Calibration Cost Functions and Derivatives

Ruby Byrne

last updated April 2024

1 Per-Frequency, Per-Polarization, and Per-Time Calibration

The simplest calibration implementation is not fully polarized. We calibrate the *pp* and *qq* visibilities separately and fit the crosspol phase only from the cross-polarized visibilities *pq* and *qp*. We also exclude the autocorrelations in calibration, assume a single time step, and calibrate each frequency independently.

1.1 Cost Function

For a single polarization, time step, and frequency, calibration consists of minimizing the chi-squared quantity

$$\chi^2(\boldsymbol{g}) = \sum_{jk} \frac{1}{\sigma_{jk}^2} \left| m_{jk} - g_j g_k^* v_{jk} \right|^2 . \tag{1}$$

1.2 Jacobian Calculation

Expanding the chi-squared gives

$$\begin{aligned} \chi^2(\boldsymbol{g}) &= \sum_{jk} \frac{1}{\sigma_{jk}^2} \left[|m_{jk}|^2 + |g_j|^2 |g_k|^2 |v_{jk}|^2 - m_{jk}^* g_j g_k^* v_{jk} - m_{jk} g_j^* g_k v_{jk}^* \right] \\ &= \sum_{jk} \frac{1}{\sigma_{jk}^2} \left[|m_{jk}|^2 + |g_j|^2 |g_k|^2 |v_{jk}|^2 - 2\,\mathrm{Re}\left(m_{jk}^* g_j g_k^* v_{jk} \right) \right] . \end{aligned} \tag{2}$$

Now taking derivatives and omitting autocorrelations (assuming $\frac{1}{\sigma_{jk}^2} = 0$ for $j = k$) gives

$$\begin{aligned} \frac{\partial \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Re}(g_a)} = \frac{\partial}{\partial\,\mathrm{Re}(g_a)} \sum_{j \neq a} \Bigg(& \frac{1}{\sigma_{aj}^2} \left[|m_{aj}|^2 + |g_a|^2 |g_j|^2 |v_{aj}|^2 - 2\,\mathrm{Re}\left(m_{aj}^* g_a g_j^* v_{aj} \right) \right] \\ & + \frac{1}{\sigma_{ja}^2} \left[|m_{ja}|^2 + |g_j|^2 |g_a|^2 |v_{ja}|^2 - 2\,\mathrm{Re}\left(m_{ja}^* g_j g_a^* v_{ja} \right) \right] \Bigg) . \end{aligned} \tag{3}$$

Evaluating the derivative gives

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)} = \sum_{j\neq a} \Bigg(&\frac{1}{\sigma_{aj}^2} \left[2\operatorname{Re}(g_a)|g_j|^2|v_{aj}|^2 - 2\operatorname{Re}\left(m_{aj}^* g_j^* v_{aj} \right) \right] \\ &+ \frac{1}{\sigma_{ja}^2} \left[2\operatorname{Re}(g_a)|g_j|^2|v_{ja}|^2 - 2\operatorname{Re}\left(m_{ja}^* g_j v_{ja} \right) \right] \Bigg). \end{aligned} \tag{4}$$

The derivative with respect to the imaginary component is

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a)} = \sum_{j\neq a} \Bigg(&\frac{1}{\sigma_{aj}^2} \left[2\operatorname{Im}(g_a)|g_j|^2|v_{aj}|^2 + 2\operatorname{Im}\left(m_{aj}^* g_j^* v_{aj} \right) \right] \\ &+ \frac{1}{\sigma_{ja}^2} \left[2\operatorname{Im}(g_a)|g_j|^2|v_{ja}|^2 - 2\operatorname{Im}\left(m_{ja}^* g_j v_{ja} \right) \right] \Bigg). \end{aligned} \tag{5}$$

We can rewrite this as

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)} = 2\operatorname{Re} \sum_{j\neq a} \Bigg(&\frac{1}{\sigma_{aj}^2} \left[g_a|g_j|^2|v_{aj}|^2 - m_{aj} g_j v_{aj}^* \right] \\ &+ \frac{1}{\sigma_{ja}^2} \left[g_a|g_j|^2|v_{ja}|^2 - m_{ja}^* g_j v_{ja} \right] \Bigg) \end{aligned} \tag{6}$$

and

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a)} = 2\operatorname{Im} \sum_{j\neq a} \Bigg(&\frac{1}{\sigma_{aj}^2} \left[g_a|g_j|^2|v_{aj}|^2 - m_{aj} g_j v_{aj}^* \right] \\ &+ \frac{1}{\sigma_{ja}^2} \left[g_a|g_j|^2|v_{ja}|^2 - m_{ja}^* g_j v_{ja} \right] \Bigg), \end{aligned} \tag{7}$$

or alternatively as

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)} = 2\operatorname{Re} \sum_{j\neq a} \Bigg(&\frac{1}{\sigma_{aj}^2} g_j v_{aj}^* \left[g_a g_j^* v_{aj} - m_{aj} \right] \\ &+ \frac{1}{\sigma_{ja}^2} g_j v_{ja} \left[g_j^* g_a v_{ja}^* - m_{ja}^* \right] \Bigg) \end{aligned} \tag{8}$$

and

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a)} = 2\operatorname{Im} \sum_{j\neq a} \Bigg(&\frac{1}{\sigma_{aj}^2} g_j v_{aj}^* \left[g_a g_j^* v_{aj} - m_{aj} \right] \\ &+ \frac{1}{\sigma_{ja}^2} g_j v_{ja} \left[g_j^* g_a v_{ja}^* - m_{ja}^* \right] \Bigg). \end{aligned} \tag{9}$$

1.3 Hessian Calculation

Taking second derivatives, we get that

$$\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}^2(g_a)} = \frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial \operatorname{Im}^2(g_a)} = 2\sum_{j\neq a}\left(\frac{1}{\sigma_{aj}^2}|g_j|^2|v_{aj}|^2 + \frac{1}{\sigma_{ja}^2}|g_j|^2|v_{ja}|^2\right) \tag{10}$$

and

$$\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)\partial \operatorname{Im}(g_a)} = \frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a)\partial \operatorname{Re}(g_a)} = 0. \tag{11}$$

For $b \neq a$, we get

$$\begin{aligned}
\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)\operatorname{Re}(g_b)} =& 2\frac{\partial}{\partial \operatorname{Re}(g_b)} \operatorname{Re}\sum_{j\neq a}\left(\frac{1}{\sigma_{aj}^2}\left[g_a|g_j|^2|v_{aj}|^2 - m_{aj}g_jv_{aj}^*\right]\right. \\
&\left. + \frac{1}{\sigma_{ja}^2}\left[g_a|g_j|^2|v_{ja}|^2 - m_{ja}^*g_jv_{ja}\right]\right) \\
=& 2\frac{\partial}{\partial \operatorname{Re}(g_b)}\left(\frac{1}{\sigma_{ab}^2}\left[\operatorname{Re}(g_a)|g_b|^2|v_{ab}|^2 - \operatorname{Re}(m_{ab}g_bv_{ab}^*)\right]\right. \\
&\left. + \frac{1}{\sigma_{ba}^2}\left[\operatorname{Re}(g_a)|g_b|^2|v_{ba}|^2 - \operatorname{Re}(m_{ba}^*g_bv_{ba})\right]\right) \\
=& 2\left(\frac{1}{\sigma_{ab}^2}\left[2\operatorname{Re}(g_a)\operatorname{Re}(g_b)|v_{ab}|^2 - \operatorname{Re}(m_{ab}v_{ab}^*)\right]\right. \\
&\left. + \frac{1}{\sigma_{ba}^2}\left[2\operatorname{Re}(g_a)\operatorname{Re}(g_b)|v_{ba}|^2 - \operatorname{Re}(m_{ba}^*v_{ba})\right]\right),
\end{aligned} \tag{12}$$

$$\begin{aligned}
\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)\operatorname{Im}(g_b)} =& 2\frac{\partial}{\partial \operatorname{Im}(g_b)} \operatorname{Re}\sum_{j\neq a}\left(\frac{1}{\sigma_{aj}^2}\left[g_a|g_j|^2|v_{aj}|^2 - m_{aj}g_jv_{aj}^*\right]\right. \\
&\left. + \frac{1}{\sigma_{ja}^2}\left[g_a|g_j|^2|v_{ja}|^2 - m_{ja}^*g_jv_{ja}\right]\right) \\
=& 2\frac{\partial}{\partial \operatorname{Im}(g_b)}\left(\frac{1}{\sigma_{ab}^2}\left[\operatorname{Re}(g_a)|g_b|^2|v_{ab}|^2 - \operatorname{Re}(m_{ab}g_bv_{ab}^*)\right]\right. \\
&\left. + \frac{1}{\sigma_{ba}^2}\left[\operatorname{Re}(g_a)|g_b|^2|v_{ba}|^2 - \operatorname{Re}(m_{ba}^*g_bv_{ba})\right]\right) \\
=& 2\left(\frac{1}{\sigma_{ab}^2}\left[2\operatorname{Re}(g_a)\operatorname{Im}(g_b)|v_{ab}|^2 + \operatorname{Im}(m_{ab}v_{ab}^*)\right]\right. \\
&\left. + \frac{1}{\sigma_{ba}^2}\left[2\operatorname{Re}(g_a)\operatorname{Im}(g_b)|v_{ba}|^2 + \operatorname{Im}(m_{ba}^*v_{ba})\right]\right),
\end{aligned} \tag{13}$$

and

$$
\begin{aligned}
\frac{\partial \chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a) \operatorname{Im}(g_b)} =& 2 \frac{\partial}{\partial \operatorname{Im}(g_b)} \operatorname{Im} \sum_{j \neq a} \left(\frac{1}{\sigma_{aj}^2} \left[g_a |g_j|^2 |v_{aj}|^2 - m_{aj} g_j v_{aj}^* \right] \right. \\
&\left. + \frac{1}{\sigma_{ja}^2} \left[g_a |g_j|^2 |v_{ja}|^2 - m_{ja}^* g_j v_{ja} \right] \right) \\
=& 2 \frac{\partial}{\partial \operatorname{Im}(g_b)} \left(\frac{1}{\sigma_{ab}^2} \left[\operatorname{Im}(g_a) |g_b|^2 |v_{ab}|^2 - \operatorname{Im}(m_{ab} g_b v_{ab}^*) \right] \right. \\
&\left. + \frac{1}{\sigma_{ba}^2} \left[\operatorname{Im}(g_a) |g_b|^2 |v_{ba}|^2 - \operatorname{Im}(m_{ba}^* g_b v_{ba}) \right] \right) \\
=& 2 \left(\frac{1}{\sigma_{ab}^2} \left[2 \operatorname{Im}(g_a) \operatorname{Im}(g_b) |v_{ab}|^2 - \operatorname{Re}(m_{ab} v_{ab}^*) \right] \right. \\
&\left. + \frac{1}{\sigma_{ba}^2} \left[2 \operatorname{Im}(g_a) \operatorname{Im}(g_b) |v_{ba}|^2 - \operatorname{Re}(m_{ba}^* v_{ba}) \right] \right),
\end{aligned} \tag{14}
$$

1.4 Phase Regularization

The overall phase of the gains is degenerate. We want to constrain it such that the average phase of the gains is zero. We can do this by adding a regularization term to the chi-squared to produce a cost function of the form

$$
L(\boldsymbol{g}) = \sum_t \sum_{jk} \frac{1}{\sigma_{jk}^2(t)} \left| m_{jk}(t) - g_j g_k^* v_{jk}(t) \right|^2 + \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 . \tag{15}
$$

Taking derivatives gives

$$\frac{\partial}{\partial \operatorname{Re}(g_a)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 \tag{16}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Re}(g_a)} \operatorname{Arg}(g_a) \tag{17}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Re}(g_a)} \tan^{-1} \left[\frac{\operatorname{Im}(g_a)}{\operatorname{Re}(g_a)} \right] \tag{18}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{1}{1 + \frac{\operatorname{Im}^2(g_a)}{\operatorname{Re}^2(g_a)}} \frac{\partial}{\partial \operatorname{Re}(g_a)} \left[\frac{\operatorname{Im}(g_a)}{\operatorname{Re}(g_a)} \right] \tag{19}$$

$$= -2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{1}{1 + \frac{\operatorname{Im}^2(g_a)}{\operatorname{Re}^2(g_a)}} \frac{\operatorname{Im}(g_a)}{\operatorname{Re}^2(g_a)} \tag{20}$$

$$= -2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a)}{|g_a|^2} \tag{21}$$

$$\tag{22}$$

and

$$\frac{\partial}{\partial \operatorname{Im}(g_a)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 \tag{23}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \operatorname{Arg}(g_a) \tag{24}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \tan^{-1} \left[\frac{\operatorname{Im}(g_a)}{\operatorname{Re}(g_a)} \right] \tag{25}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{1}{1 + \frac{\operatorname{Im}^2(g_a)}{\operatorname{Re}^2(g_a)}} \frac{\partial}{\partial \operatorname{Im}(g_a)} \left[\frac{\operatorname{Im}(g_a)}{\operatorname{Re}(g_a)} \right] \tag{26}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{1}{1 + \frac{\operatorname{Im}^2(g_a)}{\operatorname{Re}^2(g_a)}} \frac{1}{\operatorname{Re}(g_a)} \tag{27}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}(g_a)}{|g_a|^2}. \tag{28}$$

$$\tag{29}$$

This can be rewritten as

$$\frac{\partial}{\partial \operatorname{Re}(g_a)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 = 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}(i g_a)}{|g_a|^2} \tag{30}$$

and

$$\frac{\partial}{\partial \operatorname{Im}(g_a)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 = 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(i g_a)}{|g_a|^2}. \tag{31}$$

The second derivatives are given by

$$\frac{\partial}{\partial \operatorname{Re}(g_a) \partial \operatorname{Re}(g_b)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 \tag{32}$$

$$= -2\lambda \frac{\partial}{\partial \operatorname{Re}(g_b)} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a)}{|g_a|^2} \tag{33}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{\partial}{\partial \operatorname{Re}(g_b)} \left[\sum_j \operatorname{Arg}(g_j) \right] + \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Re}(g_b)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{34}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{\partial}{\partial \operatorname{Re}(g_b)} \tan^{-1} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Re}(g_a)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{35}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \frac{\partial}{\partial \operatorname{Re}(g_b)} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Re}(g_a)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{36}$$

$$= 2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}^2(g_b)} \right] + 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right) \tag{37}$$

$$= 2\lambda \left(\frac{\operatorname{Im}(g_a) \operatorname{Im}(g_b)}{|g_a|^2 |g_b|^2} + 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right), \tag{38}$$

$$\tag{39}$$

$$\frac{\partial}{\partial \operatorname{Re}(g_a) \partial \operatorname{Im}(g_b)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 \tag{40}$$

$$= -2\lambda \frac{\partial}{\partial \operatorname{Im}(g_b)} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a)}{|g_a|^2} \tag{41}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{\partial}{\partial \operatorname{Im}(g_b)} \left[\sum_j \operatorname{Arg}(g_j) \right] + \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_b)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{42}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{\partial}{\partial \operatorname{Im}(g_b)} \tan^{-1} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{43}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \frac{\partial}{\partial \operatorname{Im}(g_b)} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{44}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \left[\frac{1}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{45}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a) \operatorname{Re}(g_b)}{|g_a|^2 |g_b|^2} + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}^2(g_a) - \operatorname{Im}^2(g_a)}{|g_a|^4} \right), \tag{46}$$

$$\tag{47}$$

and

$$\frac{\partial}{\partial \operatorname{Im}(g_a) \partial \operatorname{Im}(g_b)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 \tag{48}$$

$$= 2\lambda \frac{\partial}{\partial \operatorname{Im}(g_b)} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}(g_a)}{|g_a|^2} \tag{49}$$

$$= 2\lambda \left(\frac{\operatorname{Re}(g_a)}{|g_a|^2} \frac{\partial}{\partial \operatorname{Im}(g_b)} \left[\sum_j \operatorname{Arg}(g_j) \right] + \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_b)} \frac{\operatorname{Re}(g_a)}{|g_a|^2} \right) \tag{50}$$

$$= 2\lambda \left(\frac{\operatorname{Re}(g_a)}{|g_a|^2} \frac{\partial}{\partial \operatorname{Im}(g_b)} \tan^{-1} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \frac{\operatorname{Re}(g_a)}{|g_a|^2} \right) \tag{51}$$

$$= 2\lambda \left(\frac{\operatorname{Re}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \frac{\partial}{\partial \operatorname{Im}(g_b)} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \frac{\operatorname{Re}(g_a)}{|g_a|^2} \right) \tag{52}$$

$$= 2\lambda \left(\frac{\operatorname{Re}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \left[\frac{1}{\operatorname{Re}(g_b)} \right] - 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right) \tag{53}$$

$$= 2\lambda \left(\frac{\operatorname{Re}(g_a) \operatorname{Re}(g_b)}{|g_a|^2 |g_b|^2} - 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right). \tag{54}$$

$$\tag{55}$$

In summary,

$$\frac{\partial}{\partial \operatorname{Re}(g_a) \partial \operatorname{Re}(g_b)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 = 2\lambda \left(\frac{\operatorname{Im}(g_a) \operatorname{Im}(g_b)}{|g_a|^2 |g_b|^2} + 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right), \tag{56}$$

$$\frac{\partial}{\partial \operatorname{Re}(g_a) \partial \operatorname{Im}(g_b)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 = 2\lambda \left(- \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_b)}{|g_a|^2 |g_b|^2} - \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}^2(g_a) - \operatorname{Im}^2(g_a)}{|g_a|^4} \right), \tag{57}$$

$$\tag{58}$$

and

$$\frac{\partial}{\partial \operatorname{Im}(g_a)\partial \operatorname{Im}(g_b)}\lambda\left[\sum_j \operatorname{Arg}(g_j)\right]^2 = 2\lambda\left(\frac{\operatorname{Re}(g_a)\operatorname{Re}(g_b)}{|g_a|^2|g_b|^2} - 2\delta_{ab}\left[\sum_j \operatorname{Arg}(g_j)\right]\frac{\operatorname{Im}(g_a)\operatorname{Re}(g_a)}{|g_a|^4}\right). \tag{59}$$

2 Per-Frequency, Per-Polarization Including the Time Axis

Extending the simple implementation to include time steps is trivial. We assume that the gains do not vary on the timescale of the observation, so only the data and model visibilities are time-dependent. We then get a chi-squared in the form

$$\chi^2(\boldsymbol{g}) = \sum_t\sum_{jk}\frac{1}{\sigma_{jk}^2(t)}\left|m_{jk}(t) - g_j g_k^* v_{jk}(t)\right|^2. \tag{60}$$

The derivatives calculated in §1.2 then become

$$\begin{aligned}\frac{\partial\chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)} = 2\operatorname{Re}\sum_t\sum_{j\neq a}\Bigg(&\frac{1}{\sigma_{aj}^2(t)}\left[g_a|g_j|^2|v_{aj}(t)|^2 - m_{aj}(t)g_j v_{aj}^*(t)\right]\\ &+\frac{1}{\sigma_{ja}^2(t)}\left[g_a|g_j|^2|v_{ja}(t)|^2 - m_{ja}^*(t)g_j v_{ja}(t)\right]\Bigg).\end{aligned} \tag{61}$$

and

$$\begin{aligned}\frac{\partial\chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a)} = 2\operatorname{Im}\sum_t\sum_{j\neq a}\Bigg(&\frac{1}{\sigma_{aj}^2(t)}\left[g_a|g_j|^2|v_{aj}(t)|^2 - m_{aj}(t)g_j v_{aj}^*(t)\right]\\ &+\frac{1}{\sigma_{ja}^2(t)}\left[g_a|g_j|^2|v_{ja}(t)|^2 - m_{ja}^*(t)g_j v_{ja}(t)\right]\Bigg).\end{aligned} \tag{62}$$

The Hessian quantities in §1.3 become

$$\frac{\partial^2\chi^2(\boldsymbol{g})}{\partial \operatorname{Re}^2(g_a)} = \frac{\partial^2\chi^2(\boldsymbol{g})}{\partial \operatorname{Im}^2(g_a)} = 2\sum_t\sum_{j\neq a}\left(\frac{1}{\sigma_{aj}^2(t)}|g_j|^2|v_{aj}(t)|^2 + \frac{1}{\sigma_{ja}^2(t)}|g_j|^2|v_{ja}(t)|^2\right). \tag{63}$$

Once again,

$$\frac{\partial^2\chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a)\partial \operatorname{Im}(g_a)} = \frac{\partial^2\chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a)\partial \operatorname{Re}(g_a)} = 0. \tag{64}$$

For $b \neq a$, we get

$$\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a) \operatorname{Re}(g_b)} = 2 \sum_t \left(\frac{1}{\sigma_{ab}^2(t)} \left[2 \operatorname{Re}(g_a) \operatorname{Re}(g_b) |v_{ab}(t)|^2 - \operatorname{Re}[m_{ab}(t) v_{ab}^*(t)] \right] \right. \\ \left. + \frac{1}{\sigma_{ba}^2(t)} \left[2 \operatorname{Re}(g_a) \operatorname{Re}(g_b) |v_{ba}(t)|^2 - \operatorname{Re}[m_{ba}^*(t) v_{ba}(t)] \right] \right), \tag{65}$$

$$\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial \operatorname{Re}(g_a) \operatorname{Im}(g_b)} = 2 \sum_t \left(\frac{1}{\sigma_{ab}^2(t)} \left[2 \operatorname{Re}(g_a) \operatorname{Im}(g_b) |v_{ab}(t)|^2 + \operatorname{Im}[m_{ab}(t) v_{ab}^*(t)] \right] \right. \\ \left. + \frac{1}{\sigma_{ba}^2(t)} \left[2 \operatorname{Re}(g_a) \operatorname{Im}(g_b) |v_{ba}(t)|^2 + \operatorname{Im}[m_{ba}^*(t) v_{ba}(t)] \right] \right), \tag{66}$$

and

$$\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial \operatorname{Im}(g_a) \operatorname{Im}(g_b)} = 2 \sum_t \left(\frac{1}{\sigma_{ab}^2(t)} \left[2 \operatorname{Im}(g_a) \operatorname{Im}(g_b) |v_{ab}(t)|^2 - \operatorname{Re}[m_{ab}(t) v_{ab}^*(t)] \right] \right. \\ \left. + \frac{1}{\sigma_{ba}^2(t)} \left[2 \operatorname{Im}(g_a) \operatorname{Im}(g_b) |v_{ba}(t)|^2 - \operatorname{Re}[m_{ba}^*(t) v_{ba}(t)] \right] \right). \tag{67}$$

The phase regularization term does not depend on either the data or model visibilities, so it is time-independent. We can therefore use the quantities derived in §1.4.

3 Absolute Calibration

3.1 Cost Function

The absolute calibration chi-squared for each frequency and time is

$$\chi^2(A, \Delta_x, \Delta_y) = \sum_{jk} \frac{1}{\sigma_{jk}^2} \left| A^2 e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} - m_{jk} \right|^2. \tag{68}$$

Here v_{jk}^{rel} are the relatively calibrated visibilities and x_{jk} and y_{jk} are the coordinates of baseline $\{j, k\}$.

3.2 Jacobian Calculation

Expanding the cost function from above gives

$$
\begin{aligned}
\chi^2(A,\Delta_x,\Delta_y) =& \sum_{jk} \frac{1}{\sigma_{jk}^2} \left| A^2 e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} - m_{jk} \right|^2 \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Bigg[\Big(A^2 \,\mathrm{Re}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}] - \mathrm{Re}[m_{jk}] \Big)^2 \\
&+ \Big(A^2 \,\mathrm{Im}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}] - \mathrm{Im}[m_{jk}] \Big)^2 \Bigg] \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(A^4 |v_{jk}^{\mathrm{rel}}|^2 + |m_{jk}|^2 \\
&- 2A^2 \,\mathrm{Re}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}]\,\mathrm{Re}[m_{jk}] \\
&- 2A^2 \,\mathrm{Im}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}]\,\mathrm{Im}[m_{jk}] \Big) \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(A^4 |v_{jk}^{\mathrm{rel}}|^2 + |m_{jk}|^2 \\
&- 2A^2 \,\mathrm{Re}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(A^4 |v_{jk}^{\mathrm{rel}}|^2 + |m_{jk}|^2 \\
&- 2A^2 \cos(\Delta_x x_{jk} + \Delta_y y_{jk})\,\mathrm{Re}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \\
&+ 2A^2 \sin(\Delta_x x_{jk} + \Delta_y y_{jk})\,\mathrm{Im}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) .
\end{aligned} \tag{69}
$$

Taking the derivative with respect to the overall amplitude A gives

$$
\frac{\partial \chi^2(A,\Delta_x,\Delta_y)}{\partial A} = \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(4A^3 |v_{jk}^{\mathrm{rel}}|^2 - 4A \,\mathrm{Re}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) . \tag{70}
$$

Taking the derivative with respect to the phase gradient term Δ_x gives

$$
\begin{aligned}
\frac{\partial \chi^2(A,\Delta_x,\Delta_y)}{\partial \Delta_x} =& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(2A^2 x_{jk} \sin(\Delta_x x_{jk} + \Delta_y y_{jk})\,\mathrm{Re}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \\
&+ 2A^2 x_{jk} \cos(\Delta_x x_{jk} + \Delta_y y_{jk})\,\mathrm{Im}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) \\
=& 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk} \,\mathrm{Im}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*].
\end{aligned} \tag{71}
$$

It follows that

$$
\frac{\partial \chi^2(A,\Delta_x,\Delta_y)}{\partial \Delta_y} = 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} y_{jk} \,\mathrm{Im}[e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*]. \tag{72}
$$

3.3 Hessian Calculation

Taking the second derivatives gives

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial A^2} = \sum_{jk} \frac{1}{\sigma_{jk}^2} \left(12A^2 |v_{jk}^{\rm rel}|^2 - 4\,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\rm rel} m_{jk}^*] \right), \tag{73}$$

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_x)}{\partial A \partial \Delta_y} = 4A \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk}\, \mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\rm rel} m_{jk}^*], \tag{74}$$

and

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial A \partial \Delta_y} = 4A \sum_{jk} \frac{1}{\sigma_{jk}^2} y_{jk}\, \mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\rm rel} m_{jk}^*]. \tag{75}$$

The second derivatives with respect to the phase gradients are

$$\begin{aligned}\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial \Delta_x^2} &= \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(2A^2 x_{jk}^2 \cos(\Delta_x x_{jk} + \Delta_y y_{jk})\, \mathrm{Re}[v_{jk}^{\rm rel} m_{jk}^*] \\ &\quad - 2A^2 x_{jk}^2 \sin(\Delta_x x_{jk} + \Delta_y y_{jk})\, \mathrm{Im}[v_{jk}^{\rm rel} m_{jk}^*] \Big) \\ &= 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk}^2\, \mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\rm rel} m_{jk}^*].\end{aligned} \tag{76}$$

It follows that

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial \Delta_y^2} = 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} y_{jk}^2\, \mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\rm rel} m_{jk}^*] \tag{77}$$

and

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial \Delta_x \Delta_y} = 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk} y_{jk}\, \mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\rm rel} m_{jk}^*]. \tag{78}$$