{\gamma}_i,\\ \tau_l,\lambda_l,\sigma_{l,i},\gamma_{l,i,1},\gamma_{l,i,2},\\ \overline{\tau}_g,\overline{\lambda}_g,\overline{\sigma}_{g,i},\overline{\gamma}_{g,i,1},\overline{\gamma}_{g,i,2},\underline{\tau}_g,\underline{\lambda}_g,\underline{\sigma}_{g,i},\underline{\gamma}_{g,i,1},\underline{\gamma}_{g,i,2}}} \quad \mathbf{c}^\top\mathbf{p} + \overline{\mathbf{c}}^\top\overline{\mathbf{r}} + \underline{\mathbf{c}}^\top\underline{\mathbf{r}} + \lambda\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i \tag{8a}$$

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

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

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

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

$$\mathbf{Y}\mathbf{e} + \mathbf{W}\,\mathbf{e} = \mathbf{0} \tag{8f}$$

$$\begin{cases} \mathbf{c}^\top\mathbf{Y}\hat{\boldsymbol{\xi}}_i + \boldsymbol{\gamma}_i^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\boldsymbol{\gamma}_i - \mathbf{c}^\top\mathbf{Y}||_* \leq \lambda & \forall i \in \{1,...,N\} \\ \boldsymbol{\gamma}_i \geq 0 & \forall i \in \{1,...,N\} \end{cases} \tag{8g}$$

$$\left.\begin{cases} \overline{\tau}_g + \frac{1}{\epsilon}\left(\overline{\lambda}_g\rho + \frac{1}{N}\sum_{i=1}^{N}\overline{\sigma}_{g,i}\right) \leq 0 & \\ \mathbf{Y}_g\hat{\boldsymbol{\xi}}_i - \overline{r}_g - \overline{\tau}_g + \overline{\boldsymbol{\gamma}}_{g,i,1}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \overline{\sigma}_{g,i} & \forall i \in \{1,...,N\} \\ \overline{\boldsymbol{\gamma}}_{g,i,2}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \overline{\sigma}_{g,i} & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\overline{\boldsymbol{\gamma}}_{g,i,1} - \mathbf{Y}_g||_* \leq \overline{\lambda}_g & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\overline{\boldsymbol{\gamma}}_{g,i,2}||_* \leq \overline{\lambda}_g & \forall i \in \{1,...,N\} \\ \overline{\boldsymbol{\gamma}}_{g,i,1} \geq 0; \overline{\boldsymbol{\gamma}}_{g,i,2} \geq 0 & \forall i \in \{1,...,N\} \end{cases}\right\} \forall g \in \mathcal{G} \tag{8h}$$

$$\left.\begin{cases} \underline{\tau}_g + \frac{1}{\epsilon}\left(\underline{\lambda}_g\rho + \frac{1}{N}\sum_{i=1}^{N}\underline{\sigma}_{g,i}\right) \leq 0 & \\ -\mathbf{Y}_g\hat{\boldsymbol{\xi}}_i - \underline{r}_g - \underline{\tau}_g + \underline{\boldsymbol{\gamma}}_{g,i,1}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \underline{\sigma}_{g,i} & \forall i \in \{1,...,N\} \\ \underline{\boldsymbol{\gamma}}_{g,i,2}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \underline{\sigma}_{g,i} & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\underline{\boldsymbol{\gamma}}_{g,i,1} + \mathbf{Y}_g||_* \leq \underline{\lambda}_g & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\underline{\boldsymbol{\gamma}}_{g,i,2}||_* \leq \underline{\lambda}_g & \forall i \in \{1,...,N\} \\ \underline{\boldsymbol{\gamma}}_{g,i,1} \geq 0; \underline{\boldsymbol{\gamma}}_{g,i,2} \geq 0 & \forall i \in \{1,...,N\} \end{cases}\right\} \forall g \in \mathcal{G} \tag{8i}$$

$$\left.\begin{cases} \tau_l + \frac{1}{\epsilon}\left(\lambda_l\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_{l,i}\right) \leq 0 & \\ \left(\mathbf{T}_l^{\mathcal{G}}\left(\mathbf{p} + \mathbf{Y}\boldsymbol{\xi}\right) + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu} + \boldsymbol{\xi}\right) - \mathbf{T}_l^{\mathcal{D}}\mathbf{d} - f_l^{\max}\right) - \tau_l + \boldsymbol{\gamma}_{l,i,1}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_{l,i} & \forall i \in \{1,...,N\} \\ \boldsymbol{\gamma}_{l,i,2}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_{l,i} & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\boldsymbol{\gamma}_{l,i,1} - \left(\mathbf{T}_l^{\mathcal{G}}\mathbf{Y} + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\right)||_* \leq \lambda_l & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\boldsymbol{\gamma}_{l,i,2}||_* \leq \lambda_l & \forall i \in \{1,...,N\} \\ \boldsymbol{\gamma}_i \geq 0 & \forall i \in \{1,...,N\} \end{cases}\right\} \forall l \in \mathcal{L}. \tag{8j}$$

6 Complete model formulation based on the exact MILP reformulation of DRCCs

The complete model formulation based on the exact MILP reformulation of DRCCs reads as

$$\min_{\substack{\mathbf{p},\overline{\mathbf{r}},\underline{\mathbf{r}},\mathbf{Y},\lambda,\sigma_i,\boldsymbol{\gamma_i},\\ \overline{t}_g,\overline{\boldsymbol{\beta}}_g,\overline{\boldsymbol{q}}_g,\underline{t}_g,\underline{\boldsymbol{\beta}}_g,\underline{\boldsymbol{q}}_g,t_l,\boldsymbol{\beta}_l\boldsymbol{q}_l,}} \quad \mathbf{c}^\top \mathbf{p} + \overline{\mathbf{c}}^\top \overline{\mathbf{r}} + \underline{\mathbf{c}}^\top \underline{\mathbf{r}} + \lambda\rho + \frac{1}{N}\sum_{i=1}^{N} \sigma_i \tag{9a}$$

$$\text{s.t.} \quad \mathbf{p} + \overline{\mathbf{r}} \leq \mathbf{p}^{\max} \tag{9b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \geq \mathbf{0} \tag{9c}$$

$$\mathbf{0} \leq \underline{\mathbf{r}} \leq \mathbf{r}^{\max}; \mathbf{0} \leq \overline{\mathbf{r}} \leq \mathbf{r}^{\max} \tag{9d}$$

$$\mathbf{e}^\top \mathbf{p} + \mathbf{e}^\top \mathbf{W}\boldsymbol{\mu} - \mathbf{e}^\top \mathbf{d} = 0 \tag{9e}$$

$$\mathbf{Y}\mathbf{e} + \mathbf{W}\,\mathbf{e} = \mathbf{0} \tag{9f}$$

$$\begin{cases} \mathbf{c}^\top \mathbf{Y}\hat{\boldsymbol{\xi}}_i + \boldsymbol{\gamma_i}^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i & \forall i \in \{1, ..., N\} \\ ||\mathbf{Q}^\top \boldsymbol{\gamma_i} - \mathbf{c}^\top \mathbf{Y}||_* \leq \lambda & \forall i \in \{1, ..., N\} \\ \boldsymbol{\gamma_i} \geq 0 & \forall i \in \{1, ..., N\} \end{cases} \tag{9g}$$

$$\left.\begin{cases} \epsilon N \overline{t}_g - \boldsymbol{e}^\top \overline{\boldsymbol{\beta}}_g \geq \rho N ||\mathbf{Y}_g||_* \\ \overline{r}_g - \mathbf{Y}_g \hat{\boldsymbol{\xi}}_i + M\overline{q}_{g,i} \geq \overline{t}_g - \overline{\beta}_{g,i} \;\; \forall i \in \{1, ..., N\} \\ M\left(1 - \overline{q}_{g,i}\right) \geq \overline{t}_g - \overline{\beta}_{g,i} \;\; \forall i \in \{1, ..., N\} \\ \overline{\boldsymbol{q}}_g \in \{0,1\}^N, \;\; \boldsymbol{\beta} \geq 0 \end{cases}\right\} \quad \forall g \in \mathcal{G} \tag{9h}$$

$$\left.\begin{cases} \epsilon N \underline{t}_g - \boldsymbol{e}^\top \underline{\boldsymbol{\beta}}_g \geq \rho N || - \mathbf{Y}_g||_* \\ \underline{r}_g + \mathbf{Y}_g \hat{\boldsymbol{\xi}}_i + M\underline{q}_{g,i} \geq \underline{t}_g - \underline{\beta}_{g,i} \;\; \forall i \in \{1, ..., N\} \\ M\left(1 - \underline{q}_{g,i}\right) \geq \underline{t}_g - \underline{\beta}_{g,i} \;\; \forall i \in \{1, ..., N\} \\ \underline{\boldsymbol{q}}_g \in \{0,1\}^N, \;\; \boldsymbol{\beta} \geq 0 \end{cases}\right\} \quad \forall g \in \mathcal{G} \tag{9i}$$

$$\left.\begin{cases} \epsilon N t_l - \boldsymbol{e}^\top \boldsymbol{\beta}_l \geq \rho N ||\mathbf{T}_l^{\mathcal{G}} \mathbf{Y} + \mathbf{T}_l^{\mathcal{W}} \mathbf{W}||_* \\ f_l^{\max} - \left(\mathbf{T}_l^{\mathcal{G}} \left(\mathbf{p} + \mathbf{Y}\boldsymbol{\xi}\right) + \mathbf{T}_l^{\mathcal{W}} \mathbf{W}\left(\boldsymbol{\mu} + \boldsymbol{\xi}\right) - \mathbf{T}_l^{\mathcal{D}} \mathbf{d}\right) + Mq_{l,i} \geq t_l - \beta_{l,i} \;\; \forall i \in \{1, ..., N\} \\ M\left(1 - q_{l,i}\right) \geq t_l - \beta_{l,i} \;\; \forall i \in \{1, ..., N\} \\ \boldsymbol{q}_l \in \{0,1\}^N, \;\; \boldsymbol{\beta} \geq 0 \end{cases}\right\} \quad \forall l \in \mathcal{L}. \tag{9j}$$

7 Complete model formulation based on the proposed exact and physically-bounded reformulation of DRCCs

The complete model formulation based on the proposed exact and physically-bounded reformulation of DRCCs reads as

$$\min_{\substack{\mathbf{p},\overline{\mathbf{r}},\underline{\mathbf{r}},\mathbf{Y},\lambda,\sigma_i,\boldsymbol{\gamma_i},\\ \overline{t}_g,\overline{\boldsymbol{\beta}}_g,\overline{w}_{g,i},\overline{\boldsymbol{x}}_{g,i},\underline{t}_g,\underline{\boldsymbol{\beta}}_g,\underline{w}_{g,i},\underline{\boldsymbol{x}}_{g,i},t_l,\boldsymbol{\beta}_l,w_{l,i},\boldsymbol{x}_{l,i}}} \quad \mathbf{c}^\top\mathbf{p} + \overline{\mathbf{c}}^\top\overline{\mathbf{r}} + \underline{\mathbf{c}}^\top\underline{\mathbf{r}} + \lambda\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i \tag{10a}$$

$$\text{s.t.} \quad \mathbf{p} + \overline{\mathbf{r}} \leq \mathbf{p}^{\max} \tag{10b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \geq \mathbf{0} \tag{10c}$$

$$\mathbf{0} \leq \underline{\mathbf{r}} \leq \mathbf{r}^{\max}; \mathbf{0} \leq \overline{\mathbf{r}} \leq \mathbf{r}^{\max} \tag{10d}$$

$$\mathbf{e}^\top\mathbf{p} + \mathbf{e}^\top\mathbf{W}\boldsymbol{\mu} - \mathbf{e}^\top\mathbf{d} = 0 \tag{10e}$$

$$\mathbf{Y}\mathbf{e} + \mathbf{W}\,\mathbf{e} = \mathbf{0} \tag{10f}$$

$$\begin{cases} \mathbf{c}^\top\mathbf{Y}\hat{\boldsymbol{\xi}}_i + \boldsymbol{\gamma_i}^\top\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i & \forall i \in \{1,...,N\} \\ ||\mathbf{Q}^\top\boldsymbol{\gamma_i} - \mathbf{c}^\top\mathbf{Y}||_* \leq \lambda & \forall i \in \{1,...,N\} \\ \boldsymbol{\gamma_i} \geq 0 & \forall i \in \{1,...,N\} \end{cases} \tag{10g}$$

$$\left.\begin{cases} \epsilon N\overline{t}_g - \boldsymbol{e}^\top\overline{\boldsymbol{\beta}}_g \geq \rho N \\ (\overline{r}_g - \mathbf{Y}_g\hat{\boldsymbol{\xi}}_i)\overline{w}_{g,i} + \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right)^\top\overline{\boldsymbol{x}}_{g,i} \geq \overline{t}_g - \overline{\beta}_{g,i} \;\; \forall i \in \{1,...,N\} \\ ||\mathbf{Y}_g\overline{w}_{g,i} + \mathbf{Q}^\top\overline{\boldsymbol{x}}_{g,i}||_* \leq 1 \\ \overline{w}_{g,i} \geq 0,\;\; \overline{\boldsymbol{x}}_{g,i} \leq 0,\;\; \overline{\boldsymbol{\beta}}_g \geq 0 \end{cases}\right\} \;\; \forall g \in \mathcal{G} \tag{10h}$$

$$\left.\begin{cases} \epsilon N\underline{t}_g - \boldsymbol{e}^\top\underline{\boldsymbol{\beta}}_g \geq \rho N \\ (\underline{r}_g + \mathbf{Y}_g\hat{\boldsymbol{\xi}}_i)\underline{w}_{g,i} + \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right)^\top\underline{\boldsymbol{x}}_{g,i} \geq \underline{t}_g - \underline{\beta}_{g,i} \;\; \forall i \in \{1,...,N\} \\ ||-\mathbf{Y}_g\underline{w}_{g,i} + \mathbf{Q}^\top\underline{\boldsymbol{x}}_{g,i}||_* \leq 1 \\ \underline{w}_{g,i} \geq 0,\;\; \underline{\boldsymbol{x}}_{g,i} \leq 0,\;\; \underline{\boldsymbol{\beta}}_g \geq 0 \end{cases}\right\} \;\; \forall g \in \mathcal{G} \tag{10i}$$

$$\left.\begin{cases} \epsilon Nt_l - \boldsymbol{e}^\top\boldsymbol{\beta}_l \geq \rho N \\ (f_l^{\max} - \left(\mathbf{T}_l^{\mathcal{G}}\left(\mathbf{p} + \mathbf{Y}\boldsymbol{\xi}\right) + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu} + \boldsymbol{\xi}\right) - \mathbf{T}_l^{\mathcal{P}}\mathbf{d}\right))w_{l,i} + \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right)^\top\boldsymbol{x}_{l,i} \geq t_l - \beta_{l,i} \;\; \forall i \in \{1,...,N\} \\ ||\left(\mathbf{T}_l^{\mathcal{G}}\mathbf{Y} + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\right)w_{l,i} + \mathbf{Q}^\top\boldsymbol{x}_{l,i}||_* \leq 1 \\ w_{l,i} \geq 0,\;\; \boldsymbol{x}_{l,i} \leq 0,\;\; \boldsymbol{\beta}_l \geq 0 \end{cases}\right\} \;\; \forall l \in \mathcal{L}. \tag{10j}$$

8 Real-time stage program

This section presents the real-time stage optimization problem (11) that is used to assess the ex-post performance of the different sets of decisions.

$$\min_{\mathbf{Y},\boldsymbol{\Delta}\mathbf{d},\boldsymbol{\Delta}\mathbf{w}} \mathbf{c}^\top\mathbf{Y}\tilde{\boldsymbol{\xi}}_j + \mathbf{v}_{\text{Shed}}^\top\boldsymbol{\Delta}\mathbf{d} \tag{11a}$$

$$\text{s.t.} \quad 0 \leq \boldsymbol{\Delta}\mathbf{d} \leq \mathbf{d} \tag{11b}$$

$$0 \leq \boldsymbol{\Delta}\mathbf{w} \leq \mathbf{W}(\boldsymbol{\mu} + \tilde{\boldsymbol{\xi}}_j) \tag{11c}$$

$$-\underline{\mathbf{r}} \leq \mathbf{Y}\tilde{\boldsymbol{\xi}}_j \leq \overline{\mathbf{r}} \tag{11d}$$

$$\mathbf{e}^\top\mathbf{Y}\tilde{\boldsymbol{\xi}}_j + \mathbf{e}^\top\mathbf{W}\tilde{\boldsymbol{\xi}}_j + \mathbf{e}^\top\boldsymbol{\Delta}\mathbf{d} - \mathbf{e}^\top\boldsymbol{\Delta}\mathbf{w} = 0 \tag{11e}$$

$$\mathbf{T}_l^{\mathcal{G}}\left(\mathbf{p} + \mathbf{Y}\tilde{\boldsymbol{\xi}}_{\mathbf{j}}\right) + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu} + \tilde{\boldsymbol{\xi}}_j\right) - \mathbf{T}_l^{\mathcal{P}}\mathbf{d} \leq f_l^{\max} \qquad \forall l \in \mathcal{L}, \tag{11f}$$

The objective function (11a) models the real-time operational costs incurred by the energy activation costs and the value of load shedding $\mathbf{v}_{\text{Shed}} \in \mathbb{R}^{|\mathcal{D}|}$ (the wind spillage cost is assumed to be equal to zero). Equations (11b) and (11c) 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 $\mathbf{Y}\tilde{\boldsymbol{\xi}}_j$ is limited by the capacity bound $\underline{\mathbf{r}}$ and $\overline{\mathbf{r}}$ determined in day-ahead (11d). The real-time power balance and line limit capacity are ensured respectively by (11e) and (11f).

9 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.

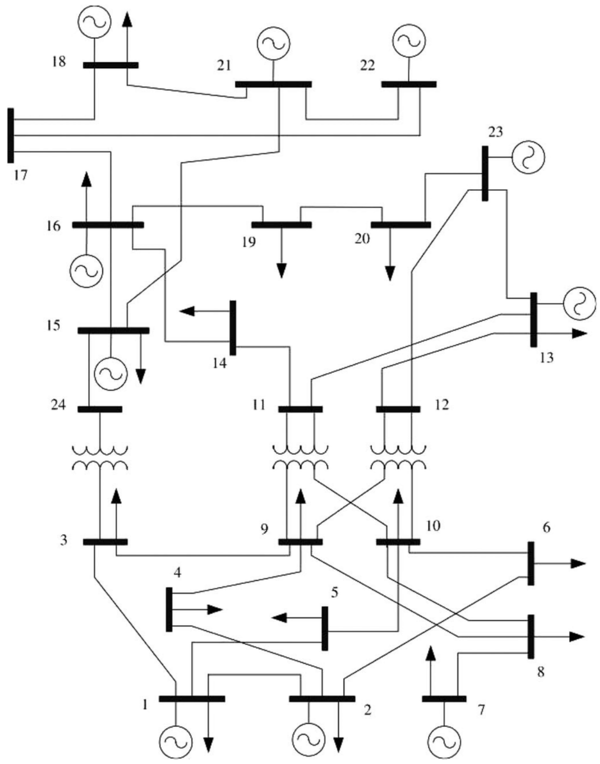

Fig. 1: IEEE RTS 24-node network case study

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	3	4
Node														3	5	16	21
$P_q^{\max}$ [MW]														500	500	300	300
Expected value [MW]														120.54	115.52	53.34	38.16
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, C. Ordoudis, J. Kazempour, Z. De Grève, J.-F. Toubeau and F. Vallée, "Distributionally robust chance constrained energy and reserve dispatch: An exact and physically bounded approach," submitted in IEEE Transactions on Power Systems, 2019.

[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.