Spectral Compressive Sensing

Marco F. Duarte

UMASS AMHERST

Portions are joint work with:

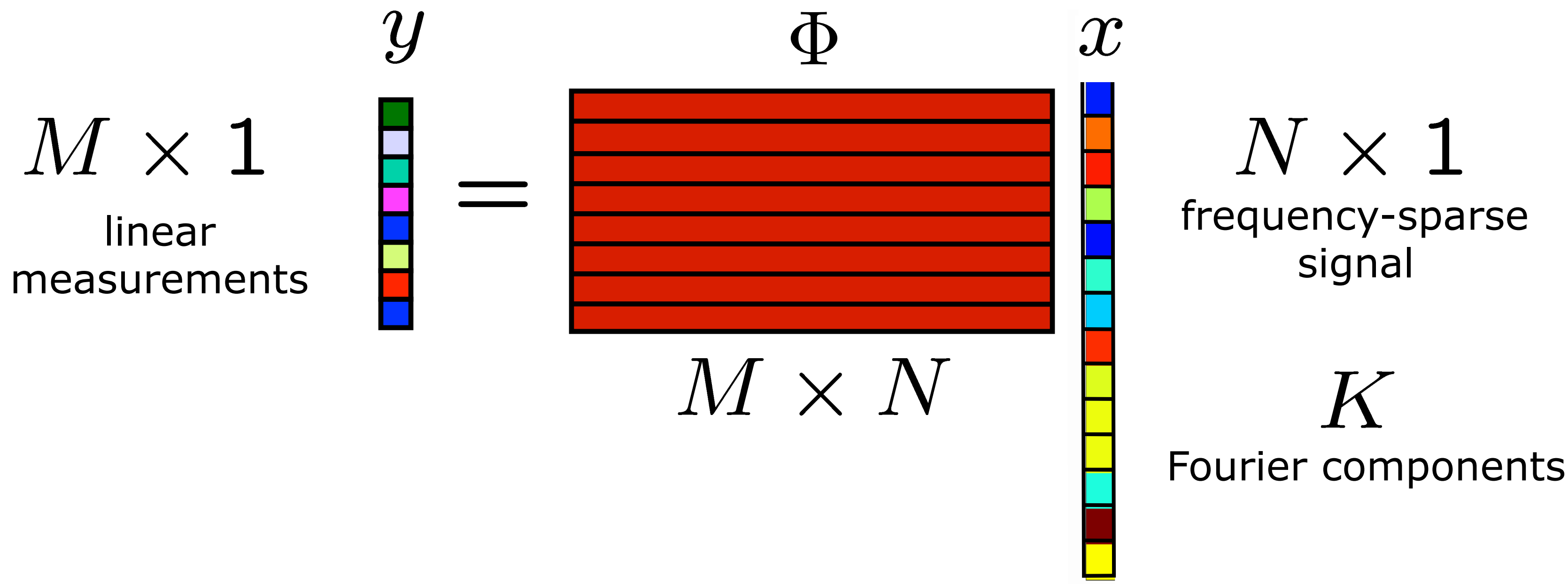
Richard G. Baraniuk
(Rice University)

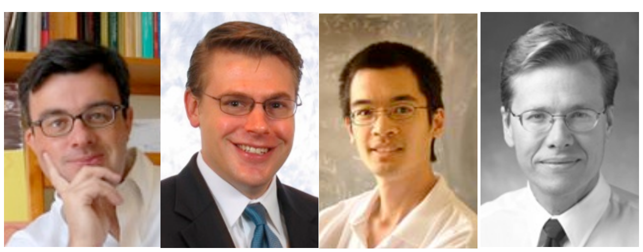
Hamid Dadkhahi
(UMass Amherst)

Karsten Fyhn
(Aalborg University)

Spectral Compressive Sensing

- Compressive sensing applied to ***frequency-sparse signals***

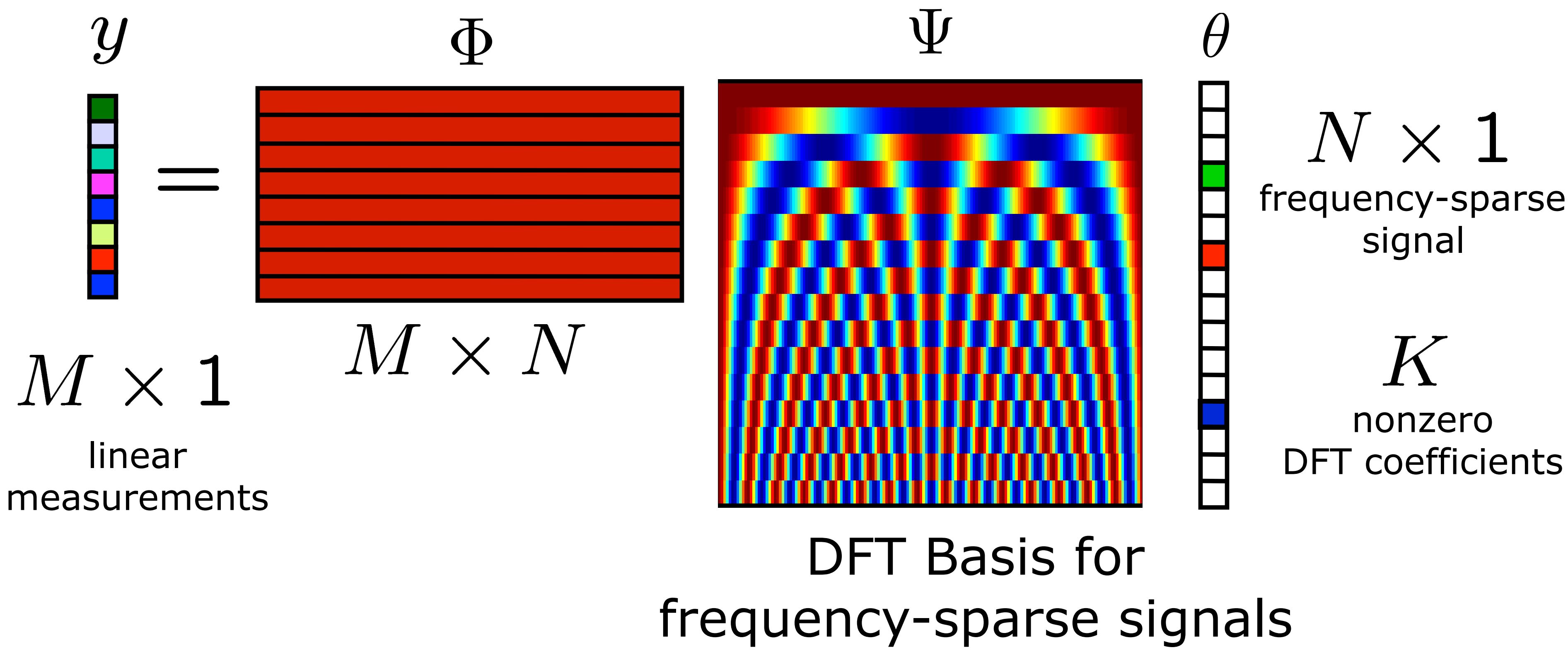




[E. Candès, J. Romberg, T. Tao; D. Donoho]

Spectral Compressive Sensing

- Compressive sensing applied to ***frequency-sparse signals***

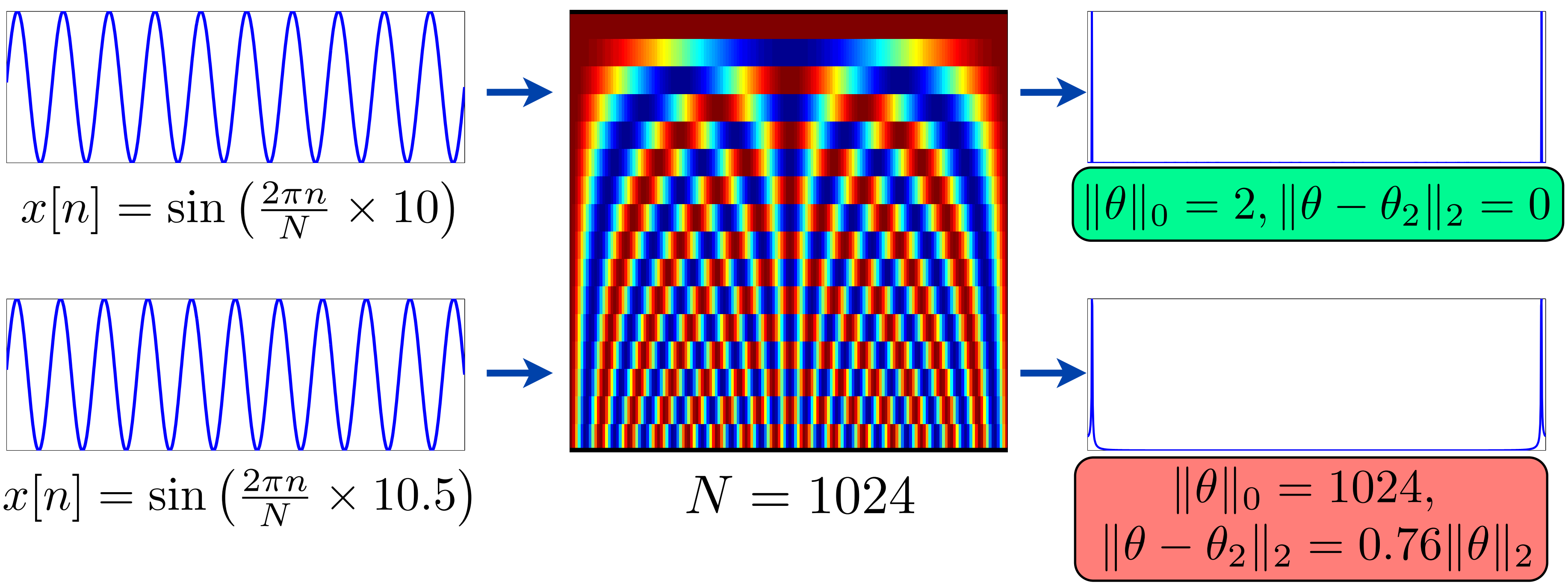


Frequency-Sparse Signals and the DFT Basis

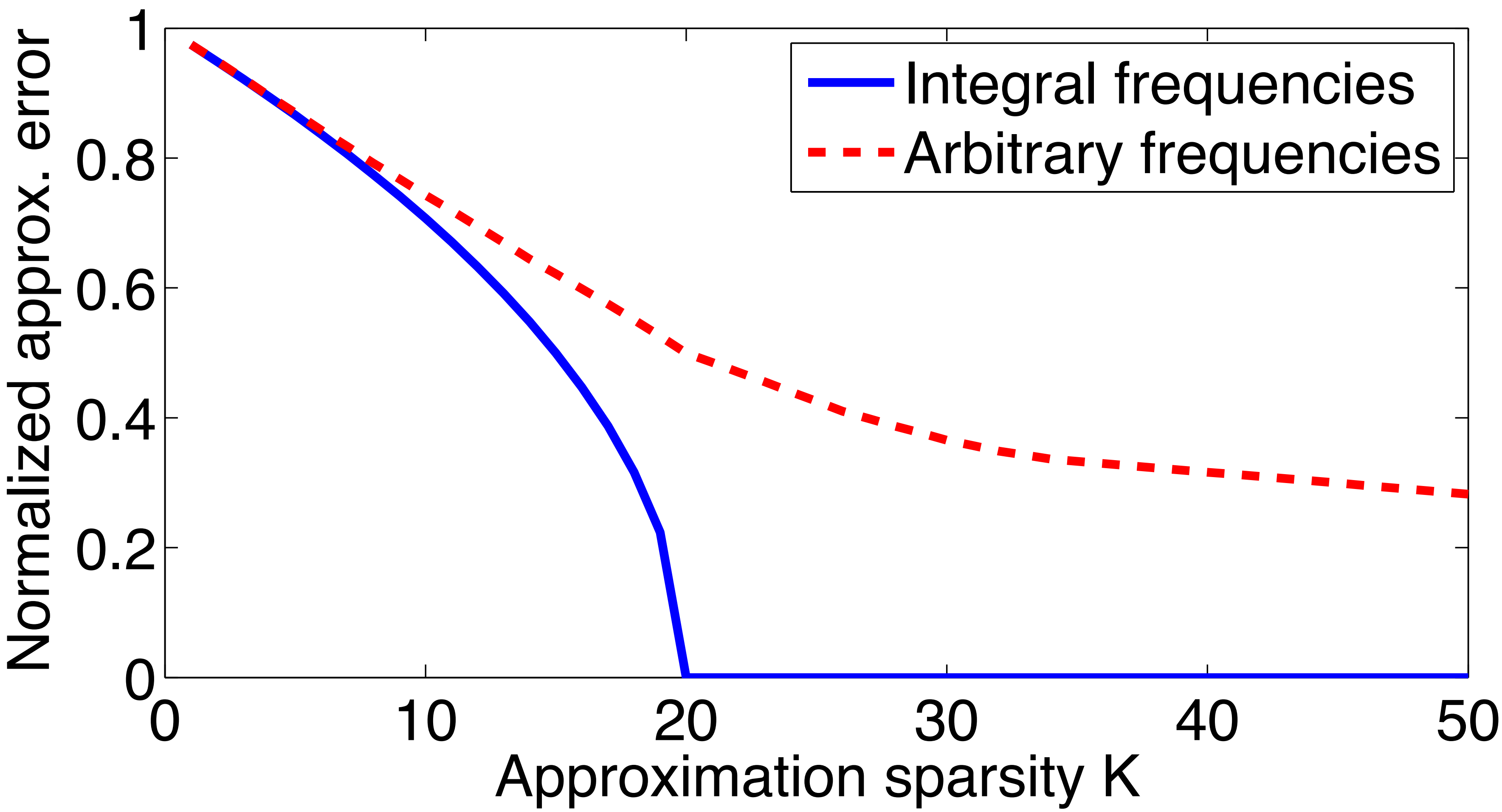
$$x = \sum_{k=1}^{K} a_k e(f_k) \qquad X(\omega) = \sum_{k=1}^{K} a_k \, \delta(\omega - \omega_k)$$

$$e(f) = \frac{1}{\sqrt{N}} \left[e^{j2\pi f/N} \;\; e^{j2\pi 2f/N} \;\; \ldots \;\; e^{j2\pi (N-1)f/N} \right]$$

$$\theta = \Psi^{-1} x$$

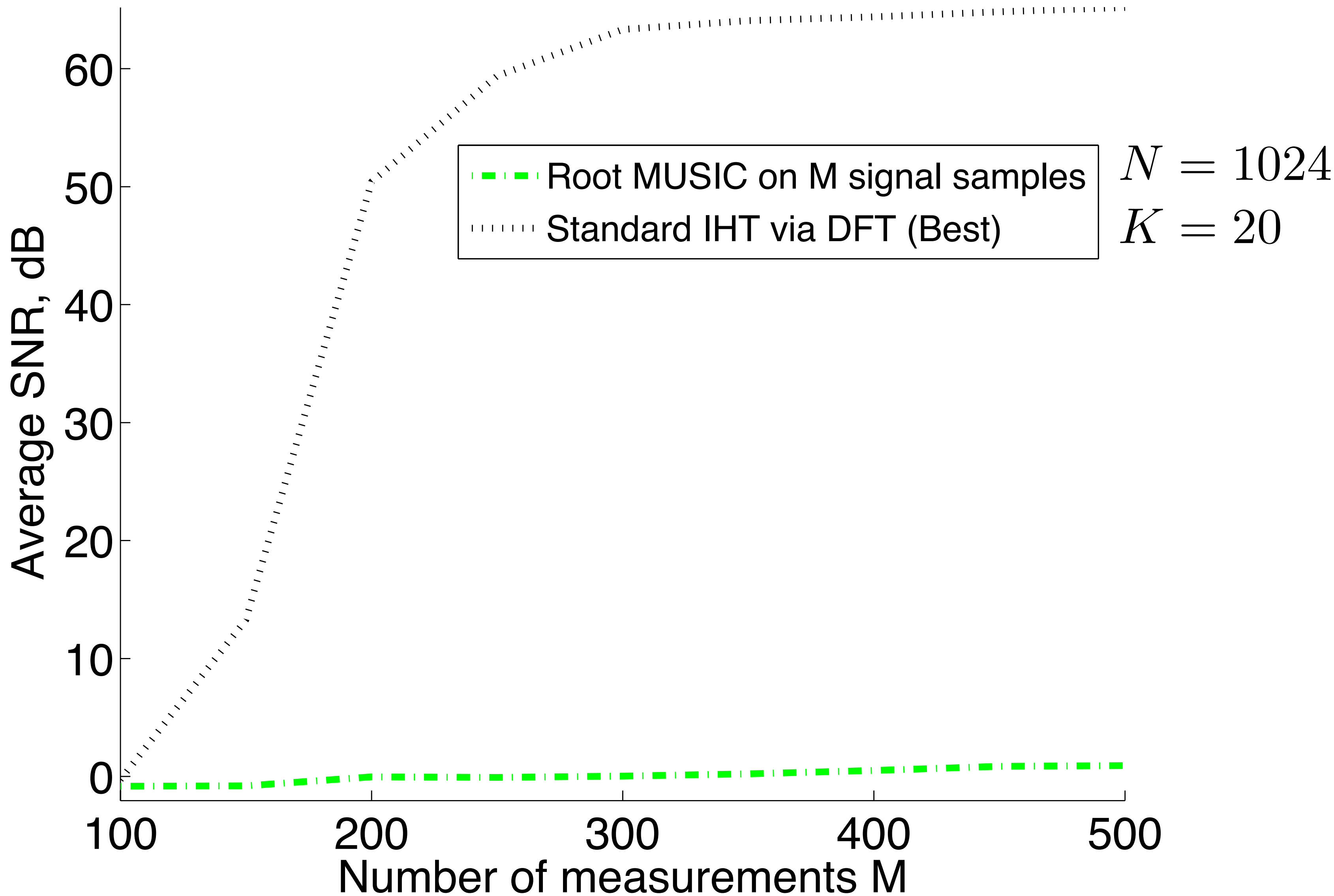


Frequency-Sparse Signals and the DFT Basis

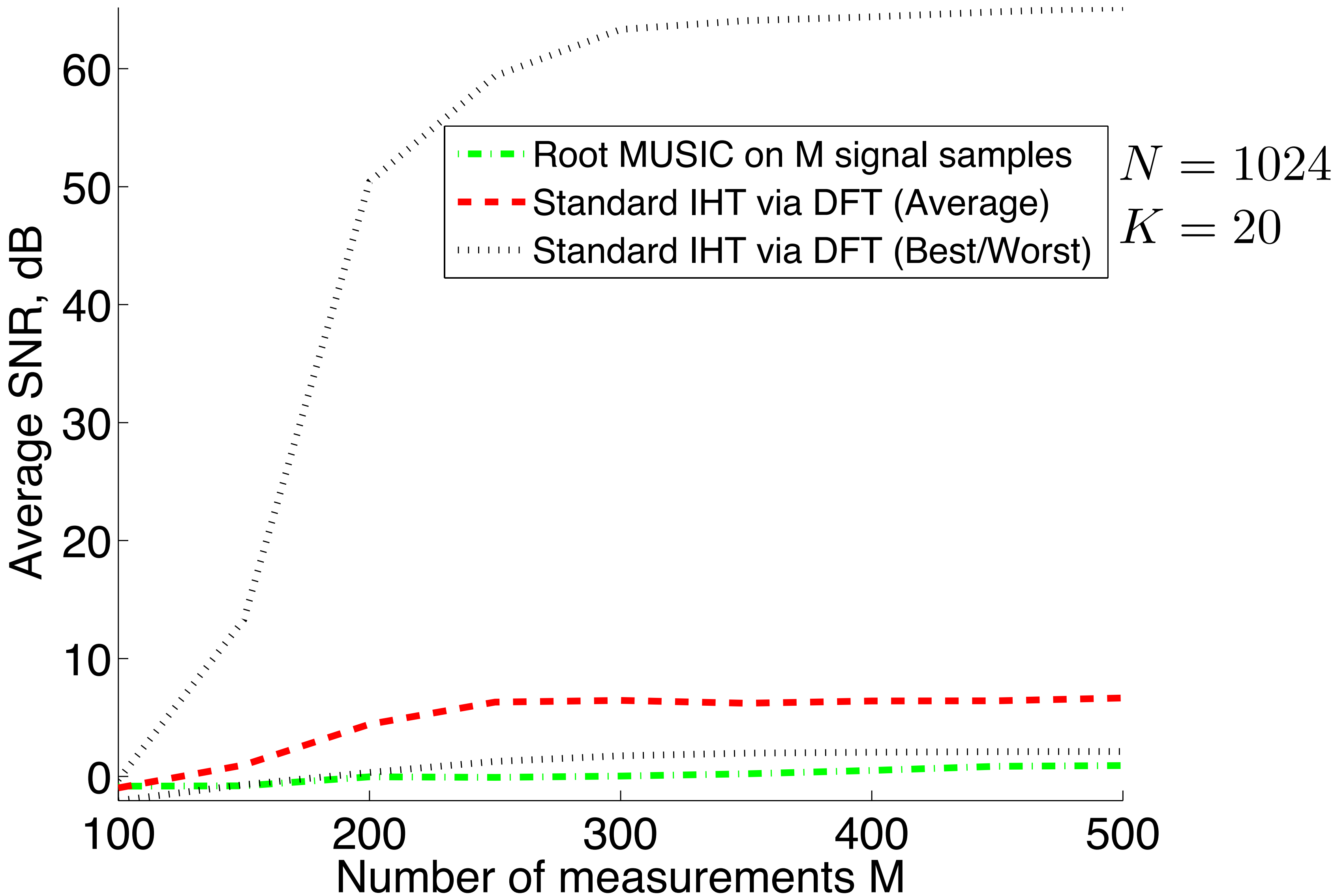


Signal is sum of 10 sinusoids

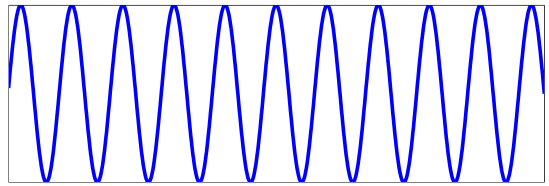
Compressive Sensing for Frequency-Sparse Signals



Compressive Sensing for Frequency-Sparse Signals

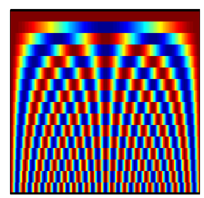


The Redundant DFT Frame

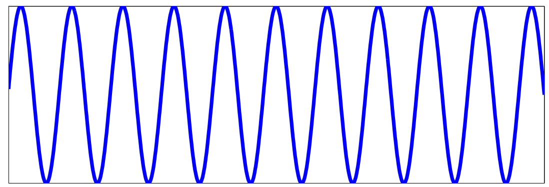


$$x[n] = \sin\left(\frac{2\pi n}{N} \times 10.5\right)$$

$$N = 1024$$



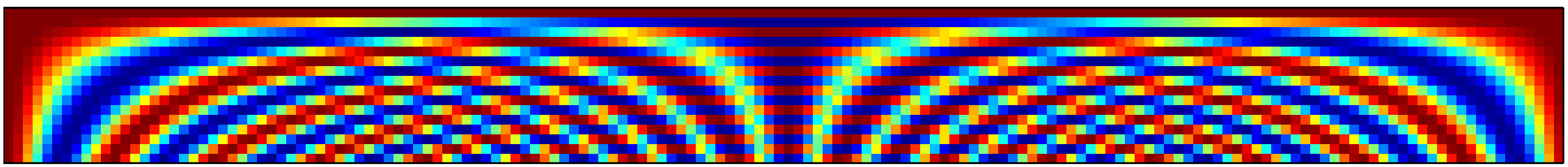
The Redundant DFT Frame



$$x[n] = \sin\left(\frac{2\pi n}{N} \times 10.5\right)$$

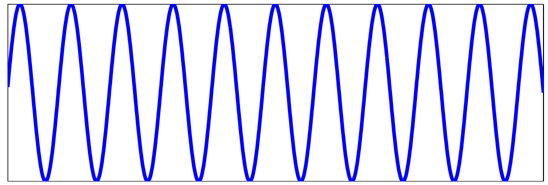
$N = 1024$

$\Psi(c),\ c = 10$



$$\Psi(c) = \left[e\left(\frac{1}{c}\right)\ e\left(\frac{2}{c}\right)\ \ldots\ e\left(\frac{N - 1/c}{c}\right) \right]$$

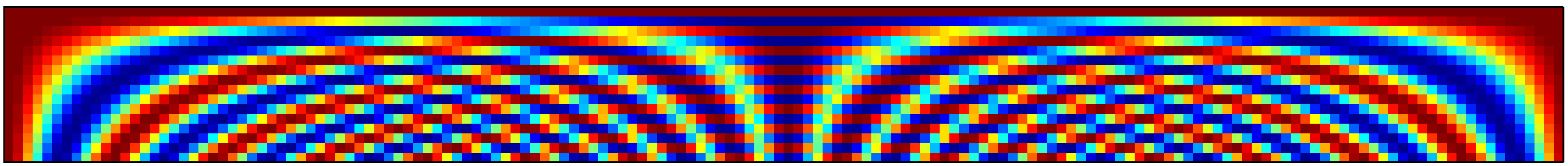
The Redundant DFT Frame



$$x[n] = \sin\left(\frac{2\pi n}{N} \times 10.5\right)$$

$$N = 1024$$

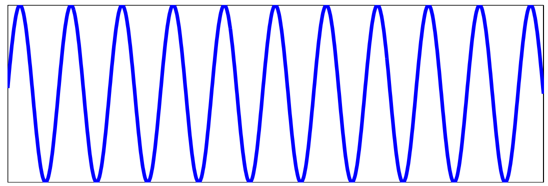
$$\Psi(c),\ c = 10$$



$$x = \Psi(c)\theta$$

$$\|\theta\|_0 = 2, \|\theta - \theta_2\|_2 = 0$$

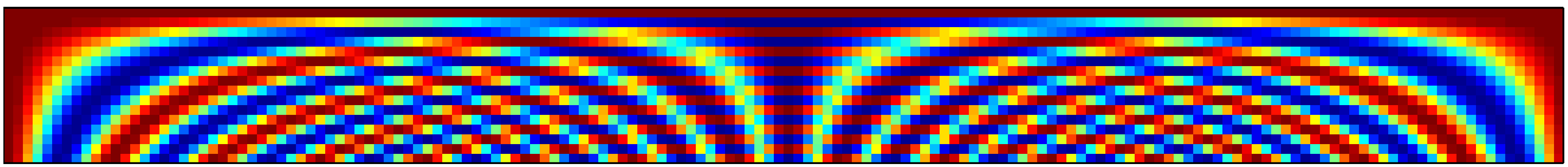
The Redundant DFT Frame



$x[n] = \sin\left(\frac{2\pi n}{N} \times 10.5\right)$

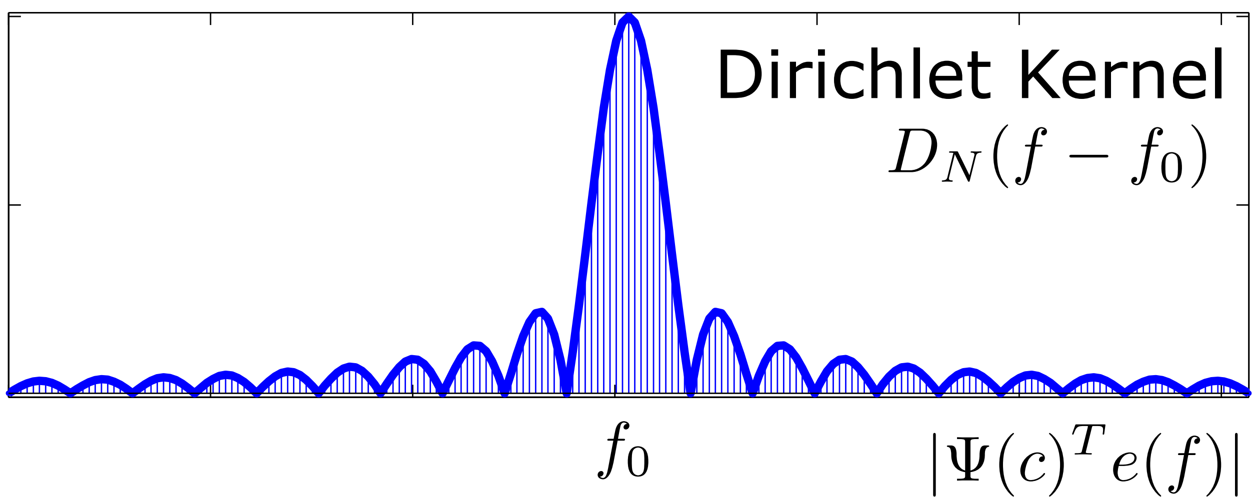
$N = 1024$

$\Psi(c),\ c = 10$

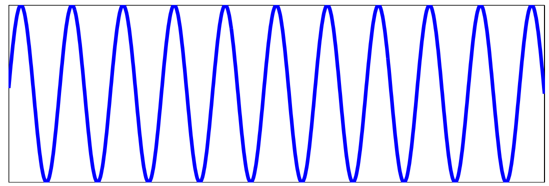


Recovery algorithms operate similarly to “matched filtering”:

$p = \Psi(c)^T x$



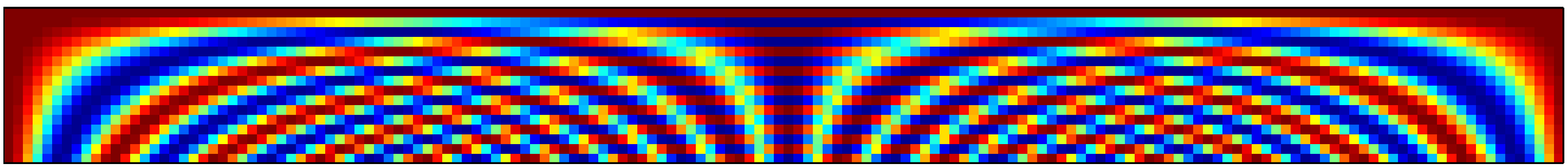
The Redundant DFT Frame



$$x[n] = \sin\left(\frac{2\pi n}{N} \times 10.5\right)$$

$N = 1024$

$\Psi(c),\ c = 10$



$$x = \Psi(c)\theta$$

$\|\theta\|_0 = 2, \|\theta - \theta_2\|_2 = 0$

$\mu(\Psi(c)) \approx 0.98$

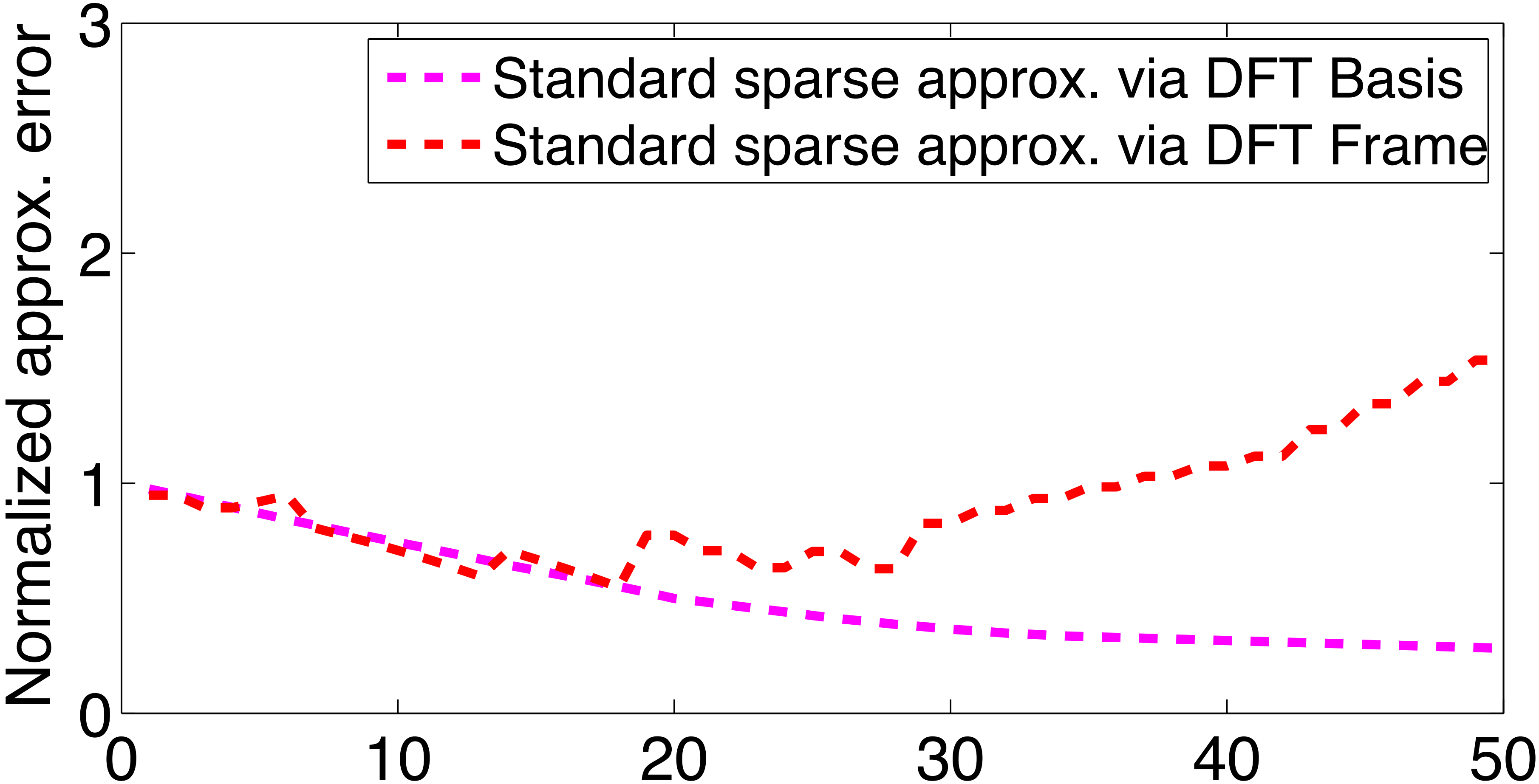
Sparse approximation algorithms fail

$$\theta' = \Psi(c)^T x$$

$$\begin{aligned} \|\theta'\|_0 &= 9218 = (c-1)N + 2, \\ \|\theta' - \theta'_2\|_2 &= 0.95\|\theta'\|_2 \end{aligned}$$

[Candès, Needell, Eldar, Randall 2011]

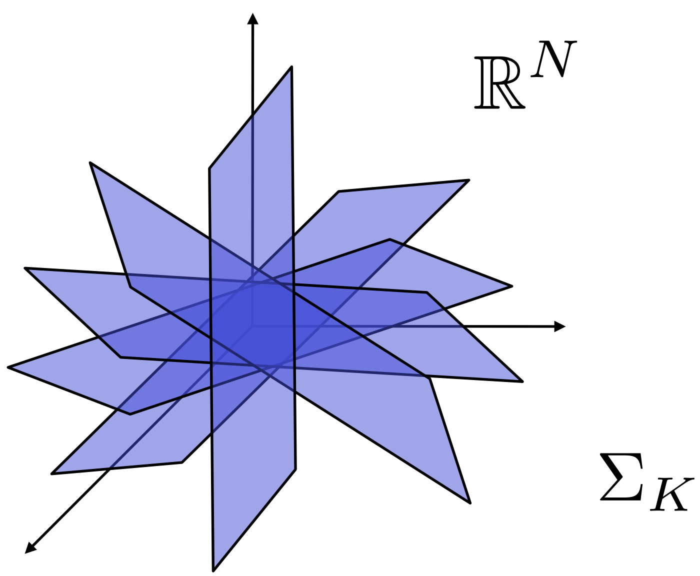
Sparse Approximation of Frequency-Sparse Signals



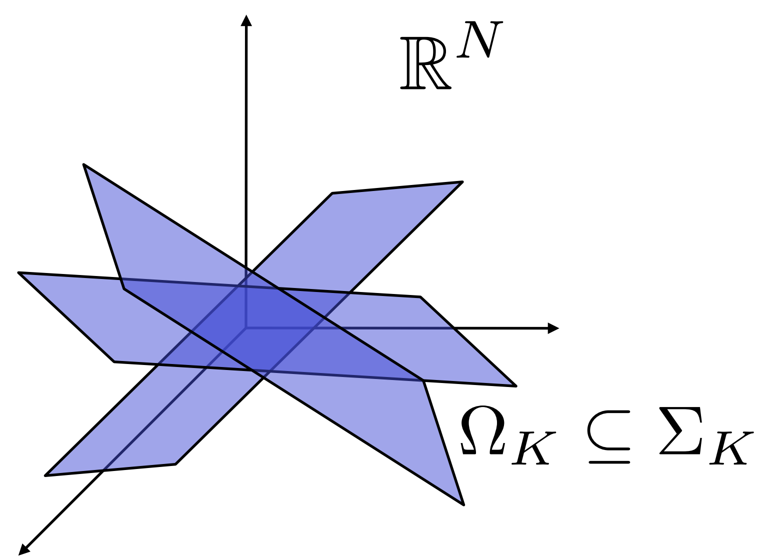
Signal is sum of 10 sinusoids at arbitrary frequencies

Structured Sparse Signals

- A K**-sparse** signal lives on the collection of K-dim subspaces aligned with coordinate axes



- A K**-structured sparse** signal lives on a particular (reduced) collection of K-dimensional canonical subspaces

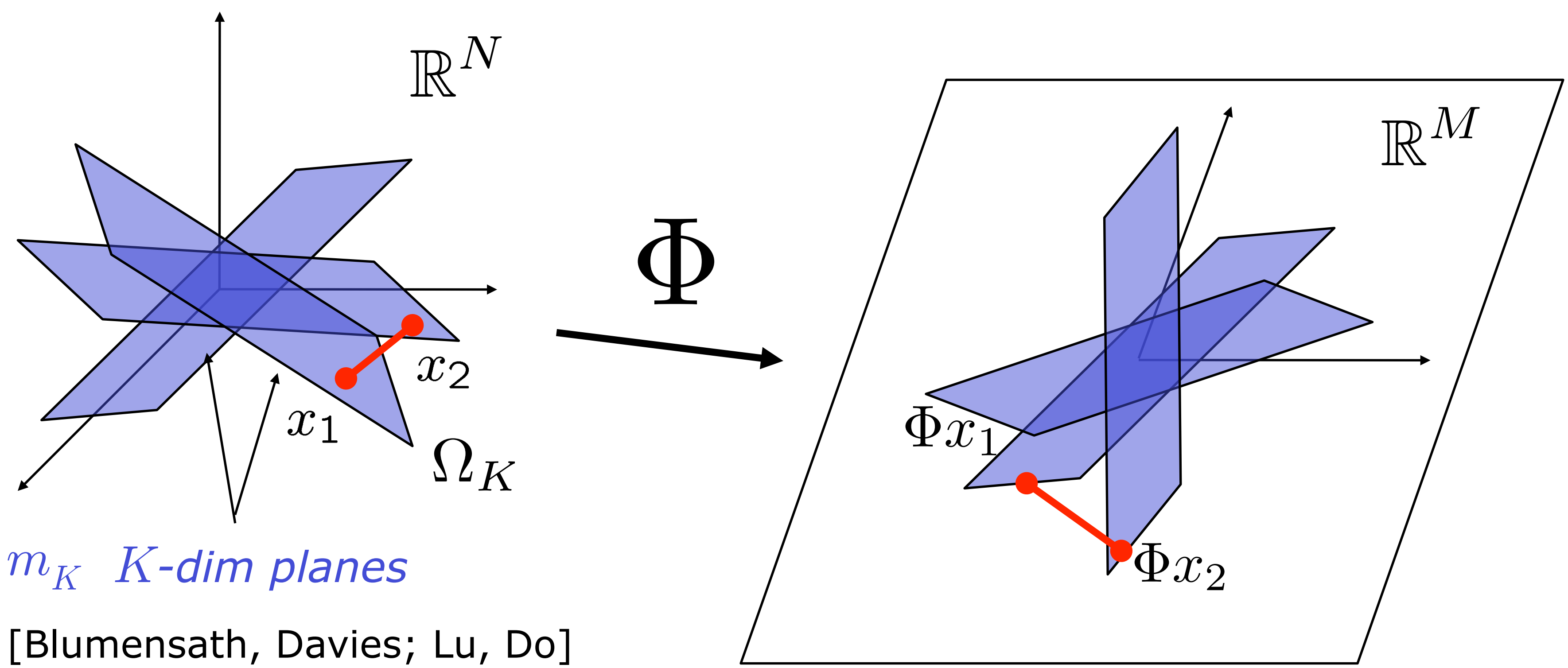


[Baraniuk, Cevher, Duarte, Hegde 2010]

Structured Restricted Isometry Property (SRIP)

- Preserve the structure **only** between sparse signals that follow the structure model
- Random (iid Gaussian, Rademacher) matrix has the SRIP with high probability if

$$M = \mathcal{O}(K + \log m_K)$$

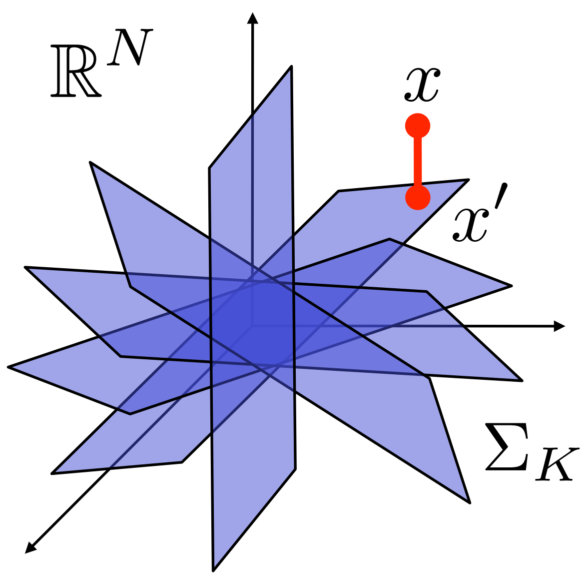


[Blumensath, Davies; Lu, Do]

Leveraging Structure in Recovery

Many state-of-the-art sparse recovery algorithms (greedy and optimization solvers) rely on **thresholding** $x' = \mathcal{T}(x, K)$ [Daubechies, Defrise, and DeMol; Nowak, Figueiredo, and Wright; Tropp and Needell; Blumensath and Davies...]

$$x'(n) = \begin{cases} x(n) & \text{if } |x(n)| \text{ is among } K \text{ largest,} \\ 0 & \text{otherwise.} \end{cases}$$

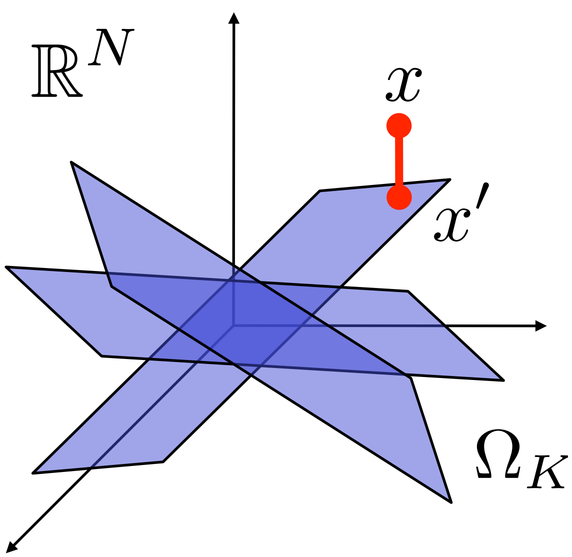


Thresholding provides the **best approximation** of x within Σ_K

$$x' = \arg\min_{\overline{x} \in \Sigma_K} \|x - \overline{x}\|_2$$

Structured Recovery Algorithms

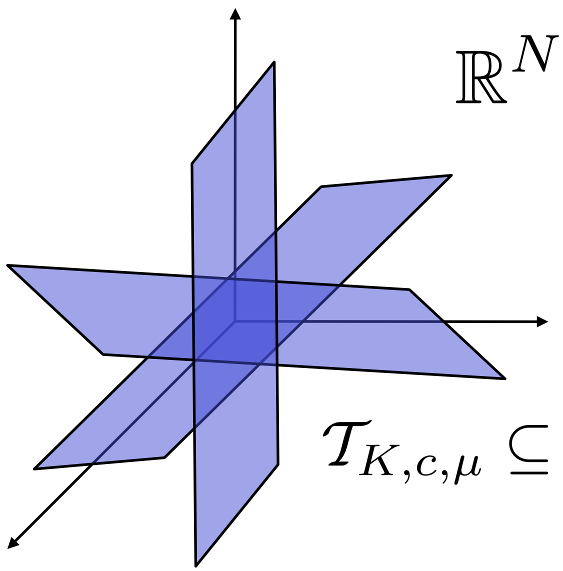
- Modify existing approaches (optimization or greedy-based) to obtain **structure-aware recovery algorithms**: replace the thresholding step in IHT, CoSaMP, SP, ... with a **best structured sparse approximation step** that finds the closest point within union of subspaces



$$x' = \mathbb{M}(x, K) = \arg\min_{\bar{x} \in \Omega_K} \|x - \bar{x}\|_2$$

Greedy structure-aware recovery algorithms **inherit guarantees** of generic counterparts *(even though feasible set may be nonconvex)*

Structured Frequency-Sparse Signals

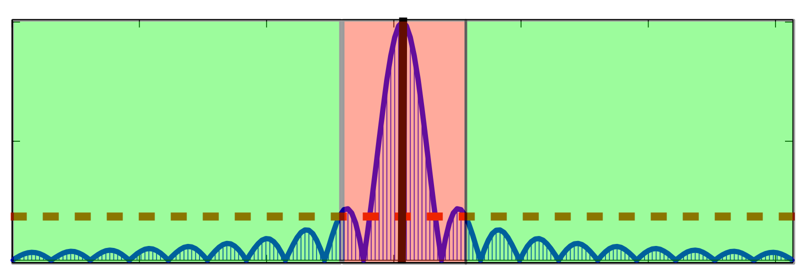


- A K**-structured frequency-sparse** signal x consists of K sinusoids that are mutually incoherent:

$$x = \sum_{k=1}^{K} a_k e(f_k) \in \mathcal{T}_{K,c,\mu} \quad \text{if}$$

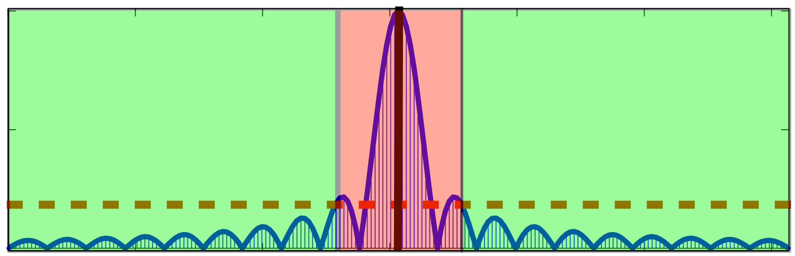
$$cf_K \in \mathbb{Z}, \ |\langle e(f_k), e(f_{k'})\rangle| \leq \mu \ \forall \ k \neq k'$$

- If x is K-structured frequency-sparse, then there exists a K-sparse vector θ such that $x = \Psi(c)\theta$ and the nonzeros in θ are spaced apart from each other (***band exclusion***).



Structured Frequency-Sparse Signals

- If x is K-structured frequency-sparse, then there exists a K-sparse vector θ such that $x = \Psi(c)\theta$ and the nonzeros in θ are spaced apart from each other.

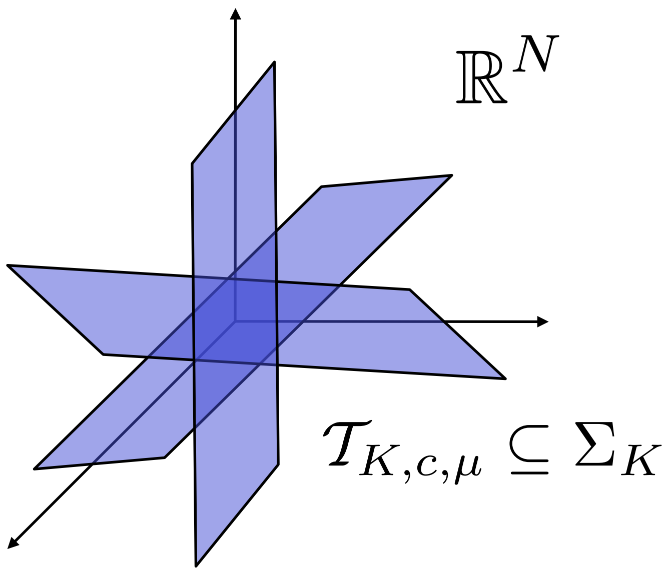


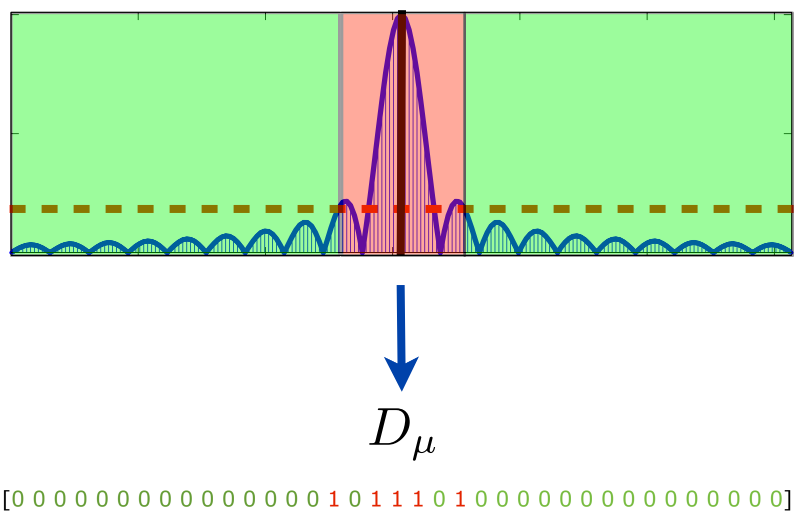
- Preserve the structure only between sparse signals that follow the structured sparsity model

- Random (iid Gaussian, Bernoulli) matrix has the structured RIP with high probability if

$$M = \mathcal{O}\left(K \log\left(\frac{c(N - K\, D_N^{-1}(\mu N))}{K}\right)\right)$$

Structured Sparse Approximation





Algorithm 1: $\mathbb{T}(x, K, c, \mu)$

Integer Program

Inputs:

- Signal vector x
- Target sparsity K
- Redundancy factor c
- Maximum coherence μ

Output:

- Approximation vector $\widehat{x}$

- Compute coefficients: $\theta = \Phi(c)^T x$
  $w_\theta[i] = \theta[i]^2,\ i = 0, \ldots, cN - 1$
- Solve support:
  $$s = \arg \max_{s \in \{0,1\}^{cN}} w_\theta^T s$$
  $$\text{s.t. } D_\mu s \leq \mathbf{1}, s^T \mathbf{1} \leq K$$
- Mask coefficients:
  $\widehat{\theta}[i] \leftarrow \theta[i] s[i], i = 0, \ldots, cN - 1$
- Return $\widehat{x} = \Phi(c)\widehat{\theta}$

Recovery with Structured Sparsity

Theorem:
Assume we obtain noisy CS measurements of a signal $y = \Phi x + n$. If Φ has the structured RIP with $\delta < 0.1$, then the output of the structured IHT algorithm obeys

$$\underbrace{\|x - \widehat{x}\|_2}_{\text{CS recovery error}} \leq C_1 \|x - \mathbb{M}(x,K)\|_2 + \frac{C_2}{\sqrt{K}} \|x - \mathbb{M}(x,K)\|_1 + C_3 \underbrace{\|n\|_2}_{\text{noise}}$$

CS recovery error

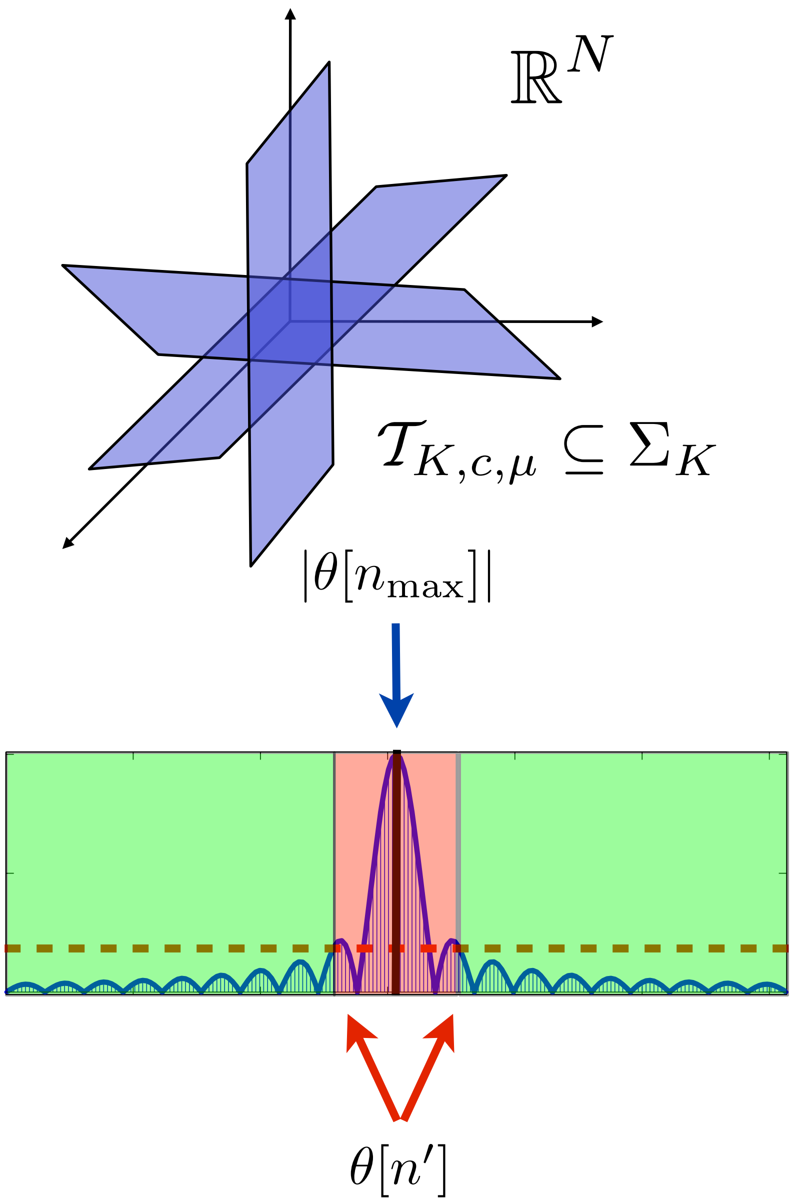
signal *K*-term **structured sparse approximation** error

noise

In words, *instance optimality* based on *structured sparse approximation*

[Baraniuk, Cevher, Duarte, Hegde 2010]

Structured Sparse Approximation



Algorithm 2: $\mathbb{T}_h(x, K, c, \mu)$

Inhibition Heuristic

Inputs:

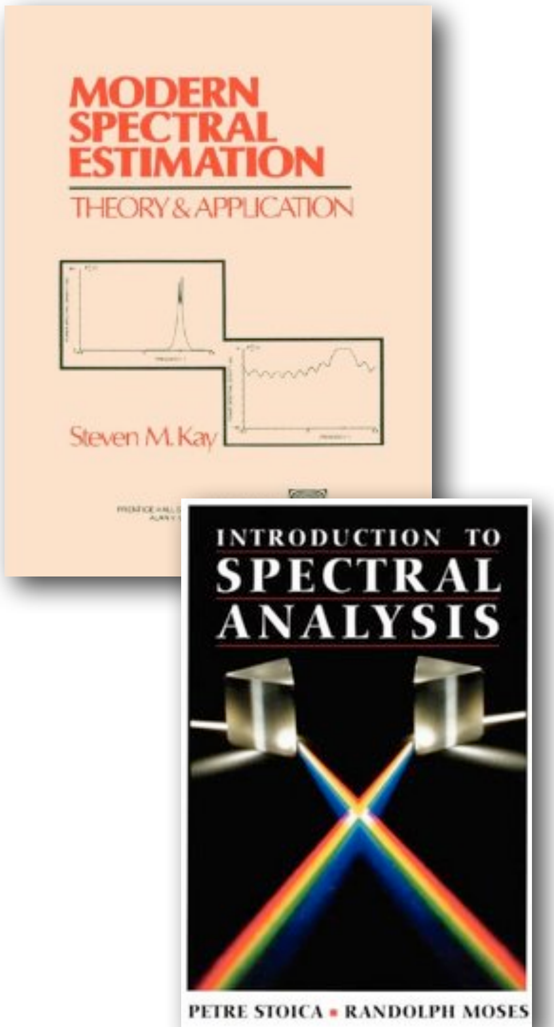
- Signal vector x
- Target sparsity K
- Redundancy factor c
- Maximum coherence μ

Output:

- Approximation vector $\widehat{x}$

- Compute coefficients: $\theta = \Phi(c)^T x$
- Initialize: $\widehat{\theta}[d] = 0, \;\; d = 0, \ldots, cN-1$
- While θ is nonzero and $\|\widehat{\theta}\|_0 \leq K$,
  - Find max abs entry $|\theta[n_{\max}]|$ of θ
  - Copy entry $\widehat{\theta}[n_{\max}] = \theta[n_{\max}]$
  - Inhibit "coherent" entries $\theta[n'] = 0$
- Return $\widehat{x} = \Phi(c)\widehat{\theta}$

Structured Sparse Approximation



$$\bar{\theta} = \mathcal{T}(\Psi(c)^T x, K)$$

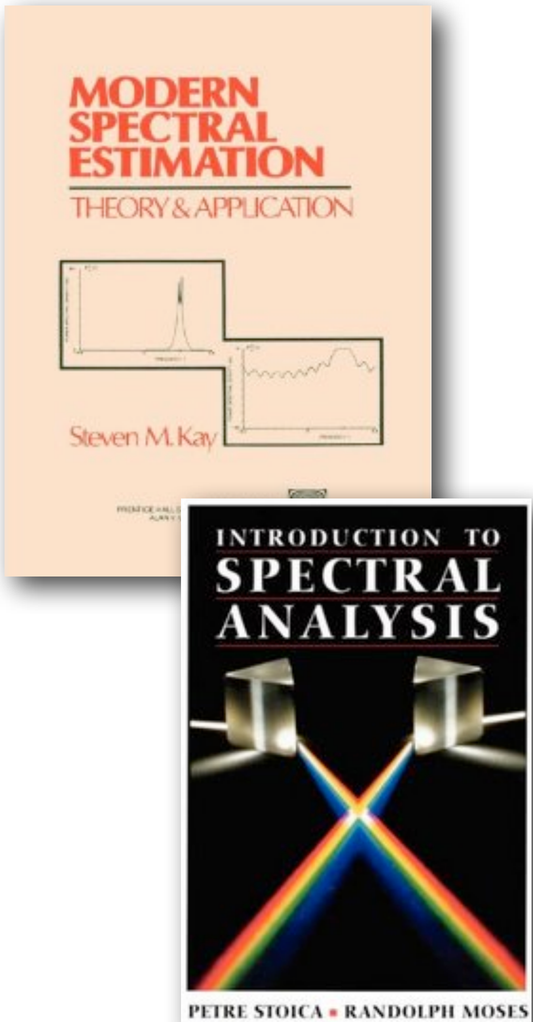
DFT Frame + Thresholding
equivalent to
Maximum Likelihood Estimate
of amplitudes and frequencies
for frequency-sparse signal
via Periodogram

$$x = \sum_{k=1}^{K} a_k e(f_k) + n$$

frequencies
amplitudes

Widely-studied problem:
Line spectral estimation

Structured Sparse Approximation



Algorithm 3: $\mathbb{T}_l(x, K)$

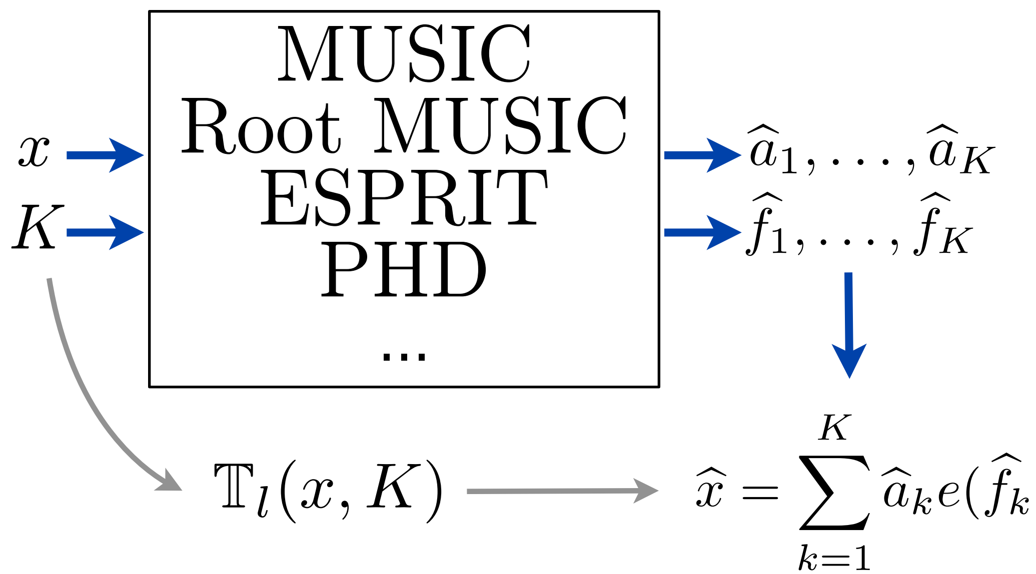
Line Spectral Estimation

Inputs:

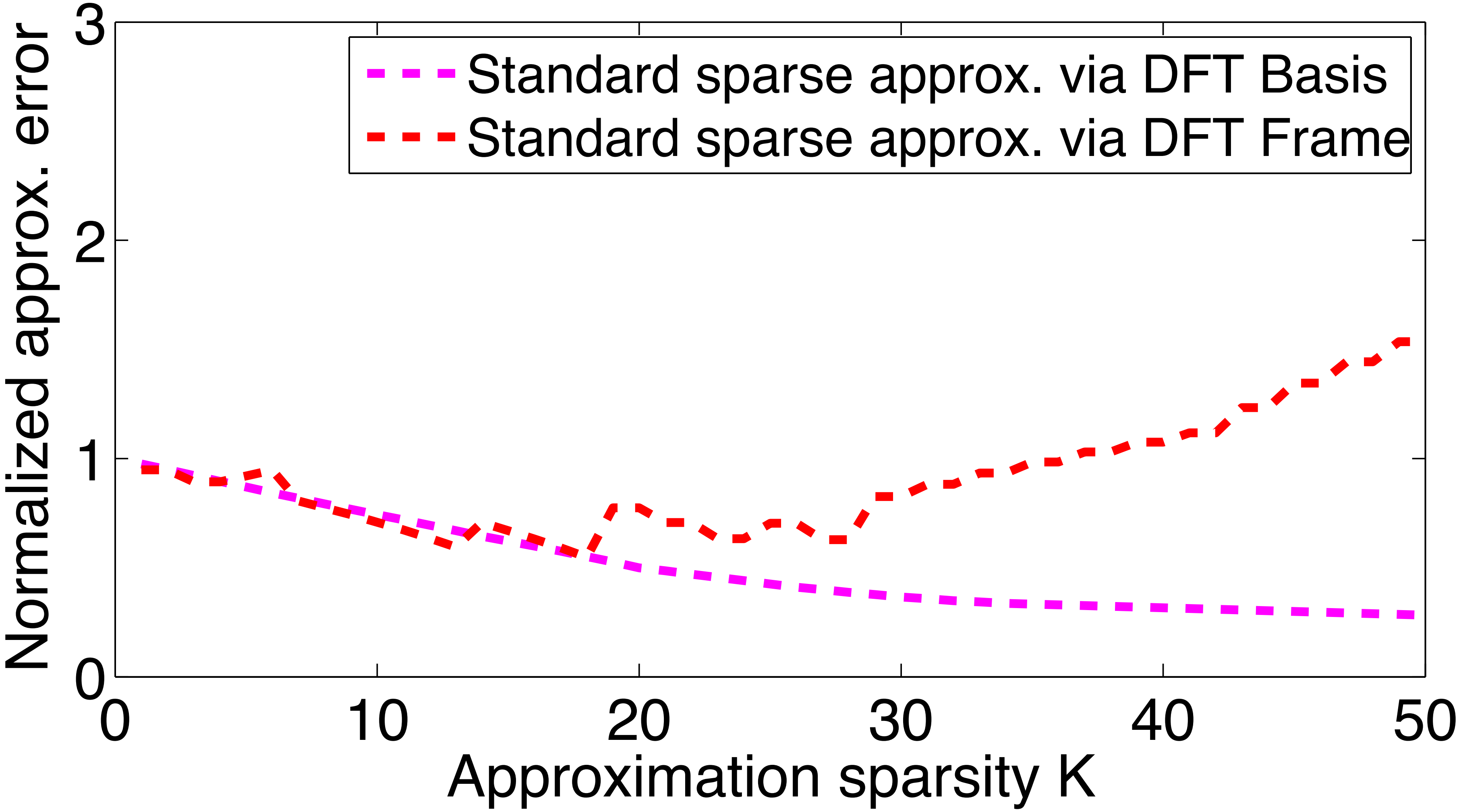
- Signal vector x
- Target sparsity K

Output:

- Parameter estimates $\widehat{a}_1, \ldots, \widehat{a}_K$ $\widehat{f}_1, \ldots, \widehat{f}_K$
- Signal estimate $\widehat{x}$

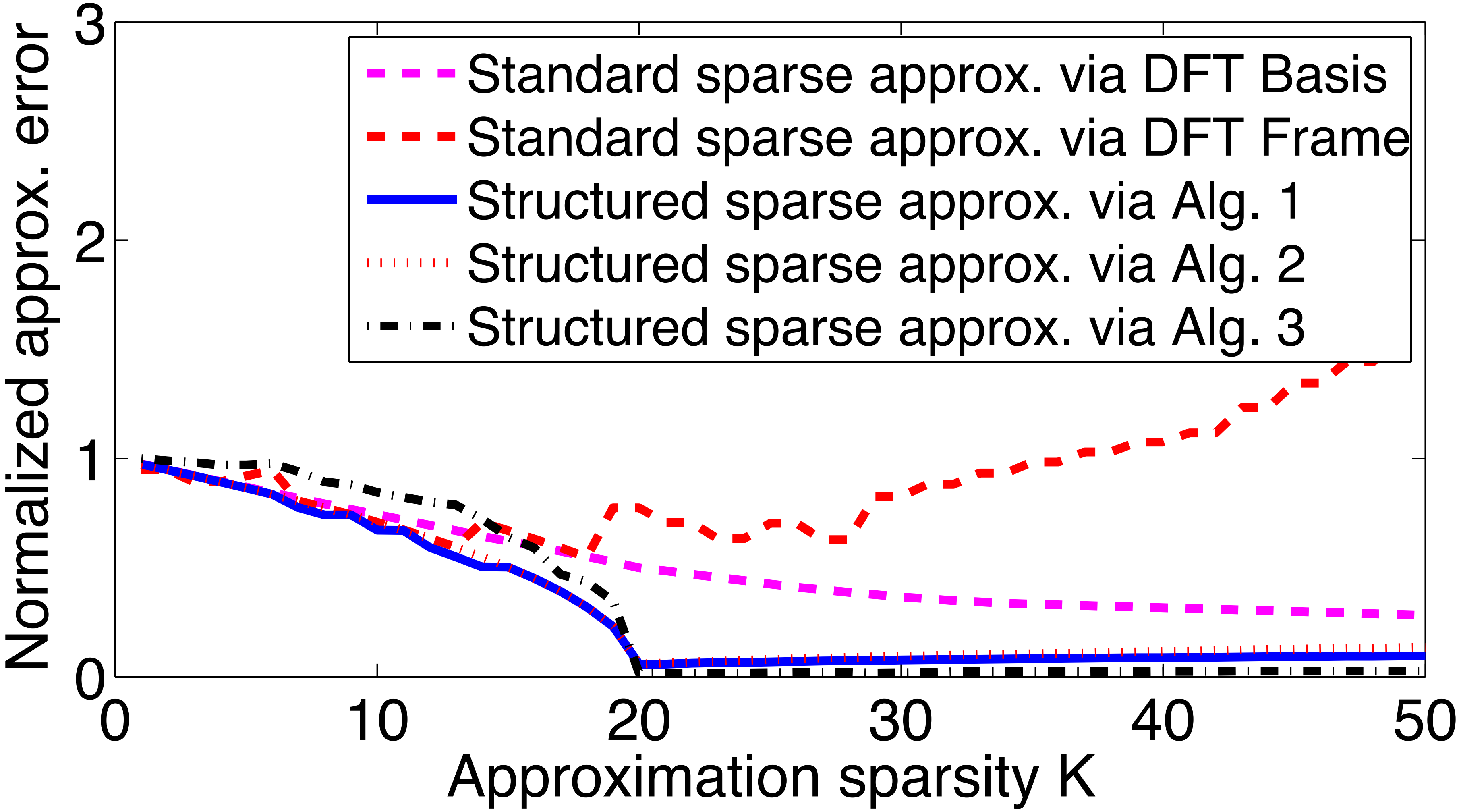


Sparse Approximation of Frequency-Sparse Signals



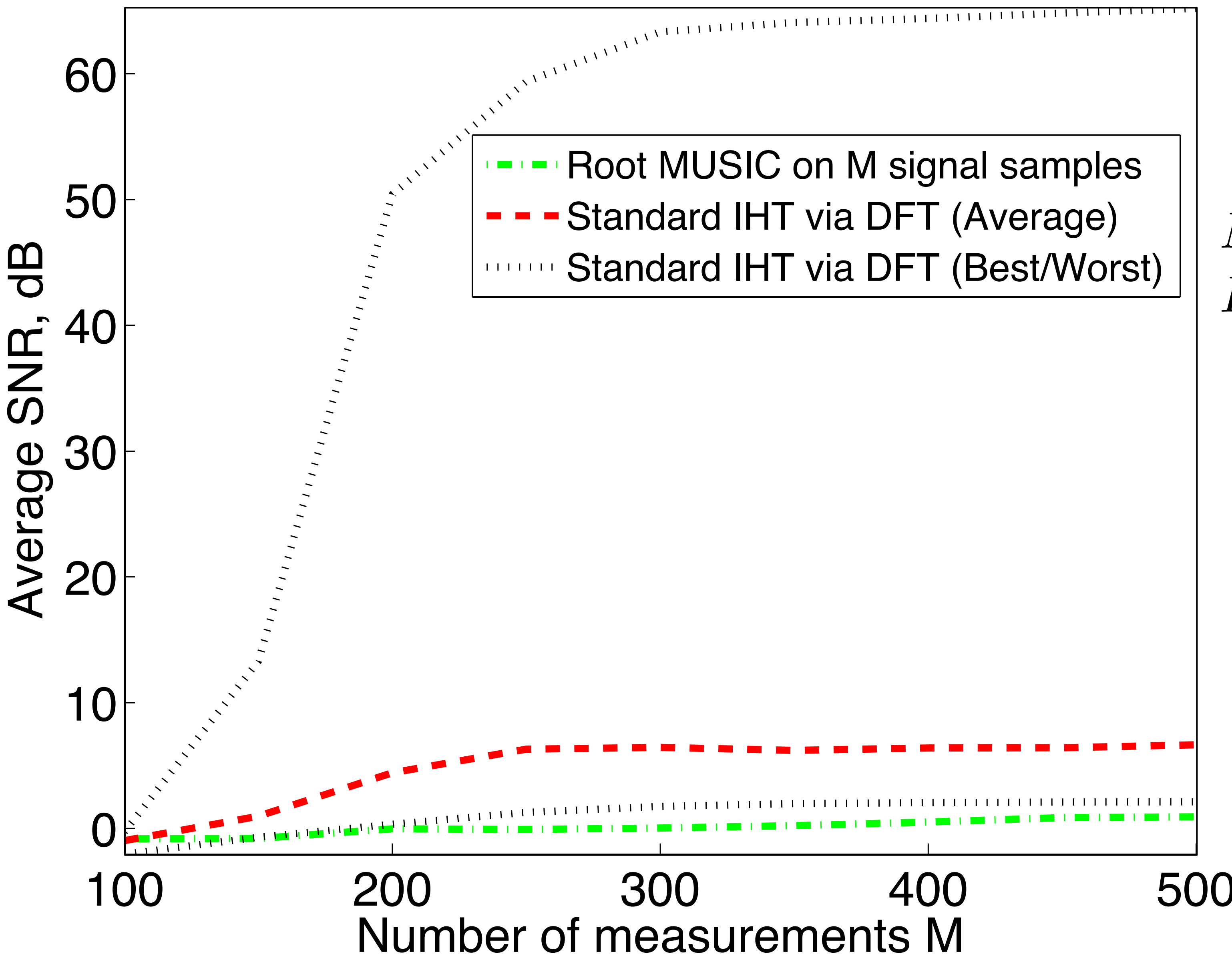
Signal is sum of 10 sinusoids at arbitrary frequencies

Sparse Approximation of Frequency-Sparse Signals



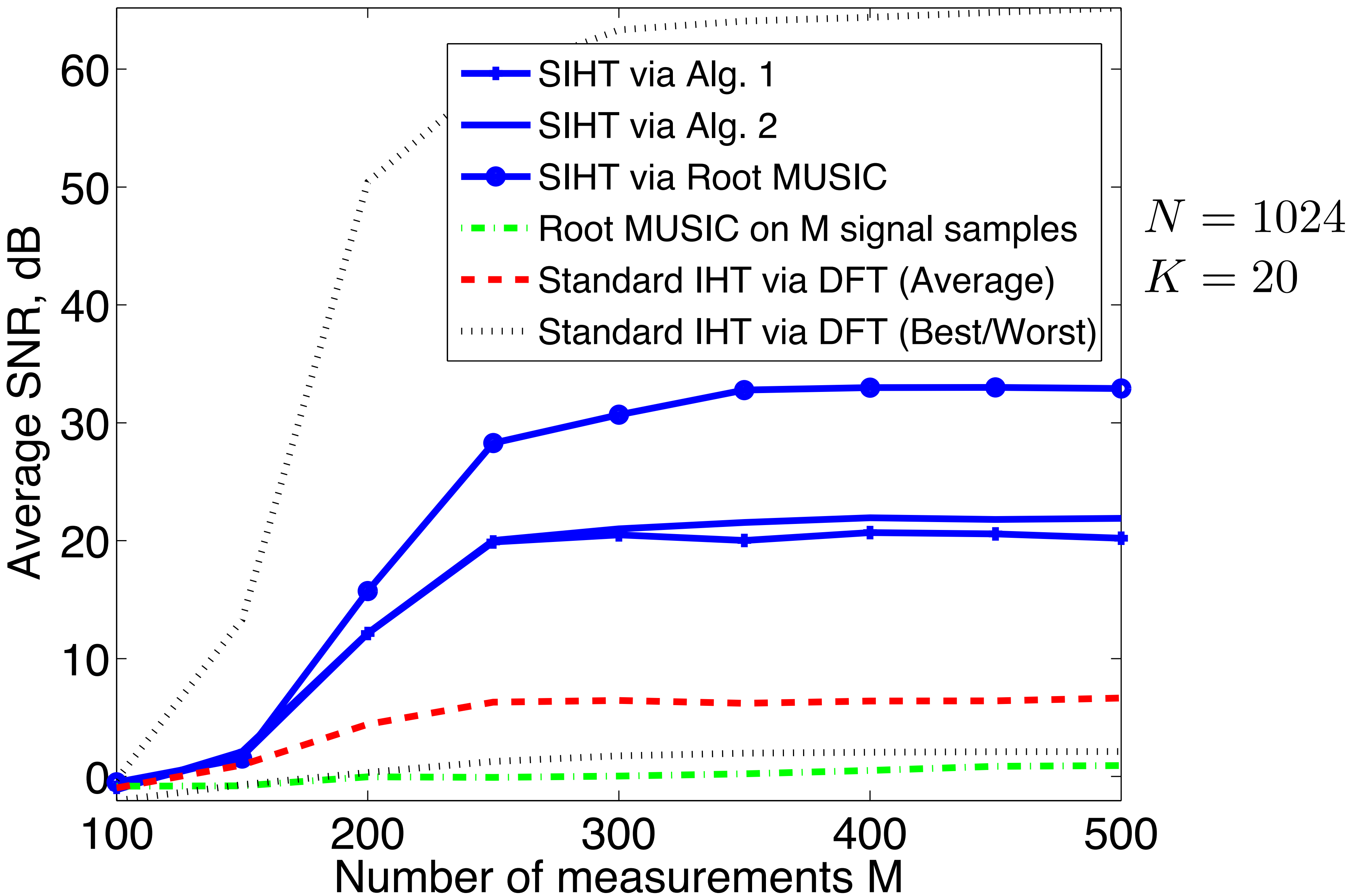
Signal is sum of 10 sinusoids at arbitrary frequencies

Structured CS: Performance



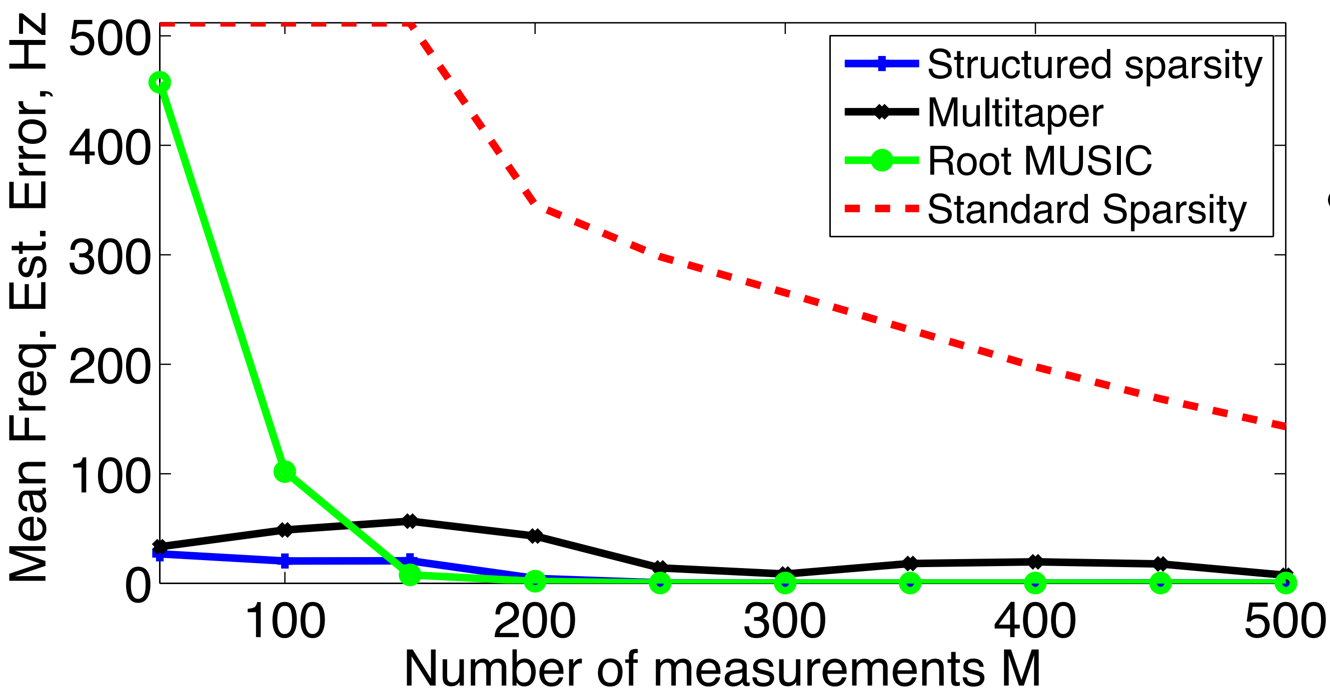
$N = 1024$
$K = 20$

Structured CS: Performance



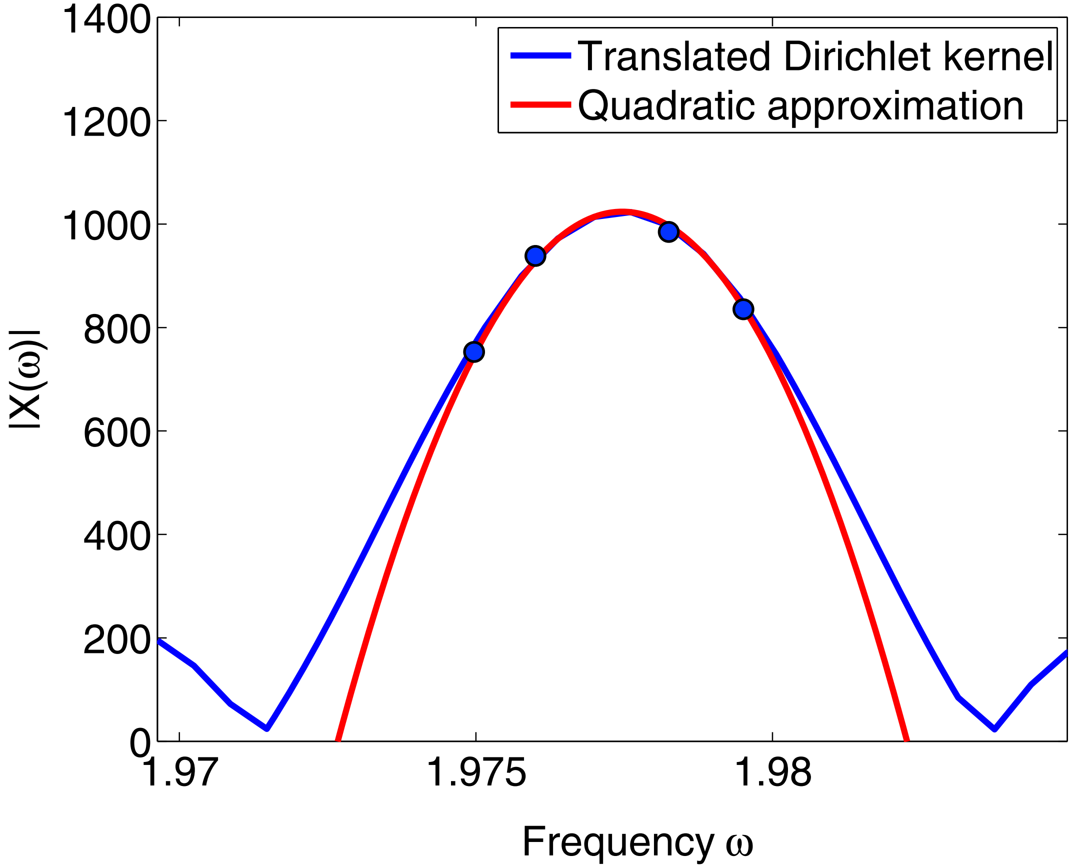
From Recovery of Sparse Signals To Line Spectral Estimation

- Can "read" indices of nonzero DFTF coefficients to obtain ***frequencies*** of frequency-sparse signal components
- Equivalence: accurate recovery = accurate estimation?
- ***Algorithms***: Alg. 3 essentially combines legacy line spectral estimation with CS recovery algorithms



- How to change ***signal model*** to further improve performance?

Interpolating the Projections (Dirichlet Kernel)



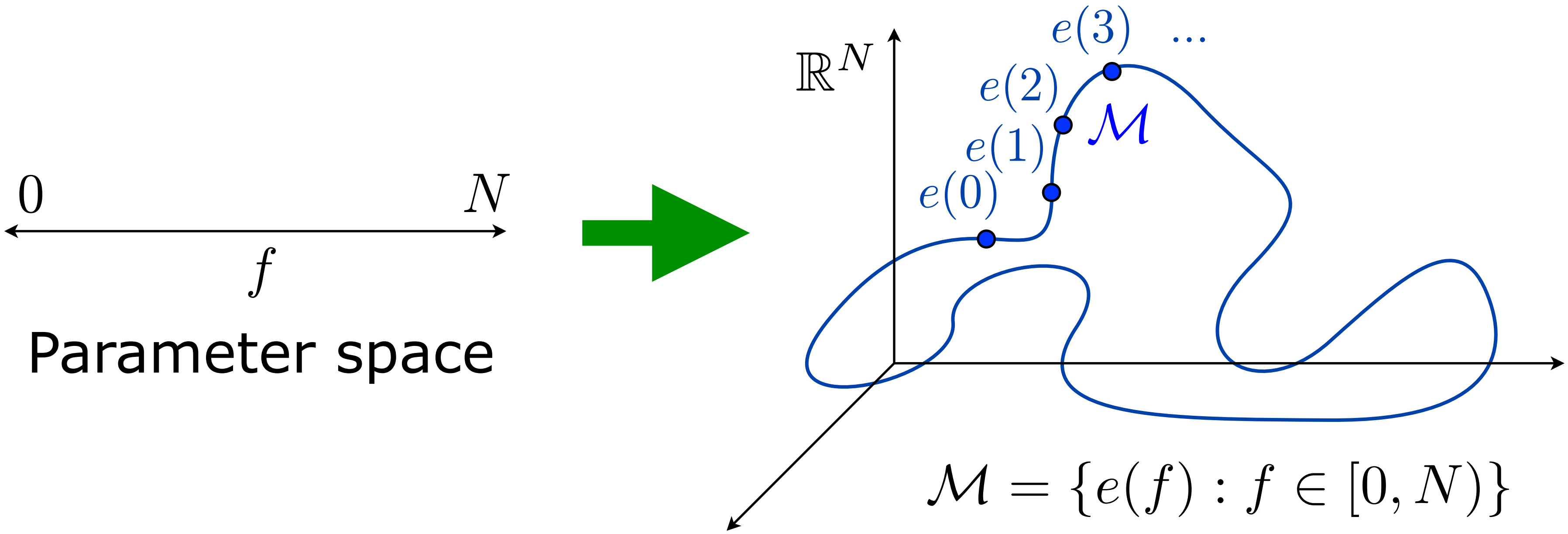
- Main lobe of Dirichlet kernel can be well approximated by a **quadratic polynomial** (parabola)
- **Three samples** around peak are required for interpolation

From Discrete to Continuous Models

- Both the DFT basis and the DFT frame can be conceived as **samplings** from an **infinite set** of signals $e(f)$ for a discrete set of values for the frequency $f \in [0, N)$

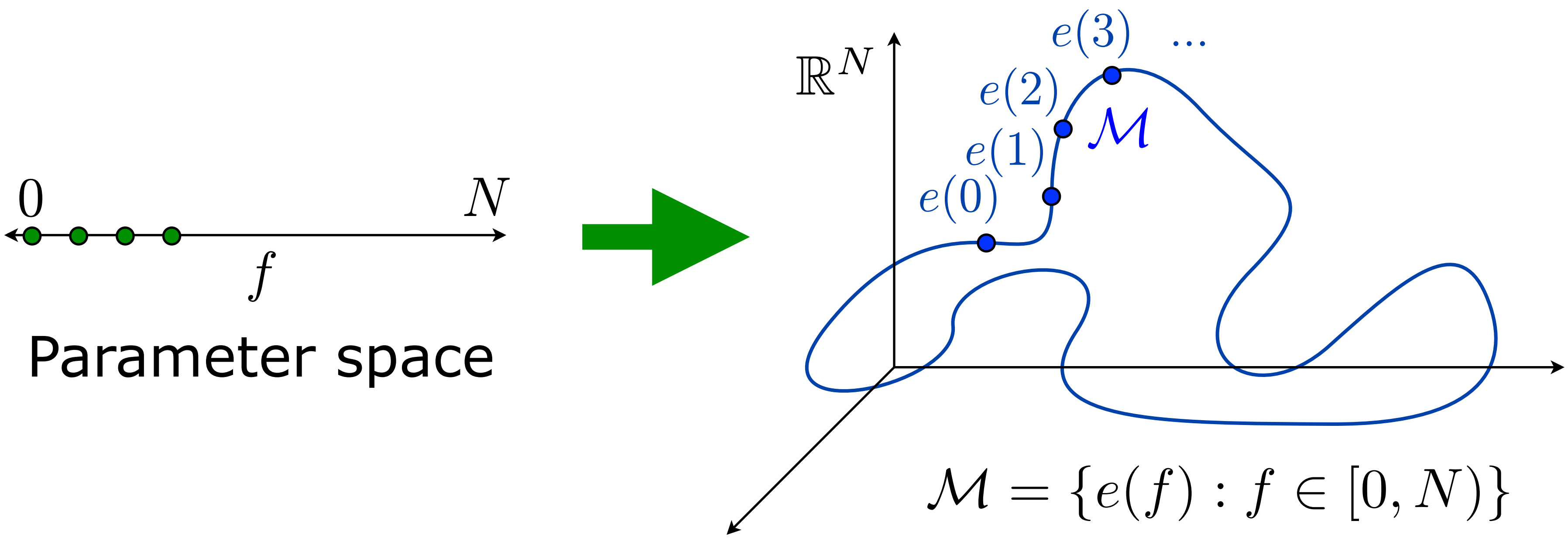
$$e(f) = \frac{1}{\sqrt{N}} \left[e^{j2\pi f/N} \ e^{j2\pi 2f/N} \ \ldots \ e^{j2\pi(N-1)f/N} \right]$$

- Since the signal vector $e(f)$ varies smoothly in each entry as a function of f, we can represent the signal set as a one-dimensional **nonlinear manifold:**

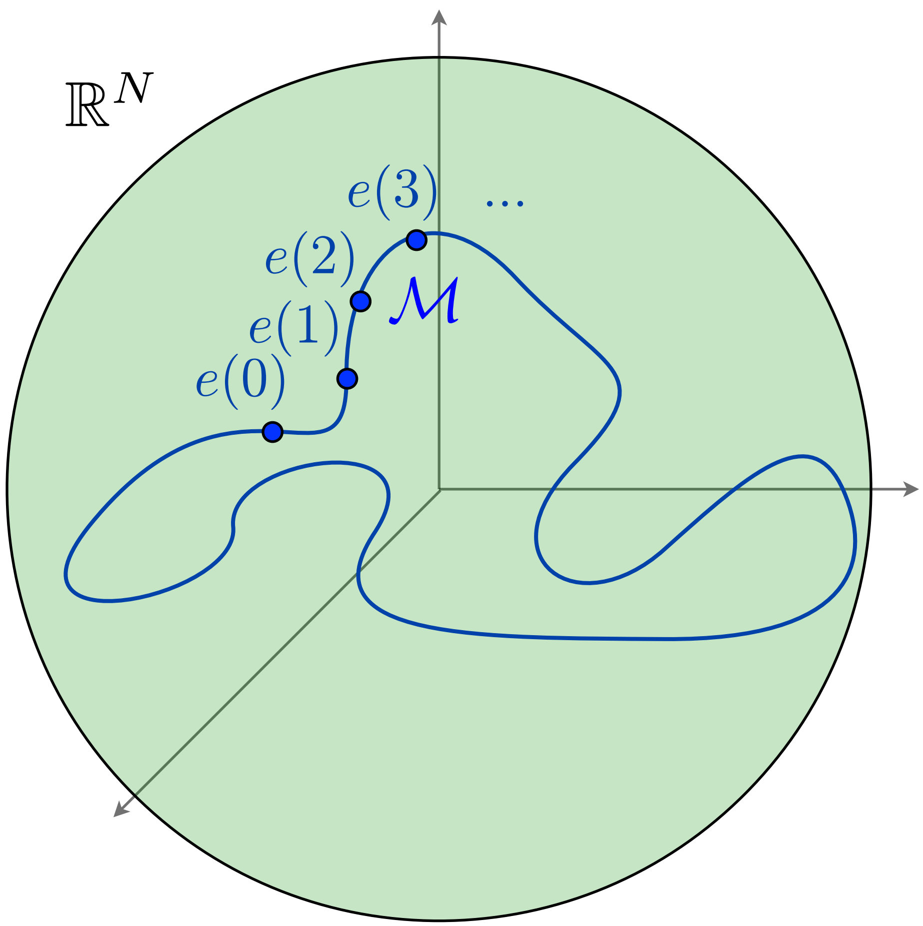


From Discrete to Continuous Models

- For computational reasons, we wish to design methods that allow us to ***interpolate*** the manifold from the samples obtained in the DFT basis/frame to increase the resolution of the frequency estimates.
- An ***interpolation-based*** compressive line spectral estimation algorithm obtains projection values for sets of manifold samples and interpolates around peak on the rest of the manifold to get frequency estimate

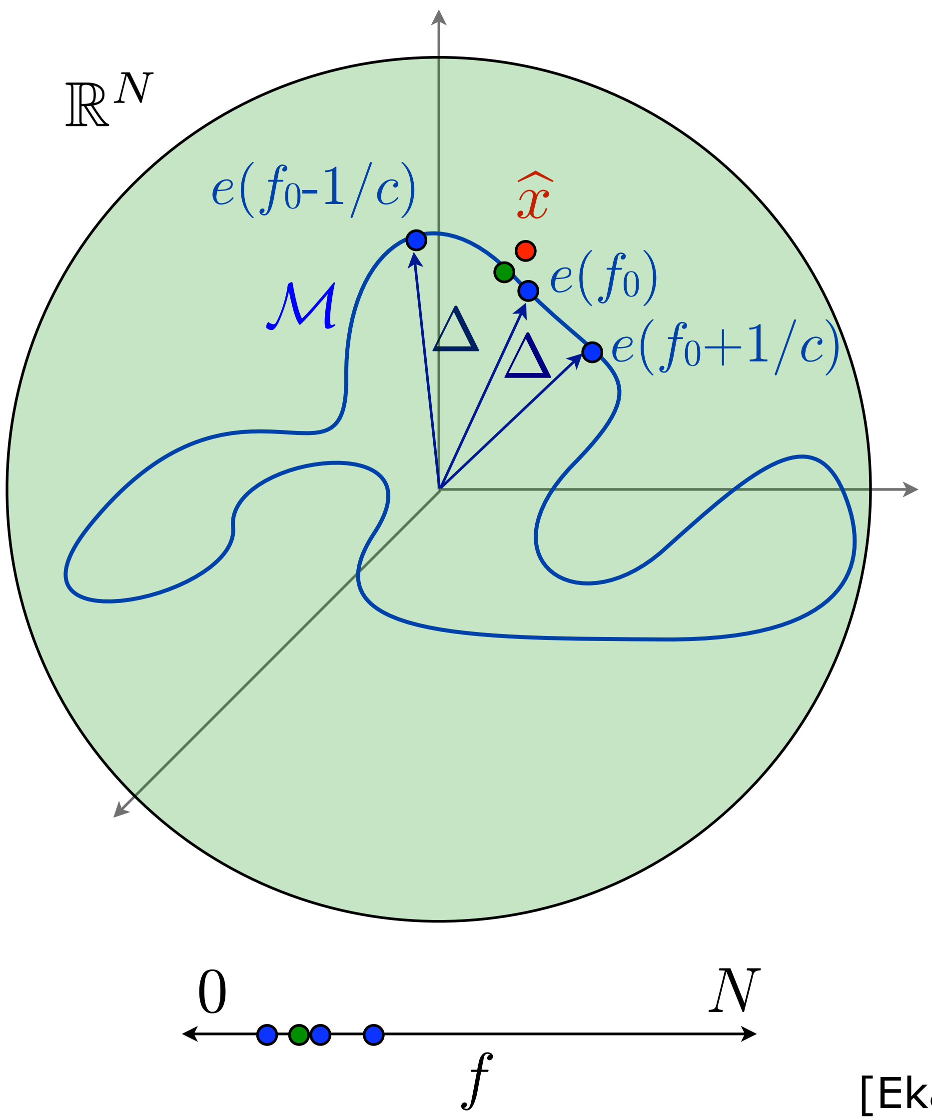


Interpolating the Manifold: Polar Interpolation



- All points in manifold have ***equal norm***; distance b/w samples is ***uniform***
- Manifold must be contained within unit Euclidean ball (***hypersphere***)
- ***Project*** signal estimates into hypersphere
- Find closest point in manifold by ***interpolating*** from closest samples with polar coordinates
- Integrate band exclusion to get ***Band-Excluding Interpolating SP*** (BISP)

Interpolating the Manifold: Polar Interpolation



- In BISP, find closest point in manifold by interpolating from closest samples with ***polar coordinates***:

$$e(f_0 - 1/c) \leftrightarrow \angle = \theta_0 - \Delta$$

$$e(f_0) \leftrightarrow \angle = \theta_0$$

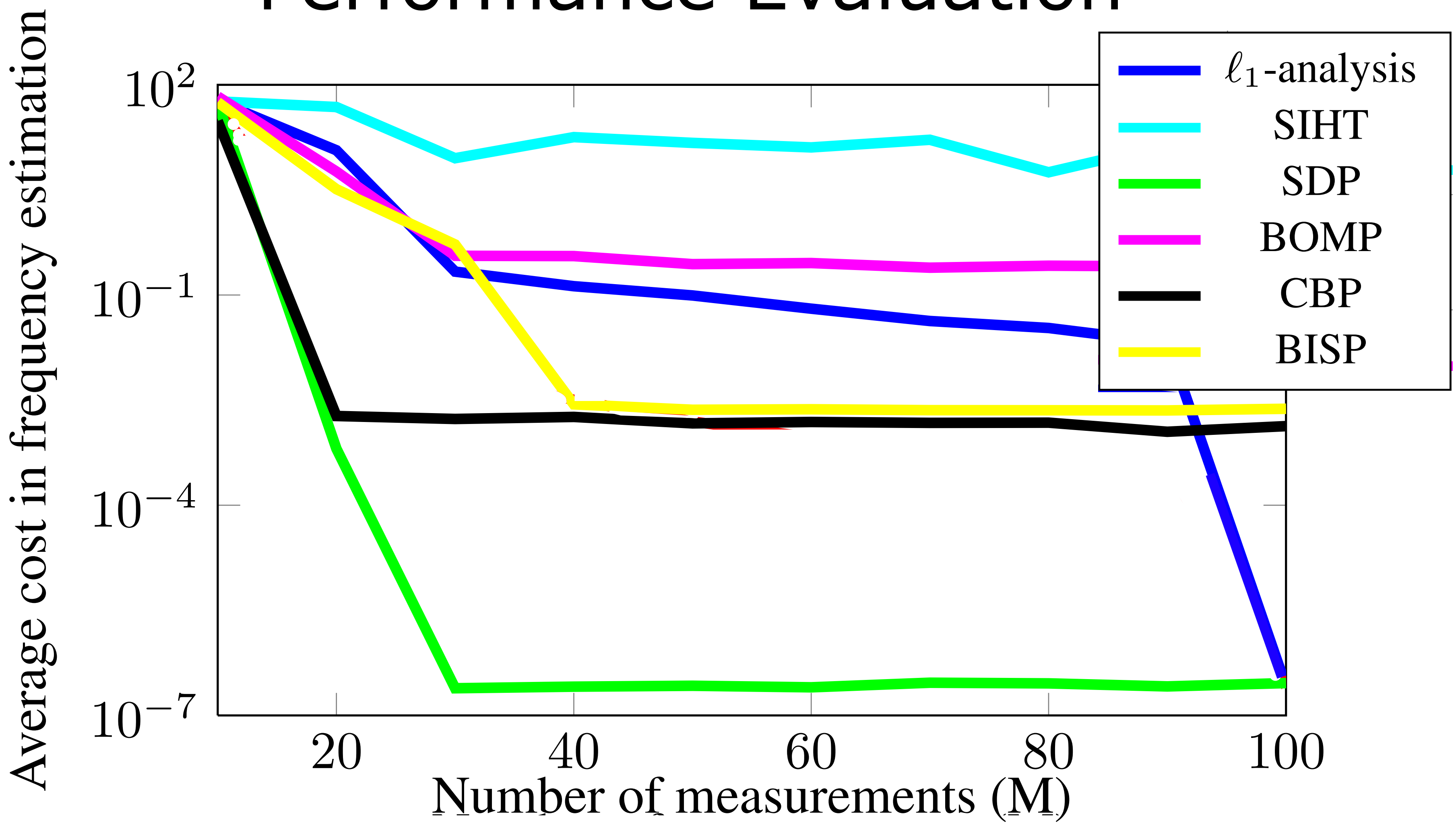
$$e(f_0 + 1/c) \leftrightarrow \angle = \theta_0 + \Delta$$

$$\widehat{x} \leftrightarrow \angle = ?$$

- Map back from manifold to frequency estimates (***parameter space***)

Akin to Continuous Basis Pursuit (CBP)
[Ekanadham, Tranchina, and Simoncelli 2011]

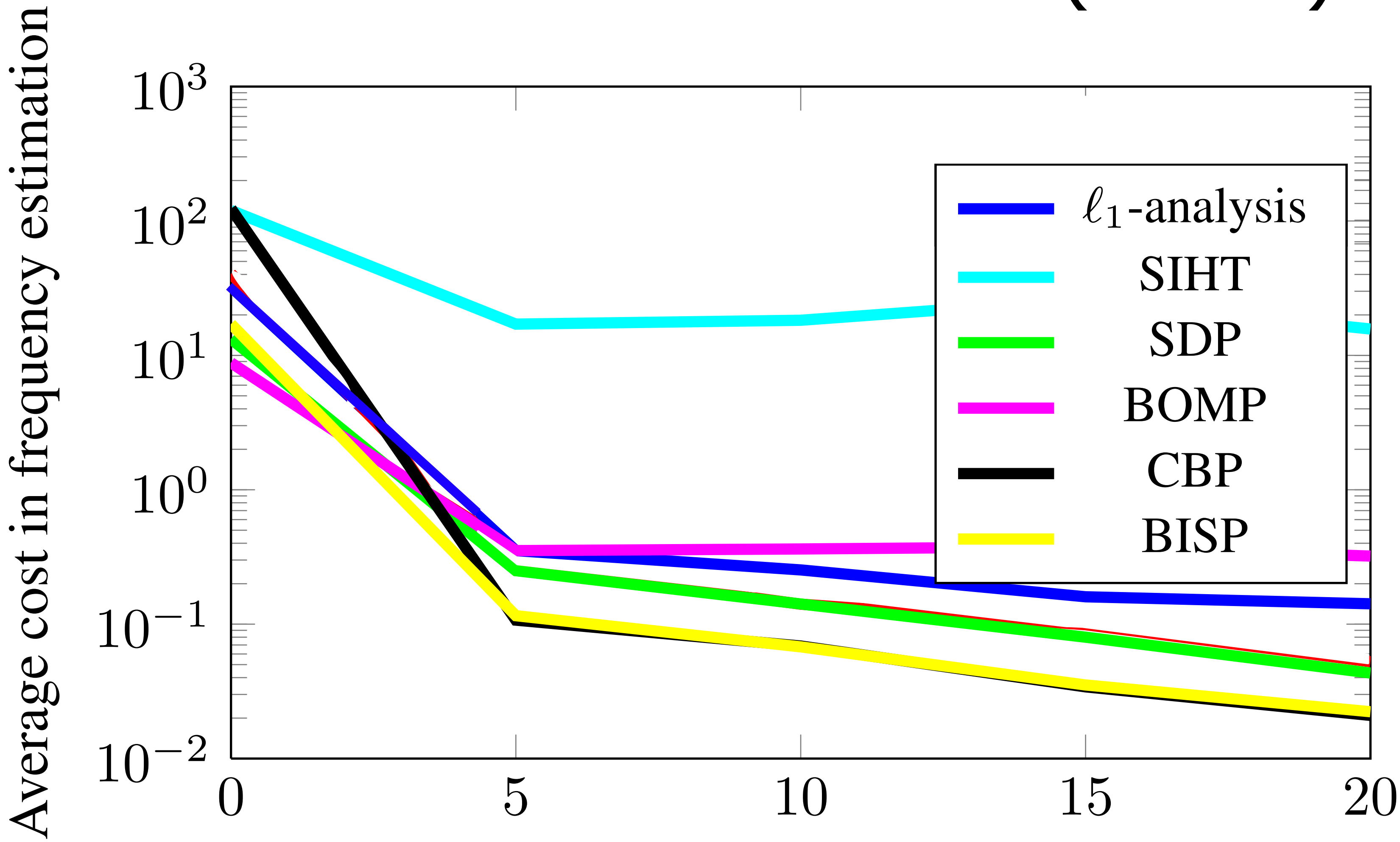
Compressive Line Spectral Estimation: Performance Evaluation



$N = 100,\ K = 4,$
$c = 5, \Delta f\ = 0.2$ Hz

BOMP [Fannjiang and Liao 2012]
SDP: Atomic Norm Minimization
[Tang, Rhaskar, Shah, Recht 2012]

Compressive Line Spectral Estimation: Performance Evaluation (Noise)



$N = 100$, $K = 4$, $M = 50$, $c = 5$, $\Delta f = 0.2$ Hz

Compressive Line Spectral Estimation: Computational Expense

Time (seconds)	Noiseless	Noisy
ℓ_1-analysis	9.5245	8.8222
SIHT	0.2628	0.1499
SDP	8.2355	9.9796
BOMP	0.0141	0.0101
CBP	46.9645	40.3477
BISP	5.4265	1.4060

Conclusions

- Spectral CS provides significant improvements on frequency-sparse signal recovery
  - address ***coherent dictionaries*** via structured sparsity
  - ***simple-to-implement*** modifications to recovery algs
  - can leverage decades of work on ***spectral estimation***
  - robust to model mismatch, presence of noise
- Compressive line spectral estimation:
  - recovery via ***parametric dictionaries*** provides compressive parameter estimation
  - dictionary elements as samples from ***manifold*** models
  - from dictionaries to manifolds via ***interpolation*** techniques
  - from recovery to ***parameter estimation*** from compressive measurements
  - localization, bearing estimation, radar imaging, ...

http://www.ecs.umass.edu/~mduarte mduarte@ecs.umass.edu