A Optimization for Multiplet Deconvolution

In this appendix we describe the mathematical formulation and solution of optimization for multiplet deconvolution. See Ref. [1] for background on the underlying mathematical methods. For convenience, we restate the key equations underlying this problem. Specifically, the objective is given by

\[ L_d = \sum_{j=1}^{M} \Delta \hat{f}_j^T \Xi^{-1} \Delta \hat{f}_j + \Delta \chi_j^T \Sigma^{-1} \Delta \chi_j, \quad (41) \]

whereas the constraint in the complex basis is

\[ \tilde{m} = \sum_j \Lambda^{-1}(c_j, v_j, R_j, \Delta t_j) [\tilde{f} + \Delta \hat{f}_j], \quad (42) \]

where \( \tilde{f} \) is the complex representation of \( f \). Our goal is to convert Eq. (42) to the same basis as \( \hat{f} \) and then simplify the optimization problem implied by Eq. (41).

Define \( \Lambda_j^{-1} = \Lambda^{-1}(c_j, v_j, R_j, \Delta t_j) \). Next, note that a vector \( \hat{f} \) has \( 2M + 1 \) elements. Because we compute them via a DFT, we take the convention that the first \( M + 1 \) modes correspond to \( k = 0, \pi, 2\pi, ..., M\pi \), whereas the last \( M \) modes correspond to \( k = -M\pi, -(M-1)\pi, ..., -\pi \). Moreover, because the signals are real in the time-domain, we know that \( \hat{f}(k) = \hat{f}(-k)^* \), where \( * \) denotes the complex conjugate. This implies that all of the relevant information about \( \hat{f} \) is contained in \( \hat{f} \), since only \( M \) complex Fourier modes are needed to describe the signal. To see this explicitly, express an arbitrary \( \hat{f} \) as

\[
\hat{f} = \begin{pmatrix}
a_1 \\ a_2 + ia_{M+2} \\ a_3 + ia_{M+3} \\ \vdots \\ a_{M+1} + ia_{2M+1} \\ a_{M+1} - ia_{2M+1} \\ a_M - ia_{2M} \\ \vdots \\ a_2 - ia_{M+2}
\end{pmatrix}, \quad (43)
\]

where \( a_j \) is the \( j \)th element of \( \hat{f} \).

To derive the transformed version of \( \Lambda \) in the basis of \( \hat{f} \), first decompose the matrix into blocks via

\[
\Lambda = \begin{bmatrix}
A & B \\
C & D
\end{bmatrix}, \quad (44)
\]

where \( A \) is \( (M + 1) \times (M + 1) \), \( B \) is \( (M + 1) \times M \), \( C \) is \( M \times (M + 1) \), and \( D \) is \( M \times M \). Clearly \( A \) couples the first \( M + 1 \) modes of \( \hat{f} \) into one another, \( B \) couples the remaining \( M \) modes into the first \( M + 1 \), and so forth. However, because the signal remains real in the time-domain after transformation by \( \Lambda \), knowledge of \( A \) and \( B \) is sufficient to determine the transformation matrix in the basis of \( \hat{f} \). Let \( F \) denote that operator that reverse the order of columns in a matrix, \( T_c (T_r) \) denote the operator that removes the first column (row) of a matrix, and \( 0_M \) denote a column vector with \( M \) zeros. Then it is straightforward to show that the operator \( \Lambda \) transforms to
\[
\Theta = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix},
\]

(45)

where

\[
\begin{align*}
\hat{A} &= \Re(A) + [0_{M+1}, \mathcal{F}(\Re(B))], \\
\hat{B} &= -\mathcal{T}_c(\Im(A)) + \mathcal{F}(\Im(B)), \\
\hat{C} &= \mathcal{T}_r(\Im(A)) + [0_M, \mathcal{F}(\Im(B))], \\
\hat{D} &= \mathcal{T}_r(\Re(A)) - \mathcal{F}(\Re(B)),
\end{align*}
\]

(46)

and \(\Re\) and \(\Im\) denote the real and imaginary components. Thus we arrive at an expression for \(\hat{m}\) expressed in the basis of \(\hat{f}\), viz,

\[
\hat{m} = \sum_j \Theta^{-1}_j [\hat{f} + \Delta \hat{f}_j].
\]

(47)

Minimizing Eq. (41) subject to Eq. (47) may entail optimizing over \(O(100)\) or more variables corresponding to: (i) the scale parameters \(c, v, R\), and \(\Delta t\); and (ii) the real and imaginary parts of the mode-weights. The latter comprise the majority of variables, although they only appear up to second order. In contrast, the transformation variables, while few in number, appear in highly non-linear function associated with the matrix \(\Theta^{-1}\) in Eq. (47). Further compounding these issues is the fact that both \(\Theta^{-1}\) and \(\Xi\) are dense matrices, the latter possibly having eigenvalues close to zero. This may yield a relatively large numerical problem that is poorly scaled, and thus challenging to solve.

Fortunately, the constraint given by Eq. (42) is linear in the mode-weights, which yields a key simplification. Without loss of generality, one finds

\[
\Delta \hat{f}_1 = \Theta_1 \left[ \hat{m} - \Theta^{-1}_1 \hat{f} - \sum_{j=2}^M \Theta^{-1}_j [\hat{f} + \Delta \hat{f}_j] \right].
\]

(48)

Equation (48) can be substituted into Eq. (39) and minimization performed over the remaining modes \(\Delta \hat{f}_j\) for \(j \geq 2\) for fixed transformation parameters associated with the \(\Theta_j\). We leave this exercise for the reader. For the case of doublets, one find that

\[
\begin{align*}
G &= \Theta_1 [\hat{m} - (\Theta^{-1}_1 + \Theta^{-1}_2)\hat{f}], \\
\Delta \hat{f}_2 &= \left[ (\Theta_1 \Theta_2^{-1})^T \Xi^{-1} \Theta_1 \Theta_2^{-1} + \Xi^{-1} \right]^{-1} \Theta_1 \Theta_2^{-2} \Xi^{-1} G, \\
\Delta \hat{f}_1^* &= \Theta_1 [G - \Theta_2^{-1} \Delta \hat{f}_2],
\end{align*}
\]

where \(\Delta \hat{f}_2^*\) and \(\Delta \hat{f}_1^*\) are the optimal mode perturbations (the * is distinct from the complex conjugate *). Having the \(\Delta \hat{f}_j\) in terms of the transformation parameters (via the \(\Theta_j\)), we may then express the objective as

\[
\mathcal{L}_d = \sum_j \left[ \Delta \hat{f}_j^* \right]^T \Xi^{-1} \Delta \hat{f}_j^* + \Delta \chi_j^T \gamma^{-1} \Delta \chi_j.
\]

(49)

This \(\mathcal{L}_d\) can then be optimized as a function of the scale transformations parameters.
Reference

1. Arfken GB, Weber HJ, Harris FE. Mathematical Methods for Physicists: A Comprehensive Guide. Elsevier Science; 2013.