AC Power Flows, Generalized OPF Costs and their Derivatives using Complex Matrix Notation

Ray D. Zimmerman

March 14, 2011*

MATPOWER *Technical Note 2*

© 2008, 2010, 2011 Power Systems Engineering Research Center (PSERC)

All Rights Reserved

*Third revision, with September 25, 2017 minor correction. First draft was February 29, 2008, second February 24, 2010.

Contents

1 Notation

n_b, n_g, n_l	number of buses, generators, branches, respectively
$\|v_i\|, \theta_i$	bus voltage magnitude and angle at bus i
v_i	complex bus voltage at bus i, that is $\|v_i\|e^{j\theta_i}$
$\mathcal{V}, \Theta$	$n_b \times 1$ vectors of bus voltage magnitudes and angles
V	$n_b \times 1$ vector of complex bus voltages v_i
I_{bus}	$n_b \times 1$ vector of complex bus current injections
I^f, I^t	$n_l \times 1$ vectors of complex branch current injections, *from* and *to* ends
S_{bus}	$n_b \times 1$ vector of complex bus power injections
S^f, S^t	$n_l \times 1$ vectors of complex branch power flows, *from* and *to* ends
S_g	$n_g \times 1$ vector of generator complex power injections
P, Q	real and reactive power flows/injections, $S = P + jQ$
M, N	real and imaginary parts of current flows/injections, $I = M + jN$
Y_{bus}	$n_b \times n_b$ system bus admittance matrix
Y_f	$n_l \times n_b$ system branch admittance matrix, *from* end
Y_t	$n_l \times n_b$ system branch admittance matrix, *to* end
C_g	$n_b \times n_g$ generator connection matrix $(i,j)^{th}$ element is 1 if generator j is located at bus i, 0 otherwise
C_f, C_t	$n_l \times n_b$ branch connection matrices, *from* and *to* ends, $(i,j)^{th}$ element is 1 if *from* end, or *to* end, respectively, of branch i is connected to bus j, 0 otherwise
$[A]$	diagonal matrix with vector A on the diagonal
A^T	(non-conjugate) transpose of matrix A
A^*	complex conjugate of A
$\mathbf{1}_n$	$n \times 1$ vector of all ones

2 Introduction

The purpose of this document is to show how the AC power balance and flow equations used in power flow and optimal power flow computations can be expressed in terms of complex matrices, and how their first and second derivatives can be computed efficiently using complex sparse matrix manipulations. Similarly, the derivatives of the generalized AC OPF cost function used by MATPOWER and the corresponding OPF Lagrangian function are developed.

We will be looking at complex functions of the real valued vector

$$X = \begin{bmatrix} \Theta \\ \mathcal{V} \\ P_g \\ Q_g \end{bmatrix} \tag{1}$$

For a complex scalar function $f\colon \mathbb{R}^n \to \mathbb{C}$ of a real vector $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^\mathsf{T}$, we use the following notation for the first derivatives (transpose of the gradient)

$$f_X = \frac{\partial f}{\partial X} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix} \tag{2}$$

The matrix of second partial derivatives, the Hessian of f, is

$$f_{XX} = \frac{\partial^2 f}{\partial X^2} = \frac{\partial}{\partial X}\left(\frac{\partial f}{\partial X}\right)^\mathsf{T} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \tag{3}$$

For a complex vector function $F\colon \mathbb{R}^n \to \mathbb{C}^m$ of a vector X, where

$$F(X) = \begin{bmatrix} f_1(X) & f_2(X) & \cdots & f_m(X) \end{bmatrix}^\mathsf{T} \tag{4}$$

the first derivatives form the Jacobian matrix, where row i is the transpose of the gradient of f_i

$$F_X = \frac{\partial F}{\partial X} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \tag{5}$$

In these derivations, the full 3-dimensional set of second partial derivatives of F will not be computed. Instead a matrix of partial derivatives will be formed by computing the Jacobian of the vector function obtained by multiplying the transpose of the Jacobian of F by a vector λ, using the following notation

$$F_{XX}(\lambda) = \frac{\partial}{\partial X}\left(F_X{}^{\mathsf{T}}\lambda\right) \tag{6}$$

Just to clarify the notation, if Y and Z are subvectors of X, then

$$F_{YZ}(\lambda) = \frac{\partial}{\partial Z}\left(F_Y{}^{\mathsf{T}}\lambda\right) \tag{7}$$

One common operation encountered in these derivations is the element-wise multiplication of a vector A by a vector B to form a new vector C of the same dimension, which can be expressed in either of the following forms

$$C = [A]\,B = [B]\,A \tag{8}$$

It is useful to note that the derivative of such a vector can be calculated by the chain rule as

$$C_X = \frac{\partial C}{\partial X} = [A]\,\frac{\partial B}{\partial X} + [B]\,\frac{\partial A}{\partial X} = [A]\,B_X + [B]\,A_X \tag{9}$$

3 Voltages

3.1 Bus Voltages

V is the $n_b \times 1$ vector of complex bus voltages. The element for bus i is $v_i = |v_i|e^{j\theta_i}$. $\mathcal{V}$ and Θ are the vectors of bus voltage magnitudes and angles. Let

$$E = [\mathcal{V}]^{-1}\,V \tag{10}$$

3.1.1 First Derivatives

$$V_\Theta = \frac{\partial V}{\partial \Theta} \quad = \quad j\,[V] \tag{11}$$

$$V_\mathcal{V} = \frac{\partial V}{\partial \mathcal{V}} \quad = \quad [V]\,[\mathcal{V}]^{-1} = [E] \tag{12}$$

$$E_\Theta = \frac{\partial E}{\partial \Theta} \quad = \quad j\,[E] \tag{13}$$

$$E_\mathcal{V} = \frac{\partial E}{\partial \mathcal{V}} \quad = \quad 0 \tag{14}$$

3.1.2 Second Derivatives

It may be useful in later derivations to note that

$$V_{\mathcal{V}\mathcal{V}}(\lambda) = \frac{\partial}{\partial \mathcal{V}} \left(\frac{\partial V}{\partial \mathcal{V}}^{\mathsf{T}} \lambda \right) = [\lambda]\, E_{\mathcal{V}} = 0 \tag{15}$$

3.2 Branch Voltages

The $n_l \times 1$ vectors of complex voltages at the *from* and *to* ends of all branches are, respectively

$$V_f = C_f V \tag{16}$$
$$V_t = C_t V \tag{17}$$

3.2.1 First Derivatives

$$\frac{\partial V_f}{\partial \Theta} = C_f \frac{\partial V}{\partial \Theta} = jC_f\,[V] \tag{18}$$
$$\frac{\partial V_f}{\partial \mathcal{V}} = C_f \frac{\partial V}{\partial \mathcal{V}} = C_f\,[V]\,[\mathcal{V}]^{-1} = C_f\,[E] \tag{19}$$

4 Bus Injections

4.1 Complex Current Injections

$$I_{\text{bus}} = Y_{\text{bus}} V \tag{20}$$

4.1.1 First Derivatives

$$\frac{\partial I_{\text{bus}}}{\partial X} = \begin{bmatrix} \frac{\partial I_{\text{bus}}}{\partial \Theta} & \frac{\partial I_{\text{bus}}}{\partial \mathcal{V}} & 0 & 0 \end{bmatrix} \tag{21}$$
$$\frac{\partial I_{\text{bus}}}{\partial \Theta} = Y_{\text{bus}} \frac{\partial V}{\partial \Theta} = jY_{\text{bus}}\,[V] \tag{22}$$
$$\frac{\partial I_{\text{bus}}}{\partial \mathcal{V}} = Y_{\text{bus}} \frac{\partial V}{\partial \mathcal{V}} = Y_{\text{bus}}\,[V]\,[\mathcal{V}]^{-1} = Y_{\text{bus}}\,[E] \tag{23}$$

4.2 Complex Power Injections

Consider the complex power balance equation, $G^s(X) = 0$, where

$$G^s(X) = S_{\text{bus}} + S_d - C_g S_g \tag{24}$$

and

$$S_{\text{bus}} = [V]\, {I_{\text{bus}}}^* \tag{25}$$

4.2.1 First Derivatives

$$G^s_X = \frac{\partial G^s}{\partial X} = \begin{bmatrix} G^s_\Theta & G^s_\mathcal{V} & G^s_{P_g} & G^s_{Q_g} \end{bmatrix} \tag{26}$$

$$\begin{aligned}
G^s_\Theta = \frac{\partial S_{\text{bus}}}{\partial \Theta} &= [{I_{\text{bus}}}^*] \frac{\partial V}{\partial \Theta} + [V] \frac{\partial {I_{\text{bus}}}^*}{\partial \Theta} && (27)\\
&= [{I_{\text{bus}}}^*]\, j\, [V] + [V] \left(j Y_{\text{bus}} [V] \right)^* && (28)\\
&= j\, [V] \left([{I_{\text{bus}}}^*] - {Y_{\text{bus}}}^* [V^*] \right) && (29)
\end{aligned}$$

$$\begin{aligned}
G^s_\mathcal{V} = \frac{\partial S_{\text{bus}}}{\partial \mathcal{V}} &= [{I_{\text{bus}}}^*] \frac{\partial V}{\partial \mathcal{V}} + [V] \frac{\partial {I_{\text{bus}}}^*}{\partial \mathcal{V}} && (30)\\
&= [{I_{\text{bus}}}^*]\, [E] + [V]\, {Y_{\text{bus}}}^* [E^*] && (31)\\
&= [V] \left([{I_{\text{bus}}}^*] + {Y_{\text{bus}}}^* [V^*] \right) [\mathcal{V}]^{-1} && (32)
\end{aligned}$$

$$G^s_{P_g} = -C_g \tag{33}$$

$$G^s_{Q_g} = -jC_g \tag{34}$$

4.2.2 Second Derivatives

$$\begin{aligned}
G^s_{XX}(\lambda) &= \frac{\partial}{\partial X} \left({G^s_X}^\mathsf{T} \lambda \right) && (35)\\
&= \begin{bmatrix} G^s_{\Theta\Theta}(\lambda) & G^s_{\Theta\mathcal{V}}(\lambda) & 0 & 0 \\ G^s_{\mathcal{V}\Theta}(\lambda) & G^s_{\mathcal{V}\mathcal{V}}(\lambda) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} && (36)
\end{aligned}$$

$$
\begin{aligned}
G^s_{\Theta\Theta}(\lambda) &= \frac{\partial}{\partial \Theta}\left(G^s_\Theta{}^{\mathsf{T}}\lambda\right) && (37)\\
&= \frac{\partial}{\partial \Theta}\left(j\left([I_{\mathrm{bus}}{}^*] - [V^*]\,Y_{\mathrm{bus}}{}^{*\mathsf{T}}\right)[V]\,\lambda\right) && (38)\\
&= j\frac{\partial}{\partial \Theta}\left([I_{\mathrm{bus}}{}^*]\,[V]\,\lambda - [V^*]\,Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,\lambda\right) && (39)\\
&= j\Bigg([V]\,[\lambda]\,\underbrace{(-jY_{\mathrm{bus}}{}^*\,[V^*])}_{\frac{\partial I_{\mathrm{bus}}{}^*}{\partial\Theta}} + [I_{\mathrm{bus}}{}^*]\,[\lambda]\,\underbrace{(j\,[V])}_{\frac{\partial V}{\partial\Theta}} \\
&\qquad - [V^*]\,Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[\lambda]\,\underbrace{(j\,[V])}_{\frac{\partial V}{\partial\Theta}} - \left[Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,\lambda\right]\underbrace{(-j\,[V^*])}_{\frac{\partial V^*}{\partial\Theta}}\Bigg) && (40)\\
&= [V^*]\left(Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,[\lambda] - \left[Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,\lambda\right]\right)\\
&\qquad + [\lambda]\,[V]\,(Y_{\mathrm{bus}}{}^*\,[V^*] - [I_{\mathrm{bus}}{}^*]) && (41)\\
&= \mathcal{E} + \mathcal{F} && (42)
\end{aligned}
$$

$$
\begin{aligned}
G^s_{\mathcal{V}\Theta}(\lambda) &= \frac{\partial}{\partial \Theta}\left(G^s_{\mathcal{V}}{}^{\mathsf{T}}\lambda\right) && (43)\\
&= \frac{\partial}{\partial \Theta}\left([E]\,[I_{\mathrm{bus}}{}^*]\,\lambda + [E^*]\,Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,\lambda\right) && (44)\\
&= [E]\,[\lambda]\,\underbrace{(-jY_{\mathrm{bus}}{}^*\,[V^*])}_{\frac{\partial I_{\mathrm{bus}}{}^*}{\partial\Theta}} + [I_{\mathrm{bus}}{}^*]\,[\lambda]\,\underbrace{j\,[E]}_{\frac{\partial E}{\partial\Theta}}\\
&\qquad + [E^*]\,Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[\lambda]\,\underbrace{j\,[V]}_{\frac{\partial V}{\partial\Theta}} + \left[Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,\lambda\right]\underbrace{(-j\,[E^*])}_{\frac{\partial E^*}{\partial\Theta}} && (45)\\
&= j\left([E^*]\left(Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,[\lambda] - \left[Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,\lambda\right]\right)\right.\\
&\qquad \left. - [\lambda]\,[E]\,(Y_{\mathrm{bus}}{}^*\,[V^*] - [I_{\mathrm{bus}}{}^*])\right) && (46)\\
&= j\,[\mathcal{V}]^{-1}\left([V^*]\left(Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,[\lambda] - \left[Y_{\mathrm{bus}}{}^{*\mathsf{T}}\,[V]\,\lambda\right]\right)\right.\\
&\qquad \left. - [\lambda]\,[V]\,(Y_{\mathrm{bus}}{}^*\,[V^*] - [I_{\mathrm{bus}}{}^*])\right) && (47)\\
&= j\mathcal{G}(\mathcal{E} - \mathcal{F}) && (48)
\end{aligned}
$$

$$
\begin{aligned}
G^s_{\Theta\mathcal{V}}(\lambda) &= \frac{\partial}{\partial \mathcal{V}}\left(G^{s\,\mathsf{T}}_{\Theta}\lambda\right) && (49)\\
&= j\left(\left([\lambda]\,[V]\,{Y_{\mathrm{bus}}}^* - \left[{Y_{\mathrm{bus}}}^{*\mathsf{T}}\,[V]\,\lambda\right]\right)[V^*]\right.\\
&\quad \left. - \left([V^*]\,{Y_{\mathrm{bus}}}^{*\mathsf{T}} - [{I_{\mathrm{bus}}}^*]\right)[V]\,[\lambda]\right)[\mathcal{V}]^{-1} && (50)\\
&= {G^s_{\mathcal{V}\Theta}}^{\mathsf{T}}(\lambda) && (51)
\end{aligned}
$$

$$
\begin{aligned}
G^s_{\mathcal{V}\mathcal{V}}(\lambda) &= \frac{\partial}{\partial \mathcal{V}}\left(G^{s\,\mathsf{T}}_{\mathcal{V}}\lambda\right) && (52)\\
&= \frac{\partial}{\partial \mathcal{V}}\left([E]\,[{I_{\mathrm{bus}}}^*]\,\lambda + [E^*]\,{Y_{\mathrm{bus}}}^{*\mathsf{T}}\,[V]\,\lambda\right) && (53)\\
&= [E]\,[\lambda]\,\underbrace{{Y_{\mathrm{bus}}}^*\,[E^*]}_{\frac{\partial {I_{\mathrm{bus}}}^*}{\partial \mathcal{V}}} + [{I_{\mathrm{bus}}}^*]\,[\lambda]\,\underbrace{0}_{\frac{\partial E}{\partial \mathcal{V}}}\\
&\quad + [E^*]\,{Y_{\mathrm{bus}}}^{*\mathsf{T}}\,[\lambda]\,\underbrace{[E]}_{\frac{\partial V}{\partial \mathcal{V}}} + \left[{Y_{\mathrm{bus}}}^{*\mathsf{T}}\,[V]\,\lambda\right]\underbrace{0}_{\frac{\partial E^*}{\partial \mathcal{V}}} && (54)\\
&= [\mathcal{V}]^{-1}\left([\lambda]\,[V]\,{Y_{\mathrm{bus}}}^*\,[V^*] + [V^*]\,{Y_{\mathrm{bus}}}^{*\mathsf{T}}\,[V]\,[\lambda]\right)[\mathcal{V}]^{-1} && (55)\\
&= \mathcal{G}(\mathcal{C} + \mathcal{C}^{\mathsf{T}})\mathcal{G} && (56)
\end{aligned}
$$

Computational savings can be achieved by storing and reusing certain intermediate terms during the computation of these second derivatives, as follows:

$$
\begin{aligned}
\mathcal{A} &= [\lambda]\,[V] && (57)\\
\mathcal{B} &= Y_{\mathrm{bus}}\,[V] && (58)\\
\mathcal{C} &= \mathcal{A}\mathcal{B}^* && (59)\\
\mathcal{D} &= {Y_{\mathrm{bus}}}^{*\mathsf{T}}\,[V] && (60)\\
\mathcal{E} &= [V^*]\,(\mathcal{D}\,[\lambda] - [\mathcal{D}\lambda]) && (61)\\
\mathcal{F} &= \mathcal{C} - \mathcal{A}\,[{I_{\mathrm{bus}}}^*] = j\,[\lambda]\,G^s_{\Theta} && (62)\\
\mathcal{G} &= [\mathcal{V}]^{-1} && (63)\\
G^s_{\Theta\Theta}(\lambda) &= \mathcal{E} + \mathcal{F} && (64)\\
G^s_{\mathcal{V}\Theta}(\lambda) &= j\mathcal{G}(\mathcal{E} - \mathcal{F}) && (65)\\
G^s_{\Theta\mathcal{V}}(\lambda) &= {G^s_{\mathcal{V}\Theta}}^{\mathsf{T}}(\lambda) && (66)\\
G^s_{\mathcal{V}\mathcal{V}}(\lambda) &= \mathcal{G}(\mathcal{C} + \mathcal{C}^{\mathsf{T}})\mathcal{G} && (67)
\end{aligned}
$$

5 Branch Flows

Consider the line flow constraints of the form $H(X) < 0$. This section examines 3 variations based on the square of the magnitude of the current, apparent power and real power, respectively. The relationships are derived first for the complex flows at the *from* ends of the branches. Derivations for the *to* end are identical (i.e. just replace all f sub/super-scripts with t).

5.1 Complex Currents

$$I^f = Y_f V \tag{68}$$
$$I^t = Y_t V \tag{69}$$

5.1.1 First Derivatives

$$I_X^f = \frac{\partial I^f}{\partial X} = \left[\begin{array}{cccc} I_\Theta^f & I_\mathcal{V}^f & I_{P_g}^f & I_{Q_g}^f \end{array} \right] \tag{70}$$

$$I_\Theta^f = Y_f \left(\frac{\partial V}{\partial \Theta} \right) = jY_f \left[V \right] \tag{71}$$
$$I_\mathcal{V}^f = Y_f \left(\frac{\partial V}{\partial \mathcal{V}} \right) = Y_f \left[V \right] \left[\mathcal{V} \right]^{-1} = Y_f \left[E \right] \tag{72}$$
$$I_{P_g}^f = 0 \tag{73}$$
$$I_{Q_g}^f = 0 \tag{74}$$

5.1.2 Second Derivatives

$$I_{XX}^f(\mu) = \frac{\partial}{\partial X} \left({I_X^f}^\mathsf{T} \mu \right) \tag{75}$$
$$= \left[\begin{array}{cccc} I_{\Theta\Theta}^f(\mu) & I_{\Theta\mathcal{V}}^f(\mu) & 0 & 0 \\ I_{\mathcal{V}\Theta}^f(\mu) & I_{\mathcal{V}\mathcal{V}}^f(\mu) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right] \tag{76}$$

$$I_{\Theta\Theta}^f(\mu) = \frac{\partial}{\partial \Theta} \left({I_\Theta^f}^\mathsf{T} \mu \right) \tag{77}$$

$$
\begin{aligned}
&= \frac{\partial}{\partial \Theta}\left(j\,[V]\,{Y_f}^{\mathsf{T}}\mu\right) && (78)\\
&= -\left[{Y_f}^{\mathsf{T}}\mu\right][V] && (79)
\end{aligned}
$$

$$
\begin{aligned}
I^f_{\mathcal{V}\Theta}(\mu) &= \frac{\partial}{\partial \Theta}\left({I^f_{\mathcal{V}}}^{\mathsf{T}}\mu\right) && (80)\\
&= \frac{\partial}{\partial \Theta}\left([E]\,{Y_f}^{\mathsf{T}}\mu\right) && (81)\\
&= j\left[{Y_f}^{\mathsf{T}}\mu\right][E] && (82)\\
&= -jI^f_{\Theta\Theta}(\mu)\,[\mathcal{V}]^{-1} && (83)
\end{aligned}
$$

$$
\begin{aligned}
I^f_{\Theta\mathcal{V}}(\mu) &= \frac{\partial}{\partial \mathcal{V}}\left({I^f_{\Theta}}^{\mathsf{T}}\mu\right) && (84)\\
&= \frac{\partial}{\partial \mathcal{V}}\left(j\,[V]\,{Y_f}^{\mathsf{T}}\mu\right) && (85)\\
&= j\left[{Y_f}^{\mathsf{T}}\mu\right][E] && (86)\\
&= I^f_{\mathcal{V}\Theta}(\mu) && (87)
\end{aligned}
$$

$$
\begin{aligned}
I^f_{\mathcal{V}\mathcal{V}}(\mu) &= \frac{\partial}{\partial \mathcal{V}}\left({I^f_{\mathcal{V}}}^{\mathsf{T}}\mu\right) && (88)\\
&= \frac{\partial}{\partial \mathcal{V}}\left([E]\,{Y_f}^{\mathsf{T}}\mu\right) && (89)\\
&= 0 && (90)
\end{aligned}
$$

5.2 Complex Power Flows

$$
\begin{aligned}
S^f &= [V_f]\,{I^f}^* && (91)\\
S^t &= [V_t]\,{I^t}^* && (92)
\end{aligned}
$$

5.2.1 First Derivatives

$$
\begin{aligned}
S^f_X = \frac{\partial S^f}{\partial X} &= \begin{bmatrix} S^f_\Theta & S^f_{\mathcal{V}} & S^f_{P_g} & S^f_{Q_g} \end{bmatrix} && (93)\\
&= \left[{I^f}^*\right]\frac{\partial V_f}{\partial X} + [V_f]\,\frac{\partial {I^f}^*}{\partial X} && (94)
\end{aligned}
$$

$$
\begin{aligned}
S^f_\Theta &= \left[I^{f*}\right] \frac{\partial V_f}{\partial \Theta} + [V_f] \frac{\partial I^{f*}}{\partial \Theta} && (95)\\
&= \left[I^{f*}\right] jC_f [V] + [C_f V] (jY_f [V])^* && (96)\\
&= j(\left[I^{f*}\right] C_f [V] - [C_f V] Y_f{}^* [V^*]) && (97)
\end{aligned}
$$

$$
\begin{aligned}
S^f_\mathcal{V} &= \left[I^{f*}\right] \frac{\partial V_f}{\partial \mathcal{V}} + [V_f] \frac{\partial I^{f*}}{\partial \mathcal{V}} && (98)\\
&= \left[I^{f*}\right] C_f [E] + [C_f V] Y_f{}^* [E^*] && (99)
\end{aligned}
$$

$$
S^f_{P_g} = 0 \quad (100)
$$

$$
S^f_{Q_g} = 0 \quad (101)
$$

5.2.2 Second Derivatives

$$
\begin{aligned}
S^f_{XX}(\mu) &= \frac{\partial}{\partial X}\left(S^f_X{}^\mathsf{T} \mu\right) && (102)\\
&= \begin{bmatrix} S^f_{\Theta\Theta}(\mu) & S^f_{\Theta\mathcal{V}}(\mu) & 0 & 0 \\ S^f_{\mathcal{V}\Theta}(\mu) & S^f_{\mathcal{V}\mathcal{V}}(\mu) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} && (103)
\end{aligned}
$$

$$
\begin{aligned}
S^f_{\Theta\Theta}(\mu) &= \frac{\partial}{\partial \Theta}\left(S^f_\Theta{}^\mathsf{T} \mu\right) && (104)\\
&= \frac{\partial}{\partial \Theta}\left(j\left([V] C_f{}^\mathsf{T} \left[I^{f*}\right] - [V^*] Y_f{}^{*\mathsf{T}} [C_f V]\right)\mu\right) && (105)\\
&= j\frac{\partial}{\partial \Theta}\left([V] C_f{}^\mathsf{T} \left[I^{f*}\right]\mu - [V^*] Y_f{}^{*\mathsf{T}} [C_f V]\mu\right) && (106)\\
&= j\Bigg([V] C_f{}^\mathsf{T} [\mu] \underbrace{(-jY_f{}^* [V^*])}_{\frac{\partial I^{f*}}{\partial \Theta}} + \left[C_f{}^\mathsf{T} \left[I^{f*}\right]\mu\right] \underbrace{j[V]}_{\frac{\partial V}{\partial \Theta}}
\end{aligned}
$$

$$
\left. - [V^*] \, {Y_f^*}^{\mathsf{T}} [\mu] \, C_f \underbrace{j [V]}_{\frac{\partial V}{\partial \Theta}} - \left[{Y_f^*}^{\mathsf{T}} [C_f V] \, \mu \right] \underbrace{(-j [V^*])}_{\frac{\partial V^*}{\partial \Theta}} \right) \tag{107}
$$

$$
\begin{aligned}
&= [V^*] \, {Y_f^*}^{\mathsf{T}} [\mu] \, C_f [V] + [V] \, {C_f}^{\mathsf{T}} [\mu] \, Y_f^* [V^*] \\
&\quad - \left[{Y_f^*}^{\mathsf{T}} [\mu] \, C_f V \right] [V^*] - \left[{C_f}^{\mathsf{T}} [\mu] \, Y_f^* V^* \right] [V] && (108) \\
&= \mathcal{F}_f - \mathcal{D}_f - \mathcal{E}_f && (109)
\end{aligned}
$$

$$
\begin{aligned}
S^f_{\mathcal{V}\Theta}(\mu) &= \frac{\partial}{\partial \Theta} \left({S^f_{\mathcal{V}}}^{\mathsf{T}} \mu \right) && (110) \\
&= \frac{\partial}{\partial \Theta} \left([E] \, {C_f}^{\mathsf{T}} \left[I^{f^*} \right] \mu + [E^*] \, {Y_f^*}^{\mathsf{T}} [C_f V] \, \mu \right) && (111) \\
&= [E] \, {C_f}^{\mathsf{T}} [\mu] \underbrace{(-j Y_f^* [V^*])}_{\frac{\partial I^{f^*}}{\partial \Theta}} + \left[{C_f}^{\mathsf{T}} \left[I^{f^*} \right] \mu \right] \underbrace{j [E]}_{\frac{\partial E}{\partial \Theta}} \\
&\quad + [E^*] \, {Y_f^*}^{\mathsf{T}} [\mu] \, C_f \underbrace{j [V]}_{\frac{\partial V}{\partial \Theta}} + \left[{Y_f^*}^{\mathsf{T}} [C_f V] \, \mu \right] \underbrace{(-j [E^*])}_{\frac{\partial E^*}{\partial \Theta}} && (112) \\
&= j \Big([E^*] \, {Y_f^*}^{\mathsf{T}} [\mu] \, C_f [V] - [E] \, {C_f}^{\mathsf{T}} [\mu] \, Y_f^* [V^*] \\
&\quad - \left[{Y_f^*}^{\mathsf{T}} [\mu] \, C_f V \right] [E^*] + \left[{C_f}^{\mathsf{T}} [\mu] \, Y_f^* V^* \right] [E] \Big) && (113) \\
&= j [\mathcal{V}]^{-1} \Big([V^*] \, {Y_f^*}^{\mathsf{T}} [\mu] \, C_f [V] - [V] \, {C_f}^{\mathsf{T}} [\mu] \, Y_f^* [V^*] \\
&\quad - \left[{Y_f^*}^{\mathsf{T}} [\mu] \, C_f V \right] [V^*] + \left[{C_f}^{\mathsf{T}} [\mu] \, Y_f^* V^* \right] [V] \Big) && (114) \\
&= j \mathcal{G} (\mathcal{B}_f - {\mathcal{B}_f}^{\mathsf{T}} - \mathcal{D}_f + \mathcal{E}_f) && (115)
\end{aligned}
$$

$$
\begin{aligned}
S^f_{\Theta\mathcal{V}}(\mu) &= \frac{\partial}{\partial \mathcal{V}} \left({S^f_{\Theta}}^{\mathsf{T}} \mu \right) && (116) \\
&= j \Big([V] \, {C_f}^{\mathsf{T}} [\mu] \, Y_f^* [V^*] - [V^*] \, {Y_f^*}^{\mathsf{T}} [\mu] \, C_f [V] \\
&\quad - \left[{Y_f^*}^{\mathsf{T}} [\mu] \, C_f V \right] [V^*] + \left[{C_f}^{\mathsf{T}} [\mu] \, Y_f^* V^* \right] [V] \Big) [\mathcal{V}]^{-1} && (117) \\
&= {S^f_{\mathcal{V}\Theta}}^{\mathsf{T}}(\mu) && (118)
\end{aligned}
$$

$$
S^f_{\mathcal{V}\mathcal{V}}(\mu) = \frac{\partial}{\partial \mathcal{V}} \left({S^f_{\mathcal{V}}}^{\mathsf{T}} \mu \right) \tag{119}
$$

$$
\begin{aligned}
&= \frac{\partial}{\partial \mathcal{V}} \left([E]\, {C_f}^{\mathsf{T}} \left[I^{f*} \right] \mu + [E^*]\, {Y_f}^{*\mathsf{T}} [C_f V]\, \mu \right) && (120) \\
&= [E]\, {C_f}^{\mathsf{T}} [\mu] \underbrace{{Y_f}^* [E^*]}_{\frac{\partial I^{f*}}{\partial \mathcal{V}}} + \left[{C_f}^{\mathsf{T}} \left[I^{f*} \right] \mu \right] \underbrace{0}_{\frac{\partial E}{\partial \mathcal{V}}} \\
&\quad + [E^*]\, {Y_f}^{*\mathsf{T}} [\mu]\, C_f \underbrace{[E]}_{\frac{\partial V}{\partial \mathcal{V}}} + \left[{Y_f}^{*\mathsf{T}} [C_f V]\, \mu \right] \underbrace{0}_{\frac{\partial E^*}{\partial \mathcal{V}}} && (121) \\
&= [\mathcal{V}]^{-1} \left([V^*]\, {Y_f}^{*\mathsf{T}} [\mu]\, C_f [V] + [V]\, {C_f}^{\mathsf{T}} [\mu]\, {Y_f}^* [V^*] \right) [\mathcal{V}]^{-1} && (122) \\
&= \mathcal{G} \mathcal{F}_f \mathcal{G} && (123)
\end{aligned}
$$

Computational savings can be achieved by storing and reusing certain intermediate terms during the computation of these second derivatives, as follows:

$$
\begin{aligned}
\mathcal{A}_f &= {Y_f}^{*\mathsf{T}} [\mu]\, C_f && (124) \\
\mathcal{B}_f &= [V^*]\, \mathcal{A}_f [V] && (125) \\
\mathcal{D}_f &= [\mathcal{A}_f V]\, [V^*] && (126) \\
\mathcal{E}_f &= \left[{\mathcal{A}_f}^{\mathsf{T}} V^* \right] [V] && (127) \\
\mathcal{F}_f &= \mathcal{B}_f + {\mathcal{B}_f}^{\mathsf{T}} && (128) \\
\mathcal{G} &= [\mathcal{V}]^{-1} && (129) \\
S^f_{\Theta\Theta}(\mu) &= \mathcal{F}_f - \mathcal{D}_f - \mathcal{E}_f && (130) \\
S^f_{\mathcal{V}\Theta}(\mu) &= j\mathcal{G}(\mathcal{B}_f - {\mathcal{B}_f}^{\mathsf{T}} - \mathcal{D}_f + \mathcal{E}_f) && (131) \\
S^f_{\Theta\mathcal{V}}(\mu) &= {S^f_{\mathcal{V}\Theta}}^{\mathsf{T}}(\mu) && (132) \\
S^f_{\mathcal{V}\mathcal{V}}(\mu) &= \mathcal{G} \mathcal{F}_f \mathcal{G} && (133)
\end{aligned}
$$

5.3 Squared Current Magnitudes

Let $I^2_{\max}$ denote the vector of the squares of the current magnitude limits. Then the flow constraint function $H(X)$ can be defined in terms of the square of the current magnitudes as follows:

$$
\begin{aligned}
H^f(X) &= \left[I^{f*} \right] I^f - I^2_{\max} && (134) \\
&= \left[M^f \right] M^f + \left[N^f \right] N^f - I^2_{\max} && (135)
\end{aligned}
$$

where $I^f = M^f + jN^f$.

5.3.1 First Derivatives

$$
\begin{aligned}
H_X^f &= \left[I^{f*}\right] I_X^f + \left[I^f\right] {I_X^f}^* && (136)\\
&= \left[I^{f*}\right] I_X^f + \left(\left[I^{f*}\right] I_X^f\right)^* && (137)\\
&= 2 \cdot \Re\left\{\left[I^{f*}\right] I_X^f\right\} && (138)\\
&= \left[M^f - jN^f\right](M_X^f + jN_X^f) + \left[M^f + jN^f\right](M_X^f - jN_X^f) && (139)\\
&= 2\left(\left[M^f\right] M_X^f + \left[N^f\right] N_X^f\right) && (140)\\
&= 2\left(\Re\left\{\left[I^f\right]\right\}\Re\left\{I_X^f\right\} + \Im\left\{\left[I^f\right]\right\}\Im\left\{I_X^f\right\}\right) && (141)
\end{aligned}
$$

5.3.2 Second Derivatives

$$
\begin{aligned}
H_{XX}^f(\mu) &= \frac{\partial}{\partial X}\left({H_X^f}^\mathsf{T} \mu\right) && (142)\\
&= \begin{bmatrix} H_{\Theta\Theta}^f(\mu) & H_{\Theta\mathcal{V}}^f(\mu) & 0 & 0 \\ H_{\mathcal{V}\Theta}^f(\mu) & H_{\mathcal{V}\mathcal{V}}^f(\mu) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} && (143)
\end{aligned}
$$

$$
\begin{aligned}
H_{XX}^f(\mu) &= \frac{\partial}{\partial X}\left({H_X^f}^\mathsf{T} \mu\right) && (144)\\
&= \frac{\partial}{\partial X}\left({I_X^f}^\mathsf{T}\left[I^{f*}\right]\mu + {I_X^f}^{*\mathsf{T}}\left[I^f\right]\mu\right) && (145)\\
&= I_{XX}^f(\left[I^{f*}\right]\mu) + {I_X^f}^\mathsf{T}\left[\mu\right]{I_X^f}^* + {I_{XX}^f}^*(\left[I^f\right]\mu) + {I_X^f}^{*\mathsf{T}}\left[\mu\right] I_X^f && (146)\\
&= 2 \cdot \Re\left\{I_{XX}^f(\left[I^{f*}\right]\mu) + {I_X^f}^\mathsf{T}\left[\mu\right]{I_X^f}^*\right\} && (147)
\end{aligned}
$$

$$
\begin{aligned}
H_{\Theta\Theta}^f(\mu) &= 2 \cdot \Re\left\{I_{\Theta\Theta}^f(\left[I^{f*}\right]\mu) + {I_\Theta^f}^\mathsf{T}\left[\mu\right]{I_\Theta^f}^*\right\} && (148)\\
H_{\mathcal{V}\Theta}^f(\mu) &= 2 \cdot \Re\left\{I_{\mathcal{V}\Theta}^f(\left[I^{f*}\right]\mu) + {I_\mathcal{V}^f}^\mathsf{T}\left[\mu\right]{I_\Theta^f}^*\right\} && (149)\\
H_{\Theta\mathcal{V}}^f(\mu) &= 2 \cdot \Re\left\{I_{\Theta\mathcal{V}}^f(\left[I^{f*}\right]\mu) + {I_\Theta^f}^\mathsf{T}\left[\mu\right]{I_\mathcal{V}^f}^*\right\} && (150)\\
H_{\mathcal{V}\mathcal{V}}^f(\mu) &= 2 \cdot \Re\left\{I_{\mathcal{V}\mathcal{V}}^f(\left[I^{f*}\right]\mu) + {I_\mathcal{V}^f}^\mathsf{T}\left[\mu\right]{I_\mathcal{V}^f}^*\right\} && (151)
\end{aligned}
$$

5.4 Squared Apparent Power Magnitudes

Let $S^2_{\max}$ denote the vector of the squares of the apparent power flow limits. Then the flow constraint function $H(X)$ can be defined in terms of the square of the apparent power flows as follows:

$$\begin{aligned} H^f(X) &= \left[S^{f*}\right] S^f - S^2_{\max} && (152) \\ &= \left[P^f\right] P^f + \left[Q^f\right] Q^f - S^2_{\max} && (153) \end{aligned}$$

where $S^f = P^f + jQ^f$.

5.4.1 First Derivatives

$$\begin{aligned} H^f_X &= \left[S^{f*}\right] S^f_X + \left[S^f\right] {S^f_X}^* && (154) \\ &= \left[S^{f*}\right] S^f_X + \left(\left[S^{f*}\right] S^f_X\right)^* && (155) \\ &= 2 \cdot \Re\left\{\left[S^{f*}\right] S^f_X\right\} && (156) \\ &= \left[P^f - jQ^f\right] (P^f_X + jQ^f_X) + \left[P^f + jQ^f\right] (P^f_X - jQ^f_X) && (157) \\ &= 2\left(\left[P^f\right] P^f_X + \left[Q^f\right] Q^f_X\right) && (158) \\ &= 2\left(\Re\left\{\left[S^f\right]\right\} \Re\left\{S^f_X\right\} + \Im\left\{\left[S^f\right]\right\} \Im\left\{S^f_X\right\}\right) && (159) \end{aligned}$$

5.4.2 Second Derivatives

$$\begin{aligned} H^f_{XX}(\mu) &= \frac{\partial}{\partial X}\left({H^f_X}^{\mathsf{T}} \mu\right) && (160) \\ &= \begin{bmatrix} H^f_{\Theta\Theta}(\mu) & H^f_{\Theta\mathcal{V}}(\mu) & 0 & 0 \\ H^f_{\mathcal{V}\Theta}(\mu) & H^f_{\mathcal{V}\mathcal{V}}(\mu) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} && (161) \end{aligned}$$

$$\begin{aligned} H^f_{XX}(\mu) &= \frac{\partial}{\partial X}\left({H^f_X}^{\mathsf{T}} \mu\right) && (162) \\ &= \frac{\partial}{\partial X}\left({S^f_X}^{\mathsf{T}} \left[S^{f*}\right] \mu + {S^f_X}^{*\mathsf{T}} \left[S^f\right] \mu\right) && (163) \\ &= S^f_{XX}(\left[S^{f*}\right] \mu) + {S^f_X}^{\mathsf{T}} \left[\mu\right] {S^f_X}^* + {S^f_{XX}}^*(\left[S^f\right] \mu) + {S^f_X}^{*\mathsf{T}} \left[\mu\right] S^f_X && (164) \\ &= 2 \cdot \Re\left\{S^f_{XX}(\left[S^{f*}\right] \mu) + {S^f_X}^{\mathsf{T}} \left[\mu\right] {S^f_X}^*\right\} && (165) \end{aligned}$$

$$
\begin{aligned}
H^f_{\Theta\Theta}(\mu) &= 2 \cdot \Re\left\{ S^f_{\Theta\Theta}(\left[S^{f*}\right]\mu) + {S^f_{\Theta}}^{\mathsf{T}}\left[\mu\right] {S^f_{\Theta}}^{*} \right\} && (166)\\
H^f_{\mathcal{V}\Theta}(\mu) &= 2 \cdot \Re\left\{ S^f_{\mathcal{V}\Theta}(\left[S^{f*}\right]\mu) + {S^f_{\mathcal{V}}}^{\mathsf{T}}\left[\mu\right] {S^f_{\Theta}}^{*} \right\} && (167)\\
H^f_{\Theta\mathcal{V}}(\mu) &= 2 \cdot \Re\left\{ S^f_{\Theta\mathcal{V}}(\left[S^{f*}\right]\mu) + {S^f_{\Theta}}^{\mathsf{T}}\left[\mu\right] {S^f_{\mathcal{V}}}^{*} \right\} && (168)\\
H^f_{\mathcal{V}\mathcal{V}}(\mu) &= 2 \cdot \Re\left\{ S^f_{\mathcal{V}\mathcal{V}}(\left[S^{f*}\right]\mu) + {S^f_{\mathcal{V}}}^{\mathsf{T}}\left[\mu\right] {S^f_{\mathcal{V}}}^{*} \right\} && (169)
\end{aligned}
$$

5.5 Squared Real Power Magnitudes

Let $P^2_{\max}$ denote the vector of the squares of the real power flow limits. Then the flow constraint function $H(X)$ can be defined in terms of the square of the real power flows as follows:

$$
\begin{aligned}
H^f(X) &= \left[\Re\left\{S^f\right\}\right] \Re\left\{S^f\right\} - P^2_{\max} && (170)\\
&= \left[P^f\right] P^f - P^2_{\max} && (171)
\end{aligned}
$$

5.5.1 First Derivatives

$$
\begin{aligned}
H^f_X &= 2\left[P^f\right] P^f_X && (172)\\
&= 2\left(\Re\left\{\left[S^f\right]\right\} \Re\left\{S^f_X\right\}\right) && (173)
\end{aligned}
$$

5.5.2 Second Derivatives

$$
\begin{aligned}
H^f_{XX}(\mu) &= \frac{\partial}{\partial X}\left({H^f_X}^{\mathsf{T}}\mu\right) && (174)\\
&= \left[\begin{array}{cccc} H^f_{\Theta\Theta}(\mu) & H^f_{\Theta\mathcal{V}}(\mu) & 0 & 0 \\ H^f_{\mathcal{V}\Theta}(\mu) & H^f_{\mathcal{V}\mathcal{V}}(\mu) & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] && (175)
\end{aligned}
$$

$$
\begin{aligned}
H^f_{XX}(\mu) &= \frac{\partial}{\partial X}\left({H^f_X}^{\mathsf{T}}\mu\right) && (176)\\
&= \frac{\partial}{\partial X}\left(2{P^f_X}^{\mathsf{T}}\left[P^f\right]\mu\right) && (177)
\end{aligned}
$$

$$= 2\left(P^f_{XX}(\left[P^f\right]\mu) + {P^f_X}^\mathsf{T}\left[\mu\right]P^f_X\right) \tag{178}$$

$$= 2\left(\Re\left\{S^f_{XX}(\left[\Re\left\{S^f\right\}\right]\mu)\right\} + \Re\left\{{S^f_X}^\mathsf{T}\right\}\left[\mu\right]\Re\left\{S^f_X\right\}\right) \tag{179}$$

$$H^f_{\Theta\Theta}(\mu) = 2\left(\Re\left\{S^f_{\Theta\Theta}(\left[\Re\left\{S^f\right\}\right]\mu)\right\} + \Re\left\{{S^f_\Theta}^\mathsf{T}\right\}\left[\mu\right]\Re\left\{S^f_\Theta\right\}\right) \tag{180}$$

$$H^f_{\mathcal{V}\Theta}(\mu) = 2\left(\Re\left\{S^f_{\mathcal{V}\Theta}(\left[\Re\left\{S^f\right\}\right]\mu)\right\} + \Re\left\{{S^f_\mathcal{V}}^\mathsf{T}\right\}\left[\mu\right]\Re\left\{S^f_\Theta\right\}\right) \tag{181}$$

$$H^f_{\Theta\mathcal{V}}(\mu) = 2\left(\Re\left\{S^f_{\Theta\mathcal{V}}(\left[\Re\left\{S^f\right\}\right]\mu)\right\} + \Re\left\{{S^f_\Theta}^\mathsf{T}\right\}\left[\mu\right]\Re\left\{S^f_\mathcal{V}\right\}\right) \tag{182}$$

$$H^f_{\mathcal{V}\mathcal{V}}(\mu) = 2\left(\Re\left\{S^f_{\mathcal{V}\mathcal{V}}(\left[\Re\left\{S^f\right\}\right]\mu)\right\} + \Re\left\{{S^f_\mathcal{V}}^\mathsf{T}\right\}\left[\mu\right]\Re\left\{S^f_\mathcal{V}\right\}\right) \tag{183}$$

6 Generalized AC OPF Costs

The generalized cost function for the AC OPF consists of three parts,

$$f(X) = f^a(X) + f^b(X) + f^c(X) \tag{184}$$

expressed as functions of the full set of optimization variables.

$$X = \begin{bmatrix} \Theta \\ \mathcal{V} \\ P_g \\ Q_g \\ Y \\ Z \end{bmatrix} \tag{185}$$

where Y is the $n_y \times 1$ vector of cost variables associated with piecewise linear generator costs and Z is an $n_z \times 1$ vector of additional linearly constrained user variables.

6.1 Polynomial Generator Costs

Let $f^i_P(p^i_g)$ and $f^i_Q(q^i_g)$ be polynomial cost functions for real and reactive power for generator i and F^P and F^Q be the $n_g \times 1$ vectors of these costs.

$$F^P(P_g) = \begin{bmatrix} f^1_P(p^1_g) \\ \vdots \\ f^{n_g}_P(p^{n_g}_g) \end{bmatrix} \tag{186}$$

$$F^Q(Q_g) = \begin{bmatrix} f_Q^1(q_g^1) \\ \vdots \\ f_Q^{n_g}(q_g^{n_g}) \end{bmatrix} \tag{187}$$

$$f^a(X) = \mathbf{1}_{n_g}^\mathsf{T} \left(F^P(P_g) + F^Q(Q_g) \right) \tag{188}$$

6.1.1 First Derivatives

We will use $F^{P'}$ and $F^{P''}$ to represent the vectors of first and second derivatives of each of these real power cost functions with respect to the corresponding generator output. Likewise for the reactive power costs.

$$f_X^a = \frac{\partial f^a}{\partial X} \tag{189}$$

$$= \begin{bmatrix} f_\Theta^a & f_\mathcal{V}^a & f_{P_g}^a & f_{Q_g}^a & f_Y^a & f_Z^a \end{bmatrix} \tag{190}$$

$$= \begin{bmatrix} 0 & 0 & (F^{P'})^\mathsf{T} & (F^{Q'})^\mathsf{T} & 0 & 0 \end{bmatrix} \tag{191}$$

6.1.2 Second Derivatives

$$f_{XX}^a = \frac{\partial {f_X^a}^\mathsf{T}}{\partial X} \tag{192}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & f_{P_gP_g}^a & 0 & 0 & 0 \\ 0 & 0 & 0 & f_{Q_gQ_g}^a & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{193}$$

where

$$f_{P_gP_g}^a = \left[F^{P''} \right] \tag{194}$$

$$f_{Q_gQ_g}^a = \left[F^{Q''} \right] \tag{195}$$

6.2 Piecewise Linear Generator Costs

$$f^b(X) = \mathbf{1}_{n_y}^\mathsf{T} Y \tag{196}$$

6.2.1 First Derivatives

$$
\begin{aligned}
f_X^b &= \frac{\partial f^b}{\partial X} && (197)\\
&= \begin{bmatrix} f_\Theta^b & f_\mathcal{V}^b & f_{P_g}^b & f_{Q_g}^b & f_Y^b & f_Z^b \end{bmatrix} && (198)\\
&= \begin{bmatrix} 0 & 0 & 0 & 0 & \mathbf{1}_{n_y}^\mathsf{T} & 0 \end{bmatrix} && (199)
\end{aligned}
$$

6.2.2 Second Derivatives

$$
f_{XX}^b = 0 \tag{200}
$$

6.3 General Cost Term

Let the general cost be defined in terms of the $n_w \times n_w$ matrix H^w and $n_w \times 1$ vector C^w of coefficients and the parameters specified in the $n_w \times n_x$ matrix $\mathcal{N}$ and the $n_w \times 1$ vectors D, $\widehat{R}$, K, and $\mathcal{M}$. The parameters $\mathcal{N}$ and $\widehat{R}$ provide a linear transformation and shift to the full set of optimization variables X, resulting in a new set of variables R.

$$
R = \mathcal{N} X - \widehat{R} \tag{201}
$$

Each element of K specifies the size of a "dead zone" in which the cost is zero for the corresponding element of R. The elements k_i are used to define $n_w \times 1$ vectors U, $\bar{K}$ and $\bar{R}$, where

$$
\begin{aligned}
u_i &= \begin{cases} 0, & -k_i \le r_i \le k_i \\ 1, & \text{otherwise} \end{cases} && (202)\\
\bar{k}_i &= \begin{cases} k_i, & r_i < -k_i \\ 0, & -k_i \le r_i \le k_i \\ -k_i, & r_i > k_i \end{cases} && (203)
\end{aligned}
$$

The "dead zone" costs are zeroed by multiplying by $[U]$. The remaining elements are shifted toward zero by the size of the "dead zone" by adding $\bar{K}$, before applying a cost.

$$
\bar{R} = R + \bar{K} \tag{204}
$$

Each element of D specifies whether to apply a linear or quadratic function to the corresponding element of $\bar{R}$. This can be done via two more $n_w \times 1$ vectors, D^L

and D^Q, defined as follows

$$d_i^L = \begin{cases} 1, & d_i = 1 \\ 0, & \text{otherwise} \end{cases} \tag{205}$$

$$d_i^Q = \begin{cases} 1, & d_i = 2 \\ 0, & \text{otherwise} \end{cases} \tag{206}$$

The result is scaled by the corresponding element of $\mathcal{M}$ to form a new $n_w \times 1$ vector

$$\begin{aligned} W &= [\mathcal{M}]\,[U]\left(\left[D^L\right] + \left[D^Q\right]\left[\bar{R}\right]\right)\bar{R} && (207) \\ &= \left(\mathcal{D}_\mathcal{L} + \mathcal{D}_\mathcal{Q}\left[\bar{R}\right]\right)\bar{R} && (208) \end{aligned}$$

where

$$\begin{aligned} \mathcal{D}_\mathcal{L} &= [\mathcal{M}]\,[U]\left[D^L\right] && (209) \\ \mathcal{D}_\mathcal{Q} &= [\mathcal{M}]\,[U]\left[D^Q\right] && (210) \end{aligned}$$

The full general cost term is then expressed as a quadratic function of W as follows

$$f^c(X) = \frac{1}{2} W^\mathsf{T} H^w W + C^{w\mathsf{T}} W \tag{211}$$

6.3.1 First Derivatives

For simplicity of derivation and computation, we defined $\mathcal{A}$ and $\mathcal{B}$ as follows

$$\mathcal{A} = W_{\bar{R}} = \frac{\partial W}{\partial \bar{R}} = \mathcal{D}_\mathcal{L} + 2\,\mathcal{D}_\mathcal{Q}\left[\bar{R}\right] \tag{212}$$

$$\mathcal{B} = f_W^c = \frac{\partial f^c}{\partial W} = W^\mathsf{T} H^w + C^{w\mathsf{T}} \tag{213}$$

$$\bar{R}_X = \frac{\partial \bar{R}}{\partial X} = \mathcal{N} \tag{214}$$

$$\begin{aligned} W_X = \frac{\partial W}{\partial X} &= W_{\bar{R}} \cdot \bar{R}_X && (215) \\ &= \mathcal{A}\mathcal{N} && (216) \end{aligned}$$

$$\begin{aligned} f_X^c = \frac{\partial f^c}{\partial X} &= f_W^c \cdot W_X && (217) \\ &= \mathcal{B}\mathcal{A}\mathcal{N} && (218) \end{aligned}$$

6.3.2 Second Derivatives

$$
\begin{aligned}
f^c_{XX} &= \frac{\partial}{\partial X}\left({f^c_X}^\mathsf{T}\right) && (219)\\
&= \frac{\partial}{\partial X}\left(\mathcal{N}^\mathsf{T}\mathcal{A}\mathcal{B}^\mathsf{T}\right) && (220)\\
&= \mathcal{N}^\mathsf{T}\left(\mathcal{A}\frac{\partial}{\partial X}\left(H^w W + C^w\right) + 2\,\mathcal{D}_\mathcal{Q}\left[\mathcal{B}^\mathsf{T}\right]\frac{\partial \bar{R}}{\partial X}\right) && (221)\\
&= \mathcal{N}^\mathsf{T}\left(\mathcal{A}H^w W_X + 2\,\mathcal{D}_\mathcal{Q}\left[\mathcal{B}^\mathsf{T}\right]\bar{R}_X\right) && (222)\\
&= \mathcal{N}^\mathsf{T}\left(\mathcal{A}H^w\mathcal{A} + 2\,\mathcal{D}_\mathcal{Q}\left[\mathcal{B}^\mathsf{T}\right]\right)\mathcal{N} && (223)
\end{aligned}
$$

6.4 Full Cost Function

$$
\begin{aligned}
f(X) &= f^a(X) + f^b(X) + f^c(X) && (224)\\
&= \mathbf{1}^\mathsf{T}_{n_g}\left(F^P(P_g) + F^Q(Q_g)\right) + \mathbf{1}^\mathsf{T}_{n_y} Y + \frac{1}{2}W^\mathsf{T} H^w W + {C^w}^\mathsf{T} W && (225)
\end{aligned}
$$

6.4.1 First Derivatives

$$
\begin{aligned}
f_X = \frac{\partial f}{\partial X} &= f^a_X + f^b_X + f^c_X && (226)\\
&= \left[\begin{array}{cccccc} 0 & 0 & \left(F^{P\prime}\right)^\mathsf{T} & \left(F^{Q\prime}\right)^\mathsf{T} & \mathbf{1}^\mathsf{T}_{n_y} & 0 \end{array}\right] + \mathcal{B}\mathcal{A}\mathcal{N} && (227)
\end{aligned}
$$

6.4.2 Second Derivatives

$$
\begin{aligned}
f_{XX} = \frac{\partial^2 f}{\partial X^2} &= f^a_{XX} + f^b_{XX} + f^c_{XX} && (228)\\
&= \left[\begin{array}{cccccc}
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & \left[F^{P\prime\prime}\right] & 0 & 0 & 0\\
0 & 0 & 0 & \left[F^{Q\prime\prime}\right] & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]\\
&\quad + \mathcal{N}^\mathsf{T}\left(\mathcal{A}H^w\mathcal{A} + 2\,\mathcal{D}_\mathcal{Q}\left[\mathcal{B}^\mathsf{T}\right]\right)\mathcal{N} && (229)
\end{aligned}
$$

7 Lagrangian of the AC OPF

Consider the following AC OPF problem formulation, where X is defined as in (185), f is the generalized cost function described above, and $\mathcal{X}$ represents the reduced form of X, consisting of only Θ, $\mathcal{V}$, P_g and Q_g, without Y and Z.

$$\min_X f(X) \tag{230}$$

subject to

$$G(X) = 0 \tag{231}$$
$$H(X) \leq 0 \tag{232}$$

where

$$G(X) = \begin{bmatrix} \Re\{G^s(\mathcal{X})\} \\ \Im\{G^s(\mathcal{X})\} \\ A_E X - B_E \end{bmatrix} \tag{233}$$

and

$$H(X) = \begin{bmatrix} H^f(\mathcal{X}) \\ H^t(\mathcal{X}) \\ A_I X - B_I \end{bmatrix} \tag{234}$$

Partitioning the corresponding multipliers λ and μ similarly,

$$\lambda = \begin{bmatrix} \lambda_P \\ \lambda_Q \\ \lambda_E \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_f \\ \mu_t \\ \mu_I \end{bmatrix} \tag{235}$$

the Lagrangian for this problem can be written as

$$\mathcal{L}(X, \lambda, \mu) = f(X) + \lambda^\mathsf{T} G(X) + \mu^\mathsf{T} H(X) \tag{236}$$

7.1 First Derivatives

$$\mathcal{L}_X(X, \lambda, \mu) = f_X + \lambda^\mathsf{T} G_X + \mu^\mathsf{T} H_X \tag{237}$$
$$\mathcal{L}_\lambda(X, \lambda, \mu) = G^\mathsf{T}(X) \tag{238}$$
$$\mathcal{L}_\mu(X, \lambda, \mu) = H^\mathsf{T}(X) \tag{239}$$

where

$$G_X = \begin{bmatrix} \Re\{G^s_\mathcal{X}\} & 0 & 0 \\ \Im\{G^s_\mathcal{X}\} & 0 & 0 \\ & A_E & \end{bmatrix} = \begin{bmatrix} \Re\{G^s_\Theta\} & \Re\{G^s_\mathcal{V}\} & -C_g & 0 & 0 & 0 \\ \Im\{G^s_\Theta\} & \Im\{G^s_\mathcal{V}\} & 0 & -C_g & 0 & 0 \\ & & A_E & & & \end{bmatrix} \tag{240}$$

and

$$H_X = \begin{bmatrix} H^f_{\mathcal{X}} & 0 & 0 \\ H^t_{\mathcal{X}} & 0 & 0 \\ & A_I & \end{bmatrix} = \begin{bmatrix} H^f_{\Theta} & H^f_{\mathcal{V}} & 0 & 0 & 0 & 0 \\ H^t_{\Theta} & H^t_{\mathcal{V}} & 0 & 0 & 0 & 0 \\ & & A_I & & & \end{bmatrix} \tag{241}$$

7.2 Second Derivatives

$$\mathcal{L}_{XX}(X, \lambda, \mu) = f_{XX} + G_{XX}(\lambda) + H_{XX}(\mu) \tag{242}$$

where

$$G_{XX}(\lambda) = \begin{bmatrix} \Re\{G^s_{\mathcal{X}\mathcal{X}}(\lambda_P)\} + \Im\{G^s_{\mathcal{X}\mathcal{X}}(\lambda_Q)\} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{243}$$

$$= \Re\left\{\begin{bmatrix} G^s_{\Theta\Theta}(\lambda_P) & G^s_{\Theta\mathcal{V}}(\lambda_P) & 0 & 0 & 0 & 0 \\ G^s_{\mathcal{V}\Theta}(\lambda_P) & G^s_{\mathcal{V}\mathcal{V}}(\lambda_P) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right\}$$

$$+ \Im\left\{\begin{bmatrix} G^s_{\Theta\Theta}(\lambda_Q) & G^s_{\Theta\mathcal{V}}(\lambda_Q) & 0 & 0 & 0 & 0 \\ G^s_{\mathcal{V}\Theta}(\lambda_Q) & G^s_{\mathcal{V}\mathcal{V}}(\lambda_Q) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right\} \tag{244}$$

and

$$H_{XX}(\mu) = \begin{bmatrix} H^f_{\mathcal{X}\mathcal{X}}(\mu_f) + H^t_{\mathcal{X}\mathcal{X}}(\mu_t) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{245}$$

$$= \begin{bmatrix} H^f_{\Theta\Theta}(\mu_f) + H^t_{\Theta\Theta}(\mu_t) & H^f_{\Theta\mathcal{V}}(\mu_f) + H^t_{\Theta\mathcal{V}}(\mu_t) & 0 & 0 & 0 & 0 \\ H^f_{\mathcal{V}\Theta}(\mu_f) + H^t_{\mathcal{V}\Theta}(\mu_t) & H^f_{\mathcal{V}\mathcal{V}}(\mu_f) + H^t_{\mathcal{V}\mathcal{V}}(\mu_t) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{246}$$