Online Appendix of "Robust Generator Maintenance Schedule for Frequency-Secure Power Systems"

Yang Yang, Quizhuang Sun, Jimmy Chih-Hsien Peng, Loon Ching Tang, and Zhisheng Ye

EC.1. The Frequency Regulation Mechanism

In modern power systems, electricity is transmitted in the form of alternating current, i.e., the direction of current changes periodically with a nominal frequency f^0, e.g., 50 Hz. Following a sudden power deficiency, power generation will be smaller than the demand. The rotational kinetic energy (*inertia*) stored in the rotor of generators will be spontaneously converted into power and injected into the system. As a result, the rotation of generators slows down, and the system frequency drops accordingly. To recover frequency, generators subsequently extract more steam from the turbine and increase their power generation. The increased power refers to the *frequency regulation reserve*. In addition, electric loads can provide frequency regulation reserve by reducing their instantaneous power demand until the end of frequency regulation process.

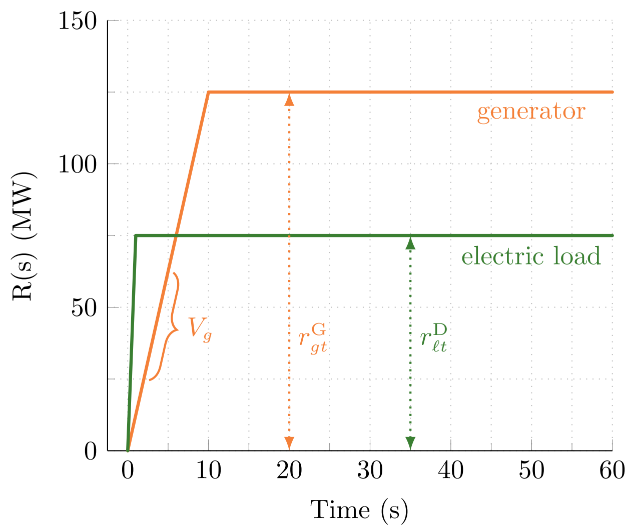


Figure EC.1: Schematic of the delivery function of the frequency regulation reserve.

During the frequency regulation process, the frequency deviation $\Delta f(s)$ at time s after the power deficiency is commonly modeled by a first-order differential equation,

$$\frac{2H_t}{f^0}\frac{d\Delta f(s)}{ds} + \eta\Delta f(s) = -\Delta P + R(s), \tag{EC.1}$$

where H_t denotes the available inertia in the system at operating hour t (H_t should be adjusted accordingly with $H_t - \bar{H}^{\mathrm{G}}$ if the loss of inertia due to generator outages is considered); η is the load damping ratio; ΔP is the power deficiency; and $R(s)$ is the total frequency regulation reserve delivered by generators and electric loads at time s, which is commonly modeled by piece-wise linear functions, as shown in Figure EC.1. With sufficient inertia and frequency regulation reserve resources, frequency will be gradually stabilized at a quasi-steady state at the end of frequency regulation process, which commonly takes 60 s. The three indices during the frequency regulation process, i.e., the initial rate of change of frequency (RoCoF), frequency nadir, and quasi-steady-state frequency can be calculated by the first-order differential equation (EC.1).

The initial RoCoF reflects how fast the frequency drops right after the occurrence of power deficiency. Let $s=0$ be the beginning time of the sudden power deficiency. Note that both $\Delta f(s)=0$ and $R(s)=0$ at time $s=0$. Plugging into the equation (EC.1), the initial RoCoF, denoted as $\texttt{RoCoF}_t$, can therefore be calculated as

$$\texttt{RoCoF}_t = d\Delta f(s)/ds\,|_{s\mapsto 0} = -f^0\Delta P/(2H_t).$$

Moreover, to avoid under-frequency load-shedding problems, the absolute frequency deviation at the nadir should be smaller than a maximum admissible value $\Delta\bar{f}^{\text{nadir}}$. The sufficient condition is that the frequency regulation reserve of generators should be fully delivered before the arrival of the frequency nadir (Wen et al. 2016). Frequency nadir time t_{nadir} can be estimated by $-\Delta\bar{f}^{\text{nadir}}/\texttt{RoCoF}_t = 2H_t\Delta\bar{f}^{\text{nadir}}/(f^0\Delta P)$. Hence, given the scheduled frequency regulation reserve r^{G}_{gt} of generator g and the ramp rate of generator V_g, the sufficient condition is equivalent to

$$r^{\text{G}}_{gt} \le -V_g\Delta\bar{f}^{\text{nadir}}/\texttt{RoCoF}_t = 2H_tV_g\Delta\bar{f}^{\text{nadir}}/(f^0\Delta P).$$

On the other hand, electric loads can provide frequency regulation reserve almost immediately by curtailing some load demand (De et al. 2020). Therefore, we assume that their reserve can always be fully delivered before the arrival of the frequency nadir. Hence, the frequency nadir condition is not imposed on the electric loads.

At the end of frequency regulation process, i.e., $s=60$, frequency is stabilized at the quasi-steady state and hence $d\Delta f(s)/ds|_{s=60}=0$. Moreover, the frequency regulation reserve from online generators and electric loads are fully delivered. Let $r^{\text{D}}_{\ell t}$ be the frequency regulation reserve of electric load ℓ at the operating hour t, and the overall delivered frequency regulation reserve $R(s=60)=\sum_{g\in[G]} r^{\text{G}}_{gt} + \sum_{\ell\in[L]} r^{\text{D}}_{\ell t}$. Plugging into equation (EC.1), the absolute frequency deviation at quasi-steady state Δf^{qss}_t can be calculated as

$$\Delta f^{\text{qss}}_t = \left(\Delta P - \sum_{g\in[G]} r^{\text{G}}_{gt} - \sum_{\ell\in[L]} r^{\text{D}}_{\ell t}\right)/\eta.$$

EC.2. Computational Intractability of Modeling Demand Uncertainty

Our test case involves a 118-bus system over a 48-week horizon, resulting in 8,064 hourly time periods. Introducing load demand uncertainty at each bus and each hour would require modeling approximately 952,752 random variables. For stochastic programming, accurately capturing this uncertainty would require a large number of scenarios, significantly increasing computational complexity in an already large-scale mixed-integer program.

Alternatively, we may consider an adaptive robust optimization model where uncertainty in the load-demand is within a budget of uncertainty set $\mathcal{D}_t \triangleq \{(D_{\ell t})_{\ell\in[L]} : D_{\ell t}\in[\underline{D}_\ell, \overline{D}_\ell],\ \forall \ell \in$

$[L],\ \sum_{\ell\in[L]}|D_{\ell t}-(\underline{D}_\ell+\overline{D}_\ell)/2|\leq B\}$. Here, $\underline{D}_\ell$ and $\overline{D}_\ell$ are the lower and upper bounds for $D_{\ell t}$, respectively, and $B\geq 0$ is the budget of uncertainty. When $B\to\infty$, the worst case is achieved at $D_{\ell t}=\overline{D}_\ell$ for all $\ell\in[L]$. For a finite budget B of uncertainty, this approach leads to the following optimization problem:

$$\begin{aligned}
\min_{\mathbf{x}}\quad & \sum_{m\in[M]}\sum_{g\in[G]}C^{\mathrm{M}}_{gm}u_{gm}+\sum_{m\in[M]}\sum_{t\in\mathcal{T}_m}\sum_{g\in[G]}C^{\mathrm{U}}_{gt}y_{gt}+\\
& \sum_{m\in[M]}\sum_{t\in\mathcal{T}_m}\max_{(D_{\ell t})_{\ell\in[L]}\in\mathcal{D}_t}\left\{\sum_{g\in[G]}F_g(p_{gt})+\max_{\Delta P\in\mathcal{U}}\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)}\mathrm{CVaR}^{\mathbb{P}}_{\varphi}\left[Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}^{\mathrm{S}}_t)\right]\right\} \qquad \text{(EC.2)}\\
\text{s.t.}\quad & \text{Generator maintenance constraints for } u_{gm}\text{: (1)–(2) in the manuscript,}\\
& \text{Binary unit commitment constraints for } y_{gt}\text{:}\ \forall t\in\mathcal{T}_m,\ m\in[M],\\
& \text{Linear operational constraints for } p_{gt}\text{:}\ \forall D_{\ell t}\in\mathcal{D}_t,\ \forall\ell\in[L],\ t\in\mathcal{T}_m,\ m\in[M],
\end{aligned}$$

where C^{M}_{gm} denotes the maintenance cost of generator $g\in[G]$ if it is under maintenance at interval m, i.e., $u_{gm}=1$, C^{U}_{gt} denotes the cost for a start-up at hour t, $F_g(\cdot)$ denotes the generation cost rate (per hour) at the power output level p_{gt}, and $Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}^{\mathrm{S}}_t)$ refers to the frequency regulation cost given the decision $\mathbf{x}$ (from both maintenance and unit commitment) and the realizations of the uncertain power deficiency ΔP and inertia $\Delta\tilde{H}^{\mathrm{S}}_t$.

Recall that the term $\max_{\Delta P\in\mathcal{U}}\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)}\mathrm{CVaR}^{\mathbb{P}}_{\varphi}\left[Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}^{\mathrm{S}}_t)\right]$ can be reformulated into a linear program based on Lemma 1. For convenience of analysis, we can rewrite Problem (EC.2) into the following form:

$$\begin{aligned}
\min_{\mathbf{x}}\quad & \mathbf{c}^\top\mathbf{x}+\sum_{m\in[M]}\sum_{t\in\mathcal{T}_m}\max_{\mathbf{d}_t\in\mathcal{D}_t}\min_{\mathbf{z}_t\in\Omega(\mathbf{x},\mathbf{d}_t)}\mathbf{b}^\top\mathbf{z}_t\\
\text{s.t.}\quad & \mathbf{F}\mathbf{x}\leq\mathbf{f},\quad \mathbf{x}\ \text{binary},
\end{aligned}$$

where $\mathbf{d}_t$ denotes the vector of $[D_{\ell t}]_{\ell\in[L]}$, $\mathbf{x}$ denotes the binary maintenance and unit commitment decisions, $\mathbf{z}_t$ denotes the continuous operational decisions of power systems at hour t, such as the power output and frequency regulation reserve of generators, and the inertia from batteries. Other symbols denote the corresponding coefficient matrix. With some effort, the feasibility set for $\mathbf{z}_t$ can be rewritten in the following form: $\Omega(\mathbf{x},\mathbf{d}_t):=\{\mathbf{z}_t:\mathbf{H}\mathbf{z}_t\leq\mathbf{h},\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{z}_t\leq\mathbf{g},\mathbf{G}\mathbf{z}_t=\mathbf{d}_t\}$. To obtain a tractable formulation, the next step would be to obtain the dual of the inner minimization problem with respect to $\mathbf{z}_t$. Eventually, the inner maximization problem can be rewritten as:

$$\begin{aligned}
\max_{\mathbf{d}_t,\boldsymbol{\varphi},\boldsymbol{\lambda},\boldsymbol{\eta}}\quad & \boldsymbol{\lambda}^\top(\mathbf{A}\mathbf{x}-\mathbf{g})-\boldsymbol{\varphi}^\top\mathbf{h}+\boldsymbol{\eta}^\top\mathbf{d}_t\\
\text{s.t.}\quad & -\boldsymbol{\lambda}^\top\mathbf{B}-\boldsymbol{\varphi}^\top\mathbf{H}+\boldsymbol{\eta}^\top\mathbf{G}=\mathbf{b}^\top,\\
& \boldsymbol{\varphi}\geq\mathbf{0},\ \boldsymbol{\lambda}\geq\mathbf{0},\ \mathbf{d}_t\in\mathcal{D}_t.
\end{aligned}$$

The above optimization problem contains a bilinear term $\boldsymbol{\eta}^\top\mathbf{d}_t$, which makes it hard to solve. To handle this term, a common approach is to use outer approximation algorithms that iteratively approximate the value of the bilinear term (Bertsimas et al. 2012). However, this will greatly aggravate the computational complexity, given that our problem is already computationally intensive.

EC.3. Joint Ambiguity Set for the Power Deficiency and Inertia

In this section, we provide the distributional robust model accounting for both the uncertain power deficiency ΔP and inertia $\Delta\tilde{H}_t^{\mathrm{S}}$, as well as the exact reformulation for the model.

Given a random sample $\{\Delta P_1, \Delta H_1^{\mathrm{S}}\}$, $\{\Delta P_2, \Delta H_2^{\mathrm{S}}\}$, ..., $\{\Delta P_N, \Delta H_N^{\mathrm{S}}\}$ for the pair of the two random variables $\Delta\tilde{H}_t^{\mathrm{S}}$ and ΔP in the $\mathbb{R}^2$ space. The empirical joint distribution $\mathbb{P}_N$ of $\Delta\tilde{H}_t^S$ and ΔP is

$$\mathbb{P}_N = \frac{1}{N}\sum_{n\in[N]} \delta_{\{\Delta P_n, \Delta H_n^S\}},$$

where $\delta_{\{\Delta P_n, \Delta H_n^S\}}$ is the Dirac function. We then construct a Wasserstein ambiguity set for $\mathbb{P}_N$:

$$\mathcal{B}_\epsilon(\mathbb{P}_N) = \left\{\mathbb{P}\in\mathcal{P}_0(\mathbb{R}^2) : \inf_{\mathbb{Q}\in\mathcal{Q}(\mathbb{P},\mathbb{P}_N)} \mathbb{E}_{\mathbb{Q}}\left[d\left(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2\right)\right] \le \epsilon\right\},$$

where $\mathcal{P}_0(\mathbb{R}^2)$ denotes the set of all probability distributions supported on $\mathbb{R}^2$, the distance metric is $d\left(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2\right) := \|\boldsymbol{\xi}_1-\boldsymbol{\xi}_2\|_1$, and $\mathcal{Q}\left(\mathbb{P}_1,\mathbb{P}_2\right)$ is the set of all joint probability distributions of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ with separate marginal distributions $\mathbb{P}_1$ and $\mathbb{P}_2$. With the above-constructed ambiguity set, we reformulate our problem as a distributional robust generator maintenance scheduling (DRGMS) model:

$$\begin{aligned}\min_{\mathbf{x}}\quad & \sum_{m\in[M]}\sum_{g\in[G]} C_{gm}^{\mathrm{M}} u_{gm} + \sum_{m\in[M]}\sum_{t\in\mathcal{T}_m}\left\{\sum_{g\in[G]}\left[C_{gt}^{\mathrm{U}} y_{gt} + F_g(p_{gt})\right]\right.\\ &\left. + \max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathrm{CVaR}_\varphi^{\mathbb{P}}\left[Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})\right]\right\} \qquad \text{(DRGMS)}\\ \text{s.t.}\quad & \text{Generator maintenance constraints: (1)–(2)},\\ & \text{Unit commitment constraints: (3)},\end{aligned}$$

where $Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})$ is slightly modified based on (9) to generate a tractable formulation. Specifically, in Equation (7), the frequency regulation reserve from generators is constrained by the ramp time $-\bar{f}_t^{\mathrm{nadir}}/\mathtt{RoCoF}_t$, i.e.,

$$r_{gt}^{\mathrm{G}} \le -V_g\Delta\bar{f}_t^{\mathrm{nadir}}/\mathtt{RoCoF}_t = 2(H_t-\bar{H}^{\mathrm{G}})V_g\Delta\bar{f}^{\mathrm{nadir}}/(f^0\Delta P), \quad \forall g\in[G],\ t\in\mathcal{T}_m,\ m\in[M].$$

Here, the nonlinear term $H_t/\Delta P$ in the above display makes it hard to derive a tractable reformulation for the distributionally robust model. To circumvent this difficulty, we use $\underline{\mathtt{RoCoF}}$ to replace $\mathtt{RoCoF}_t$, which reduces the upper bound of r_{gt}^{G} and therefore induces a more robust model. Moreover, without loss of generality, we assume the L-th load is the most expensive one for providing frequency regulation reserve. In the operation practice, the magnitude of power deficiency is always smaller than the overall amount of load demand. Hence, the frequency regulation reserve of the L-th load will be the last one to be dispatched and the dispatched amount is always smaller than its upper bound $D_{\ell t}$. Consequently, we remove its upper bound and only consider the upper bounds of other loads $\ell\in[L']=\{1,2,\ldots,L-1\}$. Following the same approach in Appendix EC.4, we obtain the equivalent reformulation for $\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)}\mathrm{CVaR}_\varphi^{\mathbb{P}}\left[Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})\right]$:

LEMMA EC.1. *The robust counterpart* $\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathrm{CVaR}_\varphi^{\mathbb{P}}[Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})]$ *is equivalent to the following linear program:*

$$\min_{\substack{\sigma_t,\tau_{nt},H_{nt}^{\mathrm{B}},\\ r_{ngt}^{\mathrm{G}},r_{n\ell t}^{\mathrm{D}}}} \quad \sigma_t+\frac{1}{\varphi}\left(\left(-\frac{f^0}{2\underline{\mathtt{RoCoF}}}C_t^{\mathrm{B}}+C_{Lt}^{\mathrm{D}}\right)\epsilon+\frac{1}{N}\sum_{n\in[N]}\tau_{nt}\right) \tag{EC.3a}$$

$$\text{s.t.} \quad \tau_{nt}\in\mathbb{R}_+,\ H_{nt}^{\mathrm{B}}\in\mathbb{R}_+, r_{n\ell t}^{\mathrm{D}}\in\mathbb{R}^+, r_{ngt}^{\mathrm{G}}\in\mathbb{R}^+, \quad \forall n\in[N], g\in[G],\ell\in[L], \tag{EC.3b}$$

$$H_{nt}=\sum_{g\in[G]} IT_g\overline{P}_g x_{gt}+\widehat{H}_t^{\mathrm{S}}+\Delta H_n^{\mathrm{S}}+H_t^{\mathrm{C}}+H_{nt}^{\mathrm{B}}, \quad \forall n\in[N], \tag{EC.3c}$$

$$-f^0\Delta P_n/(2\underline{\mathtt{RoCoF}})+\bar{H}^{\mathrm{G}}\le H_{nt}, \quad \forall n\in[N], \tag{EC.3d}$$

$$C_t^{\mathrm{B}}H_{nt}^{\mathrm{B}}+\sum_{g\in[G]}C_{gt}^{\mathrm{G}}r_{ngt}^{\mathrm{G}}+\sum_{\ell\in[L]}C_{\ell t}^{\mathrm{D}}r_{n\ell t}^{\mathrm{D}}\le\tau_{nt}+\sigma_t, \quad \forall n\in[N], \tag{EC.3e}$$

$$r_{ngt}^{\mathrm{G}}\le\overline{P}_g x_{gt}-p_{gt}, \quad \forall n\in[N],\ g\in[G], \tag{EC.3f}$$

$$r_{ngt}^{\mathrm{G}}\le V_g\Delta\bar{f}^{\mathrm{nadir}}/\underline{\mathtt{RoCoF}}, \quad \forall n\in[N],\ g\in[G], \tag{EC.3g}$$

$$r_{n\ell t}^{\mathrm{D}}\le D_{\ell t}, \quad \forall n\in[N],\ \ell\in[L'], \tag{EC.3h}$$

$$\Delta P_n-\eta\Delta\bar{f}^{\mathrm{qss}}\le\sum_{g\in[G]}r_{ngt}^{\mathrm{G}}+\sum_{\ell\in[L]}r_{n\ell t}^{\mathrm{D}}, \quad \forall n\in[N]. \tag{EC.3i}$$

Proof. For ease of exposition, we first rewrite $Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})$ as

$$Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})=\min_{\mathbf{y}_t} \quad \mathbf{q}_t^\top\mathbf{y}_t \tag{EC.4a}$$

$$\text{s.t.} \quad \mathbf{T}[\Delta P,\Delta\tilde{H}_t^{\mathrm{S}}]^\top+\mathbf{h}_t(\mathbf{x})\le\mathbf{W}\mathbf{y}_t, \tag{EC.4b}$$

where $\mathbf{y}_t=[H_t^{\mathrm{B}},r_{gt}^{\mathrm{G}},r_{\ell t}^{\mathrm{D}}]^\top_{g\in[G],\ell\in[L]}$, $\mathbf{q}_t=[C_t^{\mathrm{B}},C_{gt}^{\mathrm{G}},C_{\ell t}^{\mathrm{D}}]^\top_{g\in[G],\ell\in[L]}$, and the detailed expressions for $\mathbf{T}$, $\mathbf{h}_t(\mathbf{x})$, and $\mathbf{W}$ are provided below:

$$\mathbf{T}=\begin{bmatrix}-\frac{f^0}{2\underline{\mathtt{RoCoF}}} & -1\\ \mathbf{0}_{G\times1} & \mathbf{0}_{G\times1}\\ \mathbf{0}_{G\times1} & \mathbf{0}_{G\times1}\\ \mathbf{0}_{L'\times1} & \mathbf{0}_{L'\times1}\\ 0 & 0\\ \mathbf{0}_{L'\times1} & \mathbf{0}_{L'\times1}\\ 1 & 0\end{bmatrix}\quad \mathbf{h}_t(\mathbf{x})=\begin{bmatrix}\underline{H}^{\mathrm{B}}\\ \mathbf{0}_{G\times1}\\ -\bar{\mathbf{r}}_t\\ \mathbf{0}_{L'\times1}\\ 0\\ -\mathbf{D}_t'\\ -\eta\Delta\bar{f}_{\mathrm{qss}}\end{bmatrix}\quad \mathbf{W}=\begin{bmatrix}1 & \mathbf{0}_{1\times G} & \mathbf{0}_{1\times L'} & 0\\ \mathbf{0}_{G\times1} & \mathbf{0}_{G\times G} & \mathbf{0}_{G\times L'} & 0\\ \mathbf{0}_{G\times1} & -\boldsymbol{e}_{G\times G} & \mathbf{0}_{G\times L'} & 0\\ \mathbf{0}_{L'\times1} & \mathbf{0}_{L'\times G} & \boldsymbol{e}_{L'\times L'} & 0\\ 0 & 0 & 0 & 1\\ \mathbf{0}_{L'\times1} & \mathbf{0}_{L'\times G} & -\boldsymbol{e}_{L'\times L'} & \mathbf{0}_{L'\times1}\\ 0 & \mathbf{1}_{1\times G} & \mathbf{1}_{1\times L'} & 1\end{bmatrix},$$

where $\underline{H}^{\mathrm{B}}=-\sum_{g\in[G]}\mathtt{IT}_g\overline{P}_g x_{gt}-\widehat{H}_t^{\mathrm{S}}-H_t^{\mathrm{C}}+\bar{H}^{\mathrm{G}}$, $L'=L-1$, and $\mathbf{D}_t'=[D_{1t},D_{2t},\ldots,D_{L't}]$. Moreover, $\mathbf{0}$ denotes a matrix with all elements as 0, $\mathbf{1}$ denotes a matrix with all elements as 1, and $\boldsymbol{e}$ denotes an identity matrix.

For any fixed $\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)$, the CVaR of frequency regulation cost $\mathrm{CVaR}_\varphi^{\mathbb{P}}[Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})]$ is equivalent to $\inf_{\sigma_t\in\mathbb{R}}\{\sigma_t+\frac{1}{\varphi}\mathbb{E}_{\mathbb{P}}[\max\{Z_t(\mathbf{x},\Delta P,\Delta\tilde{H}_t^{\mathrm{S}})-\sigma_t,0\}]\}$ (Wiesemann et al. 2014). By further introducing an auxiliary variable $\tau_t \in \mathbb{R}_+$ and requiring $\tau_t \geq Z_t(\mathbf{x}, \Delta P, \Delta \tilde{H}_t^{\mathrm{S}}) - \sigma_t$, we obtain the following epigraph representation of the robust counterpart $\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathrm{CVaR}_\varphi^{\mathbb{P}}[Z_t(\mathbf{x}, \Delta P, \Delta \tilde{H}_t^{\mathrm{S}})]$:

$$\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \inf_{\sigma_t} \left\{ \sigma_t + \frac{1}{\varphi}\mathbb{E}_{\mathbb{P}} \left[\min_{\tau_t, \mathbf{y}_t} \left\{ \tau_t \mid \tau_t \in \mathbb{R}_+, \mathbf{q}_t^\top \mathbf{y}_t - \sigma_t \leq \tau_t, \mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_t \right\} \right] \right\}.$$

Next, we can switch the order of the supremum over $\mathbb{P}$ and infimum over σ_t by applying the minimax theorem (Sion 1958). This yields

$$\inf_{\sigma_t} \left\{ \sigma_t + \frac{1}{\varphi} \max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathbb{E}_{\mathbb{P}} \left[\min_{\tau_t, \mathbf{y}_t} \left\{ \tau_t \mid \tau_t \in \mathbb{R}_+, \mathbf{q}_t^\top \mathbf{y}_t - \sigma_t \leq \tau_t, \mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_t \right\} \right] \right\}.$$

In the following, we focus on deriving the equivalent model for the inner maximization problem while fixing the variable σ_t:

$$\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathbb{E}_{\mathbb{P}} \left[\min_{\tau_t, \mathbf{y}_t} \left\{ \tau_t \mid \tau_t \in \mathbb{R}_+, \mathbf{q}_t^\top \mathbf{y}_t - \sigma_t \leq \tau_t, \mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_t \right\} \right]. \tag{EC.5}$$

PROPOSITION EC.1. *Problem (EC.5) admits the following equivalent formulation:*

$$\inf_{\lambda\in\mathbb{R}_+} \epsilon\lambda + \frac{1}{N}\sum_{n\in[N]} \sup_{\Delta P,\ \Delta \tilde{H}_t^{\mathrm{S}}} \left[\Phi(\mathbf{x}, \sigma_t, \Delta P, \Delta \tilde{H}_t^{\mathrm{S}}) - \lambda|\Delta \tilde{H}_t^{\mathrm{S}} - \Delta H_n^{\mathrm{S}}| - \lambda|\Delta P - \Delta P_n| \right], \tag{EC.6}$$

where $\Phi(\mathbf{x}, \sigma_t, \Delta P, \Delta \tilde{H}_t^{\mathrm{S}}) = \min_{\tau_t, \mathbf{y}_t}\{\tau_t \mid \tau_t \in \mathbb{R}_+, \mathbf{q}_t^\top \mathbf{y}_t - \sigma_t \leq \tau_t, \mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_t\}$.

By dualizing the linear program $\Phi(\mathbf{x}, \sigma_t, \Delta P, \Delta \tilde{H}_t^{\mathrm{S}})$, we have

$$\begin{aligned} \Phi(\mathbf{x}, \sigma_t, \Delta P, \Delta \tilde{H}_t^{\mathrm{S}}) = \max_{d_1, d_2, \mathbf{d}_3} \quad & -\sigma_t d_2 + (\mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}))^\top \mathbf{d}_3, \\ \text{s.t.} \quad & d_1 \in \mathbb{R}_+,\ d_2 \in \mathbb{R}_+,\ \mathbf{d}_3 \in \mathbb{R}_+^{G+L+1}, \\ & d_1 + d_2 = 1,\ \mathbf{W}^\top \mathbf{d}_3 - d_2 \mathbf{q}_t = \mathbf{0}. \end{aligned} \tag{EC.7}$$

Let $\mathcal{D} = \{d_1 \in \mathbb{R}_+, d_2 \in \mathbb{R}_+, \mathbf{d}_3 \in \mathbb{R}_+^{G+L+1}, d_1 + d_2 = 1, \mathbf{W}^\top \mathbf{d}_3 - d_2\mathbf{q}_t = \mathbf{0}\}$ be the feasible space of the dual variables d_1, d_2, and $\mathbf{d}_3$. Problem (EC.6) is equivalent to

$$\begin{aligned} & \inf_{\lambda\in\mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N}\sum_{n\in[N]} \sup_{\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}} \sup_{d_2, \mathbf{d}_3\in\mathcal{D}} -\sigma_t d_2 + \left(\mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x})\right)^\top \mathbf{d}_3 - \lambda|\Delta P - \Delta P_n| - \lambda|\Delta \tilde{H}_t^{\mathrm{S}} - \Delta H_n^{\mathrm{S}}| \right\} \\ = & \inf_{\lambda\in\mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N}\sum_{n\in[N]} \sup_{d_2, \mathbf{d}_3\in\mathcal{D}} \sup_{\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}} \inf_{\substack{|\gamma_1|\leq\lambda \\ |\gamma_2|\leq\lambda}} \left(-\sigma_t d_2 + (\mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}))^\top \mathbf{d}_3 - \gamma_1\Delta P + \gamma_1 \Delta P_n - \gamma_2 \Delta \tilde{H}_t^{\mathrm{S}} + \gamma_2 \Delta H_n^{\mathrm{S}} \right) \right\} \\ = & \inf_{\lambda\in\mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N}\sum_{n\in[N]} \sup_{d_2, \mathbf{d}_3\in\mathcal{D}} \inf_{\substack{|\gamma_1|\leq\lambda \\ |\gamma_2|\leq\lambda}} \sup_{\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}} \left(-\sigma_t d_2 + (\mathbf{T}[\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}))^\top \mathbf{d}_3 - \gamma_1\Delta P + \gamma_1 \Delta P_n - \gamma_2 \Delta \tilde{H}_t^{\mathrm{S}} + \gamma_2 \Delta H_n^{\mathrm{S}} \right) \right\} \\ = & \inf_{\lambda\in\mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N}\sum_{n\in[N]} \sup_{d_2, \mathbf{d}_3\in\mathcal{D}} \inf_{\substack{|\gamma_1|\leq\lambda \\ |\gamma_2|\leq\lambda}} -\sigma_t d_2 + \mathbf{h}_t(\mathbf{x})^\top \mathbf{d}_3 + \gamma_1 \Delta H_n^{\mathrm{S}} + \gamma_2 \Delta P_n + \sup_{\Delta \tilde{H}_t^{\mathrm{S}}, \Delta P} \left[\Delta P(\mathbf{T}_{:1}^\top \mathbf{d}_3 - \gamma_1) + \Delta \tilde{H}_t^{\mathrm{S}}(\mathbf{T}_{:2}^\top \mathbf{d}_3 - \gamma_2) \right] \right\} \end{aligned}$$

where the first equality is achieved by switching the supremum over $d_2, \mathbf{d}_3$ and the supremum over $\Delta P, \Delta \tilde{H}_t^{\mathrm{S}}$ and the definition of the dual norm. The second equality is achieved by further switching the supremum and infimum based on the minimax theorem (Sion 1958). Observe that the conditions $\gamma_1 = \mathbf{T}_{:1}^\top \mathbf{d}_3$ and $\gamma_2 = \mathbf{T}_{:2}^\top \mathbf{d}_3$ must hold for all $\mathbf{d}_3 \in \mathcal{D}$ to obtain a bounded value for (EC.6). We thus obtain an equivalent problem for (EC.6):

$$\begin{aligned} \inf_{\lambda} \quad & \left\{ \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \sup_{d_2, \mathbf{d}_3 \in \mathcal{D}} \left[-\sigma_t d_2 + \mathbf{h}_t(\mathbf{x})^\top \mathbf{d}_3 + \mathbf{T}_{:1}^\top \mathbf{d}_3 \Delta P_n + \mathbf{T}_{:2}^\top \mathbf{d}_3 \Delta H_n^{\mathrm{S}} \right] \right\} \\ \text{s.t.} \quad & \lambda \in \mathbb{R}_+, \\ & |\mathbf{T}_{:1}^\top \mathbf{d}_3| \le \lambda, |\mathbf{T}_{:2}^\top \mathbf{d}_3| \le \lambda \quad \forall \mathbf{d}_3 \in \mathcal{D}. \end{aligned} \tag{EC.8}$$

Notice that the constraints $|\mathbf{T}_{:1}^\top \mathbf{d}_3| \le \lambda, |\mathbf{T}_{:1}^\top \mathbf{d}_3| \le \lambda, \forall \mathbf{d}_3 \in \mathcal{D}$ is equivalent to the following four linear programs: $\sup_{\mathbf{d}_3\in\mathcal{D}} \mathbf{T}_{:1}^\top \mathbf{d}_3 \le \lambda$, $\sup_{\mathbf{d}_3\in\mathcal{D}} -\mathbf{T}_{:1}^\top \mathbf{d}_3 \le \lambda$, $\sup_{\mathbf{d}_3\in\mathcal{D}} \mathbf{T}_{:2}^\top \mathbf{d}_3 \le \lambda$, and $\sup_{\mathbf{d}_3\in\mathcal{D}} -\mathbf{T}_{:2}^\top \mathbf{d}_3 \le \lambda$. By dualizing the four linear programs over $\mathbf{d}_3$, we obtain the following constraints:

$$\begin{aligned} &\mathbf{q}_t^\top \boldsymbol{\phi}_1 \le \lambda,\ \mathbf{q}^\top \boldsymbol{\psi}_1 \le \lambda,\ \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}_1,\ -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}_1,\ \boldsymbol{\phi}_1 \in \mathbb{R}^{G+L+1},\ \boldsymbol{\psi}_1 \in \mathbb{R}^{G+L+1} \\ &\mathbf{q}_t^\top \boldsymbol{\phi}_2 \le \lambda,\ \mathbf{q}^\top \boldsymbol{\psi}_2 \le \lambda,\ \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}_2,\ -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}_2,\ \boldsymbol{\phi}_2 \in \mathbb{R}^{G+L+1},\ \boldsymbol{\psi}_2 \in \mathbb{R}^{G+L+1}. \end{aligned}$$

Note that the inner optimization problem in the objective function, $\sup_{d_1,d_2,\mathbf{d}_3\in\mathcal{D}}[-\sigma_t d_2 + \mathbf{h}_t(\mathbf{x})^\top \mathbf{d}_3 + \mathbf{T}_{:1}^\top \mathbf{d}_3 \Delta P_n + \mathbf{T}_{:2}^\top \mathbf{d}_3 \Delta H_n^{\mathrm{S}}]$, is equivalent to $\Phi(\mathbf{x}, \sigma_t, \Delta P_n, \Delta H_n^{\mathrm{S}}) = \min_{\tau_{nt}, \mathbf{y}_{nt}} \{\tau_{nt} \mid \tau_{nt} \in \mathbb{R}_+, \mathbf{q}_t^\top \mathbf{y}_{nt} - \sigma_t \le \tau_{nt}, \mathbf{T}[\Delta P_n, \Delta H_n^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_{nt}\}$ by strong duality, where $\mathbf{y}_{nt} = [H_{nt}^{\mathrm{B}}, r_{ngt}^{\mathrm{G}}, r_{n\ell t}^{\mathrm{D}}]_{g\in[G],\ell\in[L]}^\top$ is defined analogously to $\mathbf{y}_t$. Hence, Problem (EC.8) is equivalent to

$$\begin{aligned} \inf_{\lambda, \tau_{nt}, \mathbf{y}_{nt}, \boldsymbol{\phi}, \boldsymbol{\psi}} \quad & \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \tau_{nt} \\ \text{s.t.} \quad & \lambda \in \mathbb{R}_+,\ \boldsymbol{\phi}_1 \in \mathbb{R}^{G+L+1},\ \boldsymbol{\psi}_1 \in \mathbb{R}^{G+L+1},\ \boldsymbol{\phi}_2 \in \mathbb{R}^{G+L+1},\ \boldsymbol{\psi}_2 \in \mathbb{R}^{G+L+1}, \\ & \tau_{nt} \in \mathbb{R}_+,\ \mathbf{y}_{nt} \in \mathbb{R}^{G+L+3}, \quad \forall n \in [N], \\ & \mathbf{q}_t^\top \mathbf{y}_{nt} - \sigma_t \le \tau_{nt},\ \mathbf{T}[\Delta P_n, \Delta H_n^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_{nt}, \quad \forall n \in [N], \\ & \mathbf{q}_t^\top \boldsymbol{\phi}_1 \le \lambda,\ \mathbf{q}_t^\top \boldsymbol{\psi}_1 \le \lambda,\ \mathbf{q}_t^\top \boldsymbol{\phi}_2 \le \lambda,\ \mathbf{q}_t^\top \boldsymbol{\psi}_2 \le \lambda, \\ & \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}_1,\ -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}_1,\ \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}_2,\ -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}_2. \end{aligned} \tag{EC.9}$$

Plugging into the detailed expressions of $\mathbf{q}_t$, $\mathbf{T}$, and $\mathbf{W}$, it is easy to see that the constraint set $\{\boldsymbol{\phi}_1 \in \mathbb{R}^{G+L+1}, \boldsymbol{\psi}_1 \in \mathbb{R}^{G+L+1}, \boldsymbol{\phi}_2 \in \mathbb{R}^{G+L+1}, \boldsymbol{\psi}_2 \in \mathbb{R}^{G+L+1}, \mathbf{q}_t^\top \boldsymbol{\phi}_1 \le \lambda, \mathbf{q}_t^\top \boldsymbol{\psi}_1 \le \lambda, \mathbf{q}_t^\top \boldsymbol{\phi}_2 \le \lambda, \mathbf{q}_t^\top \boldsymbol{\psi}_2 \le \lambda, \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}_1, -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}_1, \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}_2, -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}_2\}$ is equivalent to $\lambda \ge -\frac{-2f^0 C_t^{\mathrm{B}}}{\underline{\mathtt{RoCoF}}} + C_{Lt}^{\mathrm{D}}$. Since we minimize λ in Problem (EC.9), the optimal solution is achieved at $\lambda = \frac{-2f^0 C_t^{\mathrm{B}}}{\underline{\mathtt{RoCoF}}} + C_{Lt}^{\mathrm{D}}$. Hence, the robust counterpart $\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathrm{CVaR}_\varphi^{\mathbb{P}}[Z_t(\mathbf{x}, \Delta P, \Delta\tilde{H}_t^{\mathrm{S}})]$ is equivalent to

$$\begin{aligned} \min_{\sigma_t, \tau_{nt}, \mathbf{y}_{nt}} \quad & \sigma_t + \frac{1}{\varphi} \left(\left(\frac{-2f^0 C_t^{\mathrm{B}}}{\underline{\mathtt{RoCoF}}} + C_{Lt}^{\mathrm{D}} \right) \epsilon + \frac{1}{N} \sum_{n\in[N]} \tau_{nt} \right) \\ \text{s.t.} \quad & \sigma_t \in \mathbb{R}, \\ & \tau_{nt} \in \mathbb{R}_+,\ \mathbf{y}_{nt} \in \mathbb{R}^{G+L+1}, \quad \forall n \in [N], \\ & \mathbf{q}_t^\top \mathbf{y}_{nt} - \sigma_t \le \tau_{nt},\ \mathbf{T}[\Delta P_n, \Delta H_n^{\mathrm{S}}]^\top + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_{nt}, \quad \forall n \in [N]. \end{aligned}$$

Plugging detailed expressions of $\mathbf{q}_t, \mathbf{T}, \mathbf{h}_t(\mathbf{x}), \mathbf{W}$ into the above program finishes the proof. □

EC.4. Proof of Lemma 1

For the ease of exposition, we first rewrite $Z_t(\mathbf{x}, \Delta P, \Delta\tilde{H}_t^{\mathrm{S}})$ as

$$Z_t(\mathbf{x}, \Delta P, \Delta\tilde{H}_t^{\mathrm{S}}) = \min_{\mathbf{y}_t} \quad \mathbf{q}_t^\top \mathbf{y}_t \tag{EC.10a}$$

$$\text{s.t.} \quad \mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_t, \tag{EC.10b}$$

where Constraint (EC.10b) subsumes frequency security constraints (4)–(8) under the realization of power deficiency ΔP and inertia prediction error $\Delta\tilde{H}_t^{\mathrm{S}}$, and detailed expressions for $\mathbf{y}_t$, $\mathbf{q}_t$, $\mathbf{T}$, $\mathbf{h}_t(\mathbf{x})$, and $\mathbf{W}$ are provided in Appendix EC.5.

For any fixed $\mathbb{P} \in \mathcal{B}_\epsilon(\mathbb{P}_N)$ and realization of $\Delta P \in \mathcal{U}$, the CVaR of frequency regulation cost $\mathrm{CVaR}_\varphi^{\mathbb{P}}[Z_t(\mathbf{x}, \Delta P, \Delta\tilde{H}_t^{\mathrm{S}})]$ is equivalent to $\inf_{\sigma_t \in \mathbb{R}}\{\sigma_t + \frac{1}{\varphi}\mathbb{E}_{\mathbb{P}}[\max\{Z_t(\mathbf{x}, \Delta P, \Delta\tilde{H}_t^{\mathrm{S}}) - \sigma_t, 0\}]\}$ (Wiesemann et al. 2014). By further introducing an auxiliary variable $\tau_t \in \mathbb{R}_+$ and requiring $\tau_t \ge Z_t(\mathbf{x}, \Delta P, \Delta\tilde{H}_t^{\mathrm{S}}) - \sigma_t$, we can obtain the following epigraph representation of the robust counterpart $\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathrm{CVaR}_\varphi^{\mathbb{P}}[Z_t(\mathbf{x}, \Delta P, \Delta\tilde{H}_t^{\mathrm{S}})]$:

$$\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \inf_{\sigma_t} \left\{\sigma_t + \frac{1}{\varphi}\mathbb{E}_{\mathbb{P}}\left[\min_{\tau_t, \mathbf{y}_t}\left\{\tau_t \mid \tau_t \in \mathbb{R}_+,\ \mathbf{q}_t^\top\mathbf{y}_t - \sigma_t \le \tau_t,\ \mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_t\right\}\right]\right\}. \tag{EC.11}$$

Next, we can switch the order of the supremum over $\mathbb{P}$ and infimum over σ_t in problem (EC.11) by applying the minimax theorem (Sion 1958). This yields

$$\inf_{\sigma_t} \left\{\sigma_t + \frac{1}{\varphi}\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathbb{E}_{\mathbb{P}}\left[\min_{\tau_t, \mathbf{y}_t}\left\{\tau_t \mid \tau_t \in \mathbb{R}_+,\ \mathbf{q}_t^\top\mathbf{y}_t - \sigma_t \le \tau_t,\ \mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_t\right\}\right]\right\}.$$

In the following, we focus on deriving the equivalent model for the inner maximization problem while fixing the variable σ_t:

$$\max_{\mathbb{P}\in\mathcal{B}_\epsilon(\mathbb{P}_N)} \mathbb{E}_{\mathbb{P}}\left[\min_{\tau_t, \mathbf{y}_t}\left\{\tau_t \mid \tau_t \in \mathbb{R}_+,\ \mathbf{q}_t^\top\mathbf{y}_t - \sigma_t \le \tau_t,\ \mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_t\right\}\right]. \tag{EC.12}$$

PROPOSITION EC.2. *The problem (EC.12) admits the following equivalent problem:*

$$\inf_{\lambda\in\mathbb{R}_+} \epsilon\lambda + \frac{1}{N}\sum_{n\in[N]} \sup_{\Delta\tilde{H}_t^{\mathrm{S}}} \left[\Phi(\mathbf{x}, \sigma_t, \Delta\tilde{H}_t^{\mathrm{S}}) - \lambda|\Delta\tilde{H}_t^{\mathrm{S}} - \Delta H_n^{\mathrm{S}}|\right], \tag{EC.13}$$

where $\Phi(\mathbf{x}, \sigma_t, \Delta\tilde{H}_t^{\mathrm{S}}) = \min_{\tau_t, \mathbf{y}_t}\{\tau_t \mid \tau_t \in \mathbb{R}_+, \mathbf{q}_t^\top\mathbf{y}_t - \sigma_t \le \tau_t, \mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \le \mathbf{W}\mathbf{y}_t\}$.

By dualizing the linear program $\Phi(\mathbf{x}, \sigma_t, \Delta\tilde{H}_t^{\mathrm{S}})$, we have

$$\begin{aligned}
\Phi(\mathbf{x}, \sigma_t, \Delta\tilde{H}_t^{\mathrm{S}}) = \max_{d_1, d_2, \mathbf{d}_3} \quad & -\sigma_t d_2 + (\mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}))^\top \mathbf{d}_3, \\
\text{s.t.} \quad & d_1 \in \mathbb{R}_+, d_2 \in \mathbb{R}_+, \mathbf{d}_3 \in \mathbb{R}_+^{G+L+3}, \\
& d_1 + d_2 = 1, \mathbf{W}^\top\mathbf{d}_3 - d_2\mathbf{q}_t = \mathbf{0}.
\end{aligned} \tag{EC.14}$$

For ease of notation, we let $\mathcal{D} = \{d_1 \in \mathbb{R}_+, d_2 \in \mathbb{R}_+, \mathbf{d}_3 \in \mathbb{R}_+^{G+L+3}, d_1 + d_2 = 1, \mathbf{W}^\top \mathbf{d}_3 - d_2 \mathbf{q}_t = \mathbf{0}\}$ be the feasibility set of the dual variables d_1, d_2 and $\mathbf{d}_3$. Hence, problem (EC.13) is equivalent to

$$\begin{aligned}
&\inf_{\lambda \in \mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \sup_{\Delta \tilde{H}_t^{\mathrm{S}}} \sup_{d_1,d_2,\mathbf{d}_3 \in \mathcal{D}} -\sigma_t d_2 + (\mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}))^\top \mathbf{d}_3 - \lambda|\Delta\tilde{H}_t^{\mathrm{S}} - \Delta H_n^{\mathrm{S}}| \right\} \\
&= \inf_{\lambda \in \mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \sup_{d_1,d_2,\mathbf{d}_3 \in \mathcal{D}} \sup_{\Delta \tilde{H}_t^{\mathrm{S}}} \inf_{|\gamma|\le\lambda} \left(-\sigma_t d_2 + (\mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}))^\top \mathbf{d}_3 - \gamma\Delta\tilde{H}_t^{\mathrm{S}} + \gamma\Delta H_n^{\mathrm{S}} \right) \right\} \\
&= \inf_{\lambda \in \mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \sup_{d_1,d_2,\mathbf{d}_3 \in \mathcal{D}} \inf_{|\gamma|\le\lambda} \sup_{\Delta \tilde{H}_t^{\mathrm{S}}} \left(-\sigma_t d_2 + (\mathbf{T}\Delta\tilde{H}_t^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}))^\top \mathbf{d}_3 - \gamma\Delta\tilde{H}_t^{\mathrm{S}} + \gamma\Delta H_n^{\mathrm{S}} \right) \right\} \\
&= \inf_{\lambda \in \mathbb{R}_+} \left\{ \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \sup_{d_1,d_2,\mathbf{d}_3 \in \mathcal{D}} \inf_{|\gamma|\le\lambda} -\sigma_t d_2 + \mathbf{h}_t(\mathbf{x})^\top \mathbf{d}_3 + \gamma\Delta H_n^{\mathrm{S}} + \sup_{\Delta \tilde{H}_t^{\mathrm{S}}} \left[\Delta\tilde{H}_t^{\mathrm{S}} (\mathbf{T}^\top \mathbf{d}_3 - \gamma) \right] \right\},
\end{aligned}$$

where the first equality is achieved by switching the supremum over d_1, d_2, $\mathbf{d}_3$ and the definition of the dual norm. The second equality is achieved by further switching the supremum over $\Delta\tilde{H}_t^{\mathrm{S}}$ and infimum over γ based on the minimax theorem (Sion 1958).

Observe that the condition $\gamma = \mathbf{T}^\top \mathbf{d}_3$ must hold for all $\mathbf{d}_3 \in \mathcal{D}$ to obtain a bounded value for (EC.13). We thus obtain an equivalent problem for (EC.13):

$$\begin{aligned}
\inf_{\lambda} \quad & \left\{ \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \sup_{d_1,d_2,\mathbf{d}_3 \in \mathcal{D}} \left[-\sigma_t d_2 + \mathbf{h}_t(\mathbf{x})^\top \mathbf{d}_3 + \mathbf{T}^\top \mathbf{d}_3 \Delta H_n^{\mathrm{S}} \right] \right\} \\
\text{s.t.} \quad & \lambda \in \mathbb{R}_+, \\
& |\mathbf{T}^\top \mathbf{d}_3| \le \lambda, \ \forall \mathbf{d}_3 \in \mathcal{D}.
\end{aligned} \tag{EC.15}$$

Notice that the constraint $|\mathbf{T}^\top \mathbf{d}_3| \le \lambda, \forall \mathbf{d}_3 \in \mathcal{D}$ is equivalent to the following two linear programs: $\sup_{\mathbf{d}_3 \in \mathcal{D}} \mathbf{T}^\top \mathbf{d}_3 \le \lambda$ and $\sup_{\mathbf{d}_3 \in \mathcal{D}} -\mathbf{T}^\top \mathbf{d}_3 \le \lambda$. By dualizing the two linear programs over $\mathbf{d}_3$, we obtain the following constraints:

$$\mathbf{q}_t^\top \boldsymbol{\phi} \le \lambda, \ \mathbf{q}^\top \boldsymbol{\psi} \le \lambda, \ \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}, \ -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}, \ \boldsymbol{\phi} \in \mathbb{R}^{G+L+3}, \ \boldsymbol{\psi} \in \mathbb{R}^{G+L+3}.$$

Plugging these constraints into problem (EC.15), we have

$$\begin{aligned}
\inf_{\lambda,\boldsymbol{\phi},\boldsymbol{\psi}} \quad & \left\{ \epsilon\lambda + \frac{1}{N} \sum_{n\in[N]} \sup_{d_1,d_2,\mathbf{d}_3 \in \mathcal{D}} \left[-\sigma_t d_2 + \mathbf{h}_t(\mathbf{x})^\top \mathbf{d}_3 + \mathbf{T}^\top \mathbf{d}_3 \Delta H_n^{\mathrm{S}} \right] \right\} \\
\text{s.t.} \quad & \lambda \in \mathbb{R}_+, \boldsymbol{\phi} \in \mathbb{R}^{G+L+3}, \ \boldsymbol{\psi} \in \mathbb{R}^{G+L+3}, \\
& \mathbf{q}_t^\top \boldsymbol{\phi} \le \lambda, \ \mathbf{q}_t^\top \boldsymbol{\psi} \le \lambda, \\
& \mathbf{T} \le \mathbf{W}\boldsymbol{\phi}, \ -\mathbf{T} \le \mathbf{W}\boldsymbol{\psi}.
\end{aligned} \tag{EC.16}$$

Note that the inner optimization problem in the objective function, i.e., $\sup_{d_1,d_2,\mathbf{d}_3 \in \mathcal{D}}[-\sigma_t d_2 + \mathbf{h}_t(\mathbf{x})^\top \mathbf{d}_3 + \mathbf{T}^\top \mathbf{d}_3 \Delta H_n^{\mathrm{S}}]$, is equivalent to $\Phi(\mathbf{x}, \sigma_t, \Delta H_n^{\mathrm{S}}) = \min_{\tau_{nt}, \mathbf{y}_{nt}} \{\tau_{nt} \mid \tau_{nt} \in \mathbb{R}_+, \mathbf{q}_t^\top \mathbf{y}_{nt} - \sigma_t \le$

$\tau_{nt}, \mathbf{T}\Delta H_n^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_{nt}\}$ by the strong duality, where $\mathbf{y}_{nt} = [H_{nt}, H_{nt}^{\mathrm{G}}, H_{nt}^{\mathrm{B}}, r_{ngt}^{\mathrm{G}}, r_{n\ell t}^{\mathrm{D}}]_{g\in[G],\ell\in[L]}^{\top}$ is analogous to $\mathbf{y}_t$. Hence, problem (EC.16) is equivalent to

$$\begin{aligned}
\inf_{\lambda,\tau_{nt},\mathbf{y}_{nt},\boldsymbol{\phi},\boldsymbol{\psi}} \quad & \epsilon\lambda + \frac{1}{N}\sum_{n\in[N]} \tau_{nt} \\
\text{s.t.} \quad & \lambda \in \mathbb{R}_+,\ \boldsymbol{\phi} \in \mathbb{R}^{G+L+3},\ \boldsymbol{\psi} \in \mathbb{R}^{G+L+3}, \\
& \tau_{nt} \in \mathbb{R}_+,\ \mathbf{y}_{nt} \in \mathbb{R}^{G+L+3}, \quad \forall n \in [N], \\
& \mathbf{q}_t^{\top}\mathbf{y}_{nt} - \sigma_t \leq \tau_{nt},\ \mathbf{T}\Delta H_n^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_{nt}, \quad \forall n \in [N], \\
& \mathbf{q}_t^{\top}\boldsymbol{\phi} \leq \lambda,\ \mathbf{q}_t^{\top}\boldsymbol{\psi} \leq \lambda,\ \mathbf{T} \leq \mathbf{W}\boldsymbol{\phi},\ -\mathbf{T} \leq \mathbf{W}\boldsymbol{\psi}.
\end{aligned} \tag{EC.17}$$

Plugging into the detailed formulation of $\mathbf{q}_t$, $\mathbf{T}$, and $\mathbf{W}$, it is easy to see that the constraint set $\{\boldsymbol{\phi} \in \mathbb{R}^{G+L+3}, \boldsymbol{\psi} \in \mathbb{R}^{G+L+3}, \mathbf{q}_t^{\top}\boldsymbol{\phi} \leq \lambda, \mathbf{q}_t^{\top}\boldsymbol{\psi} \leq \lambda, \mathbf{T} \leq \mathbf{W}\boldsymbol{\phi}, -\mathbf{T} \leq \mathbf{W}\boldsymbol{\psi}\}$ is equivalent to $\lambda \geq C_t^{\mathrm{B}}$. Since we are minimizing λ in problem (EC.17), the optimal solution is achieved at $\lambda = C_t^{\mathrm{B}}$. We then obtain the following equivalent problem for (EC.12):

$$\begin{aligned}
\inf_{\tau_{nt},\mathbf{y}_{nt}} \quad & C_t^{\mathrm{B}}\epsilon + \frac{1}{N}\sum_{n\in[N]} \tau_{nt} \\
\text{s.t.} \quad & \tau_{nt} \in \mathbb{R}_+,\ \mathbf{y}_{nt} \in \mathbb{R}^{G+L+3}, \quad \forall n \in [N], \\
& \mathbf{q}_t^{\top}\mathbf{y}_{nt} - \sigma_t \leq \tau_{nt},\ \mathbf{T}\Delta H_n^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_{nt}, \quad \forall n \in [N].
\end{aligned}$$

Hence, problem (EC.11) is equivalent to

$$\begin{aligned}
\min_{\sigma_t,\tau_{nt},\mathbf{y}_{nt}} \quad & \sigma_t + \frac{1}{\varphi}\left(C_t^{\mathrm{B}}\epsilon + \frac{1}{N}\sum_{n\in[N]} \tau_{nt}\right) \\
\text{s.t.} \quad & \sigma_t \in \mathbb{R}, \\
& \tau_{nt} \in \mathbb{R}_+,\ \mathbf{y}_{nt} \in \mathbb{R}^{G+L+3}, \quad \forall n \in [N], \\
& \mathbf{q}_t^{\top}\mathbf{y}_{nt} - \sigma_t \leq \tau_{nt},\ \mathbf{T}\Delta H_n^{\mathrm{S}} + \mathbf{h}_t(\mathbf{x}) \leq \mathbf{W}\mathbf{y}_{nt}, \quad \forall n \in [N],
\end{aligned}$$

which finishes the proof by plugging into the detailed formulation of $\mathbf{q}_t$, $\mathbf{T}$, $\mathbf{h}_t(\mathbf{x})$, and $\mathbf{W}$. □

EC.5. Expressions of the Matrices

The decision variable is $\mathbf{y}_t = [H_t, H_t^{\mathrm{G}}, H_t^{\mathrm{B}}, r_{gt}^{\mathrm{G}}, r_{\ell t}^{\mathrm{D}}]_{g\in[G],\ell\in[L]}^{\top}$. Moreover, we define the following scalars and vectors for ease of exposition: $\underline{H} = -f^0\Delta\overline{P}/(2\underline{\texttt{RoCoF}}) + \bar{H}^{\mathrm{G}}$; $\texttt{RR}_g = 2V_g\Delta\bar{f}^{\mathrm{nadir}}/(f^0\Delta\overline{P})$; $\underline{R} = \Delta\overline{P} - \eta\Delta\bar{f}^{\mathrm{qss}}$; $\mathbf{q}_t = \left[0, 0, C_t^{\mathrm{B}}, C_{gt}^{\mathrm{G}}, C_{\ell t}^{\mathrm{D}}\right]_{g\in[G],\ell\in[L]}^{\top}$; $\mathbf{D}_t = [D_{1t}, D_{2t}, \ldots, D_{Lt}]^{\top}$;

$\bar{\mathbf{r}}_t = \left[\overline{P}_1 x_{1t} - p_{1t}, \overline{P}_2 x_{2t} - p_{2t}, \ldots, \overline{P}_G x_{Gt} - p_{Gt}\right]^\top$; $\mathbf{RR} = [\mathtt{RR}_1, \mathtt{RR}_2, \ldots, \mathtt{RR}_G]^\top$. The detailed formulations of matrices $\mathbf{T}$, $\mathbf{h}_t(\mathbf{x})$, and $\mathbf{W}$ are provided below:

$$\mathbf{T} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ \mathbf{0}_{G\times 1} \\ \mathbf{0}_{G\times 1} \\ \mathbf{0}_{L\times 1} \\ \mathbf{0}_{L\times 1} \\ 0 \\ \mathbf{0}_{G\times 1} \\ 0 \end{bmatrix}, \quad \mathbf{h}_t(\mathbf{x}) = \begin{bmatrix} -\widehat{H}_t^{\mathrm{S}} - H_t^{\mathrm{C}} \\ \widehat{H}_t^{\mathrm{S}} + H_t^{\mathrm{C}} \\ \sum_{g\in[G]} \mathtt{IT}_g \overline{P}_g x_{gt} \\ -\sum_{g\in[G]} \mathtt{IT}_g \overline{P}_g x_{gt} \\ 0 \\ \mathbf{0}_{G\times 1} \\ -\bar{\mathbf{r}}_t \\ \mathbf{0}_{L\times 1} \\ -\mathbf{D}_t \\ \underline{H} \\ \bar{H}^{\mathrm{G}}\mathbf{RR} \\ \underline{R} \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} -1 & 1 & 1 & \mathbf{0}_{1\times G} & \mathbf{0}_{1\times L} \\ 1 & -1 & -1 & \mathbf{0}_{1\times G} & \mathbf{0}_{1\times L} \\ 0 & 1 & 0 & \mathbf{0}_{1\times G} & \mathbf{0}_{1\times L} \\ 0 & -1 & 0 & \mathbf{0}_{1\times G} & \mathbf{0}_{1\times L} \\ 0 & 0 & 1 & \mathbf{0}_{1\times G} & \mathbf{0}_{1\times L} \\ \mathbf{0}_{G\times 1} & \mathbf{0}_{G\times 1} & \mathbf{0}_{G\times 1} & \boldsymbol{e}_{G\times G} & \mathbf{0}_{G\times L} \\ \mathbf{0}_{G\times 1} & \mathbf{0}_{G\times 1} & \mathbf{0}_{G\times 1} & -\boldsymbol{e}_{G\times G} & \mathbf{0}_{G\times L} \\ \mathbf{0}_{L\times 1} & \mathbf{0}_{L\times 1} & \mathbf{0}_{L\times 1} & \mathbf{0}_{L\times G} & \boldsymbol{e}_{L\times L} \\ \mathbf{0}_{L\times 1} & \mathbf{0}_{L\times 1} & \mathbf{0}_{L\times 1} & \mathbf{0}_{L\times G} & -\boldsymbol{e}_{L\times L} \\ 1 & 0 & 0 & \mathbf{0}_{1\times G} & \mathbf{0}_{1\times L} \\ \mathbf{RR} & \mathbf{0}_{G\times 1} & \mathbf{0}_{G\times 1} & -\boldsymbol{e}_{G\times G} & \mathbf{0}_{1\times G} \\ 0 & 0 & 0 & \mathbf{1}_{1\times G} & \mathbf{1}_{1\times L} \end{bmatrix},$$

where we use $\mathbf{0}$ to denote a matrix with all elements as 0, $\mathbf{1}$ to denote a matrix with all elements as 1, and $\boldsymbol{e}$ to denote the identity matrix.

EC.6. Proof of Proposition EC.2

Based on the definition of the Wasserstein metric, problem (EC.12) can be reformulated as

$$\begin{aligned} \max \quad & \int_{\mathbb{R}} \Phi(\mathbf{x}, \sigma_t, \Delta\tilde{H}_t^{\mathrm{S}}) \mathbb{P}(d\Delta\tilde{H}_t^{\mathrm{S}}) \\ \text{s.t.} \quad & \mathbb{P} \in \mathcal{P}_0(\mathbb{R}), \ \mathbb{Q} \in \mathcal{P}_0(\mathbb{R}, \mathbb{R}), \\ & \int_{\mathbb{R}\times\mathbb{R}} |\Delta\tilde{H}_t^{\mathrm{S}} - \xi| \mathbb{Q}(d\Delta\tilde{H}_t^{\mathrm{S}}, d\xi) \le \epsilon, \end{aligned} \tag{EC.18}$$

where $\mathcal{P}_0(\mathbb{R})$ and $\mathcal{P}_0(\mathbb{R}, \mathbb{R})$ denote, respectively, the sets of all probability distributions supported on $\mathbb{R}$ and joint probability distributions supported on $\mathbb{R}\times\mathbb{R}$. The law of total probability indicates that $\mathbb{Q}(\Delta\tilde{H}_t^{\mathrm{S}}, \xi) = \frac{1}{N}\sum_{n\in[N]} \mathbb{P}_n$, where $\mathbb{P}_n$ denotes the marginal probability distribution of $\mathbb{Q}$ conditional on $\xi = \Delta H_n^{\mathrm{S}}$. Therefore, problem (EC.18) can be reformulated as

$$\begin{aligned} \max \quad & \frac{1}{N} \sum_{n\in[N]} \int_{\mathbb{R}} \Phi(\mathbf{x}, \sigma_t, \Delta\tilde{H}_t^{\mathrm{S}}) \mathbb{P}_n(d\Delta\tilde{H}_t^{\mathrm{S}}) \\ \text{s.t.} \quad & \mathbb{P}_n \in \mathcal{P}_0(\mathbb{R}), \\ & \frac{1}{N} \sum_{n\in[N]} \int_{\mathbb{R}} |\Delta\tilde{H}_t^{\mathrm{S}} - \Delta H_n^{\mathrm{S}}| \mathbb{P}_n(d\Delta\tilde{H}_t^{\mathrm{S}}) \le \epsilon. \end{aligned} \tag{EC.19}$$

Introducing the dual variable λ for the inequality constraint in (EC.19), its dual problem is exactly (EC.13). Moreover, the strong duality between (EC.19) and (EC.13) holds due to the linearity, which finishes the proof. □

EC.7. Proof of Theorem 1

The proof follows directly by plugging Problem (10) under the realization of $\Delta\overline{P}$ into Problem (ROGMS). $\square$

EC.8. Proof of Proposition 2

By the strong duality of linear programs, the optimality cuts can be derived from the dual problem of (SSP), given by

$$\begin{aligned}
\Gamma_{nt}(\check{x}_{gt},\check{p}_{gt},\check{H}_{nt}) = \max_{\substack{\gamma_{nt},\lambda_{ngt}\\ \kappa_{ngt},\mu_{n\ell t}}} \quad & \gamma_{nt}\underline{R} - \sum_{g\in[G]}\lambda_{ngt}(\overline{P}_g\check{x}_{gt}-\check{p}_{gt}) - \sum_{g\in[G]}\kappa_{ngt}\mathtt{RR}_g(\check{H}_{nt}-\bar{H}^{\mathrm{G}}) - \sum_{\ell\in[L]}\mu_{n\ell t}D_{\ell t}\\
\text{s.t.} \quad & \gamma_{nt}\in\mathbb{R}_+,\ \lambda_{ngt}\in\mathbb{R}_+,\ \kappa_{ngt}\in\mathbb{R}_+,\ \mu_{n\ell t}\in\mathbb{R}_+, \quad \forall g\in[G],\ \ell\in[L],\\
& \gamma_{nt}-\lambda_{ngt}-\kappa_{ngt}\le C^{\mathrm{G}}_{gt}, \quad \forall g\in[G],\\
& \gamma_{nt}-\mu_{n\ell t}\le C^{\mathrm{D}}_{\ell t}, \quad \forall \ell\in[L].
\end{aligned}$$

Given the optimal value of dual variables γ^*_{nt}, λ^*_{ngt}, κ^*_{ngt}, and $\mu^*_{n\ell t}$, we have

$$\Gamma_{nt}(\check{x}_{gt},\check{p}_{gt},\check{H}_{nt}) = \gamma^*_{nt}\underline{R} - \sum_{g\in[G]}\lambda^*_{ngt}(\overline{P}_g\check{x}_{gt}-\check{p}_{gt}) - \sum_{g\in[G]}\kappa^*_{ngt}\mathtt{RR}_g(\check{H}_{nt}-\bar{H}^{\mathrm{G}}) - \sum_{\ell\in[L]}\mu^*_{n\ell t}D_{\ell t}.$$

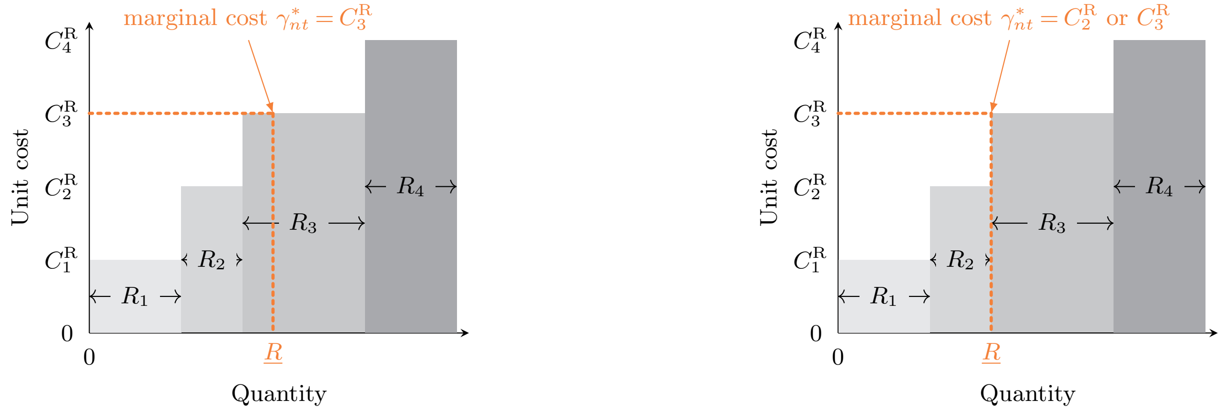
Next, based on the subgradient inequality and the convexity of the linear program, we have

$$\begin{aligned}
\Gamma_{nt}(x_{gt},p_{gt},H_{nt}) &\ge \Gamma_{nt}(\check{x}_{gt},\check{p}_{gt},\check{H}_{nt}) - \sum_{g\in[G]}\lambda^*_{ngt}(\overline{P}_g(x_{gt}-\check{x}_{gt})-(p_{gt}-\check{p}_{gt})) - \sum_{g\in[G]}\kappa^*_{ngt}\mathtt{RR}_g(H_{nt}-\check{H}_{nt})\\
&= \gamma^*_{nt}\underline{R} - \sum_{g\in[G]}\lambda^*_{ngt}(\overline{P}_g x_{gt}-p_{gt}) - \sum_{g\in[G]}\kappa^*_{ngt}\mathtt{RR}_g(H_{nt}-\bar{H}^{\mathrm{G}}) - \sum_{\ell\in[L]}\mu^*_{n\ell t}D_{\ell t},
\end{aligned}$$

which is used to generate the optimality cut. $\square$

EC.9. Procedure to obtain the optimal value of marginal cost of different frequency regulation reserves

We first sort the unit costs of frequency regulation reserve C^{G}_{gt} for all generators $g\in[G]$ and $C^{\mathrm{D}}_{\ell t}$ for all electric loads $\ell\in[L]$ in the ascending order. Without loss of generality, we assume that the unit costs of frequency regulation reserve providers differ from each other. Specifically, γ^*_{nt} is the marginal cost of the frequency regulation reserve. It reflects how much the frequency regulation reserve cost increases with a one-unit increment in the minimum quantity $\underline{R}$ of the frequency regulation reserve. Marginal cost γ^*_{nt} can be subsequently determined, as visualized in Figure EC.2a. The marginal frequency regulation reserve provider is the one whose bid overlaps with the dashed line, and marginal cost γ^*_{nt} is therefore equal to its cost. However, multiple dual optimal values may exist in some special cases, and γ^*_{nt} can have different values. For example, both the second and third reserve providers are marginal providers in Figure EC.2b, and thus lead to different cutting planes. Herein, we take both cutting planes. Frequency regulation reserve providers with unit costs greater than γ^*_{nt} will not be dispatched. In contrast, those with costs lower than γ^*_{nt} will be fully dispatched. Once marginal cost γ^*_{nt} is determined, optimal values $\lambda^*_{ngt},\kappa^*_{ngt}$ for all $g\in[G]$ and $\mu^*_{n\ell t}$ for all $\ell\in[L]$ can then be explicitly determined by Proposition 3.



(a) A general case

(b) A special case

Figure EC.2 The schematic to determine marginal cost γ^*_{nt} by sorting the unit cost of frequency regulation reserve providers, whose quantity and unit cost are respectively described by the width and height of a block. Here, each $C^{\mathrm{R}}_1, C^{\mathrm{R}}_2, \ldots$ is one of C^{G}_{gt} or $C^{\mathrm{D}}_{\ell t}$, and each $R_1, R_2, \ldots$ is one of r^{G}_{ngt} and $r^{\mathrm{D}}_{n\ell t}$.

EC.10. Proof of Proposition 3

Let $r^{\mathrm{G}*}_{ngt}$ and $r^{\mathrm{D}*}_{n\ell t}$ for all $g \in [G]$ and $\ell \in [L]$ be the optimal solution of (SSP), the Karush–Kuhn–Tucker (KKT) condition includes:

- Primal feasible constraints:

$$
\begin{aligned}
&0 \le r^{\mathrm{G}*}_{ngt}, \quad \forall g \in [G], && (\alpha^*_{ngt})\\
&0 \le r^{\mathrm{D}*}_{n\ell t}, \quad \forall \ell \in [L], && (\beta^*_{n\ell t})\\
&\underline{R} \le \sum_{g\in[G]} r^{\mathrm{G}*}_{ngt} + \sum_{\ell\in[L]} r^{\mathrm{D}*}_{n\ell t}, && (\gamma^*_{nt})\\
&r^{\mathrm{G}*}_{ngt} \le \overline{P}_g \check{x}_{gt} - \check{p}_{gt}, \quad \forall g \in [G], && (\lambda^*_{ngt})\\
&r^{\mathrm{D}*}_{n\ell t} \le D_{\ell t}, \quad \forall \ell \in [L], && (\mu^*_{n\ell t})\\
&r^{\mathrm{G}*}_{ngt} \le \mathtt{RR}_g(\check{H}_{nt} - \bar{H}^{\mathrm{G}}), \quad \forall g \in [G], && (\kappa^*_{ngt})
\end{aligned}
$$

- Dual feasible constraints:

$$
\alpha^*_{ngt} \in \mathbb{R}_+,\ \beta^*_{n\ell t} \in \mathbb{R}_+,\ \gamma^*_{nt} \in \mathbb{R}_+,\ \lambda^*_{ngt} \in \mathbb{R}_+,\ \mu^*_{n\ell t} \in \mathbb{R}_+,\ \kappa^*_{ngt} \in \mathbb{R}_+,\ \forall g \in [G],\ \ell \in [L],
$$

- Complete slackness constraints:

$$
\begin{aligned}
& -r^{\mathrm{G}*}_{ngt}\alpha^*_{ngt} = 0, && \forall g \in [G],\\
& -r^{\mathrm{D}*}_{n\ell t}\beta^*_{n\ell t} = 0, && \forall \ell \in [L],\\
& \left(\underline{R} - \sum_{g\in[G]} r^{\mathrm{G}*}_{ngt} - \sum_{\ell\in[L]} r^{\mathrm{D}*}_{n\ell t}\right)\gamma^*_{nt} = 0, && \\
& \left(r^{\mathrm{G}*}_{ngt} - \overline{P}_g\check{x}_{gt} + \check{p}_{gt}\right)\lambda^*_{ngt} = 0, && \forall g \in [G],\\
& \left(r^{\mathrm{G}*}_{ngt} - \mathtt{RR}_g(\check{H}_{nt} - \bar{H}^{\mathrm{G}})\right)\kappa^*_{ngt} = 0, && \forall g \in [G],\\
& (r^{\mathrm{D}*}_{n\ell t} - D_{\ell t})\mu^*_{n\ell t} = 0, && \forall \ell \in [L],
\end{aligned}
$$

- Optimality conditions:

$$
\begin{aligned}
& C^{\mathrm{G}}_{gt} - \alpha^*_{ngt} - \gamma^*_{nt} + \lambda^*_{ngt} + \kappa^*_{ngt} = 0, && \forall g \in [G],\\
& C^{\mathrm{D}}_{\ell t} - \beta^*_{n\ell t} - \gamma^*_{nt} + \mu^*_{n\ell t} = 0, && \forall \ell \in [L].
\end{aligned}
$$

Based on the KKT condition, we investigate the following cases.

First, for generator $g \in [G]$:

- if $\gamma^*_{nt} < C^{\mathrm{G}}_{gt}$, then it will not be dispatched because it is too expensive; i.e., $r^{\mathrm{G}*}_{ngt} = 0$. Hence, the complete slackness constraints and the optimality constraints indicate that

$$\lambda^*_{ngt} = 0,\ \kappa^*_{ngt} = 0,\ \alpha^*_{ngt} = C^{\mathrm{G}}_{gt} - \gamma^*_{nt};$$

- if $\gamma^*_{nt} = C^{\mathrm{G}}_{gt}$, then it is the marginal reserve provider; i.e., $0 < r^{\mathrm{G}*}_{ngt} < \overline{P}_g\check{x}_{gt} - \check{p}_{gt}$ and $0 < r^{\mathrm{G}*}_{ngt} < \mathtt{RR}_g(\check{H}_{nt} - \bar{H}^{\mathrm{G}})$. Hence, the complete slackness constraints indicate that

$$\alpha^*_{ngt} = 0,\ \lambda^*_{ngt} = 0,\ \kappa^*_{ngt} = 0;$$

- if $\gamma^*_{nt} > C^{\mathrm{G}}_{gt}$, then it is cheaper than the marginal reserve provider and will be fully dispatched.

—Suppose $\overline{P}_g\check{x}_{gt} - \check{p}_{gt} < \mathtt{RR}_g(\check{H}_{nt} - \bar{H}^{\mathrm{G}})$, then $r^{\mathrm{G}*}_{ngt} = \overline{P}_g\check{x}_{gt} - \check{p}_{gt}$. The complete slackness constraints and the optimality constraints indicate that

$$\alpha^*_{ngt} = 0,\ \kappa^*_{ngt} = 0,\ \lambda^*_{ngt} = \gamma^*_{nt} - C^{\mathrm{G}}_{gt};$$

—if $\mathtt{RR}_g(\check{H}_{nt} - \bar{H}^{\mathrm{G}}) \le \overline{P}_g\check{x}_{gt} - \check{p}_{gt}$, then $r^{\mathrm{G}*}_{ngt} = \mathtt{RR}_g(\check{H}_{nt} - \bar{H}^{\mathrm{G}})$. The complete slackness constraints and the optimality constraints indicate that

$$\alpha^*_{ngt} = 0,\ \lambda^*_{ngt} = 0,\ \kappa^*_{ngt} = \gamma^*_{nt} - C^{\mathrm{G}}_{gt}.$$

Next, for electric load $\ell \in [L]$:

- if $\gamma^*_{nt} < C^{\mathrm{D}}_{\ell t}$, then it will not be dispatched because it is too expensive; i.e., $r^{\mathrm{D}*}_{n\ell t} = 0$. The complete slackness constraints and the optimality constraints indicate that

$$\mu^*_{n\ell t} = 0,\ \beta^*_{n\ell t} = C^{\mathrm{D}}_{\ell t} - \gamma^*_{nt};$$

- if $\gamma_{nt}^* = C_{\ell t}^{\mathrm{D}}$, it is the marginal reserve provider; i.e., $0 < r_{n\ell t}^{\mathrm{D}*} < D_{\ell t}$. The complete slackness constraints indicate that

$$\mu_{n\ell t}^* = 0,\ \beta_{n\ell t}^* = 0;$$

- if $\gamma_{nt}^* > C_{\ell t}^{\mathrm{D}}$, then it is cheaper than the marginal reserve provider and will be fully dispatched; i.e., $r_{n\ell t}^{\mathrm{D}*} = D_{\ell t}$. The complete slackness constraints and the optimality constraints indicate that

$$\beta_{n\ell t}^* = 0,\ \mu_{n\ell t}^* = \gamma_{nt}^* - C_{\ell t}^{\mathrm{D}}.$$

The proof finishes by summarizing the above conditions. □

Online Appendices 11-21

In addition to this electronic companion, we provide a technical supplement that includes implementation details of the methods in this study and some supplementary numerical results. The technical supplement is available on the authors' website: https://drive.google.com/file/d/10nLRUQTAOjdjROJ-K3TT2f7cg7FMKvAw/view?usp=sharing

References

Birge, John R, Francois Louveaux. 2011. *Introduction to Stochastic Programming*. Springer Science & Business Media.

Dimitris Bertsimas, Eugene Litvinov, Xu Andy Sun, Jinye Zhao, and Tongxin Zheng. 2012. Adaptive robust optimization for the security constrained unit commitment problem. *IEEE Transactions on Power Systems*, 28(1), 2012.

Du, Pengwei, Nitika V Mago, Weifeng Li, Sandip Sharma, Qinran Hu, Tao Ding. 2020. New ancillary service market for ERCOT. *IEEE Access*, 8 178391-178401.

Hanasusanto, Grani A, Daniel Kuhn. 2018. Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls. *Operations Research*, 66(3), 849-869.

Owen Q. Wu, Roman Kapuscinski. 2013. Curtailing intermittent generation in electrical systems. *Manufacturing & Service Operations Management*, 15(4), 578–595.

Rahmaniani, Ragheb, Shabbir Ahmed, Teodor Gabriel Crainic, Michel Gendreau, Walter Rei. 2020. The Benders dual decomposition method. *Operations Research*, 68 (3), 878-895.

Sion, Maurice. 1958. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1), 171-176.

Douglas Thomas, Juan Fung. 2022. Measuring downstream supply chain losses due to power disturbances. *Energy Economics*, 144, 106314.

Wen, Yunfeng, Wenyuan Li, Gang Huang, Xuan Liu. 2016. Frequency dynamics constrained unit commitment with battery energy storage. *IEEE Transactions on Power Systems*, 31 (6), 5115-5125.

Wiesemann, Wolfram, Daniel Kuhn, Melvyn Sim. 2014. Distributionally robust convex optimization. *Operations Research*, 62 (6), 1358-1376.