Calibration Cost Functions and Derivatives

Ruby Byrne

last updated December 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

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

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

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

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 Gain Multiplication Convention

Equation 1 uses a convention in which the gains multiply the data v_{jk}. Another convention is to define the gains as multiplying the model visibilities. The gains would then be applied to the data by dividing the data by the gains, not multiplying. In that case, the cost function would become

$$
\chi^2(\boldsymbol{g}) = \sum_{jk} \frac{1}{\sigma_{jk}^2} \left| g_j g_k^* m_{jk} - v_{jk} \right|^2 . \tag{15}
$$

This is just equivalent to 1 where m_{jk} and v_{jk} are switched. The calibration solutions are not equivalent to those from Equation 1 because the noise properties are somewhat different.

The first and second derivatives from sections 1.2 and 1.3 can be inferred by simply switching m_{jk} and v_{jk}. This option is supported in the CaliCo software package by setting the option `gains_multiply_model=True`.

1.5 Phase Regularization

The overall phase of the gains is degenerate. Here we show how to constrain this phase by requiring that the average phase of the gains is zero. This is in contrast to an approach that uses a reference antenna.

We constrain the phase by adding a regularization term to the chi-squared to produce a cost function of the form

$$
L(\boldsymbol{g}) = \sum_t \sum_{jk} \frac{1}{\sigma_{jk}^2(t)} \left| m_{jk}(t) - g_j g_k^* v_{jk}(t) \right|^2 + \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 . \tag{16}
$$

Taking derivatives gives

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

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

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

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

$$= -2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{1}{1 + \frac{\operatorname{Im}^2(g_a)}{\operatorname{Re}^2(g_a)}} \frac{\operatorname{Im}(g_a)}{\operatorname{Re}^2(g_a)} \tag{21}$$

$$= -2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a)}{|g_a|^2} \tag{22}$$

$$\tag{23}$$

and

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

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

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

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

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{1}{1 + \frac{\operatorname{Im}^2(g_a)}{\operatorname{Re}^2(g_a)}} \frac{1}{\operatorname{Re}(g_a)} \tag{28}$$

$$= 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}(g_a)}{|g_a|^2}. \tag{29}$$

$$\tag{30}$$

This can be rewritten as

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

and

$$\frac{\partial}{\partial \operatorname{Im}(g_a)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 = 2\lambda \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(i g_a)}{|g_a|^2}. \tag{32}$$

The second derivatives are given by

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

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

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

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

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

$$= 2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \left[\frac{\operatorname{Im}(g_b)}{\operatorname{Re}^2(g_b)} \right] + 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right) \tag{38}$$

$$= 2\lambda \left(\frac{\operatorname{Im}(g_a) \operatorname{Im}(g_b)}{|g_a|^2 |g_b|^2} + 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right), \tag{39}$$

$$\tag{40}$$

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

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

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

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

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

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \left[\frac{1}{\operatorname{Re}(g_b)} \right] + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\partial}{\partial \operatorname{Im}(g_a)} \frac{\operatorname{Im}(g_a)}{|g_a|^2} \right) \tag{46}$$

$$= -2\lambda \left(\frac{\operatorname{Im}(g_a) \operatorname{Re}(g_b)}{|g_a|^2 |g_b|^2} + \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}^2(g_a) - \operatorname{Im}^2(g_a)}{|g_a|^4} \right), \tag{47}$$

(48)

and

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

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

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

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

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

$$= 2\lambda \left(\frac{\operatorname{Re}(g_a)}{|g_a|^2} \frac{1}{1 + \frac{\operatorname{Im}^2(g_b)}{\operatorname{Re}^2(g_b)}} \left[\frac{1}{\operatorname{Re}(g_b)} \right] - 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right) \tag{54}$$

$$= 2\lambda \left(\frac{\operatorname{Re}(g_a) \operatorname{Re}(g_b)}{|g_a|^2 |g_b|^2} - 2\delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Im}(g_a) \operatorname{Re}(g_a)}{|g_a|^4} \right). \tag{55}$$

$$\tag{56}$$

In summary,

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

$$\frac{\partial}{\partial \operatorname{Re}(g_a) \partial \operatorname{Im}(g_b)} \lambda \left[\sum_j \operatorname{Arg}(g_j) \right]^2 = 2\lambda \left(-\frac{\operatorname{Im}(g_a) \operatorname{Re}(g_b)}{|g_a|^2 |g_b|^2} - \delta_{ab} \left[\sum_j \operatorname{Arg}(g_j) \right] \frac{\operatorname{Re}^2(g_a) - \operatorname{Im}^2(g_a)}{|g_a|^4} \right), \tag{58}$$

$$\tag{59}$$

and

$$\frac{\partial}{\partial\,\mathrm{Im}(g_a)\partial\,\mathrm{Im}(g_b)}\lambda\left[\sum_j \mathrm{Arg}(g_j)\right]^2 = 2\lambda\left(\frac{\mathrm{Re}(g_a)\,\mathrm{Re}(g_b)}{|g_a|^2|g_b|^2} - 2\delta_{ab}\left[\sum_j \mathrm{Arg}(g_j)\right]\frac{\mathrm{Im}(g_a)\,\mathrm{Re}(g_a)}{|g_a|^4}\right). \tag{60}$$

1.6 Cross-Polarization Phase

Per-polarization calibration has a degeneracy in the cross-polarization phase, meaning the overall phase between the p-polarized and q-polarized gains. We can see this by noting that the transformation $g_{jp} \rightarrow g_{jp}e^{-i\phi/2}$ and $g_{jq} \rightarrow g_{jq}e^{i\phi/2}$ leave χ^2 unchanged, where here p and q denote the two feed polarizations.

1.6.1 From Model Visibilities

We can constrain this phase by matching the cross-polarization data and model. We do this by minimizing this chi-squared:

$$\chi^2(\phi) = \sum_{jk}\frac{1}{\sigma^2_{jkpq}}\left|e^{-i\phi}v^{\mathrm{cal}}_{jkpq} - m_{jkpq}\right|^2 + \sum_{jk}\frac{1}{\sigma^2_{jkqp}}\left|e^{i\phi}v^{\mathrm{cal}}_{jkqp} - m_{jkqp}\right|^2. \tag{61}$$

Here v^{cal}_{jkpq} and v^{cal}_{jkqp} are the cross-polarization visibilities, calibrated up to the cross-polarization phase.

Expanding the chi-squared gives

$$
\begin{aligned}
\chi^2(\phi) = &\sum_{jk} \frac{1}{\sigma^2_{jkpq}} \left[\left(\mathrm{Re}[e^{-i\phi} v^{\mathrm{cal}}_{jkpq}] - \mathrm{Re}[m_{jkpq}] \right)^2 + \left(\mathrm{Im}[e^{-i\phi} v^{\mathrm{cal}}_{jkpq}] - \mathrm{Im}[m_{jkpq}] \right)^2 \right] \\
&+ \sum_{jk} \frac{1}{\sigma^2_{jkqp}} \left[\left(\mathrm{Re}[e^{i\phi} v^{\mathrm{cal}}_{jkqp}] - \mathrm{Re}[m_{jkqp}] \right)^2 + \left(\mathrm{Im}[e^{i\phi} v^{\mathrm{cal}}_{jkqp}] - \mathrm{Im}[m_{jkqp}] \right)^2 \right] \\
= &\sum_{jk} \frac{1}{\sigma^2_{jkpq}} \left[|v^{\mathrm{cal}}_{jkpq}|^2 + |m_{jkpq}|^2 - 2\,\mathrm{Re}[e^{-i\phi} v^{\mathrm{cal}}_{jkpq}]\,\mathrm{Re}[m_{jkpq}] - 2\,\mathrm{Im}[e^{-i\phi} v^{\mathrm{cal}}_{jkpq}]\,\mathrm{Im}[m_{jkpq}] \right] \\
&+ \sum_{jk} \frac{1}{\sigma^2_{jkqp}} \left[|v^{\mathrm{cal}}_{jkqp}|^2 + |m_{jkqp}|^2 - 2\,\mathrm{Re}[e^{i\phi} v^{\mathrm{cal}}_{jkqp}]\,\mathrm{Re}[m_{jkqp}] - 2\,\mathrm{Im}[e^{i\phi} v^{\mathrm{cal}}_{jkqp}]\,\mathrm{Im}[m_{jkqp}] \right] \\
= &\sum_{jk} \frac{1}{\sigma^2_{jkpq}} \left[|v^{\mathrm{cal}}_{jkpq}|^2 + |m_{jkpq}|^2 - 2\,\mathrm{Re}[e^{-i\phi} v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] \right] \\
&+ \sum_{jk} \frac{1}{\sigma^2_{jkqp}} \left[|v^{\mathrm{cal}}_{jkqp}|^2 + |m_{jkqp}|^2 - 2\,\mathrm{Re}[e^{i\phi} v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] \right] \\
= &\sum_{jk} \frac{1}{\sigma^2_{jkpq}} \left[|v^{\mathrm{cal}}_{jkpq}|^2 + |m_{jkpq}|^2 - 2\cos[-\phi]\,\mathrm{Re}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] + 2\sin[-\phi]\,\mathrm{Im}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] \right] \\
&+ \sum_{jk} \frac{1}{\sigma^2_{jkqp}} \left[|v^{\mathrm{cal}}_{jkqp}|^2 + |m_{jkqp}|^2 - 2\cos\phi\,\mathrm{Re}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] + 2\sin\phi\,\mathrm{Im}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] \right] \\
= &\sum_{jk} \frac{1}{\sigma^2_{jkpq}} \left[|v^{\mathrm{cal}}_{jkpq}|^2 + |m_{jkpq}|^2 - 2\cos\phi\,\mathrm{Re}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] - 2\sin\phi\,\mathrm{Im}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] \right] \\
&+ \sum_{jk} \frac{1}{\sigma^2_{jkqp}} \left[|v^{\mathrm{cal}}_{jkqp}|^2 + |m_{jkqp}|^2 - 2\cos\phi\,\mathrm{Re}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] + 2\sin\phi\,\mathrm{Im}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] \right].
\end{aligned}
\tag{62}
$$

Taking the derivative gives

$$
\begin{aligned}
\frac{\partial \chi^2(\phi)}{\partial \phi} = &\sum_{jk} \frac{1}{\sigma^2_{jkpq}} \left[2\sin\phi\,\mathrm{Re}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] - 2\cos\phi\,\mathrm{Im}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] \right] \\
&+ \sum_{jk} \frac{1}{\sigma^2_{jkqp}} \left[2\sin\phi\,\mathrm{Re}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] + 2\cos\phi\,\mathrm{Im}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] \right].
\end{aligned}
\tag{63}
$$

Using the extremum condition $\left.\frac{\partial \chi^2(\phi)}{\partial \phi}\right|_{\phi=\phi_0} = 0$ gives

$$
\begin{aligned}
&\sin\phi_0 \sum_{jk} \left(\frac{1}{\sigma^2_{jkpq}}\,\mathrm{Re}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] + \frac{1}{\sigma^2_{jkqp}}\,\mathrm{Re}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] \right) \\
&= \cos\phi_0 \sum_{jk} \left(\frac{1}{\sigma^2_{jkpq}}\,\mathrm{Im}[v^{\mathrm{cal}}_{jkpq} m^*_{jkpq}] - \frac{1}{\sigma^2_{jkqp}}\,\mathrm{Im}[v^{\mathrm{cal}}_{jkqp} m^*_{jkqp}] \right)
\end{aligned}
\tag{64}
$$

or

$$
\begin{aligned}
\tan\phi_0 &= \frac{\sum_{jk}\left(\frac{1}{\sigma^2_{jkpq}}\,\mathrm{Im}[v^{\mathrm{cal}}_{jkpq}m^*_{jkpq}] - \frac{1}{\sigma^2_{jkqp}}\,\mathrm{Im}[v^{\mathrm{cal}}_{jkqp}m^*_{jkqp}]\right)}{\sum_{jk}\left(\frac{1}{\sigma^2_{jkpq}}\,\mathrm{Re}[v^{\mathrm{cal}}_{jkpq}m^*_{jkpq}] + \frac{1}{\sigma^2_{jkqp}}\,\mathrm{Re}[v^{\mathrm{cal}}_{jkqp}m^*_{jkqp}]\right)} \\
&= \frac{\sum_{jk}\left(\frac{1}{\sigma^2_{jkpq}}\,\mathrm{Im}[v^{\mathrm{cal}}_{jkpq}m^*_{jkpq}] + \frac{1}{\sigma^2_{jkqp}}\,\mathrm{Im}[(v^{\mathrm{cal}}_{jkqp})^*m_{jkqp}]\right)}{\sum_{jk}\left(\frac{1}{\sigma^2_{jkpq}}\,\mathrm{Re}[v^{\mathrm{cal}}_{jkpq}m^*_{jkpq}] + \frac{1}{\sigma^2_{jkqp}}\,\mathrm{Re}[(v^{\mathrm{cal}}_{jkqp})^*m_{jkqp}]\right)} \\
&= \frac{\mathrm{Im}\left[\sum_{jk}\left(\frac{1}{\sigma^2_{jkpq}}v^{\mathrm{cal}}_{jkpq}m^*_{jkpq} + \frac{1}{\sigma^2_{jkqp}}(v^{\mathrm{cal}}_{jkqp})^*m_{jkqp}\right)\right]}{\mathrm{Re}\left[\sum_{jk}\left(\frac{1}{\sigma^2_{jkpq}}v^{\mathrm{cal}}_{jkpq}m^*_{jkpq} + \frac{1}{\sigma^2_{jkqp}}(v^{\mathrm{cal}}_{jkqp})^*m_{jkqp}\right)\right]}.
\end{aligned}
\tag{65}
$$

It follows that

$$
\phi_0 = \mathrm{Arg}\left[\sum_{jk}\left(\frac{1}{\sigma^2_{jkpq}}v^{\mathrm{cal}}_{jkpq}m^*_{jkpq} + \frac{1}{\sigma^2_{jkqp}}(v^{\mathrm{cal}}_{jkqp})^*m_{jkqp}\right)\right], \tag{66}
$$

where Arg denotes the complex argument.

1.6.2 From Pseudo-V

An alternative approach to minimizing the cross-polarization phase is appropriate when the cross-polarization model visibilities are untrustworthy. This approach leverages the assumption that the sky has minimal instrinsic Stokes V polarization and that psuedo Stokes V is a reasonable approximation for true Stokes V.

Pseudo Stokes V visibilities are given by $iv_{jkpq} - iv_{jkqp}$. We can minimize pseudo Stokes V by minimizing the quantity

$$
\chi^2(\phi) = \sum_{jk}\frac{1}{\sigma^2_{jk}}\left|ie^{-i\phi}v^{\mathrm{cal}}_{jkpq} - ie^{i\phi}v^{\mathrm{cal}}_{jkqp}\right|^2, \tag{67}
$$

where v^{cal}_{jkpq} and v^{cal}_{jkqp} are the cross-polarization visibilities, calibrated up to the cross-polarization phase ϕ. This is equivalent to minimizing

$$
\chi^2(\phi) = \sum_{jk}\frac{1}{\sigma^2_{jk}}\left|e^{-i\phi}v^{\mathrm{cal}}_{jkpq} - e^{i\phi}v^{\mathrm{cal}}_{jkqp}\right|^2. \tag{68}
$$

Expanding this expression gives

$$\begin{aligned}
\chi^2(\phi) &= \sum_{jk} \frac{1}{\sigma_{jk}^2} \left[\mathrm{Re}^2 \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} - e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) + \mathrm{Im}^2 \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} - e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) \right] \\
&= \sum_{jk} \frac{1}{\sigma_{jk}^2} \left[\mathrm{Re}^2 \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) + \mathrm{Re}^2 \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) - 2\,\mathrm{Re} \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) \mathrm{Re} \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) \right. \\
&\left. + \mathrm{Im}^2 \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) + \mathrm{Im}^2 \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) - 2\,\mathrm{Im} \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) \mathrm{Im} \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) \right] \\
&= \sum_{jk} \frac{1}{\sigma_{jk}^2} \left[\left| v_{jkpq}^{\mathrm{cal}} \right|^2 + \left| v_{jkqp}^{\mathrm{cal}} \right|^2 \right. \\
&\left. -2\,\mathrm{Re} \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) \mathrm{Re} \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) - 2\,\mathrm{Im} \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) \mathrm{Im} \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) \right].
\end{aligned} \tag{69}$$

Minimizing this expression is equivalent to minimizing

$$\begin{aligned}
\chi^2 &= \sum_{jk} \frac{1}{\sigma_{jk}^2} \left[-\mathrm{Re} \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) \mathrm{Re} \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) - \mathrm{Im} \left(e^{-i\phi} v_{jkpq}^{\mathrm{cal}} \right) \mathrm{Im} \left(e^{i\phi} v_{jkqp}^{\mathrm{cal}} \right) \right] \\
&= -\sum_{jk} \frac{1}{\sigma_{jk}^2} \mathrm{Re} \left(e^{-2i\phi} v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) \\
&= -\sum_{jk} \frac{1}{\sigma_{jk}^2} \left[\cos(-2\phi)\,\mathrm{Re} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) - \sin(-2\phi)\,\mathrm{Im} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) \right] \\
&= -\sum_{jk} \frac{1}{\sigma_{jk}^2} \left[\cos(2\phi)\,\mathrm{Re} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) + \sin(2\phi)\,\mathrm{Im} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) \right].
\end{aligned} \tag{70}$$

Taking the derivative gives

$$\frac{\partial \chi^2(\phi)}{\partial \phi} = -\sum_{jk} \frac{1}{\sigma_{jk}^2} \left[-2\sin(2\phi)\,\mathrm{Re} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) + 2\cos(2\phi)\,\mathrm{Im} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) \right]. \tag{71}$$

Using the extremum condition $\left.\frac{\partial \chi^2(\phi)}{\partial \phi}\right|_{\phi=\phi_0} = 0$ gives

$$\sin(2\phi_0) \sum_{jk} \frac{1}{\sigma_{jk}^2} \mathrm{Re} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right) = \cos(2\phi_0) \sum_{jk} \frac{1}{\sigma_{jk}^2} \mathrm{Im} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right), \tag{72}$$

so

$$\tan(2\phi_0) = \frac{\sum_{jk} \frac{1}{\sigma_{jk}^2} \mathrm{Im} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right)}{\sum_{jk} \frac{1}{\sigma_{jk}^2} \mathrm{Re} \left(v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^* \right)}. \tag{73}$$

We then get that

$$\phi_0 = \frac{1}{2}\,\mathrm{Arg}\left[\sum_{jk} \frac{1}{\sigma_{jk}^2} v_{jkpq}^{\mathrm{cal}} (v_{jkqp}^{\mathrm{cal}})^*\right] \tag{74}$$

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

$$\chi^2(\boldsymbol{g}) = \sum_t \sum_{jk} \frac{1}{\sigma_{jk}^2(t)} \left| m_{jk}(t) - g_j g_k^* v_{jk}(t) \right|^2 . \tag{75}$$

The derivatives calculated in §1.2 then become

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Re}(g_a)} = 2\,\mathrm{Re} \sum_t \sum_{j\neq a} \Bigg(& \frac{1}{\sigma_{aj}^2(t)} \left[g_a |g_j|^2 |v_{aj}(t)|^2 - m_{aj}(t) g_j v_{aj}^*(t) \right] \\ & + \frac{1}{\sigma_{ja}^2(t)} \left[g_a |g_j|^2 |v_{ja}(t)|^2 - m_{ja}^*(t) g_j v_{ja}(t) \right] \Bigg) .\end{aligned} \tag{76}$$

and

$$\begin{aligned}\frac{\partial \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Im}(g_a)} = 2\,\mathrm{Im} \sum_t \sum_{j\neq a} \Bigg(& \frac{1}{\sigma_{aj}^2(t)} \left[g_a |g_j|^2 |v_{aj}(t)|^2 - m_{aj}(t) g_j v_{aj}^*(t) \right] \\ & + \frac{1}{\sigma_{ja}^2(t)} \left[g_a |g_j|^2 |v_{ja}(t)|^2 - m_{ja}^*(t) g_j v_{ja}(t) \right] \Bigg) .\end{aligned} \tag{77}$$

The Hessian quantities in §1.3 become

$$\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Re}^2(g_a)} = \frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Im}^2(g_a)} = 2 \sum_t \sum_{j\neq a} \left(\frac{1}{\sigma_{aj}^2(t)} |g_j|^2 |v_{aj}(t)|^2 + \frac{1}{\sigma_{ja}^2(t)} |g_j|^2 |v_{ja}(t)|^2 \right) . \tag{78}$$

Once again,

$$\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Re}(g_a) \partial\,\mathrm{Im}(g_a)} = \frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Im}(g_a) \partial\,\mathrm{Re}(g_a)} = 0. \tag{79}$$

For $b \neq a$, we get

$$\begin{aligned}\frac{\partial^2 \chi^2(\boldsymbol{g})}{\partial\,\mathrm{Re}(g_a)\,\mathrm{Re}(g_b)} = 2 \sum_t \Bigg(& \frac{1}{\sigma_{ab}^2(t)} \left[2\,\mathrm{Re}(g_a)\,\mathrm{Re}(g_b) |v_{ab}(t)|^2 - \mathrm{Re}[m_{ab}(t) v_{ab}^*(t)] \right] \\ & + \frac{1}{\sigma_{ba}^2(t)} \left[2\,\mathrm{Re}(g_a)\,\mathrm{Re}(g_b) |v_{ba}(t)|^2 - \mathrm{Re}[m_{ba}^*(t) v_{ba}(t)] \right] \Bigg) ,\end{aligned} \tag{80}$$

$$\frac{\partial^2\chi^2(\boldsymbol{g})}{\partial\,\mathrm{Re}(g_a)\,\mathrm{Im}(g_b)} = 2\sum_t\left(\frac{1}{\sigma_{ab}^2(t)}\left[2\,\mathrm{Re}(g_a)\,\mathrm{Im}(g_b)|v_{ab}(t)|^2 + \mathrm{Im}[m_{ab}(t)v_{ab}^*(t)]\right]\right.$$
$$\left. + \frac{1}{\sigma_{ba}^2(t)}\left[2\,\mathrm{Re}(g_a)\,\mathrm{Im}(g_b)|v_{ba}(t)|^2 + \mathrm{Im}[m_{ba}^*(t)v_{ba}(t)]\right]\right), \tag{81}$$

and

$$\frac{\partial^2\chi^2(\boldsymbol{g})}{\partial\,\mathrm{Im}(g_a)\,\mathrm{Im}(g_b)} = 2\sum_t\left(\frac{1}{\sigma_{ab}^2(t)}\left[2\,\mathrm{Im}(g_a)\,\mathrm{Im}(g_b)|v_{ab}(t)|^2 - \mathrm{Re}[m_{ab}(t)v_{ab}^*(t)]\right]\right.$$
$$\left. + \frac{1}{\sigma_{ba}^2(t)}\left[2\,\mathrm{Im}(g_a)\,\mathrm{Im}(g_b)|v_{ba}(t)|^2 - \mathrm{Re}[m_{ba}^*(t)v_{ba}(t)]\right]\right). \tag{82}$$

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 §1.5.

3 Absolute Calibration

3.1 Cost Function

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

$$\chi^2(A,\Delta_x,\Delta_y) = \sum_{jk}\frac{1}{\sigma_{jk}^2}\left|A^2 e^{i(\Delta_x x_{jk}+\Delta_y y_{jk})}v_{jk}^{\mathrm{rel}} - m_{jk}\right|^2. \tag{83}$$

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

$$\begin{aligned}
\chi^2(A,\Delta_x,\Delta_y) =& \sum_{jk} \frac{1}{\sigma_{jk}^2} \left| A^2 e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} - m_{jk} \right|^2 \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Bigg[\Big(A^2 \,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}] - \mathrm{Re}[m_{jk}] \Big)^2 \\
&+ \Big(A^2 \,\mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}] - \mathrm{Im}[m_{jk}] \Big)^2 \Bigg] \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(A^4 |v_{jk}^{\mathrm{rel}}|^2 + |m_{jk}|^2 \\
&- 2A^2 \,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}] \,\mathrm{Re}[m_{jk}] \\
&- 2A^2 \,\mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}}] \,\mathrm{Im}[m_{jk}] \Big) \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(A^4 |v_{jk}^{\mathrm{rel}}|^2 + |m_{jk}|^2 \\
&- 2A^2 \,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) \\
=& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(A^4 |v_{jk}^{\mathrm{rel}}|^2 + |m_{jk}|^2 \\
&- 2A^2 \cos(\Delta_x x_{jk} + \Delta_y y_{jk}) \,\mathrm{Re}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \\
&+ 2A^2 \sin(\Delta_x x_{jk} + \Delta_y y_{jk}) \,\mathrm{Im}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) .
\end{aligned} \tag{84}$$

Taking the derivative with respect to the overall amplitude A gives

$$\frac{\partial \chi^2(A,\Delta_x,\Delta_y)}{\partial A} = \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(4A^3 |v_{jk}^{\mathrm{rel}}|^2 - 4A \,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) . \tag{85}$$

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

$$\begin{aligned}
\frac{\partial \chi^2(A,\Delta_x,\Delta_y)}{\partial \Delta_x} =& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(2A^2 x_{jk} \sin(\Delta_x x_{jk} + \Delta_y y_{jk}) \,\mathrm{Re}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \\
&+ 2A^2 x_{jk} \cos(\Delta_x x_{jk} + \Delta_y y_{jk}) \,\mathrm{Im}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) \\
=& 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk} \,\mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*].
\end{aligned} \tag{86}$$

It follows that

$$\frac{\partial \chi^2(A,\Delta_x,\Delta_y)}{\partial \Delta_y} = 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} y_{jk} \,\mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*]. \tag{87}$$

3.3 Hessian Calculation

Taking the second derivatives gives

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial A^2} = \sum_{jk} \frac{1}{\sigma_{jk}^2} \left(12A^2 |v_{jk}^{\mathrm{rel}}|^2 - 4\,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*] \right), \tag{88}$$

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_x)}{\partial A \partial \Delta_y} = 4A \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk}\,\mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*], \tag{89}$$

and

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial A \partial \Delta_y} = 4A \sum_{jk} \frac{1}{\sigma_{jk}^2} y_{jk}\,\mathrm{Im}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*]. \tag{90}$$

The second derivatives with respect to the phase gradients are

$$\begin{aligned}\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial \Delta_x^2} =& \sum_{jk} \frac{1}{\sigma_{jk}^2} \Big(2A^2 x_{jk}^2 \cos(\Delta_x x_{jk} + \Delta_y y_{jk})\,\mathrm{Re}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \\ &- 2A^2 x_{jk}^2 \sin(\Delta_x x_{jk} + \Delta_y y_{jk})\,\mathrm{Im}[v_{jk}^{\mathrm{rel}} m_{jk}^*] \Big) \\ =& 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk}^2\,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*].\end{aligned} \tag{91}$$

It follows that

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial \Delta_y^2} = 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} y_{jk}^2\,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*] \tag{92}$$

and

$$\frac{\partial^2 \chi^2(A, \Delta_x, \Delta_y)}{\partial \Delta_x \Delta_y} = 2A^2 \sum_{jk} \frac{1}{\sigma_{jk}^2} x_{jk} y_{jk}\,\mathrm{Re}[e^{i(\Delta_x x_{jk} + \Delta_y y_{jk})} v_{jk}^{\mathrm{rel}} m_{jk}^*]. \tag{93}$$

3.4 Applying Abscal Solutions

For per-polarization abscal, the outputs are parameters A_p, Δ_{xp}, Δ_{yp}, A_q, Δ_{xq}, and Δ_{yq}. Applying these solutions to the single-polarization visibilities is simple:

$$v_{jk\,pp}^{\mathrm{cal}} = A_p^2 e^{i(\Delta_{xp} x_{jk} + \Delta_{yp} y_{jk})} v_{jk\,pp}^{\mathrm{rel}} \tag{94}$$

and

$$v_{jk\,qq}^{\mathrm{cal}} = A_q^2 e^{i(\Delta_{xq} x_{jk} + \Delta_{yq} y_{jk})} v_{jk\,qq}^{\mathrm{rel}}, \tag{95}$$

where $v_{jk\,pp}^{\mathrm{cal}}$ and $v_{jk\,qq}^{\mathrm{cal}}$ are the calibrated visibilities.

However, calibrating the crosspol visibilities is slightly more complicated. Here we get that

$$v^{\mathrm{cal}}_{jk\,pq} = A_p A_q e^{i(\Delta_{xp} x_j - \Delta_{xq} x_k + \Delta_{yp} y_j - \Delta_{yq} y_k)} v^{\mathrm{rel}}_{jk\,pq} \tag{96}$$

and

$$v^{\mathrm{cal}}_{jk\,qp} = A_p A_q e^{i(\Delta_{xq} x_j - \Delta_{xp} x_k + \Delta_{yq} y_j - \Delta_{yp} y_k)} v^{\mathrm{rel}}_{jk\,qp}. \tag{97}$$

Here $\{x_j, y_j\}$ and $\{x_k, y_k\}$ are the coordinates of antennas j and k, respectively. Note that we require calibrating with the antenna positions, not with the baseline locations.

4 Delay-Weighted Absolute Calibration

4.1 Cost Function

For delay-weighted absolute calibration, we need to include the frequency axis in our cost function calculation. The basic cost would then be

$$\begin{aligned} &\chi^2[A(f), \Delta_x(f), \Delta_y(f)] = \\ &\sum_f \sum_{jk} \frac{1}{\sigma^2_{jk}(f)} \left| A^2(f) e^{i[\Delta_x(f) x_{jk} + \Delta_y(f) y_{jk}]} v^{\mathrm{rel}}_{jk}(f) - m_{jk}(f) \right|^2. \end{aligned} \tag{98}$$

With the weighting function $W_{jk}(f)$, this becomes

$$\begin{aligned} &\chi^2[A(f), \Delta_x(f), \Delta_y(f)] = \\ &\sum_{f_1} \sum_{f_2} \sum_{jk} \frac{1}{\sigma_{jk}(f_1)\sigma_{jk}(f_2)} W_{jk}(f_2 - f_1) \times \\ &\left[A^2(f_1) e^{i[\Delta_x(f_1) x_{jk} + \Delta_y(f_1) y_{jk}]} v^{\mathrm{rel}}_{jk}(f_1) - m_{jk}(f_1) \right]^* \times \\ &\left[A^2(f_2) e^{i[\Delta_x(f_2) x_{jk} + \Delta_y(f_2) y_{jk}]} v^{\mathrm{rel}}_{jk}(f_2) - m_{jk}(f_2) \right]. \end{aligned} \tag{99}$$

4.2 Jacobian

We now take derivatives of the cost function:

$$
\begin{aligned}
&\frac{\partial \chi^2[A(f), \Delta_x(f), \Delta_y(f)]}{\partial A(f_0)} \\
&= \frac{\partial}{\partial A(f_0)} \sum_{f_1} \sum_{f_2} \sum_{jk} \frac{1}{\sigma_{jk}(f_1)\sigma_{jk}(f_2)} W_{jk}(f_2 - f_1) \times \\
&\left[A^2(f_1) e^{i[\Delta_x(f_1)x_{jk} + \Delta_y(f_1)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_1) - m_{jk}(f_1)\right]^* \times \\
&\left[A^2(f_2) e^{i[\Delta_x(f_2)x_{jk} + \Delta_y(f_2)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_2) - m_{jk}(f_2)\right] \\
&= \sum_{f_1 \neq f_0} \sum_{jk} \frac{1}{\sigma_{jk}(f_1)\sigma_{jk}(f_0)} W_{jk}(f_0 - f_1) \times \\
&\left[A^2(f_1) e^{i[\Delta_x(f_1)x_{jk} + \Delta_y(f_1)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_1) - m_{jk}(f_1)\right]^* \times \\
&\left[2A(f_0) e^{i[\Delta_x(f_0)x_{jk} + \Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right] \\
&+ \sum_{f_2 \neq f_0} \sum_{jk} \frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_2)} W_{jk}(f_2 - f_0) \times \\
&\left[2A(f_0) e^{i[\Delta_x(f_0)x_{jk} + \Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right]^* \times \\
&\left[A^2(f_2) e^{i[\Delta_x(f_2)x_{jk} + \Delta_y(f_2)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_2) - m_{jk}(f_2)\right] \\
&+ \sum_{jk} \frac{1}{\sigma_{jk}^2(f_0)} W_{jk}(0) \times \\
&\left[A^2(f_0) e^{i[\Delta_x(f_0)x_{jk} + \Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0) - m_{jk}(f_0)\right]^* \times \\
&\left[2A(f_0) e^{i[\Delta_x(f_0)x_{jk} + \Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right] \\
&+ \sum_{jk} \frac{1}{\sigma_{jk}^2(f_0)} W_{jk}(0) \times \\
&\left[2A(f_0) e^{i[\Delta_x(f_0)x_{jk} + \Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right]^* \times \\
&\left[A^2(f_0) e^{i[\Delta_x(f_0)x_{jk} + \Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0) - m_{jk}(f_0)\right].
\end{aligned}
\tag{100}
$$

This reduces to

$$
\begin{aligned}
&\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial A(f_0)} \\
&=\sum_{f_1}\sum_{jk}\frac{1}{\sigma_{jk}(f_1)\sigma_{jk}(f_0)}W_{jk}(f_0-f_1)\times \\
&\left[A^2(f_1)e^{i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_1)-m_{jk}(f_1)\right]^*\times \\
&\left[2A(f_0)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right] \\
&+\sum_{f_2}\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_2)}W_{jk}(f_2-f_0)\times \\
&\left[2A(f_0)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right]^*\times \\
&\left[A^2(f_2)e^{i[\Delta_x(f_2)x_{jk}+\Delta_y(f_2)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_2)-m_{jk}(f_2)\right].
\end{aligned}
\tag{101}
$$

Using the fact that the weighting matrix must be Hermitian, such that $W_{jk}(f)=W_{jk}^*(-f)$, we get that

$$
\begin{aligned}
&\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial A(f_0)} \\
&=4A(f_0)\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)}\left[e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right]^*\times\right. \\
&\left.\sum_{f}\frac{1}{\sigma_{jk}(f)}W_{jk}(f-f_0)\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]\right).
\end{aligned}
\tag{102}
$$

Now taking derivatives with respect to the phase gradient terms gives

$$
\begin{aligned}
&\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial \Delta_x(f_0)} \\
&=\sum_{f_1}\sum_{jk}\frac{1}{\sigma_{jk}(f_1)\sigma_{jk}(f_0)}W_{jk}(f_0-f_1)\times \\
&\left[A^2(f_1)e^{i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_1)-m_{jk}(f_1)\right]^*\times \\
&\left[ix_{jk}A^2(f_0)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right] \\
&+\sum_{f_2}\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_2)}W_{jk}(f_2-f_0)\times \\
&\left[-ix_{jk}A^2(f_0)e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\times \\
&\left[A^2(f_2)e^{i[\Delta_x(f_2)x_{jk}+\Delta_y(f_2)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_2)-m_{jk}(f_2)\right].
\end{aligned}
\tag{103}
$$

Once again using the fact that the weighting matrix is Hermitian, this becomes

$$\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial\Delta_x(f_0)}$$
$$= 2A^2(f_0)\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)}\left[-ix_{jk}e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\right.$$
$$\left.\times\sum_{f}\frac{1}{\sigma_{jk}(f)}W_{jk}(f-f_0)\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]\right). \tag{104}$$

It follows that

$$\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial\Delta_y(f_0)}$$
$$= 2A^2(f_0)\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)}\left[-iy_{jk}e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\right.$$
$$\left.\times\sum_{f}\frac{1}{\sigma_{jk}(f)}W_{jk}(f-f_0)\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]\right). \tag{105}$$

4.3 Hessian

Beginning with the second derivative with respect to the amplitude terms, we get

$$
\begin{aligned}
&\frac{\partial^2 \chi^2[A(f), \Delta_x(f), \Delta_y(f)]}{\partial A(f_0)\partial A(f_1)} \\
&= 4\frac{\partial A(f_0)}{\partial A(f_1)} \operatorname{Re}\left(\sum_{jk} \frac{1}{\sigma_{jk}(f_0)} \left[e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right]^* \times \right. \\
&\left. \sum_f \frac{1}{\sigma_{jk}(f)} W_{jk}(f-f_0) \left[A^2(f) e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]} v_{jk}^{\mathrm{rel}}(f) - m_{jk}(f)\right]\right) \\
&+ 4A(f_0) \operatorname{Re}\left(\sum_{jk} \frac{1}{\sigma_{jk}(f_0)} \left[e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right]^* \times \right. \\
&\left. \sum_f \frac{1}{\sigma_{jk}(f)} W_{jk}(f-f_0) \left[\frac{\partial}{\partial A(f_1)} A^2(f) e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]} v_{jk}^{\mathrm{rel}}(f) - m_{jk}(f)\right]\right) \\
&= 4\delta_{f_0 f_1} \operatorname{Re}\left(\sum_{jk} \frac{1}{\sigma_{jk}(f_0)} \left[e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right]^* \times \right. \\
&\left. \sum_f \frac{1}{\sigma_{jk}(f)} W_{jk}(f-f_0) \left[A^2(f) e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]} v_{jk}^{\mathrm{rel}}(f) - m_{jk}(f)\right]\right) \\
&+ 8A(f_0)A(f_1) \operatorname{Re}\left(\sum_{jk} \frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)} W_{jk}(f_1-f_0) \left[e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right]^* \times \right. \\
&\left. \left[e^{i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]} v_{jk}^{\mathrm{rel}}(f_1)\right]\right)
\end{aligned}
\tag{106}
$$

Now taking the derivatives of the phase Jacobians we get

$$
\begin{aligned}
&\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial\Delta_x(f_0)\partial A(f_1)} \\
&= 4\delta_{f_0 f_1}A(f_0)\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)}\left[-ix_{jk}e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\right. \\
&\left.\times\sum_f\frac{1}{\sigma_{jk}(f)}W_{jk}(f-f_0)\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]\right) \\
&+4A^2(f_0)A(f_1)\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)}W_{jk}(f_1-f_0)\right. \\
&\left.\times\left[-ix_{jk}e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\left[e^{i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_1)\right]\right)
\end{aligned}
\tag{107}
$$

and

$$
\begin{aligned}
&\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial\Delta_y(f_0)\partial A(f_1)} \\
&= 4\delta_{f_0 f_1}A(f_0)\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)}\left[-iy_{jk}e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\right. \\
&\left.\times\sum_f\frac{1}{\sigma_{jk}(f)}W_{jk}(f-f_0)\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]\right) \\
&+4A^2(f_0)A(f_1)\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)}W_{jk}(f_1-f_0)\right. \\
&\left.\times\left[-iy_{jk}e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\left[e^{i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_1)\right]\right).
\end{aligned}
\tag{108}
$$

The phase-phase derivative is given by

$$
\begin{aligned}
&\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial\Delta_x(f_0)\partial\Delta_x(f_1)} \\
&= \frac{\partial}{\partial\Delta_x(f_1)}\left(\sum_{jk}\left[ix_{jk}A^2(f_0)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right]\times\right. \\
&\sum_f\frac{1}{\sigma_{jk}(f)\sigma_{jk}(f_0)}W_{jk}(f_0-f)\times \\
&\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]^* \\
&+\sum_{jk}\left[-ix_{jk}A^2(f_0)e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\times \\
&\sum_f\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f)}W_{jk}(f-f_0)\times \\
&\left.\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]\right) \\
&= \delta_{f_0f_1}\sum_{jk}\left[-x_{jk}^2A^2(f_0)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right]\times \\
&\sum_f\frac{1}{\sigma_{jk}(f)\sigma_{jk}(f_0)}W_{jk}(f_0-f)\times \\
&\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]^* \\
&+\delta_{f_0f_1}\sum_{jk}\left[-x_{jk}^2A^2(f_0)e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)\right]\times \\
&\sum_f\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f)}W_{jk}(f-f_0)\times \\
&\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right] \\
&+\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)}W_{jk}(f_0-f_1)\times \\
&x_{jk}^2A^2(f_0)A^2(f_1)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]-i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)v_{jk}^{\mathrm{rel}*}(f_1) \\
&+\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)}W_{jk}(f_1-f_0)\times \\
&x_{jk}^2A^2(f_0)A^2(f_1)e^{-i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]+i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}*}(f_0)v_{jk}^{\mathrm{rel}}(f_1).
\end{aligned} \tag{109}
$$

Using the fact that the weighting matrix is Hermitian, we get that

$$\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial\Delta_x(f_0)\partial\Delta_x(f_1)}$$
$$= 2\,\mathrm{Re}\left(\delta_{f_0 f_1}\sum_{jk}\left[-x_{jk}^2 A^2(f_0)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right]\times\right.$$
$$\sum_f \frac{1}{\sigma_{jk}(f)\sigma_{jk}(f_0)}W_{jk}(f_0-f)\times$$
$$\left.\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]^*\right)$$
$$+\,2\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)}W_{jk}(f_0-f_1)\times\right.$$
$$\left. x_{jk}^2 A^2(f_0)A^2(f_1)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]-i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)v_{jk}^{\mathrm{rel}*}(f_1)\right). \tag{110}$$

It follows that

$$\frac{\partial\chi^2[A(f),\Delta_x(f),\Delta_y(f)]}{\partial\Delta_y(f_0)\partial\Delta_y(f_1)}$$
$$= 2\,\mathrm{Re}\left(\delta_{f_0 f_1}\sum_{jk}\left[-y_{jk}^2 A^2(f_0)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)\right]\times\right.$$
$$\sum_f \frac{1}{\sigma_{jk}(f)\sigma_{jk}(f_0)}W_{jk}(f_0-f)\times$$
$$\left.\left[A^2(f)e^{i[\Delta_x(f)x_{jk}+\Delta_y(f)y_{jk}]}v_{jk}^{\mathrm{rel}}(f)-m_{jk}(f)\right]^*\right)$$
$$+\,2\,\mathrm{Re}\left(\sum_{jk}\frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)}W_{jk}(f_0-f_1)\times\right.$$
$$\left. y_{jk}^2 A^2(f_0)A^2(f_1)e^{i[\Delta_x(f_0)x_{jk}+\Delta_y(f_0)y_{jk}]-i[\Delta_x(f_1)x_{jk}+\Delta_y(f_1)y_{jk}]}v_{jk}^{\mathrm{rel}}(f_0)v_{jk}^{\mathrm{rel}*}(f_1)\right). \tag{111}$$

and

$$\frac{\partial \chi^2[A(f), \Delta_x(f), \Delta_y(f)]}{\partial \Delta_x(f_0) \partial \Delta_y(f_1)}$$

$$= 2\,\mathrm{Re}\left(\delta_{f_0 f_1} \sum_{jk} \left[-x_{jk} y_{jk} A^2(f_0) e^{i[\Delta_x(f_0) x_{jk} + \Delta_y(f_0) y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0)\right] \times \right.$$

$$\sum_f \frac{1}{\sigma_{jk}(f)\sigma_{jk}(f_0)} W_{jk}(f_0 - f) \times$$

$$\left.\left[A^2(f) e^{i[\Delta_x(f) x_{jk} + \Delta_y(f) y_{jk}]} v_{jk}^{\mathrm{rel}}(f) - m_{jk}(f)\right]^*\right)$$

$$+ 2\,\mathrm{Re}\left(\sum_{jk} \frac{1}{\sigma_{jk}(f_0)\sigma_{jk}(f_1)} W_{jk}(f_0 - f_1) \times \right.$$

$$\left. x_{jk} y_{jk} A^2(f_0) A^2(f_1) e^{i[\Delta_x(f_0) x_{jk} + \Delta_y(f_0) y_{jk}] - i[\Delta_x(f_1) x_{jk} + \Delta_y(f_1) y_{jk}]} v_{jk}^{\mathrm{rel}}(f_0) v_{jk}^{\mathrm{rel}*}(f_1)\right). \tag{112}$$