The Path to the Optimal Filter

Eli Cohen

McGill University

October 29, 2024

Presentation Overview

1. Data
2. Filters
   - Wiener Filter
   - Matched Filter
   - Debiased Matched Filter
   - Optimal Filter
   - Filter Summary
3. Performance

Discrete Time Signals

Definition

A discrete time signal is a sequence of samples defined only at discrete points in time. For example, if x is a discrete time signal, $x[n]$ denotes the sample at time index n

Definition

A filter is Linear Time-Invariant (LTI) if it exhibits both linearity and time-invariance, meaning its output is directly proportional to its input and it responds the same way at any point in time.

Linearity

A filter, T, is linear if, when acting on inputs $x_1[n]$ and $x_2[n]$ with constants a, and b, it obeys the equation
$T\{ax_1[n] + bx_2[n]\} = aT\{x_1[n]\} + bT\{x_2[n]\}$.

Time Invariance

A filter, T, is time invariant if, for input $x[n]$, output $y[n] = T\{x[n]\}$, and time shift n_0, $T\{x[n - n_0]\} = y[n - n_0]$.

Data

Theorem

The output of any LTI filter, $y[n]$ can be represented as the linear convolution of an impulse response, $h[n]$, with and input $x[n]$:
$y[n] = (x * h)[n] = \sum_{m=-\infty}^{\infty} x[m]h[n-m]$.

Definition

A filter is called FIR if for some constants n_1 and n_2 with $n_1 < n_2$, if $n < n_1$ and $n > n_2$, $h[n] = 0$. A FIR filter can be represented by a finite convolution: $y[n] = \sum_{m=0}^{N-1} x[m]h[n-m]$ where N is the length of $h[n]$.

The ATLAS Liquid-Argon Calorimeter (LAr) is used to reconstruct the energy of charged particles created through the proton-proton collisions at the LHC. When a particle passes through one of the cells in the calorimeter a peak in voltage is created. This voltage is proportional to the energy of the particle which allows for the reconstruction of energy from voltages. The peak is then shaped by analog electronics (a band-pass CR-RC^2 filter) and it is subsequently sampled at 40 MHz and digitized.

Of important note is that the analog pulse shaping is an LTI and FIR process. Although the shaping occurs in the continuum we can consider the sampled version and write the following.

The digital signal produced by the LAr calorimeter is denoted $d[n]$. We can write $d[n]$ in terms of the delayed energies of the particles (in order to gather more information about the deposition), $s[n] = E[n - n_d]$, where $E[n]$ is the true energy, convolved with a known pulse shape, $h[n]$. with the addition of zero-mean noise, $n[n]$: $d[n] = (h * s)[n] + n[n]$.

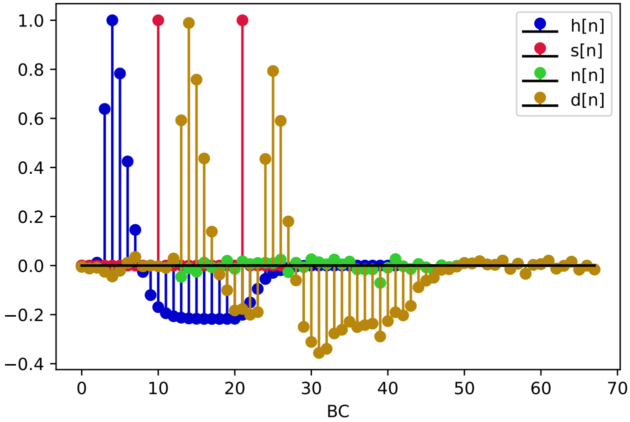


Figure: An example of a possible calorimeter output.

For this presentation we use data corresponding to one calorimeter cell using simulated using AREUS. In particular the data has the following specifications

- Calorimeter cell located in the middle of the LAr electromagnetic barrel
- $\eta = 0.5125$ (Specifies location)
- $\phi = 140$ (Specifies location)
- $\mu = 140$ (Specifies out of time pileup)
- Random distribution of target signals (In both time and amplitude)

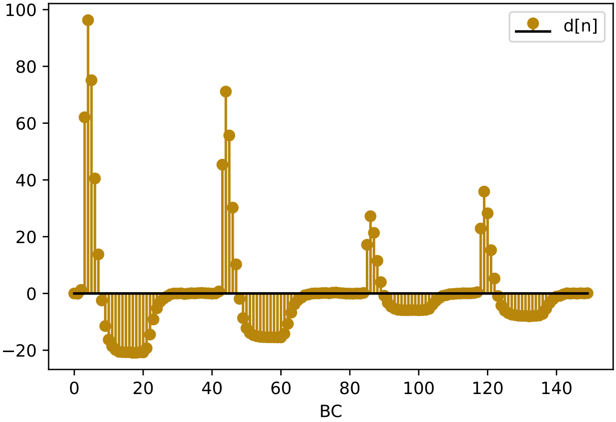


Figure: The first 200 samples of the simulated $d[n]$ dataset.

Temporal Wiener Filter 1

The Wiener filter is a FIR implementation seeks to minimize the mean squared error, MSE, of P coefficients, denoted $c[n]$, and the data $d[n]$, $\hat{s}[n] = (c * d)[n] = \sum_{m=0}^{P-1} c[m]d[n-m]$:

$$MSE = \mathbb{E}\left\{(\hat{s}[n] - s[n])^2\right\} = \mathbb{E}\left\{(e[n])^2\right\} \tag{1}$$

This is done by taking the derivative of the MSE wrt to the filter coefficients and setting it to 0 to find a minimum. The resulting expression is

$$\mathbb{E}\left\{e[n]d[n-m]\right\} = 0 \text{ or } R_{ed}[m] = 0 \tag{2}$$

Here $R_{ed}[m]$ denotes the cross-correlation sequence between the error, e, and d. We can rewrite $R_{ed}[m]$ as

$$\begin{aligned} R_{ed}[m] &= \mathbb{E}\left\{e[n]d[n-m]\right\} \\ &= \mathbb{E}\left\{(\hat{s}[n] - s[n])d[n-m]\right\} \\ &= R_{\hat{s}d}[m] - R_{sd}[m] \end{aligned}$$

Temporal Wiener Filter 2

We find that $R_{\hat{s}d}[m] = R_{sd}[m]$ for all m. Therefore, for the Wiener filter we seek, the cross-correlation between the result of the estimator, $\hat{s}[n] = (c * d)[n]$, and the data, $d[n]$, matches the cross-correlation between the true signal, $s[n]$, and the data $d[n]$.

Since $\hat{s}[n]$ is generated by filtering $d[n]$ through an LTI system with a response $c[n]$ we find that

$$R_{\hat{s}d}[m] = (c * R_{dd})[m] \tag{3}$$

This allows us to retrieve the coefficients,

$$(c * R_{dd})[m] = R_{sd}[m] \tag{4}$$

Frequency Wiener Filter

We take the z-transform of Eq. 4 we find

$$C(z) = S_{sd}(z)/S_{ss}(z) \tag{5}$$

where $S_{ss}(z)$ is the transform of the autocorrelation R_{dd} and $S_{ds}(z)$ is the transform of the R_{yx} cross correlation. We can recognize S_{SD} as the cross power spectral density between s and d, and S_{ss} as the PSD of the input signal.

This is the frequency space representation of the Wiener filter. Using the fact that the noise and signal are uncorrelated we can rewrite the coefficients, $C(z)$ in a more familiar form:

$$C(z) = \frac{H^*(z) \cdot S_{ss}(z)}{|H(z)|^2 S_{ss}(z) + S_{nn}(z)} \tag{6}$$

Here $S_{ss}(z)$ and $S_{nn}(z)$ are the PSDs of the signal and noise respectively, and $H(z)$ is the transform of the system's impulse response, $h[n]$.

Linear Algebra Wiener Filter 1

In the linear algebra domain we can rewrite a convolution, $y[n] = x[n] * h[n]$ as the result of a dot product:

$$y[n] = x[n] * h[n] \leftrightarrow \mathbf{y} = \mathbf{x}^\mathsf{T}\tilde{\mathbf{h}} \tag{7}$$

Here a^T represents a transposition and $\tilde{a}$ is the reversing of vector indices with respect to sequence indices. Of critical importance is that we must express either $\mathbf{x}^\mathsf{T}$ or $\tilde{\mathbf{h}}$ as a square (PxP) circulant matrix with each row consisting of a segment of the data time shifted by one with each increasing row.
We can express Eq. 4 as follows:

$$\tilde{\mathbf{c}} = \mathbf{R_{dd}}^{-1}\tilde{\mathbf{R}}_{\mathbf{sd}} \tag{8}$$

Linear Algebra Wiener Filter 2

$$\mathbf{R_{dd}} = (R_{dd_{ij}}) \in \mathbb{R}^{P\times P}, \quad \tilde{\mathbf{R_{sd}}} = (R_{sd_i}) \in \mathbb{R}^{P}, \quad \tilde{\mathbf{c}} = (c_i) \in \mathbb{R}^{P} \tag{9}$$

$$R_{dd_{ij}} := R_{dd}[j-i], \quad \tilde{R_{sd_i}} := R_{sd}[-i], \quad \tilde{c}_i := c[P-1-i] \tag{10}$$

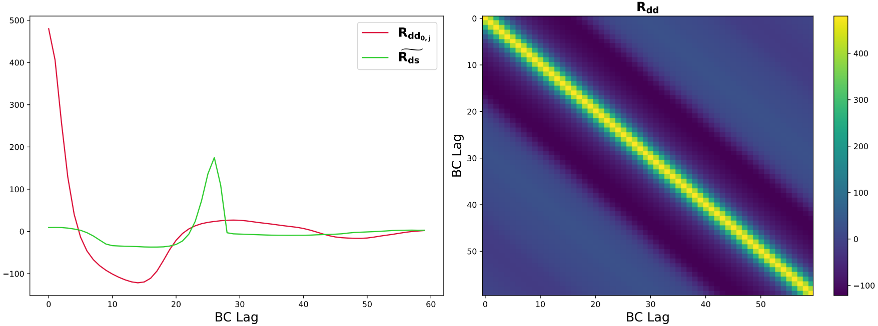


Figure: 60 coefficient values for $\mathbf{R_{dd}}$ and $\tilde{\mathbf{R_{sd}}}$ with a delay $n_d = 25$

Linear Algebra Wiener Filter 3

In the time domain, we can derive an equivalent formula to Eq. 6 for the coefficients (assuming that the signal depositions are uncorrelated and the signal and the noise are uncorrelated):

$$\tilde{\mathbf{c}} = (\mathbf{H}\sigma_s^2 + \mathbf{R_{nn}})^{-1}\mathbf{h}\sigma_s^2 \tag{11}$$

$$\mathbf{H} = (h_{ij}) \in \mathbb{R}^{P\times P}, \quad \mathbf{R_{nn}} = (R_{nn_{ij}}) \in \mathbb{R}^{P\times P}, \quad \mathbf{h} = (h_i) \in \mathbb{R}^{P}, \tag{12}$$

$$h_{ij} := (h \star h)[j-i], \quad R_{nn_{ij}} := R_{nn}[j-i], \quad h_i := h[i-P-1+d] \tag{13}$$

Linear Algebra Wiener Filter 4

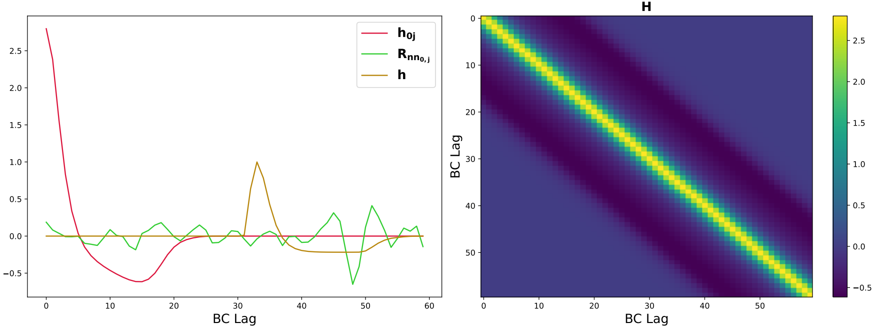


Figure: 60 coefficient values for $\mathbf{H}$, $\mathbf{h}$ and $\tilde{\mathbf{R}}_{\mathbf{nn}}$.

Wiener Filter Coefficients

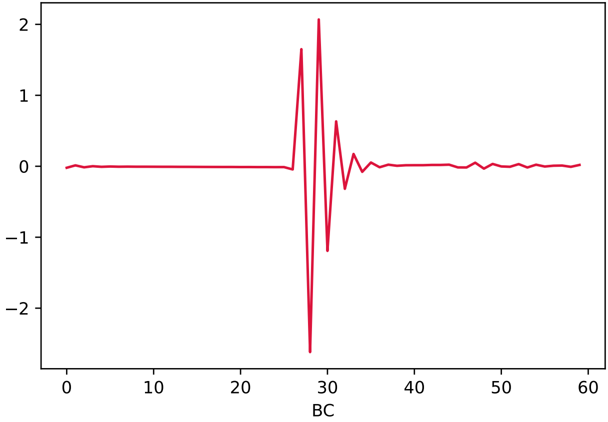


Figure: 60 coefficient values for the Wiener Filter with a delay $n_d = 25$.

Linear Algebra Matched Filter 1

If we are interested in reconstructing *just* the deposited energies, we can employ a matched filter that minimizes the error in the BCs where we assume a hit to be (as opposed to the Wiener filter that minimizes error in *all* BCs). This decreases variance in cells where energy has actually been deposited, but results in bias around said BC. As mentioned n_d (which denotes a delay of n_d samples) now appears explicitly

$$\tilde{\mathbf{c}} = \mathbf{A}^{-1}\tilde{\mathbf{v}} \tag{14}$$

$$\mathbf{A} = (a_{ij}) \in \mathbb{R}^{P\times P}, \quad \tilde{\mathbf{v}} = (\tilde{v}_i) \in \mathbb{R}^{P}, \quad \tilde{\mathbf{c}} = (\tilde{c}_i) \in \mathbb{R}^{P}, \tag{15}$$

$$\begin{aligned} a_{ij} &:= \mathbb{E}\left[d[i+n_d-P+1]d[j+n_d-P+1]\right], \\ \tilde{v}_i &:= \mathbb{E}\left[d[i+n_d-P+1]s[n_d]\right], \\ \tilde{c}_i &:= c[P-1-i] \end{aligned} \tag{16}$$

Linear Algebra Matched Filter 2

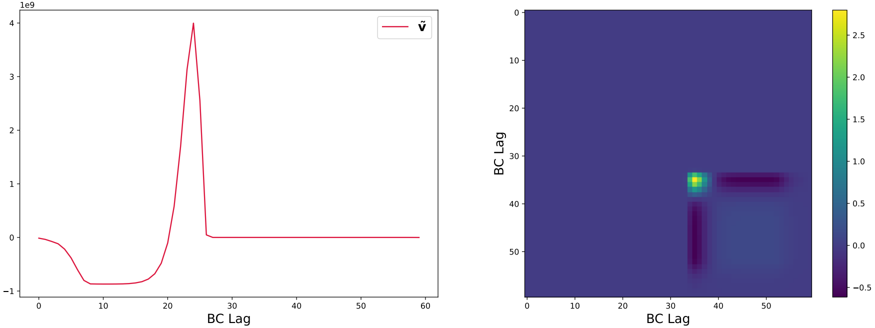


Figure: 60 coefficient values for **A**, and $\tilde{\mathbf{v}}$.

Linear Algebra Matched Filter 3

We can also generate coefficients for the matched filter using $\mathbf{R_{nn}}$ and $\mathbf{h}$:

$$\tilde{\mathbf{c}} = \frac{\left(\mu_s^2 + \sigma_s^2\right) \mathbf{R_{nn}}^{-1} \mathbf{h}}{1 + \left(\mu_s^2 + \sigma_s^2\right) \mathbf{h}^\top \mathbf{R_{nn}}^{-1} \mathbf{h}}. \tag{17}$$

Matched Filter Coefficients

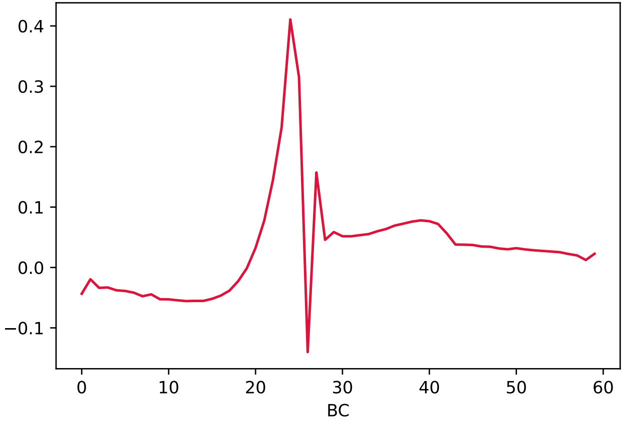


Figure: 60 coefficient values for the matched filter with a delay $n_d = 25$.

Linear Algebra Debiased Matched Filter 1

If we minimize the MSE of the error (in cells where energy deposition occurs) subject to the constraint that the error has zero bias (using Lagrange multipliers) we recover the following coefficients:

$$\tilde{\mathbf{c}} = \frac{\mathbf{R_{nn}}^{-1}\mathbf{h}}{\mathbf{h}^{\mathsf{T}}\mathbf{R_{nn}}^{-1}\mathbf{h}} \tag{18}$$

We can also identify the unbiased matched filter in the limit as $\mu_s^2 + \sigma_s^2 >> \sigma_n^2$: where the contribution of our signal dominates the noise.

Debiased Matched Filter Coefficients

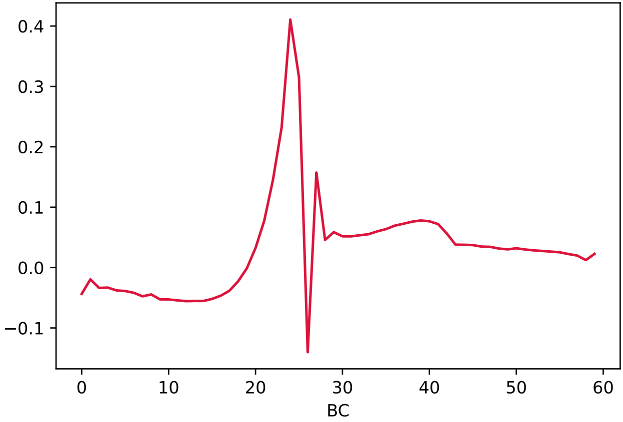


Figure: 60 coefficient values for the debiased matched filter with a delay $n_d = 25$.

Linear Algebra Optimal Filter 1

We can be more specific with the shape of our known signal. Since energy depositions may not exactly align with when the signal is sampled, we can introduce a parameter τ to reflect this shift in time. Now, one of our signals may look like $S_i = Ah(t_i - \tau)$ where A is the amplitude of the signal, h is our impulse response and t_i are the sampled times. The Optimal Filter relies on a first-order approximation to this signal, $S_i = Ah_i - A\tau g_i' + n_i$ where the prime denotes a derivative. The coefficients for the optimal filter can be obtained by plugging the hit-timing term into the unbiased matched filter.

Linear Algebra Optimal Filter 2

We obtain two sets of coefficients. One to estimate the amplitude of the signal, **a** and the other to estimate the hit timing, **b**

$$\tilde{\mathbf{a}} := \frac{Q_2 \mathbf{R_{nn}}^{-1}\mathbf{h} - Q_3 \mathbf{R_{nn}}^{-1}\mathbf{h}'}{Q_1 Q_2 - Q_3^2}, \quad \tilde{\mathbf{b}} := \frac{Q_3 \mathbf{R_{nn}}^{-1}\mathbf{h} - Q_1 \mathbf{R_{nn}}^{-1}\mathbf{h}'}{Q_1 Q_2 - Q_3^2} \tag{19}$$

$$Q_1 := \mathbf{h}^\top \mathbf{R_{nn}}^{-1}\mathbf{h}, \quad Q_2 := \mathbf{h}'^\top \mathbf{R_{nn}}^{-1}\mathbf{h}', \quad Q_3 := \mathbf{h}'^\top \mathbf{R_{nn}}^{-1}\mathbf{h} \tag{20}$$

In the limit where Q_3 goes to 0 we recover the unbiased matched filter. This is the case where the underlying signal and the timing offset have little interaction.

Optimal Filter Coefficients

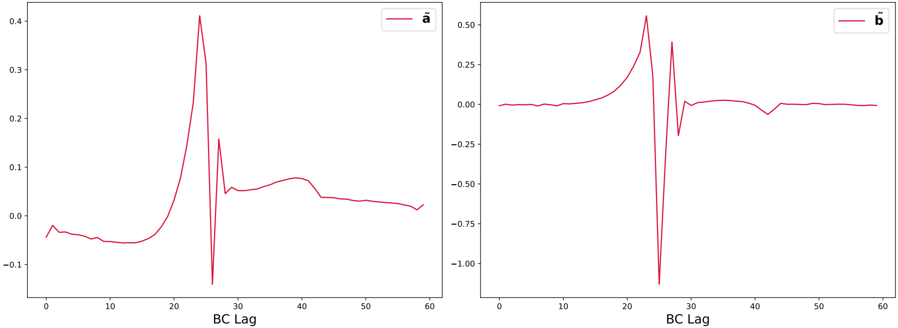


Figure: 60 coefficient values for the two sets of optimal filter coefficients, $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$, with a delay $n_d = 25$.

Filter Comparison

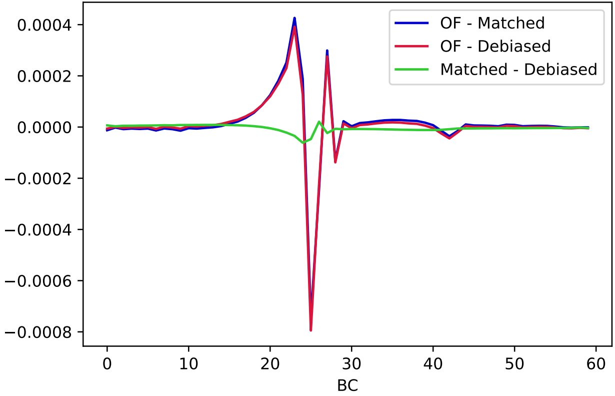
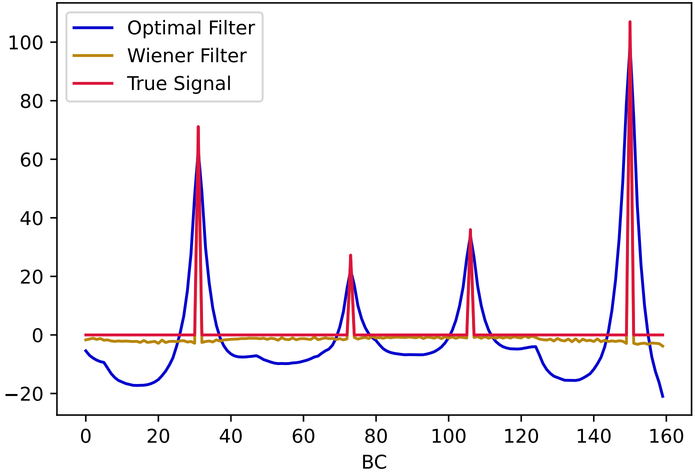


Figure: A comparison of different residuals for the various matched filters. In practice, with the signals we are considering, the difference between the filters are negligible, so only the Optimal Filter will be used to evaluate performances.

Filter Concluson

We have discussed have two classes of filters. One aimed at reconstructing the entire signal and the other aimed at reconstructing only BCs of interest. The former provides a poor reconstruction of energy but better identification of BC, whereas in the latter there is better reconstruction of energy but due to bias there is poor resolution of the temporal component. Moreover, filters in the latter category struggle to disentangle depositions in quick succession, leading to issues with out-of-time pileup.

Filter Output



Max Finder

In order to draw conclusions about the performance of the filter we employ a non-linear filter to pick out peaks in the data. If a sample is greater than a given threshold (T=0.25 GeV) and it is greater than the two surrounding samples we return the value of the sample, otherwise we return zero.

$$x_{mf}[n] = \begin{cases} x[n], & \text{if } x[n] > \text{T and } x[n] > x[n-1] \text{ and } x[n] > x[n+1] \\ 0, & \text{else} \end{cases} \tag{21}$$

Max Finder Output