Data Unfolding with Wiener-SVD Method

arXiv:1705.03568

W. Tang,[a,1] ***X. Li,***[b,1] X. Qian,[a,2] H. Wei,[a] C. Zhang[a]

[a] Physics Department, Brookhaven National Laboratory, Upton, NY, USA

[b] State University of New York at Stony Brook, Department of Physics and Astronomy, Stony Brook, NY, USA

xiaoyue.li@stonybrook.edu

DPF – FNAL, Aug. 2nd, 2017

Outline

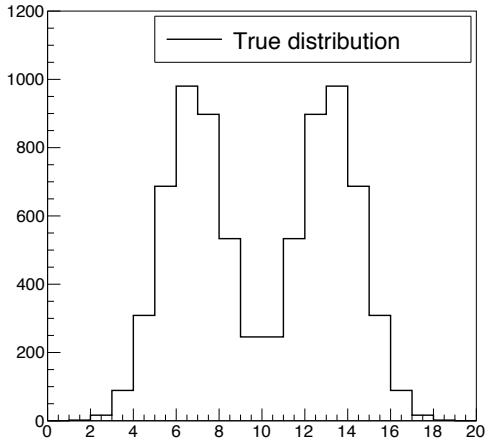
- The unfolding problem
- Wiener filter in digital signal processing
- SVD unfolding with Wiener filter
- Data unfolding example
  - Cross-section
  - Reactor neutrino flux
- Summary and recommendation

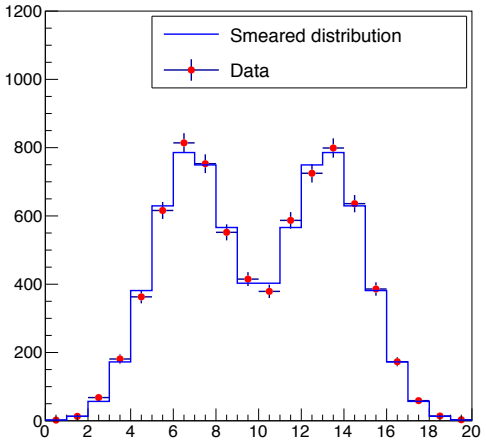
The data unfolding problem

- The problem:

$$f_{smear}(x) = \int R(x|y) f_{true}(y)\, \mathrm{d}y$$

  - $R(x|y)$: response function
  - $m_i = \sum_j R_{ij} s_j$
  - $R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j)$



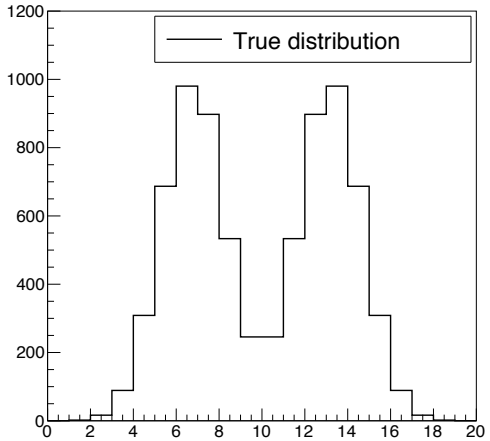


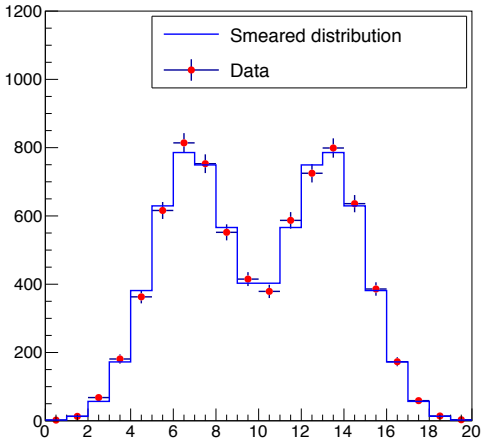
The data unfolding problem

- The problem:

$$f_{smear}(x) = \int R(x|y) f_{true}(y)\, \mathrm{d}y$$

  - $R(x|y)$: response function
  - $m_i = \sum_j R_{ij} s_j$
  - $R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j)$
- Why not do $\boldsymbol{s} = R^{-1} \cdot \boldsymbol{m}$?



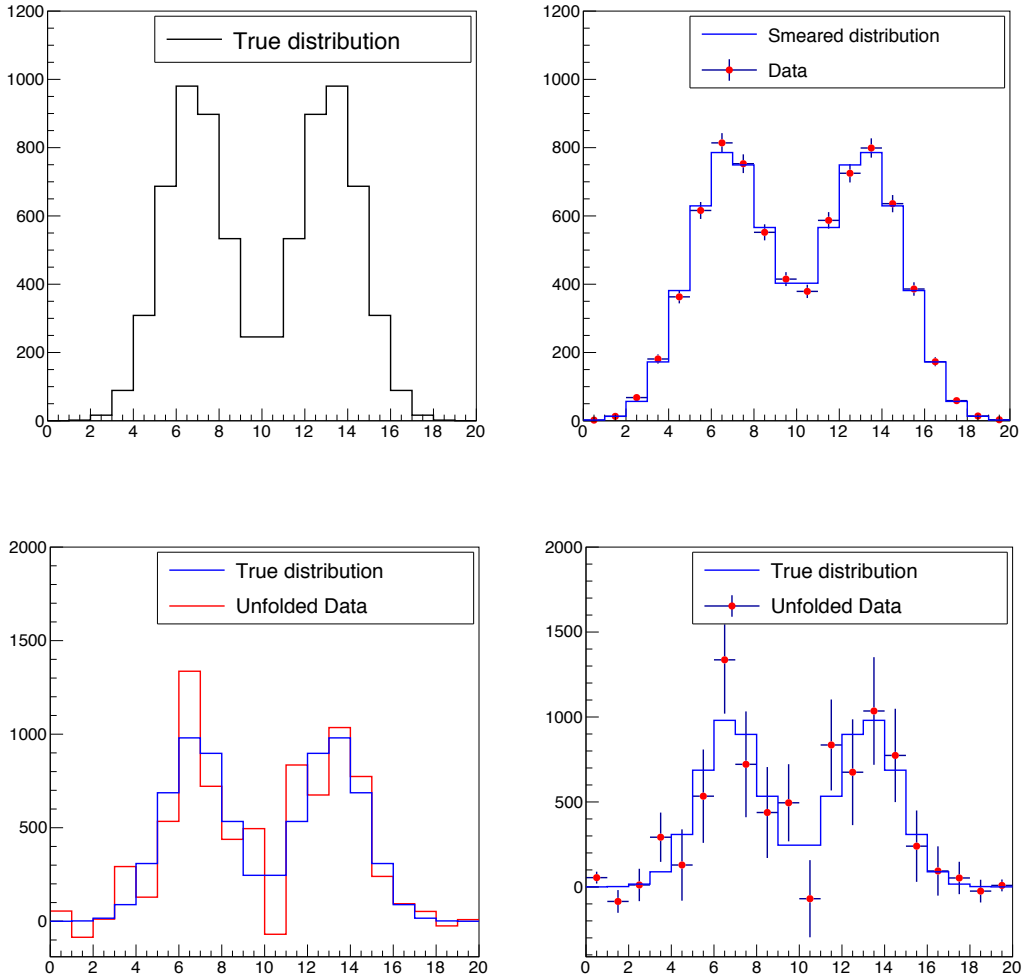


The data unfolding problem

- The problem:

$$f_{smear}(x) = \int R(x|y) f_{true}(y)\, \mathrm{d}y$$

  - $R(x|y)$: response function
  - $m_i = \sum_j R_{ij} s_j$
  - $R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j)$
- Why not do $\boldsymbol{s} = R^{-1} \cdot \boldsymbol{m}$?

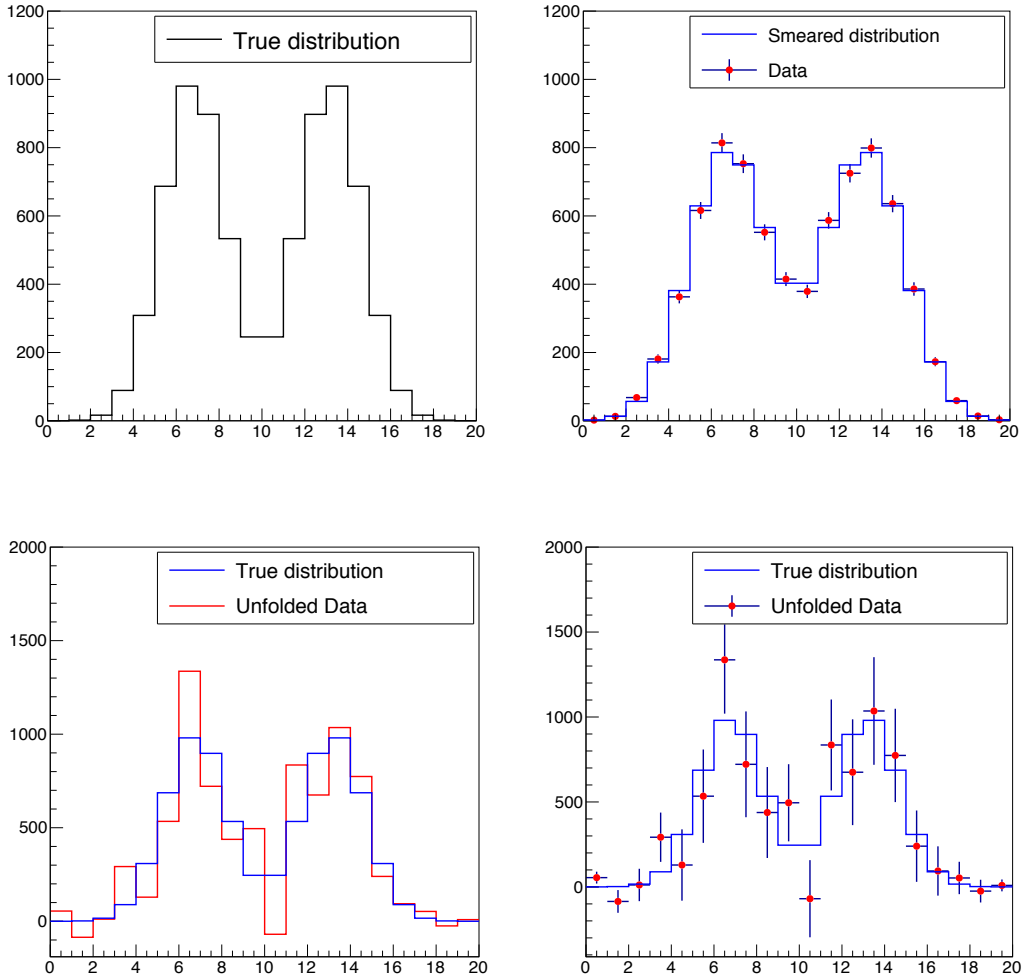


The data unfolding problem

- The problem:

$$f_{smear}(x) = \int R(x|y) f_{true}(y)\, \mathrm{d}y$$

  - $R(x|y)$: response function
  - $m_i = \sum_j R_{ij} s_j$
  - $R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j)$
- Why not do $\boldsymbol{s} = R^{-1} \cdot \boldsymbol{m}$?
- Due to statistical fluctuation, the unfolded spectrum from direct inversion of response matrix bares no resemblance to the true spectrum.

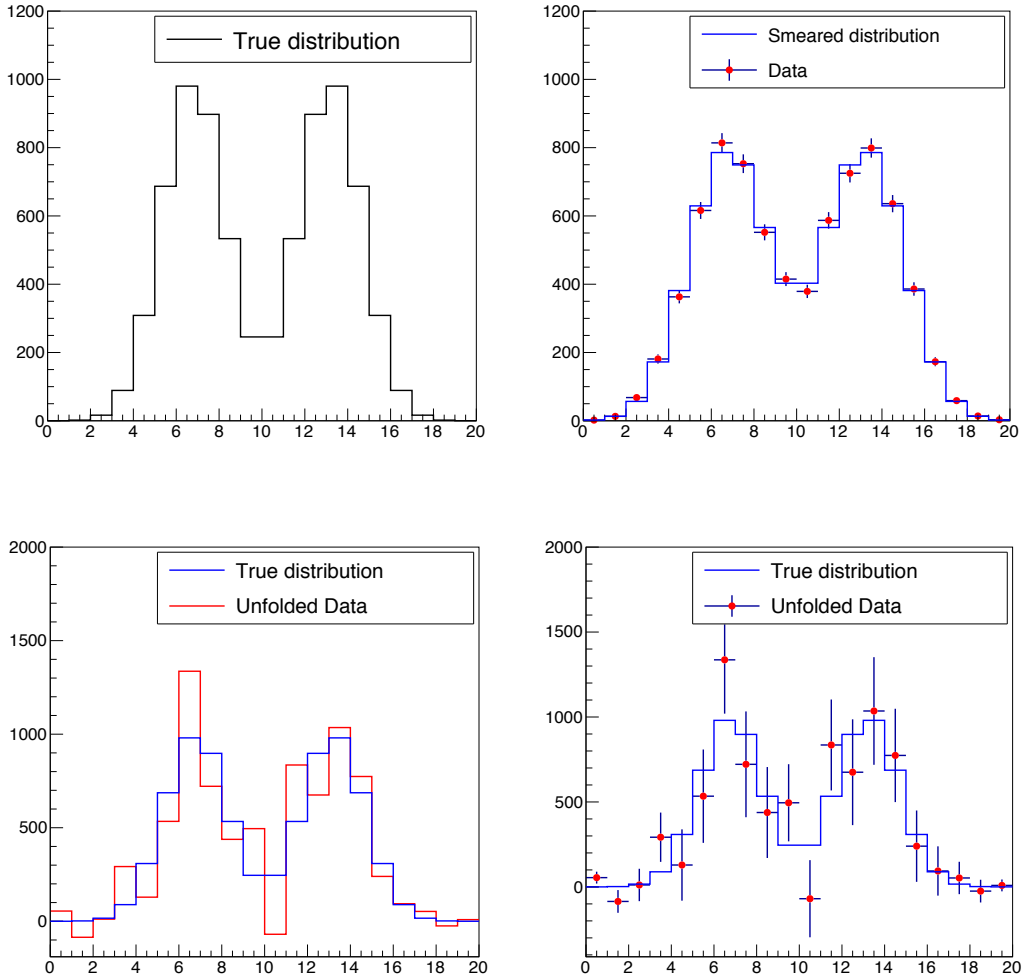


The data unfolding problem

- The problem:

$$f_{smear}(x) = \int R(x|y) f_{true}(y)\, \mathrm{d}y$$

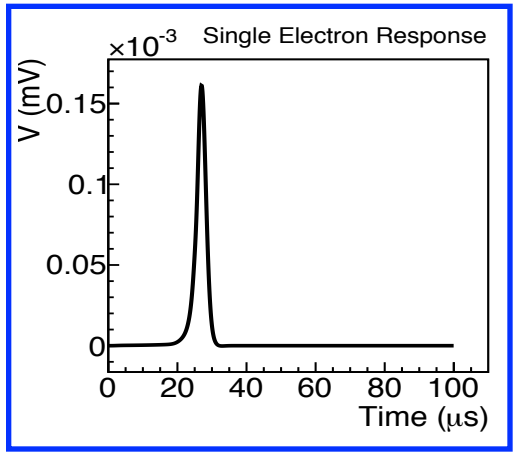
  - $R(x|y)$: response function
  - $m_i = \sum_j R_{ij} s_j$
  - $R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j)$
- Why not do $\boldsymbol{s} = R^{-1} \cdot \boldsymbol{m}$?
- Due to statistical fluctuation, the unfolded spectrum from direct inversion of response matrix bares no resemblance to the true spectrum.
- Introduce additional constraints.
  - E.g. trade bias for smoothness
  - Bayesian analysis



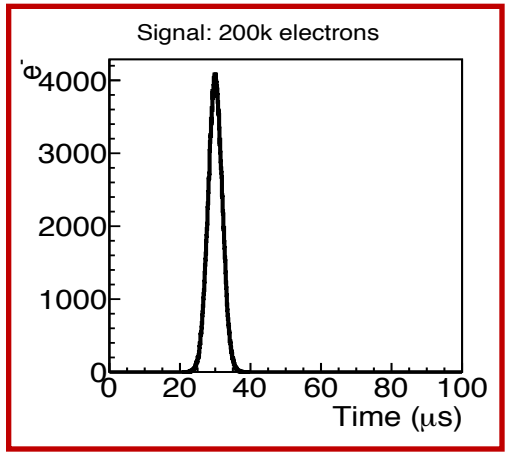
Wiener filter in digital signal processing (1)

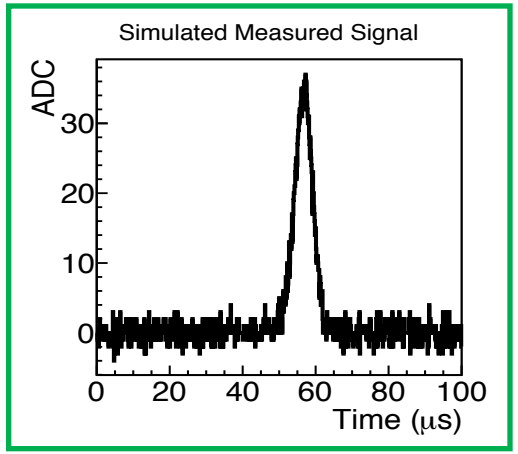
- Deconvolution problem:
  - $M(t') = \int_{-\infty}^{\infty} R(t,t') \cdot S(t)\, \mathrm{d}t$
  - True signal $S(t)$, measured signal $M(t')$, response function $R(t,t') \equiv R(t-t')$

Response function $R(t,t')$



True signal $S(t)$





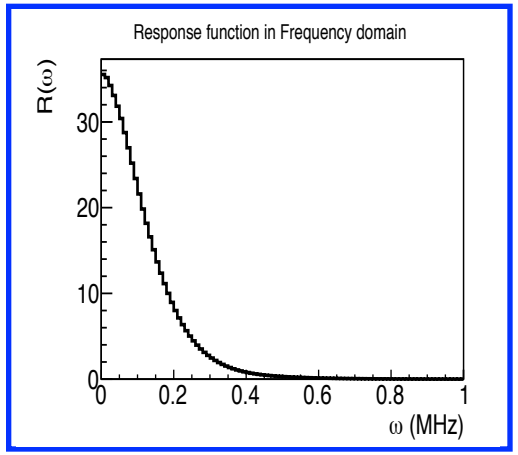
Measured signal, with statistical fluctuation

Wiener filter in digital signal processing (1)

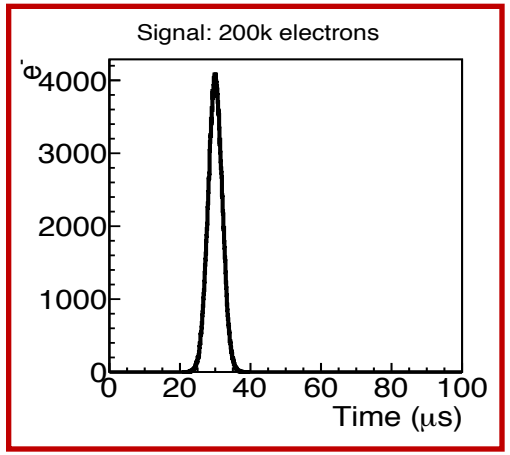
- Deconvolution problem:
  - $M(t') = \int_{-\infty}^{\infty} R(t,t') \cdot S(t)\, \mathrm{d}t$
  - True signal $S(t)$, measured signal $M(t')$, response function $R(t,t') \equiv R(t-t')$
- Fourier transform:

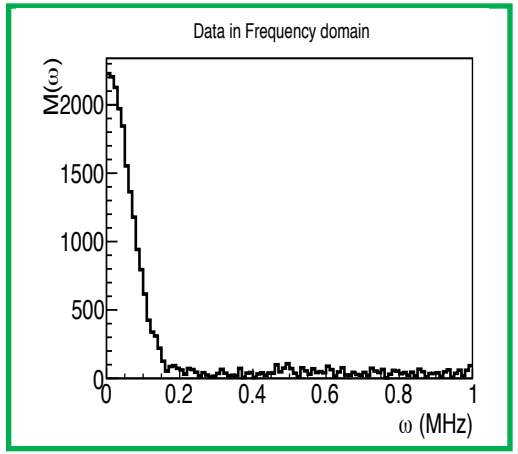
$$M(\omega) = R(\omega) \cdot S(\omega) \rightarrow S(\omega) = M(\omega)/R(\omega)$$

Response function $R(\omega)$



True signal $S(t)$





Measured signal in frequency domain $M(\omega)$

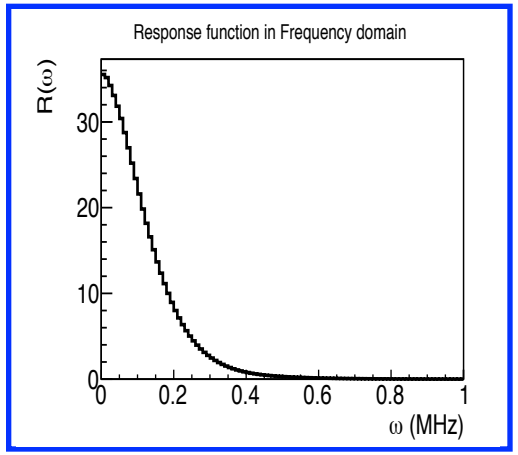
Wiener filter in digital signal processing (1)

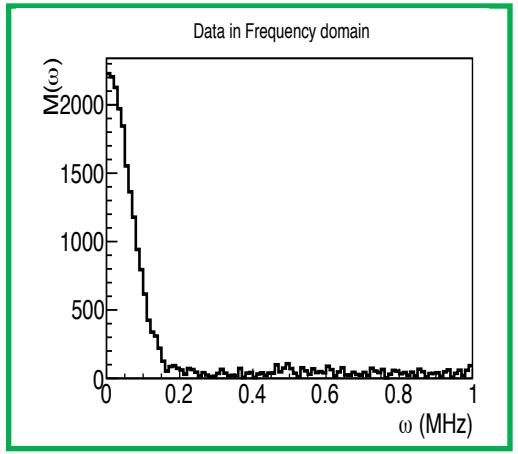
- Deconvolution problem:
  - $M(t') = \int_{-\infty}^{\infty} R(t,t') \cdot S(t)\, \mathrm{d}t$
  - True signal $S(t)$, measured signal $M(t')$, response function $R(t,t') \equiv R(t-t')$
- Fourier transform:

$$M(\omega) = R(\omega) \cdot S(\omega) \rightarrow S(\omega) = M(\omega)/R(\omega)$$

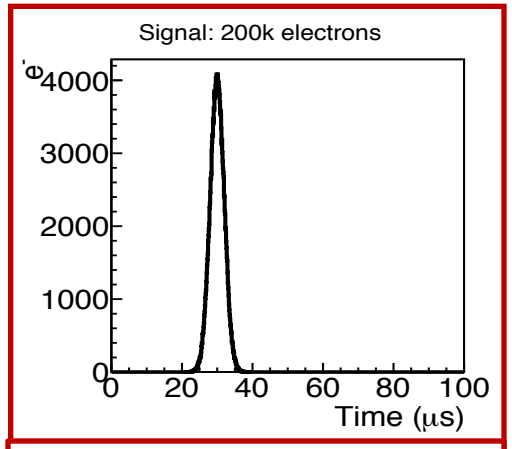
$$\rightarrow \text{Inverse FFT } S(\omega) \rightarrow S(t)$$

Response function $R(\omega)$

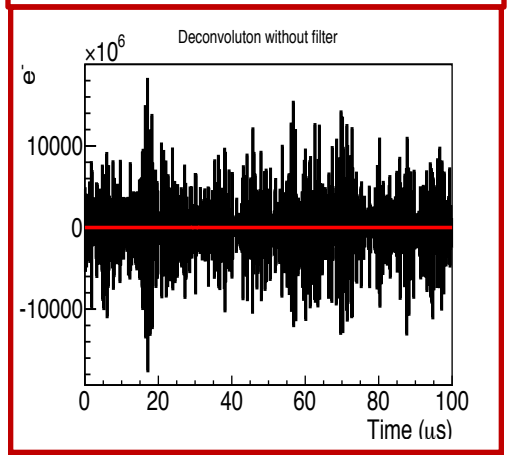




True signal $S(t)$



Deconvoluted signal $S(t)$



Wiener filter in digital signal processing (1)

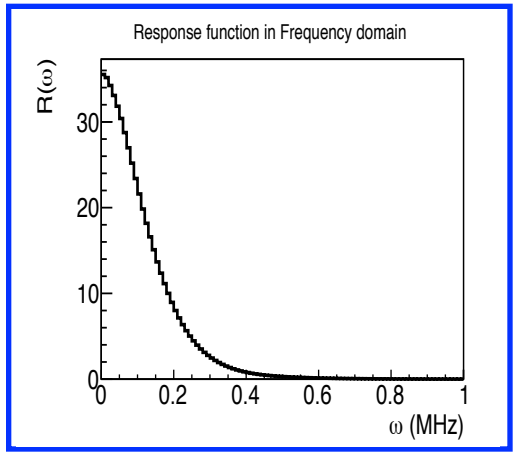
- Deconvolution problem:
  - $M(t') = \int_{-\infty}^{\infty} R(t,t') \cdot S(t)\, \mathrm{d}t$
  - True signal $S(t)$, measured signal $M(t')$, response function $R(t,t') \equiv R(t-t')$
- Fourier transform:

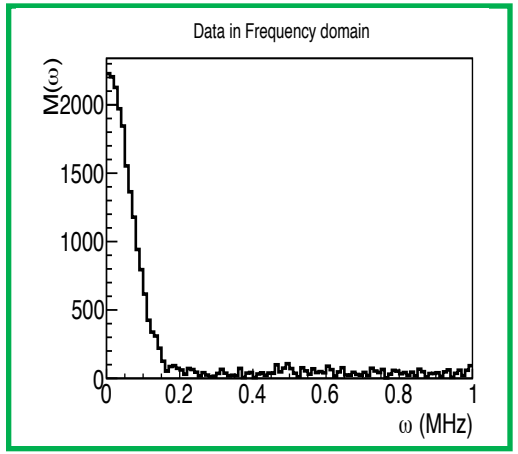
$$M(\omega) = R(\omega) \cdot S(\omega) \rightarrow S(\omega) = M(\omega)/R(\omega)$$

$$\rightarrow \text{Inverse FFT } S(\omega) \rightarrow S(t)$$

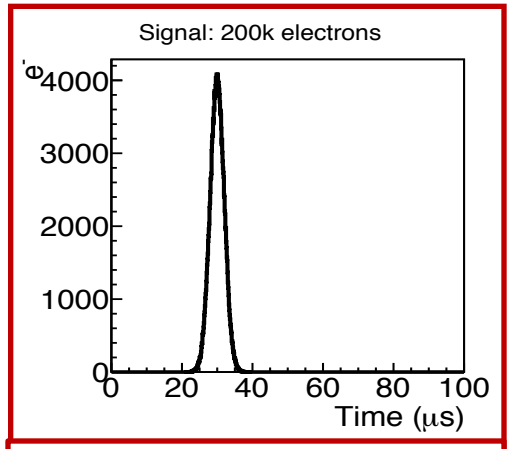
- The response function $R(\omega)$ does not address noise contributions to the measured signal. Worse still, $R(\omega)$ is generally smaller at higher frequencies due to the shaping features of electronics, resulting in amplification of noises.

Response function $R(\omega)$

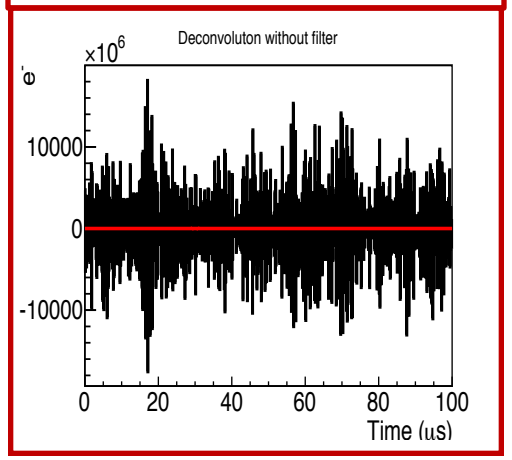




True signal $S(t)$



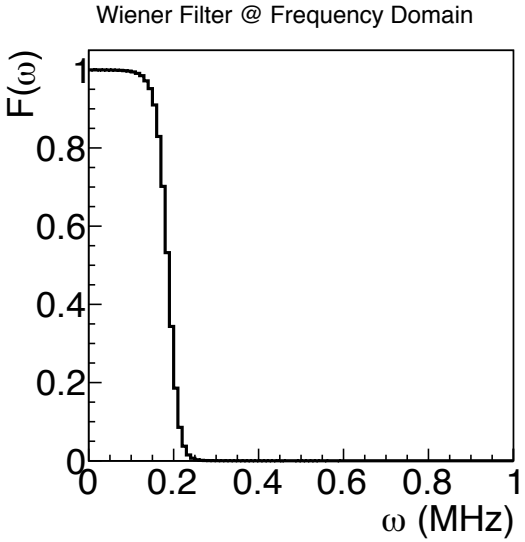
Deconvoluted signal $S(t)$



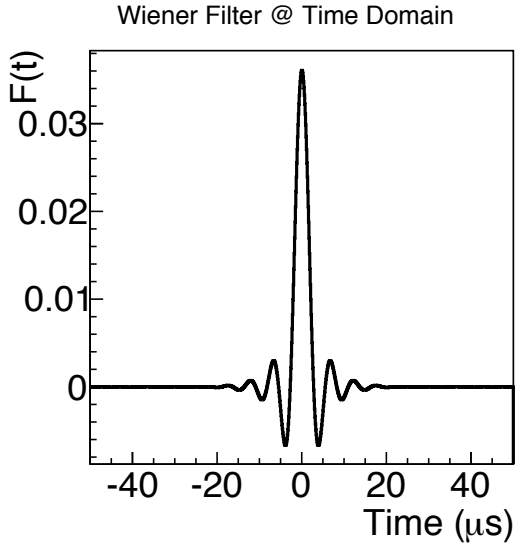
Wiener filter in digital signal processing (2)

- To address the issue with noise:
  - $\hat{S}(\omega) = \frac{M(\omega)}{R(\omega)} \cdot \boxed{F(\omega)}$ Noise filter
- Wiener filter:
  - $F(\omega) = \frac{\overline{S^2(\omega)}}{\overline{S^2(\omega)} + \boxed{N^2(\omega)}}$ Noise
  - The functional form of Wiener filter is obtained by minimizing

    $$E[(F(\omega) \cdot M(\omega) - S(\omega))^2]$$
    $$= E[(F(\omega) \cdot (S(\omega) + N(\omega)) - S(\omega))^2]$$
  - ***Wiener filter is designed to achieve the best signal/noise.***



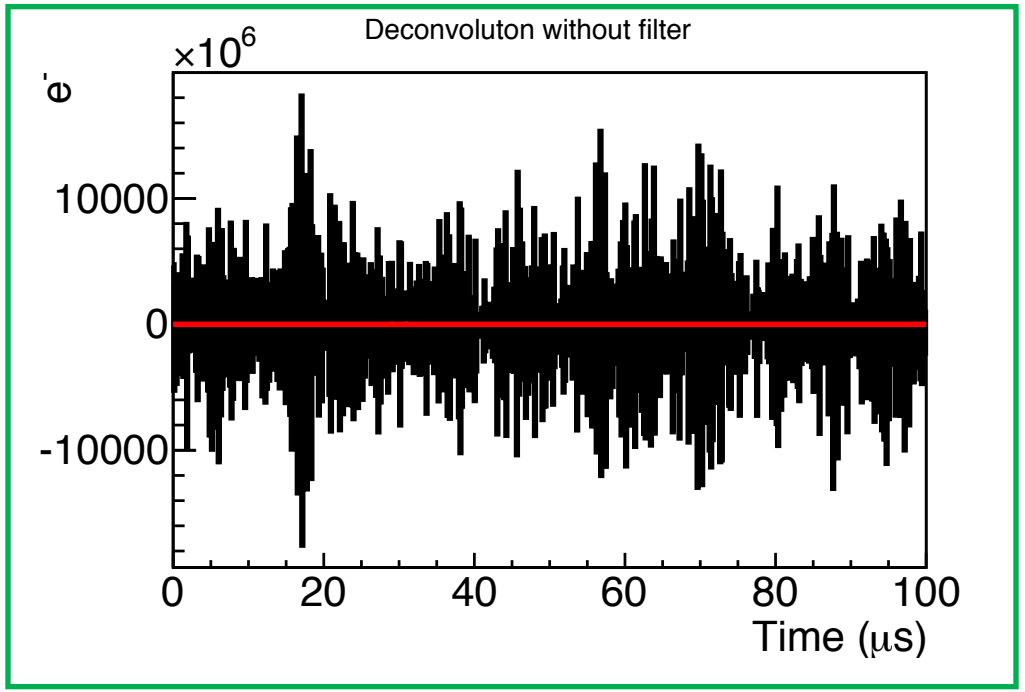
Wiener filter in ***frequency*** domain



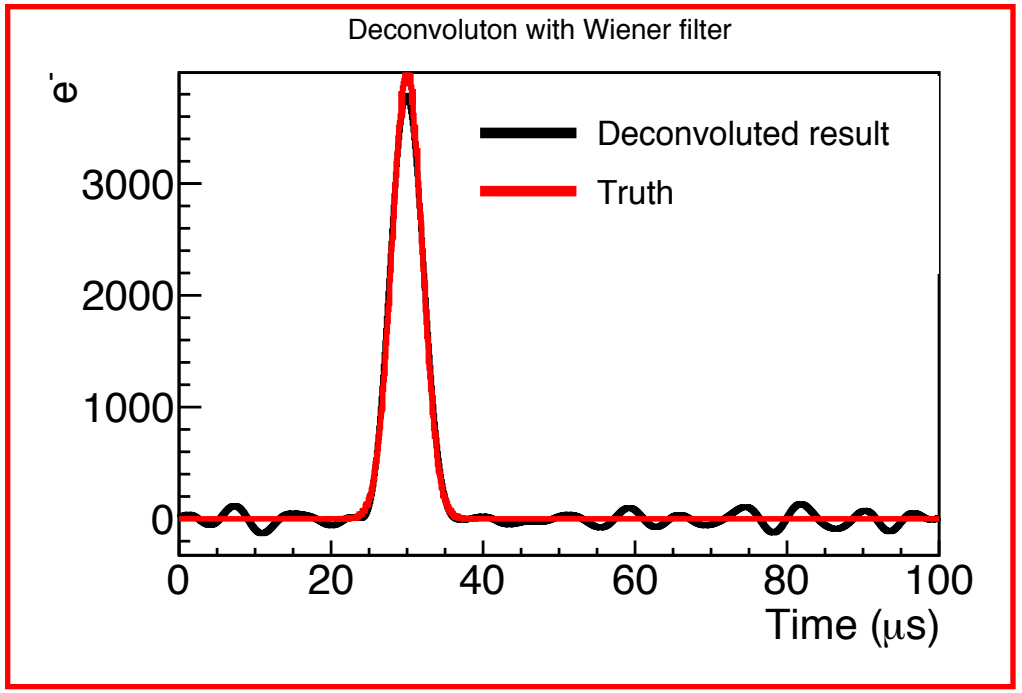
Wiener filter in ***time*** domain

Wiener filter in digital signal processing (3)

Deconvolution without filter



Deconvolution with Wiener filter



Reminder of SVD Method

- Minimize χ^2:

$$\chi^2(s) = (\mathbf{m} - \mathbf{r} \cdot s)^T \cdot Cov^{-1} \cdot (\mathbf{m} - \mathbf{r} \cdot s)$$

  - Cholesky decomposition: $Cov^{-1} = Q^T \cdot Q$
  - Pre-scaling: $M := Q \cdot \mathbf{m}, R := Q \cdot \mathbf{r}$
  - ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

- Equivalently, the estimator can be written as

$$\hat{s} = (R^T R)^{-1} \cdot R^T \cdot (R \cdot s_{\text{true}} + N)$$

where N, the "noise" coming from uncertainties, follows a normal distribution.

Reminder of SVD Method

- Minimize χ^2:

$$\chi^2(s) = (\mathbf{m} - \mathbf{r} \cdot s)^T \cdot Cov^{-1} \cdot (\mathbf{m} - \mathbf{r} \cdot$$

  - Cholesky decomposition: $Cov^{-1} = Q^T \cdot Q$
  - Pre-scaling: $M := Q \cdot \mathbf{m}, R := Q \cdot \mathbf{r}$
  - ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

Singular value decomposition (SVD)

$$R = U \cdot D \cdot V^T$$

U, V: orthogonal

D: diagonal (diagonal elements ranked by magnitude)

- Equivalently, the estimator can be written as

$$\hat{s} = (R^T R)^{-1} \cdot R^T \cdot (R \cdot s_{\text{true}} + N)$$

where N, the "noise" coming from uncertainties, follows a normal distribution.

Reminder of SVD Method

- Minimize χ^2:

$$\chi^2(s) = (\mathbf{m} - \mathbf{r} \cdot s)^T \cdot Cov^{-1} \cdot (\mathbf{m} - \mathbf{r} \cdot$$

  - Cholesky decomposition: $Cov^{-1} = Q^T \cdot Q$
  - Pre-scaling: $M := Q \cdot \mathbf{m}, R := Q \cdot \mathbf{r}$
  - ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

Singular value decomposition (SVD)

$$R = U \cdot D \cdot V^T$$

U, V: orthogonal

D: diagonal (diagonal elements ranked by magnitude)

- Equivalently, the estimator can be written as

$N_U = U^T \cdot N$
$M_U = U^T \cdot M$

$$\hat{s} = V \cdot D^{-1} \cdot U^T \cdot (R \cdot s_{true} + N)$$
$$= V \cdot D^{-1} \cdot (R_U \cdot s_{true} + N_U) = V \cdot D^{-1} \cdot M_U$$

  - Because U is an orthogonal matrix, N_U are still uncorrelated and follows a normal distribution.
  - $M_U = U^T \cdot M$ is the measurement in the *effective frequency domain*

Reminder of SVD Method

- Minimize χ^2:

$$\chi^2(s) = (\mathbf{m} - \mathbf{r} \cdot s)^T \cdot Cov^{-1} \cdot (\mathbf{m} - \mathbf{r} \cdot$$

  - Cholesky decomposition: $Cov^{-1} = Q^T \cdot Q$
  - Pre-scaling: $M := Q \cdot \mathbf{m}, R := Q \cdot \mathbf{r}$
  - ML estimator: $\hat{s} = (R^T R)^{-1} \cdot R^T \cdot M$

Singular value decomposition (SVD)

$$R = U \cdot D \cdot V^T$$

U, V: orthogonal

D: diagonal (diagonal elements ranked by magnitude)

- Equivalently, the estimator can be written as

$N_U = U^T \cdot N$
$M_U = U^T \cdot M$

$$\hat{s} = V \cdot D^{-1} \cdot U^T \cdot (R \cdot s_{true} + N)$$
$$= V \cdot D^{-1} \cdot (R_U \cdot s_{true} + N_U) = \boxed{V \cdot D^{-1} \cdot M_U}$$

Recall in signal processing

$$S(\omega) = [R(\omega)]^{-1} M(\omega)$$

  - Because U is an orthogonal matrix, N_U are still uncorrelated and follows a normal distribution.
  - $M_U = U^T \cdot M$ is the measurement in the *effective frequency domain*

Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of

$$\phi(s) = \log \boxed{L(s)} + \tau \boxed{\Sigma(s)}$$

Likelihood function

Regularization function

Regularization strength

- E.g. Tikhonov regularization

$$\Sigma^k\big(s(E)\big) = -\int \left(\frac{\mathrm{d}^k s(E)}{\mathrm{d}^k E}\right)^2 \mathrm{d}E, k = 0, 1, 2, \ldots$$

- When $k = 0$, the unfolded result can be written as

$$\hat{s} = (R^T R + \tau I)^{-1} \cdot R^T \cdot M$$

where I is an identity matrix and τ is the regularization strength

Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of

$$\phi(s) = \log \boxed{L(s)} + \tau \boxed{\Sigma(s)}$$

Likelihood function

Regularization function

Regularization strength

- E.g. Tikhonov regularization

$$\Sigma^k\big(s(E)\big) = -\int \left(\frac{\mathrm{d}^k s(E)}{\mathrm{d}^k E}\right)^2 \mathrm{d}E, k = 0, 1, 2, \ldots$$

$$R = U \cdot D \cdot V^T$$

- When $k = 0$, the unfolded result can be written as

$$\hat{s} = (R^T R + \tau I)^{-1} \cdot R^T \cdot M$$

where I is an identity matrix and τ is the regularization strength

Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of

$$\phi(s) = \log \boxed{L(s)} + \tau \boxed{\Sigma(s)}$$

Likelihood function

Regularization function

Regularization strength

- E.g. Tikhonov regularization

$$\Sigma^k\big(s(E)\big) = -\int \left(\frac{\mathrm{d}^k s(E)}{\mathrm{d}^k E}\right)^2 \mathrm{d}E, k = 0, 1, 2, \ldots$$

$$R = U \cdot D \cdot V^T$$

- When $k = 0$, the unfolded result can be written as

$$\hat{s} = A \cdot (R^T R)^{-1} \cdot R^T \cdot M$$

where $A = V \cdot F \cdot V^T$, $F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij}$ (d_i is the diagonal terms of D)

Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of

$$\phi(s) = \log L(s) + \tau \Sigma(s)$$

Likelihood function

Regularization function

Regularization strength

- E.g. Tikhonov regularization

Regularization is effectively introducing an additional smearing to the unfolded results

$$R = U \cdot D \cdot V^T$$

- When k

$$\hat{s} = A \cdot (R^T R)^{-1} \cdot R^T \cdot M$$

where $A = V \cdot F \cdot V^T$, $F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij}$ (d_i is the diagonal terms of D)

Reminder of Regularization unfolding

- The estimator can be obtained by finding the maximum of

$$\phi(s) = \log \boxed{L(s)} + \tau \boxed{\Sigma(s)}$$

Likelihood function

Regularization function

Regularization strength

- E.g. Tikhonov regularization

$$\Sigma^k\big(s(E)\big) = -\int \left(\frac{\mathrm{d}^k s(E)}{\mathrm{d}^k E}\right)^2 \mathrm{d}E, k = 0, 1, 2, \ldots$$

$$R = U \cdot D \cdot V^T$$

- When $k = 0$, the unfolded result can be written as

$$\hat{s} = V \cdot F \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U), \text{where } F_{ij} = \frac{d_i^2}{d_i^2 + \tau}\delta_{ij}$$

N_U are suppressed by $d_i/(d_i^2 + \tau)$, regardless of signal strength

- Recall: $\hat{s} = V \cdot F \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U)$, $F_{ij} = \dfrac{d_i^2}{d_i^2 + \tau} \delta_{ij}$

Wiener-SVD

- Recall: $\hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\mathrm{true}} + N_U)$, W: Wiener filter

Wiener-SVD

- Recall: $\hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U)$, W: Wiener filter
- In Wiener-SVD method, the functional form of W (replace F as filter) also factors in the signal expectation in the effective frequency domain

$$\overline{M_U} := U^T \cdot \bar{M} = U^T \cdot R \cdot \bar{s} = D \cdot V^T \cdot \bar{s}$$

$$W_{ij} = \frac{\overline{M_U}_i^2}{\overline{M_U}_i^2 + \overline{N_U}_i^2} \delta_{ij}$$

Wiener-SVD

- Recall: $\hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U)$, W: Wiener filter
- In Wiener-SVD method, the functional form of W (replace F as filter) also factors in the signal expectation in the effective frequency domain

$$\overline{M_U} := U^T \cdot \overline{M} = U^T \cdot R \cdot \bar{s} = D \cdot V^T \cdot \bar{s}$$

$$W_{ij} = \frac{\overline{M_U}_i^2}{\overline{M_U}_i^2 + \overline{N_U}_i^2} \delta_{ij}$$

- Signal: $\overline{M_U}_i^2 = d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2$
- Noise: 1 (normal distribution)

Wiener-SVD

- Recall: $\hat{s} = V \cdot W \cdot D^{-1} \cdot (R_U \cdot s_{\text{true}} + N_U)$, W: Wiener filter
- In Wiener-SVD method, the functional form of W (replace F as filter) also factors in the signal expectation in the effective frequency domain

$$\overline{M_U} := U^T \cdot \overline{M} = U^T \cdot R \cdot \bar{s} = D \cdot V^T \cdot \bar{s}$$

- Signal: $\overline{M_U}_i^2 = d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2$
- Noise: 1 (normal distribution)

$$W_{ij} = \frac{\overline{M_U}_i^2}{\overline{M_U}_i^2 + \overline{N_U}_i^2} \delta_{ij}$$

$$W_{ij} = \frac{d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2}{d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2 + 1} \delta_{ij}$$

- In Tikhonov regularization, when $k \neq 0$, the solution is

$$\hat{s} = (R^T R + \tau\, C^T C)^{-1} \cdot R^T \cdot \bar{M}$$

where in the case of $k = 2$,

$$C = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & -2 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is

$$\hat{s} = (R^T R + \tau\, C^T C)^{-1} \cdot R^T \cdot \bar{M}$$

where in the case of $k = 2$,

$$C = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & -2 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

- Similarly to Tikhonov regularization, we can add an additional matrix C to the equation

$$\bar{M} = (R \cdot C^{-1}) \cdot (C \cdot \bar{s}\,)$$

Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is

$$\hat{s} = (R^T R + \tau\, C^T C)^{-1} \cdot R^T \cdot \bar{M}$$

where in the case of $k = 2$,

$$C = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & -2 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

- Similarly to Tikhonov regularization, we can add an additional matrix C to the equation

$$\bar{M} = \boxed{(R \cdot C^{-1})} \cdot (C \cdot \bar{s})$$

Since the effective frequency domain is determined by the smearing matrix R, the inclusion of C would alter the basis of the effective frequency domain.

Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is

$$\hat{s} = (R^T R + \tau\, C^T C)^{-1} \cdot R^T \cdot \bar{M}$$

where in the case of $k = 2$,

$$C = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & -2 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}$$

- Similarly to Tikhonov regularization, we can add an additional matrix C to the equation

$$\bar{M} = \boxed{(R \cdot C^{-1})} \cdot (C \cdot \bar{s})$$

SVD

$$RC^{-1} = U_C \cdot D_C \cdot V_C^T$$

Generalized Wiener-SVD

- In Tikhonov regularization, when $k \neq 0$, the solution is

$$\hat{s} = (R^T R + \tau\, C^T C)^{-1} \cdot R^T \cdot \bar{M}$$

The Wiener-SVD solution becomes

$$\hat{s} = C^{-1} \cdot V_C \cdot W_C \cdot V_C^T \cdot C \cdot (R^T R)^{-1} \cdot R^T \cdot \bar{M}$$

$$(W_C)_{ij} = \frac{d_i^2 \cdot \left(\sum_j (V_C^T)_{ij} \cdot \left(\sum_l C_{jl} \cdot \bar{s}_l\right)\right)^2}{d_i^2 \cdot \left(\sum_j (V_C^T)_{ij} \cdot \left(\sum_l C_{jl} \cdot \bar{s}_l\right)\right)^2 + 1}$$

SVD

$$RC^{-1} = U_C \cdot D_C \cdot V_C^T$$

- Similarly to Tikhonov regularization, we can add an additional matrix C to the equation

$$\bar{M} = \boxed{(R \cdot C^{-1})} \cdot (C \cdot \bar{s})$$

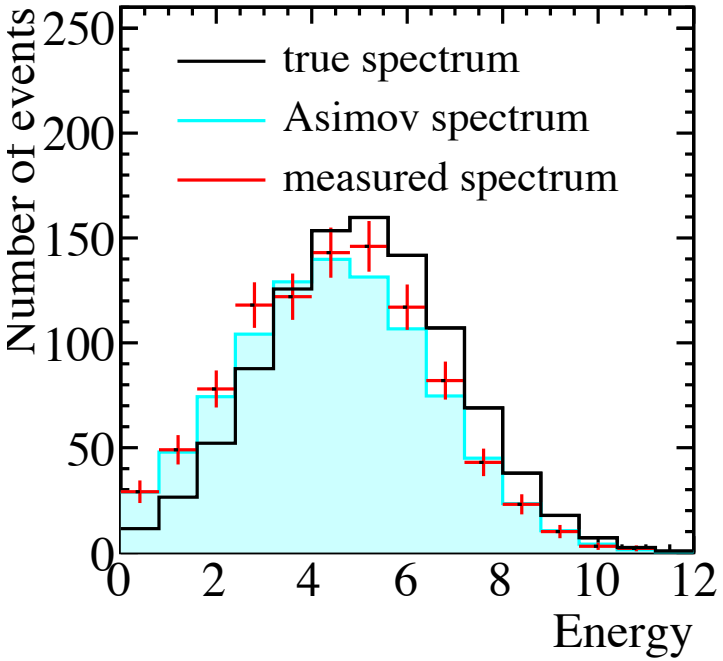
Regularization interpretation of Wiener-SVD method

- The Wiener-SVD unfolding method is equivalent to regularization w.r.t. minimizing the signal to noise ratio in the effective frequency domain $(M_U)_i = \sum_j V_{ij}^T s_j$:

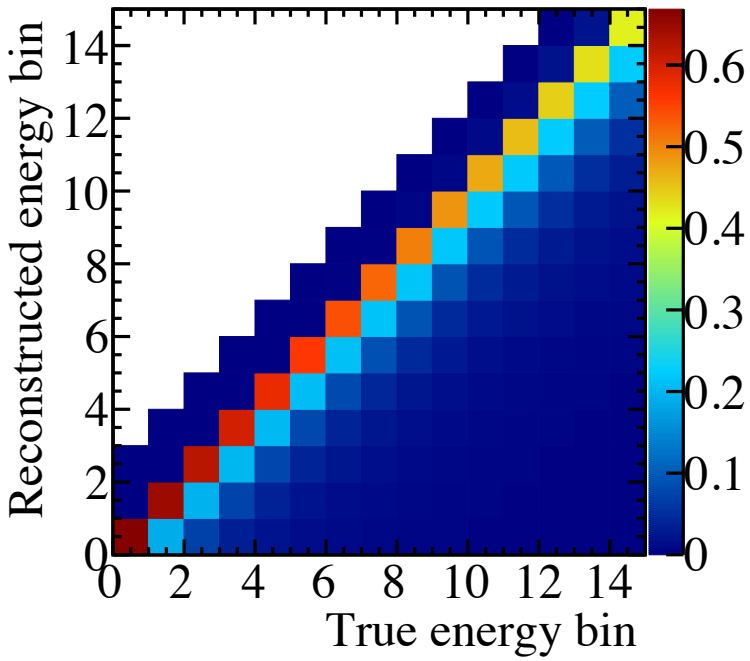
$$\phi(s) = \log L(s) + \frac{1}{2}\sum_i \log\frac{(M_U)_i^2}{\overline{N^2}} = \log L(s) + \frac{1}{2}\sum_i \log\frac{\left(\sum_j V_{ij}^T s_j\right)^2}{1}$$

Application on neutrino cross-section measurements (1)

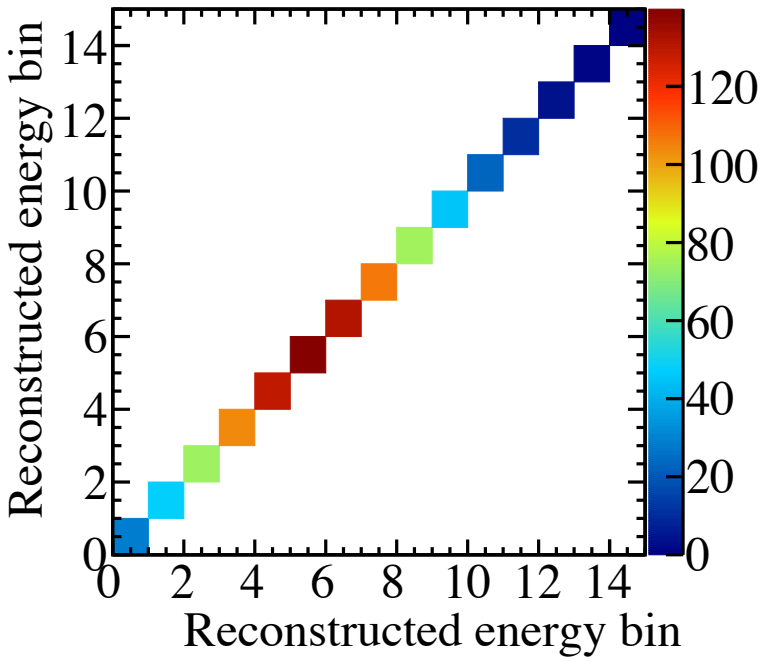
Toy experiment



Detector response



Statistics only uncertainties



Application on neutrino cross-section measurements (2)

Regularization

$$\hat{s} = V \cdot F \cdot D^{-1} \cdot M$$

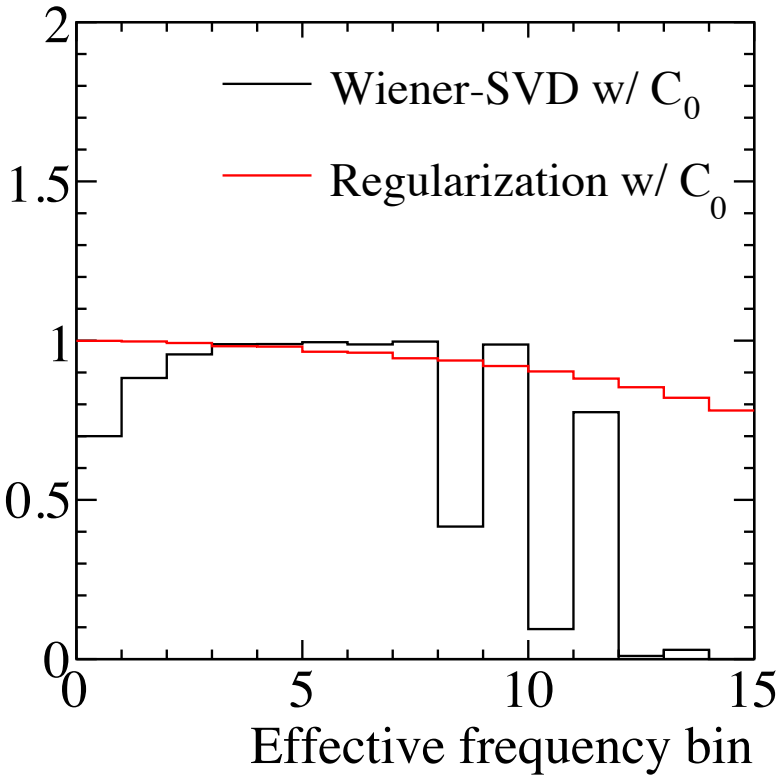
$$F_{ij} = \frac{d_i^2}{d_i^2 + \tau} \delta_{ij}$$

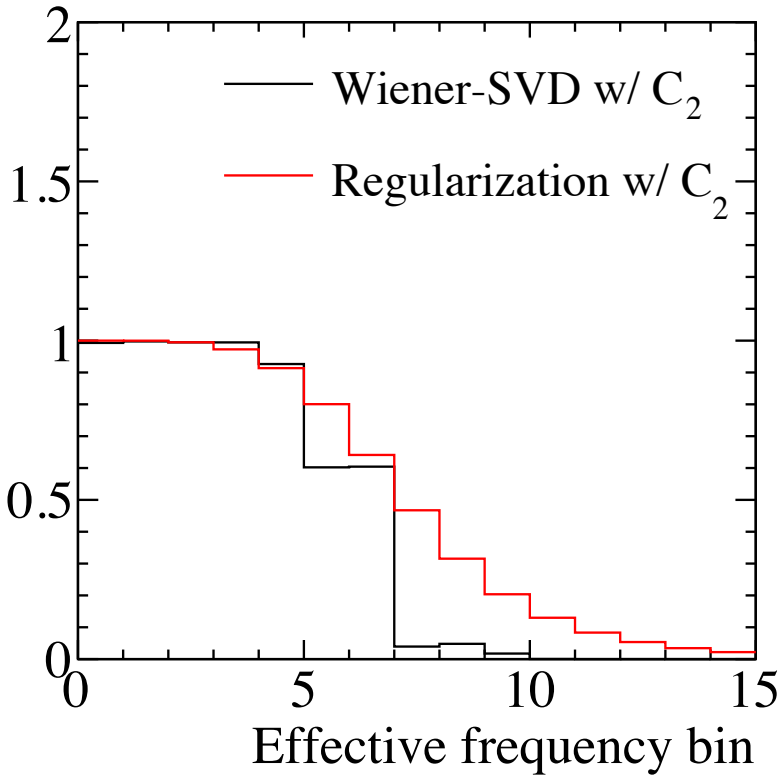
Wiener-SVD

$$\hat{s} = V \cdot W \cdot D^{-1} \cdot M$$

$$W_{ii} = \frac{d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2}{d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2 + 1}$$

Filter shape



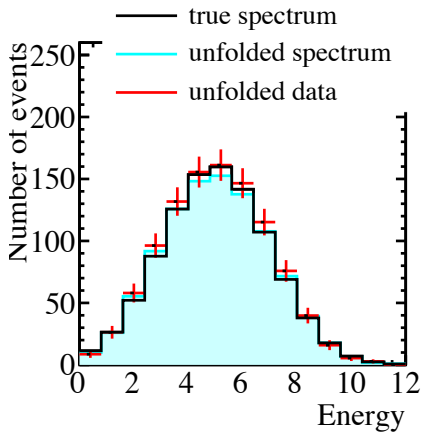


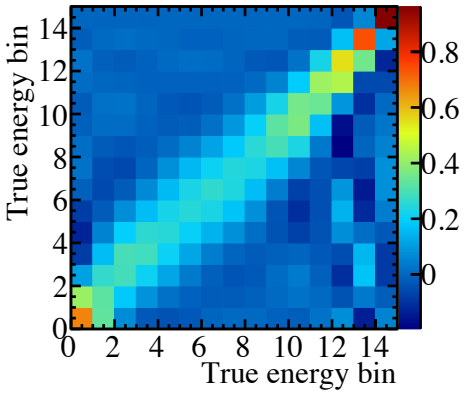
Application on neutrino cross-section measurements (3)

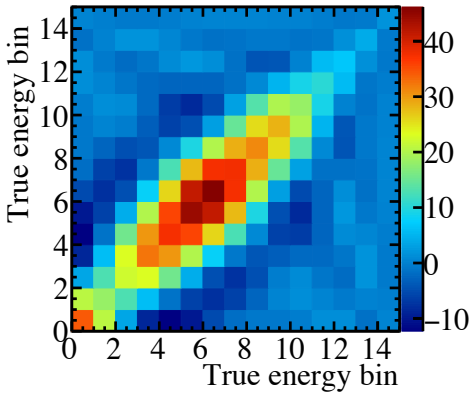
Additional smearing matrix A_C

Unfolded covariance

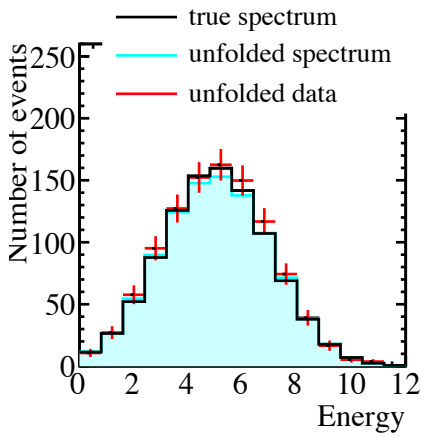
$$\hat{s} = A_C \cdot (R^T R)^{-1} \cdot R^T \cdot \bar{M}$$

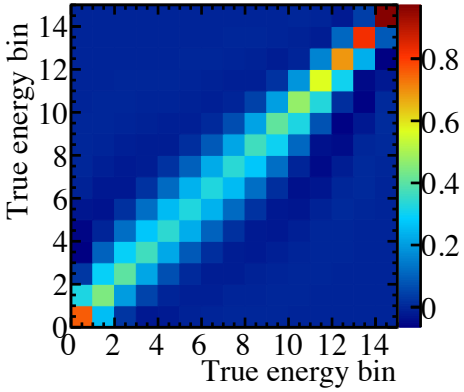


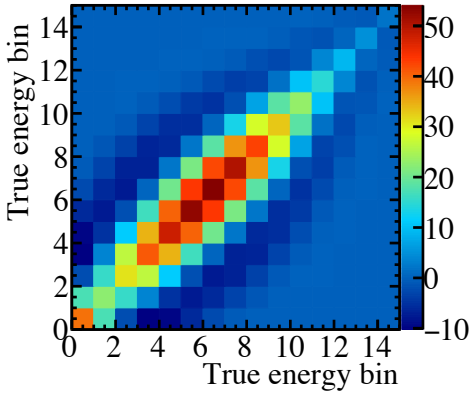




Wiener-SVD w/ C_2



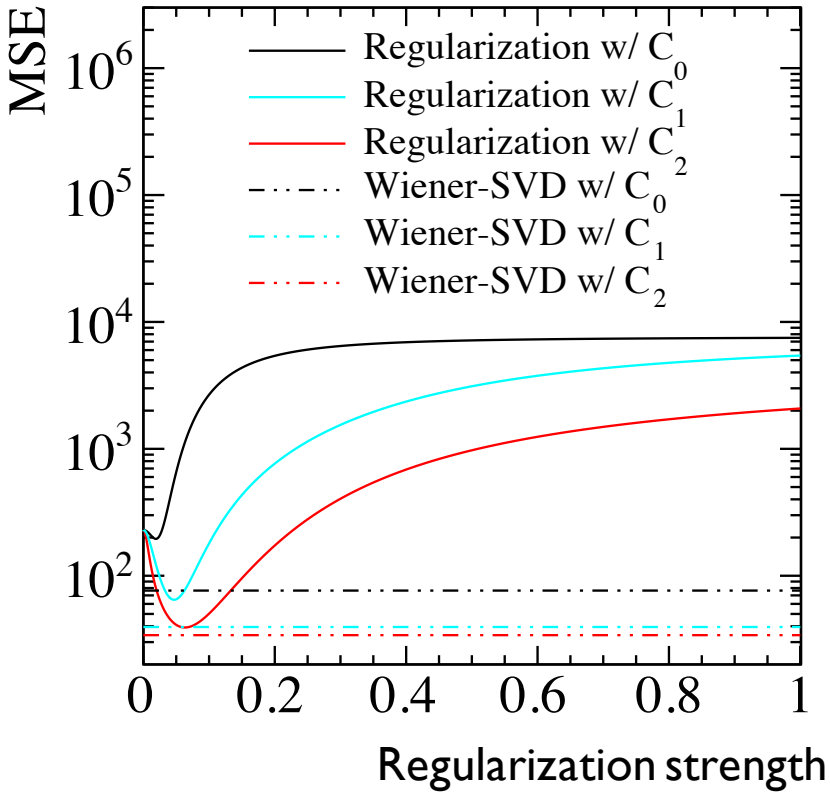


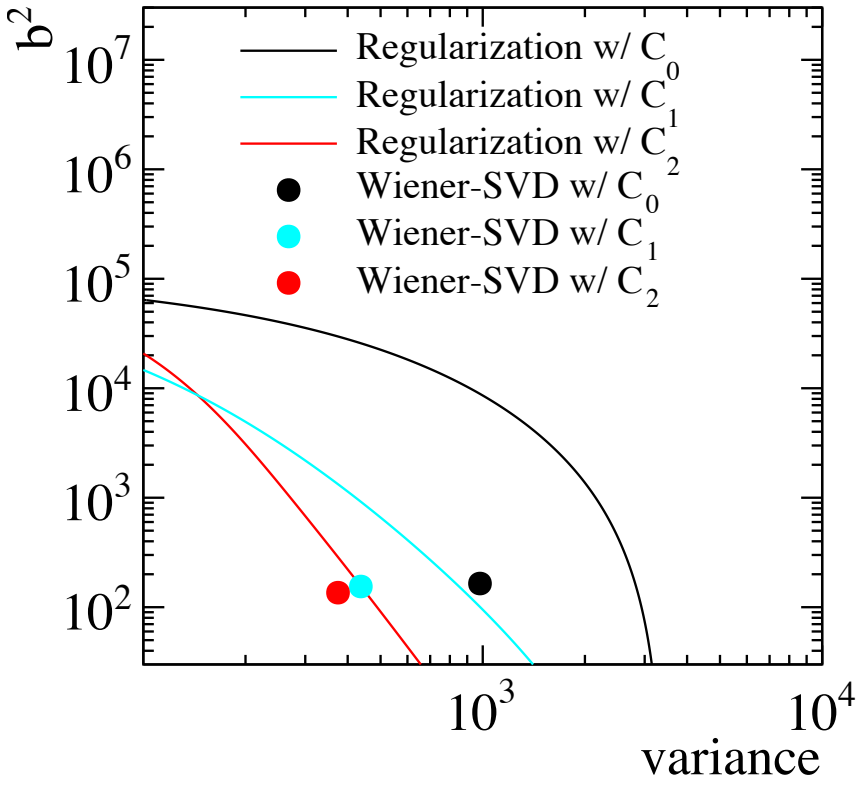


Tikhonov regularization w/ C_2

Application on neutrino cross-section measurements (4)

$$\mathrm{MSE} = \frac{1}{n}\left(\sigma^2 + b^2\right)$$



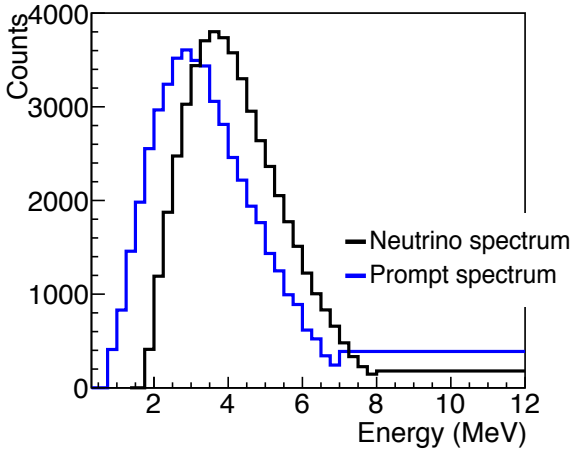


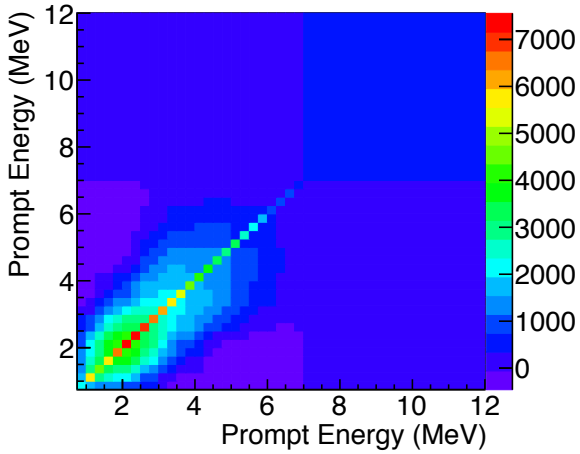
$$C_0 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix},$$

$$C_1 = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 \\ 0 & 0 & 0 & \dots & 0 & -1 \end{bmatrix},$$

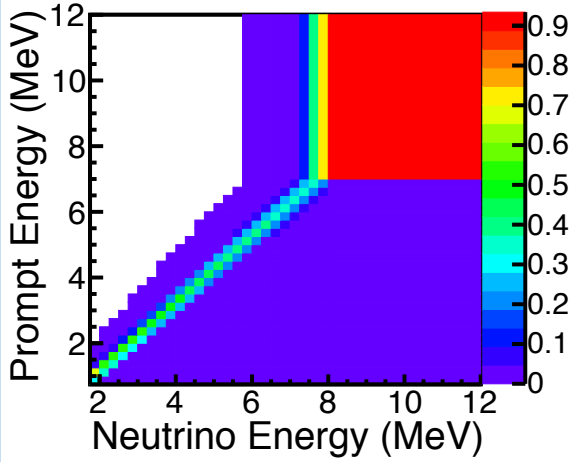
$$C_2 = \begin{bmatrix} -1+\epsilon & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2+\epsilon & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2+\epsilon & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & -1+\epsilon \end{bmatrix},$$

Application on reactor neutrino measurement (1)

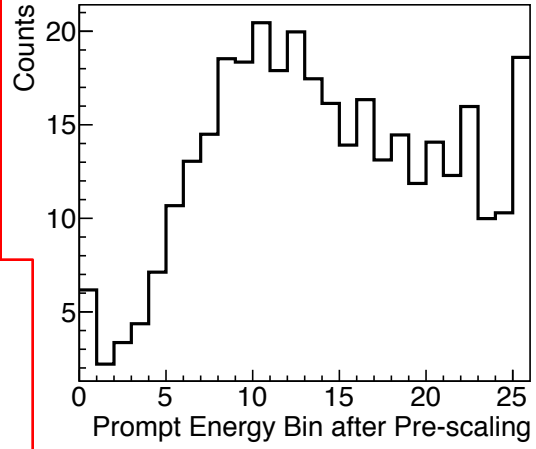




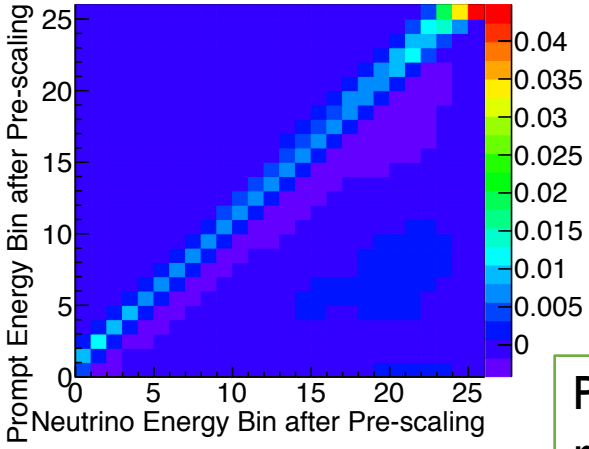
Covariance matrix includes both statistical and systematic uncertainties



Detector energy response matrix $\mathbf{r}$ as reported by the Daya Bay collaboration



Pre-scaled prompt energy distribution $M := Q \cdot \mathbf{m}$

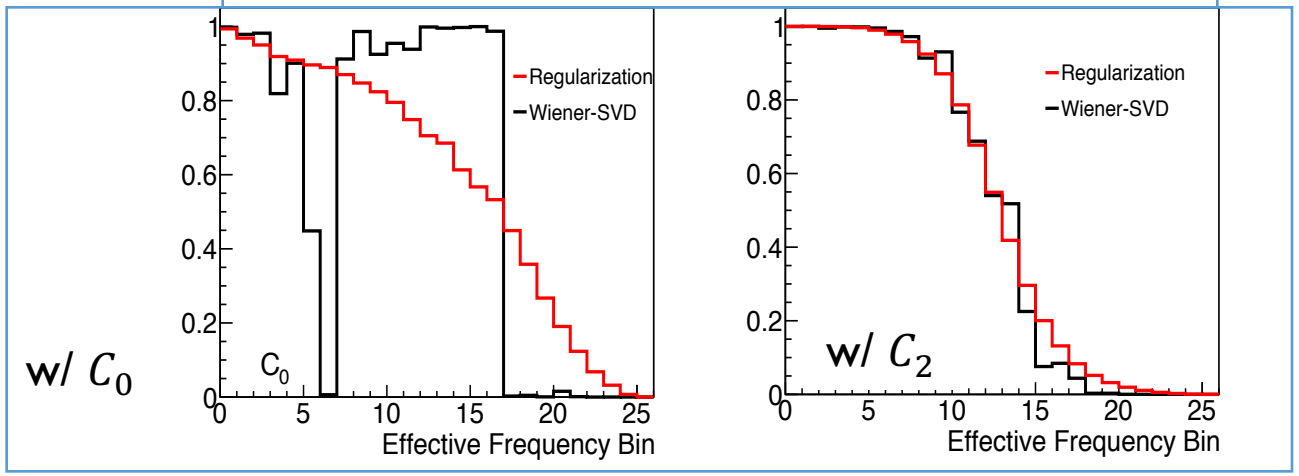


Pre-scaled detector energy response matrix $R := Q \cdot \mathbf{r}$

$$Cov^{-1} = Q^T \cdot Q$$

Application on reactor neutrino measurement (2)

"Filter" shape in the effective frequency domain

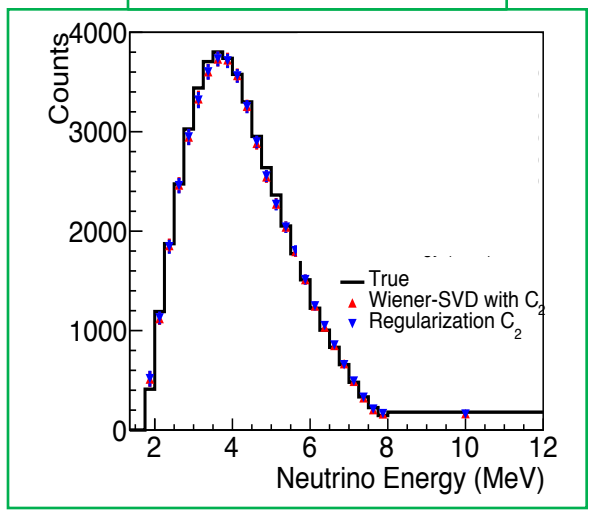


- In practice, the "true" energy spectrum is unknown; all we have are various models → take average

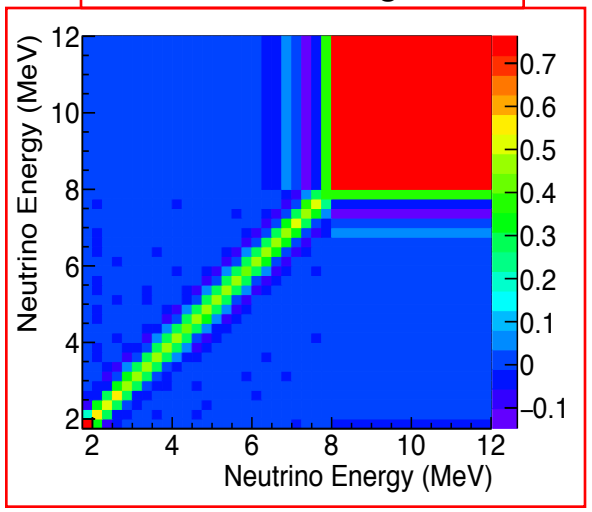
$$\text{“}\overline{S^2}\text{”} = \overline{M}_U^2 = d_i^2 \cdot \frac{\sum_k \left(\sum_j V_{ij}^T \cdot \overline{s(k)}_j\right)^2 \cdot e^{-\frac{x_k^2}{2}}}{\sum_k e^{-\frac{x_k^2}{2}}}$$

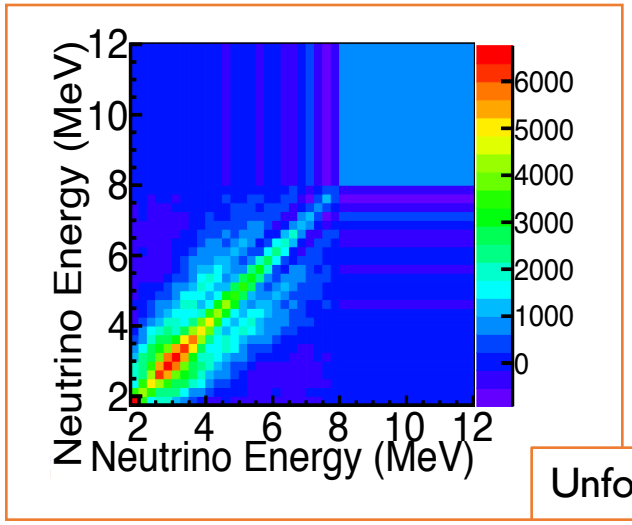
- The χ_k^2 for model k is calculated by comparing the predicted spectrum $\boldsymbol{m}_k := \boldsymbol{r} \cdot \boldsymbol{s}_k$ and measurement $\boldsymbol{m}$

Unfolded distributions



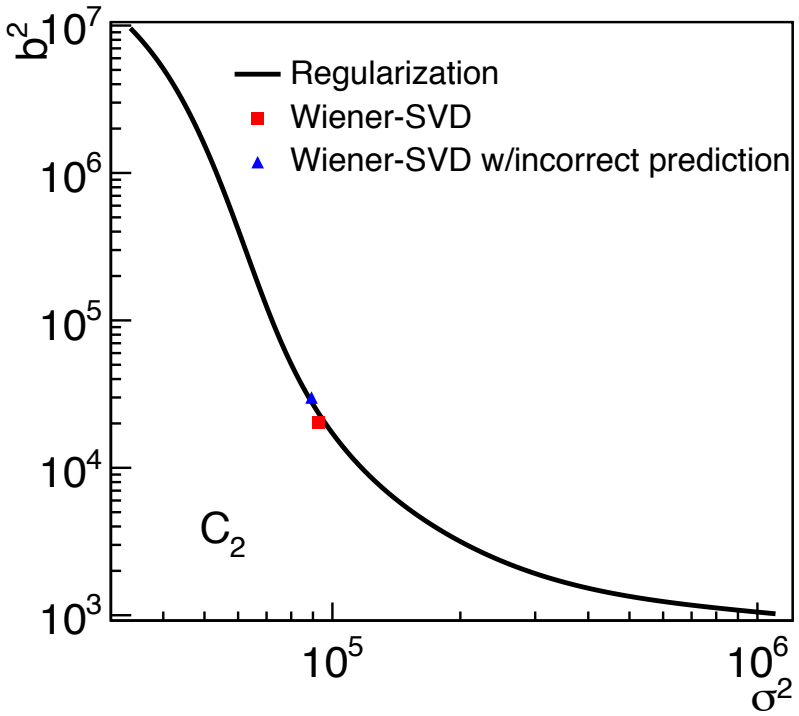
Additional smearing matrix

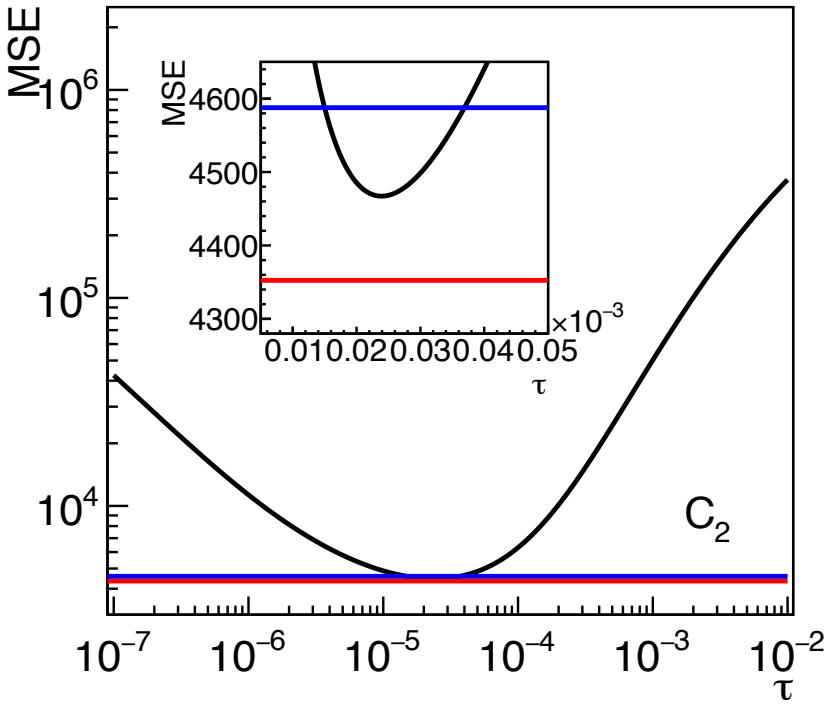




Unfolded covariance matrix

Application on reactor neutrino measurement (3)





Using $\bar{s}$ to represent s_{true}

Using an incorrect s to represent s_{true}

$$\text{“}\overline{S^2}\text{”} = \overline{M}_U^2 = d_i^2 \cdot \frac{\sum_k \left(\sum_j V_{ij}^T \cdot \overline{s(k)}_j\right)^2 \cdot e^{-\frac{x_k^2}{2}}}{\sum_k e^{-\frac{x_k^2}{2}}}$$

Discussions and recommendations (1)

- Pros & cons of Wiener-SVD
  - It does not require the evaluation of regularization strength.
  - It is model-dependent. However, the dependence can be mitigated by taking into account a number of different models, and by reporting the additional smearing matrix.
- Both the traditional regularization or Wiener-SVD are equivalent to applying a new smearing matrix, which suppresses the large oscillation (high variance) of the direct matrix inversion unfolded results, but also introduces bias.

$$\hat{s} = A_C \cdot (R^T R)^{-1} \cdot R^T \cdot \overline{M}$$

Where $A_C = C^{-1} \cdot V_C \cdot W_C \cdot V_C^T \cdot C$ $\quad (W_C)_{ij} = \dfrac{d_i^2 \cdot \left(\sum_j (V_C^T)_{ij} \cdot \left(\sum_l C_{jl} \cdot \bar{s}_l\right)\right)^2}{d_i^2 \cdot \left(\sum_j (V_C^T)_{ij} \cdot \left(\sum_l C_{jl} \cdot \bar{s}_l\right)\right)^2 + 1}$

Discussions and recommendations (2)

- The Wiener filter should be constructed in ways which takes into account a range of prior expectations.
- C_2 matrix generally yields better results than C_0 and C_1.
- We recommend reporting this additional smearing matrix A_C in the publication together with the unfolded results to enable a more direct comparison of expectations (e.g. from new theoretical calculations) with the unfolded results. In practice, the new smearing matrix should be applied to the theoretical calculation before comparing to the unfolded results.

arXiv:1705.03568

https://github.com/BNLIF/Wiener-SVD-Unfolding

Supplemental

Regularization interpretation of Wiener-SVD method

- The Wiener-SVD unfolding method is equivalent to regularization w.r.t. minimizing the signal to noise ratio in the effective frequency domain $(M_U)_i = \sum_j V_{ij}^T s_j$:

$$\phi(s) = \log L(s) + \frac{1}{2}\sum_i \log\frac{(M_U)_i^2}{\overline{N^2}} = \log L(s) + \frac{1}{2}\sum_i \log\frac{\left(\sum_j V_{ij}^T s_j\right)^2}{1}$$

Minimizing Eq. (7) yields the following estimator

$$\hat{s} = -X^{-1} \cdot Y \cdot M_U$$

Where

$$X_{ij} = \frac{\partial^2 \phi^2}{\partial s_i \partial s_j} = -(R^T R)_{ij} - \sum_k V_{ik} \cdot \frac{1}{(M_U)_k^2} \cdot V_{kj}^T$$

$$Y_{ij} = \frac{\partial^2 \phi^2}{\partial s_i \partial M_{U_j}} = R_{ij}^T$$

- With X and Y evaluated at the expectation of $\bar{s}$ and $\bar{M}$, one recovers Eq.(6)

$$\hat{s} = V \cdot W \cdot V^T \cdot R^T \cdot \bar{M}$$

With

$$W_{ij} = \frac{d_i^2}{d_i^2 + \frac{1}{\overline{M_{U_i}^2}}}\delta_{ij} = \frac{d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2}{d_i^2 \cdot \left(\sum_j V_{ij}^T \cdot \bar{s}_j\right)^2 + 1}\delta_{ij}$$