A Parallelized, Adam-Based Solver for Reserve and Security Constrained AC Unit Commitment

Supplemental Information (SI)

Introduction

The general architecture of the "quasiGrad" solver is inspired by the latest neural network verification routines (e.g., α,β-CROWN), which have been massively successful. Meanwhile, the workhorse behind quasiGrad's numerical iterations is Adam, which is the most successful gradient-based algorithm for training large scale machine learning models (e.g., GPT-3, whose optimization hyperparameters we utilize exactly).

Gradient-Based Optimization via Backpropagation

In GO3, we seek to compute

$$\nabla_{\Omega} z^{\mathrm{ms}}. \tag{1}$$

Manually keeping track of the cascade of derivatives, however, is challenging, since multiple functions can be dependent on the same variable. For example, the function z might depend on variables z_1 *and* x_1: thus, their contributions sum together. To keep track of these sums, we generate a hierarchy of derivatives which all cascade back to the final objective function z^{ms}.

Penalized scoring functions

The upper-most scoring function is given by the scalar market surplus variable z^{ms}:

$$z^{\mathrm{ms}} \triangleq z^{\mathrm{base}} + z^{\mathrm{ctg,min}} + z^{\mathrm{ctg,avg}}. \tag{2}$$

As posed, the challenge is to *maximize* z^{ms}. However, in this formulation, we choose to *minimize* its negative:

$$\min\ z^{\mathrm{ms}'}$$

where

$$z^{\mathrm{ms}'} \triangleq -z^{\mathrm{base}} - z^{\mathrm{ctg,min}} - z^{\mathrm{ctg,avg}}. \tag{3}$$

The gradient of the market surplus variable is given by

$$\nabla_{z^{\mathrm{base}}} z^{\mathrm{ms}'} = -1 \tag{4}$$

$$\nabla_{z^{\mathrm{ctg,min}}} z^{\mathrm{ms}'} = -1 \tag{5}$$

$$\nabla_{z^{\mathrm{ctg,avg}}} z^{\mathrm{ms}'} = -1. \tag{6}$$

Next, we have the base objective:

$$z^{\mathrm{base}} = \sum_{t\in T} z_t^{\mathrm{t}} + \sum_{j\in J^{\mathrm{pr,cs}}} z_j^{\mathrm{en,max}} + \sum_{j\in J^{\mathrm{pr,cs}}} z_j^{\mathrm{en,min}} - {\color{blue}\delta} \sum_{j\in J^{\mathrm{pr,cs}}} \sum_{w\in W_j^{\mathrm{su,max}}} {\color{blue}z_{jw}^{\mathrm{mxst}}} \tag{7}$$

with an associated gradient

$$\nabla_{z_t^{\mathrm{t}}} z^{\mathrm{base}} = 1 \tag{8}$$

$$\nabla_{z_j^{\mathrm{en,max}}} z^{\mathrm{base}} = 1 \tag{9}$$

$$\nabla_{z_j^{\mathrm{en,min}}} z^{\mathrm{base}} = 1 \tag{10}$$

$$\nabla_{{\color{blue}z_{jw}^{\mathrm{mxst}}}} z^{\mathrm{base}} = -{\color{blue}\delta}. \tag{11}$$

The upper-most contingency scores are

$$z^{\text{ctg,min}} = \sum_{t\in T} z_t^{\text{ctg,min}} \tag{12}$$

$$z^{\text{ctg ,avg}} = \sum_{t\in T} z_t^{\text{ctg,avg}} \tag{13}$$

with associated gradients

$$\nabla_{z_t^{\text{ctg,min}}} z^{\text{ctg,min}} = 1 \tag{14}$$

$$\nabla_{z_t^{\text{ctg,avg}}} z^{\text{ctg ,avg}} = 1. \tag{15}$$

The interval contingency scores are given by

$$z_t^{\text{ctg,min}} = \min_{k\in K} z_{tk}^{\text{ctg}} \text{ if } |K| > 0 \text{ else } 0 \forall t \in T \tag{16}$$

$$z_t^{\text{ctg, avg}} = 1/|K| \sum_{k\in K} z_{tk}^{\text{ctg}} \text{ if } |K| > 0 \text{ else } 0 \forall t \in T \tag{17}$$

with gradients

$$\nabla_{z_{tk}^{\text{ctg}}} z_t^{\text{ctg,min}} = \begin{cases} 1, & z_{tk}^{\text{ctg}} = \min_{k\in K} z_{tk}^{\text{ctg}} \\ 0, & \text{else} \end{cases} \tag{18}$$

$$\nabla_{z_{tk}^{\text{ctg}}} z_t^{\text{ctg, avg}} = 1/|K|. \tag{19}$$

The base case minimum and maximum energy constraint violations are

$$z_j^{\text{en,max}} = -\sum_{w\in W_j^{\text{en,max}}} z_w^{\text{en,max}} \forall j \in J^{\text{pr,cs}} \tag{20}$$

$$z_j^{\text{en,min}} = -\sum_{w\in W_j^{\text{en,min}}} z_w^{\text{en,min}} \forall j \in J^{\text{pr,cs}} \tag{21}$$

with gradients

$$\nabla_{z_w^{\text{en,max}}} z_j^{\text{en,max}} = -1 \tag{22}$$

$$\nabla_{z_w^{\text{en,min}}} z_j^{\text{en,min}} = -1. \tag{23}$$

The time-dependent base-case market surplus is given by

$$\begin{aligned} z_t^{\text{t}} = &\sum_{j\in J^{\text{cs}}} z_{jt}^{\text{en}} - \sum_{j\in J^{\text{pr}}} z_{jt}^{\text{en}} - \sum_{j\in J^{\text{pr,cs,ac}}} \left(z_{jt}^{\text{su}} + z_{jt}^{\text{sd}}\right) - \sum_{j\in J^{\text{pr,cs}}} \left(z_{jt}^{\text{on}} + z_{jt}^{\text{sus}}\right) - \sum_{j\in J^{\text{ac}}} z_{jt}^{\text{s}} \\ &- \sum_{j\in J^{\text{pr,cs}}} \left(z_{jt}^{\text{rgu}} + z_{jt}^{\text{rgd}} + z_{jt}^{\text{scr}} + z_{jt}^{\text{nsc}} + z_{jt}^{\text{rru}} + z_{jt}^{\text{rrd}} + z_{jt}^{\text{qru}} + z_{jt}^{\text{qrd}}\right) - \sum_{i\in I} \left(z_{it}^{\text{p}} + z_{it}^{\text{q}}\right) \\ &- \sum_{n\in N^{\text{p}}} \left(z_{nt}^{\text{rgu}} + z_{nt}^{\text{rgd}} + z_{nt}^{\text{scr}} + z_{nt}^{\text{nsc}} + z_{nt}^{\text{rru}} + z_{nt}^{\text{rrd}}\right) - \sum_{n\in N^{\text{q}}} \left(z_{nt}^{\text{qru}} + z_{nt}^{\text{qrd}}\right) \\ &- \delta \sum_{j\in J^{\text{pr,cs}}} \hat{z}_{jt}, \forall\, t \in T, \end{aligned} \tag{24}$$

where $\hat{z}$ has been introduced in order to penalize hard constraint deviation. We set $\delta = 1$ in stage 1, where constraints are penalized, and $\delta = 0$ in stage 2, where constraints are projected. The definition of z_{jt} is given by

$$\hat{z}_{jt} = \hat{z}_{jt}^{\text{mndn}} + \hat{z}_{jt}^{\text{mnup}} + \hat{z}_{jt}^{\text{rup}} + \hat{z}_{jt}^{\text{rd}} + \hat{z}_{jt}^{\text{rgu}} + \hat{z}_{jt}^{\text{rgd}} + \hat{z}_{jt}^{\text{scr}} + \hat{z}_{jt}^{\text{nsc}} + \hat{z}_{jt}^{\text{rruon}} + \hat{z}_{jt}^{\text{rruoff}} + \hat{z}_{jt}^{\text{rrdon}} + \hat{z}_{jt}^{\text{rrdoff}} +$$

$$\hat{z}_{jt}^{\text{pmax,pr}} + \hat{z}_{jt}^{\text{pmin,pr}} + \hat{z}_{jt}^{\text{pmaxoff,pr}} + \hat{z}_{jt}^{\text{qmax,pr}} + \hat{z}_{jt}^{\text{qmin,pr}} + \hat{z}_{jt}^{\text{qmax,beta,pr}} + \hat{z}_{jt}^{\text{qmin,beta,pr}} +$$
$$\hat{z}_{jt}^{\text{pmax,cs}} + \hat{z}_{jt}^{\text{pmin,cs}} + \hat{z}_{jt}^{\text{pmaxoff,cs}} + \hat{z}_{jt}^{\text{qmax,cs}} + \hat{z}_{jt}^{\text{qmin,cs}} + \hat{z}_{jt}^{\text{qmax,beta,cs}} + \hat{z}_{jt}^{\text{qmin,beta,cs}}. \tag{25}$$

Notably, we do not penalize the hard constraints which are associated with trivially projected variables; these include all non-device variables:

$$\Omega^{\text{TPVs}} = v \cup \theta \cup \phi \cup \tau \cup p^{\text{fr,dc}} \cup q^{\text{fr,dc}} \cup q^{\text{to,dc}} \cup u^{\text{sh}} \cup u^{\text{on,ac}} \cup u^{\text{on,xfm}}, \tag{26}$$

since enforcing feasibility of the constraints associated with these variables is trivial. The *non*-trivially projected variables (NTPVs) are associated with the device level constraints.

For clarity, we split the AC start up costs, shutdown costs, and overload penalties into separate sets associated with ac line and transformers:

$$z_{jt}^{\text{su}} = \left\{ z_{jt}^{\text{su,ln}}, j \in J^{\text{ln}} \right\} \cup \left\{ z_{jt}^{\text{su,xf}}, j \in J^{\text{xf}} \right\} \tag{27}$$
$$z_{jt}^{\text{sd}} = \left\{ z_{jt}^{\text{sd,ln}}, j \in J^{\text{ln}} \right\} \cup \left\{ z_{jt}^{\text{sd,xf}}, j \in J^{\text{xf}} \right\} \tag{28}$$
$$z_{jt}^{\text{s}} = \left\{ z_{jt}^{\text{s,ln}}, j \in J^{\text{ln}} \right\} \cup \left\{ z_{jt}^{\text{s,xf}}, j \in J^{\text{xf}} \right\}. \tag{29}$$

The gradient of z_t^{t} is given below:

$$\nabla_x z_t^{\text{t}} = \begin{cases} +1, & x = z_{jt}^{\text{en}}, j \in J^{\text{cs}} \\ +\delta, & x = \hat{z} \\ -1, & \text{else.} \end{cases} \tag{30}$$

The contingency low-level objective terms are given by

$$z_{tk}^{\text{ctg}} = -\sum_{j \in J_k^{\text{ac}}} z_{jtk}^{\text{s}} \forall t \in T, k \in K. \tag{31}$$

The gradients are given with respect to the individual overloads across all contingencies:

$$\nabla_{z_{jtk}^{\text{s}}} z_{tk}^{\text{ctg}} = -1 \forall t \in T, k \in K. \tag{32}$$

In this case, the elements z_{jtk}^{s} correspond to both ac lines and transformers (since they are not significantly distinguished in the contingencies).

Power Balance

The active and reactive power balance equations are given by

$$p_{it} = \sum_{j\in J_i^{\mathrm{cs}}} p_{jt} + \sum_{j\in J_i^{\mathrm{sh}}} p_{jt} + \sum_{j\in J_i^{\mathrm{fr}}} p_{jt}^{\mathrm{fr}} + \sum_{j\in J_i^{\mathrm{to}}} p_{jt}^{\mathrm{to}} - \sum_{j\in J_i^{\mathrm{pr}}} p_{jt} \forall t \in T, i \in I \tag{33}$$

$$q_{it} = \sum_{j\in J_i^{\mathrm{cs}}} q_{jt} + \sum_{j\in J_i^{\mathrm{sh}}} q_{jt} + \sum_{j\in J_i^{\mathrm{fr}}} q_{jt}^{\mathrm{fr}} + \sum_{j\in J_i^{\mathrm{to}}} q_{jt}^{\mathrm{to}} - \sum_{j\in J_i^{\mathrm{pr}}} q_{jt} \forall t \in T, i \in I, \tag{34}$$

where they have been rearranged to isolate the signed mismatch variables p_{it} and q_{it}. We use an absolute value operator in order to compute the penalized mismatch variables p_{it}^+ and q_{it}^+:

$$p_{it}^+ = |p_{it}| \,\forall t \in T, i \in I \tag{35}$$

$$q_{it}^+ = |q_{it}| \,\forall t \in T, i \in I. \tag{36}$$

The mismatch penalties are computed as

$$z_{it}^{\mathrm{p}} = d_t c^{\mathrm{p}} p_{it}^+ \forall t \in T, i \in I \tag{37}$$

$$z_{it}^{\mathrm{q}} = d_t c^{\mathrm{q}} q_{it}^+ \forall t \in T, i \in I. \tag{38}$$

NOTE: we don't actually compute p_{it}^+ and q_{it}^+ in the code, unless specifically requested. Finally, voltage magnitudes are bounded by

$$v_i^{\min} \le v_{it} \le v_i^{\max} \forall t \in T, i \in I. \tag{39}$$

The gradients of the balance equations are given by

$$\nabla_{p_{jt}} p_{it} = \begin{cases} +1, & p_{jt} \in \{J_i^{\mathrm{cs}}, J_i^{\mathrm{sh}}, J_i^{\mathrm{fr}}, J_i^{\mathrm{to}}\} \\ -1, & p_{jt} \in \{J_i^{\mathrm{pr}}\} \end{cases} \tag{40}$$

$$\nabla_{q_{jt}} q_{it} = \begin{cases} +1, & q_{jt} \in \{J_i^{\mathrm{cs}}, J_i^{\mathrm{sh}}, J_i^{\mathrm{fr}}, J_i^{\mathrm{to}}\} \\ -1, & q_{jt} \in \{J_i^{\mathrm{pr}}\}. \end{cases} \tag{41}$$

The gradients of the absolute mismatches are given by

$$\nabla_{p_{it}} p_{it}^+ = \operatorname{sgn}(p_{it}) \,\forall t \in T, i \in I \tag{42}$$

$$\nabla_{q_{it}} q_{it}^+ = \operatorname{sgn}(q_{it}) \,\forall t \in T, i \in I. \tag{43}$$

The gradients of the penalties are given by

$$\nabla_{p_{it}^+} z_{it}^{\mathrm{p}} = d_t c^{\mathrm{p}} \forall t \in T, i \in I \tag{44}$$

$$\nabla_{q_{it}^+} z_{it}^{\mathrm{q}} = d_t c^{\mathrm{q}} \forall t \in T, i \in I. \tag{45}$$

Notably, voltage magnitude bounds are enforced by clipping, so their derivatives are not needed.

AC Branch Flow

The AC branch flow equations are given by

$$
\begin{aligned}
p_{jt}^{\mathrm{fr}} &= u_{jt}^{\mathrm{on}}\left(\left(g_j^{\mathrm{sr}}+g_j^{\mathrm{fr}}\right)v_{it}^2/\tau_{jt}^2+\left(-g_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right.\right.\\
&\quad\left.\left.-b_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}\right)\forall t\in T, j\in J^{\mathrm{ac}}, i=i_j^{\mathrm{fr}}, i'=i_j^{\mathrm{to}} && (46)\\
q_{jt}^{\mathrm{fr}} &= u_{jt}^{\mathrm{on}}\left(\left(-b_j^{\mathrm{sr}}-b_j^{\mathrm{fr}}-b_j^{\mathrm{ch}}/2\right)v_{it}^2/\tau_{jt}^2+\left(b_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right.\right.\\
&\quad\left.\left.-g_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}\right)\forall t\in T, j\in J^{\mathrm{ac}}, i=i_j^{\mathrm{fr}}, i'=i_j^{\mathrm{to}} && (47)\\
p_{jt}^{\mathrm{to}} &= u_{jt}^{\mathrm{on}}\left(\left(g_j^{\mathrm{sr}}+g_j^{\mathrm{to}}\right)v_{i't}^2+\left(-g_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right.\right.\\
&\quad\left.\left.+b_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}\right)\forall t\in T, j\in J^{\mathrm{ac}}, i=i_j^{\mathrm{fr}}, i'=i_j^{\mathrm{to}} && (48)\\
q_{jt}^{\mathrm{to}} &= u_{jt}^{\mathrm{on}}\left(\left(-b_j^{\mathrm{sr}}-b_j^{\mathrm{to}}-b_j^{\mathrm{ch}}/2\right)v_{i't}^2+\left(b_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right.\right.\\
&\quad\left.\left.+g_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}\right)\forall t\in T, j\in J^{\mathrm{ac}}, i=i_j^{\mathrm{fr}}, i'=i_j^{\mathrm{to}}. && (49)
\end{aligned}
$$

Reformulated flow limits and the associated penalty function are given by

$$
\begin{aligned}
s_{jt}^{\mathrm{fr},+} &= \left(\left(p_{jt}^{\mathrm{fr}}\right)^2+\left(q_{jt}^{\mathrm{fr}}\right)^2\right)^{1/2}-s_j^{\max}\forall t\in T, j\in J^{\mathrm{ac}} && (50)\\
s_{jt}^{\mathrm{to},+} &= \left(\left(p_{jt}^{\mathrm{to}}\right)^2+\left(q_{jt}^{\mathrm{to}}\right)^2\right)^{1/2}-s_j^{\max}\forall t\in T, j\in J^{\mathrm{ac}} && (51)\\
s_{jt}^{+} &= \max\left(s_{jt}^{\mathrm{fr},+}, s_{jt}^{\mathrm{to},+}, 0\right) && (52)\\
z_{jt}^{\mathrm{s}} &= d_t c^{\mathrm{s}} s_{jt}^{+}\forall t\in T, j\in J^{\mathrm{ac}}. && (53)
\end{aligned}
$$

Note: in the code, we do not define s_{jt}^{+}. Instead, we directly utilize $z_{jt}^{\mathrm{s}} = d_t c^{\mathrm{s}}\max\left(s_{jt}^{\mathrm{fr},+}, s_{jt}^{\mathrm{to},+}, 0\right)\forall t\in T, j\in J^{\mathrm{ac}}$.

We take the gradient associated with p_{jt}^{fr} with respect to all variables (the other flow derivatives are highly similar):

$$
\begin{aligned}
\nabla_{v_{it}}p_{jt}^{\mathrm{fr}} &= 2u_{jt}^{\mathrm{on}}\left(\left(g_j^{\mathrm{sr}}+g_j^{\mathrm{fr}}\right)v_{it}/\tau_{jt}^2+\left(-g_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)-b_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{i't}/\tau_{jt}\right) && (54)\\
\nabla_{v_{i't}}p_{jt}^{\mathrm{fr}} &= u_{jt}^{\mathrm{on}}\left(\left(-g_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)-b_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}/\tau_{jt}\right) && (55)\\
\nabla_{\theta_{it}}p_{jt}^{\mathrm{fr}} &= u_{jt}^{\mathrm{on}}\left(\left(g_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)-b_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}\right) && (56)\\
\nabla_{\theta_{i't}}p_{jt}^{\mathrm{fr}} &= u_{jt}^{\mathrm{on}}\left(\left(-g_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)+b_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}\right) && (57)\\
\nabla_{\tau_{jt}}p_{jt}^{\mathrm{fr}} &= u_{jt}^{\mathrm{on}}\left(-2\left(g_j^{\mathrm{sr}}+g_j^{\mathrm{fr}}\right)v_{it}^2/\tau_{jt}^3-\left(-g_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)-b_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}^2\right) && (58)\\
\nabla_{\phi_{jt}}p_{jt}^{\mathrm{fr}} &= u_{jt}^{\mathrm{on}}\left(\left(-g_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)+b_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}\right) && (59)\\
\nabla_{u_{jt}^{\mathrm{on}}}p_{jt}^{\mathrm{fr}} &= \left(g_j^{\mathrm{sr}}+g_j^{\mathrm{fr}}\right)v_{it}^2/\tau_{jt}^2+\left(-g_j^{\mathrm{sr}}\cos\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)-b_j^{\mathrm{sr}}\sin\left(\theta_{it}-\theta_{i't}-\phi_{jt}\right)\right)v_{it}v_{i't}/\tau_{jt}. && (60)
\end{aligned}
$$

Next, we take the gradient with respect to the flow limits and penalty function:

$$
\begin{aligned}
\nabla_{p_{jt}^{\mathrm{fr}}}s_{jt}^{\mathrm{fr},+} &= \frac{1}{2}\left(\left(p_{jt}^{\mathrm{fr}}\right)^2+\left(q_{jt}^{\mathrm{fr}}\right)^2\right)^{-1/2}\left(2p_{jt}^{\mathrm{fr}}\right) && (61)\\
&= \frac{p_{jt}^{\mathrm{fr}}}{s_{jt}^{\mathrm{fr},+}} && (62)\\
\nabla_{q_{jt}^{\mathrm{fr}}}s_{jt}^{\mathrm{fr},+} &= \frac{q_{jt}^{\mathrm{fr}}}{s_{jt}^{\mathrm{fr},+}} && (63)\\
\nabla_{p_{jt}^{\mathrm{to}}}s_{jt}^{\mathrm{to},+} &= \frac{p_{jt}^{\mathrm{to}}}{s_{jt}^{\mathrm{to},+}} && (64)\\
\nabla_{q_{jt}^{\mathrm{to}}}s_{jt}^{\mathrm{to},+} &= \frac{q_{jt}^{\mathrm{to}}}{s_{jt}^{\mathrm{to},+}} && (65)
\end{aligned}
$$

$$\nabla_{s^{\mathrm{fr},+}_{jt}} s^{+}_{jt} = \begin{cases} 1, & s^{\mathrm{fr},+}_{jt} = \max\left(s^{\mathrm{fr},+}_{jt}, s^{\mathrm{to},+}_{jt}, 0\right) \\ 0, & \text{else} \end{cases} \tag{66}$$

$$\nabla_{s^{\mathrm{to},+}_{jt}} s^{+}_{jt} = \begin{cases} 1, & s^{\mathrm{to},+}_{jt} = \max\left(s^{\mathrm{fr},+}_{jt}, s^{\mathrm{to},+}_{jt}, 0\right) \\ 0, & \text{else} \end{cases} \tag{67}$$

$$\nabla_{s^{+}_{jt}} z^{\mathrm{s}}_{jt} = d_t c^{\mathrm{s}} \tag{68}$$

Producing and Consuming Devices

In this subsection, we consider the constraints and costs associated with producing and consuming devices. Sub-subsection II also includes the costs associated with AC transmission lines.

I. Energy costs

We begin by considering the cost associated with the energy provided by producing and consuming devices. The energy cost is computed by piecewise linear convex (or concave) cost (or value) function.

$$0 \le p_{jtm} \le p^{\max}_{jtm} \forall t \in T, j \in J^{\mathrm{pr,cs}}, m \in M_{jt} \tag{69}$$

$$p_{jt} = \sum_{m \in M_{jt}} p_{jtm} \forall t \in T, j \in J^{\mathrm{pr,cs}} \tag{70}$$

$$z^{\mathrm{en}}_{jt} = d_t \sum_{m \in M_{jt}} c^{\mathrm{en}}_{jtm} p_{jtm} \forall t \in T, j \in J^{\mathrm{pr,cs}} \tag{71}$$

We transform these costs by first defining a total sum across all blocks

$$p^{\mathrm{max,total}}_{jt} = \sum_{m \in M_{jt}} p^{\max}_{jtm} \forall t \in T, j \in J^{\mathrm{pr,cs}} \tag{72}$$

along with the new associated constraint

$$0 \le p_{jt} \le p^{\mathrm{max,total}}_{jt} \ \forall t \in T, j \in J^{\mathrm{pr,cs}}. \tag{73}$$

Note: this last constraint isn't explicitly enforced, since it just represents the sum of energy produced across all blocks; the value of p_{jt} is enforced elsewhere.

The total cost associated with p_{jt} is computed by first defining the cumulative block size $p^{\mathrm{cum,max}}_{jtm_L}$, which assumes blocks $m \in M_{jt}$ are ordered via $1, 2, ..., l$ such that

$$c^{\mathrm{en}}_{jtm_1} \le c^{\mathrm{en}}_{jtm_2} \le ... \le c^{\mathrm{en}}_{jtm_L}, j \in J^{\mathrm{pr}} \tag{74}$$

$$c^{\mathrm{en}}_{jtm_1} \ge c^{\mathrm{en}}_{jtm_2} \ge ... \ge c^{\mathrm{en}}_{jtm_L}, j \in J^{\mathrm{cs}}. \tag{75}$$

Cumulative block size $p^{\mathrm{cum,max}}_{jtm_L}$ is given by

$$p^{\mathrm{cum,max}}_{jtm_L} = \sum_{l=1}^{L} p^{\max}_{jtm_l}. \tag{76}$$

The energy cost may be computed as

$$z^{\mathrm{en}}_{jt} = d_t \sum_{l=1}^{|M_{jt}|} c^{\mathrm{en}}_{jtm_l} \max\left(\min\left(p_{jt} - p^{\mathrm{cum,max}}_{jtm_{L=l-1}}, p_{jtm_l}\right), 0\right) \forall t \in T, j \in J^{\mathrm{pr,cs}} \tag{77}$$

where $p^{\mathrm{cum,max}}_{jtm_{L=0}} = 0$ (**note:** $l-1$ in the subscript). Furthermore, the gradient of the energy cost is computed as

$$\nabla_{p_{jt}} z^{\mathrm{en}}_{jt} = d_t c^{\mathrm{en}}_{jtm_l}, \ l = \underset{l}{\operatorname{argmin}} \left(p_{jt} - p^{\mathrm{cum,max}}_{jtm_{L=l}}, p_{jtm_l}\right) > 0. \tag{78}$$

II. On and off relationships (including AC lines)

Our goal here is to define two key relations: the mapping from the IIVs Ω to the associated market surplus function terms, and hard inequality device constraints. We begin by stating the binary "online" variable:

$$u_{jt}^{\text{on}} \in \{0,1\} \forall t \in T, j \in J^{\text{pr,cs,ac}}. \tag{79}$$

Its relaxed counterpart is given by

$$0 \le u_{jt}^{\text{on}} \le 1 \forall t \in T, j \in J^{\text{pr,cs,ac}}. \tag{80}$$

The online variable is used to explicitly define the startup and shutdown variables:

$$u_{jt}^{\text{su}} \triangleq +\max\left(u_{jt}^{\text{on}} - u_{j,t-1}^{\text{on}}, 0\right) \forall t > t^{\text{start}}, j \in J^{\text{pr,cs,ac}} \tag{81}$$

$$u_{jt}^{\text{sd}} \triangleq -\min\left(u_{jt}^{\text{on}} - u_{j,t-1}^{\text{on}}, 0\right) \forall t > t^{\text{start}}, j \in J^{\text{pr,cs,ac}} \tag{82}$$

The costs associated with online, startup, and shutdown binary variables are given by

$$z_{jt}^{\text{on}} = d_t c_j^{\text{on}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{pr,cs}} \tag{83}$$

$$z_{jt}^{\text{su}} = c_j^{\text{su}} u_{jt}^{\text{su}} \forall t \in T, j \in J^{\text{pr,cs,ac}} \tag{84}$$

$$z_{jt}^{\text{sd}} = c_j^{\text{sd}} u_{jt}^{\text{sd}} \forall t \in T, j \in J^{\text{pr,cs,ac}}. \tag{85}$$

The cost associated with starting up can be discounted if the device is still be in one (or several) of the "startup state(s)". Rather than defining an inequality constrained startup variable, we capture the same effect by first defining the bounding variable $u_{jft}^{\text{sus,bnd}}$:

$$u_{jft}^{\text{sus,bnd}} = \begin{cases} \max_{t' \in T_{jft}^{\text{sus}}} \left\{u_{jt'}^{\text{on}}\right\}, & \forall t \in T_{jf}^{\text{sus}} \\ 1, & \forall t \notin T_{jf}^{\text{sus}} \end{cases}, j \in J^{\text{pr,cs}}, f \in F_j, \tag{86}$$

$$u_{jft}^{\text{sus,bnd}} = \begin{cases} \max\left(\sum_{t' \in T_{jft}^{\text{sus}}} u_{jt'}^{\text{on}}, 1\right), & \forall t \in T_{jf}^{\text{sus}} \\ 1, & \forall t \notin T_{jf}^{\text{sus}} \end{cases}, j \in J^{\text{pr,cs}}, f \in F_j, \tag{87}$$

where the second equality indicates that if a device does not have a sufficient amount of accumulated downtime at the start of the optimization window, then the startup state bound is automatically 1.

Using $u_{jft}^{\text{sus,bnd}}$, the startup reward (i.e., discount) can be computed by taking the minimum between the startup variable (i.e., are we starting up?) and the sus-bound (i.e., can we discount the startup?) – both must be true for a discount to be applied:

$$z_{jt}^{\text{sus}} = \sum_{f \in F_j} c_{jf}^{\text{sus}} \min\left(u_{jt}^{\text{su}}, u_{jft}^{\text{sus,bnd}}\right) \forall t \in T, j \in J^{\text{pr,cs}}. \tag{88}$$

III. Total active power

The total power injected into the network p_{jt} is equal to the dispatchable power p_{jt}^{on}

$$p_{jt} = p_{jt}^{\text{on}} + p_{jt}^{\text{su}} + p_{jt}^{\text{sd}} \forall t \in T, j \in J^{\text{pr,cs}}, \tag{89}$$

where p_{jt}, p_{jt}^{su}, and p_{jt}^{sd} are intermediate variables, with the startup and shutdown variables given as

$$p_{jt}^{\text{su}} = \sum_{t' \in T_{jt}^{\text{supc}}} p_{jtt'}^{\text{supc}} u_{jt'}^{\text{su}} \forall t \in T, j \in J^{\text{pr,cs}} \tag{90}$$

$$p_{jt}^{\text{sd}} = \sum_{t' \in T_{jt}^{\text{sdpc}}} p_{jtt'}^{\text{sdpc}} u_{jt'}^{\text{sd}} \forall t \in T, j \in J^{\text{pr,cs}}. \tag{91}$$

The parameters $p^{\mathrm{supc}}_{jtt'}$ and $p^{\mathrm{sdpc}}_{jtt'}$ are constant valued, constructed as

$$p^{\mathrm{supc}}_{jtt'} = p^{\mathrm{min}}_{jt'} - p^{\mathrm{ru,su}}_{j}\left(a^{\mathrm{end}}_{t'} - a^{\mathrm{end}}_{t}\right) \forall j \in J^{\mathrm{pr,cs}}, t < t' \in T \tag{92}$$

$$T^{\mathrm{supc}}_{jt} = \left\{t' > t : p^{\mathrm{supc}}_{jtt'} > 0\right\} \forall j \in J^{\mathrm{pr,cs}}, t \in T \tag{93}$$

$$p^{\mathrm{sdpc}}_{jtt'} = p^{\mathrm{min}}_{j,t'-1} - p^{\mathrm{rd,sd}}_{j}\left(a^{\mathrm{end}}_{t} - a^{\mathrm{start}}_{t'}\right) \forall j \in J^{\mathrm{pr,cs}}, t, t' \in T, t \geq t' > t^{\mathrm{start}} \tag{94}$$

$$p^{\mathrm{sdpc}}_{jtt'} = p^{0}_{j} - p^{\mathrm{rd,sd}}_{j}\left(a^{\mathrm{end}}_{t} - a^{\mathrm{start}}_{t'}\right) \forall j \in J^{\mathrm{pr,cs}}, t, t' \in T, t \geq t' = t^{\mathrm{start}} \tag{95}$$

$$T^{\mathrm{sdpc}}_{jt} = \left\{t' \leq t : p^{\mathrm{sdpc}}_{jtt'} > 0\right\} \forall j \in T^{\mathrm{pr,cs}}, t \in T. \tag{96}$$

IV. Maximum and minimum energy penalties

The penalty equations associated with maximum and minimum energy values are reformulated in typical fashion via the max operator.

$$z^{\mathrm{en,max}}_{w} = c^{\mathrm{e}} \max\left(\sum_{t \in T^{\mathrm{en,max}}_{w}} d_t p_{jt} - e^{\mathrm{max}}_{w}, 0\right) \forall j \in J^{\mathrm{pr,cs}}, w \in W^{\mathrm{en,max}}_{j} \tag{97}$$

$$z^{\mathrm{en,min}}_{w} = c^{\mathrm{e}} \max\left(e^{\mathrm{min}}_{w} - \sum_{t \in T^{\mathrm{en,min}}_{w}} d_t p_{jt}, 0\right) \forall j \in J^{\mathrm{pr,cs}}, w \in W^{\mathrm{en,min}}_{j}. \tag{98}$$

V. Projected constraints (stage 1)

In stage 1, a subset of the device level constraints are enforced via projection. First, we project the binaries to lie within [0, 1] (this includes AC lines):

$$u^{\mathrm{on}}_{jt} = \min(\max(u^{\mathrm{on}}_{jt}, 0), 1) \forall t \in T, j \in J^{\mathrm{pr,cs,ac}}. \tag{99}$$

Next, we project all device reserve variables nonnegative; each one of these projections is applied $\forall t \in T, j \in J^{\mathrm{pr,cs}}$:

$$p^{\mathrm{rgu}}_{jt} = \max(p^{\mathrm{rgu}}_{jt}, 0) \tag{100}$$

$$p^{\mathrm{rgd}}_{jt} = \max(p^{\mathrm{rgd}}_{jt}, 0) \tag{101}$$

$$p^{\mathrm{scr}}_{jt} = \max(p^{\mathrm{scr}}_{jt}, 0) \tag{102}$$

$$p^{\mathrm{nsc}}_{jt} = \max(p^{\mathrm{nsc}}_{jt}, 0) \tag{103}$$

$$p^{\mathrm{rru,on}}_{jt} = \max(p^{\mathrm{rru,on}}_{jt}, 0) \tag{104}$$

$$p^{\mathrm{rru,off}}_{jt} = \max(p^{\mathrm{rru,off}}_{jt}, 0) \tag{105}$$

$$p^{\mathrm{rrd,on}}_{jt} = \max(p^{\mathrm{rrd,on}}_{jt}, 0) \tag{106}$$

$$p^{\mathrm{rrd,off}}_{jt} = \max(p^{\mathrm{rrd,off}}_{jt}, 0) \tag{107}$$

$$q^{\mathrm{qru}}_{jt} = \max(q^{\mathrm{qru}}_{jt}, 0) \tag{108}$$

$$q^{\mathrm{qrd}}_{jt} = \max(q^{\mathrm{qrd}}_{jt}, 0). \tag{109}$$

VI. Penalized constraints (stage 1)

In this subsection, we enumerate the constraints which are initially penalized via a penalty function and later enforced via Lagrange Duality. The relaxed on/off binary variables associated with devices fall into a special class, since they are initially "trivially projected" during the penalization stage, and later enforced via LD.

1. **Minimum downtime:** The original constraint is given by

$$u^{\mathrm{su}}_{jt} \leq 1 - \sum_{t' \in T^{\mathrm{dn,min}}_{jt}} u^{\mathrm{sd}}_{jt'} \forall t \in T, j \in J^{\mathrm{pr,cs}}. \tag{110}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{mndn}} = d_t \hat{c}^{\text{mndn}} \max\left(u_{jt}^{\text{su}} + \sum_{t' \in T_{jt}^{\text{dn,min}}} u_{jt'}^{\text{sd}} - 1, 0 \right)^2. \tag{111}$$

2. Minimum uptime: The original constraint is given by

$$u_{jt}^{\text{sd}} \leq 1 - \sum_{t' \in T_{jt}^{\text{up,min}}} u_{jt'}^{\text{su}} \forall t \in T, j \in J^{\text{pr,cs}} \tag{112}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{mnup}} = d_t \hat{c}^{\text{mnup}} \max\left(u_{jt}^{\text{sd}} + \sum_{t' \in T_{jt}^{\text{up,min}}} u_{jt'}^{\text{su}} - 1, 0 \right)^2. \tag{113}$$

3. Ramping limits (up): The original constraint is given by

$$p_{jt} - p_{j,t-1} \leq d_t \left(p_j^{\text{ru}} \left(u_{jt}^{\text{on}} - u_{jt}^{\text{su}} \right) + p_j^{\text{ru,su}} \left(u_{jt}^{\text{su}} + 1 - u_{jt}^{\text{on}} \right) \right) \forall t > t^{\text{start}}, j \in J^{\text{pr,cs}} \tag{114}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rup}} = d_t \hat{c}^{\text{rup}} \max\left(p_{jt} - p_{j,t-1} - d_t \left(p_j^{\text{ru}} \left(u_{jt}^{\text{on}} - u_{jt}^{\text{su}} \right) + p_j^{\text{ru,su}} \left(u_{jt}^{\text{su}} + 1 - u_{jt}^{\text{on}} \right) \right), 0 \right)^2. \tag{115}$$

4. Ramping limits (down): The original constraint is given by

$$p_{jt} - p_{j,t-1} \geq -d_t \left(p_j^{\text{rd}} u_{jt}^{\text{on}} + p_j^{\text{rd,sd}} \left(1 - u_{jt}^{\text{on}} \right) \right) \forall t > t^{\text{start}}, j \in J^{\text{pr,cs}} \tag{116}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rd}} = d_t \hat{c}^{\text{rd}} \max\left(p_{j,t-1} - p_{jt} - d_t \left(p_j^{\text{rd}} u_{jt}^{\text{on}} + p_j^{\text{rd,sd}} \left(1 - u_{jt}^{\text{on}} \right) \right), 0 \right)^2. \tag{117}$$

5. Regulation up: The original constraint is given by

$$p_{jt}^{\text{rgu}} \leq p_j^{\text{rgu,max}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{pr,cs}}. \tag{118}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rgu}} = d_t \hat{c}^{\text{rgu}} \max\left(p_{jt}^{\text{rgu}} - p_j^{\text{rgu,max}} u_{jt}^{\text{on}}, 0 \right)^2. \tag{119}$$

6. Regulation down: The original constraint is given by

$$p_{jt}^{\text{rgd}} \leq p_j^{\text{rgd,max}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{pr,cs}}. \tag{120}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rgd}} = d_t \hat{c}^{\text{rgd}} \max\left(p_{jt}^{\text{rgd}} - p_j^{\text{rgd,max}} u_{jt}^{\text{on}}, 0 \right)^2. \tag{121}$$

7. Synchronized reserve: The original constraint is given by

$$p_{jt}^{\text{rgu}} + p_{jt}^{\text{scr}} \leq p_j^{\text{scr,max}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{pr,cs}} \tag{122}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{scr}} = d_t \hat{c}^{\text{scr}} \max\left(p_{jt}^{\text{rgu}} + p_{jt}^{\text{scr}} - p_j^{\text{scr,max}} u_{jt}^{\text{on}}, 0 \right)^2. \tag{123}$$

8. Non-synchronized reserve: The original constraint is given by

$$p_{jt}^{\text{nsc}} \leq p_j^{\text{nsc, max}} \left(1 - u_{jt}^{\text{on}}\right) \quad \forall t \in T, j \in J^{\text{pr,cs}}. \tag{124}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{nsc}} = d_t \hat{c}^{\text{nsc}} \max\left(p_{jt}^{\text{nsc}} - p_j^{\text{nsc, max}} \left(1 - u_{jt}^{\text{on}}\right), 0\right)^2. \tag{125}$$

9. Ramping reserve up (on): The original constraint is given by

$$p_{jt}^{\text{rgu}} + p_{jt}^{\text{scr}} + p_{jt}^{\text{rru,on}} \leq p_j^{\text{rru,on,max}} u_{jt}^{\text{on}} \; \forall t \in T, j \in J^{\text{pr,cs}}. \tag{126}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rru,on}} = d_t \hat{c}^{\text{rru,on}} \max\left(p_{jt}^{\text{rgu}} + p_{jt}^{\text{scr}} + p_{jt}^{\text{rru,on}} - p_j^{\text{rru,on,max}} u_{jt}^{\text{on}}, 0\right)^2. \tag{127}$$

10. Ramping reserve up (off): The original constraint is given by

$$p_{jt}^{\text{nsc}} + p_{jt}^{\text{rru,off}} \leq p_j^{\text{rru,off,max}} \left(1 - u_{jt}^{\text{on}}\right) \quad \forall t \in T, j \in J^{\text{pr,cs}}. \tag{128}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rru,off}} = d_t \hat{c}^{\text{rru,off}} \max\left(p_{jt}^{\text{nsc}} + p_{jt}^{\text{rru,off}} - p_j^{\text{rru,off,max}} \left(1 - u_{jt}^{\text{on}}\right), 0\right)^2. \tag{129}$$

11. Ramping reserve down (on): The original constraint is given by

$$p_{jt}^{\text{rgd}} + p_{jt}^{\text{rrd,on}} \leq p_j^{\text{rrd,on,max}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{pr,cs}}. \tag{130}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rrd,on}} = d_t \hat{c}^{\text{rrd,on}} \max\left(p_{jt}^{\text{rgd}} + p_{jt}^{\text{rrd,on}} - p_j^{\text{rrd,on,max}} u_{jt}^{\text{on}}, 0\right)^2 \tag{131}$$

12. Ramping reserve down (off): The original constraint is given by

$$p_{jt}^{\text{rrd,off}} \leq p_j^{\text{rrd, off,max}} \left(1 - u_{jt}^{\text{on}}\right) \forall t \in T, j \in J^{\text{pr,cs}}. \tag{132}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{rrd,off}} = d_t \hat{c}^{\text{rrd,off}} \max\left(p_{jt}^{\text{rrd,off}} - p_j^{\text{rrd, off,max}} \left(1 - u_{jt}^{\text{on}}\right), 0\right)^2 \tag{133}$$

13p. Maximum reserve limits (producer): The original constraint is given by

$$p_{jt}^{\text{on}} + p_{jt}^{\text{rgu}} + p_{jt}^{\text{scr}} + p_{jt}^{\text{rru,on}} \leq p_{jt}^{\text{max}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{pr}}. \tag{134}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{pmax,pr}} = d_t \hat{c}^{\text{pmax,pr}} \max\left(p_{jt}^{\text{on}} + p_{jt}^{\text{rgu}} + p_{jt}^{\text{scr}} + p_{jt}^{\text{rru,on}} - p_{jt}^{\text{max}} u_{jt}^{\text{on}}, 0\right)^2. \tag{135}$$

13c. Maximum reserve limits (consumer): The original constraint is given by

$$p_{jt}^{\text{on}} + p_{jt}^{\text{rgd}} + p_{jt}^{\text{rrd,on}} \leq p_{jt}^{\text{max}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{cs}}. \tag{136}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{pmax,cs}} = d_t \hat{c}^{\text{pmax,cs}} \max\left(p_{jt}^{\text{on}} + p_{jt}^{\text{rgd}} + p_{jt}^{\text{rrd,on}} - p_{jt}^{\text{max}} u_{jt}^{\text{on}}, 0\right)^2 \tag{137}$$

14p. Minimum reserve limits (producer): The original constraint is given by

$$p_{jt}^{\text{on}} - p_{jt}^{\text{rgd}} - p_{jt}^{\text{rrd,on}} \geq p_{jt}^{\text{min}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{pr}}. \tag{138}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{pmin,pr}} = d_t \hat{c}^{\text{pmin,pr}} \max\left(p_{jt}^{\text{min}} u_{jt}^{\text{on}} + p_{jt}^{\text{rrd,on}} + p_{jt}^{\text{rgd}} - p_{jt}^{\text{on}}, 0\right)^2. \tag{139}$$

14c. Minimum reserve limits (consumer): The original constraint is given by

$$p_{jt}^{\text{on}} - p_{jt}^{\text{rgu}} - p_{jt}^{\text{scr}} - p_{jt}^{\text{rru,on}} \geq p_{jt}^{\text{min}} u_{jt}^{\text{on}} \forall t \in T, j \in J^{\text{cs}}. \tag{140}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{pmin,cs}} = d_t \hat{c}^{\text{pmin,cs}} \max\left(p_{jt}^{\text{min}} u_{jt}^{\text{on}} + p_{jt}^{\text{rru,on}} + p_{jt}^{\text{scr}} + p_{jt}^{\text{rgu}} - p_{jt}^{\text{on}}, 0\right)^2. \tag{141}$$

15p. Off reserve limits (producer): The original constraint is given by

$$p_{jt}^{\text{su}} + p_{jt}^{\text{sd}} + p_{jt}^{\text{nsc}} + p_{jt}^{\text{rru,off}} \leq p_{jt}^{\text{max}} \left(1 - u_{jt}^{\text{on}}\right) \forall t \in T, j \in J^{\text{pr}}. \tag{142}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{pmax,off,pr}} = d_t \hat{c}^{\text{pmax,off,pr}} \max\left(p_{jt}^{\text{su}} + p_{jt}^{\text{sd}} + p_{jt}^{\text{nsc}} + p_{jt}^{\text{rru,off}} - p_{jt}^{\text{max}} \left(1 - u_{jt}^{\text{on}}\right), 0\right)^2 \tag{143}$$

15c. Off reserve limits (consumer): The original constraint is given by

$$p_{jt}^{\text{su}} + p_{jt}^{\text{sd}} + p_{jt}^{\text{rrd,off}} \leq p_{jt}^{\text{max}} \left(1 - u_{jt}^{\text{on}}\right) \forall t \in T, j \in J^{\text{cs}}. \tag{144}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{pmax,off,pr}} = d_t \hat{c}^{\text{pmax,off,pr}} \max\left(p_{jt}^{\text{su}} + p_{jt}^{\text{sd}} + p_{jt}^{\text{rrd,off}} - p_{jt}^{\text{max}} \left(1 - u_{jt}^{\text{on}}\right), 0\right)^2. \tag{145}$$

16p. Maximum reactive power reserves (producer): The original constraint is given by

$$q_{jt} + q_{jt}^{\text{qru}} \leq q_{jt}^{\text{max}} \left(u_{jt}^{\text{on}} + \sum_{t' \in T_{jt}^{\text{supc}}} u_{jt'}^{\text{su}} + \sum_{t' \in T_{jt}^{\text{sdpc}}} u_{jt'}^{\text{sd}}\right) \forall t \in T, j \in J^{\text{pr}}. \tag{146}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{qmax,pr}} = d_t \hat{c}^{\text{qmax,pr}} \max\left(q_{jt} + q_{jt}^{\text{qru}} - q_{jt}^{\text{max}} \left(u_{jt}^{\text{on}} + \sum_{t' \in T_{jt}^{\text{supc}}} u_{jt'}^{\text{su}} + \sum_{t' \in T_{jt}^{\text{sdpc}}} u_{jt'}^{\text{sd}}\right), 0\right)^2. \tag{147}$$

16c. Maximum reactive power reserves (consumer): The original constraint is given by

$$q_{jt} + q_{jt}^{\text{qrd}} \leq q_{jt}^{\text{max}} \left(u_{jt}^{\text{on}} + \sum_{t' \in T_{jt}^{\text{supc}}} u_{jt'}^{\text{su}} + \sum_{t' \in T_{jt}^{\text{sdpc}}} u_{jt'}^{\text{sd}}\right) \quad \forall t \in T, j \in J^{\text{cs}}. \tag{148}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}_{jt}^{\text{qmax,cs}} = d_t \hat{c}^{\text{qmax,cs}} \max\left(q_{jt} + q_{jt}^{\text{qrd}} - q_{jt}^{\text{max}} \left(u_{jt}^{\text{on}} + \sum_{t' \in T_{jt}^{\text{supc}}} u_{jt'}^{\text{su}} + \sum_{t' \in T_{jt}^{\text{sdpc}}} u_{jt'}^{\text{sd}}\right), 0\right)^2. \tag{149}$$

17p. Minimum reactive power reserves (producer): The original constraint is given by

$$q_{jt} - q^{\text{qrd}}_{jt} \geq q^{\min}_{jt}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) \forall t \in T, j \in J^{\text{pr}}. \tag{150}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}^{\text{qmin,pr}}_{jt} = d_t \hat{c}^{\text{qmin,pr}} \max\left(q^{\text{qrd}}_{jt} + q^{\min}_{jt}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) - q_{jt}, 0\right)^2 \tag{151}$$

17c. Minimum reactive power reserves (consumer): The original constraint is given by

$$q_{jt} - q^{\text{qru}}_{jt} \geq q^{\min}_{jt}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) \forall t \in T, j \in J^{\text{cs}} \tag{152}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}^{\text{qmin,cs}}_{jt} = d_t \hat{c}^{\text{qmin,cs}} \max\left(q^{\text{qru}}_{jt} + q^{\min}_{jt}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) - q_{jt}, 0\right)^2. \tag{153}$$

18p. Linked maximum reactive power reserves (producer): The original constraint is given by

$$q_{jt} + q^{\text{qru}}_{jt} \leq q^{\text{max,p0}}_{j}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) + \beta^{\max}_j p_{jt} \tag{154}$$

$$\forall t \in T, j \in J^{\text{pr}} \cap J^{\text{pqmax}}. \tag{155}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}^{\text{qmax,beta,pr}}_{jt} = d_t \hat{c}^{\text{qmax,beta,pr}} \times$$
$$\max\left(q_{jt} + q^{\text{qru}}_{jt} - q^{\text{max,p0}}_{j}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) - \beta^{\max}_j p_{jt}, 0\right)^2. \tag{156}$$

18c. Linked maximum reactive power reserves (consumer): The original constraint is given by

$$q_{jt} + q^{\text{qrd}}_{jt} \leq q^{\text{max,p0}}_{j}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) + \beta^{\max}_j p_{jt}$$
$$\forall t \in T, j \in J^{\text{cs}} \cap J^{\text{pq}\max}. \tag{157}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}^{\text{qmax,beta,cs}}_{jt} = d_t \hat{c}^{\text{qmax,beta,cs}} \times$$
$$\max\left(q_{jt} + q^{\text{qrd}}_{jt} - q^{\text{max,p0}}_{j}\left(u^{\text{on}}_{jt} + \sum_{t'\in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t'\in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'}\right) - \beta^{\max}_j p_{jt}, 0\right)^2. \tag{158}$$

19p. Linked minimum reactive power reserves (producer): The original constraint is given by

$$q_{jt} - q^{\text{qrd}}_{jt} \geq q^{\text{min,p0}}_{j} \left(u^{\text{on}}_{jt} + \sum_{t' \in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t' \in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'} \right) + \beta^{\text{min}}_{j} p_{jt}$$
$$\forall t \in T, j \in J^{\text{pr}} \cap J^{\text{pqmin}}. \tag{159}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}^{\text{qmin,beta,pr}}_{jt} = d_t \hat{c}^{\text{qmin,beta,pr}} \times$$
$$\max\left(q^{\text{min,p0}}_{j} \left(u^{\text{on}}_{jt} + \sum_{t' \in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t' \in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'} \right) + \beta^{\text{min}}_{j} p_{jt} + q^{\text{qrd}}_{jt} - q_{jt}, 0 \right)^2. \tag{160}$$

19c. Linked minimum reactive power reserves (consumer): The original constraint is given by

$$q_{jt} - q^{\text{qru}}_{jt} \geq q^{\text{min,p0}}_{j} \left(u^{\text{on}}_{jt} + \sum_{t' \in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t' \in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'} \right) + \beta^{\text{min}}_{j} p_{jt}$$
$$\forall t \in T, j \in J^{\text{cs}} \cap J^{\text{pqmin}}. \tag{161}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}^{\text{qmin,beta,cs}}_{jt} = d_t \hat{c}^{\text{qmin,beta,cs}} \times$$
$$\max\left(q^{\text{min,p0}}_{j} \left(u^{\text{on}}_{jt} + \sum_{t' \in T^{\text{supc}}_{jt}} u^{\text{su}}_{jt'} + \sum_{t' \in T^{\text{sdpc}}_{jt}} u^{\text{sd}}_{jt'} \right) + \beta^{\text{min}}_{j} p_{jt} + q^{\text{qru}}_{jt} - q_{jt}, 0 \right)^2. \tag{162}$$

20. Maximum starts over multiple intervals: The final penalized constraint is not time indexed; instead, it is indexed by interval w. The original constraint is given by

$$\sum_{t \in T^{\text{su,max}}_{w}} u^{\text{su}}_{jt} \leq u^{\text{su, max}}_{w} \; \forall j \in J^{\text{pr,cs}}, w \in W^{\text{su,max}}_{j}. \tag{163}$$

This is reformulated using a quadratic penalty function:

$$\hat{z}^{\text{mxst}}_{jw} = \hat{c}^{\text{mxst}} \max\left(\sum_{t \in T^{\text{su,max}}_{w}} u^{\text{su}}_{jt} - u^{\text{su, max}}_{w}, 0 \right)^2. \tag{164}$$

Zonal Reserve Requirements

The original zonal requirements are stated as

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} p_{jt}^{\mathrm{rgu}} + p_{nt}^{\mathrm{rgu},+} \geq p_{nt}^{\mathrm{rgu,req}} \forall t \in T, n \in N^{\mathrm{p}} \tag{165}$$

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} p_{jt}^{\mathrm{rgd}} + p_{nt}^{\mathrm{rgd},+} \geq p_{nt}^{\mathrm{rgd,req}} \forall t \in T, n \in N^{\mathrm{p}} \tag{166}$$

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rgu}} + p_{jt}^{\mathrm{scr}}\right) + p_{nt}^{\mathrm{scr},+} \geq p_{nt}^{\mathrm{rgu,req}} + p_{nt}^{\mathrm{scr,req}} \forall t \in T, n \in N^{\mathrm{p}} \tag{167}$$

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rgu}} + p_{jt}^{\mathrm{scr}} + p_{jt}^{\mathrm{nsc}}\right) + p_{nt}^{\mathrm{nsc},+} \geq p_{nt}^{\mathrm{rgu,req}} + p_{nt}^{\mathrm{scr,req}} + p_{nt}^{\mathrm{nsc,req}} \forall t \in T, n \in N^{\mathrm{p}} \tag{168}$$

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rru,on}} + p_{jt}^{\mathrm{rru,off}}\right) + p_{nt}^{\mathrm{rru},+} \geq p_{nt}^{\mathrm{rru,min}} \forall t \in T, n \in N^{\mathrm{p}} \tag{169}$$

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rrd,on}} + p_{jt}^{\mathrm{rrd,off}}\right) + p_{nt}^{\mathrm{rrd},+} \geq p_{nt}^{\mathrm{rrd,min}} \forall t \in T, n \in N^{\mathrm{p}} \tag{170}$$

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} q_{jt}^{\mathrm{qru}} + q_{nt}^{\mathrm{qru},+} \geq q_{nt}^{\mathrm{qru,min}} \forall t \in T, n \in N^{\mathrm{q}} \tag{171}$$

$$\sum_{j \in J_n^{\mathrm{pr,cs}}} q_{jt}^{\mathrm{qrd}} + q_{nt}^{\mathrm{qrd},+} \geq q_{nt}^{\mathrm{qrd,min}} \forall t \in T, n \in N^{\mathrm{q}}. \tag{172}$$

We reformulate these expressions in order to capture the shortfall slack variable explicitly using a max operator. This reformulation hinges on the assumption that there is always downward pressure on the slack variable, since it will incur a penalty in the objective function. For example, consider the following penalized "inequality". In this expression, we assume downward pressure on the slack variable, s^+. Accordingly, the following relations are equivalent:

$$x + s^+ \geq x^{\min} \quad \Leftrightarrow \quad s^+ = \max\left(x^{\min} - x, 0\right).$$

Analogously reformulating the zonal requirements, we have:

$$p_{nt}^{\mathrm{rgu},+} = \max\left(p_{nt}^{\mathrm{rgu,req}} - \sum_{j \in J_n^{\mathrm{pr,cs}}} p_{jt}^{\mathrm{rgu}}, 0\right) \forall t \in T, n \in N^{\mathrm{p}} \tag{173}$$

$$p_{nt}^{\mathrm{rgd},+} = \max\left(p_{nt}^{\mathrm{rgd,req}} - \sum_{j \in J_n^{\mathrm{pr,cs}}} p_{jt}^{\mathrm{rgd}}, 0\right) \forall t \in T, n \in N^{\mathrm{p}} \tag{174}$$

$$p_{nt}^{\mathrm{scr},+} = \max\left(p_{nt}^{\mathrm{rgu,req}} + p_{nt}^{\mathrm{scr,req}} - \sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rgu}} + p_{jt}^{\mathrm{scr}}\right), 0\right) \forall t \in T, n \in N^{\mathrm{p}} \tag{175}$$

$$p_{nt}^{\mathrm{nsc},+} = \max\left(p_{nt}^{\mathrm{rgu,req}} + p_{nt}^{\mathrm{scr,req}} + p_{nt}^{\mathrm{nsc,req}} - \sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rgu}} + p_{jt}^{\mathrm{scr}} + p_{jt}^{\mathrm{nsc}}\right), 0\right) \forall t \in T, n \in N^{\mathrm{p}} \tag{176}$$

$$p_{nt}^{\mathrm{rru},+} = \max\left(p_{nt}^{\mathrm{rru,min}} - \sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rru,on}} + p_{jt}^{\mathrm{rru,off}}\right), 0\right) \forall t \in T, n \in N^{\mathrm{p}} \tag{177}$$

$$p_{nt}^{\mathrm{rrd},+} = \max\left(p_{nt}^{\mathrm{rrd,min}} - \sum_{j \in J_n^{\mathrm{pr,cs}}} \left(p_{jt}^{\mathrm{rrd,on}} + p_{jt}^{\mathrm{rrd,off}}\right), 0\right) \forall t \in T, n \in N^{\mathrm{p}} \tag{178}$$

$$q_{nt}^{\mathrm{qru},+} = \max\left(q_{nt}^{\mathrm{qru,min}} - \sum_{j \in J_n^{\mathrm{pr,cs}}} q_{jt}^{\mathrm{qru}}, 0\right) \forall t \in T, n \in N^{\mathrm{q}} \tag{179}$$

$$q_{nt}^{\text{qrd},+} = \max\left(q_{nt}^{\text{qrd,min}} - \sum_{j \in J_n^{\text{pr,cs}}} q_{jt}^{\text{qrd}}, 0\right) \forall t \in T, n \in N^{\text{q}}. \tag{180}$$

We seek to take the gradients of the reserves requirement expressions. To do so, we start at the top with the shortfall penalties:

$$z_{nt}^{\text{rgu}} = d_t c_n^{\text{rgu}} p_{nt}^{\text{rgu},+} \forall t \in T, n \in N^{\text{p}} \tag{181}$$
$$z_{nt}^{\text{rgd}} = d_t c_n^{\text{rgd}} p_{nt}^{\text{rgd},+} \forall t \in T, n \in N^{\text{p}} \tag{182}$$
$$z_{nt}^{\text{scr}} = d_t c_n^{\text{scr}} p_{nt}^{\text{scr},+} \forall t \in T, n \in N^{\text{p}} \tag{183}$$
$$z_{nt}^{\text{nsc}} = d_t c_n^{\text{nsc}} p_{nt}^{\text{nsc},+} \forall t \in T, n \in N^{\text{p}} \tag{184}$$
$$z_{nt}^{\text{rru}} = d_t c_n^{\text{rru}} p_{nt}^{\text{rru},+} \forall t \in T, n \in N^{\text{p}} \tag{185}$$
$$z_{nt}^{\text{rrd}} = d_t c_n^{\text{rrd}} p_{nt}^{\text{rrd},+} \forall t \in T, n \in N^{\text{p}} \tag{186}$$
$$z_{nt}^{\text{qru}} = d_t c_n^{\text{qru}} q_{nt}^{\text{qru},+} \forall t \in T, n \in N^{\text{q}} \tag{187}$$
$$z_{nt}^{\text{qrd}} = d_t c_n^{\text{qrd}} q_{nt}^{\text{qrd},+} \forall t \in T, n \in N^{\text{q}}. \tag{188}$$

Their gradients (with respect to the reserve shortfall) are given by

$$\nabla_{p_{nt}^{\text{rgu},+}} z_{nt}^{\text{rgu}} = d_t c_n^{\text{rgu}} \forall t \in T, n \in N^{\text{p}} \tag{189}$$
$$\nabla_{p_{nt}^{\text{rgd},+}} z_{nt}^{\text{rgd}} = d_t c_n^{\text{rgd}} \forall t \in T, n \in N^{\text{p}} \tag{190}$$
$$\nabla_{p_{nt}^{\text{scr},+}} z_{nt}^{\text{scr}} = d_t c_n^{\text{scr}} \forall t \in T, n \in N^{\text{p}} \tag{191}$$
$$\nabla_{p_{nt}^{\text{nsc},+}} z_{nt}^{\text{nsc}} = d_t c_n^{\text{nsc}} \forall t \in T, n \in N^{\text{p}} \tag{192}$$
$$\nabla_{p_{nt}^{\text{rru},+}} z_{nt}^{\text{rru}} = d_t c_n^{\text{rru}} \forall t \in T, n \in N^{\text{p}} \tag{193}$$
$$\nabla_{p_{nt}^{\text{rrd},+}} z_{nt}^{\text{rrd}} = d_t c_n^{\text{rrd}} \forall t \in T, n \in N^{\text{p}} \tag{194}$$
$$\nabla_{q_{nt}^{\text{qru},+}} z_{nt}^{\text{qru}} = d_t c_n^{\text{qru}} \forall t \in T, n \in N^{\text{q}} \tag{195}$$
$$\nabla_{q_{nt}^{\text{qrd},+}} z_{nt}^{\text{qrd}} = d_t c_n^{\text{qrd}} \forall t \in T, n \in N^{\text{q}}. \tag{196}$$

Next, we have the "endogenous" reserve requirements:

$$p_{nt}^{\text{rgu,req}} = \sigma_n^{\text{rgu}} \sum_{j \in J_n^{\text{cs}}} p_{jt} \forall t \in T, n \in N^{\text{p}} \tag{197}$$
$$p_{nt}^{\text{rgd,req}} = \sigma_n^{\text{rgd}} \sum_{j \in J_n^{\text{cs}}} p_{jt} \forall t \in T, n \in N^{\text{p}} \tag{198}$$
$$p_{nt}^{\text{scr,req}} = \sigma_n^{\text{scr}} \max_{j \in J_n^{\text{pr}}} p_{jt} \forall t \in T, n \in N^{\text{p}} \tag{199}$$
$$p_{nt}^{\text{nsc,req}} = \sigma_n^{\text{nsc}} \max_{j \in J_n^{\text{pr}}} p_{jt} \forall t \in T, n \in N^{\text{p}}. \tag{200}$$

Their gradients, with respect to active power, are given by

$$\nabla_{p_{jt}} p_{nt}^{\text{rgu,req}} = \sigma_n^{\text{rgu}} \gamma_j,\ \gamma_j = \begin{cases} 1, & j \in J_n^{\text{cs}} \\ 0, & \text{else} \end{cases}, \forall t \in T, n \in N^{\text{p}} \tag{201}$$

$$\nabla_{p_{jt}} p_{nt}^{\text{rgd,req}} = \sigma_n^{\text{rgd}} \gamma_j,\ \gamma_j = \begin{cases} 1, & j \in J_n^{\text{cs}} \\ 0, & \text{else} \end{cases}, \forall t \in T, n \in N^{\text{p}} \tag{202}$$

$$\nabla_{p_{jt}} p_{nt}^{\text{scr,req}} = \sigma_n^{\text{scr}} \gamma_j,\ \gamma_j = \begin{cases} 1, & j \in J_n^{\text{cs}} \wedge p_{jt} = \max_{\text{j} \in J_n^{\text{pr}}} \text{p}_{\text{jt}} \\ 0, & \text{else} \end{cases}, \forall t \in T, n \in N^{\text{p}} \tag{203}$$

$$\nabla_{p_{jt}} p_{nt}^{\text{nsc,req}} = \sigma_n^{\text{nsc}} \gamma_j,\ \gamma_j = \begin{cases} 1, & j \in J_n^{\text{cs}} \wedge p_{jt} = \max_{\text{j} \in J_n^{\text{pr}}} \text{p}_{\text{jt}} \\ 0, & \text{else} \end{cases}, \forall t \in T, n \in N^{\text{p}}. \tag{204}$$

Finally, we have the derivatives of the reserve balance equations themselves. To compute these, we first consider the derivative of a max operator with respect to its summed inputs:

$$f = \alpha \cdot \max(x + y) \tag{205}$$

$$\nabla_{x,y} f = \begin{cases} \begin{bmatrix} 0 \\ 0 \end{bmatrix}, & x + y \leq 0 \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \alpha, & x + y > 0. \end{cases} \tag{206}$$

Analogously, we have the following active power reserve zone derivatives, where all are applied $\forall t \in T, n \in N^{\text{p}}$:

$$\nabla_{p_{nt}^{\text{rgu,req}}} p_{nt}^{\text{rgu},+} = \begin{cases} 0, & p_{nt}^{\text{rgu},+} \leq 0 \\ 1, & \text{else} \end{cases} \tag{207}$$

$$\nabla_{p_{jt}^{\text{rgu}}} p_{nt}^{\text{rgu},+} = \begin{cases} 0, & p_{nt}^{\text{rgu},+} \leq 0 \\ -1, & \text{else} \end{cases} \tag{208}$$

$$\nabla_{p_{nt}^{\text{rgd,req}}} p_{nt}^{\text{rgd},+} = \begin{cases} 0, & p_{nt}^{\text{rgd},+} \leq 0 \\ 1, & \text{else} \end{cases} \tag{209}$$

$$\nabla_{p_{jt}^{\text{rgd}}} p_{nt}^{\text{rgd},+} = \begin{cases} 0, & p_{nt}^{\text{rgd},+} \leq 0 \\ -1, & \text{else} \end{cases} \tag{210}$$

$$\nabla_{p_{nt}^{\text{rgu,req}}} p_{nt}^{\text{scr},+} = \begin{cases} 0, & p_{nt}^{\text{scr},+} \leq 0 \\ 1, & \text{else} \end{cases} \tag{211}$$

$$\nabla_{p_{nt}^{\text{scr,req}}} p_{nt}^{\text{scr},+} = \begin{cases} 0, & p_{nt}^{\text{scr},+} \leq 0 \\ 1, & \text{else} \end{cases} \tag{212}$$

$$\nabla_{p_{jt}^{\text{rgu}}} p_{nt}^{\text{scr},+} = \begin{cases} 0, & p_{nt}^{\text{scr},+} \leq 0 \\ -1, & \text{else} \end{cases} \tag{213}$$

$$\nabla_{p_{jt}^{\text{scr}}} p_{nt}^{\text{scr},+} = \begin{cases} 0, & p_{nt}^{\text{scr},+} \leq 0 \\ -1, & \text{else} \end{cases} \tag{214}$$

$$\vdots \tag{215}$$

We have the following derivatives associated with the reactive reserve zones, where all are applied $\forall t \in T, n \in N^{\text{q}}$:

$$\nabla_{q_{nt}^{\text{qru,min}}} q_{nt}^{\text{qru},+} = \begin{cases} 0, & q_{nt}^{\text{qru},+} \leq 0 \\ 1, & \text{else} \end{cases} \tag{216}$$

$$\nabla_{q_{jt}^{\text{qru}}} q_{nt}^{\text{qru},+} = \begin{cases} 0, & q_{nt}^{\text{qru},+} \leq 0 \\ -1, & \text{else} \end{cases} \tag{217}$$

$$\nabla_{q_{nt}^{\text{qrd,min}}} q_{nt}^{\text{qrd},+} = \begin{cases} 0, & q_{nt}^{\text{qrd},+} \leq 0 \\ 1, & \text{else} \end{cases} \tag{218}$$

$$\nabla_{q_{jt}^{\text{qrd}}} q_{nt}^{\text{qrd},+} = \begin{cases} 0, & q_{nt}^{\text{qrd},+} \leq 0 \\ -1, & \text{else.} \end{cases} \tag{219}$$

Shunt Devices

Injected shunt power is given by

$$p_{jt} = g^{\mathrm{sh}}_{jt} v^2_{it} \forall t \in T, j \in J^{\mathrm{sh}}, i = i_j \tag{220}$$

$$q_{jt} = -b^{\mathrm{sh}}_{jt} v^2_{it} \forall t \in T, j \in J^{\mathrm{sh}}, i = i_j. \tag{221}$$

The associated gradients are given by

$$\nabla_{v_{it}} p_{jt} = 2 g^{\mathrm{sh}}_{jt} v_{it} \forall t \in T, j \in J^{\mathrm{sh}}, i = i_j \tag{222}$$

$$\nabla_{g^{\mathrm{sh}}_{jt}} p_{jt} = v^2_{it} \forall t \in T, j \in J^{\mathrm{sh}}, i = i_j \tag{223}$$

$$\nabla_{v_{it}} q_{jt} = -2 b^{\mathrm{sh}}_{jt} v_{it} \forall t \in T, j \in J^{\mathrm{sh}}, i = i_j \tag{224}$$

$$\nabla_{b^{\mathrm{sh}}_{jt}} q_{jt} = -v^2_{it} \forall t \in T, j \in J^{\mathrm{sh}}, i = i_j. \tag{225}$$

The values of the time-varying shunts are given by

$$g^{\mathrm{sh}}_{jt} = g^{\mathrm{sh}}_{j} u^{\mathrm{sh}}_{jt} \forall t \in T, j \in J^{\mathrm{sh}} \tag{226}$$

$$b^{\mathrm{sh}}_{jt} = b^{\mathrm{sh}}_{j} u^{\mathrm{sh}}_{jt} \forall t \in T, j \in J^{\mathrm{sh}} \tag{227}$$

with associated gradients (in the context of LP relaxation)

$$\nabla_{u^{\mathrm{sh}}_{jt}} g^{\mathrm{sh}}_{jt} = g^{\mathrm{sh}}_{j} \forall t \in T, j \in J^{\mathrm{sh}} \tag{228}$$

$$\nabla_{u^{\mathrm{sh}}_{jt}} b^{\mathrm{sh}}_{jt} = b^{\mathrm{sh}}_{j} \forall t \in T, j \in J^{\mathrm{sh}}. \tag{229}$$

In the LP relaxation, the shunts live in a continuous bounding box:

$$u^{\mathrm{sh,min}}_{j} \leq u^{\mathrm{sh}}_{jt} \leq u^{\mathrm{sh,max}}_{j} \forall t \in T, j \in J^{\mathrm{sh}}. \tag{230}$$

Integrality constraints in the non-relaxed context are given by

$$u^{\mathrm{sh}}_{jt} \in \{\ldots, -1, 0, 1, \ldots\} \forall t \in T, j \in J^{\mathrm{sh}}. \tag{231}$$

N-1 Contingency Costs

We now consider the N-1 contingency costs, which are induced by line overloads. when a line goes out and system reverts to DC power flow model, the line losses in the network are no longer captured. Therefore, we compute the sum of these losses as

$$p_t^{\mathrm{sl}} = \sum_{j \in J^{\mathrm{pr}}} p_{jt} - \sum_{j \in J^{\mathrm{cs}}} p_{jt} - \sum_{j \in J^{\mathrm{sh}}} p_{jt} \forall t \in T. \tag{232}$$

These losses are then distributed across all buses according to the slack parameter α:

$$\sum_{j \in J_k^{J_k^{\mathrm{ac}} \cap J_i^{\mathrm{fr}}}} p_{jtk} - \sum_{j \in J_k^{J_k^{\mathrm{ac}} \cap J_i^{\mathrm{to}}}} p_{jtk} = \sum_{j \in J_i^{\mathrm{pr}}} p_{jt} - \sum_{j \in J_i^{\mathrm{s}}} p_{jt} - \sum_{j \in J_i^{\mathrm{sh}}} p_{jt} \tag{233}$$

$$- \sum_{j \in J_k^{\mathrm{dc}} \cap J_i^{\mathrm{fr}}} p_{jt}^{\mathrm{fr}} - \sum_{j \in J_k^{\mathrm{dc}} \cap J_i^{\mathrm{to}}} p_{jt}^{\mathrm{to}} - \alpha_i \ \mathrm{s}_t^{\mathrm{sl}} \forall t \in T, k \in K, i \in I. \tag{234}$$

For convenience, we define a nodal injection variable:

$$p_{jt}^{\mathrm{inj,DC}} \triangleq \sum_{j \in J_i^{\mathrm{pr}}} p_{jt} - \sum_{j \in J_i^{\mathrm{s}}} p_{jt} - \sum_{j \in J_i^{\mathrm{sh}}} p_{jt} - \sum_{j \in J_k^{\mathrm{dc}} \cap J_i^{\mathrm{fr}}} p_{jt}^{\mathrm{fr}} - \sum_{j \in J_k^{\mathrm{dc}} \cap J_i^{\mathrm{to}}} p_{jt}^{\mathrm{to}} - \alpha_i \ \mathrm{s}_t^{\mathrm{sl}}. \tag{235}$$

Our goal is to solve the linear DC power flow equation, where a flow DC flow is given by

$$p_{jtk} = -b_j^{\mathrm{sr}} u_{jt}^{\mathrm{on}} \left(\theta_{itk} - \theta_{i'tk} - \phi_{jt}\right) \quad \forall t \in T, k \in K, j \in J_k^{\mathrm{ac}}, i = i_j^{\mathrm{fr}}, i' = i_j^{\mathrm{to}} \ . \tag{236}$$

We vectorize these equations by building the directed incidence matrix $E \in \mathbb{R}^{n_l \times n_b}$ which encodes the connections between all AC devices. We also build the diagonal matrix $Y_x \in \mathbb{R}^{n_l \times n_l}$ whose diagonal values correspond to the negative susceptances of all AC components. Notably, we parameterize this matrix with the on/off binary variable $u_t^{\mathrm{on,ac}}$:

$$Y_{x,u} \triangleq Y_x(u_t^{\mathrm{on,ac}}) \tag{237}$$

$$Y_{x,ii} = -b_j^{\mathrm{sr}} \tag{238}$$

where $Y_{x,u,ii} = 0$ if the corresponding binary in $u_t^{\mathrm{on,ac}}$ is 0, and b_j^{sr} is the series susceptance associated with ac device j.

Using these terms, we may build the DC “Ybus” matrix:

$$Y_b = E^T Y_{x,u} E, \tag{239}$$

where $Y_{x,u}E$ computes flows directly. To utilize these structures, we peel apart the DC flow equation into an explicit flow term and a bias term:

$$p_{jtk} = -b_j^{\mathrm{sr}} u_{jt}^{\mathrm{on}} \left(\theta_{itk} - \theta_{i'tk} - \phi_{jt}\right) \tag{240a}$$

$$= \underbrace{-b_j^{\mathrm{sr}} u_{jt}^{\mathrm{on}} \left(\theta_{itk} - \theta_{i'tk}\right)}_{f_{jtk}} + \underbrace{b_j^{\mathrm{sr}} u_{jt}^{\mathrm{on}} \phi_{jt}}_{b_{jt}}. \tag{240b}$$

Next, we “vectorize” by stacking these values into vectors:

$$\boldsymbol{p}_{tk} = \boldsymbol{f}_{tk} + \boldsymbol{b}_t \tag{241}$$

$$\boldsymbol{f}_{tk} = Y_{x,u} E \boldsymbol{\theta}_{tk} \tag{242}$$

$$\boldsymbol{p}_t^{\mathrm{inj}} = E^T \boldsymbol{p}_{tk} \tag{243}$$

$$= E^T \left(\boldsymbol{f}_{tk} + \boldsymbol{b}_t\right) \tag{244}$$

$$= E^T \left(Y_{x,u} E \boldsymbol{\theta}_{tk} + \boldsymbol{b}_t\right) \tag{245}$$

$$= E^T Y_{x,u} E \boldsymbol{\theta}_{tk} + E^T \boldsymbol{b}_t \tag{246}$$

$$= Y_b \boldsymbol{\theta}_{tk} + E^T \boldsymbol{b}_t \tag{247}$$

In order to assess contingency penalty, we consider the original penalty equations:

$$\left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{fr}}\right)^2\right)^{1/2} \leq s_j^{\mathrm{max,ctg}} + s_{jtk}^{+} \forall t \in T, k \in K, j \in J_k^{\mathrm{ac}} \tag{248}$$

$$\left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{to}}\right)^2\right)^{1/2} \leq s_j^{\mathrm{max,ctg}} + s_{jtk}^{+} \forall t \in T, k \in K, j \in J_k^{\mathrm{ac}}. \tag{249}$$

We reformulate these in a standard way, and we compute the associated penalty costs:

$$s_{jtk}^{\mathrm{fr},+} = \left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{fr}}\right)^2\right)^{1/2} - s_j^{\mathrm{max,ctg}} \tag{250}$$

$$s_{jtk}^{\mathrm{to},+} = \left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{to}}\right)^2\right)^{1/2} - s_j^{\mathrm{max,ctg}} \tag{251}$$

$$s_{jtk}^{+} = \max\left(s_{jtk}^{\mathrm{fr},+}, s_{jtk}^{\mathrm{to},+}, 0\right) \tag{252}$$

$$z_{jtk}^{\mathrm{s}} = d_t c^{\mathrm{s}} s_{jtk}^{+} \;\forall t \in T, k \in K, j \in J_k^{\mathrm{ac}}. \tag{253}$$

Note: in the code, we do not define s_{jtk}^{+}. Instead, we directly utilize $z_{jtk}^{\mathrm{s}} = d_t c^{\mathrm{s}} \max\left(s_{jtk}^{\mathrm{fr},+}, s_{jtk}^{\mathrm{to},+}, 0\right) \forall t \in T, j \in J^{\mathrm{ac}}$. Let's take a gradient. Define

$$s_{jtk}^{+} = \max\left(\left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{to}}\right)^2\right)^{1/2} - s_j^{\mathrm{max,ctg}}, \left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{fr}}\right)^2\right)^{1/2} - s_j^{\mathrm{max,ctg}}, 0\right). \tag{254}$$

Next, we define an indicator function function given by

$$\gamma_{\mathrm{fr}} = \begin{cases} 1, & s_{jtk}^{\mathrm{fr},+} > \{s_{jtk}^{\mathrm{to},+} \cup 0\} \\ 0, & \text{else} \end{cases} \tag{255}$$

$$\gamma_{\mathrm{to}} = \begin{cases} 1, & s_{jtk}^{\mathrm{to},+} > \{s_{jtk}^{\mathrm{fr},+} \cup 0\} \\ 0, & \text{else.} \end{cases} \tag{256}$$

Now, the gradients are given by

$$\frac{\partial s_{jtk}^{+}}{\partial p_{jtk}} = \gamma_{\mathrm{fr}} \cdot \frac{1}{2}\left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{fr}}\right)^2\right)^{-1/2} 2p_{jtk} + \gamma_{\mathrm{to}} \cdot \frac{1}{2}\left((p_{jtk})^2 + \left(q_{jt}^{\mathrm{to}}\right)^2\right)^{-1/2} 2p_{jtk} \tag{257}$$

$$= \gamma_{\mathrm{fr}} \cdot \frac{p_{jtk}}{\sqrt{(p_{jtk})^2 + \left(q_{jt}^{\mathrm{fr}}\right)^2}} + \gamma_{\mathrm{to}} \cdot \frac{p_{jtk}}{\sqrt{(p_{jtk})^2 + \left(q_{jt}^{\mathrm{to}}\right)^2}} \tag{258}$$

$$\frac{\partial s_{jtk}^{+}}{\partial q_{jt}^{\mathrm{fr}}} = \gamma_{\mathrm{fr}} \cdot \frac{q_{jt}^{\mathrm{fr}}}{\sqrt{(p_{jtk})^2 + \left(q_{jt}^{\mathrm{fr}}\right)^2}} \tag{259}$$

$$\frac{\partial s_{jtk}^{+}}{\partial q_{jt}^{\mathrm{to}}} = \gamma_{\mathrm{to}} \cdot \frac{q_{jt}^{\mathrm{to}}}{\sqrt{(p_{jtk})^2 + \left(q_{jt}^{\mathrm{to}}\right)^2}}, \tag{260}$$

where q_{jt}^{fr} and q_{jt}^{to} are simply functions of nodal voltages and adjustable transformer variables.

The challenge here is that the vector of contingency phase angles $\boldsymbol{\theta}_{tk}$ is unknown. Ostensibly, we can just solve for them in order to compute flows (singularity issues will be dealt with later):

$$Y_b \boldsymbol{\theta}_{tk} = \boldsymbol{p}_t^{\mathrm{inj}} - E^T \boldsymbol{b}_t \tag{261}$$

$$\boldsymbol{\theta}_{tk} = Y_b^{-1}\left(\boldsymbol{p}_t^{\mathrm{inj}} - E^T \boldsymbol{b}_t\right) \tag{262}$$

and then compute the line flows:

$$\boldsymbol{p}_{tk} = Y_{x,u} E \boldsymbol{\theta}_{tk} \tag{263}$$

$$= Y_{x,u} E Y_b^{-1} \left(\boldsymbol{p}_t^{\text{inj}} - E^T \boldsymbol{b}_t \right). \tag{264}$$

This, of course, is a very bad idea, since explicitly solving this linear system of equations at each step is totally intractable. Instead, to solve this system, we take the following general approach:

1. compute the reduced admittance matrix
2. compute the incomplete cholesky factorization of the reduced matrix (associated with the base case!)
3. use the factorization as a preconditioner for conjugate gradient descent
4. take conjugate gradient steps
5. at each step, update the flows on each contingency network via Woodbury matrix identity
6. use a numerical approximation to estimate the gradient (or, use the incomplete Cholesky factors somehow?)

Taking the gradient: How do we get the gradient associated with this system of equations? We need the sensitivity of the flows with respect to the injections and the transformer phase angles.

Say we have the following linear system and its partials:

$$\boldsymbol{y} = A\boldsymbol{x} \tag{265}$$

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = A. \tag{266}$$

Now, assume we actually want the gradient of some function like $\|\boldsymbol{y}\|_2^2 = \sum y_i^2$. Then, we have

$$\nabla_{\boldsymbol{x}} \|\boldsymbol{y}\|_2^2 = \begin{bmatrix} 2y_1 \frac{\partial y_1}{\partial x_1} + 2y_2 \frac{\partial y_2}{\partial x_1} + \cdots + 2y_n \frac{\partial y_n}{\partial x_1} \\ 2y_1 \frac{\partial y_1}{\partial x_2} + 2y_2 \frac{\partial y_2}{\partial x_2} + \cdots + 2y_n \frac{\partial y_n}{\partial x_2} \\ \vdots \\ 2y_1 \frac{\partial y_1}{\partial x_n} + 2y_2 \frac{\partial y_2}{\partial x_n} + \cdots + 2y_n \frac{\partial y_n}{\partial x_n} \end{bmatrix} \tag{267a}$$

$$= \left(\left(\underbrace{\begin{bmatrix} 2y_1 \\ 2y_2 \\ \vdots \\ 2y_n \end{bmatrix}}_{\boldsymbol{g}} \right)^T A \right)^T \tag{267b}$$

$$= A^T \boldsymbol{g}. \tag{267c}$$

Thus, we need to compute $A^T \boldsymbol{g}$. Now, however, imagine that A represents an inverse matrix:

$$\nabla_{\boldsymbol{x}} \|\boldsymbol{y}\|_2^2 = \boldsymbol{g}^T B^{-1}. \tag{268}$$

In this case, we can observe that

$$\left(B^{-1}\right)^T \boldsymbol{g} = B^{-1} \boldsymbol{g}, \tag{269}$$

since the reduced admittance matrix (in our application) is symmetric, and the inverse of a symmetric matrix is also symmetric (via $BB^{-1} = I \rightarrow I^T = (BB^{-1})^T = (B^{-1})^T B^T = (B^{-1})^T B = B^{-1} B$).

Using these observations, let us compute the gradient of the contingency penalty z_{tk}^{ctg} (for a given time and a given contingency) with respect to nodal injections and transformer phase shifts:

$$\nabla_{z_{jtk}^{\text{s}}} z_{tk}^{\text{ctg}} = - \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n_l} \tag{270}$$

$$\nabla_{s^+_{jtk}} z^{\mathrm{s}}_{jtk} = -\begin{bmatrix} d_t c^{\mathrm{s}} \\ \vdots \\ d_t c^{\mathrm{s}} \end{bmatrix}_{n_l} \tag{271}$$

$$\nabla_{s^{\mathrm{fr},+}_{jtk}} s^+_{jtk} = \begin{bmatrix} 1 \text{ or } 0 \\ \vdots \\ 1 \text{ or } 0 \end{bmatrix}_{n_l} \tag{272}$$

$$\nabla_{s^{\mathrm{to},+}_{jtk}} s^+_{jtk} = \begin{bmatrix} 1 \text{ or } 0 \\ \vdots \\ 1 \text{ or } 0 \end{bmatrix}_{n_l} \tag{273}$$

$$\nabla_{p_{jtk}} s^{\mathrm{to},+}_{jtk} = \begin{bmatrix} \frac{p_{jtk}}{\left((p_{jtk})^2+\left(q^{\mathrm{to}}_{jt}\right)^2\right)^{1/2}} \\ \vdots \\ \frac{p_{jtk}}{\left((p_{jtk})^2+\left(q^{\mathrm{to}}_{jt}\right)^2\right)^{1/2}} \end{bmatrix}_{n_l} \tag{274}$$

$$\nabla_{p_{jtk}} s^{\mathrm{fr},+}_{jtk} = \begin{bmatrix} \frac{p_{jtk}}{\left((p_{jtk})^2+\left(q^{\mathrm{fr}}_{jt}\right)^2\right)^{1/2}} \\ \vdots \\ \frac{p_{jtk}}{\left((p_{jtk})^2+\left(q^{\mathrm{fr}}_{jt}\right)^2\right)^{1/2}} \end{bmatrix}_{n_l}. \tag{275}$$

Thus, from the scalar z^{ctg}_{tk} to the DC flow vector, we have

$$\nabla_{p_{jtk}} z^{\mathrm{ctg}}_{tk} = \underbrace{\nabla_{z^{\mathrm{s}}_{jtk}} z^{\mathrm{ctg}}_{tk} \odot \nabla_{s^+_{jtk}} z^{\mathrm{s}}_{jtk} \odot \left(\nabla_{s^{\mathrm{fr},+}_{jtk}} s^+_{jtk} \odot \nabla_{p_{jtk}} s^{\mathrm{fr},+}_{jtk} + \nabla_{s^{\mathrm{to},+}_{jtk}} s^+_{jtk} \odot \nabla_{p_{jtk}} s^{\mathrm{to},+}_{jtk}\right)}_{\boldsymbol{v}}. \tag{276}$$

This vector, $\boldsymbol{v}$, gives us the derivative of the contingency penalty with respect to the flows. Next, we consider the flow equations:

$$\boldsymbol{p}_{tk} = Y_{x,u} E \boldsymbol{\theta}_{tk} \tag{277a}$$

$$= Y_{x,u} E Y_b^{-1} \left(\boldsymbol{p}^{\mathrm{inj}}_t - E^T \boldsymbol{b}_t\right) \tag{277b}$$

$$= Y_{x,u} E Y_b^{-1} \boldsymbol{p}^{\mathrm{inj}}_t - Y_{x,u} E Y_b^{-1} E^T \boldsymbol{b}_t. \tag{277c}$$

The DC flow gradients are given by

$$\nabla_{\boldsymbol{p}^{\mathrm{inj}}_t} \boldsymbol{p}_{tk} = Y_{x,u} E Y_b^{-1} \tag{278}$$

$$\nabla_{\boldsymbol{b}_t} \boldsymbol{p}_{tk} = -Y_{x,u} E Y_b^{-1} E^T. \tag{279}$$

Therefore, the partial from injection and phase to the contingency penalty is given by

$$\nabla_{p^{\mathrm{inj}}_{jt}} z^{\mathrm{ctg}}_{tk} = \left(Y_{x,u} E Y_b^{-1}\right)^T \boldsymbol{v} \tag{280}$$

$$= Y_b^{-1} E^T Y_{x,u} \boldsymbol{v} \tag{281}$$

$$\nabla_{p^{\mathrm{inj}}_{jt}} z^{\mathrm{ctg}}_{tk} = \left(-Y_{x,u} E Y_b^{-1} E^T\right)^T \boldsymbol{v} \tag{282}$$

$$= -E Y_b^{-1} E^T Y_{x,u} \boldsymbol{v} \tag{283}$$

$$= -E \left(\nabla_{p^{\mathrm{inj}}_{jt}} z^{\mathrm{ctg}}_{tk}\right). \tag{284}$$

From the injection and phase angle partial, we may compute all other partial information (e.g., sensitivity of injection to the unit commitment binaries). So, the fundamental equation we must solve is given by

$$Y_b \underbrace{\nabla_{p_{jt}^{\text{inj}}} z_{tk}^{\text{ctg}}}_{\boldsymbol{x}} = \underbrace{E^T Y_{x,u} \boldsymbol{v}}_{\boldsymbol{b}}. \tag{285}$$

Neglecting singularity, in an ideal world, we would simply take a Cholesky decomposition: $Y_b = LL^T$. Then, we would solve

$$LL^T \boldsymbol{x} = \boldsymbol{b}. \tag{286}$$

However, this is still too expensive to solve at each step. Instead, we take a permuted, incomplete Cholesky decomposition:

$$\hat{L}\hat{L}^T \approx Y_b. \tag{287}$$

We then use $\hat{L}$ and $\hat{L}^T$ as the left and right preconditioners:

$$\begin{aligned}
Y_b \boldsymbol{x} &= \boldsymbol{b} && (288)\\
LL^T \boldsymbol{x} &= \boldsymbol{b} && (289)\\
\hat{L}^{-1} LL^T \boldsymbol{x} &= \hat{L}^{-1} \boldsymbol{b} && (290)\\
\hat{L}^{-1} LL^T \left(\left(\hat{L}^T\right)^{-1} \hat{L}^T\right) \boldsymbol{x} &= \hat{L}^{-1} \boldsymbol{b} && (291)\\
\underbrace{\left(\hat{L}^{-1} LL^T \left(\hat{L}^T\right)^{-1}\right)}_{\hat{Y}_b} \underbrace{\left(\hat{L}^T \boldsymbol{x}\right)}_{\hat{\boldsymbol{x}}} &= \underbrace{\hat{L}^{-1} \boldsymbol{b}}_{\hat{\boldsymbol{b}}}. && (292)
\end{aligned}$$

Thus, we effectively solve the following:

$$\text{approximate with conjugate gradient: } \hat{Y}_b \hat{\boldsymbol{x}} = \hat{\boldsymbol{b}}. \tag{293}$$

Low Rank Updates

Fundamentally, we need to solve

$$Y_b \boldsymbol{\theta}_{tk} = \boldsymbol{p}_t^{\text{inj}} - E^T \boldsymbol{b}_t. \tag{294}$$

This is singular, so instead, we set $\boldsymbol{\theta}_{tk,1} \triangleq 0$, and then we seek to solve the associated reduced system (where hats indicate reduction):

$$\begin{aligned}
\hat{Y}_b \hat{\boldsymbol{\theta}}_{tk} &= \hat{\boldsymbol{p}}_t^{\text{inj}} - \hat{E}^T \boldsymbol{b}_t && (295)\\
\left(\hat{E}^T Y_{x,u} \hat{E}\right) \hat{\boldsymbol{\theta}}_{tk} &= \hat{\boldsymbol{p}}_t^{\text{inj}} - \hat{E}^T \boldsymbol{b}_t. && (296)
\end{aligned}$$

Now, we assume the loss of a line or transformer. This is modeled by adding a negative reactance on the associated line(s) in order to build the reduced contingency matrix $\hat{Y}_{b,k}$:

$$\begin{aligned}
\hat{Y}_b &= \hat{E}^T Y_{x,u} \hat{E} && (297)\\
\hat{Y}_{b,k} &= \hat{E}^T Y_{x,u} \hat{E} + \underbrace{\hat{E}^T Y_k \hat{E}}_{\text{rank1}}. && (298)
\end{aligned}$$

For example, if line 3 is lost in the particular contingency, then

$$Y_k = \begin{bmatrix} 0 & & & & 0 \\ & 0 & & & \\ & & -Y_{x,u,33} & & \\ & & & 0 & \\ 0 & & & & 0 \end{bmatrix}. \tag{299}$$

The low rank update may be written as the sum of outer products via

$$\hat{E}^T Y_k \hat{E} = \sum (\hat{E}_i)^T \beta_i \hat{E}_i, \tag{300}$$

where β_i are the nonzero elements from Y_k. We may now use the WMI to compute the updated solution. We seek to solve

$$\left(\hat{E}^T Y_{x,u} \hat{E} + \sum (\hat{E}_i)^T \beta_i \hat{E}_i\right) \hat{\boldsymbol{\theta}}_{tk} = \underbrace{\hat{\boldsymbol{p}}_t^{\text{inj}} - \hat{E}^T \boldsymbol{b}_{tk}}_{\boldsymbol{c}}, \tag{301}$$

where the "base case" solution $\hat{\boldsymbol{\theta}}_{tk}^{(b)} = \left(\hat{E}^T Y_{x,u} \hat{E}\right)^{-1} \boldsymbol{c}$ is assumed known. Setting $A = \hat{E}^T Y_{x,u} \hat{E}$, we have

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{tk} &= \left(A + \hat{E}_1^T \beta_1 \hat{E}_1 + \hat{E}_2^T \beta_2 \hat{E}_2 + \cdots\right)^{-1} \boldsymbol{c} && (302)\\
&= \left(A + \underbrace{\begin{bmatrix} \hat{E}_1^T & \hat{E}_2^T & \cdots \end{bmatrix}}_{\hat{E}_s} \mathrm{d}(\beta) \underbrace{\begin{bmatrix} \hat{E}_1 \\ \hat{E}_2 \\ \vdots \end{bmatrix}}_{\hat{E}_s^T}\right)^{-1} \boldsymbol{c} && (303)\\
&= A^{-1} - A^{-1} \hat{E}_s \left(\mathrm{d}(\beta)^{-1} + \hat{E}_s^T A^{-1} \hat{E}_s\right)^{-1} \hat{E}_s^T A^{-1} \boldsymbol{c}. && (304)
\end{aligned}$$

Since $A^{-1}\boldsymbol{c}$ is known, the bottleneck is the term $A^{-1}\hat{E}_s$. In the following, assume $\gamma = A^{-1}\hat{E}_s$. Then, we have

$$\hat{\boldsymbol{\theta}}_{tk} = A^{-1}\boldsymbol{c} - \gamma \left(\mathrm{d}(\beta)^{-1} + \hat{E}_s^T \gamma\right)^{-1} \gamma^T \boldsymbol{c}, \tag{305}$$

assuming the number of perturbations away from the base case is fairly small. For a given time, we need to solve

$$\left(\hat{Y}_b + v_k \beta_k v_k^T\right) \hat{\boldsymbol{\theta}}_k = \boldsymbol{c}, \tag{306}$$

where $\hat{\boldsymbol{\theta}}_0 = \hat{Y}_b^{-1}\boldsymbol{c}$ is already known. By WMI, the solution is given by

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= \left(\hat{Y}_b^{-1} - \frac{\hat{Y}_b^{-1} v_k \beta_k v_k^T \hat{Y}_b^{-1}}{1 + v_k^T \hat{Y}_b^{-1} v_k \beta_k}\right) \boldsymbol{c} && (307)\\
&= \hat{Y}_b^{-1}\boldsymbol{c} - \frac{\hat{Y}_b^{-1} v_k \beta_k v_k^T \hat{Y}_b^{-1}}{1 + v_k^T \hat{Y}_b^{-1} v_k \beta_k} \boldsymbol{c}. && (308)
\end{aligned}$$

Next, we define $u_k = \hat{Y}_b^{-1} v_k$, yielding the following simplification:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_k &= \hat{\boldsymbol{\theta}}_0 - u_k \underbrace{\frac{\beta_k u_k^T}{1 + v_k^T u_k \beta_k}}_{w_k^T} \boldsymbol{c} && (309)\\
&= \hat{\boldsymbol{\theta}}_0 - u_k \left(w_k^T \boldsymbol{c}\right), && (310)
\end{aligned}$$

where

$$\begin{aligned}
u_k &= \hat{Y}_b^{-1} v_k && (311)\\
w_k^T &= \frac{\beta_k u_k^T}{1 + v_k^T u_k \beta_k}. && (312)
\end{aligned}$$

Alternatively, we use instead:

$$\hat{\boldsymbol{\theta}}_k = \hat{\boldsymbol{\theta}}_0 - u_k u_k^T \underbrace{\frac{\beta_k}{1 + v_k^T u_k \beta_k}}_{g_k} \boldsymbol{c} \tag{313}$$

$$= \hat{\boldsymbol{\theta}}_0 - u_k \left(g_k \left(u_k^T \boldsymbol{c} \right) \right) \tag{314}$$

Appendix I: Gradient Hierarchies

Chain rule is used to compute the gradient of the objective function z^{ms}.

Back-propagation: Flow Gradients

The back-propagation from variables which influence line flows to the market surplus function is given by

$$\begin{aligned}\nabla_x z^{\mathrm{ms}} =& \nabla_{z^{\mathrm{base}}} z^{\mathrm{ms}} \cdot \nabla_{z_t^{\mathrm{t}}} z^{\mathrm{base}} \cdot \nabla_{z_{jt}^{\mathrm{s}}} z_t^{\mathrm{t}} \cdot \nabla_{s_{jt}^{+}} z_{jt}^{\mathrm{s}} \cdot \nabla_{s_{jt}^{\mathrm{fr/to},+}} s_{jt}^{+} \cdot \nabla_{p/q_{jt}^{\mathrm{fr/to},+}} s_{jt}^{\mathrm{fr/to},+} \\ & \cdot \nabla_x p/q_{jt}^{\mathrm{fr/to},+},\ x \in \{v_{it}, v_{i't}, \theta_{it}, \theta_{i't}, \tau_{jt}, \phi_{jt}, u_{jt}^{\mathrm{on}}\}. \end{aligned} \tag{315}$$

Line Flow

Appendix II: Transformer Model

The current flows through an AC transmission line with shunts and an off-nominal transformer are given according to

$$\begin{bmatrix} \tilde{i}^{\mathrm{fr}} \\ \tilde{i}^{\mathrm{to}} \end{bmatrix} = \begin{bmatrix} \frac{\left(\frac{1}{r^{\mathrm{sr}}+jx^{\mathrm{sr}}} + \frac{jb^{\mathrm{ch}}}{2} + g^{\mathrm{fr,sh}} + jb^{\mathrm{fr,sh}}\right)}{\tau^2} & \frac{-\left(\frac{1}{r^{\mathrm{sr}}+jx^{\mathrm{sr}}}\right)}{\tau e^{-j\phi}} \\ \frac{-\left(\frac{1}{r^{\mathrm{sr}}+jx^{\mathrm{sr}}}\right)}{\tau e^{j\phi}} & \frac{\left(\frac{1}{r^{\mathrm{sr}}+jx^{\mathrm{sr}}} + \frac{jb^{\mathrm{ch}}}{2} + g^{\mathrm{to,sh}} + jb^{\mathrm{to,sh}}\right)}{1} \end{bmatrix} \begin{bmatrix} \tilde{v}^{\mathrm{fr}} \\ \tilde{v}^{\mathrm{to}} \end{bmatrix}. \tag{316}$$

Subsequent power flows may be computed as

$$p^{\mathrm{fr}} + q^{\mathrm{fr}} = \tilde{v}^{\mathrm{fr}} \left(\tilde{i}^{\mathrm{fr}}\right)^* \tag{317}$$

$$s^{\mathrm{fr}} = \sqrt{\left(p^{\mathrm{fr}}\right)^2 + \left(q^{\mathrm{fr}}\right)^2} \tag{318}$$

$$p^{\mathrm{to}} + q^{\mathrm{to}} = \tilde{v}^{\mathrm{to}} \left(\tilde{i}^{\mathrm{to}}\right)^* \tag{319}$$

$$s^{\mathrm{to}} = \sqrt{\left(p^{\mathrm{to}}\right)^2 + \left(q^{\mathrm{to}}\right)^2}. \tag{320}$$