Appendix A: Nomenclature

A. Indices

$d \in \{1,...,\mathcal{D}\}$	Index for representative days.
$i \in \{1,...,G\}$	Index for conventional generating units.
$j \in \{1,...,\mathcal{J}\}$	Index for transmission lines.
$n \in \{1,...,\mathcal{N}\}$	Index for long-term scenarios.
$t \in \{1,...,24\}$	Index for operating hours.

B. Uncertainty modeling

$\boldsymbol{\mu} \in \mathbb{R}^Z$	Mean vector of wind forecast error $\boldsymbol{\xi}$.
$\boldsymbol{\nu} \in \mathbb{R}^Z$	Mode vector of wind forecast error $\boldsymbol{\xi}$.
$\boldsymbol{\Sigma} \in \mathbb{R}^{Z\times Z}$	Covariance matrix of wind forecast error $\boldsymbol{\xi}$.
Ω^α	Set of potential α-unimodal distributions.
$\mathbb{P} \in \Pi$	Worst-case distribution of the short-term uncertainty selected from the ambiguity set Π.

C. Parameters

$\iota_n \in \mathbb{R}_+$	Probability of long-term scenario n.
$\mathbf{C} \in \mathbb{R}_+^G$	Vector of production cost of conventional units [\$/MWh].
$\mathbf{K} \in \mathbb{R}_+^{G^C}$	Vector of annualized capital cost of candidate conventional units [\$].
$\mathbf{l}_{ndt} \in \mathbb{R}_+^L$	Vector of load level under long-term scenario n in representative day d at hour t [MW].
$\mathbf{M}^D \in \mathbb{R}^{\mathcal{J}\times L}$	Matrix of power transfer distribution factor for loads.
$\mathbf{M}^G \in \mathbb{R}^{\mathcal{J}\times G}$	Matrix of power transfer distribution factor for conventional units.
$\mathbf{M}^W \in \mathbb{R}^{\mathcal{J}\times Z}$	Matrix of power transfer distribution factor for wind farms.
$\mathbf{S} \in \mathbb{R}_+^G$	Vector of start-up cost of conventional units [\$].
$\mathbf{w}_{ndt} \in \mathbb{R}_+^Z$	Mean vector of power production of wind farms under long-term scenario n in representative day d at hour t [MW].
$\omega_d \in \mathbf{N}$	Number of days in cluster of representative day d [day].
$F_j^{\max} \in \mathbb{R}_+$	Capacity of transmission line j [MW].
$P_i^{\max} \in \mathbb{R}_+$	Capacity of conventional unit i [MW].
$P_i^{\min} \in \mathbb{R}_+$	Lower bound for production level of conventional unit i [MW].
$R_i^{\text{st}} \in \mathbb{R}_+$	Start-up and shut-down ramp rate capability of unit i [MW].
$R_i^{\text{up}}, R_i^{\text{dn}} \in \mathbb{R}_+$	Ramp-up and ramp-down capabilities of conventional unit i [MW].
$U_i^{\text{up}}, U_i^{\text{dn}} \in \mathbb{R}_+$	Minimum up and down time of conventional unit i [hour].

E. Variables

$\boldsymbol{\beta}_{ndt} \in \mathbb{R}^G$	Vector of participation factor of conventional unit i under long-term scenario n in representative day d at hour t [per-unit].
$\mathbf{p}_{ndt} \in \mathbb{R}_+^G$	Vector of nominal dispatch of conventional unit i under long-term scenario n in representative day d at hour t [MW].
$\mathbf{u}_{ndt} \in \{0,1\}^G$	Vector of start-up status of conventional unit i under long-term scenario n in representative day d at hour t.
$\mathbf{x}_{ndt} \in \{0,1\}^G$	Vector of on/off commitment status of conventional unit i under long-term scenario n in representative day d at hour t.
$\mathbf{y} \in \{0,1\}^{G^C}$	Investment decision of candidate units.

Appendix B: Relaxation of Unit Commitment Integrality Constraints

In order to tightly relax the operational binary variables x_{indt} and v_{indt} to lie between zero and one, we use the convex relaxation approach proposed in [24], where by relaxing the binary variables and adding extra constraints, the feasible set of the unit commitment problem is substituted by an approximation of its convex hull [8]. By doing so, the additional constraints for tightening the relaxation are considered as probabilistic constraints. As the first alternative, they can be written in form of chance constraints as

$$\min_{\mathbb{P}\in\Pi} \mathbb{P}\{p_{ind(t-1)}(\boldsymbol{\xi}) \leqslant R_i^{\text{st}}\ x_{ind(t-1)} + (P_i^{\max} - R_i^{\text{st}})(x_{indt} - v_{indt})\} \geqslant 1-\epsilon, \forall i,n,d,t \tag{18a}$$

$$\min_{\mathbb{P}\in\Pi} \mathbb{P}\{p_{indt}(\boldsymbol{\xi}) \leqslant P_i^{\max} x_{isrt} - (P_i^{\max} - R_i^{\text{st}})v_{indt}\} \geqslant 1-\epsilon, \quad \forall i,n,d,t \tag{18b}$$

$$\min_{\mathbb{P}\in\Pi} \mathbb{P}\{p_{indt}(\boldsymbol{\xi}) - p_{ind(t-1)}(\boldsymbol{\xi}) \leqslant (P_i^{\min} + R_i^{\text{up}})x_{indt} - P_i^{\min} x_{ind(t-1)} - (P_i^{\min} + R_i^{\text{up}} - R_i^{\text{st}})v_{indt}\} \geqslant 1-\epsilon, \quad \forall i,n,d,t \tag{18c}$$

$$\min_{\mathbb{P}\in\Pi} \mathbb{P}\{p_{isr(t-1)}(\boldsymbol{\xi}) - p_{indt}(\boldsymbol{\xi}) \leqslant R_i^{\text{st}} x_{ind(t-1)} - (P_i^{\min} + R_i^{\text{dn}} - R_i^{\text{st}})v_{indt} - (R_i^{\text{st}} - R_i^{\text{dn}})x_{indt}\} \geqslant 1-\epsilon, \quad \forall i,n,d,t. \tag{18d}$$

As the second alternative, they can be in form of CVaR constraints as

$$\max_{\mathbb{P}\in\Pi} \text{CVaR}_{\mathbb{P}}^{\epsilon}[p_{ind(t-1)}(\boldsymbol{\xi})] \leqslant R_i^{\text{st}}\ x_{ind(t-1)} + (P_i^{\max} - R_i^{\text{st}})(x_{indt} - v_{indt}), \quad \forall i,n,d,t \tag{19a}$$

$$\max_{\mathbb{P}\in\Pi} \text{CVaR}_{\mathbb{P}}^{\epsilon}[p_{indt}(\boldsymbol{\xi})] \leqslant P_i^{\max} x_{isrt} - (P_i^{\max} - R_i^{\text{st}})v_{indt}, \quad \forall i,n,d,t \tag{19b}$$

$$\max_{\mathbb{P}\in\Pi} \text{CVaR}_{\mathbb{P}}^{\epsilon}[p_{indt}(\boldsymbol{\xi}) - p_{ind(t-1)}(\boldsymbol{\xi})] \leqslant (P_i^{\min} + R_i^{\text{up}})x_{indt} - P_i^{\min} x_{ind(t-1)} - (P_i^{\min} + R_i^{\text{up}} - R_i^{\text{st}})v_{indt}] \geqslant 1-\epsilon_i, \quad \forall i,n,d,t \tag{19c}$$

$$\max_{\mathbb{P}\in\Pi} \text{CVaR}_{\mathbb{P}}^{\epsilon}[p_{isr(t-1)}(\boldsymbol{\xi}) - p_{indt}(\boldsymbol{\xi})] \leqslant R_i^{\text{st}} x_{ind(t-1)} - (P_i^{\min} + R_i^{\text{dn}} - R_i^{\text{st}})v_{indt} - (R_i^{\text{st}} - R_i^{\text{dn}})x_{indt}, \quad \forall i,n,d,t. \tag{19d}$$

Appendix C: Final Model

The final tractable model for distributionally robust generation expansion planning problem (7) is the mixed-integer second order cone program presented below:

$$\begin{aligned}&\text{Objective function: (15)}\\&\text{Subject to:}\\&\quad \text{(7b)} - \text{(7g)}, \text{(14)},\\&\quad \text{Reformulated probabilistic constraints.}\end{aligned}$$

As mentioned earlier, probabilistic constraints can be enforced in the form of distributionally robust chance constraints

(8), (18) or in the form of distributionally robust CVaR constraints (9), (19). In the following, we present their resulting reformulations. Note that due to the ramping constraints of generating units, enforcing inter-temporal coupling between hours t and $t-1$, for these constraints we define new uncertainty parameter vector $\hat{\boldsymbol{\xi}}_{ndt} \in \mathbb{R}^{2Z}$ as

$$\hat{\boldsymbol{\xi}}_{ndt} = \begin{bmatrix} \boldsymbol{\xi}_{ndt} \\ \boldsymbol{\xi}_{nd(t-1)} \end{bmatrix}. \tag{20}$$

Parameters $\hat{\boldsymbol{\mu}}_{ndt} \in \mathbb{R}^{2Z}$ and $\hat{\boldsymbol{\Sigma}}_{ndt} \in \mathbb{R}^{2Z\times 2Z}$ represent the mean and covariance of $\hat{\boldsymbol{\xi}}_{ndt}$, respectively. In addition, $\hat{\boldsymbol{\nu}}_{ndt} \in \mathbb{R}^{2Z}$ refers to the mode location. Similar to the definition of $\boldsymbol{\Phi}_{ndt}$, matrix $\hat{\boldsymbol{\Phi}}_{ndt}$ is defined as

$$\hat{\boldsymbol{\Phi}}_{ndt} = \frac{\alpha+2}{\alpha}(\hat{\boldsymbol{\Sigma}}_{ndt} - \hat{\boldsymbol{\mu}}_{ndt}\hat{\boldsymbol{\mu}}_{ndt}^\top) - \frac{1}{\alpha^2}(\hat{\boldsymbol{\mu}}_{ndt} - \hat{\boldsymbol{\nu}}_{ndt})(\hat{\boldsymbol{\mu}}_{ndt} - \hat{\boldsymbol{\nu}}_{ndt})^\top. \tag{21}$$

For notational clarity, we drop indices n and d from $\{p_{indt}, x_{indt}, \beta_{indt}\}$ and all auxiliary variables and also indices n, d and t from $\{\boldsymbol{\mu}_{ndt}, \boldsymbol{\nu}_{ndt}, \boldsymbol{\Sigma}_{ndt}, \boldsymbol{\Phi}_{ndt}, \hat{\boldsymbol{\mu}}_{ndt}, \hat{\boldsymbol{\nu}}_{ndt}, \hat{\boldsymbol{\Sigma}}_{ndt}, \hat{\boldsymbol{\Phi}}_{ndt}\}$.

A. Distributionally Robust Chance Constraints

According to the reformulation algorithm explained in Section IV-A, the resulting constraints from the reformulation of (8) and (18) are written as

$$\sqrt{\frac{1-\epsilon-\eta_{its}^{-\alpha}}{\epsilon}}\|\beta_{it}\mathbf{1}^\top\boldsymbol{\Phi}\| \leqslant \eta_{its}[P_i^{\max}x_{it} - p_{it} - \beta_{it}\mathbf{1}^\top\boldsymbol{\nu}] - (\frac{\alpha+1}{\alpha})(\beta_{it}\mathbf{1}^\top)(\boldsymbol{\mu}-\boldsymbol{\nu}),\ \forall i,t,s \tag{22a}$$

$$\sqrt{\frac{1-\epsilon-\gamma_{its}^{-\alpha}}{\epsilon}}\|\beta_{it}\mathbf{1}^\top\boldsymbol{\Phi}\| \leqslant \gamma_{its}[p_{it} - P_i^{\min}x_{it} + \beta_{it}\mathbf{1}^\top\boldsymbol{\nu}] + (\frac{\alpha+1}{\alpha})(\beta_{it}\mathbf{1}^\top)(\boldsymbol{\mu}-\boldsymbol{\nu}),\ \forall i,t,s \tag{22b}$$

$$\sqrt{\frac{1-\epsilon-\delta_{its}^{-\alpha}}{\epsilon}}\|\hat{\boldsymbol{\beta}}_{it}^\top\hat{\boldsymbol{\Phi}}\| \leqslant \delta_{its}[p_{i(t-1)} - p_{it} + R_i^{\mathrm{up}}x_{i(t-1)} + R_i^{\mathrm{st}}(1-x_{i(t-1)}) - \hat{\boldsymbol{\beta}}_i^\top\hat{\boldsymbol{\nu}}] - (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_i^\top(\hat{\boldsymbol{\mu}}-\boldsymbol{\nu}), \forall i,t,s \tag{22c}$$

$$\sqrt{\frac{1-\epsilon-\zeta_{its}^{-\alpha}}{\epsilon}}\|\hat{\boldsymbol{\beta}}_{it}^\top\hat{\boldsymbol{\Phi}}\| \leqslant \zeta_{its}[p_{it} - p_{i(t-1)} + R_i^{\mathrm{dn}}x_{it} + R_i^{\mathrm{st}}(1-x_{it}) + \hat{\boldsymbol{\beta}}_i^\top\hat{\boldsymbol{\nu}}] + (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_i^\top(\hat{\boldsymbol{\mu}}-\hat{\boldsymbol{\nu}}), \forall i,t,s \tag{22d}$$

$$\sqrt{\frac{1-\epsilon-\varphi_{jts}^{-\alpha}}{\epsilon}}\|(\mathbf{M}_j^{\mathrm{G}}\boldsymbol{\beta}_t\mathbf{1}^\top + \mathbf{M}_j^{\mathrm{W}})\boldsymbol{\Phi}\| \leqslant \varphi_{jts}[F_j^{\max} + \mathbf{M}_j^{\mathrm{D}}\mathbf{l}_t - \mathbf{M}_j^{\mathrm{G}}\mathbf{p}_t - \mathbf{M}_j^{\mathrm{W}}\mathbf{w}_t - (\mathbf{M}_j^{\mathrm{G}}\boldsymbol{\beta}_t\mathbf{1}^\top + \mathbf{M}_j^{\mathrm{W}})\mathbf{1}^\top\boldsymbol{\nu}] - (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\mathrm{G}}\boldsymbol{\beta}_t\mathbf{1}^\top + \mathbf{M}_j^{\mathrm{W}})(\boldsymbol{\mu}-\boldsymbol{\nu}), \quad \forall j,t,s \tag{22e}$$

$$\sqrt{\frac{1-\epsilon-\sigma_{jts}^{-\alpha}}{\epsilon}}\|(\mathbf{M}_j^{\mathrm{G}}\boldsymbol{\beta}_t\mathbf{1}^\top + \mathbf{M}_j^{\mathrm{W}})\boldsymbol{\Phi}\| \leqslant \sigma_{jts}[F_j^{\max} - \mathbf{M}_j^{\mathrm{D}}\mathbf{l}_t + \mathbf{M}_j^{\mathrm{G}}\mathbf{p}_t + \mathbf{M}_j^{\mathrm{W}}\mathbf{w}_t + (\mathbf{M}_j^{\mathrm{G}}\boldsymbol{\beta}_t\mathbf{1}^\top + \mathbf{M}_j^{\mathrm{W}})\boldsymbol{\nu}] + (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\mathrm{G}}\boldsymbol{\beta}_t\mathbf{1}^\top + \mathbf{M}_j^{\mathrm{W}})(\boldsymbol{\mu}-\boldsymbol{\nu}), \quad \forall j,t,s \tag{22f}$$

$$\sqrt{\frac{1-\epsilon-\lambda_{its}^{-\alpha}}{\epsilon}}\|\beta_{i(t-1)}\mathbf{1}^\top\boldsymbol{\Phi}\| \leqslant \lambda_{its}[R_i^{\mathrm{st}}x_{i(t-1)} + (P_i^{\max} - R_i^{\mathrm{st}})(x_{it} - v_{it}) - p_{i(t-1)} - \beta_{i(t-1)}\mathbf{1}^\top\boldsymbol{\nu}] - (\frac{\alpha+1}{\alpha})(\beta_{i(t-1)}\mathbf{1}^\top)(\boldsymbol{\mu}-\boldsymbol{\nu}), \quad \forall i,t,s \tag{22g}$$

$$\sqrt{\frac{1-\epsilon-\pi_{its}^{-\alpha}}{\epsilon}}\|\beta_{it}\mathbf{1}^\top\boldsymbol{\Phi}\| \leqslant \pi_{its}[(P_i^{\max} - R_i^{\mathrm{st}})v_{it} - p_{it} - \beta_{it}\mathbf{1}^\top\boldsymbol{\nu}] - (\frac{\alpha+1}{\alpha})(\beta_{it}\mathbf{1}^\top)(\boldsymbol{\mu}-\boldsymbol{\nu}), \forall i,t,s \tag{22h}$$

$$\sqrt{\frac{1-\epsilon-\rho_{its}^{-\alpha}}{\epsilon}}\|\hat{\boldsymbol{\beta}}_{it}^\top\hat{\boldsymbol{\Phi}}\| \leqslant \rho_{its}[p_{i(t-1)} - p_{it} + (P_i^{\min} + R_i^{\mathrm{up}})x_{it} - (P_i^{\min} + R_i^{\mathrm{up}} - R_i^{\mathrm{st}})v_{it} - P_i^{\min}x_{i(t-1)} - \hat{\boldsymbol{\beta}}_i^\top\hat{\boldsymbol{\nu}}] - (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_i^\top(\hat{\boldsymbol{\mu}}-\hat{\boldsymbol{\nu}}), \forall i,t,s \tag{22i}$$

$$\sqrt{\frac{1-\epsilon-\psi_{its}^{-\alpha}}{\epsilon}}\|\hat{\boldsymbol{\beta}}_{it}^\top\hat{\boldsymbol{\Phi}}\| \leqslant \psi_{its}[p_{it} - p_{i(t-1)} + R_i^{\mathrm{st}}x_{i(t-1)} - (P_i^{\min} + R_i^{\mathrm{dn}} - R_i^{\mathrm{st}})v_{it} - (R_i^{\mathrm{st}} - R_i^{\mathrm{dn}})x_{it} + R_i^{\mathrm{st}}(1-x_{it}) + \hat{\boldsymbol{\beta}}_i^\top\hat{\boldsymbol{\nu}}] + (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_i^\top(\hat{\boldsymbol{\mu}}-\hat{\boldsymbol{\nu}}), \forall i,t,s, \tag{22j}$$

where index s indicates iteration s and auxiliary variables $\{\eta_{its}, \gamma_{its}, \delta_{its}, \zeta_{its}, \varphi_{jts}, \sigma_{jts}, \lambda_{its}, \pi_{its}, \rho_{its}, \psi_{its}\}$ are used to reformulate constraints (8a)-(8f),(18a)-(19d), respectively. In (22c)-(22d) and (22i)-(22j), variable vector $\hat{\boldsymbol{\beta}_{it}}$ is defined as $\hat{\boldsymbol{\beta}_{it}} = [\beta_{it}\mathbf{1}^\top \quad -\beta_{i(t-1)}\mathbf{1}^\top]^\top$.

B. Distributionally Robust CVaR Constraints

In a case that probabilistic constraints are considered in the form of distributionally robust CVaR constraints, i.e. (9) and (19), they are approximately reformulated as

$$\left\|\begin{matrix}\theta_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu} \\ \boldsymbol{\Phi}\beta_i\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](P_i^{\max}x_{it} - p_{it}) - [\frac{2\epsilon(\alpha+1)}{\alpha} - 1]\theta_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu}, \quad \forall i,t \tag{23a}$$

$$\left\|\begin{matrix}\theta_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu} \\ \boldsymbol{\Phi}\beta_i\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](P_i^{\max}x_{it} - p_{it}) - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\theta_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu}, \quad \forall i,t \tag{23b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\theta_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu} \\ \boldsymbol{\Phi}\beta_i\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](P_i^{\max}x_{it} - p_{it}) - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\theta_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu}, \quad \forall i,t \tag{23c}$$

$$\left\|\begin{matrix}\gamma_{it} + (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu} \\ \boldsymbol{\Phi}\beta_i\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{it} - P_i^{\min}x_{it}) - [\frac{2\epsilon(\alpha+1)}{\alpha} - 1]\gamma_{it} + (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu}, \quad \forall i,t \tag{24a}$$

$$\left\|\begin{matrix}\gamma_{it} + (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu} \\ \boldsymbol{\Phi}\beta_i\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{it} - P_i^{\min}x_{it}) - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\gamma_{it} + (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu}, \quad \forall i,t \tag{24b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\gamma_{it} + (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu} \\ \boldsymbol{\Phi}\beta_i\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{it} - P_i^{\min}x_{it}) - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\gamma_{it} + (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^\top\boldsymbol{\mu}, \quad \forall i,t \tag{24c}$$

$$\left\|\begin{matrix}\delta_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{i(t-1)}-p_{it} + R_i^{\text{up}}x_{i(t-1)} + R_i^{\text{st}}(1-x_{i(t-1)})) - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\delta_{it} - (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{25a}$$

$$\left\|\begin{matrix}\delta_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{i(t-1)}-p_{it} + R_i^{\text{up}}x_{i(t-1)} + R_i^{\text{st}}(1-x_{i(t-1)})) - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\delta_{it} - (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{25b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\delta_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{i(t-1)}-p_{it} + R_i^{\text{up}}x_{i(t-1)} + R_i^{\text{st}}(1-x_{i(t-1)})) - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\delta_{it} - (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{25c}$$

$$\left\|\begin{matrix}\zeta_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ -\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{it}-p_{i(t-1)} + R_i^{\text{dn}}x_{it} + R_i^{\text{st}}(1-x_{it})) - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\zeta_{it} + (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{26a}$$

$$\left\|\begin{matrix}\zeta_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ -\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{it}-p_{i(t-1)} + R_i^{\text{dn}}x_{it} + R_i^{\text{st}}(1-x_{it})) - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\zeta_{it} + (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{26b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\zeta_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ -\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](p_{it}-p_{i(t-1)} + R_i^{\text{dn}}x_{it} + R_i^{\text{st}}(1-x_{it})) - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\zeta_{it} + (\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{26c}$$

$$\left\|\begin{matrix}\varphi_{jt}-(\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu}\\ \boldsymbol{\Phi}(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](F_j^{\max}+\mathbf{M}_j^{\text{D}}\mathbf{l}_t-\mathbf{M}_j^{\text{G}}\mathbf{p}_t-\mathbf{M}_j^{\text{W}}\mathbf{w}_t) - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\varphi_{jt} - (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu}, \forall j,t \tag{27a}$$

$$\left\|\begin{matrix}\varphi_{jt}-(\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\boldsymbol{\mu}\\ \boldsymbol{\Phi}(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\end{matrix}\right\| \leqslant -[\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\varphi_{jt} - (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu} + [\frac{2\epsilon(\alpha+1)}{\alpha}](F_j^{\max}+\mathbf{M}_j^{\text{D}}\mathbf{l}_t-\mathbf{M}_j^{\text{G}}\mathbf{p}_t-\mathbf{M}_j^{\text{W}}\mathbf{w}_t), \forall j,t \tag{27b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\varphi_{jt}-(\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu}\\ \boldsymbol{\Phi}(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\end{matrix}\right\| \leqslant - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\varphi_{jt} - (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu} \; [\frac{2\epsilon(\alpha+1)}{\alpha}](F_j^{\max}+\mathbf{M}_j^{\text{D}}\mathbf{l}_t-\mathbf{M}_j^{\text{G}}\mathbf{p}_t-\mathbf{M}_j^{\text{W}}\mathbf{w}_t), \forall j,t \tag{27c}$$

$$\left\|\begin{matrix}\sigma_{jt}+(\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu}\\ -\boldsymbol{\Phi}(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}](F_j^{\max}-\mathbf{M}_j^{\text{D}}\mathbf{l}_t+\mathbf{M}_j^{\text{G}}\mathbf{p}_t+\mathbf{M}_j^{\text{W}}\mathbf{w}_t) - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\sigma_{jt} + (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu}, \forall j,t \tag{28a}$$

$$\left\|\begin{matrix}\sigma_{jt}+(\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\boldsymbol{\mu}\\ -\boldsymbol{\Phi}(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\end{matrix}\right\| \leqslant -[\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\sigma_{jt} + (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu} + [\frac{2\epsilon(\alpha+1)}{\alpha}](F_j^{\max}-\mathbf{M}_j^{\text{D}}\mathbf{l}_t+\mathbf{M}_j^{\text{G}}\mathbf{p}_t+\mathbf{M}_j^{\text{W}}\mathbf{w}_t), \forall j,t \tag{28b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\sigma_{jt}+(\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu}\\ -\boldsymbol{\Phi}(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})^{\top}\end{matrix}\right\| \leqslant - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\sigma_{jt} + (\frac{\alpha+1}{\alpha})(\mathbf{M}_j^{\text{G}}\boldsymbol{\beta}_t\mathbf{1}^{\top}+\mathbf{M}_j^{\text{W}})\boldsymbol{\mu} \; [\frac{2\epsilon(\alpha+1)}{\alpha}](F_j^{\max}-\mathbf{M}_j^{\text{D}}\mathbf{l}_t+\mathbf{M}_j^{\text{G}}\mathbf{p}_t+\mathbf{M}_j^{\text{W}}\mathbf{w}_t), \forall j,t \tag{28c}$$

$$\left\|\begin{matrix}\lambda_{it}-(\frac{\alpha+1}{\alpha})\beta_{i(t-1)}\mathbf{1}^{\top}\boldsymbol{\mu}\\ \boldsymbol{\Phi}\beta_{i(t-1)}\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][R_i^{\text{st}}x_{i(t-1)}+(P_i^{\max}-R_i^{\text{st}})(x_{it}-v_{it})-p_{i(t-1)}] - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\lambda_{it} - (\frac{\alpha+1}{\alpha})\beta_{i(t-1)}\mathbf{1}^{\top}\boldsymbol{\mu}, \forall i,t \tag{29a}$$

$$\left\|\begin{matrix}\lambda_{it}-(\frac{\alpha+1}{\alpha})\beta_{i(t-1)}\mathbf{1}^{\top}\boldsymbol{\mu}\\ \boldsymbol{\Phi}\beta_{i(t-1)}\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][R_i^{\text{st}}x_{i(t-1)}+(P_i^{\max}-R_i^{\text{st}})(x_{it}-v_{it})-p_{i(t-1)}] - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\lambda_{it} - (\frac{\alpha+1}{\alpha})\beta_{i(t-1)}\mathbf{1}^{\top}\boldsymbol{\mu}, \forall i,t \tag{29b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\lambda_{it}-(\frac{\alpha+1}{\alpha})\beta_{i(t-1)}\mathbf{1}^{\top}\boldsymbol{\mu}\\ \boldsymbol{\Phi}\beta_{i(t-1)}\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][R_i^{\text{st}}x_{i(t-1)}+(P_i^{\max}-R_i^{\text{st}})(x_{it}-v_{it})-p_{i(t-1)}] - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\lambda_{it} - (\frac{\alpha+1}{\alpha})\beta_{i(t-1)}\mathbf{1}^{\top}\boldsymbol{\mu}, \forall i,t \tag{29c}$$

$$\left\|\begin{matrix}\pi_{it}-(\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^{\top}\boldsymbol{\mu}\\ \boldsymbol{\Phi}\beta_{it}\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][(P_i^{\max}-R_i^{\text{st}})v_{it}-p_{it}] - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\pi_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^{\top}\boldsymbol{\mu}, \forall i,t \tag{30a}$$

$$\left\|\begin{matrix}\pi_{it}-(\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^{\top}\boldsymbol{\mu}\\ \boldsymbol{\Phi}\beta_{it}\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][(P_i^{\max}-R_i^{\text{st}})v_{it}-p_{it}] - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\pi_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^{\top}\boldsymbol{\mu}, \forall i,t \tag{30b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\pi_{it}-(\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^{\top}\boldsymbol{\mu}\\ \boldsymbol{\Phi}\beta_{it}\mathbf{1}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][(P_i^{\max}-R_i^{\text{st}})v_{it}-p_{it}] - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\pi_{it} - (\frac{\alpha+1}{\alpha})\beta_{it}\mathbf{1}^{\top}\boldsymbol{\mu}, \forall i,t \tag{30c}$$

$$\left\|\begin{matrix}\rho_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][p_{i(t-1)}-p_{it}$$

$$+(P_i^{\min}+R_i^{\mathrm{up}})x_{it}-(P_i^{\min}+R_i^{\mathrm{up}}-R_i^{\mathrm{st}})v_{it}-P_i^{\min}x_{i(t-1)}] - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\rho_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{31a}$$

$$\left\|\begin{matrix}\rho_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][p_{i(t-1)}-p_{it} +(P_i^{\min}+R_i^{\mathrm{up}})x_{it}-(P_i^{\min}+R_i^{\mathrm{up}}-R_i^{\mathrm{st}})v_{it}-P_i^{\min}x_{i(t-1)}] - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\rho_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{31b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\rho_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ \hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][p_{i(t-1)}-p_{it} +(P_i^{\min}+R_i^{\mathrm{up}})x_{it}-(P_i^{\min}+R_i^{\mathrm{up}}-R_i^{\mathrm{st}})v_{it}-P_i^{\min}x_{i(t-1)}] - [\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\rho_{it}-(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{31c}$$

$$\left\|\begin{matrix}\psi_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ -\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][p_{it}+p_{i(t-1)}+R_i^{\mathrm{st}}x_{i(t-1)} -(P_i^{\min}+R_i^{\mathrm{dn}}-R_i^{\mathrm{st}})v_{it}-(R_i^{\mathrm{st}}-R_i^{\mathrm{dn}})x_{it}+R_i^{\mathrm{st}}(1-x_{it})] - [\frac{2\epsilon(\alpha+1)}{\alpha}-1]\psi_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{32a}$$

$$\left\|\begin{matrix}\psi_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ -\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][p_{it}+p_{i(t-1)}+R_i^{\mathrm{st}}x_{i(t-1)} -(P_i^{\min}+R_i^{\mathrm{dn}}-R_i^{\mathrm{st}})v_{it}-(R_i^{\mathrm{st}}-R_i^{\mathrm{dn}})x_{it}+R_i^{\mathrm{st}}(1-x_{it})] - [\frac{(2\epsilon-1)(\alpha+1)-1}{\alpha}]\psi_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \quad \forall i,t \tag{32b}$$

$$\left\|\begin{matrix}(\frac{\alpha+1}{\alpha})\psi_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}\\ -\hat{\boldsymbol{\Phi}}\hat{\boldsymbol{\beta}}_{it}\end{matrix}\right\| \leqslant [\frac{2\epsilon(\alpha+1)}{\alpha}][p_{it}+p_{i(t-1)} +R_i^{\mathrm{st}}x_{i(t-1)}-(P_i^{\min}+R_i^{\mathrm{dn}}-R_i^{\mathrm{st}})v_{it}-(R_i^{\mathrm{st}}-R_i^{\mathrm{dn}})x_{it} +R_i^{\mathrm{st}}(1-x_{it})]-[\frac{(2\epsilon-1)(\alpha+1)}{\alpha}]\psi_{it}+(\frac{\alpha+1}{\alpha})\hat{\boldsymbol{\beta}}_{it}^{\top}\hat{\boldsymbol{\mu}}, \forall i,t, \tag{32c}$$

where auxiliary variables $\{\theta_{it},\gamma_{it},\delta_{it},\zeta_{it},\varphi_{jt},\sigma_{jt},\lambda_{it},\pi_{it},\rho_{it},\psi_{it}\}$ are used to reformulate constraints (9a)-(9f) and (19a)-(19d), resulting in constraints (23)-(32), respectively.