Continuous Basis Pursuit

Eero Simoncelli

Howard Hughes Medical Institute,
Center for Neural Science, and
Courant Institute of Mathematical Sciences

New York University

Continuous Basis Pursuit

Eero Simoncelli

Howard Hughes Medical Institute,
Center for Neural Science, and
Courant Institute of Mathematical Sciences

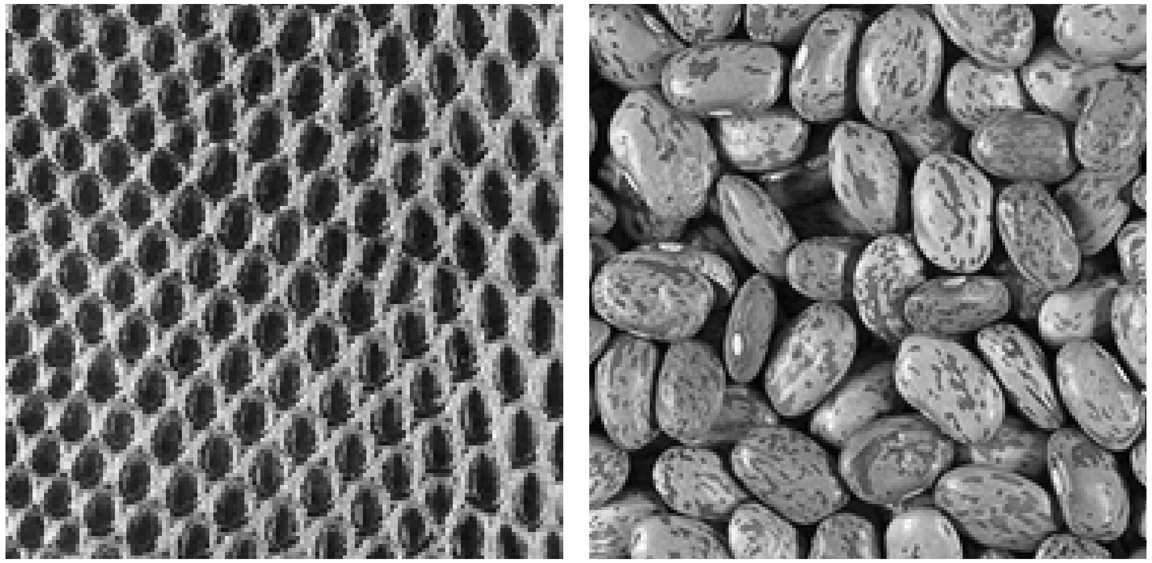
New York University

Chaitu
Ekanadham

Ubiquitous signal property: Superposition of prototype features, at arbitrary (**continuous**) times/positions/orientations/sizes etc ...

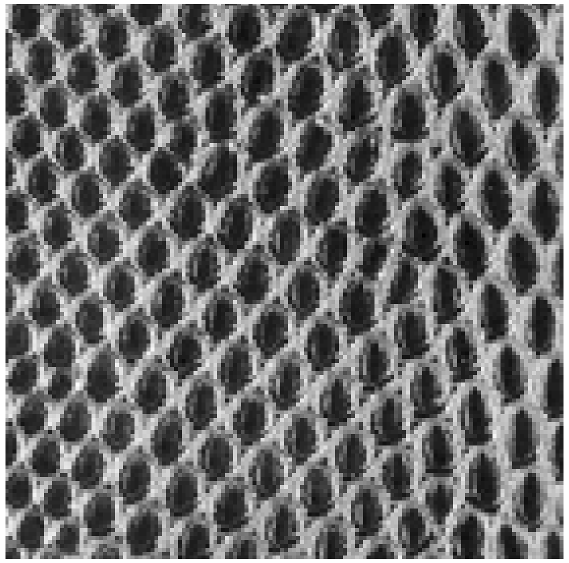
Ubiquitous signal property: Superposition of prototype features, at arbitrary (**continuous**) times/positions/orientations/sizes etc ...

images (e.g., texture)



Ubiquitous signal property: Superposition of prototype features, at arbitrary (**continuous**) times/positions/orientations/sizes etc ...

images (e.g., texture)



sounds (e.g., speech / music)

radar or seismic signals

spikes in neurons

Continuous sparse inverse problem:

$$y(t) = \sum_j a_j f(t - \tau_j) + \eta(t)$$

f localized waveform, assume known (for now)

τ_j independent arrival times

a_j i.i.d., distributed according to $p(a)$

$\eta(t)$ AWGN, variance σ^2 (for now)

Goal: recover a_j's and τ_j's

Classical solutions (from the 1940’s) for related problems...

- Matched filter
  (isolated occurrences, known waveform)

- Wiener-Kolmogorov filter
  (dense occurrences, known waveform)

Modern sparse inverse solution

- Re-formulate as a **discrete** linear inverse problem with dictionary of Δ-shifted functions:

$$\{f(t - n\Delta)\}$$

Modern sparse inverse solution

- Re-formulate as a **discrete** linear inverse problem with dictionary of Δ-shifted functions:

$$\{f(t-n\Delta)\}$$

- Enforce sparsity with a convex objective function for coefficients $\vec{x}$:

$$\arg\min_{\vec{x}} ||y(t) - \sum_{n=1}^{N} x_n f(t-n\Delta)||_2^2 + \lambda||\vec{x}||_1$$

LASSO [Tibshirani, 1996]   Basis Pursuit [Chen, Donoho, Sanders, 1998]

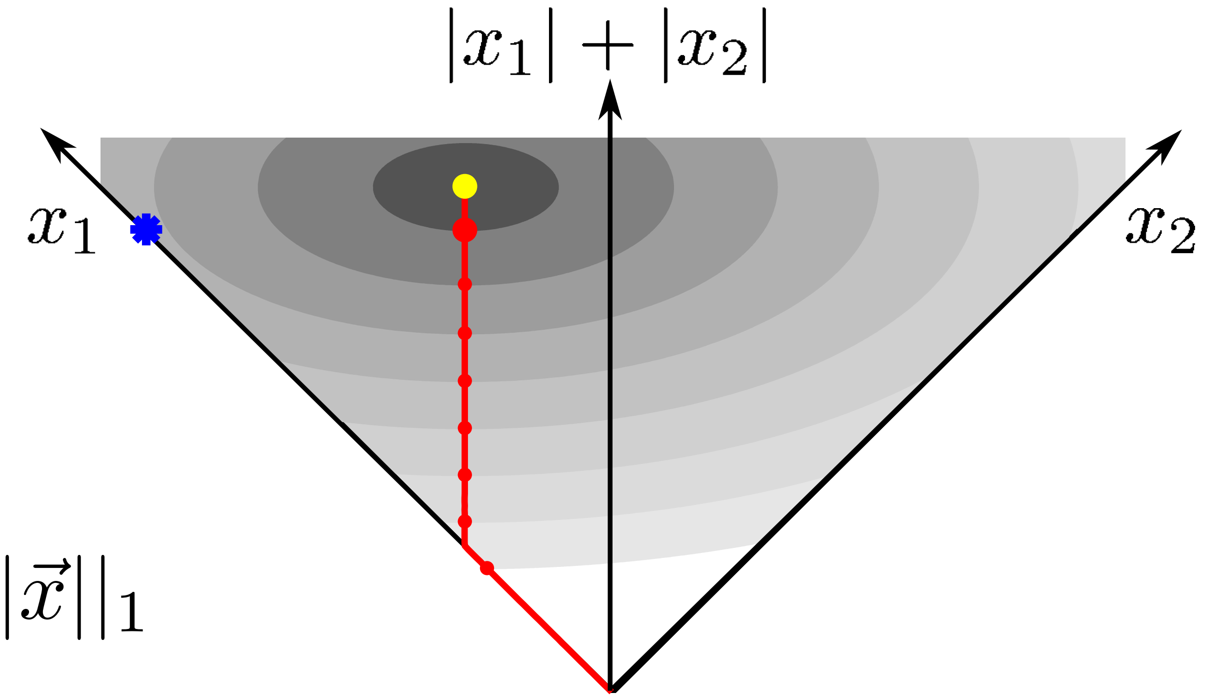
Drawbacks of BP

$$\arg\min_{\vec{x}} ||y(t) - \sum_{n=1}^{N} x_n f(t - n\Delta)||_2^2 + \lambda ||\vec{x}||_1$$

- Δ should be small
  - approximation of f at non-integer $\{\tau_j\}$
  - Mapping of $\{x_n\}$ to $\{a_j, \tau_j\}$?
- But Δ can't be too small
  - Basis redundancy => convex approximation fails to produce sparsity [Candes, Romberg, Tau, 2006]
  - High computational cost

2D illustration of objective function and solutions

$$||y(t) - \sum_{n=1}^{N} x_n f(t - n\Delta)||_2^2 + \lambda||\vec{x}||_1$$



gray - reconstruction error

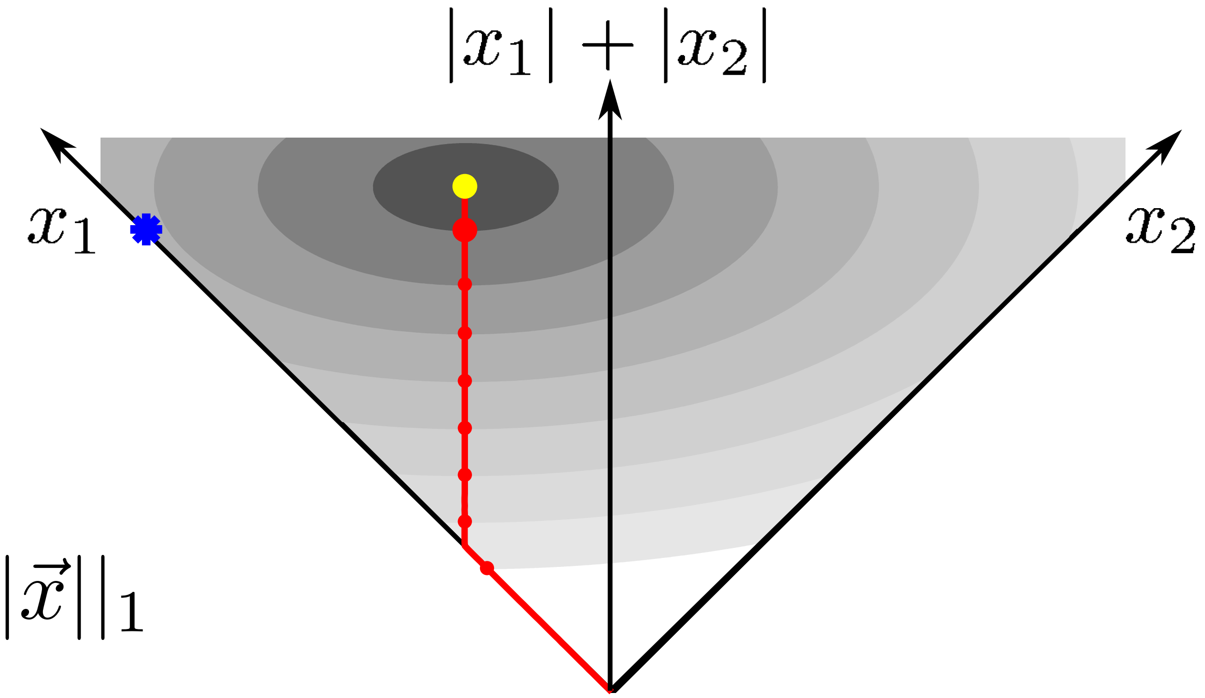
blue - L0 solution

yellow - LS solution ($\lambda = 0$)

red - solutions for different λ

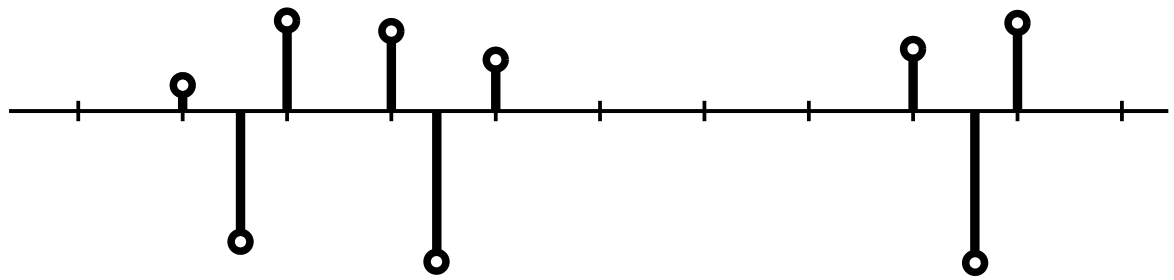
2D illustration of objective function and solutions

$$||y(t) - \sum_{n=1}^{N} x_n f(t - n\Delta)||_2^2 + \lambda||\vec{x}||_1$$



gray - reconstruction error
blue - L0 solution
yellow - LS solution ($\lambda = 0$)
red - solutions for different λ

Signal simulation



upward: BP solution
downward: true waveform locations/amplitudes

Continuous basis pursuit - Taylor

Choose spacing such that:

$$af(t+\tau) \approx af(t) + a\tau f'(t), \qquad |\tau| < \Delta/2$$

Continuous basis pursuit - Taylor

Choose spacing such that:

$$af(t+\tau) \approx af(t) + a\tau f'(t), \qquad |\tau| < \Delta/2$$

Construct *interpolative* dictionary:

$$\{f(t-n\Delta), f'(t-n\Delta)\}$$

Continuous basis pursuit - Taylor

Choose spacing such that:

$$af(t+\tau) \approx af(t) + a\tau f'(t), \qquad |\tau| < \Delta/2$$

Construct *interpolative* dictionary:

$$\{f(t-n\Delta), f'(t-n\Delta)\}$$

Optimize (convex) objective:

$$\arg\min_{\vec{x},\vec{d}} ||y(t) - \sum_{n=1}^{N} x_n f(t-n\Delta) + d_n f'(t-n\Delta)||_2^2 + \lambda ||\vec{x}||_1$$

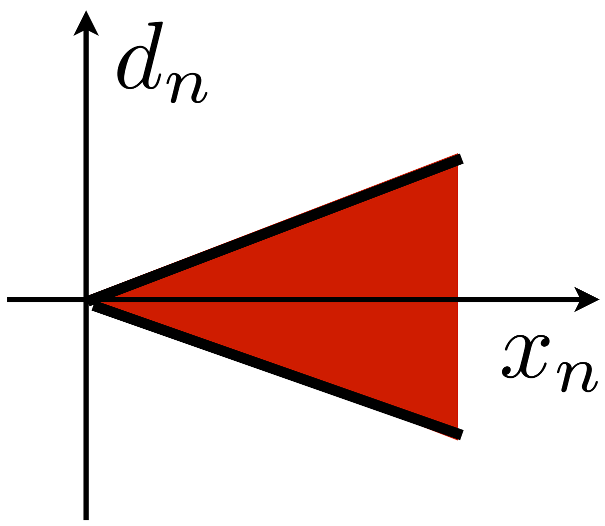
Continuous basis pursuit - Taylor

Choose spacing such that:

$$af(t+\tau) \approx af(t) + a\tau f'(t), \qquad |\tau| < \Delta/2$$

Construct *interpolative* dictionary:

$$\{f(t-n\Delta), f'(t-n\Delta)\}$$

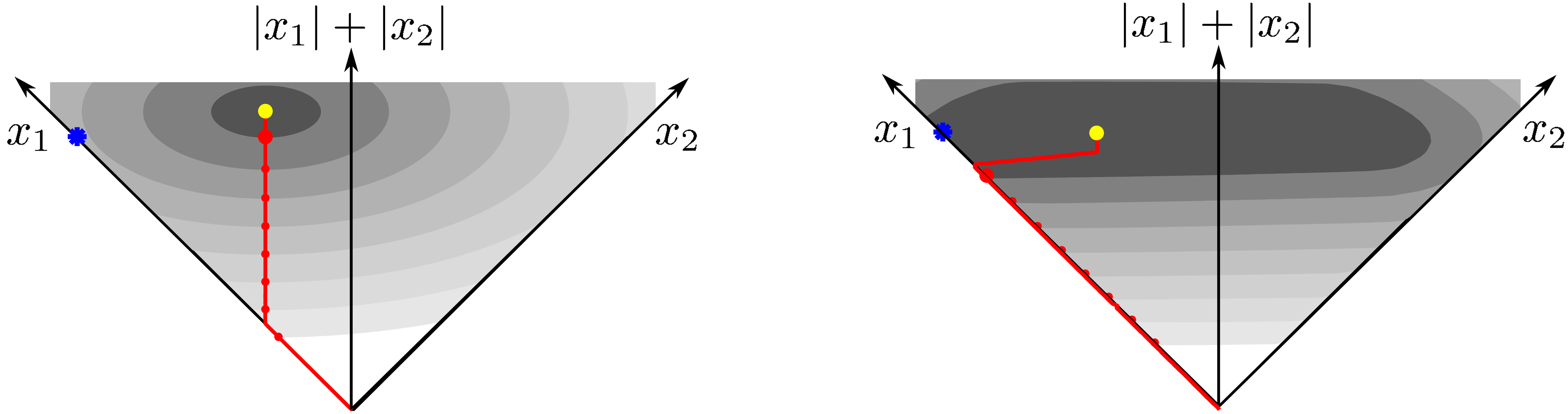


Optimize (convex) objective:

$$\arg\min_{\vec{x},\vec{d}} ||y(t) - \sum_{n=1}^{N} x_n f(t-n\Delta) + d_n f'(t-n\Delta)||_2^2 + \lambda||\vec{x}||_1$$

with $|d_n| < |x_n|\Delta/2$

BP | CBP - Taylor1



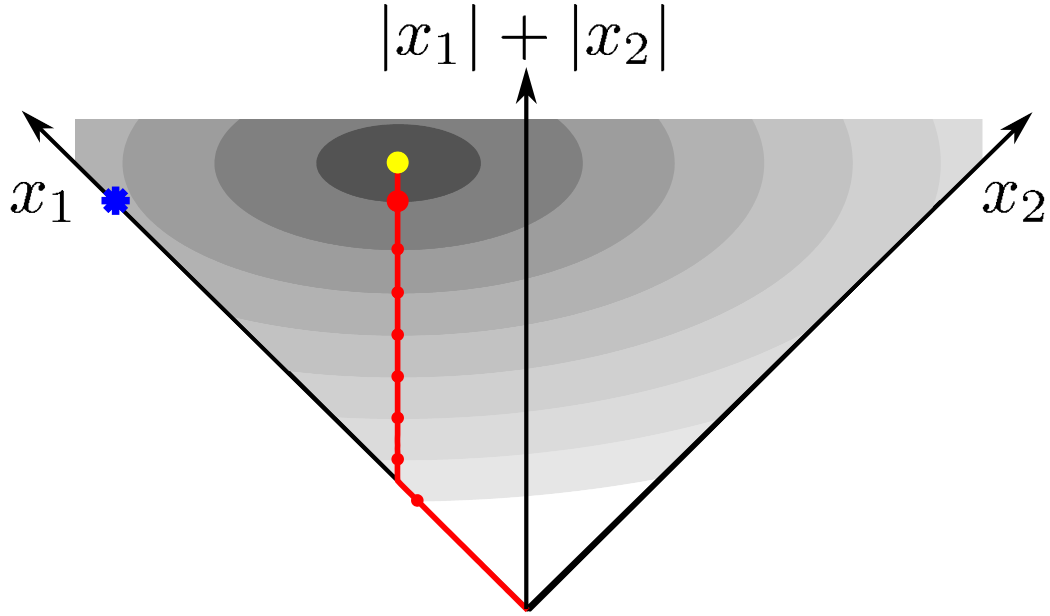
gray - reconstruction error

blue - L0 solution

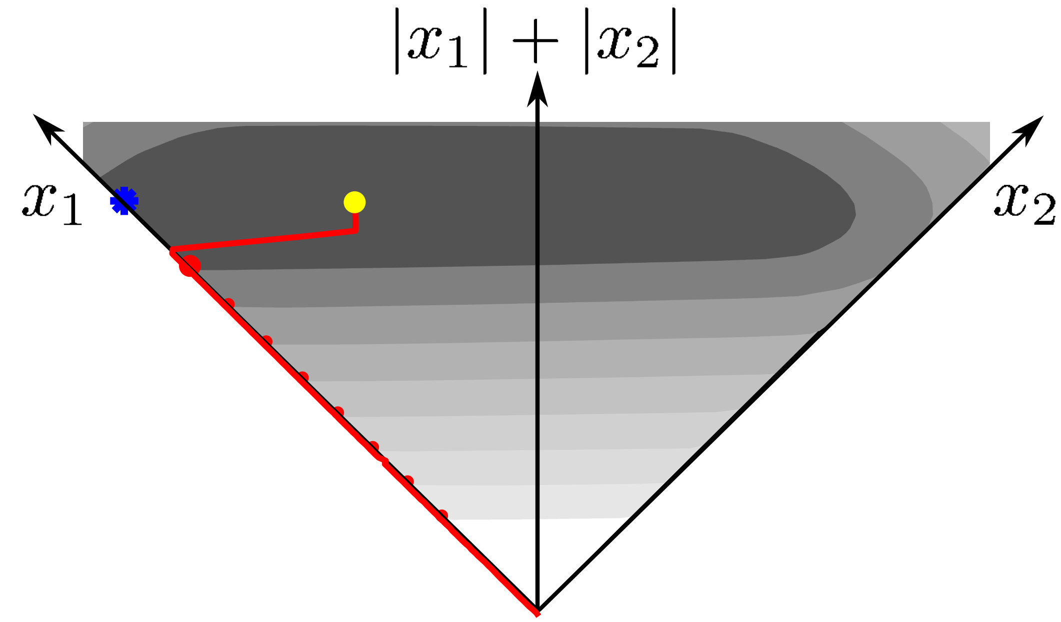
yellow - LS solution ($\lambda = 0$)

red - solutions for different λ

BP



CBP - Taylor1

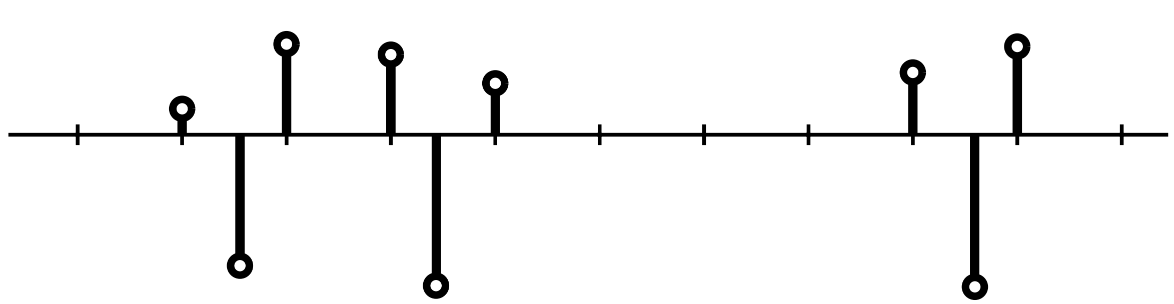


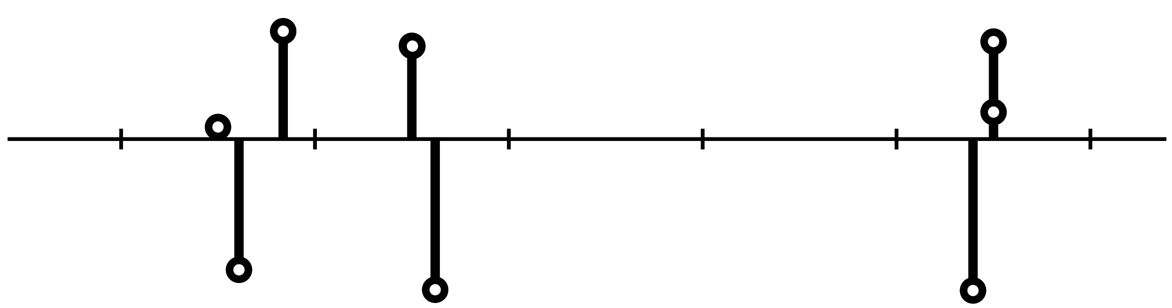
gray - reconstruction error

blue - L0 solution

yellow - LS solution ($\lambda = 0$)

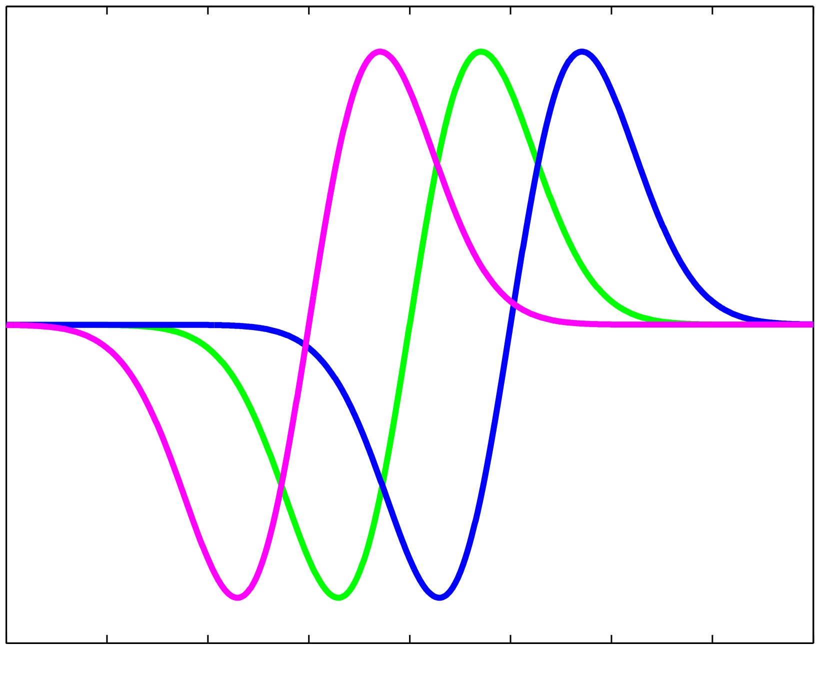
red - solutions for different λ



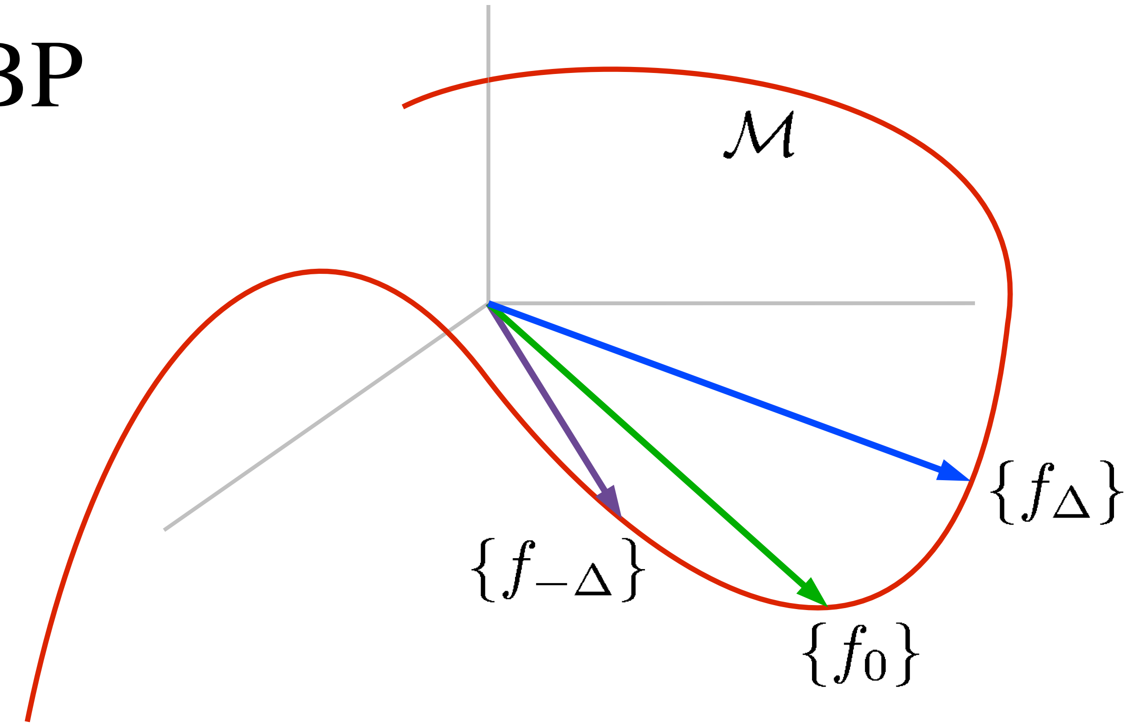


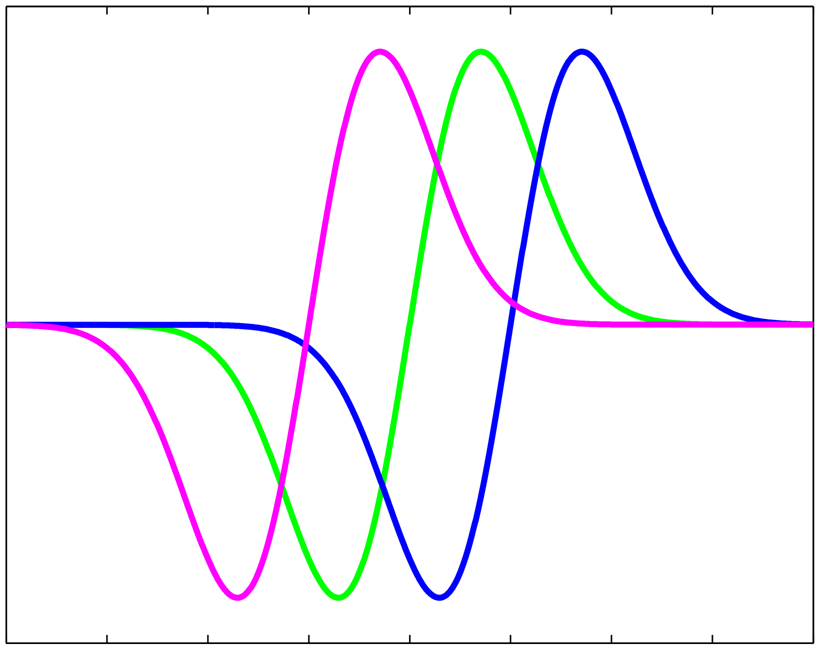
upward: estimated locations/amplitudes, with optimized $\{\lambda, \Delta\}$

downward: true waveform locations/amplitudes

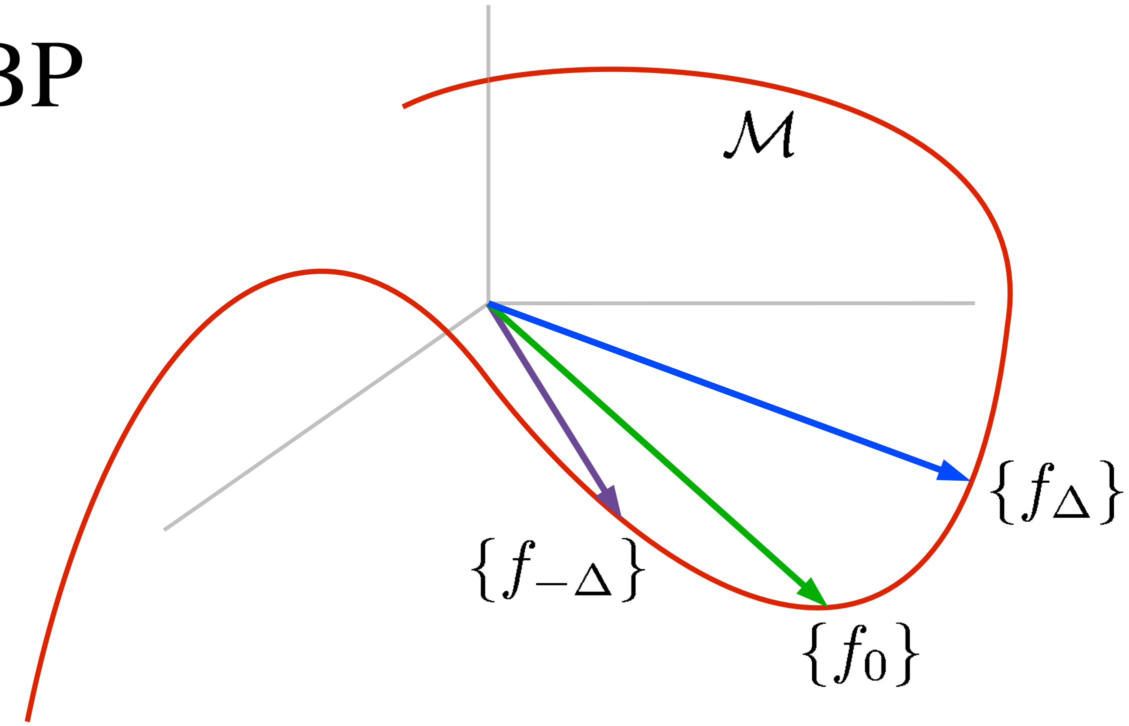


BP

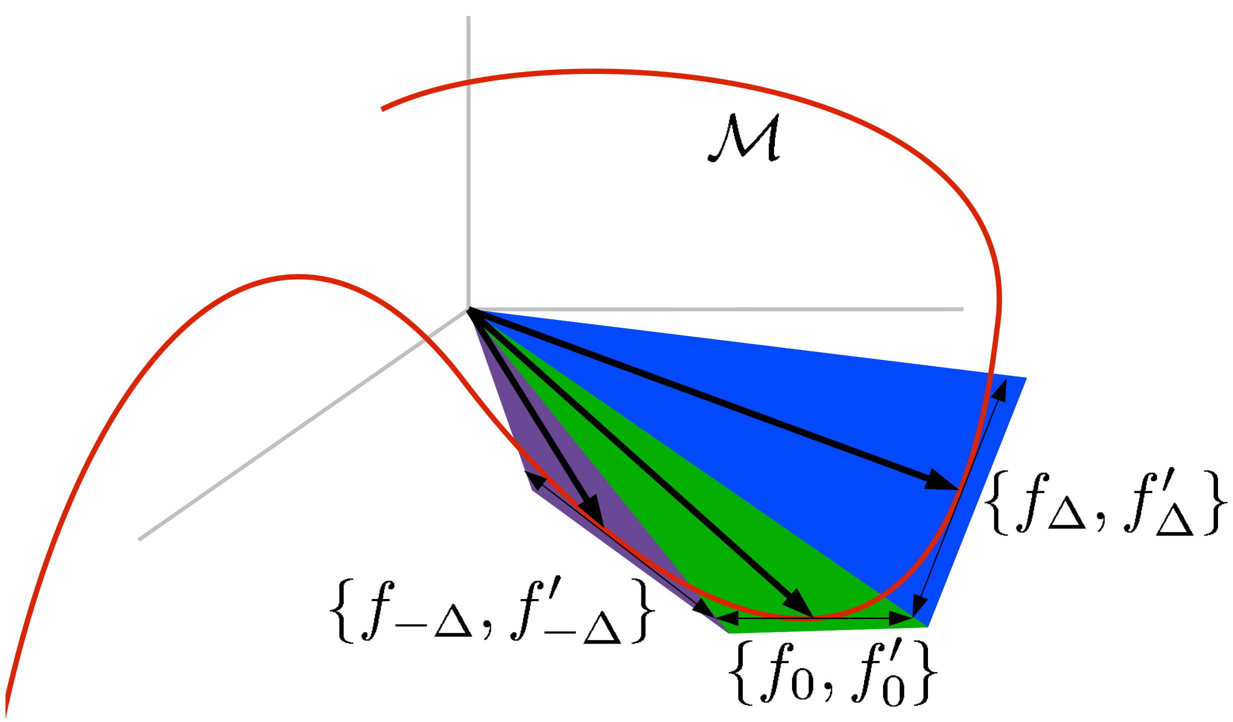




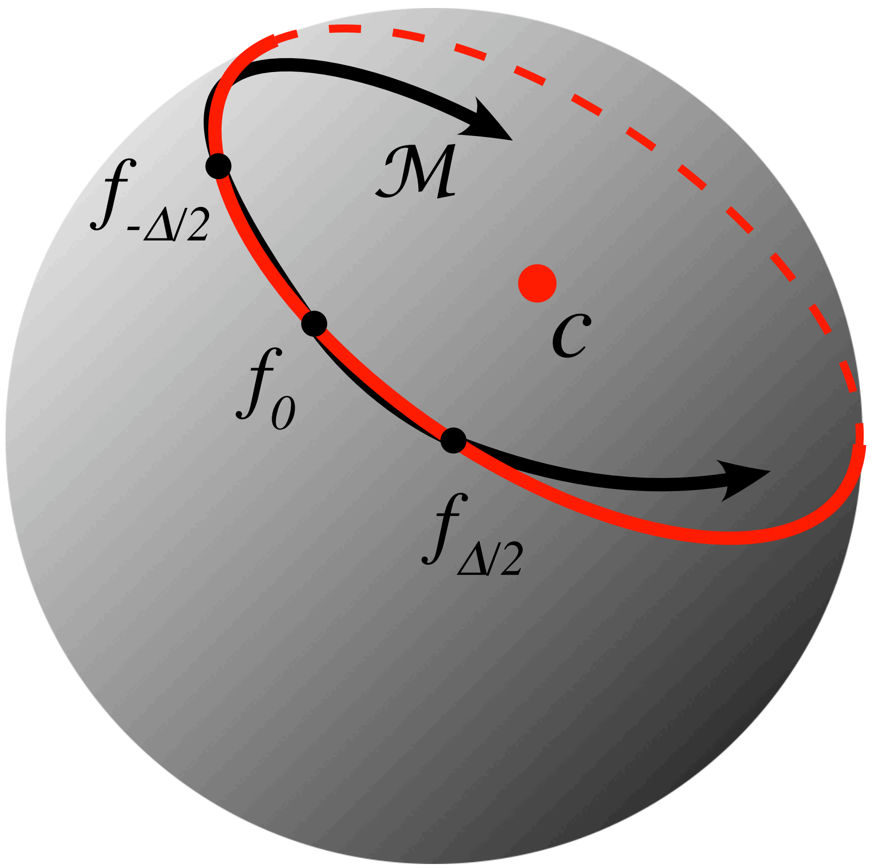
BP



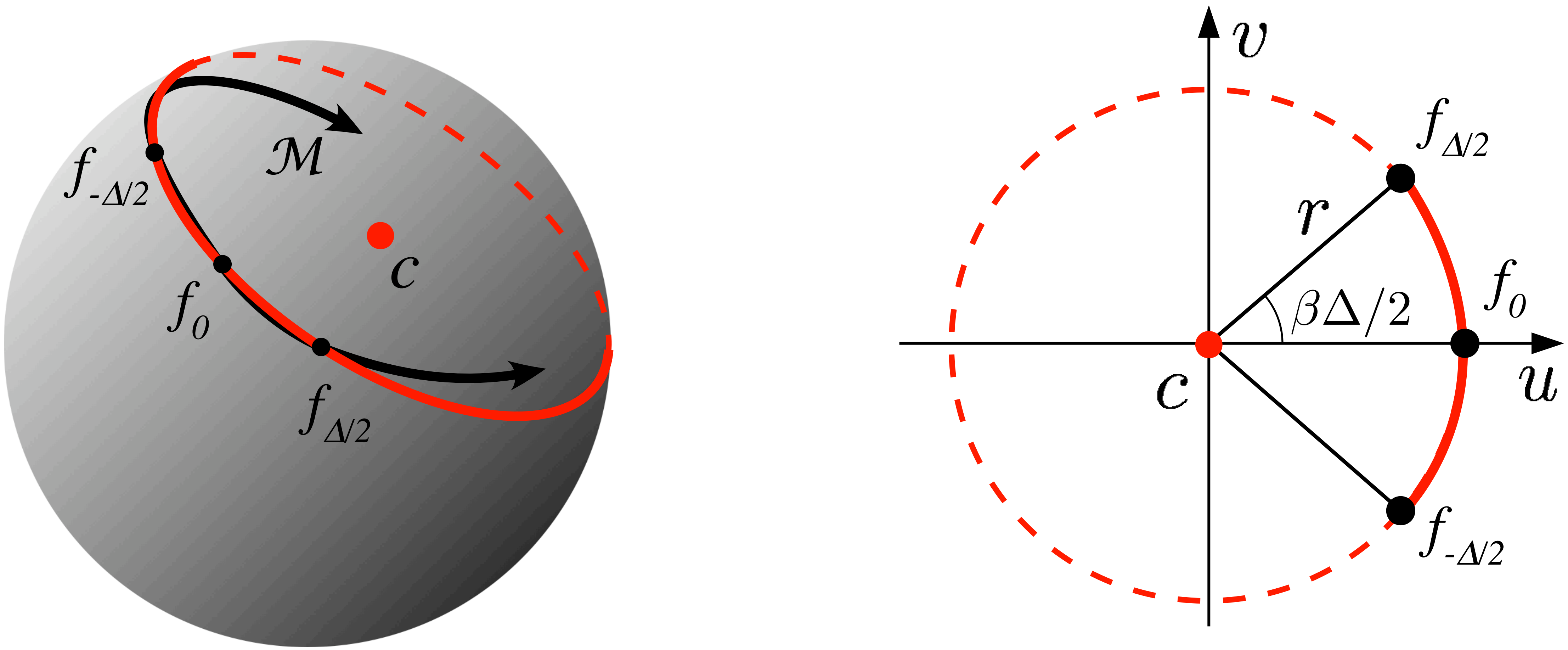
CBP - Taylor1



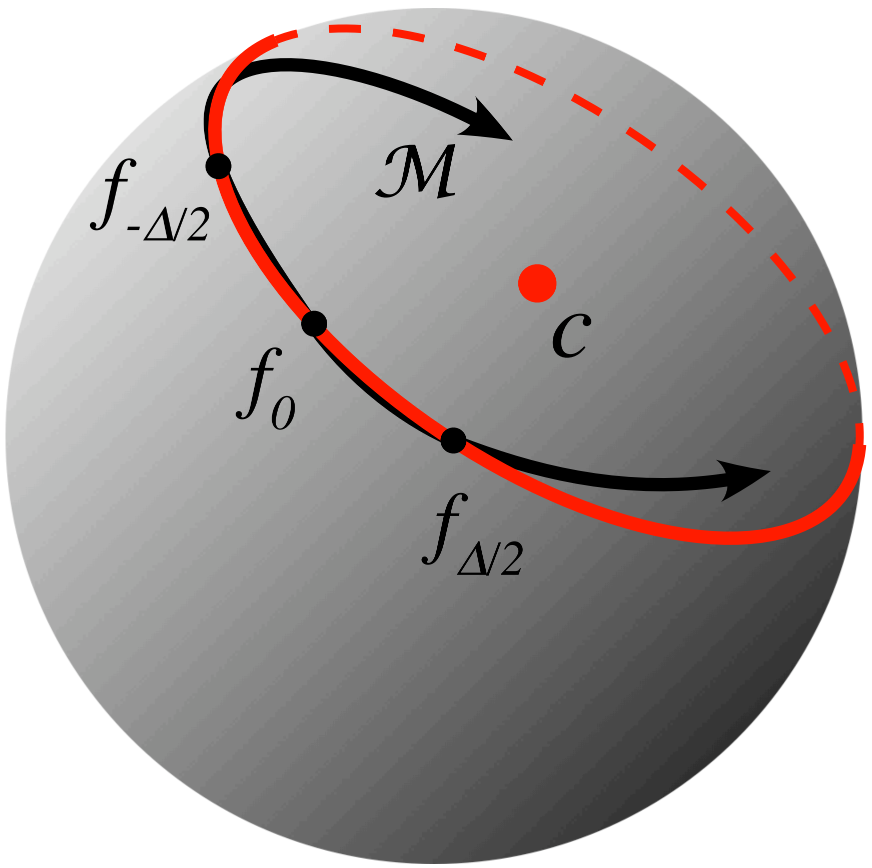
Polar interpolation

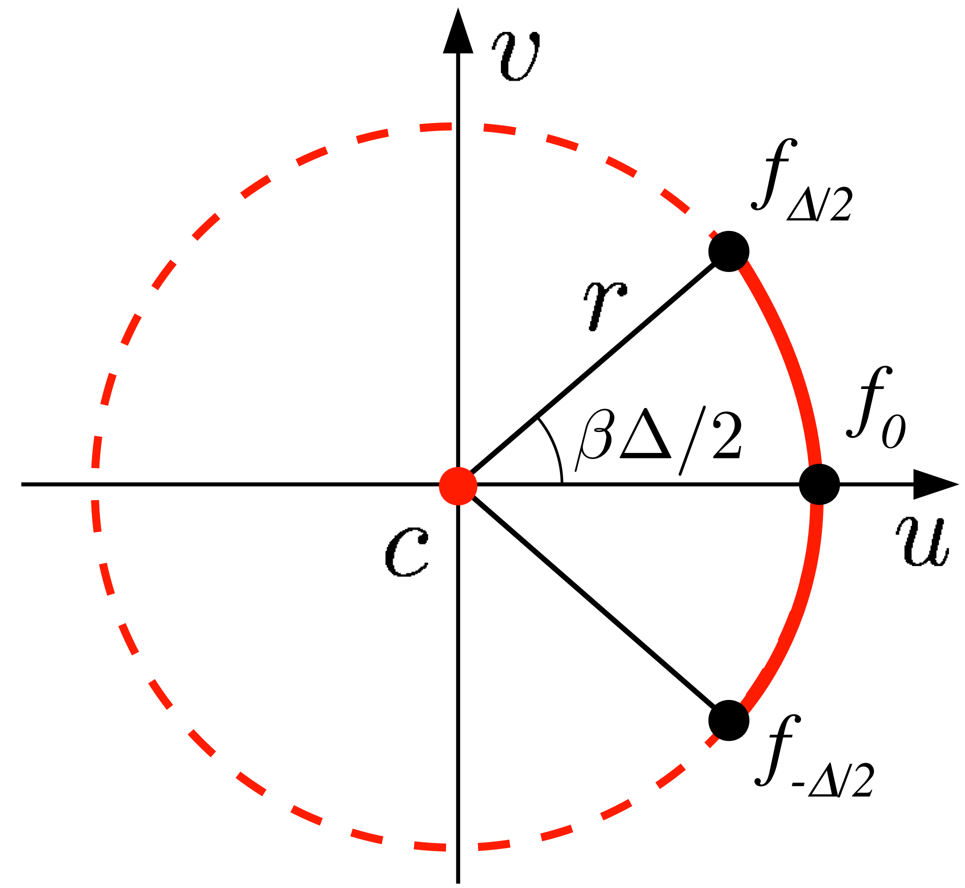


Polar interpolation



Polar interpolation

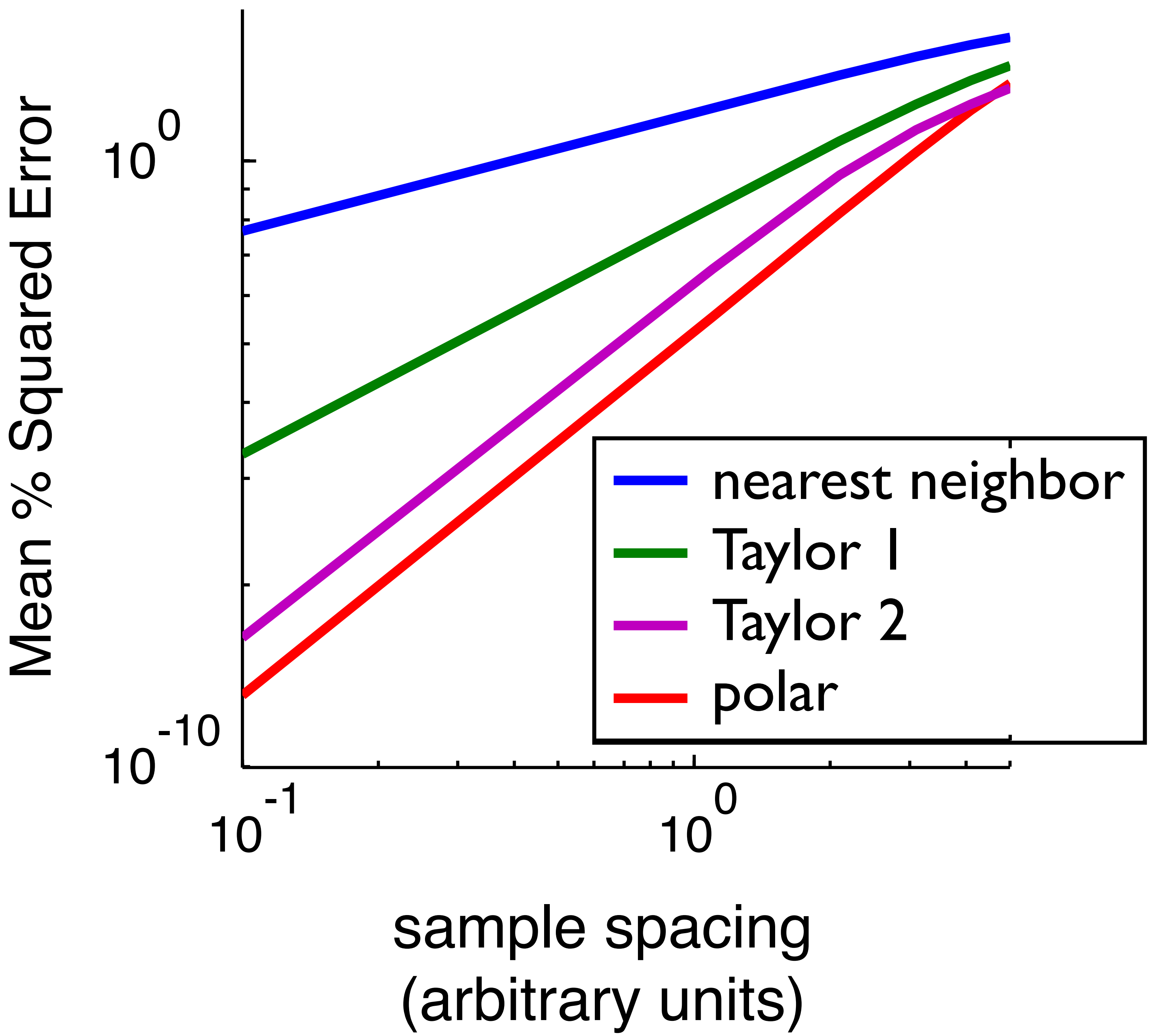




$$f(t+\tau) \approx c(t) + r\cos(\beta\tau)u(t) + r\sin(\beta\tau)v(t), \quad |\tau| < \Delta/2$$

$$\begin{pmatrix} f(t-\frac{\Delta}{2}) \\ f(t) \\ f(t-\frac{\Delta}{2}) \end{pmatrix} = \begin{pmatrix} 1 & r\cos(\beta\Delta/2) & -r\sin(\beta\Delta/2) \\ 1 & r & 0 \\ 1 & r\cos(\beta\Delta/2) & r\sin(\beta\Delta/2) \end{pmatrix} \begin{pmatrix} c(t) \\ u(t) \\ v(t) \end{pmatrix}$$

Interpolator convergence rates



Continuous basis pursuit - polar

Choose $\{\Delta, \beta, r\}$ such that

$$f(t+\tau) \approx c(t) + r\cos(\beta\tau)u(t) + r\sin(\beta\tau)v(t), \quad |\tau| < \Delta/2$$

Continuous basis pursuit - polar

Choose $\{\Delta, \beta, r\}$ such that

$$f(t+\tau) \approx c(t) + r\cos(\beta\tau)u(t) + r\sin(\beta\tau)v(t), \quad |\tau| < \Delta/2$$

Construct interpolative dictionary:

$$\{c(t-n\Delta), u(t-n\Delta), v(t-n\Delta)\}$$

Continuous basis pursuit - polar

Choose $\{\Delta, \beta, r\}$ such that

$$f(t+\tau) \approx c(t) + r\cos(\beta\tau)u(t) + r\sin(\beta\tau)v(t), \quad |\tau| < \Delta/2$$

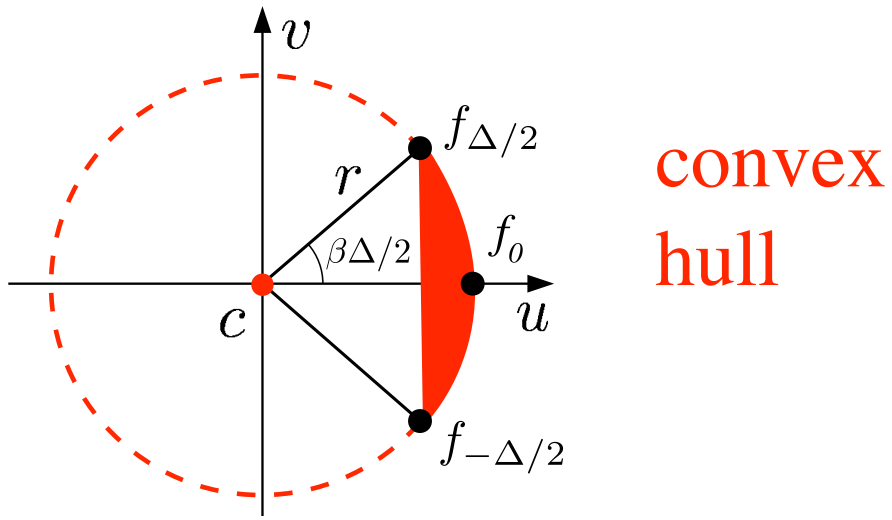
Construct interpolative dictionary:

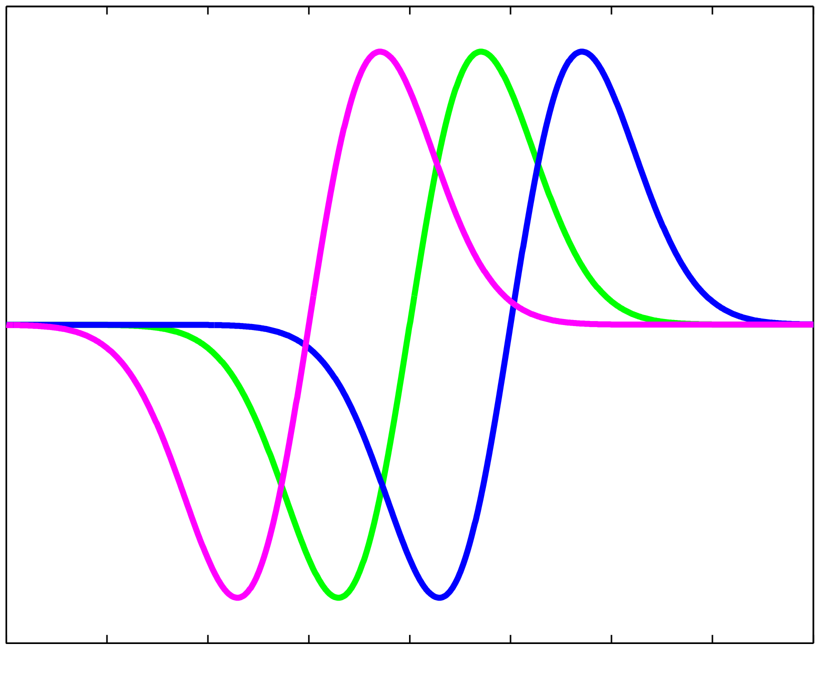
$$\{c(t-n\Delta), u(t-n\Delta), v(t-n\Delta)\}$$

Solve:

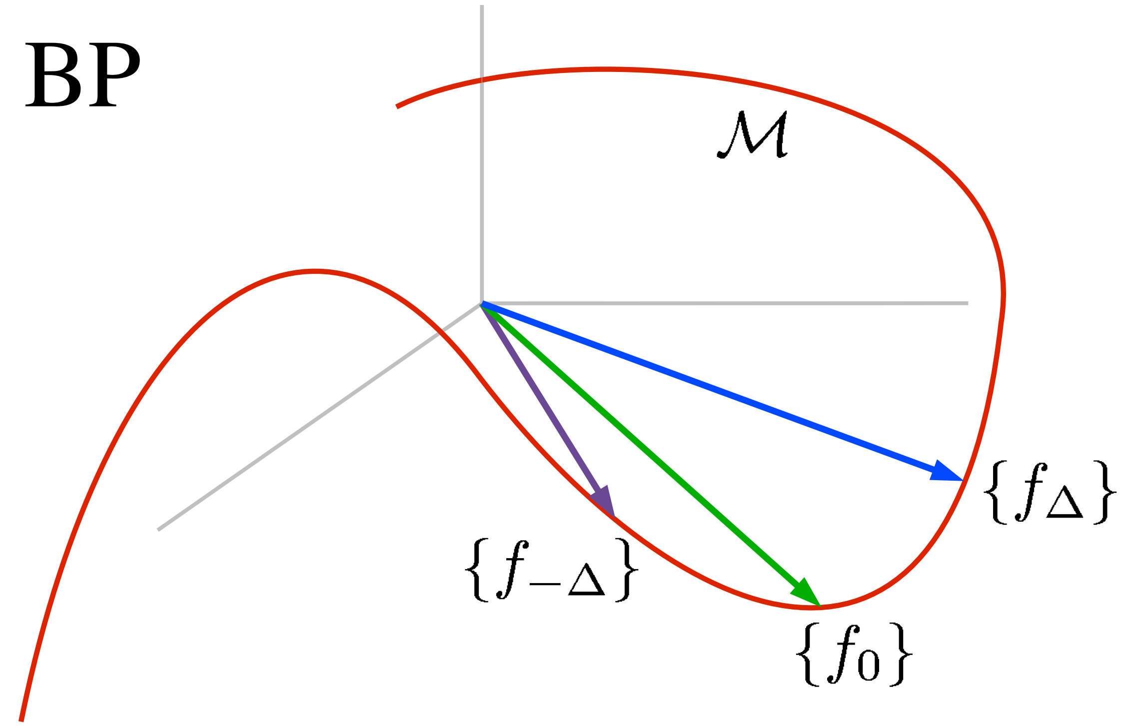
$$\arg\min_{\vec{x},\vec{w},\vec{z}} ||y(t) - \sum_{n=1}^{N} x_n c(t-n\Delta) + w_n u(t-n\Delta) + z_n v(t-n\Delta)||_2^2 + \lambda||\vec{x}||_1$$

with
$$x_n \geq 0,$$
$$\sqrt{w_n^2 + z_n^2} \leq r x_n,$$
$$x_n rcos(\beta\Delta/2) \leq w_n$$

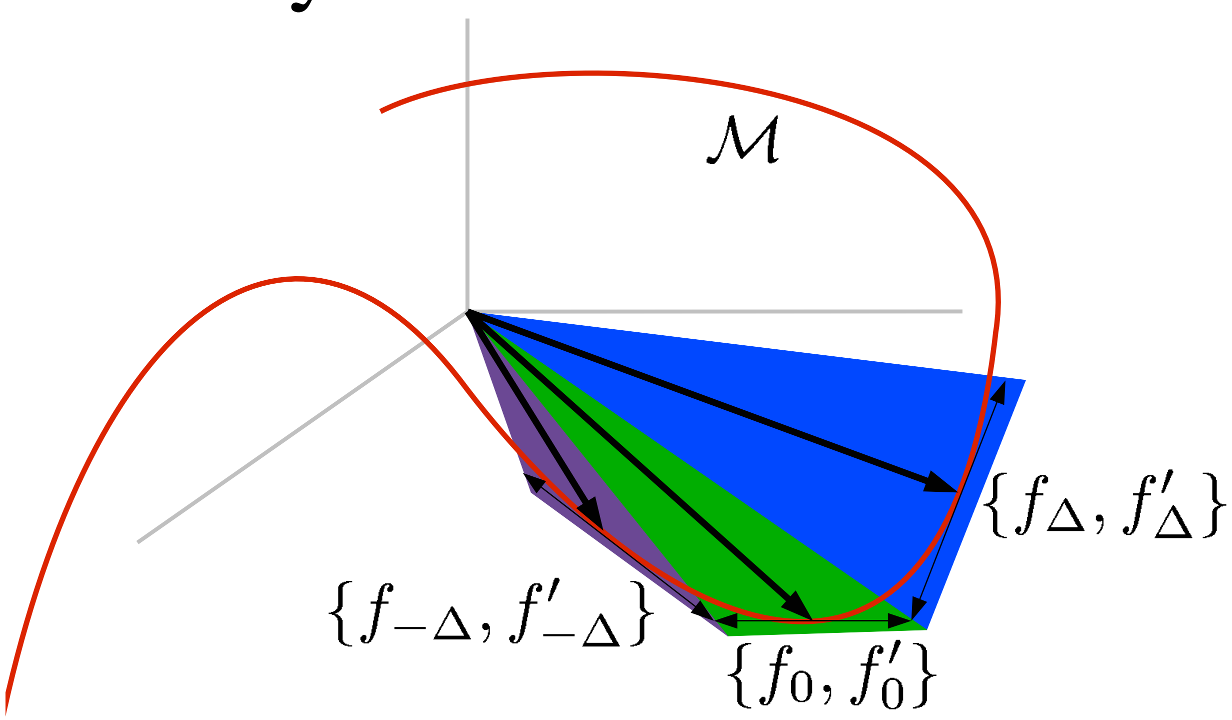


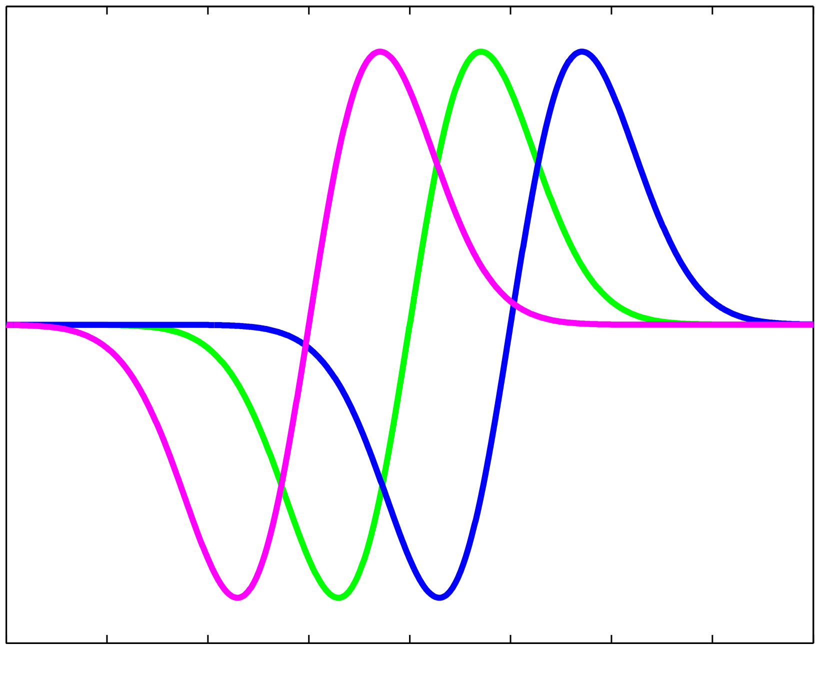


BP

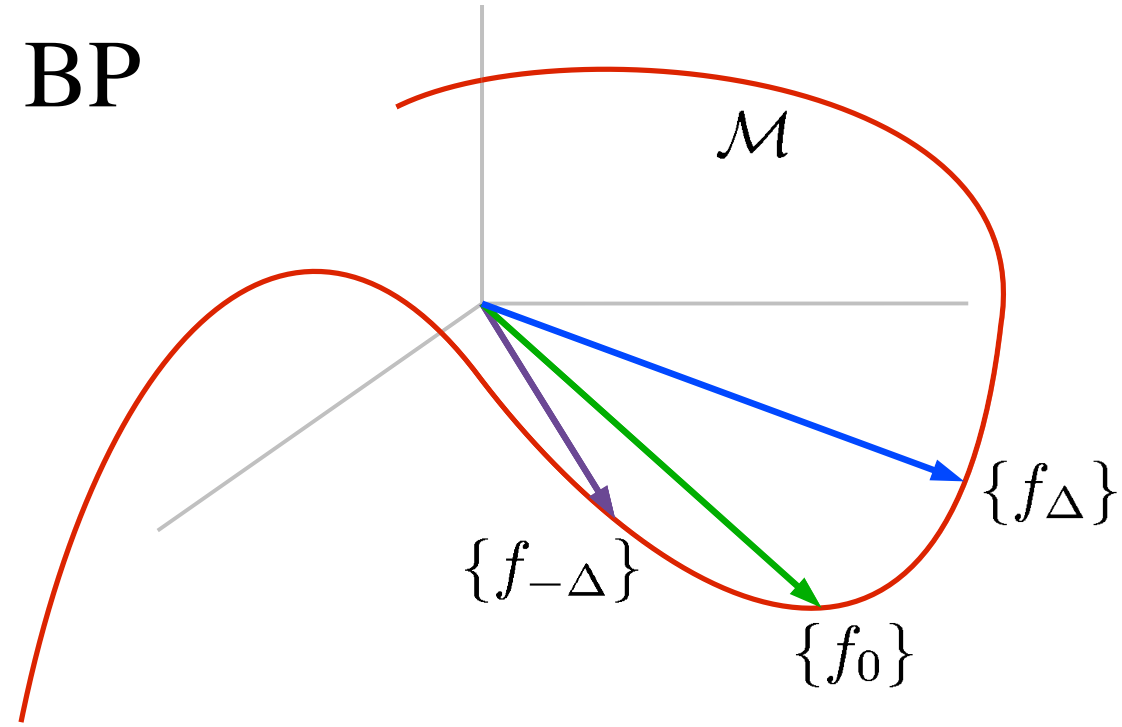


CBP - Taylor1

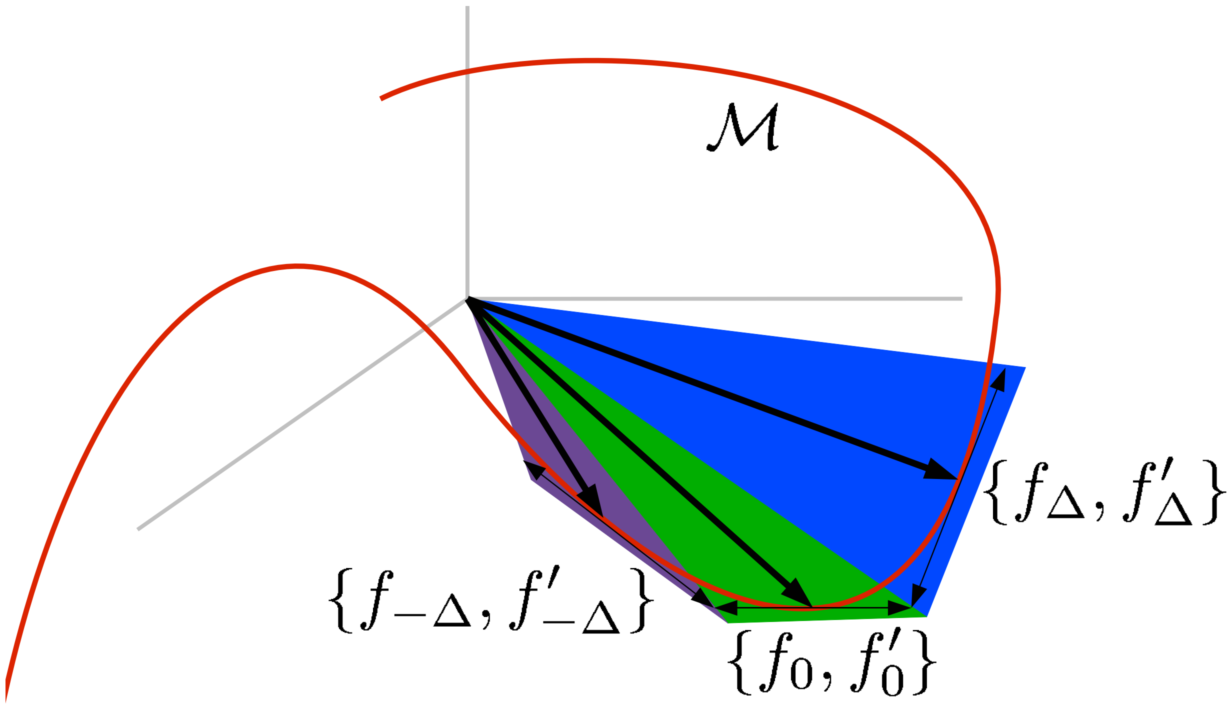




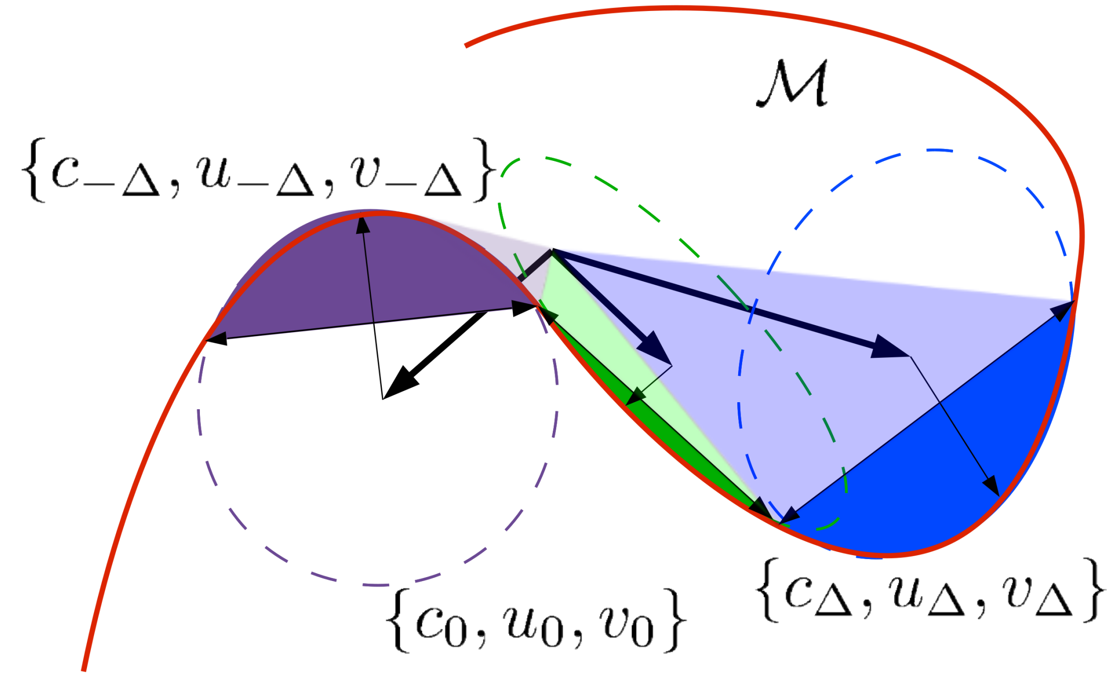
BP

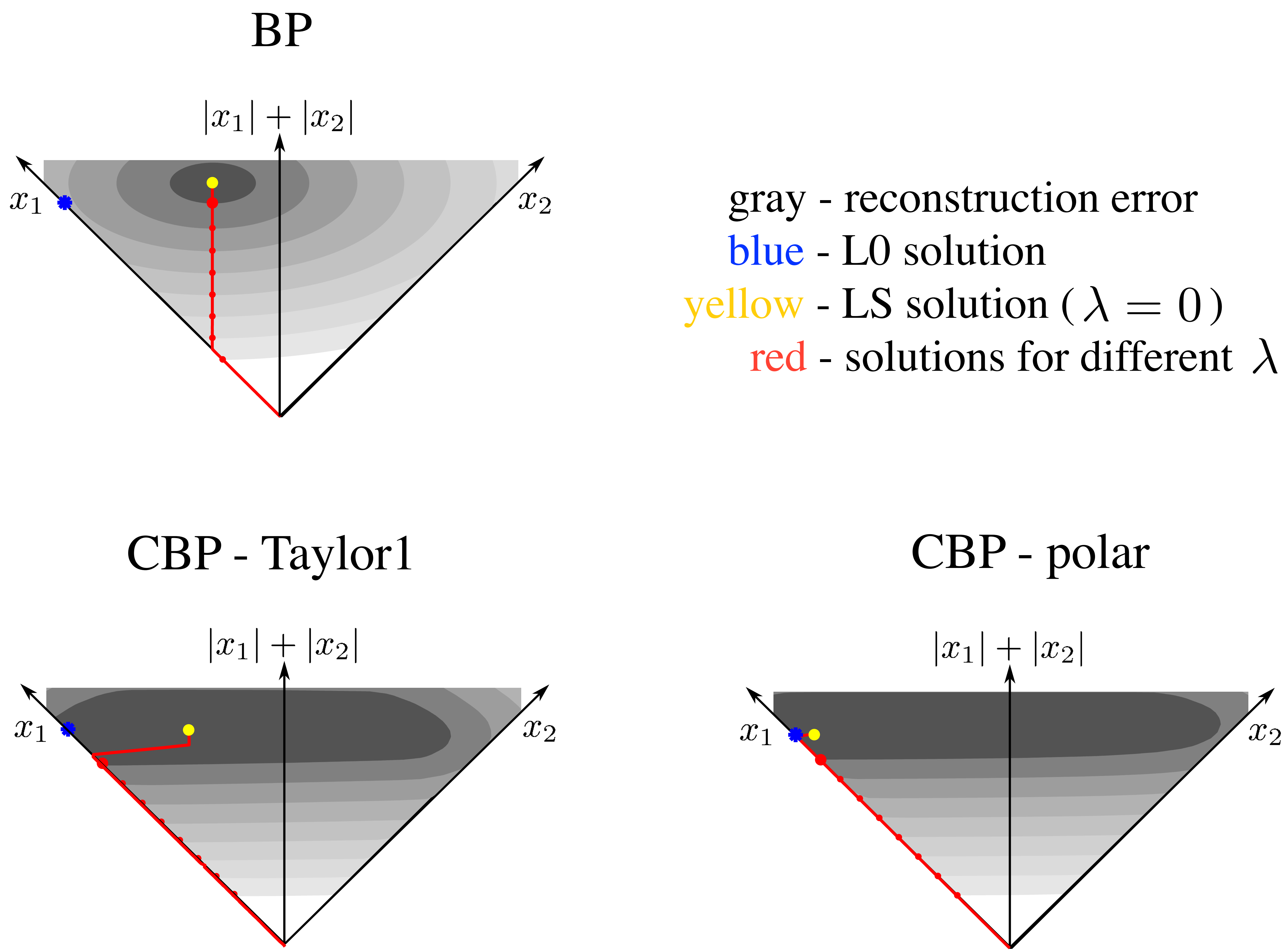


CBP - Taylor1

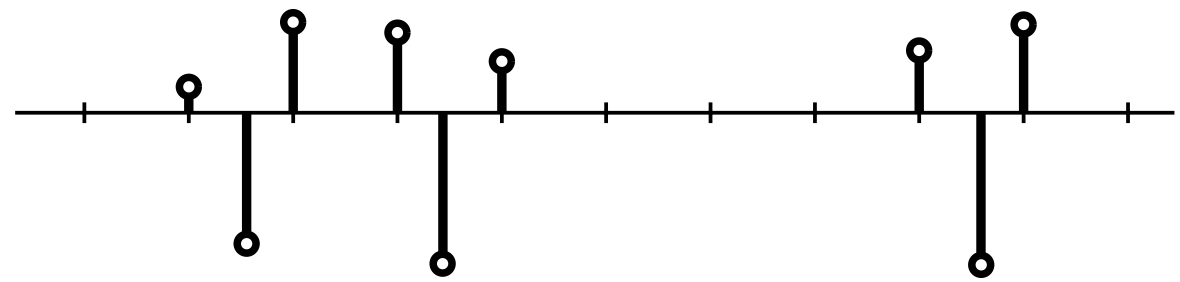


CBP - polar

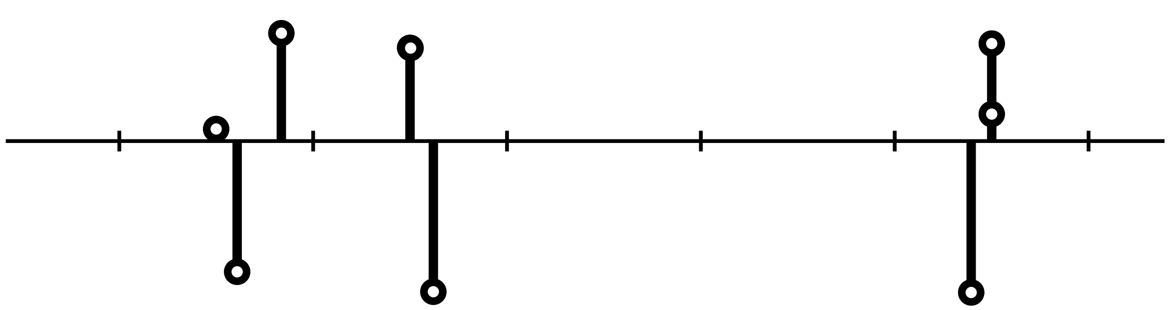



BP
$|x_1| + |x_2|$
x_1
x_2
gray - reconstruction error
blue - L0 solution
yellow - LS solution ($\lambda = 0$)
red - solutions for different λ
CBP - Taylor1
$|x_1| + |x_2|$
x_1
x_2
CBP - polar
$|x_1| + |x_2|$
x_1
x_2

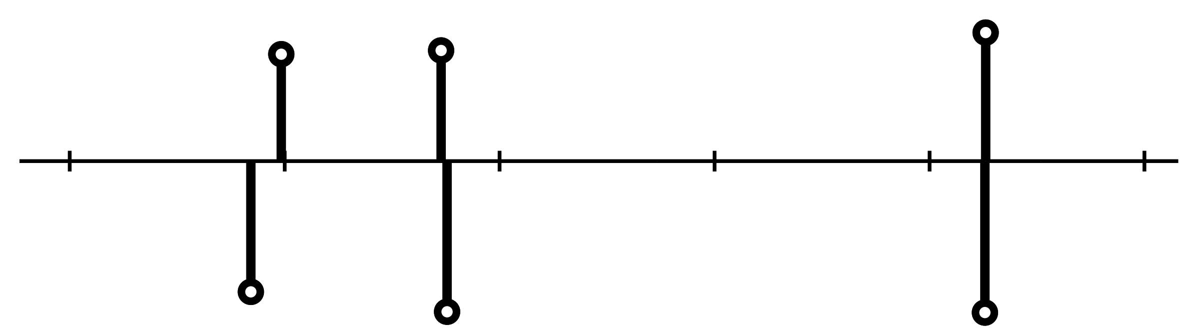
BP



CBP - Taylor1

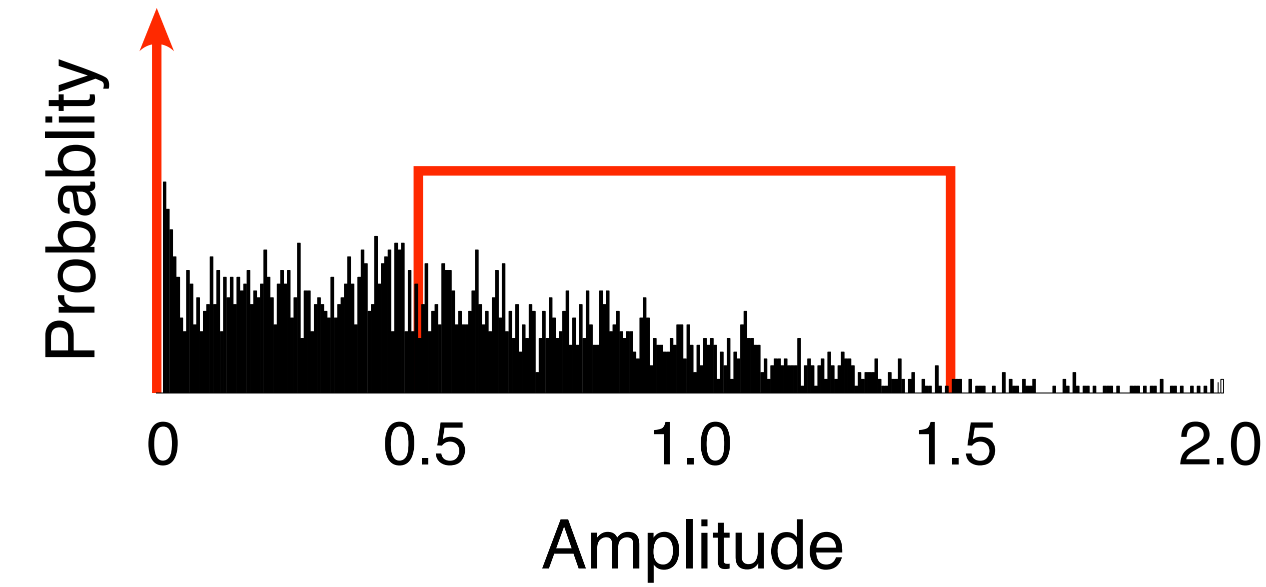


CBP - polar

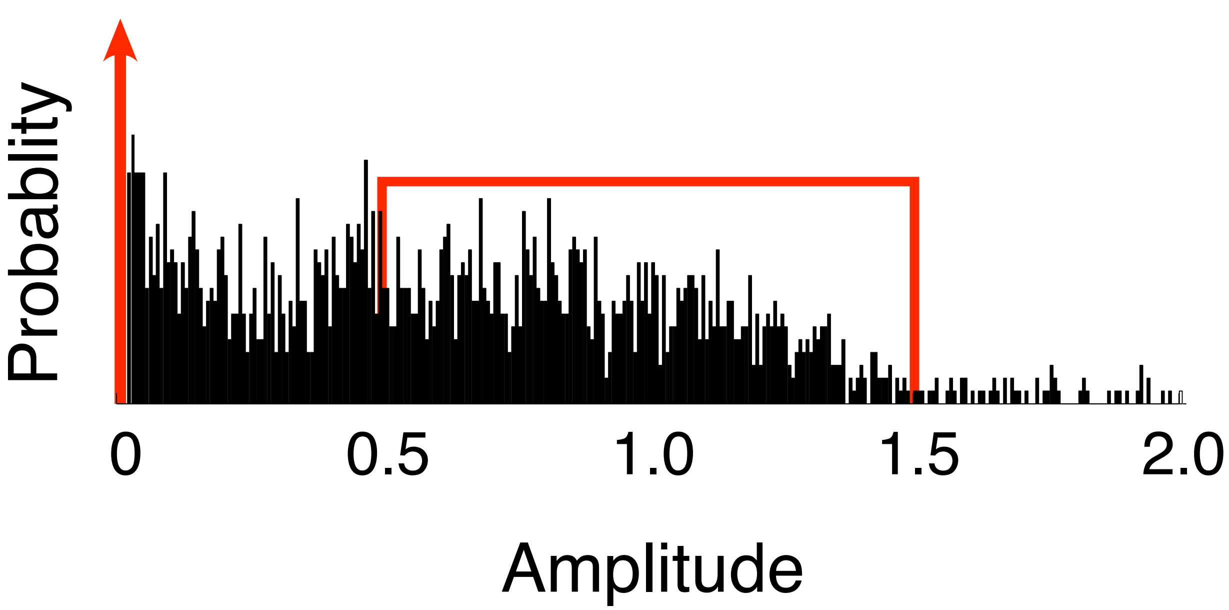


upward: estimated locations/amplitudes, with optimized $\{\lambda, \Delta\}$
downward: true waveform locations/amplitudes

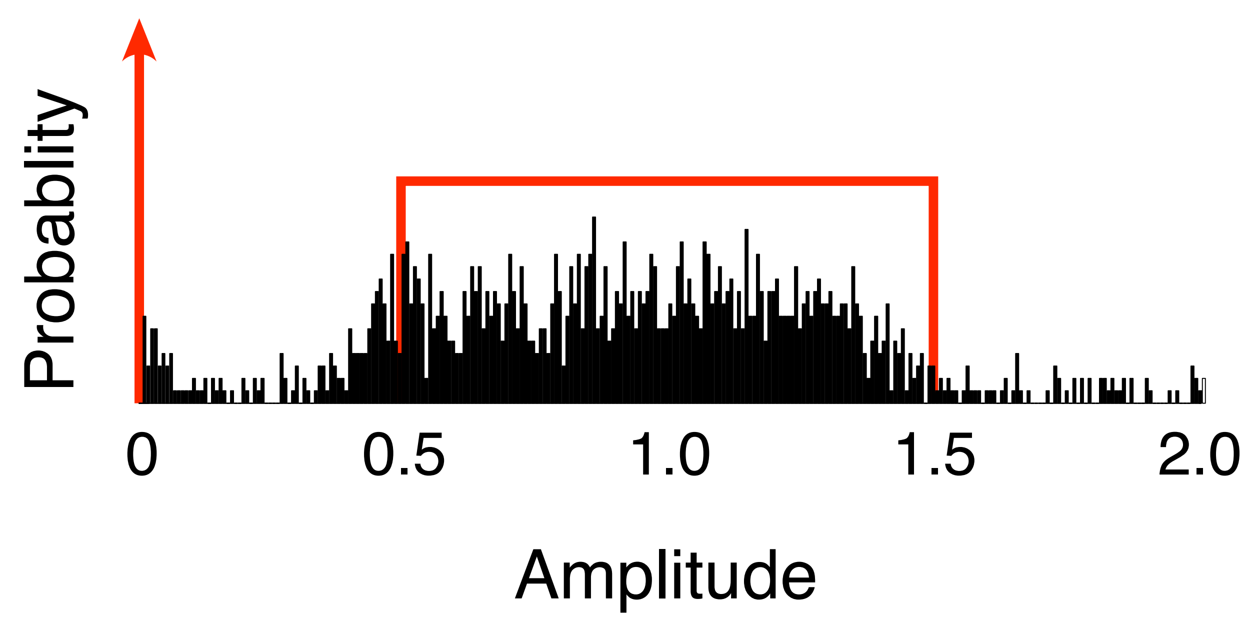
BP



CBP - Taylor1



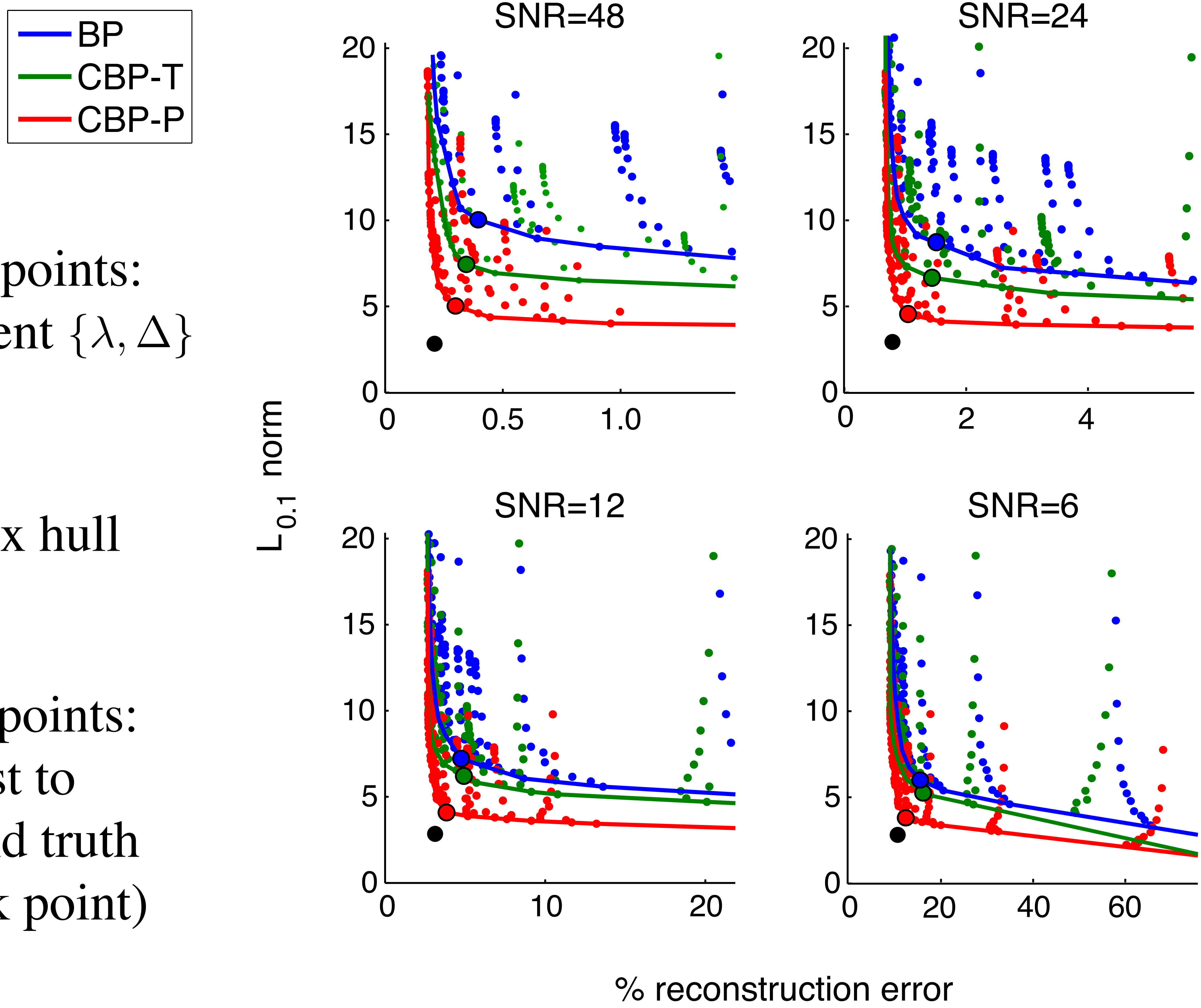
CBP - polar



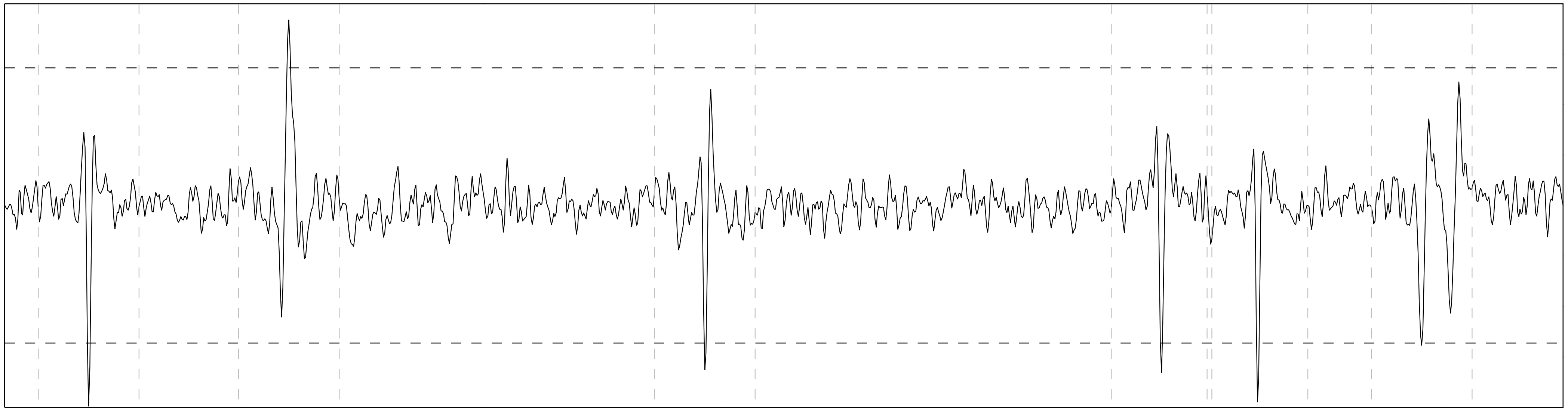
small points: different $\{\lambda, \Delta\}$

lines: convex hull

large points: closest to ground truth (black point)



Application to neural “spike sorting”



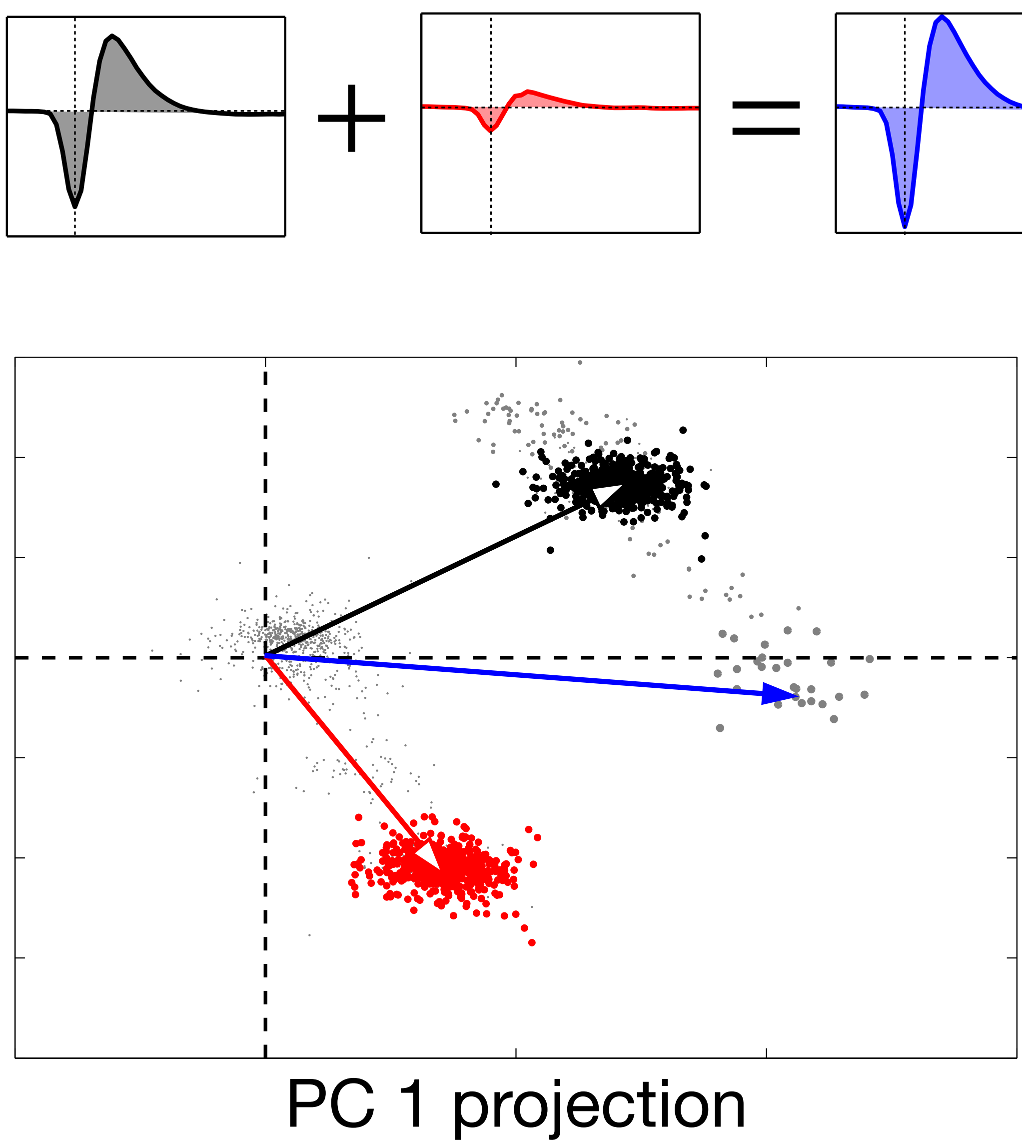
Widely-used solution:

1. Threshold to find segments containing spikes
2. Reduce dimensionality of segments using PCA
3. Identify spikes using clustering (e.g., K-means)

Guaranteed failure for overlapping spikes!

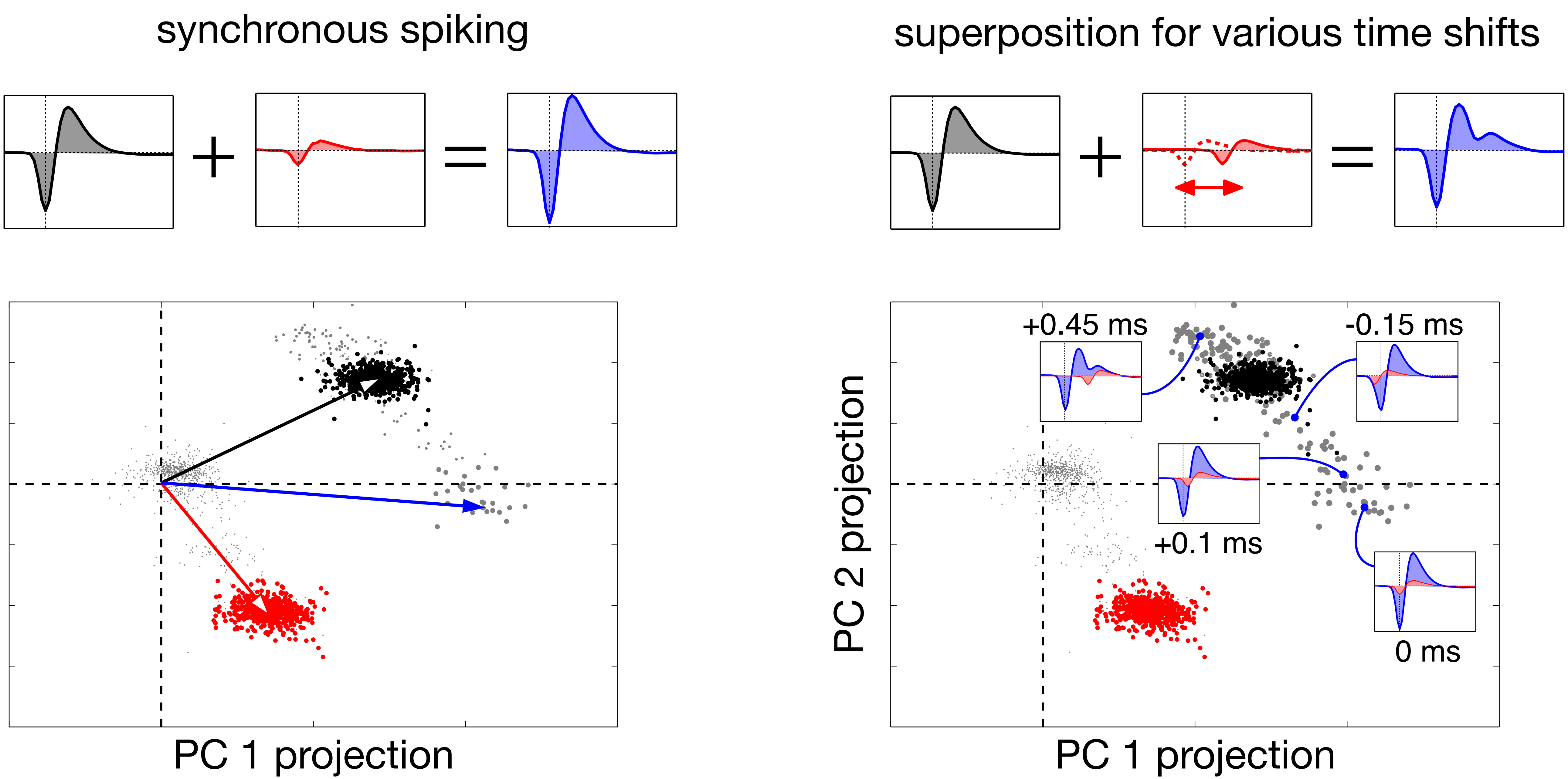
Failures of clustering for near-synchronous spikes

synchronous spiking



[Pillow et. al. 2013]

Failures of clustering for near-synchronous spikes



[Pillow et. al. 2013]

Initialize f (random, or use K-means cluster centers)

[Ekanadham et al, 2013]

Initialize f (random, or use K-means cluster centers)

Use CBP to solve for $\{x_n, w_n, z_n\}$:

$$\arg\min_{\vec{x},\vec{w},\vec{z}} ||y(t) - \sum_{n=1}^{N} x_n c(t-n\Delta) + w_n u(t-n\Delta) + z_n v(t-n\Delta)||_2^2 + \lambda||\vec{x}||_1$$

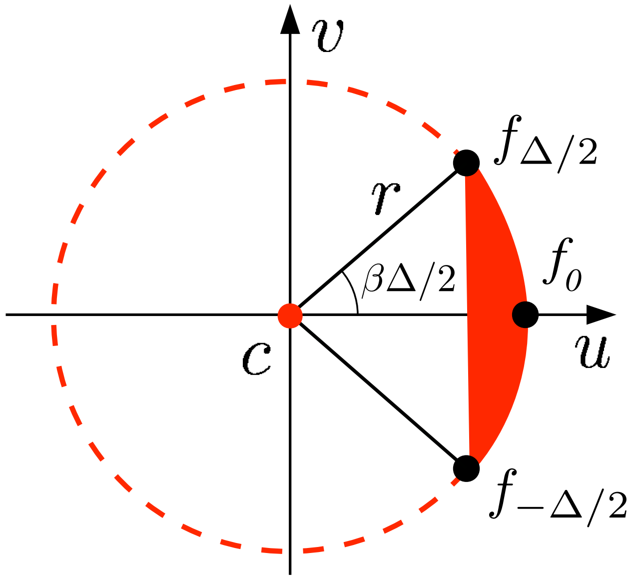
[Ekanadham et al, 2013]

Initialize f (random, or use K-means cluster centers)

Use CBP to solve for $\{x_n, w_n, z_n\}$:

$$\arg\min_{\vec{x},\vec{w},\vec{z}} ||y(t) - \sum_{n=1}^{N} x_n c(t - n\Delta) + w_n u(t - n\Delta) + z_n v(t - n\Delta)||_2^2 + \lambda ||\vec{x}||_1$$

Convert $\{x_n, w_n, z_n\}$ back to $\{a_j, \tau_j\}$



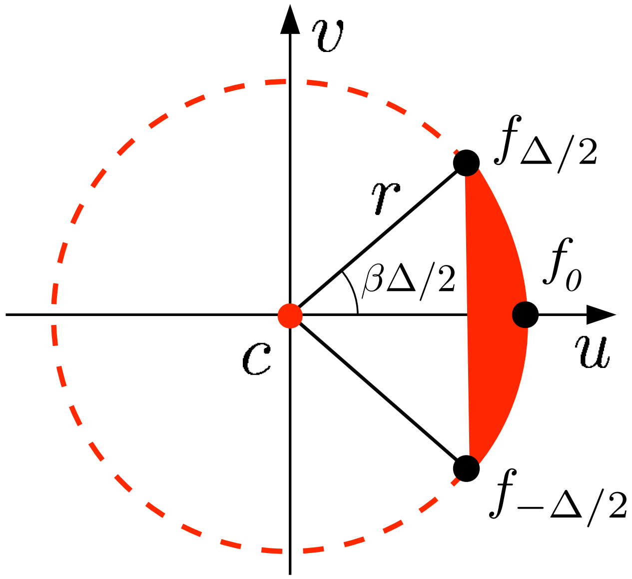
[Ekanadham et al, 2013]

Initialize f (random, or use K-means cluster centers)

Use CBP to solve for $\{x_n, w_n, z_n\}$:

$$\arg\min_{\vec{x},\vec{w},\vec{z}} ||y(t) - \sum_{n=1}^{N} x_n c(t - n\Delta) + w_n u(t - n\Delta) + z_n v(t - n\Delta)||_2^2 + \lambda||\vec{x}||_1$$

Convert $\{x_n, w_n, z_n\}$ back to $\{a_j, \tau_j\}$



Solve for f (least squares):

$$\arg\min_{f} ||y(t) - \sum_{j} a_j f(t - \tau_j)||_2^2$$

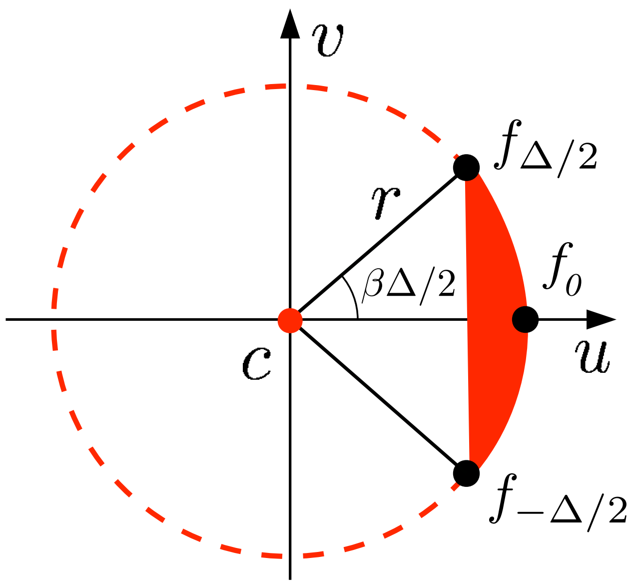
[Ekanadham et al, 2013]

Initialize f (random, or use K-means cluster centers)

Use CBP to solve for $\{x_n, w_n, z_n\}$:

$$\arg\min_{\vec{x},\vec{w},\vec{z}} ||y(t) - \sum_{n=1}^{N} x_n c(t - n\Delta) + w_n u(t - n\Delta) + z_n v(t - n\Delta)||_2^2 + \lambda ||\vec{x}||_1$$

Convert $\{x_n, w_n, z_n\}$ back to $\{a_j, \tau_j\}$



Solve for f (least squares):

$$\arg\min_{f} ||y(t) - \sum_{j} a_j f(t - \tau_j)||_2^2$$

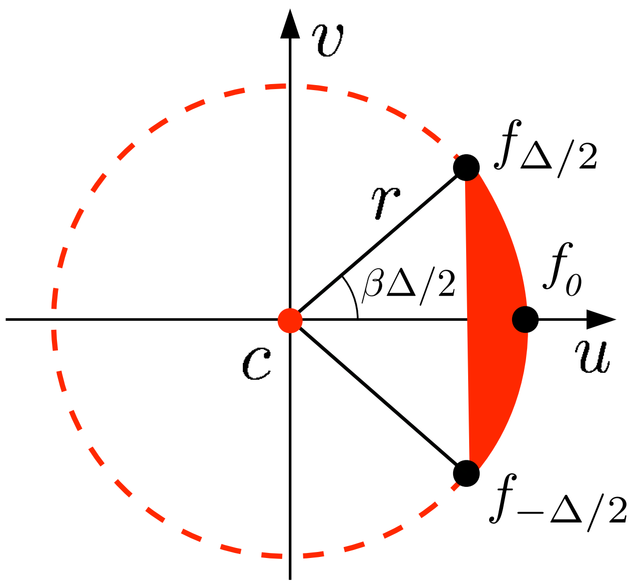
[Ekanadham et al, 2013]

Initialize f (random, or use K-means cluster centers)

Use CBP to solve for $\{x_n, w_n, z_n\}$:

$$\arg\min_{\vec{x},\vec{w},\vec{z}} ||y(t) - \sum_{n=1}^{N} x_n c(t-n\Delta) + w_n u(t-n\Delta) + z_n v(t-n\Delta)||_2^2 + \lambda||\vec{x}||_1$$

Convert $\{x_n, w_n, z_n\}$ back to $\{a_j, \tau_j\}$



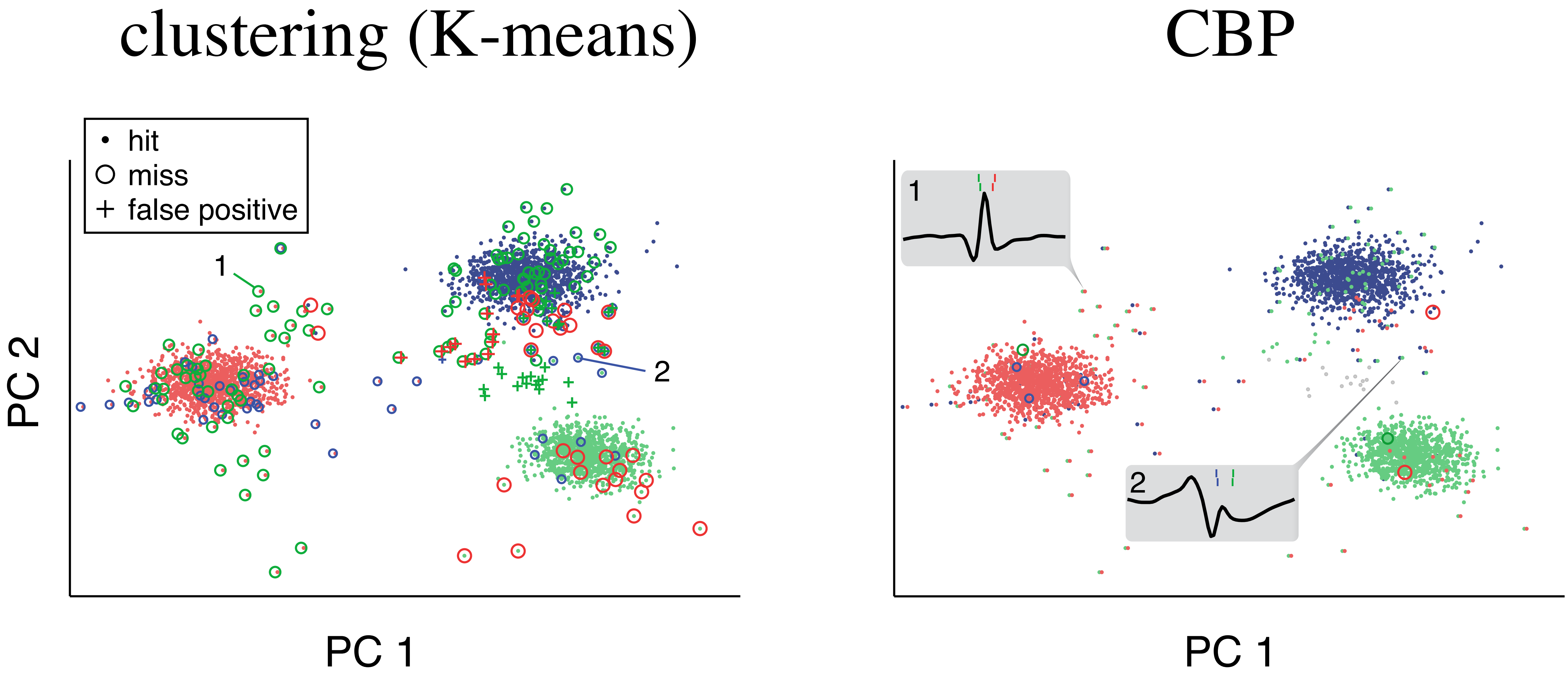
Solve for f (least squares):

$$\arg\min_{f} ||y(t) - \sum_{j} a_j f(t-\tau_j)||_2^2$$

Binarize to obtain final spike estimates

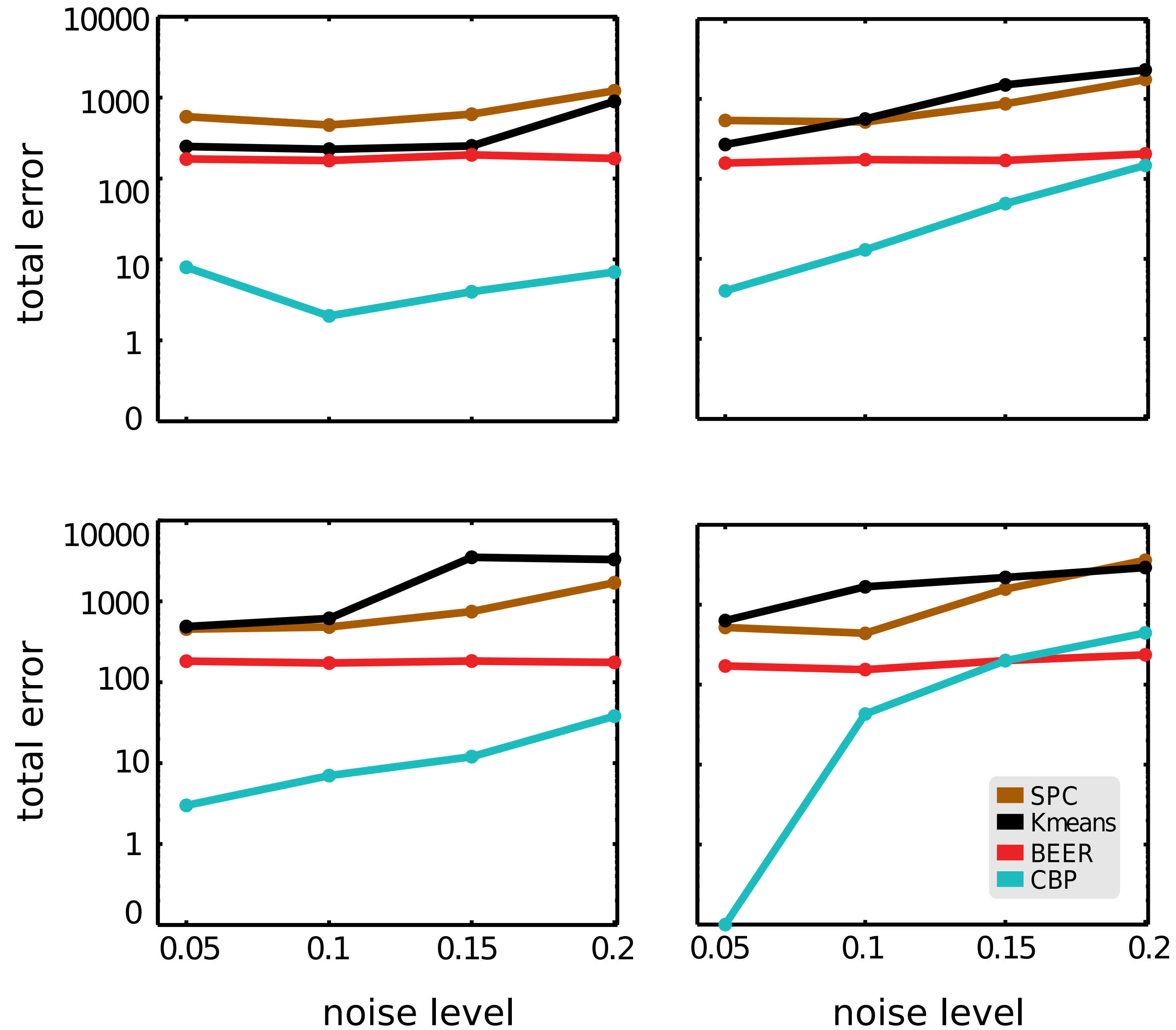
[Ekanadham et al, 2013]

Simulated data [Quiroga et. al. 2004]



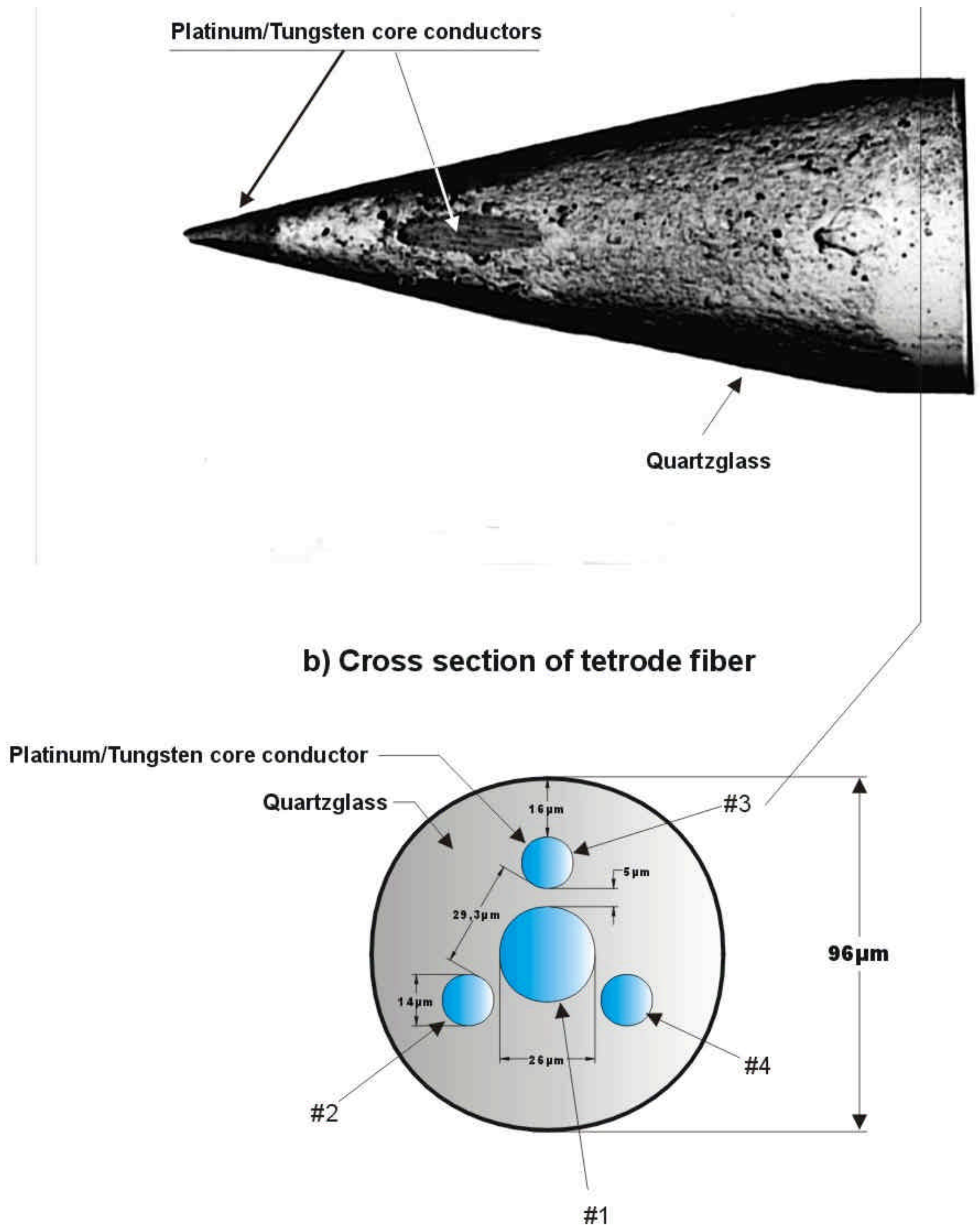
[Ekanadham et al, 2013]

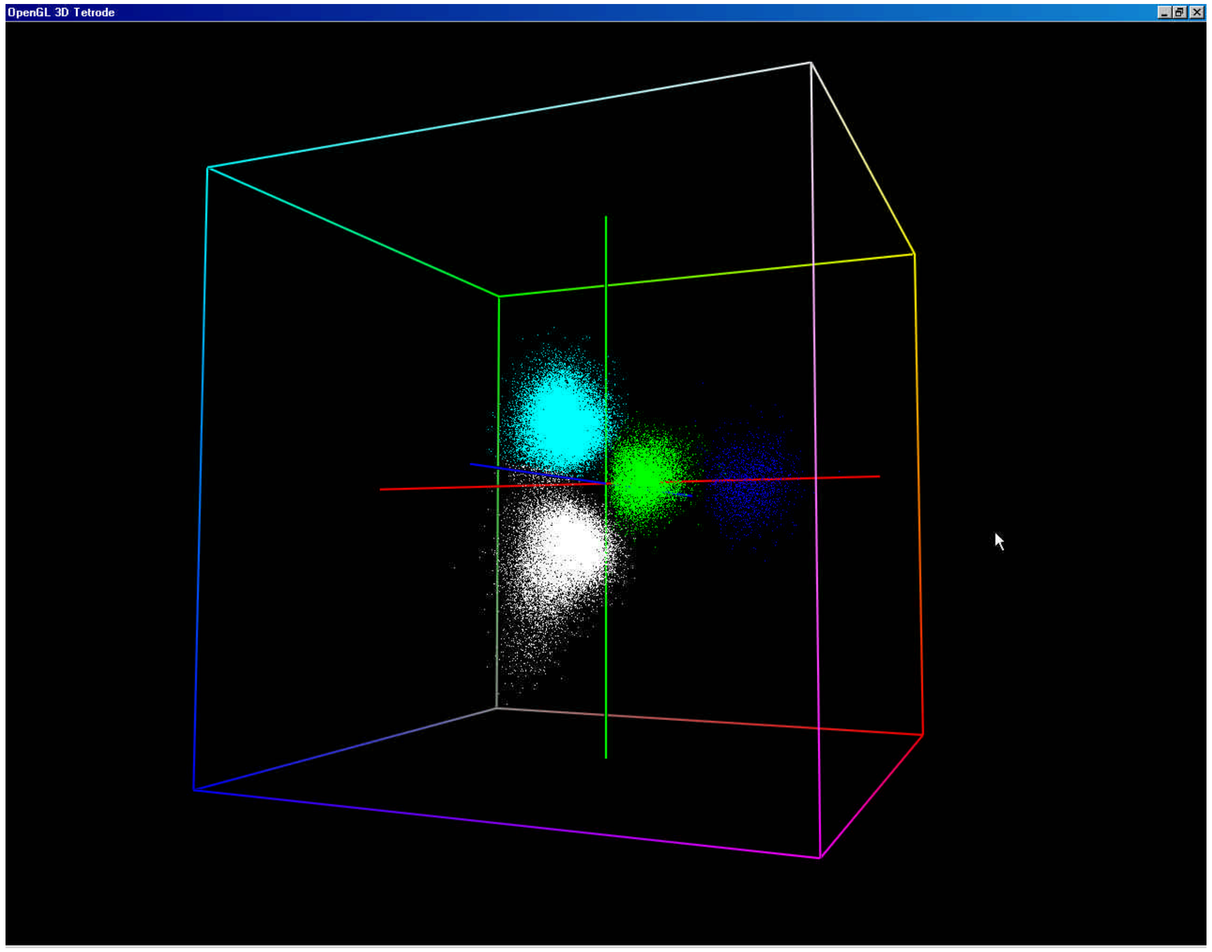
Simulated data (4 data sets) [Quiroga et. al. 2004]



SPC = super-paramagnetic clustering [Quiroga et. al. 2004]

BEER = “best ellipsoid error rate” - elliptical clustering, trained on ground truth data [Harris et. al. 2000]



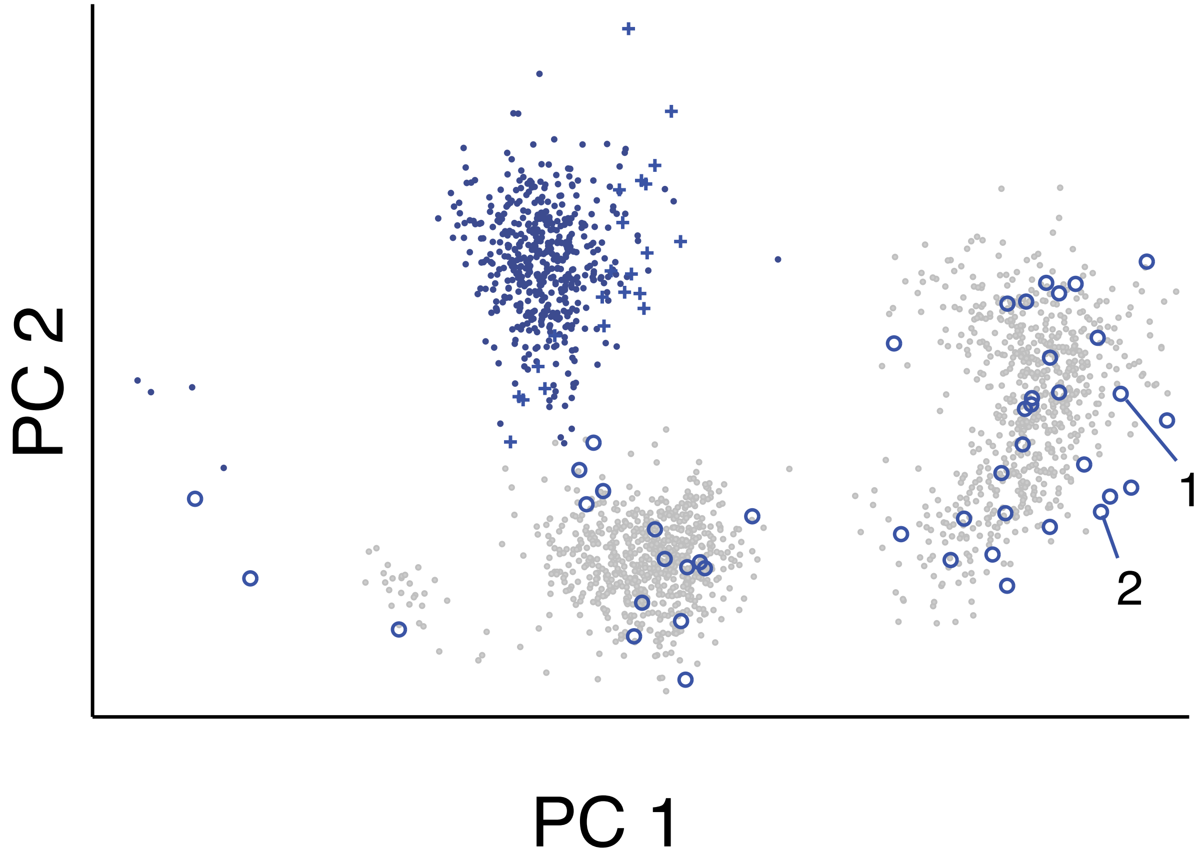


Tetrode electrode assembly

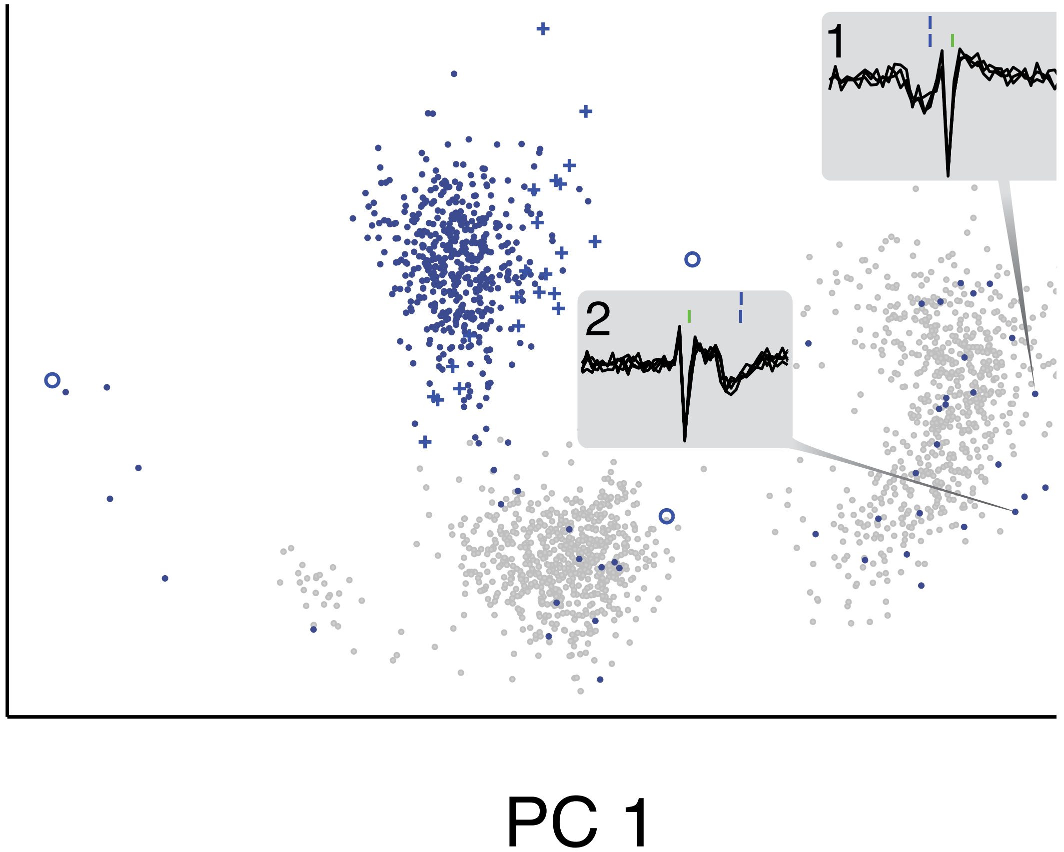
Thomson Recording GmbH, Giessen Germany

Tetrode data, with intracellular ground truth (for one cell) [Harris et. al., 2000]

clustering (K-means)



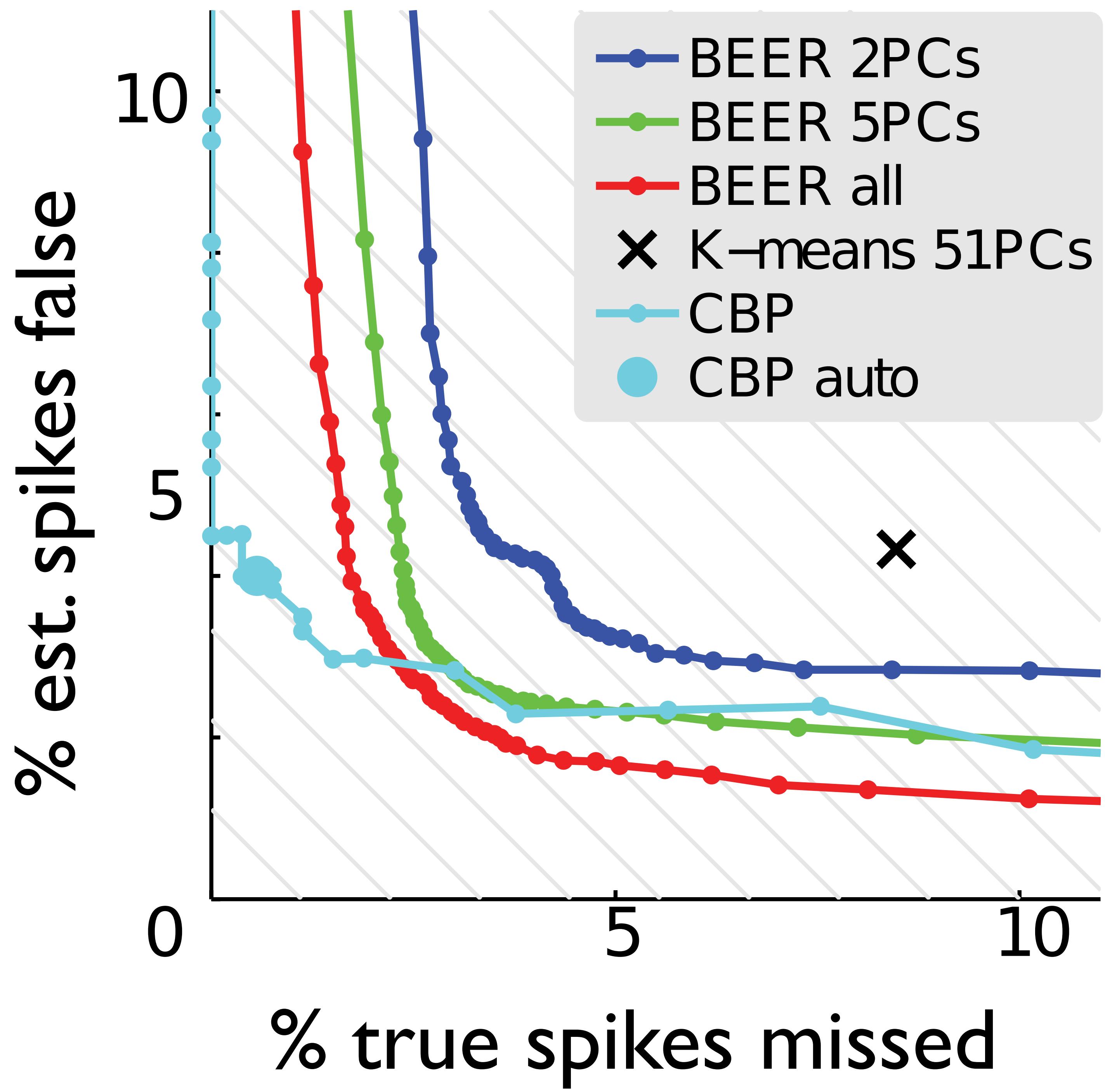
CBP



[Ekanadham et al, 2013]

Tetrode data, intracellular ground truth (for one cell)

[Harris et. al., 2000]



[Ekanadham et al, 2013]

Conclusions

- CBP: Sparse signal decomposition with continuously shifted feature waveforms
  - Translation manifold approximated with circular arcs
  - Can significantly outperform standard BP/LASSO
  - State-of-the-art results in neural spike identification
- Extensions:
  - Priors on coefficient amplitudes
  - Proximal methods / ADMM
  - Other transformations (dilation, rotation, modulation)