European Journal of Operational Research 296 (2022) 304–322

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Wasserstein distributionally robust chance-constrained optimization for energy and reserve dispatch: An exact and physically-bounded formulation

Adriano Arrigo[a,*], Christos Ordoudis[b], Jalal Kazempour[b], Zacharie De Grève[a], Jean-François Toubeau[a], François Vallée[a]

[a] Electrical Power Engineering Unit, University of Mons, 31 Boulevard Dolez, Mons 7000, Belgium
[b] Department of Electrical Engineering, Technical University of Denmark, 325 Elektrovej, Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 21 April 2020
Accepted 7 April 2021
Available online 19 April 2021

Keywords:
OR in energy
Distributionally robust chance-constrained optimization
Energy and reserve dispatch
Physically-bounded exact reformulation
Out-of-sample analysis

ABSTRACT

In the context of transition towards sustainable, cost-efficient and reliable energy systems, the improvement of current energy and reserve dispatch models is crucial to properly cope with the uncertainty of weather-dependent renewable power generation. In contrast to traditional approaches, distributionally robust optimization offers a risk-aware framework that provides performance guarantees when the distribution of uncertain parameters is not perfectly known. In this paper, we develop a distributionally robust chance-constrained optimization with a Wasserstein ambiguity set for energy and reserve dispatch, and provide an exact reformulation. While preserving the exactness, we then improve the model by enforcing physical bounds on the uncertainty space, resulting in a bilinear program. We solve the resulting bilinear model with an iterative algorithm which is computationally efficient and has convergence guarantee. A thorough out-of-sample analysis is performed to compare the proposed model against a scenario-based stochastic program. We also compare the performance of the proposed exact reformulation against an existing approximate technique in the literature, built upon a conditional-value-at-risk measure. We eventually show that the proposed physically-bounded exact reformulation outperforms the other methods by achieving a cost-optimal yet reliable trade-off between reserve procurement and load curtailment.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

The growing integration of renewable energy sources with variable and uncertain production, e.g., wind and solar units, in power systems increases the need for operational flexibility (NERC, 2010). To cope with such variability and uncertainty, the power system operators *reserve* in practice a fraction of the capacity of *flexible resources*, such as hydro and fast-start gas-fired generators, in advance, e.g., in day ahead. This reserved capacity, if needed, will be exploited later in the real-time operation when the uncertainty is realized. This gives rise to an "energy and reserve dispatch" optimization problem, whose objective is to cost-efficiently dispatch the power system with high share of renewable energy sources for the following day (Morales, Zugno, Pineda, & Pinson, 2014; Morales, Conejo, & Perez-Ruiz, 2009; Papavasiliou, Oren, & O'Neill, 2011).

In the current real-world electricity markets, two distinct paradigms for energy and reserve dispatch exist (Domínguez, Oggioni, & Smeers, 2019; González, Villar, Díaz, & Campos, 2014). The first one, which is consistent with the current regulatory policy of European electricity markets, is to design two separate day-ahead markets for energy and reserve. In this paradigm, the reserve market is cleared independently, before or after the energy market (Abbaspourtorbati & Zima, 2016; Toubeau, Grève, & Vallée, 2018). The second approach is to design a single day-ahead market for energy and reserve, which is aligned with the current regulatory policy of the U.S. electricity markets (Martin, Smeers, & Aguado, 2015). This market determines the energy and reserve dispatch jointly in a co-optimization manner.

The common practice in both aforementioned approaches is to define a *minimum reserve requirement* for the whole system or for each area, while using a *deterministic* dispatch model with a single-point forecast of renewable power generation. However, the growing penetration of renewables challenges this practice, requiring a proper incorporation of uncertainty into the energy and reserve dispatch problem (Hobbs & Oren, 2019; Litvinov, Zhao, &

* Corresponding author.
E-mail address: adriano.arrigo@umons.ac.be (A. Arrigo).

https://doi.org/10.1016/j.ejor.2021.04.015
0377-2217/© 2021 Elsevier B.V. All rights reserved.

Zheng, 2019). Following the U.S. practice, this paper considers a co-optimization approach for energy and reserve dispatch, improved by a *probabilistic* characterization of the uncertainty.

The probabilistic energy and reserve dispatch problem has been extensively addressed in the literature, where the uncertainty is modeled using scenario-based stochastic programming (Morales et al., 2009; Papavasiliou et al., 2011), robust optimization (Bertsimas, Litvinov, Sun, Zhao, & Zheng, 2013; Zugno & Conejo, 2015), or chance-constrained programming (Bienstock, Chertkov, & Harnett, 2014; Lubin, Dvorkin, & Backhaus, 2016). Any inadequate representation of uncertainty in this probabilistic optimization problem leads to a sub-optimal solution in terms of the dispatch of flexible resources, and therefore an increase in the total operational cost of the system (Jonsson, Pinson, & Madsen, 2010). The challenge is that the *true* probability distribution of the renewable power generation uncertainty is not necessarily known (Bottieau, Hubert, Grève, Vallée, & Toubeau, 2020; Pinson, 2013). Therefore, any uncertainty modeling technique relying on a specific distribution may fail in achieving the optimal dispatch of flexible resources.

An appealing technique that has been recently developed for modeling uncertainty is Distributionally Robust Optimization (DRO), which enables incorporating a *family* of potential distributions of the uncertainty, the so-called *ambiguity set*, into the problem, and solving it in a computationally tractable manner. An extensive technical survey for DRO is available in Shapiro (2017), Kuhn, Mohajerin Esfahani, Nguyen, and Shafieezadeh-Abadeh (2019), and Rahimian and Mehrotra (2019). An extension of DRO to a multi-stage stochastic program is provided in Shapiro (2021). DRO provides the decision-maker with a degree of freedom to impose her risk attitude by varying the size of the ambiguity set. It can also be seen as a proxy representing the confidence level of the decision-maker to her knowledge about the underlying uncertainty. Under two extreme cases, i.e., the smallest ambiguity set containing a unique distribution and the largest one containing all potential distributions, the outcomes of DRO would be similar to those in a scenario-based stochastic program and in a robust optimization, respectively. The choice for the size of the ambiguity set between those two extreme cases enables the decision-maker to take a risk attitude in between. In addition, DRO generally outperforms scenario-based stochastic programming and robust optimization due to their inherent shortcomings (Ordoudis, Nguyen, Kuhn, & Pinson, 2021). On the one hand, scenario-based stochastic programming provides poor out-of-sample performance unless the number of scenarios is very high, which in turn, increases the computational burden. On the other hand, robust optimization provides a conservative solution for a given uncertainty set. It is also challenging to describe all potential distributions using a single uncertainty set. Therefore, our objective in this paper is to develop an efficient and computationally tractable distributionally robust energy and reserve dispatch model.

Within the DRO framework, it is desirable to incorporate the available historical data of uncertainty realization as much as possible within the model, aiming at making *data-driven* dispatch decisions. In the literature, two distinct paradigms for DRO exist, which differ on how the historical data are embedded. The first one is the so-called *moment-based* DRO (Delage & Ye, 2010; Wiesemann, Kuhn, & Sim, 2014), which defines the ambiguity set according to the moments, e.g., mean and covariance, achieved from the historical data.[1] The shortcoming is that the historical data may not be used efficiently, since the whole historical data are represented via moments only. For example, any new additional data samples will not be exploited for improving energy and reserve dispatch decisions if they do not modify the estimation of moments. The second paradigm is the *metric-based* DRO where the ambiguity set encompasses all distributions, whose probability distance from an empirical distribution including the historical data is lower than or equal to a predefined value, called *radius*. Different probability measures exist for capturing the distance between distributions, e.g., the Wasserstein distance (Kantorovich & Rubinshtein, 1958) and Kullback–Leibler divergence (Hu & Hong, 2013). The advantages of Wasserstein distance in the metric-based DRO over other options have been discussed in Mohajerin Esfahani and Kuhn (2018) and Kuhn et al. (2019). Unlike the moment-based DRO, the metric-based one takes the advantage of existing additional data, but at the potential cost of increased computational burden. In the metric-based DRO, the power system operator may view the radius of the ambiguity set as a tuning parameter to adjust her risk attitude. For instance, a relatively large value for radius allows considering more potential distributions for renewable power generation uncertainty, resulting in a more conservative dispatch solution.

Furthermore, we model each probabilistic constraint[2] of the energy and reserve dispatch optimization problem in the form of a Distributionally Robust Chance Constraint (DRCC). The reasons are twofold: The probabilistic constraints make the optimization problem infinite-dimensional and therefore computationally intractable. By using the chance-constrained programming, the resulting optimization problem becomes tractable. Second, each DRCC allows the violation of the underlying probabilistic constraint up to a predefined extent for the worst-case distribution, ignoring the recourse actions and their potential cost. This brings an extra degree of freedom for the decision-maker to impose her risk attitude by setting a violation probability for each probabilistic constraint *a priori*. This will result in a metric-based distributionally robust chance-constrained energy and reserve dispatch optimization problem with Wasserstein distance, which is the focus of this paper.

Several works exist in the literature that use metric-based distributionally robust chance-constrained programs for power system applications, e.g., Duan, Fang, Jiang, Yao, and Liu (2018), Guo, Baker, Dall'Anese, Hu, and Summers (2019), Poolla, Hota, Bolognani, Callaway, and Cherukuri (2021) and Ordoudis et al. (2021). To the best of our knowledge, all these works *approximate* DRCCs, e.g., using a Conditional Value-at-Risk (CVaR) reformulation, as proposed by Zymler, Kuhn, and Rustem (2013) and Mohajerin Esfahani and Kuhn (2018). This may result in a sub-optimal or an over-conservative solution. Our main contribution is to reformulate DRCCs within the underlying energy and reserve dispatch problem in an *exact* manner, resulting in optimal dispatch decisions. In the following, we explain our methodological steps, and further clarify the contributions of this paper.

Following Chen, Kuhn, and Wiesemann (2018) and Xie (2021), we first provide an *exact* reformulation for the worst-case expectation in the objective function and for all DRCCs. The resulting model is a mixed-integer linear program (MILP). While keeping the exact reformulation, we improve the proposed model by including the *physical bounds* of uncertain parameters, the so-called *support*. For this purpose, we limit the renewable power dispatch to lie within zero and the installed capacity of the corresponding renewable energy unit. We provide a stylized illustrative example, showing the importance of including these additional limits. We

[1] We refer to Zhang, Shen, and Mathieu (2017), Zhao and Jiang (2018), Mieth and Dvorkin (2018), Pourahmadi, Kazempour, Ordoudis, Pinson, and Hosseini (2020) for various applications of moment-based DRO to different problems in power systems. This approach has also been widely applied in other research disciplines – see, for example, Basciftci, Ahmed, and Shen (2021), Shehadeh and Padman (2021), Xin and Goldberg (2021).

[2] By probabilistic constraints, we refer to those operational constraints within the model that include uncertain parameters.

revisit the exact reformulation process for including support, resulting in a bilinear program, instead of the MILP model without support. We solve the resulting bilinear model using an iterative algorithm, which is similar to the one in Zymler et al. (2013). Our various numerical experiments show that the algorithm is computationally efficient and converges to a solution, although a theoretical guarantee for convergence to optimality does not exist. We provide an extensive numerical convergence analysis based on a case study that allows us to quantify the gap between the optimal solution obtained by the iterative algorithm and that achieved by the non-linear solver IPOPT (Wachter & Biegler, 2006), which serves as a benchmark in this paper.

Finally, we provide a thorough numerical analysis based on an extensive out-of-sample simulation to compare different alternatives for DRCC reformulation, including (*i*) the CVaR approximation as proposed in Zymler et al. (2013), (*ii*) the exact MILP reformulation without support as proposed in Chen et al. (2018) and Xie (2021), and (*iii*) our proposed exact physically-bounded bilinear formulation. Such a comparison is extended by taking into account the scenario-based stochastic programming as an additional alternative technique for modeling renewable power generation uncertainty.

As our main finding, we show that our proposed physically-bounded DRO model with the exact reformulation outperforms all other techniques in terms of the total expected operational cost of the system and its standard deviation. Furthermore, our numerical analysis stresses the importance of incorporating the physical bounds of uncertain parameters in order to fully leverage the potential benefits of the DRO approach.

Notation: In this paper, vectors are represented by bold lower cases and matrices are distinguished by bold upper cases. Vector $\mathbf{e}$ is a vector of ones. The set of free, non-negative and non-positive real numbers are represented by $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{R}_-$, respectively. The symbols with a hat, e.g., $\hat{\boldsymbol{\xi}}$, refer to historical observations. The operator $|.|$ returns the cardinality of the underlying set. Furthermore, $(.)^\top$ is the transpose operator, whereas $\mathbb{E}^{\mathbb{P}}[.]$ refers to the expected value with respect to the probability distribution $\mathbb{P}$. The operator $||\mathbf{a}||_*$ represents the dual norm of $||\mathbf{a}||$. This dual norm is defined as $||\mathbf{a}||_* = \{\max_{\mathbf{v}} \ \mathbf{a}^\top\mathbf{v} \ \text{ s.t. } \ ||\mathbf{v}|| \leq 1\}$, where $\mathbf{v}$ is a vector with the same dimension as vector $\mathbf{a}$. All symbols are introduced throughout the paper. However, for convenience, a list of notation is also provided in Appendix A.

2. Model

The Wasserstein distributionally robust chance-constrained energy and reserve dispatch problem reads as

$$\min_{\mathbf{p},\bar{\mathbf{r}},\underline{\mathbf{r}},\mathbf{Y}} \ \mathbf{c}^\top\mathbf{p} + \bar{\mathbf{c}}^\top\bar{\mathbf{r}} + \underline{\mathbf{c}}^\top\underline{\mathbf{r}} + \max_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[\mathbf{c}^\top\mathbf{Y}\tilde{\boldsymbol{\xi}}\right] \tag{1a}$$

$$\text{s.t. } \ \mathbf{p} + \bar{\mathbf{r}} \leq \mathbf{p}^{\max} \tag{1b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \geq \mathbf{0} \tag{1c}$$

$$\mathbf{0} \leq \bar{\mathbf{r}} \leq \mathbf{r}^{\max} \tag{1d}$$

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

$$\mathbf{Y}^\top\mathbf{e} + \mathbf{W}^\top\mathbf{e} = \mathbf{0} \tag{1f}$$

$$\min_{\mathbb{P}\in\mathcal{P}} \ \mathbb{P}\left(-\underline{r}_g \leq \mathbf{Y}_g\tilde{\boldsymbol{\xi}}\right) \geq 1 - \underline{\epsilon}_g \quad \forall g \in \mathcal{G} \tag{1g}$$

$$\min_{\mathbb{P}\in\mathcal{P}} \ \mathbb{P}\left(\mathbf{Y}_g\tilde{\boldsymbol{\xi}} \leq \bar{r}_g\right) \geq 1 - \bar{\epsilon}_g \quad \forall g \in \mathcal{G} \tag{1h}$$

$$\min_{\mathbb{P}\in\mathcal{P}} \ \mathbb{P}\left(\mathbf{T}_l^{\mathcal{G}}\left(\mathbf{p} + \mathbf{Y}\tilde{\boldsymbol{\xi}}\right) + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu} + \tilde{\boldsymbol{\xi}}\right) - \mathbf{T}_l^{\mathcal{D}}\mathbf{d} \leq f_l^{\max}\right) \geq 1 - \epsilon_l \quad \forall l \in \mathcal{L}, \tag{1i}$$

where $g \in \mathcal{G}$, $w \in \mathcal{W}$, $d \in \mathcal{D}$ and $l \in \mathcal{L}$ respectively represent the set of conventional generators, renewable generators, inelastic demands and transmission lines, with their corresponding indices. Note that conventional generators are dispachable, whereas renewable generators are not, meaning that their production level is weather-dependent. Objective function (1a) minimizes the total expected operational cost of the power system, including the day-ahead scheduling cost (the first three terms) and the worst-case expectation of the real-time adjustment cost (the last term). Parameter vector $\mathbf{c} \in \mathbb{R}_+^{|\mathcal{G}|}$ represents the production cost of conventional generators. The production cost of renewable generators is assumed to be zero. In addition, $\bar{\mathbf{c}} \in \mathbb{R}_+^{|\mathcal{G}|}$ and $\underline{\mathbf{c}} \in \mathbb{R}_+^{|\mathcal{G}|}$ refer to the procurement cost of upward and downward reserve from conventional generators, respectively. The upward (downward) reserve is required to cope with renewable power deficit (excess) in the real-time operation. The decision variables are the energy dispatch of conventional generators $\mathbf{p} \in \mathbb{R}_+^{|\mathcal{G}|}$ and their upward and downward reserve capacity dispatch, i.e., $\bar{\mathbf{r}} \in \mathbb{R}_+^{|\mathcal{G}|}$ and $\underline{\mathbf{r}} \in \mathbb{R}_+^{|\mathcal{G}|}$. The recourse actions of conventional generators are approximated using linear decision rules (Kuhn, Wiesemann, & Georghiou, 2011), implying that each conventional generator responds linearly to per-unit renewable power deviations $\tilde{\boldsymbol{\xi}} \in \mathbb{R}^{|\mathcal{W}|}$ in real time.[3] Decision variable matrix $\mathbf{Y} \in \mathbb{R}^{|\mathcal{G}|\times|\mathcal{W}|}$ provides the *participation factor* of conventional generators, such that $\mathbf{Y}\tilde{\boldsymbol{\xi}}$ is the recourse action of those generators. Therefore, the final production level of conventional generators is $\mathbf{p} + \mathbf{Y}\tilde{\boldsymbol{\xi}}$, i.e., their day-ahead energy dispatch plus recourse actions in real time. The dispatch decisions are optimal in expectation for the worst-case distribution $\mathbb{P}$ that resides in the Wasserstein ambiguity set $\mathcal{P}$. This set will be defined later by Eq. (3). The worst-case distribution $\mathbb{P}$ is endogenously determined by the proposed DRO approach. Without loss of generality, we consider deviations $\tilde{\boldsymbol{\xi}}$ with respect to the per-unit day-ahead forecasts $\boldsymbol{\mu} \in \mathbb{R}_+^{|\mathcal{W}|}$ as the sole source of uncertainty.[4] However, it is straightforward to consider other potential sources of uncertainty.

Constraints (1b) to (1d) enforce the lower and upper bounds of power and reserve dispatch by capacity $\mathbf{p}^{\max} \in \mathbb{R}_+^{|\mathcal{G}|}$ and maximum reserve provision capability $\mathbf{r}^{\max} \in \mathbb{R}_+^{|\mathcal{G}|}$ of conventional generators. Constraint (1e) ensures the day-ahead balance between total power production and consumption. Parameter vector $\mathbf{d} \in \mathbb{R}_+^{|\mathcal{D}|}$ represents the consumption level of demands. In addition, diagonal matrix $\mathbf{W} \in \mathbb{R}_+^{|\mathcal{W}|\times|\mathcal{W}|}$ includes the installed capacity of renewable generators. Note that $\mathbf{W}\boldsymbol{\mu}$ gives the day-ahead power production forecast of renewable generators. Similarly, (1f) imposes the power balance in real time. This constraint ensures that the total renewable power imbalances $\mathbf{e}^\top\mathbf{W}\tilde{\boldsymbol{\xi}}$ is adjusted by total recourse actions of conventional generators $\mathbf{e}^\top\mathbf{Y}\tilde{\boldsymbol{\xi}}$. The uncertain parameter $\tilde{\boldsymbol{\xi}}$ is dropped throughout the equality constraint (1f). The probabilistic constraints including $\tilde{\boldsymbol{\xi}}$ are formulated as DRCCs (1g) to (1i). Each individual DRCC ensures fulfilling the underlying constraint with the minimum probability of $1 - \epsilon_{(.)}$ under the worst-case dis-

[3] It is straightforward to improve this affine approximation via generalized decision rules (Georghiou, Wiesemann, & Kuhn, 2015), but at the cost of increased complexity. This extension is left for future research.

[4] The day-ahead forecast $\boldsymbol{\mu}$ and real-time deviations $\tilde{\boldsymbol{\xi}}$ are average values over one hour, such that the hourly balance between total production and consumption is ensured. The intra-hour fluctuations are usually balanced in practice via other market products such as frequency reserves, which are outside the scope of this work.

tribution $\mathbb{P}$. The violation probability $\epsilon_{(.)}$ is a non-negative parameter. This formulation provides the system operator with a degree of freedom to separately adjust the violation probability for each conventional generator and transmission line.[5] Another potential alternative is to use a *joint* chance-constrained framework[6] that ensures the overall system reliability with a high probability. In such a framework, the system operator assigns a single violation probability for the set of all chance constraints. Our focus in this paper is on individual chance constraints. It is worth noting that the worst-case distribution $\mathbb{P}$ for each DRCC is not necessarily identical to that of other DRCCs, or to the one of objective function (1a). For each conventional generator g, (1g) and (1h) restrict the downward and upward adjustment actions, respectively. The subscript in $\mathbf{Y}_g$ picks the row corresponding to conventional generator g in matrix $\mathbf{Y}$. Finally, (1i) enforces the transmission capacity limits $\mathbf{f}^{\max} \in \mathbb{R}_{+}^{|\mathcal{L}|}$ using the power transfer distribution factor matrices $\mathbf{T}^{\mathcal{G}} \in \mathbb{R}^{|\mathcal{L}|\times|\mathcal{G}|}$, $\mathbf{T}^{\mathcal{W}} \in \mathbb{R}^{|\mathcal{L}|\times|\mathcal{W}|}$ and $\mathbf{T}^{\mathcal{D}} \in \mathbb{R}^{|\mathcal{L}|\times|\mathcal{D}|}$. These three mapping matrices relate the nodal injections and offtakes by conventional and renewable generators and demands to the power flow over the lines, respectively (Christie, Wollenberg, & Wangensteen, 2000). The subscript in $\mathbf{T}_l^{(.)}$ selects the row corresponding to transmission line l in matrix $\mathbf{T}^{(.)}$.

Note that the extreme recourse actions, i.e., wind spillage[7] and load curtailment[8], can be incorporated into the energy and reserve dispatch problem (1) in the same way the recourse actions of conventional generators are included. In other words, one can define participation factors for such extreme recourse actions, too. This will complicate the resulting model, although it is less likely that the participation factor of those extreme actions take a non-zero value, as they are comparatively expensive recourse actions for the system. Alternatively, one can see the chance-constrained programming as a way to simplify the dispatch model by not including the extreme recourse actions, and allowing the probabilistic constraints to be violated up to some predefined extent. The extreme recourse actions will be used in the real-time operation to restore infeasibility if the flexibility of conventional generators is insufficient.

In model (1), the ambiguity set $\mathcal{P}$ collects all distributions in the neighborhood of a central empirical distribution $\hat{\mathbb{P}}_N$. This empirical distribution is constructed by assigning a $\frac{1}{N}$ probability mass to each of the N available historical observations $\hat{\boldsymbol{\xi}}$ of uncertain parameters (Mohajerin Esfahani & Kuhn, 2018). To assess the distance of a distribution $\mathbb{P}$ to $\hat{\mathbb{P}}_N$, we use the Wasserstein metric (Kantorovich & Rubinshtein, 1958). The Wasserstein metric d_W : $\mathcal{M}(\Xi) \times \mathcal{M}(\Xi) \to \mathbb{R}$ is defined as

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

where the objective function minimizes the cost of transporting the probability mass from the empirical distribution $\hat{\mathbb{P}}_N$ to the worst-case distribution $\mathbb{P}$. The joint distribution $\Pi \in \mathcal{M}(\Xi \times \Xi)$ reflects the optimal transportation plan. We thereby mathematically define the Wasserstein ambiguity set as

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{M}(\Xi) : d_W(\mathbb{P}, \hat{\mathbb{P}}_N) \leq \rho \right\}, \tag{3}$$

where the non-negative parameter ρ represents the displacement budget, and limits the distance between distributions inside the ambiguity set and the empirical one.

Model (1) with ambiguity set (3) has a min-max structure with constraints involving the min operator. This problem cannot be handled directly by the existing off-the-shelf solvers. The next section provides different reformulation alternatives, including our proposed physically-bounded exact reformulation.

3. Model reformulation

This section reformulates distributionally robust objective function (1a) and DRCCs (1g)–(1i). First, we reformulate objective function (1a). Following the exact reformulation technique proposed by Mohajerin Esfahani and Kuhn (2018), we replace the inner maximization problem $\max\limits_{\mathbb{P}\in\mathcal{P}} \mathbb{E}^{\mathbb{P}}\left[\mathbf{c}^\top \mathbf{Y}\tilde{\boldsymbol{\xi}}\right]$ by a linear minimization problem, i.e.,

$$\left\{ \begin{array}{lll} \min\limits_{\lambda,\sigma_i,\boldsymbol{\gamma}_i} & \lambda\rho + \frac{1}{N}\sum\limits_{i=1}^{N} \sigma_i & \\ \text{s.t.} & \mathbf{c}^\top \mathbf{Y}\hat{\boldsymbol{\xi}}_i + \boldsymbol{\gamma}_i^\top \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i & \forall i \in \{1,\ldots,N\} \\ & ||\mathbf{Q}^\top \boldsymbol{\gamma}_i - \mathbf{c}^\top Y||_* \leq \lambda & \forall i \in \{1,\ldots,N\} \\ & \boldsymbol{\gamma}_i \geq 0 & \forall i \in \{1,\ldots,N\}, \end{array} \right. \tag{4}$$

where $\lambda \in \mathbb{R}$, $\boldsymbol{\sigma} \in \mathbb{R}^N$ and $\boldsymbol{\gamma}_i \in \mathbb{R}^{2|\mathcal{W}|}$ are auxiliary variables. The support, defined as $\mathbf{Q}\tilde{\boldsymbol{\xi}} \leq \mathbf{h}$, restricts the worst-case distribution to take realistic values. The resulting min-min objective function is then collapsed to a single minimization problem.

Second, we reformulate DRCCs (1g)–(1i) using three different alternatives, namely

- The CVaR approximation[9] as proposed in Zymler et al. (2013),
- The exact MILP without support as proposed in Chen et al. (2018) and Xie (2021),
- Our proposed physically-bounded exact bilinear reformulation.

Pursuing completeness, the first alternative is described in Appendix B. This alternative will be used as a benchmark in Section 4. The advantage of CVaR approximation is that it preserves linearity and allows for the incorporation of support information, but at the potential cost of a conservative solution (Nemirovski & Shapiro, 2007). The reason for the increased conservativeness is that the CVaR inherently accounts for the severity of the violation, resulting in a lower violation probability compared to the predefined one. Using the CVaR approximation, the energy and reserve dispatch model (1) is eventually reformulated as a linear program, which is provided in Appendix B.

In the rest of this section, Section 3.1 details the second alternative, i.e., an exact MILP reformulation for DRCCs without support.

[5] For instance, transmission lines within urban areas or those connecting two neighbouring countries are usually of more importance and have to be more reliable than other lines, e.g., those in rural areas. In our proposed model with individual chance constraints, the system operator has freedom to assign comparatively low violation probabilities to critical transmission lines.

[6] The joint chance constraint is usually approximated with a set of individual chance constraints, e.g., using a Bonferroni approximation (Bonferroni, 1936). As such, our proposed model with individual chance constraints can be interpreted as an approximation to a joint chance-constrained program. This program can also be efficiently reformulated using an optimized CVaR approximation (Ordoudis et al., 2021), but at the potential cost of increased computational burden. Although the optimized CVaR approximation may be exact under some certain circumstances, the increased computational burden might be restricting. The reason for this is that the energy and reserve dispatch problem in practice (e.g., in Belgium) is usually solved within the computational time restriction of one hour.

[7] Wind spillage refers to an operational action, which happens if there is a potential for additional wind power generation in real time, but it cannot be realized due to the lack of flexible resources to provide additional downward adjustment services.

[8] Load curtailment refers to the involuntarily curtailment of loads that occurs in operational conditions with wind power deficit, if existing flexible resources cannot provide additional upward adjustment services.

[9] Chen et al. (2018, Corollary 1) show that the CVaR approximation is exact under a condition for which $\epsilon \leq \frac{1}{N}$.

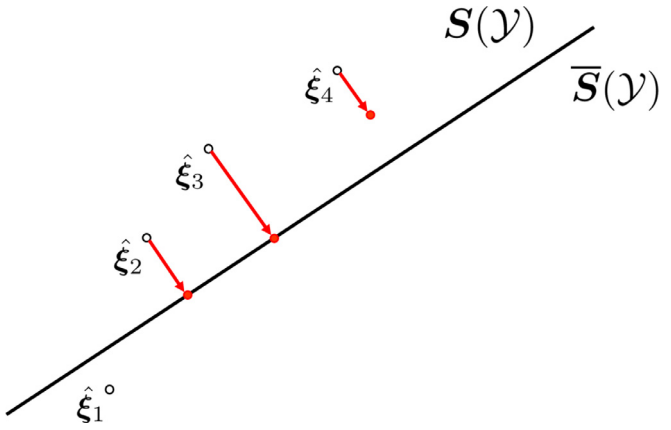


Fig. 1. Historical observations and safe set corresponding to an arbitrarily selected distributionally robust chance constraint.

Unlike the first alternative, we include the modeling framework of the second alternative in the main body of the paper, since it provides the basis for our proposed model. Section 3.2 provides a stylized example, illustrating why the MILP reformulation without support may not determine the optimal power and energy dispatch in a desirable way. Section 3.3 presents our proposed exact reformulation with support, and highlights the advantages.

Remark 1. Hereinafter, for the sake of notational clarity, each distributionally robust chance constraint (1g), (1h) and (1i) is rewritten in a generic way as $\mathbb{P}\big(\tilde{\boldsymbol{\xi}} \in \boldsymbol{S}(\mathcal{Y})\big) \geq 1-\epsilon$, where $\boldsymbol{S}(\mathcal{Y}) = \big\{\tilde{\boldsymbol{\xi}} \in \mathbb{R}^{|\mathcal{W}|} |\ \mathbf{a}^\top \tilde{\boldsymbol{\xi}} \leq b\big\}$ is the so-called *safe set*. This set represents the geometrical region where the underlying constraint is realized. Note that variable vector $\mathbf{a} = \mathbf{A}\mathcal{Y} + \check{\mathbf{a}}$ and variable scalar $b = \mathbf{b}^\top \mathcal{Y} + \check{b}$ are both affine functions of decision variables $\mathcal{Y} = \{\mathbf{p}, \bar{\mathbf{r}}, \underline{\mathbf{r}}, \mathbf{Y}\}$ with $\mathbf{a}, \check{\mathbf{a}} \in \mathbb{R}^{|\mathcal{W}|}$, $\mathbf{A} \in \mathbb{R}^{|\mathcal{W}|\times|\mathcal{Y}|}$, $b, \check{b} \in \mathbb{R}$ and $\mathbf{b} \in \mathbb{R}^{|\mathcal{Y}|}$. The vector $\mathcal{Y}$ collects all day-ahead dispatch decisions including those within vectors $\mathbf{p}, \bar{\mathbf{r}}, \underline{\mathbf{r}}$ and matrix $\mathbf{Y}$ in a vector form. We denote the *unsafe set* as $\overline{\boldsymbol{S}}(\mathcal{Y}) = \big\{\tilde{\boldsymbol{\xi}} \in \mathbb{R}^{|\mathcal{W}|} |\ \mathbf{a}^\top \tilde{\boldsymbol{\xi}} > b\big\}$, which is the complementarity region to $\boldsymbol{S}(\mathcal{Y})$.

3.1. Exact reformulation of DRCCs without physical bounds

We consider the availability of N samples, each corresponding to a historical observation of renewable power deviation in real time with respect to the day-ahead forecast. These samples are collected in the set $\big\{\hat{\boldsymbol{\xi}}_i |\ i = 1, \ldots, N\big\}$. Fig. 1 depicts these samples as well as a half space representing the safe set $\boldsymbol{S}(\mathcal{Y})$ for an individual DRCC. Note that one of the samples in Fig. 1 is outside the safe set, as the chance constraint allows violating the constraint to some extent. We also consider a *displacement budget* as a function of radius $\rho \in \mathbb{R}_+$, whose value is assigned by the system operator. The rationale behind DRO is to determine the worst-case distribution $\mathbb{P}$ within the ambiguity set $\mathcal{P}$ that moves as much as possible samples $\hat{\boldsymbol{\xi}}_i$ outside the safe set with the given displacement budget (Blanchet, Kang, & Murthy, 2019). These moves are illustrated by red arrows in Fig. 1, where the worst-case distribution, for example, could fully move two samples $\hat{\boldsymbol{\xi}}_2$ and $\hat{\boldsymbol{\xi}}_3$ to the boundary of the safe set, while the fourth sample $\hat{\boldsymbol{\xi}}_4$ is partially moved due to the limited displacement budget. In Section 3.3, we will propose a new framework to move the samples while accounting for the support.

In order to express the mathematical equivalence of each DRCC, we first introduce index $\pi_i(\mathcal{Y})$ that reorders samples $\hat{\boldsymbol{\xi}}_i$ in increasing distance to the boundary as a function of decision variables $\mathcal{Y}$. We also define distance function $\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y}))$ that computes the distance between the underlying sample $\hat{\boldsymbol{\xi}}_i$ within the safe set and the boundary of the unsafe set $\overline{\boldsymbol{S}}(\mathcal{Y})$ in the geometrical space. The samples that are already unsafe are assigned with a distance equal to zero. According to Chen et al. (2018), each DRCC, when $\rho > 0$, is mathematically equivalent to a regular inequality constraint of the form

$$\sum_{i=1}^{\epsilon N} \text{dist}\big(\hat{\boldsymbol{\xi}}_{\pi_i(\mathcal{Y})}, \overline{\boldsymbol{S}}(\mathcal{Y})\big) \geq \rho N. \tag{5}$$

The proof is available in Chen et al. (2018). The sum operator in (5) picks the first ϵ times N number of samples. However, this number is not necessarily an integer value. For example, assume the distance of four samples $\hat{\boldsymbol{\xi}}_1$ to $\hat{\boldsymbol{\xi}}_4$ close to the boundary in Fig. 1 are 5, 10, 12 and 15, respectively. If $\epsilon N = 2$, then the left-hand side of (5) is equal to 15 (i.e., 5 + 10). However, in case $\epsilon N = 2.5$, then it would be equal to 21 (i.e., 5 + 10 + 6), taking into account the half of the distance corresponding to the third sample. To get rid of the partial sum operator and the permutation index $\pi_i(\mathcal{Y})$, (5) is replaced by the set of Eq. (6) without approximation, but at the cost of adding extra variables $t \in \mathbb{R}$ and $\boldsymbol{\beta} \in \mathbb{R}^N_+$ (Chen et al., 2018; Xie, 2021):

$$\epsilon N t - \boldsymbol{e}^\top \boldsymbol{\beta} \geq \rho N, \tag{6a}$$

$$\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y})) \geq t - \beta_i \quad \forall i \in \{1, \ldots, N\}. \tag{6b}$$

The last task is to reformulate the distance function $\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y}))$. Following Chen et al. (2018), we replace (6) by (7) with no approximation, but at the cost of adding binary variables $\boldsymbol{q} \in \{0,1\}^N$:

$$\epsilon N t - \boldsymbol{e}^\top \boldsymbol{\beta} \geq \rho N ||\mathbf{a}||_*, \tag{7a}$$

$$b - \mathbf{a}^\top \hat{\boldsymbol{\xi}}_i + M q_i \geq t - \beta_i \quad \forall i \in \{1, \ldots, N\}, \tag{7b}$$

$$M(1 - q_i) \geq t - \beta_i \quad \forall i \in \{1, \ldots, N\}. \tag{7c}$$

It is worth mentioning that (7) holds for $\mathbf{a} \neq 0$, although Remark 2 in Chen et al. (2018) explains how it generalizes to a case where $\mathbf{a} = 0$. In order to maintain linearity, we pick an ∞-norm,[10] i.e., $||\mathbf{a}||_\infty$, whose dual is a 1-norm. Constraints (7b) and (7c) include a sufficiently large constant $M \in \mathbb{R}_+$. It is of importance to select a proper value for M, as a small value may affect optimality, while a very large one may result in numerical ill-conditioning. Note that if $q_i = 0$, then (7b) becomes binding while (7c) is inactive. On the contrary, $q_i = 1$ makes (7c) active while (7b) is inactive. The collection of (7b) and (7c) ensures that the value of $\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y}))$ is always non-negative.

The DRCCs (1h), (1g) and (1i) can now be replaced in an exact way by their generic equivalence (7). Using this exact technique, the energy and reserve dispatch model (1) is eventually reformulated as an MILP, which is provided in Appendix C.

3.2. Illustrative example: Why is the support important?

The illustrative example in this section clarifies why the exact reformulation (7) may give unreasonable dispatch results. We use a two-node power system as depicted in Fig. 2, including a conventional generator, a wind farm, an inelastic demand, and a transmission line. The technical characteristics of the conventional generator are $p^{\max} = 1200$ MW, $r^{\max} = 500$ MW, $c = \$15$/MWh, $\bar{c} = \$2$/MW and $\underline{c} = \$3$/MW. The installed capacity of the wind farm is

[10] This is an arbitrary choice. Alternatively, one can select the 1-norm to maintain linearity.

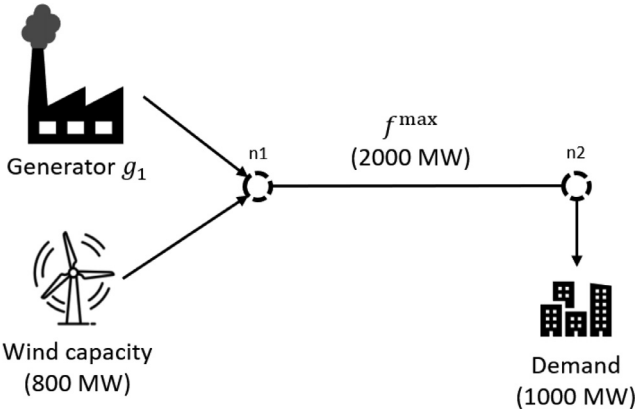


Fig. 2. An illustrative 2-node power system connecting an inelastic demand to a conventional generator and a wind farm through a transmission line.

Table 1
Optimal dispatch of the conventional generator.

	p [MW]	$\bar{r}$ [MW]	$\underline{r}$ [MW]
$\rho = 0.03$	680	480	480
$\rho = 0.05$		Infeasible	

800 MW. The demand is 1000 MW. The transmission line, whose capacity is 2000 MW, is never congested.

Consider a case where the day-ahead forecast of wind power generation is 320 MW. Given the day-ahead forecast and the installed capacity of wind farm, i.e., 320 MW and 800 MW, it is trivial that the maximum required upward and downward reserves are 320 MW and 480 MW, respectively. Assuming a specific case where the sole historical observation is $\hat{\boldsymbol{\xi}}_i = 0$, we solve the energy and reserve dispatch model with an exact reformulation for DRCCs. The ambiguity around the predicted single-point distribution is modeled through the Wasserstein radius $\rho = 0.03$, while the violation probability of each DRCC is set to $\epsilon = 0.05$. We set the value of constant M in (7b) and (7c) to be 10^4, which acts as an upper bound for $\mathbf{a}^\top \tilde{\boldsymbol{\xi}} - b$. The distributionally robust optimal dispatch of the conventional generator is given in Table 1, obtained by Gurobi solver, implemented in Julia using the JuMP package.

The interesting numerical finding is that the upward reserve dispatch of the conventional generator is $\bar{r} = 480$ MW, which is way larger than the 320-MW dispatch of the wind farm. Obviously, the upward reserve is overbooked by 160 MW of useless reserve, which has therefore increased the total operational cost of the underlying power system. Furthermore, the program may even become infeasible, e.g., for $\rho = 0.05$, owing to the need for dispatching a large amount of reserves that are not available in the system.

The reason for such naive results is that the optimization model does not impose the limits of 320 MW and 480 MW for the upward and downward reserves. These constraints can be interpreted as physical bounds, i.e., support, that should be imposed on the displacements of $\hat{\boldsymbol{\xi}}_i$ towards the boundary of the safe set associated with each DRCC under the worst-case distribution. In this direction, we develop an exact reformulation for DRCCs in the next section while taking into account the support information.

3.3. The proposed exact and physically-bounded reformulation of DRCCs

We add the information of physical bounds to the safe set, as schematically illustrated in Fig. 3 by a general green line. This line represents one of the boundaries of a convex polyhedral support. The gray area is within the safe set but outside the support. If the worst-case distribution displaces samples $\hat{\boldsymbol{\xi}}_i$ inside the gray area,

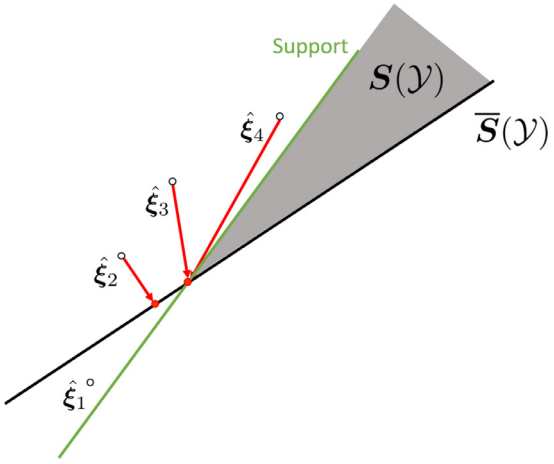


Fig. 3. Illustration of the support (green line), enforcing the gray area to be outside the feasibility space. In comparison to Fig. 1, the support changes the displacement path (red arrows) of historical observations $\hat{\boldsymbol{\xi}}_i$ towards the unsafe set $\overline{\boldsymbol{S}}(\mathcal{Y})$.

the resulting dispatch solution will be overly naive as explained in the previous section. The inclusion of support modifies the sample displacement strategy to find the worst-case distribution: samples must be moved towards the unsafe set boundary without crossing the gray area. For example, samples $\hat{\boldsymbol{\xi}}_3$ and $\hat{\boldsymbol{\xi}}_4$ now take another path towards the unsafe set compared to the one in Fig. 1. This path is characterized by the direction of the shorter path towards the boundary, while being restricted by physical bounds.

The key point is to enhance the distance function in (6b) to be able to take into account the support information. To this purpose, we develop a mathematical framework with three consecutive steps. First, we derive an analytical expression to compute the geometrical distance between a sample and the boundary of the unsafe set while considering the support. Second, we reformulate an individual DRCC based on the updated distance function, resulting in a bilinear program. Third, we leverage an iterative algorithm to efficiently solve the resulting program.

3.3.1. Deriving the expression of the updated distance function

The following Lemma 1 incorporates the support information into the function $\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y}))$.

Lemma 1. *Let* $\left\{\mathbf{Q}\tilde{\boldsymbol{\xi}} \leq \mathbf{h}\right\}$ *be the set of* $|\phi|$ *linear constraints describing the physical bounds of uncertainty, where* $\mathbf{Q} \in \mathbb{R}^{|\phi| \times |\mathcal{W}|}$ and $\mathbf{h} \in \mathbb{R}^{|\phi|}$. *The distance between sample* $\hat{\boldsymbol{\xi}}_i$ *and boundary of the unsafe set, including the support information, is obtained as*

$$\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y})) = \max_{w_i, \boldsymbol{x}_i} \ (b - \mathbf{a}^\top \hat{\boldsymbol{\xi}}_i) w_i - \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right)^\top \boldsymbol{x}_i \tag{8a}$$

$$\text{s.t.} \ \ ||\mathbf{a} w_i - \mathbf{Q}^\top \boldsymbol{x}_i||_* \leq 1, \tag{8b}$$

where $w_i \in \mathbb{R}_+$ *and* $\boldsymbol{x}_i \in \mathbb{R}_+^{|\phi|}$ *are additional auxiliary variables.*

Proof. The distance between sample $\hat{\boldsymbol{\xi}}_i$ and boundary of the unsafe set $\overline{\boldsymbol{S}}(\mathcal{Y})$, including the physical bounds information, can be formulated as

$$\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y})) = \min_{\zeta_i, \boldsymbol{\xi}_i} \ \zeta_i \tag{9a}$$

$$\text{s.t.} \ \ b - \mathbf{a}^\top \boldsymbol{\xi}_i \leq 0 \qquad : w_i \geq 0, \tag{9b}$$

$$\mathbf{Q}\boldsymbol{\xi}_i - \mathbf{h} \leq 0 \qquad : \boldsymbol{x}_i \geq 0 \tag{9c}$$

$$||\boldsymbol{\xi}_i - \hat{\boldsymbol{\xi}}_i|| \leq \zeta_i \quad : (\boldsymbol{v}, u) \in \mathcal{K}_*, \tag{9d}$$

where $\zeta_i \in \mathbb{R}$ and $\boldsymbol{\xi}_i \in \mathbb{R}^{|\mathcal{W}|}$ are the primal variables. The dual variables corresponding to each constraint are given alongside a colon. These dual variables are $w_i \in \mathbb{R}_+$ and $\boldsymbol{x}_i \in \mathbb{R}_+^{|\phi|}$ as well as $(\boldsymbol{\nu}, u) \in \mathbb{R}^{|\mathcal{W}|} \times \mathbb{R}$ residing in the dual cone $\mathcal{K}_*$. The objective is to transfer the given sample $\hat{\boldsymbol{\xi}}_i$ to the unsafe set imposed by (9b) while respecting support (9c) with the minimum displacement budget. This introduces the minimum distance ζ_i and variable $\boldsymbol{\xi}_i$. The latter is the transferred $\hat{\boldsymbol{\xi}}_i$, which is now located at the boundary of the unsafe set. The distance between $\hat{\boldsymbol{\xi}}_i$ and $\boldsymbol{\xi}_i$ is calculated by conic constraint (9d). Inspired by the proof of Lemma 2 in Chen et al. (2018), we aim at dualizing (9) to obtain a maximization problem. The resulting Lagrangian problem is

$$\max_{u,\boldsymbol{\nu},w_i,\boldsymbol{x}_i} \min_{\zeta_i,\boldsymbol{\xi}_i} \zeta_i - \boldsymbol{\nu}^\top\left(\boldsymbol{\xi}_i - \hat{\boldsymbol{\xi}}_i\right) - u\zeta_i + w_i\left(b - \mathbf{a}^\top\boldsymbol{\xi}_i\right) + \boldsymbol{x}_i^\top(\mathbf{Q}\boldsymbol{\xi}_i - \mathbf{h}) \tag{10a}$$

$$\text{s.t. } ||\boldsymbol{\nu}||_* \le u, \quad w_i \ge 0, \quad \boldsymbol{x}_i \ge 0. \tag{10b}$$

Deriving the Karush–Kuhn–Tucker conditions of the inner minimization problem allows us to find the dual problem. This finally renders

$$\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y})) = \max_{u,\boldsymbol{\nu},w_i,\boldsymbol{x}_i} \hat{\boldsymbol{\xi}}_i^\top\boldsymbol{\nu} + bw_i - \mathbf{h}^\top\boldsymbol{x}_i \tag{11a}$$

$$\text{s.t. } ||\boldsymbol{\nu}||_* \le u \tag{11b}$$

$$u = 1 \tag{11c}$$

$$\boldsymbol{\nu} = -\mathbf{a}w_i + \mathbf{Q}^\top\boldsymbol{x}_i \tag{11d}$$

$$w_i \ge 0, \quad \boldsymbol{x}_i \ge 0. \tag{11e}$$

Substituting $\boldsymbol{\nu}$ and u as defined by (11c) and (11d) in (11a) and (11b) yields (8) in Lemma 1. □

3.3.2. Reformulation

Using Lemma 1, Theorem 2 provides our proposed exact and physically-bounded reformulation of DRCCs.

Theorem 2. *The exact and physically-bounded reformulation of each DRCC is*

$$\epsilon Nt - \boldsymbol{e}^\top\boldsymbol{\beta} \ge \rho N, \tag{12a}$$

$$(b - \mathbf{a}^\top\hat{\boldsymbol{\xi}}_i)w_i - \left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right)^\top\boldsymbol{x}_i \ge t - \beta_i \quad \forall i \in \{1, \dots, N\}, \tag{12b}$$

$$||\mathbf{a}w_i - \mathbf{Q}^\top\boldsymbol{x}_i||_* \le 1, \tag{12c}$$

$$w_i \ge 0, \quad \boldsymbol{x}_i \ge 0, \quad \boldsymbol{\beta} \ge 0. \tag{12d}$$

Proof. Replacing $\text{dist}(\hat{\boldsymbol{\xi}}_i, \overline{\boldsymbol{S}}(\mathcal{Y}))$ in (6b) by the maximization problem (8) results in a constraint in the form of $\{\max_{w_i,\boldsymbol{x}} f(w_i, \boldsymbol{x}_i)$ s.t. (8b) $\} \ge t - \beta_i \ \forall i$. This implies that there exist optimal values for w and $\boldsymbol{x}$ which respect (8b) and the optimal value of $f_i(w_i, \boldsymbol{x}_i)$ is greater than or equal to $t - \beta_i$. Equivalently, we can drop the max operator, treat w_i and $\boldsymbol{x}_i$ as variables of min operator in (1a), and impose $\{f(w_i, \boldsymbol{x}_i) \ge t - \beta_i$ and (8b)$\}\forall i$. This results in (12) in Theorem 2. □

We replace individual DRCCs (1h), (1g) and (1i) by their corresponding reformulation (12) and provide the final model in Appendix D. The final set of decision variables is now $\{\mathcal{Y}, t, \boldsymbol{\beta}, w_i, \boldsymbol{x}_i\}$. It should be noted that the set of Eq. (12) contains bilinear terms bw_i and $\mathbf{a}w_i$. Recall that $\mathbf{a}$ and b are affine functions of the dispatch decision variable $\mathcal{Y}$. However, unlike (7), the resulting Eq. (12) do not include any binary variable. For the sake of clarity, we rewrite the final dispatch model in a compact form as

$$\min_{\mathcal{Y},t,\boldsymbol{\beta},w_i,\boldsymbol{x}_i} \mathcal{J}(\mathcal{Y}) \tag{13a}$$

$$\text{s.t. } \mathcal{H}_\alpha(\mathcal{Y}, t, \boldsymbol{\beta}, w_i, \boldsymbol{x}_i) \le 0 \ , \alpha = \{1, \dots, |\alpha|\}, \tag{13b}$$

$$\mathcal{R}_\eta(\mathcal{Y}, t, \boldsymbol{\beta}, w_i, \boldsymbol{x}_i) \le 0 \ , \eta = \{1, \dots, |\eta|\}, \tag{13c}$$

where $\mathcal{J}(\mathcal{Y})$ represents the objective function (1a) after its exact reformulation. The linear and bilinear constraints are gathered in (13b) and (13c), respectively. Note that $|\alpha|$ and $|\eta|$ represent the number of linear and bilinear constraints, respectively.

3.3.3. Solution algorithm

Inspired by Zymler et al. (2013) and Ordoudis et al. (2021), we solve the bilinear model (13) using an iterative sequential algorithm, as explained in Algorithm 1. The iterative procedure is straightforward. Step 1 sets initial variables $\mathcal{Y}$ and $\mathcal{J}^0 = 10^{-3}$ which ensures that the algorithm will not terminate at the first iteration, since $\mathcal{J}^0$ appears in the denominator of the convergence criterion. In Step 2, for given decision variables $\mathcal{Y}^{\text{fixed}}$, we solve (14), aiming at enlarging the feasible space of bilinear constraints as a function of $\{t, \boldsymbol{\beta}, w_i, \boldsymbol{x}_i\}$. Note that (14) is linear as the decision variable set $\mathcal{Y}$ is fixed. Then, for the obtained optimal values of w_i, the linear optimization (15) updates the value of $\mathcal{Y}$ in Step 3. Finally, Step 4 evaluates whether the convergence criterion is fulfilled and goes back to Step 2 if it is not. We select the convergence threshold to be equal to 10^{-4}, meaning that the iterative algorithm will stop if two consecutive items of the sequence do not differ by more than 0.01%.

Algorithm 1 Find a solution to (13).

Step 1: Set initial (feasible) values for dispatch decision variables $\mathcal{Y} = \mathcal{Y}^{\text{fixed}}$. Set k = 1 and $\mathcal{J}^0 = 10^{-3}$.

Step 2: Solve

$$\min_{t,\boldsymbol{\beta},w_i,\boldsymbol{x}_i} \sum_\eta \mathcal{R}_\eta\left(\mathcal{Y}^{\text{fixed}}, t, \boldsymbol{\beta}, w_i, \boldsymbol{x}_i\right) \tag{14a}$$

$$\text{s.t. } \mathcal{H}_\alpha\left(\mathcal{Y}^{\text{fixed}}, t, \boldsymbol{\beta}, w_i, \boldsymbol{x}_i\right) \le 0 \ , \alpha = \{1, \dots, |\alpha|\}, \tag{14b}$$

and fix variables w_i to their optimal values $w_i^{\text{fixed}} = w_i^*$.

Step 3: Solve

$$\min_{\mathcal{Y},t,\boldsymbol{\beta},\boldsymbol{x}_i} \mathcal{J}(\mathcal{Y}) \tag{15a}$$

$$\text{s.t. } \mathcal{H}_\alpha\left(\mathcal{Y}, t, \beta, w_i^{\text{fixed}}, \boldsymbol{x}_i\right) \le 0 \ , \alpha = \{1, \dots, |\alpha|\}, \tag{15b}$$

$$\mathcal{R}_\eta\left(\mathcal{Y}, t, \boldsymbol{\beta}, w_i^{\text{fixed}}, \boldsymbol{x}_i\right) \le 0 \ , \eta = \{1, \dots, |\eta|\}, \tag{15c}$$

and fix dispatch variables to their optimal values $\mathcal{Y}^{\text{fixed}} = \mathcal{Y}^*$.

Step 4: Check if the optimal objective function $\mathcal{J}^k(\mathcal{Y})$ in (15a) verifies the convergence criterion $\frac{\mathcal{J}^k(\mathcal{Y}) - \mathcal{J}^{k-1}(\mathcal{Y})}{\mathcal{J}^{k-1}(\mathcal{Y})} \le 10^{-4}$. If not, set $k \leftarrow k + 1$ and return to step 2.

Algorithm 1 is similar to the iterative procedure proposed in Zymler et al. (2013). The only difference is that instead of one bilinear constraint, we introduce $|\eta|$ number of bilinear constraints $\mathcal{R}_\eta(.)$, and then minimize the sum of their left-hand sides in Step 2, as suggested in Ordoudis et al. (2021). The sequence of generated objective values $\{\mathcal{J}^k\}_{k\in\mathbb{N}}$ monotonically decreases and converges to a finite limit. The proof of convergence is provided in Zymler et al. (2013), although the convergence to global optimality

may not be achieved. In addition, since the convergence criterion stops the algorithm when two consecutive objective values do not differ more than a given threshold (i.e., 10^{-4} in our case), a local optimum might not be achieved, too. However, our extensive numerical study in the next section confirms the satisfactory performance of this algorithm in terms of optimality and convergence speed. This study shows that Algorithm 1 provides outcomes that are more reliable compared to a benchmark solution obtained by the non-linear solver IPOPT.

4. Numerical study

As our case study, we consider a slightly adapted version of the IEEE 24-node reliability test system (Grigg et al., 1999) with 34 transmission lines, 12 conventional generators, 4 wind farms and 17 inelastic demands. The total conventional generation capacity is 2362.5 MW, among which a maximum of 798 MW can be counted as reserve capacity. Furthermore, the total wind power capacity is 1600 MW, and the total inelastic consumption is 2207 MW. Further details about input data and a figure illustrating the network topology of this case study are available in Appendix F.

We use a dataset of 10000 hourly wind power samples, which is available in the online companion (Arrigo et al., 2021). We split this dataset into two parts. The first one contains the arbitrarily selected 1000 samples, which are used in our in-sample simulations. The second part embodies the remaining 9000 samples, which are used a posteriori to assess the quality of decisions through an out-of-sample analysis. Using the in-sample data, we derive the mean forecast[11] $\boldsymbol{\mu}$ and 1000 wind power deviation samples $\hat{\boldsymbol{\xi}}_i$, where $i = \{1, \ldots, 1000\}$. These samples allow constructing the ambiguity set. In addition, we compute the maximum and minimum deviations, i.e., $\boldsymbol{\xi}^{\max}$ and $\boldsymbol{\xi}^{\min}$, based on physical bounds 0 and 1 p.u. By doing so, we subtract the mean wind power generation $\boldsymbol{\mu}$ to establish a rectangular support for the forecast error uncertainty that reflects the actual physical limits of uncertain parameters.

We solve the energy and reserve dispatch problem (1) using three different reformulation alternatives of DRCCs, namely (*i*) CVaR approximation, (*ii*) exact MILP reformulation without support, and (*iii*) our proposed support-based approach. As described in Section 3, the CVaR approximation incorporates the support information as well. Recall that the final reformulation of problem (1) with these three alternatives are available in Appendixes B, C and D, respectively. These three problems are linear, mixed-integer linear, and bilinear programs, respectively. We run all aforementioned models with different numbers of historical observations, i.e., $N = \{50, 100\}$ samples selected from 1000 in-sample data. We also solve these models with different risk-attitudes, i.e., $\epsilon = \{0.03, 0.05\}$, which are considered to be equal for all DRCCs. For each set of parameters, we solve three models for Wasserstein radii[12] ρ ranging from 10^{-4} to 10^{-1}, where the exponent increases linearly with a step of 0.2.

Furthermore, we perform a computational analysis to explore the computational efficiency of Algorithm 1 in terms of both convergence speed and the gap achieved between the solution obtained by Algorithm 1 and the one obtained by the non-linear solver IPOPT. Although the bilinear reformulation is exact, the solution found by Algorithm 1 is not necessarily optimal. Our numerical analyses suggest that Algorithm 1 is more reliable than the non-linear solver. In addition, our results clearly show the importance of support information to be accounted for when one uses DRO to model uncertainties in the energy and reserve dispatch problem.

All linear programs and MILPs are solved under Julia 1.1.1 using JuMP package 0.19 with Gurobi solver 8.1.1, on a 16 GB-RAM computer clocking at 3.40 GHz. All source codes are publicly available in the online companion (Arrigo et al., 2021).

4.1. Computational performance

As the first experiment, we retrieve the evolution of computational time for each model when the number of in-sample scenarios N gradually varies from 10 to 100. Since problem (1) refers to a day-ahead operational problem that should typically be solved in maximum one hour in practice, e.g., in Belgian electricity market, we set a computational time limit of 1 hour. This means that we count a model computationally intractable if its computational time exceeds one hour. We realize that the exact MILP model rapidly reaches to this time limit even for a case with relatively small number of samples, e.g., $N = 20$. From now on, when a comparatively large number of scenarios is considered, the results reported for the exact MILP technique represent sub-optimal solutions obtained within the computational time limit of 1 hour. Our next simulations account for either $N = 50$ or $N = 100$ in-sample scenarios.

Given $\rho = 0.001$ and $\epsilon = 0.05$, Fig. 4(a) shows the computational time as a function of N for the CVaR approximation and our proposed support-based approach. We observe that the computational time increases linearly with the number of in-sample scenarios for the CVaR approximation reformulation. This appealing computational time for the CVaR approximation comes from the fact that the corresponding model is a linear program to be solved in one iteration. In the case of our proposed support-based approach, the increasing trend of computational time is non-smooth, but almost linear with a steeper slope compared to the CVaR approximation. Although our proposed reformulation has a higher computational time compared to the CVaR approximation (about 6 times higher), the results suggest that the iterative procedure of Algorithm 1 converges in a polynomial time. Fig. 4(b) illustrates the number of iterations required by the algorithm when the convergence criterion δ, i.e., the relative difference between the value of objective function in two consecutive iterations, increases from 10^{-6} to 10^{-2}. This figure shows that the number of iterations decreases when δ increases, as the stopping criteria becomes less strict. This decrease is less significant as the value of δ increases from 10^{-6} to 10^{-3}. However, in case of $\delta = 10^{-2}$, the algorithm converges in two iterations only.

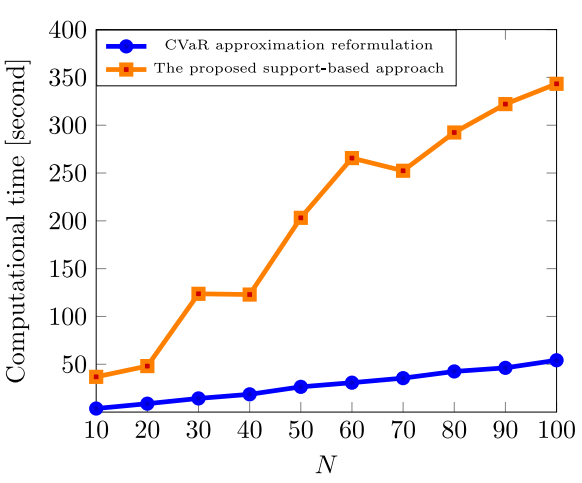
We then study the impact of values assigned for ρ and ϵ on the computational time, whose values are given in Table 2. We observe that the computational time increases with the number of samples N, but decreases as the value assigned for the Wasserstein radius ρ increases. This is consistent with the findings of Chen et al. (2018), showing that by increasing ρ the feasible region of exactly reformulated problems could be potentially convexified.

In what follows, we intend to assess the gap[13] between the solution obtained by the proposed iterative algorithm and the global

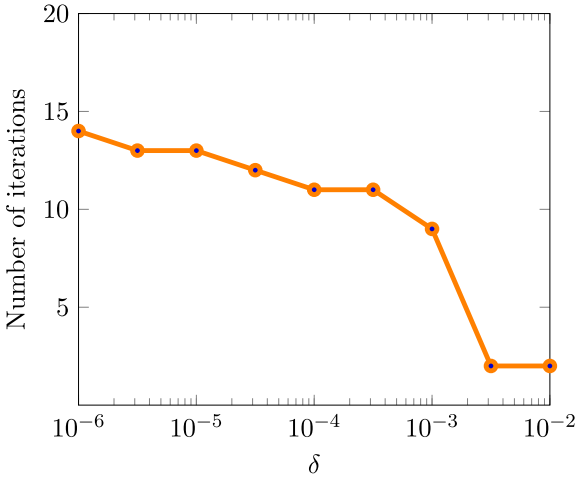
[11] We assume the availability of past observations. In real-life applications, the system operator would forecast the wind power generation based on available observations and run the optimization-based decision-making tool. In this paper, the focus is on the optimization model. We assume that the forecast is calculated as the mean of the past observations.

[12] Using the same Wasserstein radius to compare different DRO models may seem improper, since they are based on different definitions of the ambiguity set. However, this is likely the most practical and efficient way to compare the techniques between each other. Furthermore, to the best of our knowledge, there does not exist a technique to calculate an equivalent value for radius, given its value defined for another type of ambiguity set.

[13] The optimality gap is mathematically defined as the difference between the primal best obtained solution and the dual best obtained solution. In this paper, since the dual best obtained solution is not mathematically achievable, our practical alternative to assess the gap (but not necessarily the optimality gap) is to compute lower bounds on the optimal value of the objective function. We still use the term



(a) Computational time as a function of the number of in-sample scenarios N (fixed values: $\rho = 0.001$ and $\epsilon = 0.05$)



(b) Number of iterations for the iterative algorithm as a function of the stopping criteria δ (fixed values: $N = 100$, $\rho = 0.001$ and $\epsilon = 0.05$)

Fig. 4. Computational study.

Table 2
Computational time [second].

		CVaR approximation		The proposed support-based approach	
		$N = 50$	$N = 100$	$N = 50$	$N = 100$
$\rho = 0.0001$	$\epsilon = 0.05$	33.3	76.8	196.4	424.3
	$\epsilon = 0.03$	29.3	74.6	352.6	464.3
$\rho = 0.001$	$\epsilon = 0.05$	26.4	54.7	203.1	343.3
	$\epsilon = 0.03$	29.7	62.6	304.4	357.4
$\rho = 0.01$	$\epsilon = 0.05$	5.2	14.3	51.6	122.4
	$\epsilon = 0.03$	5.6	15.2	50.9	189.3

optimum. Recall that the underlying problem is a bilinear program which makes it non-convex, and therefore the global optimal point is not necessarily accessible. In general, one potential alternative to compute the optimality gap is to derive the dual problem and compare the dual best solution to our best obtained solution. However, the dual problem in our setting would not allow to find a non-trivial lower bound[14] As the next alternative, a McCormick relaxation (McCormick, 1976) of the bilinear terms would allow us to assess the optimality gap between the solution of the iterative algorithm and the one of the relaxed problem. However, variables w within the bilinear terms have a lower bound only, which is equal to zero, but there is no upper bound that is also required for the McCormick relaxation. We have checked the performance of McCormick relaxation by fixing the upper bound to the value given by the iterative algorithm's solution. However, by doing so, we have observed that the problem is infeasible. We hypothesize that the McCormick relaxation jeopardizes the proper displacements of samples within the physical bounds, which in turn, makes the program infeasible, as already observed in the illustrative example of Section 3.2.

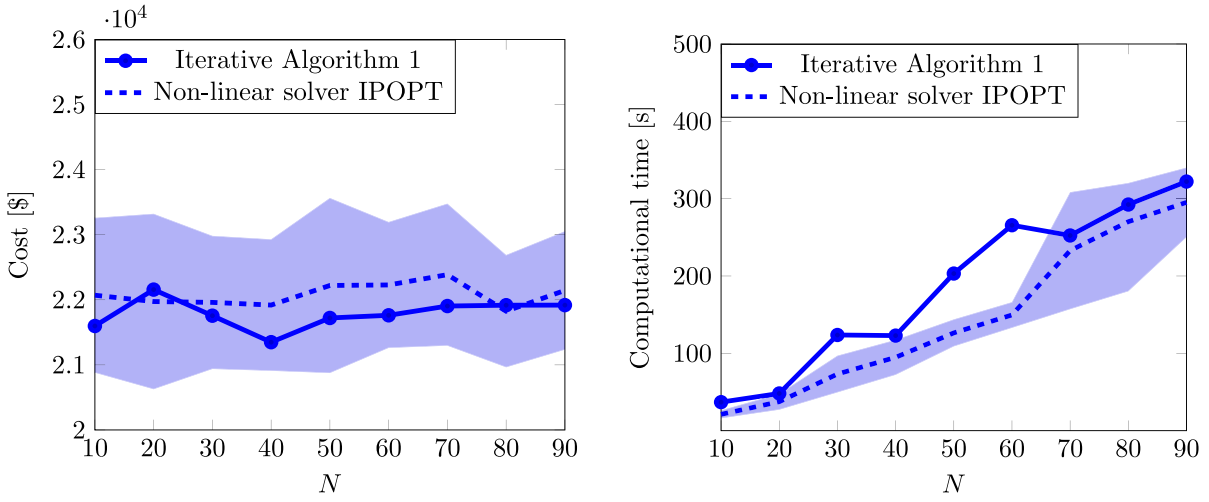
As a pragmatic alternative, we compare the optimal value of the total expected operational cost of the system achieved from the iterative Algorithm 1 to that obtained from the non-linear solver IPOPT. This solver provides a solution which we could use as a benchmark against the solution of the iterative algorithm and assess the gap. The resulting cost profiles as a function of N are presented in Fig. 5(a). We run the IPOPT solver several times with a different initialization for each training sample size N in a way that the number of training sample multiplied by the number of runs stays constant[15] and equal to 200. Fig. 5(a) reports the expected value of objective function (dashed line) achieved in all these runs as well as the standard deviation (shadow area).

We observe that the cost profile obtained from Algorithm 1 takes lower values more frequently. In some cases, IPOPT finds a solution with a lower cost compared to the solution of Algorithm 1, but the gap between two solutions is relatively small. The variability around the outcome of IPOPT stems from multiple

"gap" for describing the difference between our best obtained primal solution and a potential lower bound.

[14] According to Boyd and Vandenberghe (2004), the dual Lagrangian function is always concave even though the primal problem might be non-convex. Consider a set of quadratic constraints in form of $x^\top P_i x \leq d_i, \quad \forall i \in \{1, \ldots, M\}$ with corresponding dual variables λ_i. In (13), all quadratic constraints are bilinear, meaning that P_i is not positive semidefinite. An immediate conclusion is that, $P(\lambda) = \sum_{i=1}^{M} \lambda_i P_i$ is positive semidefinite in our bilinear problem, if and only if all $\lambda_i = 0$. This refers to a very special condition under which all bilinear constraints are non-binding. In the case there is a non-zero λ_i, the dual problem gives us a trivial conclusion that the lower bound is minus infinity. In order to get a potentially non-trivial lower bound from dual problem, one has to drop *all* bilinear constraints in the primal problem. However, nearly 98% of primal constraints in (13) are bilinear, e.g., in the case where $N = 50$. If we drop all those constraints, the resulting problem will be highly relaxed that does not represent well the original problem (13). For further details on how to derive the dual problem in a bilinear program, we refer the interested reader to Andersen (2021).

[15] When the theoretical number of runs, e.g., $\frac{200}{N}$, takes a non-integer value, we round it to the smallest integer greater than or equal to the theoretical number of runs.



(a) Total expected operational cost of the system as a function of the number of in-sample scenarios N, resulting from Algorithm 1 and the non-linear solver IPOPT (fixed values: $\rho = 0.001$ and $\epsilon = 0.05$)

(b) Computational time as a function of the number of in-sample scenarios N (fixed values: $\rho = 0.001$ and $\epsilon = 0.05$)

Fig. 5. Comparison of results achieved by the iterative Algorithm 1 and the non-linear solver IPOPT. For different runs of IPOPT with different initializations, the shadow area covers the area ranging from the expected value minus the standard deviation to the expected value plus the standard deviation.

solutions achieved by IPOPT, each with a different initialization. This variability shows that this solver usually fails to provide reliable results. We retrieve the computational times and report the mean (dashed line) and the standard deviation (shadow area) in Fig. 5(b). The computational times in IPOPT and Algorithm 1 are comparable. Algorithm 1 slightly takes more time to find an optimal solution when $\delta = 10^{-4}$. Based on our numerical observations, we conclude that Algorithm 1 is more reliable than IPOPT, and performs well in the scope of our application.

4.2. In-sample outcomes: dispatch results

In this section, we present the optimal dispatch decisions $\{p_g, \bar{r}_g, \underline{r}_g, \mathbf{Y}_g\}$ for a arbitrarily chosen conventional generator $g = 2$. For all three DRCC reformulation techniques, we compute the actual power production $p_g + \mathbf{Y}_g \hat{\boldsymbol{\xi}}_i$, which corresponds to the sum of the day-ahead power dispatch and the real-time recourse action for each historical observation $\hat{\boldsymbol{\xi}}_i$ in the in-sample database. The resulting distribution is shown with blue bars in Fig. 6 for a case with $N = 100$, $\rho = 10^{-4}$ and $\epsilon = 0.05$. The day-ahead energy as well as the upward and downward reserve dispatches are shown with green number and red arrows, respectively. The level of production is therefore restricted to lie within the extremity of the arrows, representing the production adjustment capability of the generator.

For all three reformulations, we observe that the corresponding distribution mostly takes values around the day-ahead energy dispatch p_g. However, DRCCs allow constraint violations, resulting in a total power production that lies outside the restricted zone. The number of cases with constraint violation is given in blue circles. We notice that our proposed support-based approach exhibits a higher number of cases with constraint violation, compared to the CVaR approximation. The reason for this is that the CVaR approximation is a conservative reformulation alternative and therefore the predefined level of violation probability does not necessarily occur. Finally, the exact MILP method overbooks downward reserve, which is due to the fact that the support is not included within this technique, as explained in our illustrative example in Section 3.2. Therefore, constraints are less prone to be violated.

4.3. Out-of-sample outcomes: Total system cost without re-optimization

We derive the total operational cost of the system, i.e., the optimal value of objective function (1a), using an out-of-sample analysis as follows: we collect the optimal decision variables $\{\mathbf{p}, \bar{\mathbf{r}}, \underline{\mathbf{r}}, \mathbf{Y}\}$ obtained from each model, and calculate the optimal value of day-ahead costs $\mathbf{c}^\top \mathbf{p}$, $\bar{\mathbf{c}}^\top \bar{\mathbf{r}}$ and $\underline{\mathbf{c}}^\top \underline{\mathbf{r}}$. For the recourse cost, we arbitrarily select first 1000 samples from the out-of-sample data[16], say $\hat{\boldsymbol{\xi}}_j$, where $j = \{1, \ldots, 1000\}$, and calculate the average recourse cost $10^{-3} \sum_j \mathbf{c}^\top \mathbf{Y} \hat{\boldsymbol{\xi}}_j$. The sum of day-ahead and recourse costs gives the average of out-of-sample total system cost. We also calculate the standard deviation of this cost as a measure of variability. Here, the system operator treats the participation factors $\mathbf{Y}$ as informed decisions for recourse actions, and do not solve another optimization in real time under each out-of-sample realization. Consequently, the resulting cost neglects the potential cost associated with a constraint violation.

Fig. 7 depicts the average out-of-sample total system cost (thick curves) and the cost variability (shaded zones around the thick curves) for each model. For clarity of the figures, the shaded zones around the thick curves exhibit 10% of the standard deviation only. In most of the cases, the average cost obtained from the proposed support-based model is lower than the one achieved by the other two models. This difference is significant in some cases, e.g., in the one with $N = 100$ and $\epsilon = 0.05$. The cost achieved by the exact MILP model oscillates between the costs obtained by the other two approaches, depending on the optimality gap of the solution reached after the one-hour computational time limit. By including the support information, the average cost profile saturates for higher values of ρ. Without such support information, the program either becomes infeasible, e.g., in plots 7(a) and 7(c), or optimizes for an unrealistic representation of the uncertainty.

However, it is worth noting that these results are obtained by assuming that real-time decisions are fixed in day ahead. In

[16] In our numerical study, we have observed that the out-of-sample outcomes using either arbitrarily selected 1000 or all 9000 test samples are similar. For the sake of reduced computational time in our out-of-sample study, we have only picked the first 1000 arbitrarily selected samples from the testing dataset.

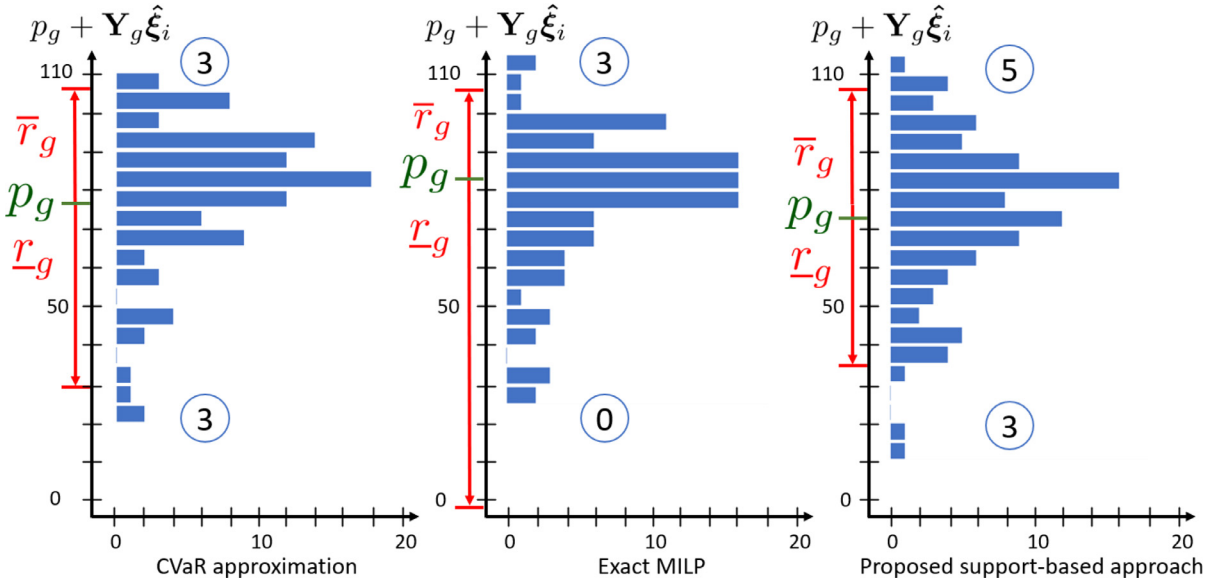


Fig. 6. In-sample energy and reserve dispatch outcomes for conventional generator $g = 2$ in three models: Blue bars show the distribution of actual power production, i.e, the day-ahead energy dispatch p_g plus recourse action $\mathbf{Y}_g\hat{\boldsymbol{\xi}}_i$, across 100 in-sample observations $\hat{\boldsymbol{\xi}}_1$, ..., $\hat{\boldsymbol{\xi}}_{100}$. The energy dispatch p_g as well as the upward and downward reserve dispatches $\overline{r}_g$ and $\underline{r}_g$ are shown with green number and red arrows, respectively. The number of observations with violated downward and upward reserve constraints is given in blue circles (fixed values: $\rho = 10^{-4}$ and $\epsilon = 0.05$).

practice, the system operator has the opportunity to re-optimize decisions in real time to keep the balance between generation and consumption as well as to restore the feasibility by introducing extreme recourse actions such as costly load curtailment and wind spillage. The next section allows the system operator to re-optimize the recourse actions in real time.

4.4. Out-of-sample outcomes: Total system cost with re-optimization

For a given day-ahead dispatch and under each out-of-sample realization of uncertainty $\hat{\boldsymbol{\xi}}_j$, we solve a deterministic optimization problem, the so-called *re-optimization*, in real time to make decisions on the optimal recourse actions including load curtailment and wind spillage. The formulation of such an optimization problem is available in Appendix E. We take into account a cost of \$500/MWh for load curtailment,[17] while assuming a zero cost for wind spillage. There is no energy storage system in our case study to store the excess wind energy. Similar to Fig. 7, we report in Fig. 8 the average out-of-sample total system cost and its standard deviation with consideration of the re-optimization process.

We compare all three models against the Sample Average Approximation (SAA) corresponding to scenario-based stochastic programming. To apply this method, we treat the N number of in-sample observations as equiprobable scenarios and each DRCC is replaced by N number of recourse constraints, one per scenario.[18] It is worth mentioning that the SAA does not directly model the potential risk attitude of the system operator, while such an attitude is adjusted by ρ and ϵ in DRO. This is the reason why the results of SAA in Fig. 8 are unchanged as a function of ρ. Moreover, one can interpret the DRO with a large value of ρ as a robust counterpart, in the sense that extreme realizations of the uncertainty are considered.

These results suggest that DRO outperforms SAA when the number of training samples N is comparatively low. This is the case where there is the lack of sufficient information about uncertainty or the number of training samples is intentionally kept below due to computational issues. To the best of our knowledge, the reserve market in Switzerland (Abbaspourtorbati & Zima, 2016) is the sole electricity market in practice with a stochastic framework, which takes into account a very limited number of scenarios. In such a market with limited information about uncertain parameters, we hypothesize a DRO model has potential to outperform other techniques, including the scenario-based stochastic programming.

One can draw three additional important observations from Fig. 8. First, similar to Fig. 7, our proposed support-based approach provides in general better results in terms of average total operational cost of the system than the CVaR approximation and the exact MILP method without support, while the cost variability in all these three models are similar. Second, there exist several values for ρ under which the proposed DRO model outperforms the SAA method. For example, in Fig. 8(d), the proposed DRO for a range of ρ from 10^{-4} to around 10^{-3} obtains a lower average cost than the SAA. Third, the proposed DRO provides a flexibility for the decision-maker to easily adjust ρ in such a way that it achieves either a solution better or similar to SAA, or a solution close to the robust one. For example, the value of ρ in Fig. 8(d) between 10^{-3} and 10^{-2} provides a solution that is lying between SAA and robust solutions. For values larger than 10^{-2}, the proposed DRO provides fully robust solutions.

Next, we provide three-dimensional plots of the out-of-sample average total operational cost of the system as a function of ρ and ϵ, which are presented in Figs. 9(a)–(c). These results suggest that the total operational cost decreases when ϵ increases. This confirms our initial intuition that the chance-constrained programming allows to reduce the total operational cost of the system by allowing constraint violation to some extent. This implies that the

[17] A change in recourse action costs may affect the out-of-sample performance. If recourse actions become too expensive, the distributionally robust chance-constrained formulation may even result in a higher total operational cost. In such a case, the optimal choice for (ρ, ϵ) will be $(\infty, 0)$. However, this case may not be foreseen, which in turn makes the use of DRCCs useful to find the optimal trade-off between reserve procurement and load curtailment costs.

[18] Contrary to this approach, the chance-constrained SAA models the constraints as classical chance constraints. In view of Remark 1 in Chen et al. (2018), the latter technique may not be achieved by setting $\rho = 0$. In addition, chance-constrained models are known to be computationally expensive and require analytical reformulations, which remain outside the scope of this paper.

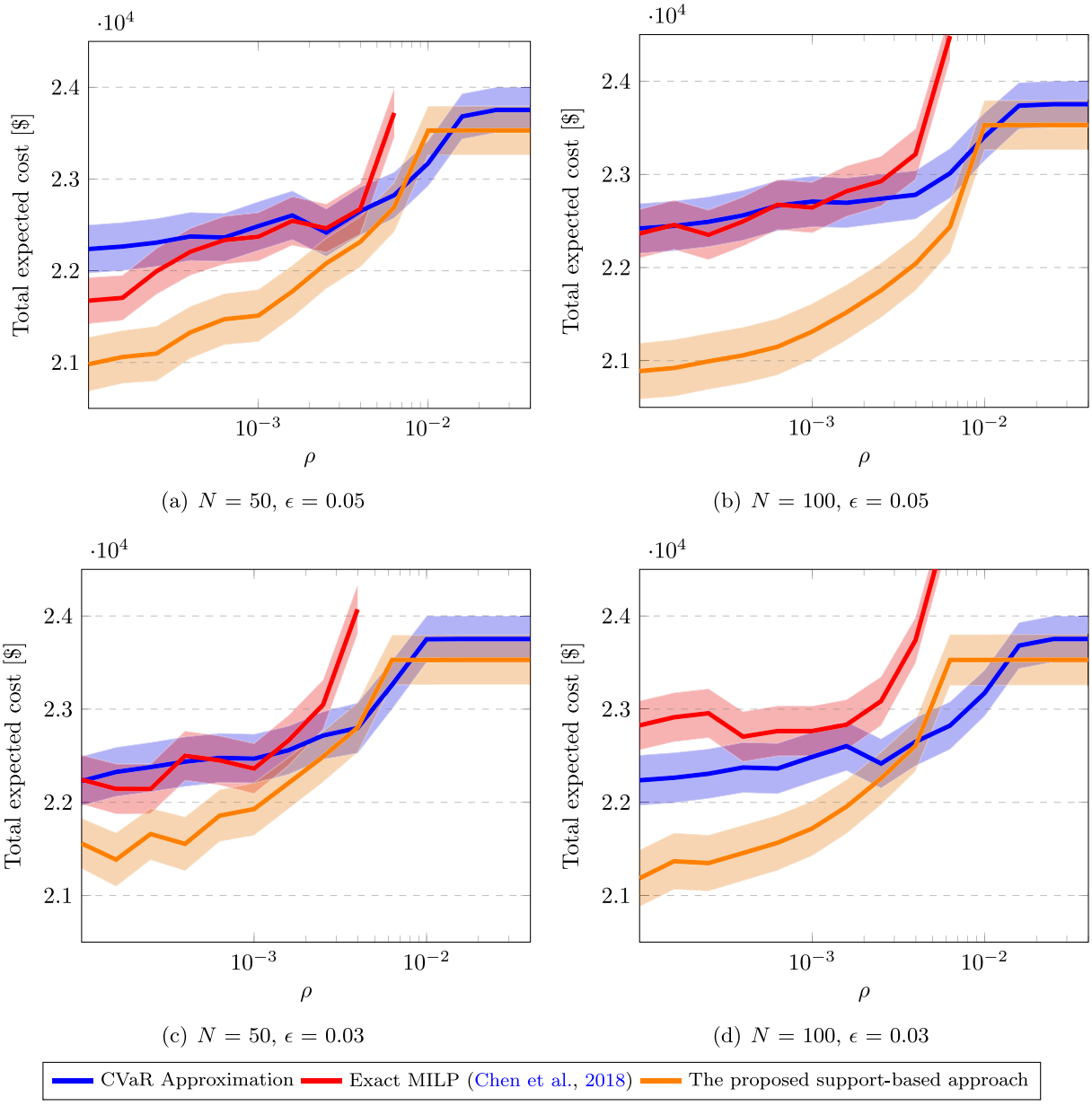


Fig. 7. Out-of-sample analysis without re-optimization: Evolution of the total average system cost and its standard deviation as a function of ρ for various sets of parameters N and ϵ.

distributionally robust chance-constrained program finds an optimal trade-off between operational cost and reliability. These figures further show that the exact MILP model becomes infeasible when the value assigned for the constraint violation probability ϵ is comparatively low.

We emphasize that the optimal ex-ante selection of radius ρ is not straightforward and may require learning from experience, i.e., past outcome of the DRO model. We have observed that there exists an optimal value for ρ that leads to the minimum out-of-sample expected operational cost. Accordingly, we suggest that the decision-maker may select a value for ρ ranging from the value that provides the best out-of-sample performance to infinity, leading to similar decisions to those of robust optimization. By doing so, the decision-maker is able to adapt her risk attitude. In general, choosing the value of ρ that correctly mimics the risk attitude of the decision-maker is challenging and is closely dependent on her preferences in terms of cost optimality, operational reliability, computational performance, etc.

4.5. Expected energy not served

In this section, we present additional numerical results in terms of the Expected Energy Not Served (EENS). This indicator shows how much load (in MWh) is expected to be curtailed. It also illustrates the violation severity of upward reserve constraints. When the upward reserve is not sufficient, the system operator loses her ability to deliver energy to the consumer and therefore load curtailment occurs. We calculate the EENS for each model when $\epsilon = \{0.03, 0.05\}$ and for Wasserstein radii ranging from 10^{-4} to 10^{-1}. The results are given in Fig. 10.

We observe that EENS tends to decrease when ρ increases as shown in Fig. 10. These results reflect the relation that exists between ρ and ϵ, which respectively set the distributional robustness of the optimal decisions and the maximum allowed violation probability of a given constraint. Even though their interpretation is different, those parameters are closely connected in the way they influence the scheduling decisions. For instance, a given vio-

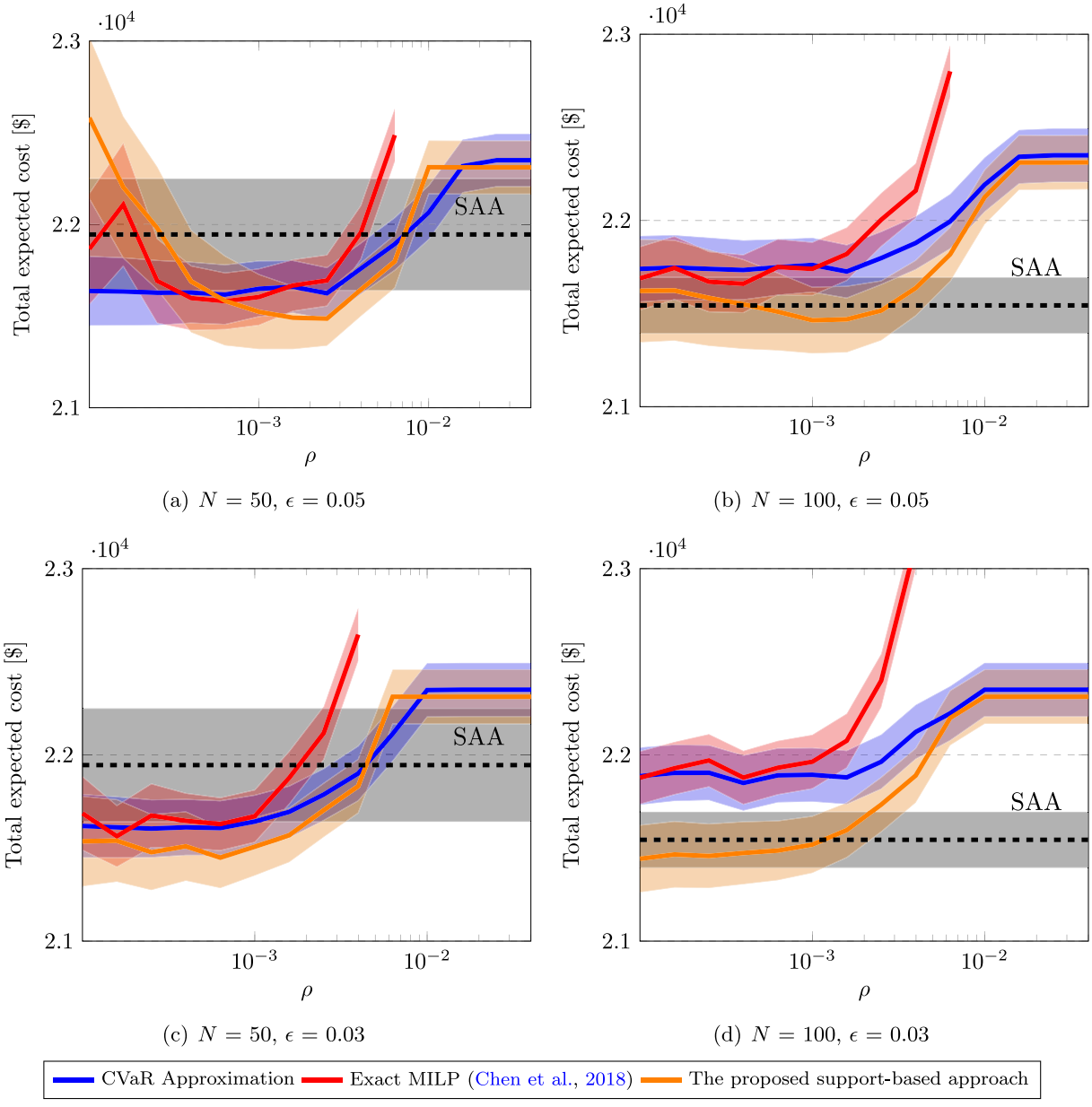


Fig. 8. Out-of-sample analysis with re-optimization: Evolution of the total average system cost and its standard deviation as a function of ρ for various set of parameters N and ϵ.

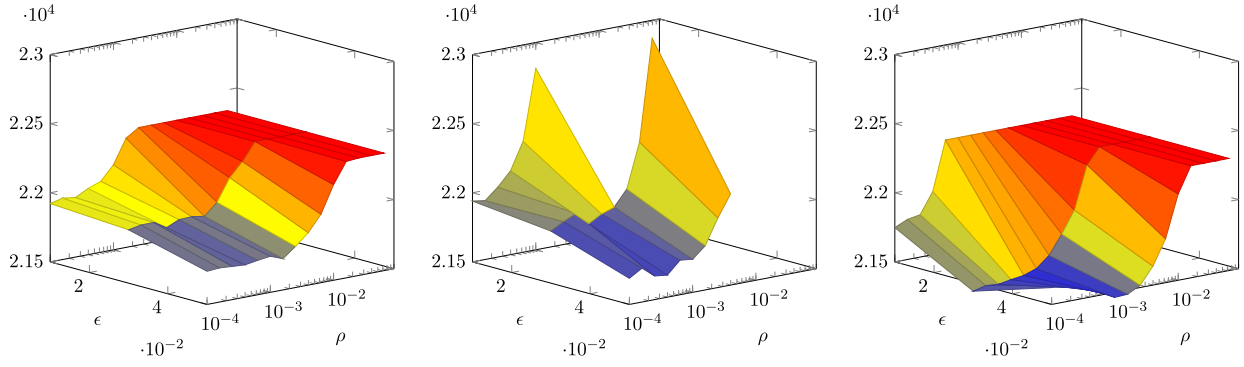


Fig. 9. The average total operational cost of the system with re-optimization as a function of ρ and ϵ.

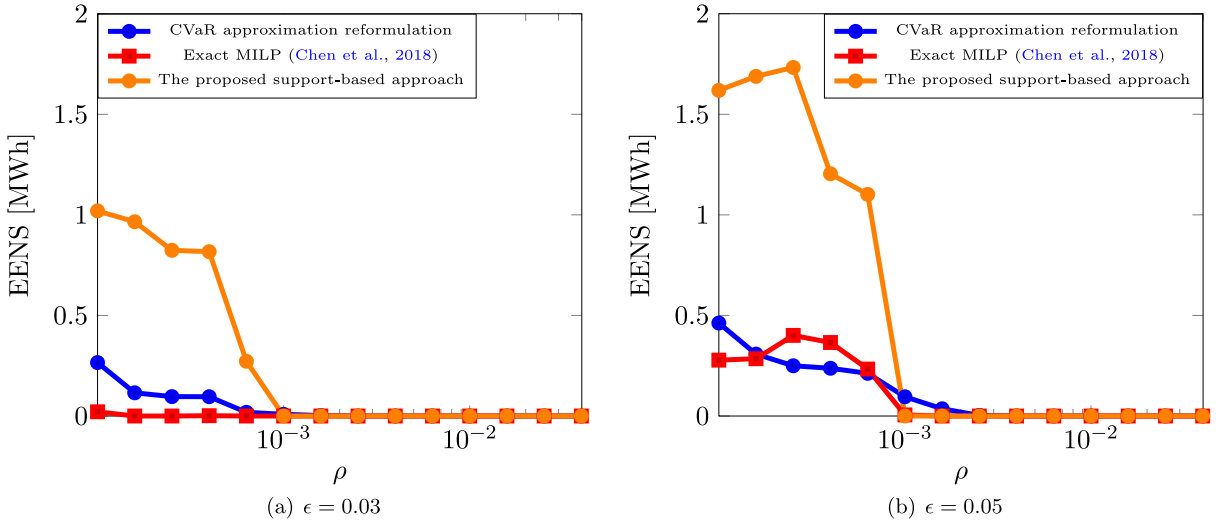


Fig. 10. Expected energy not served (EENS) as a function of ρ (fixed value: $N = 100$).

lation probability ϵ will not result in the same empirical violation probabilities for two different distributions within the ambiguity set, e.g., the empirical one and the worst-case one. When the distributional robustness increases, i.e., ρ takes a comparatively high value, the EENS decreases for a given violation probability ϵ. The reason for this is that the worst-case distribution in the ambiguity set (for which the decisions are optimized) results in more conservative solutions when ρ increases.

Finally, it is worth mentioning that the overall system reliability violation, i.e., the probability of violating at least one constraint, always reaches to 100%. This naive result comes from the fact that the allowed *individual* violation probability is comparatively high, e.g., 3% or 5% in our case study, with respect to the overall system. Our numerical case study contains 58 individual chance constraints, which means that the overall violation probability would evaluate to 100% if a Bonferroni approximation would be used. Hence, even if the overall violation probability is comparatively high, the magnitude of violation remains acceptable. To support our claim, we observe that the EENS never takes a value higher than 1.8 MWh per hour. We numerically conclude that the reliability of our approach based on individual chance constraints is satisfactory from the system operator's perspective.

5. Conclusion

In this paper, we develop a Wasserstein distributionally robust chance-constrained energy and reserve dispatch model, and provide three different reformulations: (*i*) a CVaR approximation resulting in a linear program, (*ii*) an exact reformulation without support, yielding an MILP, and (*iii*) a physically-bounded exact reformulation. The last reformulation is proposed in this paper, which results in a bilinear program. This problem is solved using an efficient iterative algorithm, and its computational performance is compared with that of a non-linear solver, IPOPT. Through an extensive out-of-sample study, we show that our proposed reformulation enables the power system operator to make more informed dispatch decisions by including physical bounds within the exact reformulation of DRCCs. We observe that our proposed support-based method outperforms other techniques specially in cases wherein the value for ρ is properly selected.

One interesting future research path is to explore methodologies for ex-ante or endogenous determination of the optimal value for ρ to further improve the applicability of the methodology for actual field operations. This, for example, requires a cross-fold validation or a dedicated machine learning technique, leveraging information from past realizations. As another potential research path, a numerical discussion could shed light on how the tempo-spatial correlations affect the risk attitude of the decision-maker. It is of interest to explore how such correlations could be included within the problem formulation.

Acknowledgment

We would like to thank Martin S. Andersen and Anubhav Ratha (Technical University of Denmark) for our helpful discussions on duality principles applied to conic, quadratic, and bilinear optimization problems.

Appendix A. Nomenclature

Sets

$g \in \mathcal{G}$	:	Set of conventional generators
$l \in \mathcal{L}$	:	Set of transmission lines
$d \in \mathcal{D}$	:	Set of demands
$w \in \mathcal{W}$	:	Set of renewable generators
$i \in \{1, \ldots, N\}$	:	Set of samples (historical observations) for in-sample simulation
$j \in \{1, \ldots, Z\}$	:	Set of samples (historical observations) for out-of-sample simulation
$\mathbb{P} \in \mathcal{P}$	:	Worst-case distribution $\mathbb{P}$ residing within ambiguity set $\mathcal{P}$
$\mathcal{S}(\mathcal{Y})$	:	Safe set
$\overline{\mathcal{S}}(\mathcal{Y})$	:	Unsafe set

Parameters

$\mathbf{c} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of production cost of conventional generators [$/MWh]
$\overline{\mathbf{c}} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of procurement cost of upward reserve from conventional generators [$/MW]
$\underline{\mathbf{c}} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of procurement cost of downward reserve from conventional generators [$/MW]
$\mathbf{d} \in \mathbb{R}^{\lvert\mathcal{D}\rvert}$	:	Vector of consumption level of demands [MW]
$\mathbf{f}^{\max} \in \mathbb{R}^{\lvert\mathcal{L}\rvert}$	:	Vector of capacity of transmission lines [MW]
$\mathbf{r}^{\max} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of maximum reserve provision capability of conventional generators [MW]
$\mathbf{p}^{\max} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of capacity of conventional generators [MW]

(*continued on next column*)

$\mathbf{W} \in \mathbb{R}^{\lvert\mathcal{W}\rvert\times\lvert\mathcal{W}\rvert}$	:	Diagonal matrix of installed capacity of renewable generators [MW]
$\mathbf{T}^{\mathcal{G}} \in \mathbb{R}^{\lvert\mathcal{L}\rvert\times\lvert\mathcal{G}\rvert}$	:	Matrix of power transfer distribution factor for conventional generators
$\mathbf{T}^{\mathcal{W}} \in \mathbb{R}^{\lvert\mathcal{L}\rvert\times\lvert\mathcal{W}\rvert}$	:	Matrix of power transfer distribution factor for renewable generators
$\mathbf{T}^{\mathcal{D}} \in \mathbb{R}^{\lvert\mathcal{L}\rvert\times\lvert\mathcal{D}\rvert}$	:	Matrix of power transfer distribution factor for demands
$\overline{\epsilon}_g, \underline{\epsilon}_g, \epsilon_l \in \mathbb{R}$	:	Violation probabilities
$\boldsymbol{\mu} \in \mathbb{R}^{\lvert\mathcal{W}\rvert}$	:	Vector of day-ahead power production forecast of renewable generators [p.u.]
$\mathbf{v}_{\text{Shed}} \in \mathbb{R}^{\lvert\mathcal{D}\rvert}$	:	Vector of curtailment cost of demands [$/MWh]
$\rho \in \mathbb{R}$	:	Wasserstein ball radius
$\mathbf{Q} \in \mathbb{R}^{2\lvert\mathcal{W}\rvert\times\lvert\mathcal{W}\rvert}$	:	Matrix for the support definition
$\mathbf{h} \in \mathbb{R}^{2\lvert\mathcal{W}\rvert}$	:	Vector for the support definition
$\tilde{\boldsymbol{\xi}} \in \mathbb{R}^{\lvert\mathcal{W}\rvert}$	:	Vector of renewable power deviations from day-ahead forecast (random variables) [p.u.]
$\hat{\boldsymbol{\xi}}_i \in \mathbb{R}^{\lvert\mathcal{W}\rvert}$	:	Vector of historical observation (sample) i of random variables $\tilde{\boldsymbol{\xi}}$[p.u.]
$\hat{\boldsymbol{\xi}}_j \in \mathbb{R}^{\lvert\mathcal{W}\rvert}$	:	Vector of realization j of random variables $\tilde{\boldsymbol{\xi}}$ in the real-time operation
Variables		
$\mathbf{p} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of power dispatch of conventional generators [MW]
$\overline{\mathbf{r}} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of upward reserve dispatch of conventional generators [MW]
$\underline{\mathbf{r}} \in \mathbb{R}^{\lvert\mathcal{G}\rvert}$	:	Vector of downward reserve dispatch of conventional generators [MW]
$\mathbf{Y} \in \mathbb{R}^{\lvert\mathcal{G}\rvert\times\lvert\mathcal{W}\rvert}$	:	Matrix of participation factor of conventional generators [MW]
$\mathcal{Y}$	:	Vector of day-ahead decisions $\{\mathbf{p}, \overline{\mathbf{r}}, \underline{\mathbf{r}}, \mathbf{Y}\}$

Appendix B. Approximate linear reformulation of model (1) with CVaR constraints

As thoroughly discussed in Zymler et al. (2013), a DRCC can be conservatively approximated by a constraint including the CVaR at level ϵ with respect to $\mathbb{P}$. This is mathematically stated by the following implication:

$$\max_{\mathbb{P}\in\mathcal{P}} \ \mathbb{P}\text{-CVaR}_{\epsilon}\left(\mathbf{a}^{\top}\boldsymbol{\xi} - b\right) \leq 0 \Rightarrow \min_{\mathbb{P}\in\mathcal{P}} \ \mathbb{P}\left(\mathbf{a}^{\top}\boldsymbol{\xi} \leq b\right) \geq 1-\epsilon. \tag{B.1}$$

Eq. (B.1) states that the CVaR formulation (on the left-hand side) is sufficient to impose the DRCC. It is a conservative approximation, because the CVaR accounts for the violation magnitude and will eventually impose the constraint with a higher probability than *a priori* required. Using the definition of the CVaR, the left-hand side of (B.1) is cast as

$$\max_{\mathbb{P}\in\mathcal{P}} \ \min_{\tau\in\mathbb{R}} \ \tau + \frac{1}{\epsilon} \ \mathbb{E}^{\mathbb{P}}\left[\left\lceil \mathbf{a}^{\top}\boldsymbol{\xi} - b - \tau \right\rceil^{+}\right] \leq 0, \tag{B.2}$$

where $\tau \in \mathbb{R}$ is an auxiliary variable and $\lceil . \rceil^{+} = \max(0, .)$. After rearranging the order of the optimization operators, this formulation is equivalent to

$$\min_{\tau\in\mathbb{R}} \ \tau + \frac{1}{\epsilon} \ \max_{\mathbb{P}\in\mathcal{P}} \ \mathbb{E}^{\mathbb{P}}\left[\left\lceil \mathbf{a}^{\top}\boldsymbol{\xi} - b - \tau \right\rceil^{+}\right] \leq 0. \tag{B.3}$$

We next reformulate the worst-case expectation, following a similar procedure as the one described by Eq. (4) in Section 3. This approach results in a min-min formulation, where we merge the min operators. Eventually, we equivalently drop the min operators and add auxiliary variables, resulting in the following set of equations, which represents the CVaR approximation of a DRCC:

$$\tau + \frac{1}{\epsilon}\left(\lambda^{\text{CVaR}}\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i^{\text{CVaR}}\right) \leq 0 \tag{B.4a}$$

$$\mathbf{a}^{\top}\hat{\boldsymbol{\xi}}_i - b - \tau + \boldsymbol{\gamma}_{i,1}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i^{\text{CVaR}} \qquad \forall i \in \{1,\ldots,N\} \tag{B.4b}$$

$$\boldsymbol{\gamma}_{i,2}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i^{\text{CVaR}} \qquad \forall i \in \{1,\ldots,N\} \tag{B.4c}$$

$$\left|\left|\mathbf{Q}^{\top}\boldsymbol{\gamma}_{i,1} - \mathbf{a}\right|\right|_{*} \leq \lambda^{\text{CVaR}} \qquad \forall i \in \{1,\ldots,N\} \tag{B.4d}$$

$$\left|\left|\mathbf{Q}^{\top}\boldsymbol{\gamma}_{i,2}\right|\right|_{*} \leq \lambda^{\text{CVaR}} \qquad \forall i \in \{1,\ldots,N\} \tag{B.4e}$$

$$\boldsymbol{\gamma}_i \geq 0 \qquad \forall i \in \{1,\ldots,N\}, \tag{B.4f}$$

where $\boldsymbol{\lambda}^{\text{CVaR}} \in \mathbb{R}^N$, $\boldsymbol{\sigma}^{\text{CVaR}} \in \mathbb{R}^N$, $\boldsymbol{\gamma}_{i,1} \in \mathbb{R}^{2|\mathcal{W}|}, \boldsymbol{\gamma}_{i,2} \in \mathbb{R}^{2|\mathcal{W}|}$ and $\tau \in \mathbb{R}$ are auxiliary variables. Note that support $\mathbf{Q}\boldsymbol{\xi} \leq \mathbf{h}$ is incorporated into formulation (B.4).

Recall that the objective function in model (1) has been already reformulated by Eq. (4) in Section 3. Given the CVaR approximation of DRCCs in (B.4), the energy and reserve dispatch model (1) is eventually reformulated into a linear program as

$$\min_{\Theta^{\text{CVaR}}} \ \mathbf{c}^{\top}\mathbf{p} + \overline{\mathbf{c}}^{\top}\overline{\mathbf{r}} + \underline{\mathbf{c}}^{\top}\underline{\mathbf{r}} + \lambda\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i \tag{B.5a}$$

$$\text{s.t.} \ \ \mathbf{p} + \overline{\mathbf{r}} \leq \mathbf{p}^{\max} \tag{B.5b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \geq \mathbf{0} \tag{B.5c}$$

$$\mathbf{0} \leq \underline{\mathbf{r}} \leq \mathbf{r}^{\max}; \mathbf{0} \leq \overline{\mathbf{r}} \leq \mathbf{r}^{\max} \tag{B.5d}$$

$$\mathbf{e}^{\top}\mathbf{p} + \mathbf{e}^{\top}\mathbf{W}\boldsymbol{\mu} - \mathbf{e}^{\top}\mathbf{d} = 0 \tag{B.5e}$$

$$\mathbf{Y}^{\top}\mathbf{e} + \mathbf{W}^{\top}\mathbf{e} = \mathbf{0} \tag{B.5f}$$

$$\begin{cases} \mathbf{c}^{\top}\mathbf{Y}\hat{\boldsymbol{\xi}}_i + \boldsymbol{\gamma}_i^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_i & \forall i \in \{1,\ldots,N\} \\ \left|\left|\mathbf{Q}^{\top}\boldsymbol{\gamma}_i - \mathbf{c}^{\top}\mathbf{Y}\right|\right|_{*} \leq \lambda & \forall i \in \{1,\ldots,N\} \\ \boldsymbol{\gamma}_i \geq 0 & \forall i \in \{1,\ldots,N\} \end{cases} \tag{B.5g}$$

$$\left.\begin{cases} \overline{\tau}_g + \frac{1}{\epsilon}\left(\overline{\lambda}_g\rho + \frac{1}{N}\sum_{i=1}^{N}\overline{\sigma}_{g,i}\right) \leq 0 & \\ \mathbf{Y}_g\hat{\boldsymbol{\xi}}_i - \overline{r}_g - \overline{\tau}_g + \overline{\boldsymbol{\gamma}}_{g,i,1}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \overline{\sigma}_{g,i} & \forall i \in \{1,\ldots,N\} \\ \overline{\boldsymbol{\gamma}}_{g,i,2}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \overline{\sigma}_{g,i} & \forall i \in \{1,\ldots,N\} \\ \left|\left|\mathbf{Q}^{\top}\overline{\boldsymbol{\gamma}}_{g,i,1} - \mathbf{Y}_g\right|\right|_{*} \leq \overline{\lambda}_g & \forall i \in \{1,\ldots,N\} \\ \left|\left|\mathbf{Q}^{\top}\overline{\boldsymbol{\gamma}}_{g,i,2}\right|\right|_{*} \leq \overline{\lambda}_g & \forall i \in \{1,\ldots,N\} \\ \overline{\boldsymbol{\gamma}}_{g,i,1} \geq 0; \overline{\boldsymbol{\gamma}}_{g,i,2} \geq 0 & \forall i \in \{1,\ldots,N\} \end{cases}\right\} \ \forall g \in \mathcal{G} \tag{B.5h}$$

$$\left.\begin{cases} \underline{\tau}_g + \frac{1}{\epsilon}\left(\underline{\lambda}_g\rho + \frac{1}{N}\sum_{i=1}^{N}\underline{\sigma}_{g,i}\right) \leq 0 & \\ -\mathbf{Y}_g\hat{\boldsymbol{\xi}}_i - \underline{r}_g - \underline{\tau}_g + \underline{\boldsymbol{\gamma}}_{g,i,1}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \underline{\sigma}_{g,i} & \forall i \in \{1,\ldots,N\} \\ \underline{\boldsymbol{\gamma}}_{g,i,2}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \underline{\sigma}_{g,i} & \forall i \in \{1,\ldots,N\} \\ \left|\left|\mathbf{Q}^{\top}\underline{\boldsymbol{\gamma}}_{g,i,1} + \mathbf{Y}_g\right|\right|_{*} \leq \underline{\lambda}_g & \forall i \in \{1,\ldots,N\} \\ \left|\left|\mathbf{Q}^{\top}\underline{\boldsymbol{\gamma}}_{g,i,2}\right|\right|_{*} \leq \underline{\lambda}_g & \forall i \in \{1,\ldots,N\} \\ \underline{\boldsymbol{\gamma}}_{g,i,1} \geq 0; \underline{\boldsymbol{\gamma}}_{g,i,2} \geq 0 & \forall i \in \{1,\ldots,N\} \end{cases}\right\} \ \forall g \in \mathcal{G} \tag{B.5i}$$

$$\left.\begin{cases} \tau_l + \frac{1}{\epsilon}\left(\lambda_l\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_{l,i}\right) \leq 0 & \\ \left(\mathbf{T}_l^{\mathcal{G}}(\mathbf{p} + \mathbf{Y}\boldsymbol{\xi}) + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}(\boldsymbol{\mu} + \boldsymbol{\xi}) - \mathbf{T}_l^{\mathcal{D}}\mathbf{d} - f_l^{\max}\right) - \tau_l + \boldsymbol{\gamma}_{l,i,1}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_{l,i} & \forall i \in \{1,\ldots,N\} \\ \boldsymbol{\gamma}_{l,i,2}^{\top}\left(\mathbf{h} - \mathbf{Q}\hat{\boldsymbol{\xi}}_i\right) \leq \sigma_{l,i} & \forall i \in \{1,\ldots,N\} \\ \left|\left|\mathbf{Q}^{\top}\boldsymbol{\gamma}_{l,i,1} - \left(\mathbf{T}_l^{\mathcal{G}}\mathbf{Y} + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\right)\right|\right|_{*} \leq \lambda_l & \forall i \in \{1,\ldots,N\} \\ \left|\left|\mathbf{Q}^{\top}\boldsymbol{\gamma}_{l,i,2}\right|\right|_{*} \leq \lambda_l & \forall i \in \{1,\ldots,N\} \\ \boldsymbol{\gamma}_i \geq 0 & \forall i \in \{1,\ldots,N\} \end{cases}\right\} \ \forall l \in \mathcal{L}, \tag{B.5j}$$

where the variable set Θ^{CVaR} includes $\mathbf{p}$, $\overline{\mathbf{r}}$, $\underline{\mathbf{r}}$, $\mathbf{Y}$, $\overline{\tau}_g$, $\overline{\lambda}_g$, $\overline{\sigma}_{g,i}$, $\overline{\boldsymbol{\gamma}}_{g,i,1}$, $\overline{\boldsymbol{\gamma}}_{g,i,2}$, $\underline{\tau}_g$, $\underline{\lambda}_g$, $\underline{\sigma}_{g,i}$, $\underline{\boldsymbol{\gamma}}_{g,i,1}$, $\underline{\boldsymbol{\gamma}}_{g,i,2}$ τ_l, λ_l, $\sigma_{l,i}$, $\boldsymbol{\gamma}_{l,i,1}$, $\boldsymbol{\gamma}_{l,i,2}$, λ, σ_i and $\boldsymbol{\gamma}_i$. Note that constraints (B.5b)–(B.5f) come from the original model,

whereas constraints (B.5g) correspond to the objective function reformulation. Finally, constraints (B.5h), (B.5i) and (B.5j) present the CVaR approximation of DRCCs (1g), (1h) and (1i), respectively.

Appendix C. Exact MILP reformulation of model (1) without support

Given the exact reformulation of DRCCs presented in Section 3.1, the energy and reserve dispatch model (1) is reformulated into a mixed-integer linear program as

$$\min_{\Theta^{\text{MILP}}} \ \mathbf{c}^\top \mathbf{p} + \overline{\mathbf{c}}^\top \overline{\mathbf{r}} + \underline{\mathbf{c}}^\top \underline{\mathbf{r}} + \lambda\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i \tag{C.1a}$$

$$\text{s.t.} \ \ \mathbf{p} + \overline{\mathbf{r}} \le \mathbf{p}^{\max} \tag{C.1b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \ge \mathbf{0} \tag{C.1c}$$

$$\mathbf{0} \le \underline{\mathbf{r}} \le \mathbf{r}^{\max};\ \mathbf{0} \le \overline{\mathbf{r}} \le \mathbf{r}^{\max} \tag{C.1d}$$

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

$$\mathbf{Y}^\top \mathbf{e} + \mathbf{W}^\top \mathbf{e} = \mathbf{0} \tag{C.1f}$$

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

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

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

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

where the variable set Θ^{MILP} includes $\mathbf{p}$, $\overline{\mathbf{r}}$, $\underline{\mathbf{r}}$, $\mathbf{Y}$, $\overline{t}_g$, $\overline{\boldsymbol{\beta}}_g$, $\overline{\boldsymbol{q}}_g$, $\underline{t}_g$, $\underline{\boldsymbol{\beta}}_g$, $\underline{\boldsymbol{q}}_g$, t_l, $\boldsymbol{\beta}_l$, $\boldsymbol{q}_l$, λ, σ_i and $\boldsymbol{\gamma}_i$. Note that constraints (C.1b)–(C.1f) are identical to those within the original model (1). Constraints (C.1g) are associated with the objective function reformulation. Finally, constraints (C.1h), (C.1i) and (C.1j) present the exact MILP reformulation of DRCCs (1g), (1h) and (1i), respectively.

Appendix D. Physically-bounded exact reformulation of model (1)

Following our proposed physically-bounded exact reformulation of DRCCs presented in Section 3.3, the energy and reserve dispatch model (1) is reformulated into a bilinear program as

$$\min_{\Theta^{\text{Exact}}} \ \mathbf{c}^\top \mathbf{p} + \overline{\mathbf{c}}^\top \overline{\mathbf{r}} + \underline{\mathbf{c}}^\top \underline{\mathbf{r}} + \lambda\rho + \frac{1}{N}\sum_{i=1}^{N}\sigma_i \tag{D.1a}$$

$$\text{s.t.} \ \ \mathbf{p} + \overline{\mathbf{r}} \le \mathbf{p}^{\max} \tag{D.1b}$$

$$\mathbf{p} - \underline{\mathbf{r}} \ge \mathbf{0} \tag{D.1c}$$

$$\mathbf{0} \le \underline{\mathbf{r}} \le \mathbf{r}^{\max};\ \mathbf{0} \le \overline{\mathbf{r}} \le \mathbf{r}^{\max} \tag{D.1d}$$

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

$$\mathbf{Y}^\top \mathbf{e} + \mathbf{W}^\top \mathbf{e} = \mathbf{0} \tag{D.1f}$$

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

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

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

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

where the variable set Θ^{Exact} includes $\mathbf{p}$, $\overline{\mathbf{r}}$, $\underline{\mathbf{r}}$, $\mathbf{Y}$, $\overline{t}_g$, $\overline{\boldsymbol{\beta}}_g$, $\overline{w}_{g,i}$, $\overline{\boldsymbol{x}}_{g,i}$, $\underline{t}_g$, $\underline{\boldsymbol{\beta}}_g$, $\underline{w}_{g,i}$, $\underline{\boldsymbol{x}}_{g,i}$, t_l, $\boldsymbol{\beta}_l$, $w_{l,i}$, $\boldsymbol{x}_{l,i}$, λ, σ_i and $\boldsymbol{\gamma}_i$. Constraints (D.1b)-(D.1f) are identical to the operational limits of conventional generators and power balance conditions within the original model (1). Constraints (D.1g) pertain to the objective function reformulation. Finally, constraints (D.1h), (D.1i) and (D.1j) present the physically-bounded exact reformulation of DRCCs (1g), (1h) and (1i), respectively.

Appendix E. Optimization problem in the real-time operation

This appendix presents the optimization problem used in Section 4.4 for the out-of-sample simulations. For given day-ahead dispatch decisions ($\mathbf{p}^{\text{fixed}}, \underline{\mathbf{r}}^{\text{fixed}}, \overline{\mathbf{r}}^{\text{fixed}}$) and realized renewable power deviation $\hat{\boldsymbol{\xi}}_j$, the power system operator solves a deterministic linear program during the real-time operation as

$$\min_{\mathbf{Y},\boldsymbol{\Delta}\mathbf{d},\boldsymbol{\Delta}\mathbf{w}} \ \mathbf{c}^\top \mathbf{Y}\hat{\boldsymbol{\xi}}_j + \mathbf{v}_{\text{Shed}}^\top \boldsymbol{\Delta}\mathbf{d} \tag{E.1a}$$

$$\text{s.t.} \ \ 0 \le \boldsymbol{\Delta}\mathbf{d} \le \mathbf{d} \tag{E.1b}$$

$$0 \leq \boldsymbol{\Delta w} \leq \mathbf{W}(\boldsymbol{\mu} + \hat{\boldsymbol{\xi}}_j) \tag{E.1c}$$

$$-\underline{\mathbf{r}}^{\text{fixed}} \leq \mathbf{Y}\hat{\boldsymbol{\xi}}_j \leq \overline{\mathbf{r}}^{\text{fixed}} \tag{E.1d}$$

$$\mathbf{e}^\top \mathbf{Y}\hat{\boldsymbol{\xi}}_j + \mathbf{e}^\top \mathbf{W}\hat{\boldsymbol{\xi}}_j + \mathbf{e}^\top \boldsymbol{\Delta d} - \mathbf{e}^\top \boldsymbol{\Delta w} = 0 \tag{E.1e}$$

$$\mathbf{T}_l^{\mathcal{G}}\left(\mathbf{p}^{\text{fixed}} + \mathbf{Y}\hat{\boldsymbol{\xi}}_j\right) + \mathbf{T}_l^{\mathcal{W}}\mathbf{W}\left(\boldsymbol{\mu} + \hat{\boldsymbol{\xi}}_j\right) - \mathbf{T}_l^{\mathcal{D}}\mathbf{d} \leq f_l^{\max} \qquad \forall l \in \mathcal{L}. \tag{E.1f}$$

The objective function (E.1a) minimizes the total operational cost of the system for adjusting imbalances, including the recourse cost of conventional generators (first term) and the load curtailment cost (second term). Note that the recourse action of conventional generators $\mathbf{Y} \in \mathbb{R}^{|\mathcal{G}|\times|\mathcal{W}|}$ is a decision variable, meaning that those generators are not imposed to stick to their participation factors obtained in the day-ahead stage. Parameter vector $\mathbf{v}_{\text{Shed}} \in \mathbb{R}^{|\mathcal{D}|}$ gives the load curtailment cost of demands, whereas variable vector $\boldsymbol{\Delta d} \in \mathbb{R}^{|\mathcal{D}|}$ provides the optimal load curtailment quantities. In addition to the recourse action of conventional generators and the load curtailment of demands, the third adjustment option for the system operator is the renewable power spillage $\boldsymbol{\Delta w} \in \mathbb{R}^{|\mathcal{W}|}$, which is assumed to be cost-free. Constraint (E.1b) restricts the load curtailment quantity to lie within zero and the consumption level. Similarly, (E.1c) enforces the renewable power spillage quantity to be non-negative, and not exceed the realized level of renewable power generation. The recourse action of conventional generators is limited in (E.1d) by their reserve dispatch fixed in the day-ahead stage. The power balance is ensured by (E.1e). Finally, the capacity limit of transmission lines is enforced by (E.1f).

Appendix F. Input data

Our case study in Section 4 is built upon the IEEE 24-node reliability test system (Grigg et al., 1999). The data for wind farms is borrowed from Ordoudis, Pinson, Morales, and Zugno (2016). The network topology of the system is illustrated in Fig. F.11. In addi-

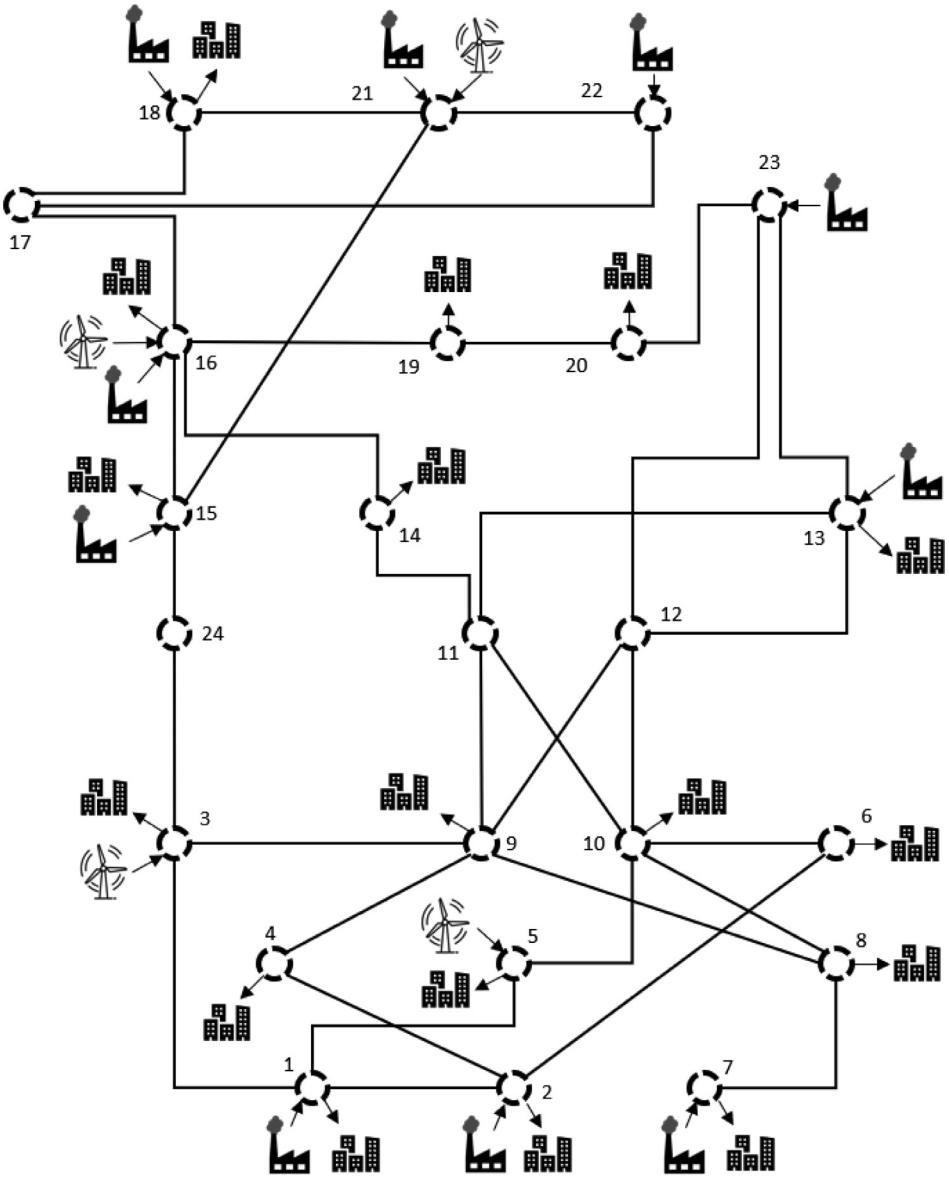


Fig. F.11. The network topology of the IEEE 24-node reliability test system (Grigg et al., 1999).

Table F3
Technical data for conventional generators, wind farms, demands and transmission lines.

Conventional generators	1	2	3	4	5	6	7	8	9	10	11	12
Location [node]	1	2	7	13	15	15	16	18	21	22	23	23
Production cost [$/MWh]	13.32	13.32	20.7	20.93	26.11	10.52	10.52	6.02	5.47	7	10.52	10.89
Upward reserve cost [$/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
Downward reserve cost [$/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
Capacity [MW]	106.4	106.4	245	413.7	42	108.5	108.5	280	280	210	217	245
Maximum upward reserve provision capability [MW]	48	48	84	216	42	36	36	60	60	48	72	48
Maximum downward reserve provision capability [MW]	48	48	84	216	42	36	36	60	60	48	72	48

Wind farms	1	2	3	4
Location [node]	3	5	16	21
Installed capacity [MW]	500	500	300	300
Day-ahead forecast [MW]	120.54	115.52	53.34	38.16

Demands	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Location [node]	1	2	3	4	5	6	7	8	9	10	13	14	15	16	18	19	20
Consumption [MW]	84	75	139	58	55	106	97	132	135	150	205	150	245	77	258	141	100
Curtailment cost [$/MWh]	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500

Transmission 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
Reactance [per-unit]	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
Capacity [MW]	175	175	350	175	175	175	400	175	350	175	350	175	175	400	400	400	400

Transmission 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
Reactance [per-unit]	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
Capacity [MW]	500	500	500	500	250	250	500	400	500	500	500	500	500	1000	1000	1000	500

tion, all technical data corresponding to conventional generators, wind farms, demands and transmission lines are given in Table F.3.

References

Abbaspourtorbati, F., & Zima, M. (2016). The swiss reserve market: Stochastic programming in practice. *IEEE Transactions on Power Systems, 31*(2), 1188–1194.

Andersen, M. S. (2021). Bilinear constraints, Hackmd note. https://hackmd.io/@martinandersen/S1KaTEYmu.

Arrigo, A., Ordoudis, C., Kazempour, J., Grève, Z. D., Toubeau, J.-F., & Vallée, F. (2021). Online companion - distributionally robust chance constrained energy and reserve dispatch: An exact and physically bounded approach. https://doi.org/10.5281/zenodo.3540810.

Basciftci, B., Ahmed, S., & Shen, S. (2021). Distributionally robust facility location problem under decision-dependent stochastic demand. *European Journal of Operational Research, 292*(2), 548–561.

Bertsimas, D., Litvinov, E., Sun, X. A., Zhao, J., & Zheng, T. (2013). Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Transactions on Power Systems, 28*(1), 52–63.

Bienstock, D., Chertkov, M., & Harnett, S. (2014). Chance-constrained optimal power flow: Risk-aware network control under uncertainty. *SIAM Review, 56*(3), 461–495.

Blanchet, J., Kang, Y., & Murthy, K. (2019). Robust Wasserstein profile inference and applications to machine learning. *Journal of Applied Probability, 56*(3), 830–857.

Bonferroni, C. E. (1936). *Teoria statistica delle classi e calcolo delle probabilita*. Libreria internazionale Seeber.

Bottieau, J., Hubert, L., Grève, Z. D., Vallée, F., & Toubeau, J.-F. (2020). Very short-term probabilistic forecasting for a risk-aware participation in the single price imbalance settlement. *IEEE Transactions on Power Systems, 35*(2), 1218–1230.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Chen, Z., Kuhn, D., & Wiesemann, W. (2018). Data-driven chance constrained programs over Wasserstein balls. http://www.optimization-online.org/DB_FILE/2018/06/6671.pdf.

Christie, R. D., Wollenberg, B. F., & Wangensteen, I. (2000). Transmission management in the deregulated environment. *Proceedings of the IEEE, 88*(2), 170–195.

Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research, 58*(3), 595–612.

Domínguez, R., Oggioni, G., & Smeers, Y. (2019). Reserve procurement and flexibility services in power systems with high renewable capacity: Effects of integration on different market designs. *International Journal of Electrical Power & Energy Systems, 113*, 1014–1034.

Duan, C., Fang, W., Jiang, L., Yao, L., & Liu, J. (2018). Distributionally robust chance–constrained approximate AC-OPF with Wasserstein metric. *IEEE Transactions on Power Systems, 33*(5), 4924–4936.

Georghiou, A., Wiesemann, W., & Kuhn, D. (2015). Generalized decision rule approximations for stochastic programming via liftings. *Mathematical Programming, 152*(1), 301–338.

González, P., Villar, J., Díaz, C. A., & Campos, F. A. (2014). Joint energy and reserve markets: Current implementations and modeling trends. *Electric Power Systems Research, 109*, 101–111.

Grigg, C., Wong, P., Albrecht, P., Allan, R., Bhavaraju, M., Billinton, R., … Singh, C. (1999). The IEEE reliability test system 1996. A report prepared by the reliability test system task force of the application of probability methods subcommittee. *IEEE Transactions on Power Systems, 14*(3), 1010–1020.

Guo, Y., Baker, K., Dall'Anese, E., Hu, Z., & Summers, T. H. (2019). Data-based distributionally robust stochastic optimal power flow. *IEEE Transactions on Power Systems, 34*(2), 1483–1492.

Hobbs, B. F., & Oren, S. S. (2019). Three waves of US reforms: Following the path of wholesale electricity market restructuring. *IEEE Power and Energy Magazine, 17*(1), 73–81.

Hu, Z., & Hong, J. (2013). Kullback–Leibler divergence constrained distributionally robust optimization. http://www.optimization-online.org/DB_FILE/2012/11/3677.pdf.

Jonsson, T., Pinson, P., & Madsen, H. (2010). On the market impact of wind energy forecasts. *Energy Economics, 32*(2), 313–320.

Kantorovich, L. V., & Rubinshtein, G. S. (1958). On a space of totally additive functions. *Vestnik Leningradskogo Universiteta, 13*, 52–59.

Kuhn, D., Mohajerin Esfahani, P., Nguyen, V. A., & Shafieezadeh-Abadeh, S. (2019). Wasserstein distributionally robust optimization: Theory and applications in machine learning. *Operations Research & Management Science in the Age of Analytics*, 130–166.

Kuhn, D., Wiesemann, W., & Georghiou, A. (2011). Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming, 130*(1), 177–209.

Litvinov, E., Zhao, F., & Zheng, T. (2019). Electricity markets in the United States: Power industry restructuring processes for the present and future. *IEEE Power & Energy Magazine, 17*(1), 32–42.

Lubin, M., Dvorkin, Y., & Backhaus, S. (2016). A robust approach to chance constrained optimal power flow with renewable generation. *IEEE Transactions on Power Systems, 31*(5), 3840–3849.

Martin, S., Smeers, Y., & Aguado, J. A. (2015). A stochastic two settlement equilib-

rium model for electricity markets with wind generation. *IEEE Transactions on Power Systems, 30*(1), 233–245.
McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part I - Convex underestimating problems. *Mathematical Programming, 10*(1), 147–175.
Mieth, R., & Dvorkin, Y. (2018). Data-driven distributionally robust optimal power flow for distribution systems. *IEEE Control Systems Letters, 2*(3), 363–368.
Mohajerin Esfahani, P., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming, 171*(1-2), 115–166.
Morales, J., Zugno, M., Pineda, S., & Pinson, P. (2014). Electricity market clearing with improved scheduling of stochastic production. *European Journal of Operational Research, 235*(3), 765–774.
Morales, J. M., Conejo, A. J., & Perez-Ruiz, J. (2009). Economic valuation of reserves in power systems with high penetration of wind power. *IEEE Transactions on Power Systems, 24*(2), 900–910.
Nemirovski, A., & Shapiro, A. (2007). Convex approximations of chance constrained programs. *SIAM Journal of Optimization, 17*(4), 969–996.
NERC (2010). Technical report of North American electric reliability corporation - Flexibility requirements and metrics for variable generation: Implications for system planning studies.https://www.nerc.com/files/IVGTF_Task_1_4_Final.pdf.
Ordoudis, C., Nguyen, V. A., Kuhn, D., & Pinson, P. (2021). Energy and reserve dispatch with distributionally robust joint chance constraints. *Operations Research Letters, 49*(3), 291–299.
Ordoudis, C., Pinson, P., Morales, J. M., & Zugno, M. (2016). An updated version of the IEEE RTS 24-bus system for electricity market and power system operation studies. http://orbit.dtu.dk/files/120568114/An.
Papavasiliou, A., Oren, S. S., & O'Neill, R. P. (2011). Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Transactions on Power Systems, 26*(4), 2197–2206.
Pinson, P. (2013). Wind energy: Forecasting challenges for its operational management. *Statistical Science, 28*(4), 564–585.
Poolla, B. K., Hota, A. R., Bolognani, S., Callaway, D. S., & Cherukuri, A. (2020). Wasserstein distributionally robust look-ahead economic dispatch. *IEEE Transactions on Power Systems, 36*(3), 2010–2022.
Pourahmadi, F., Kazempour, J., Ordoudis, C., Pinson, P., & Hosseini, S. H. (2020). Distributionally robust chance-constrained generation expansion planning. *IEEE Transactions on Power Systems, 35*(4), 2888–2903.
Rahimian, H., & Mehrotra, S. (2019). Distributionally robust optimization: A review. https://arxiv.org/abs/1908.05659.
Shapiro, A. (2017). Distributionally robust stochastic programming. *SIAM Journal on Optimization, 27*(4), 2258–2275.
Shapiro, A. (2021). Tutorial on risk neutral, distributionally robust and risk averse multistage stochastic programming. *European Journal of Operational Research, 288*, 1–13.
Shehadeh, K. S., & Padman, R. (2021). A distributionally robust optimization approach for stochastic elective surgery scheduling with limited intensive care unit capacity. *European Journal of Operational Research, 290*(3), 901–913.
Toubeau, J., Grève, Z. D., & Vallée, F. (2018). Medium-term multimarket optimization for virtual power plants: A stochastic-based decision environment. *IEEE Transactions on Power Systems, 33*(2), 1399–1410.
Wachter, A., & Biegler, L. (2006). On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Mathematical Programming, 106*(1), 25–57.
Wiesemann, W., Kuhn, D., & Sim, M. (2014). Distributionally robust convex optimization. *Operations Research, 62*(6), 1358–1376.
Xie, W. (2021). On distributionally robust chance constrained program with Wasserstein distance. *Mathematical Programming, 186*, 115–155.
Xin, L., & Goldberg, D. A. (2021). Time (in)consistency of multistage distributionally robust inventory models with moment constraints. *European Journal of Operational Research, 289*(3), 1127–1141.
Zhang, Y., Shen, S., & Mathieu, J. L. (2017). Distributionally robust chance-constrained optimal power flow with uncertain renewables and uncertain reserves provided by loads. *IEEE Transactions on Power Systems, 32*(2), 1378–1388.
Zhao, C., & Jiang, R. (2018). Distributionally robust contingency-constrained unit commitment. *IEEE Transactions on Power Systems, 33*(1), 94–102.
Zugno, M., & Conejo, A. J. (2015). A robust optimization approach to energy and reserve dispatch in electricity markets. *European Journal of Operational Research, 247*, 659–671.
Zymler, S., Kuhn, D., & Rustem, B. (2013). Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming, 137*(1-2), 167–198.