Supplement II: Spectral Reconstruction Methods

Derek M. Kita[1] and Brando Miranda[2]

[1]Department of Materials Science & Engineering, Massachusetts Institute of Technology
[2]Center for Brains, Minds, and Machines, Massachusetts Institute of Technology

February 1, 2018

Given the measured interferogram y (of size $N \times 1$) and a calibration matrix A ($N \times D$), we seek to accurately reconstruct the input optical signal x that obeys:

$$y = Ax \tag{1}$$

where $D \gg N$, and in our case $D = 801$, $N = 64$. For our 64-channel device, there are two types of signals available to us for testing the quality of optical reconstruction: (1) laser lines that produce sparse spectra, and (2) broadband sources (like an EDFA) with a broad spectrum (not-sparse).

Since the problem we are solving is underconstrained, there are infinite solutions x that solve Eq. 1. However, we can place constraints on the sparsity and magnitude of the spectrum and prevent over-fitting issues by minimizing the L1- and L2-norm of x:

$$\min_{x} \left\{ ||y - Ax||^2 + \alpha_1 ||x||_1 + \alpha_2 ||x||_2^2 \right\} \tag{2}$$

where α_1 and α_2 are the corresponding hyperparameters. For an arbitrary optical input, we find that the "smoothness" of the spectrum is an important characteristic of the spectra and a good regularizer. To induce such smoothness, characterized by the first-derivative of the spectrum, we used the finite difference matrix D to define the following regularizer $\|Dx\|_2^2$. The matrix D is defined as follows:

$$D = \begin{bmatrix} -1 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & -1 & 1 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & -1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ldots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & -1 \end{bmatrix}$$

Also note that this method can be easily generalized to spectral reconstruction on a non-equidistant grid by simply taking the real distance between points. We cast our reconstruction problem with L1-norm, L2-norm and the first-derivative smoothness prior as follows:

$$\min_{x, x>0} \left\{ ||y - Ax||^2 + \alpha_1 ||x||_1 + \alpha_2 ||x||_2^2 + \alpha_3 \|Dx\|_2^2 \right\} \tag{3}$$

Using $\|Mx\|_2^2 = x^\top Mx$, and the fact that our spectrum is non-negative (and thus $\|x\|_1 = \sum_{d=1}^{D} x_i = \mathbf{1}^\top x$), we may rewrite Equation 3 as a non-negative quadratic program:

$$\min_{x, x>0} \left\{ x^\top \left(A^\top A + \alpha_2 I + \alpha_3 D^\top D \right) x + (\alpha_1 \mathbf{1} - A^\top y)^\top x \right\} \tag{4}$$

The above form is easily computed with standard quadratic program solvers. With this method of solving for the signal x, the last step is to determine suitable hyperparameters α_1, α_2, and α_3 that correspond to the correct input. However, since we assume no prior information about our input signal, we use a standard leave-one-out cross-validation technique, which requires only two successive measurements of the interferogram, characterized by the same input signal with different noise. With two independent measurements y_1 and y_2 of the same source, and given two measurements of the basis A_1 and A_2 (performed only once in advance as a calibration step), we solve for x_1 via Equation 4 for a suitably large range of hyperparameter values, and arguments y_1 and A_2. We then choose the x_1 corresponding to the unique set of α's that maximizes the coefficient of determination between the second measurement y_2 and the value $A_2 \cdot x_1$.

$$\max_{\alpha_{1,2,3}} \left\{ R^2(y_2, A_2 \cdot x_1) \right\} \tag{5}$$