Matched Filtering from Limited Frequency Samples

Armin Eftekhari
Michael Wakin

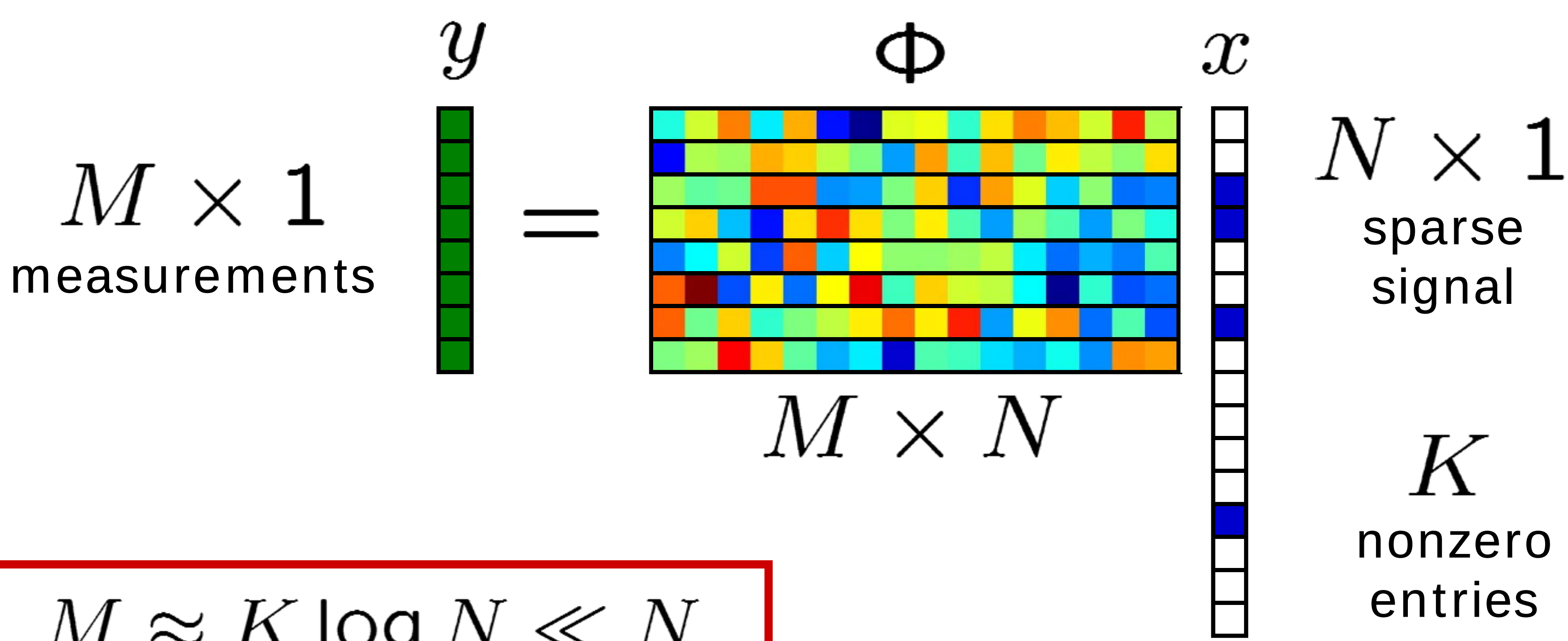
Justin Romberg

Colorado School of Mines

Georgia Tech

Compressive Sensing

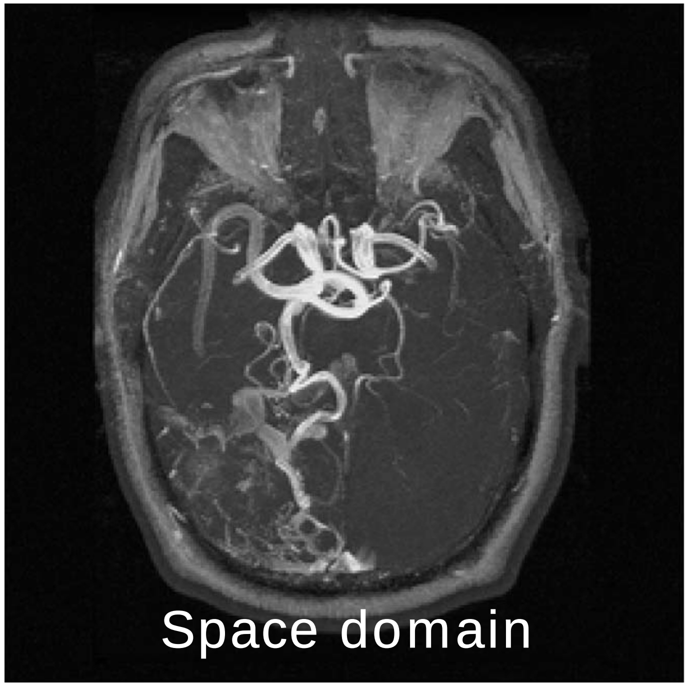
- Signal x is K-***sparse***
- Collect linear measurements $y = \Phi x$
  - ***random*** measurement operator Φ
- Recover x from y by exploiting assumption of sparsity

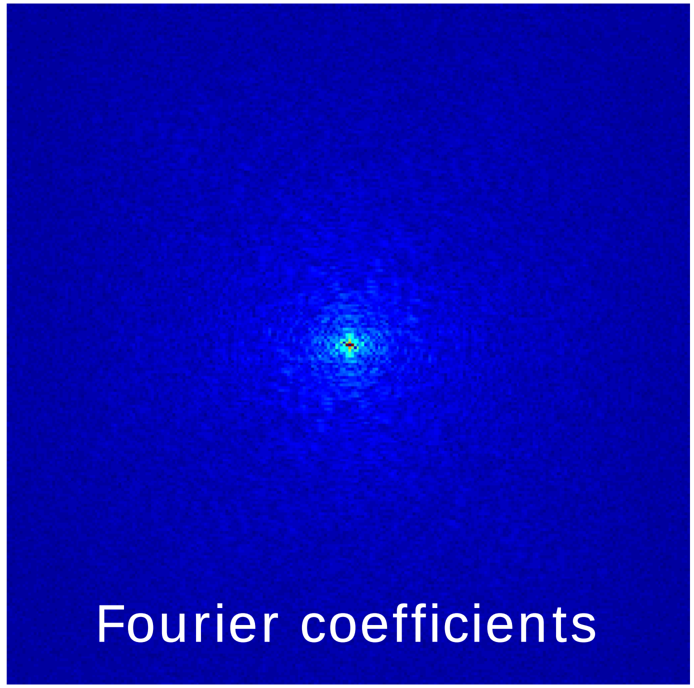


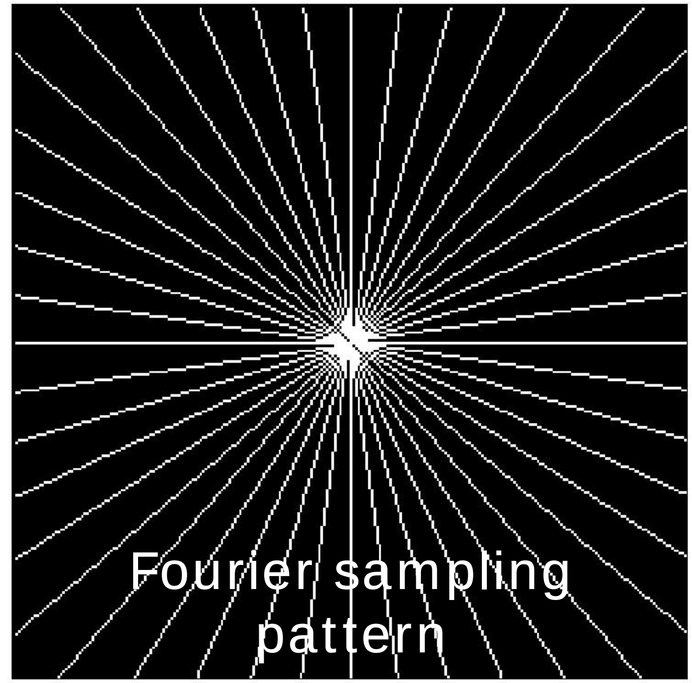
$$M \approx K \log N \ll N$$

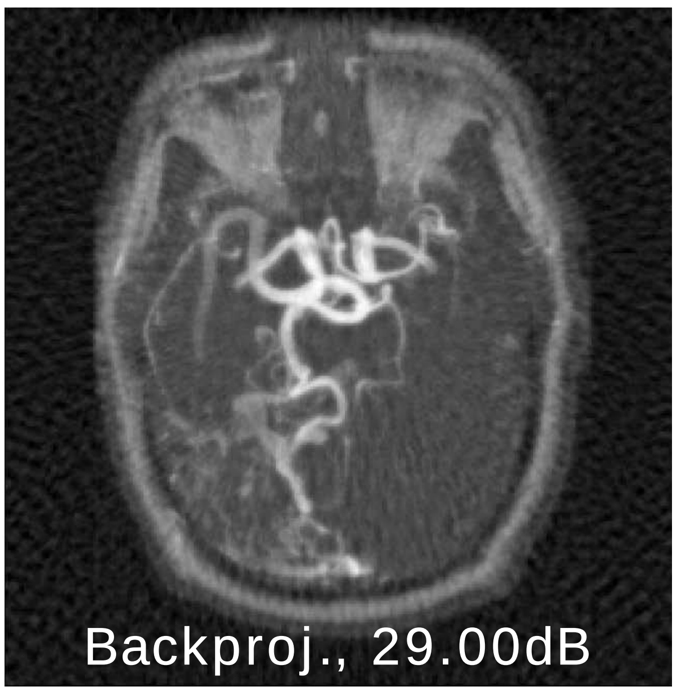
[Candès et al., Donoho, ...]

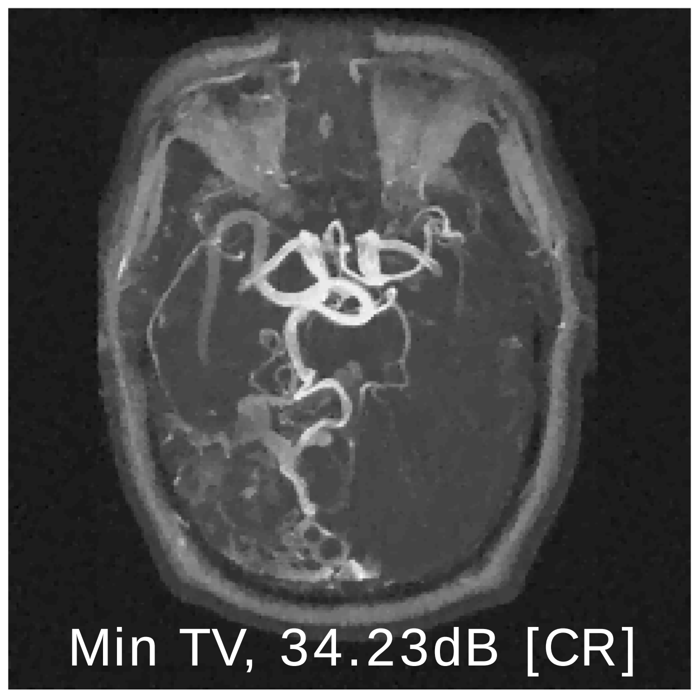
Application 1: Medical Imaging



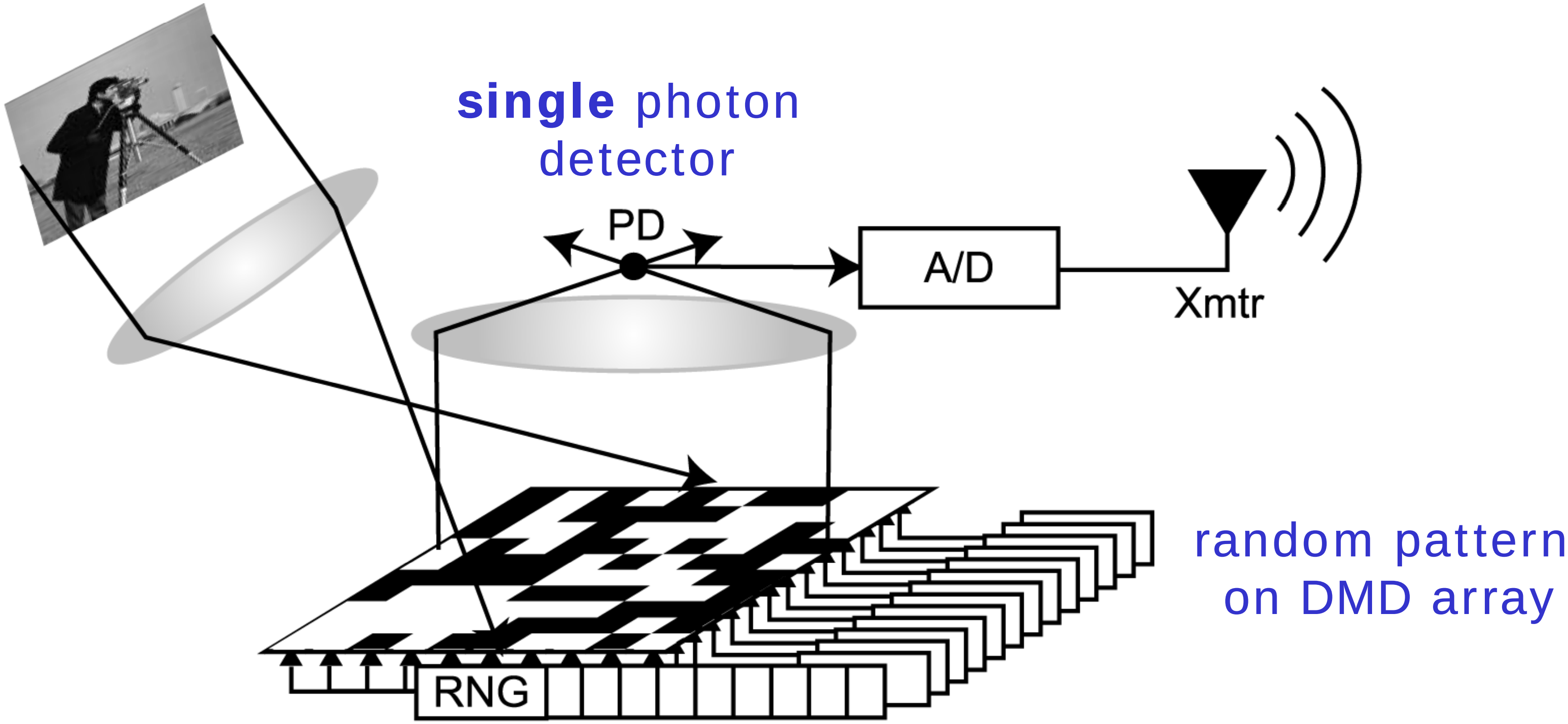




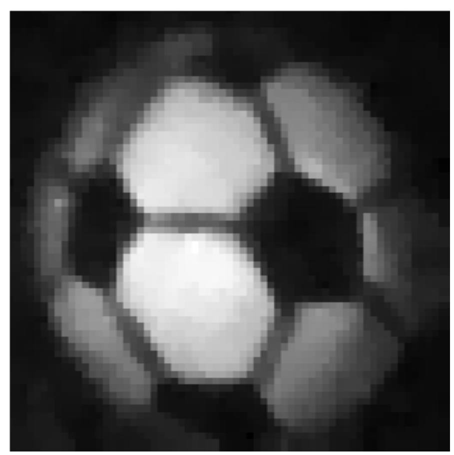


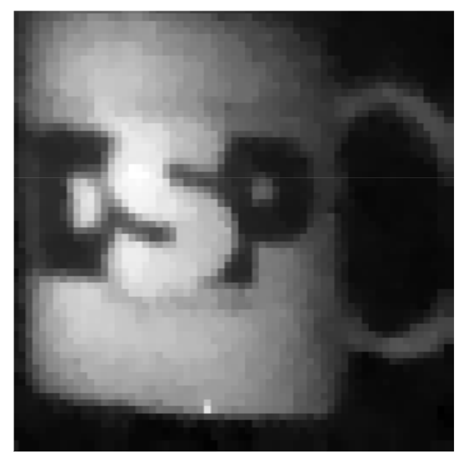


Application 2: Digital Photography



4096 pixels
1600 measurements
(40%)

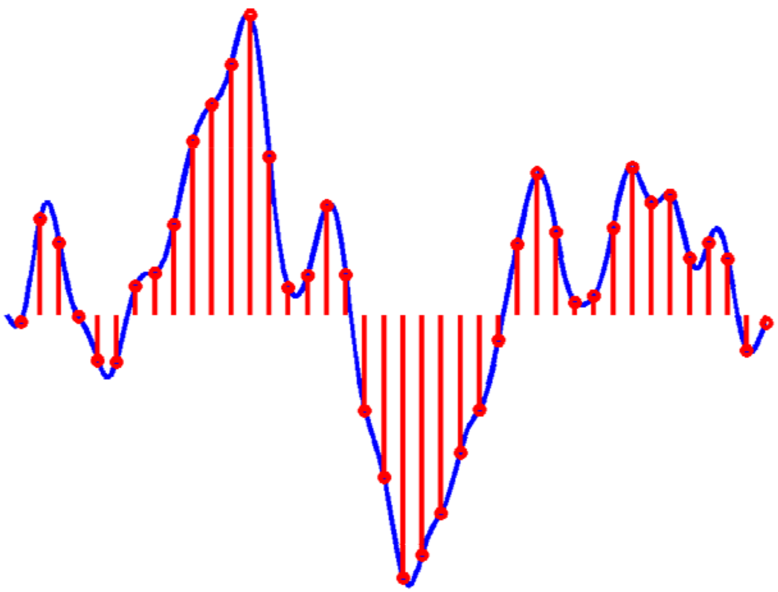




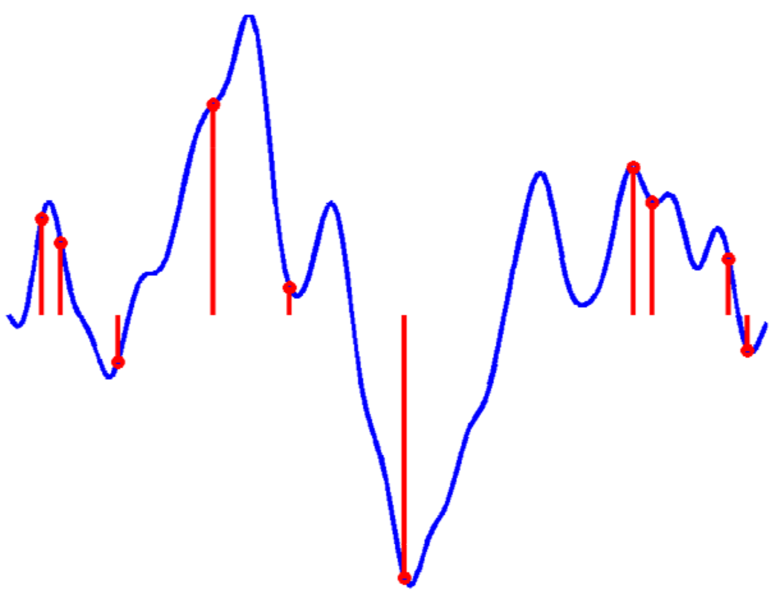
[with R. Baraniuk + Rice CS Team]

Application 3: Analog-to-Digital Conversion

- Sampling analog signals ***at the information level***



Nyquist samples



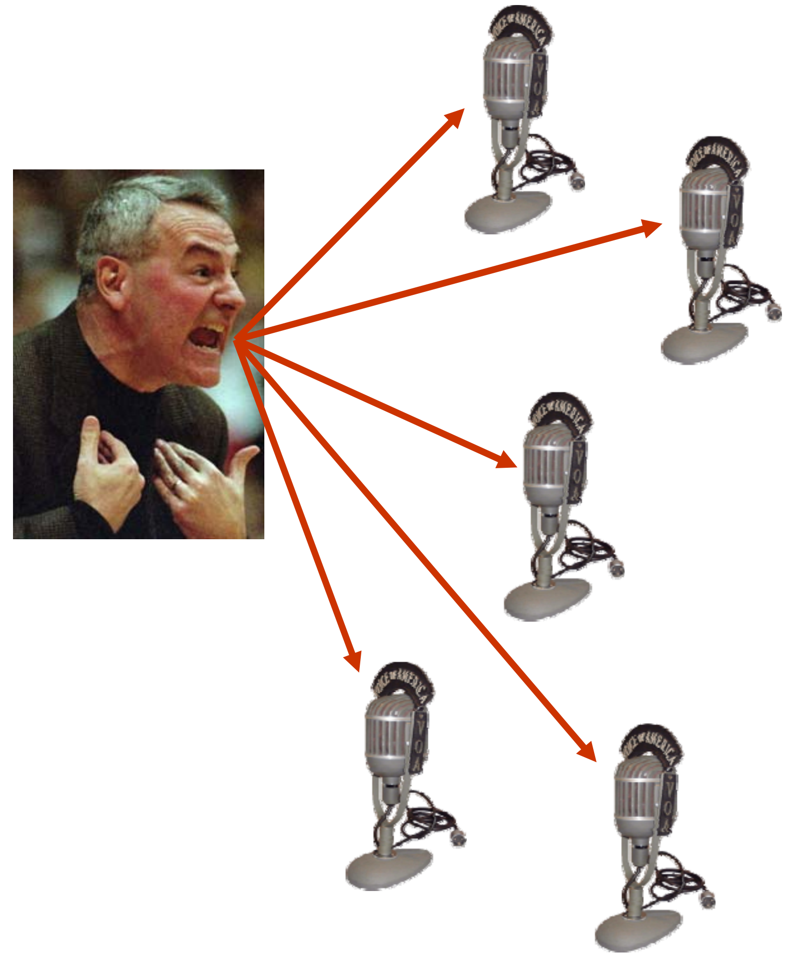
Compressive samples

Application 4: Sensor Networks

- Joint sparsity
- Distributed CS:

    measure separately,
    reconstruct jointly

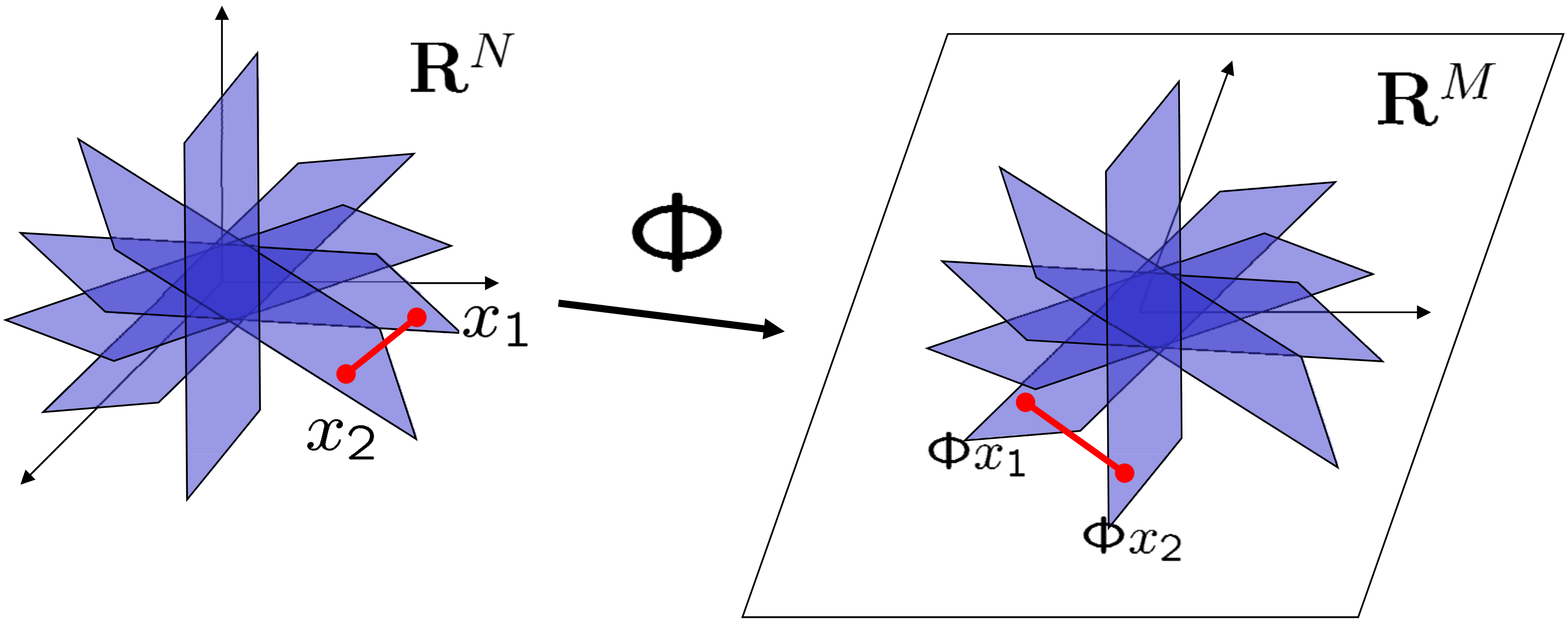
  distributed source coding
- Robust, scalable



Restricted Isometry Property (RIP)

- RIP requires: for all K-sparse x_1 and x_2,

$$(1-\delta) \leq \frac{\|\Phi x_1 - \Phi x_2\|_2^2}{\|x_1 - x_2\|_2^2} \leq (1+\delta)$$



- Stable embedding of the sparse signal family

Proving that Random Matrices Work

- Goal is to prove that for all $(2K)$-sparse $x \in R^N$

$$(1-\delta) \leq \frac{\|\Phi x\|_2^2}{\|x\|_2^2} \leq (1+\delta)$$

- Recast as a bound on a random process

$$\sup_{x \in \Sigma} \left| \|\Phi x\|_2^2 - 1 \right|$$

where Σ is the set of all $(2K)$-sparse signals x with $\|x\|_2^2 = 1.$

Bounding a Random Process

- Common techniques:
  - Dudley inequality relates the expected supremum of a random process to the geometry of its index set

$$\mathbf{E} \sup_{x \in \Sigma} \left| \|\Phi x\|_2^2 - 1 \right| \leq \gamma_1$$

  - strong tail bounds control deviation from average

$$\mathbf{Pr} \left\{ \sup_{x \in \Sigma} \left| \|\Phi x\|_2^2 - 1 \right| > \gamma_1 + \gamma_2 \right\} \leq \gamma_3$$

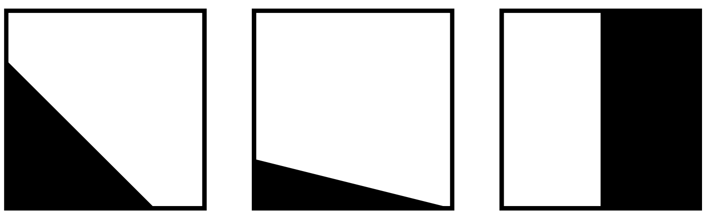
- Works for a variety of random matrix types
  - Gaussian and Fourier matrices [Rudelson and Vershynin]
  - circulant and Toeplitz matrices [Rauhut, Romberg, Tropp]
  - incoherent matrices [Candès and Plan]

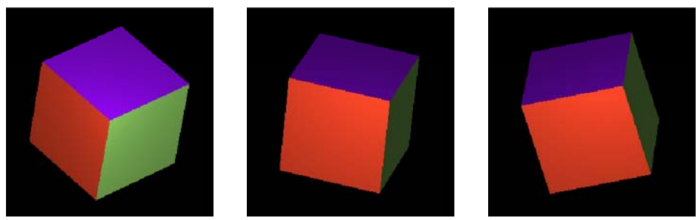
Low-Complexity Inference

- In many problems of interest,

$$\textbf{\textit{information level}} \ll \textbf{\textit{sparsity level}}$$

- Example: unknown signal parameterizations

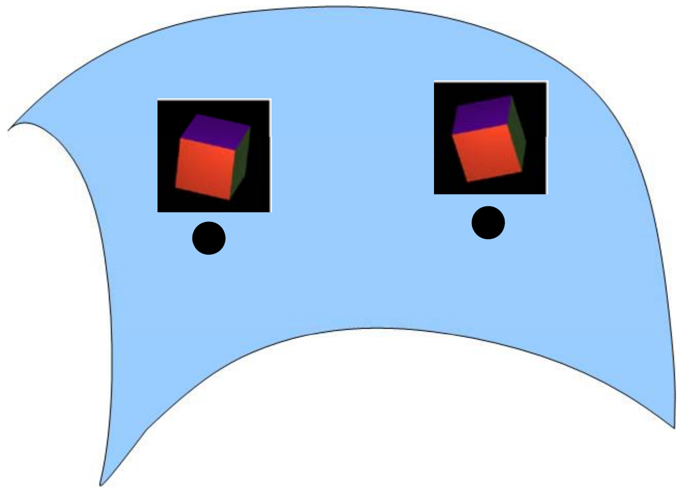




- Is it necessary to fully recover a signal in order to estimate some low-dimensional parameter?
  - Can we somehow exploit the lower information level?
  - Can we exploit the concentration of measure phenomenon?

Compressive Signal Processing

- Low-complexity inference
  - detection/classification
    [Haupt and Nowak; Davenport, W., et al.]
  - estimation ("smashed filtering")
    [Davenport, W., et al.]
    - generic analysis based on stable manifold embeddings



- This talk:
  - focus on simple technique for estimating unknown signal translations from random measurements
    - efficient alternative to conventional matched filter designs
  - special case: pure tone estimation from random samples
  - what's new
    - analysis sharply focused on the estimation statistics
    - analog front end

Tone Estimation

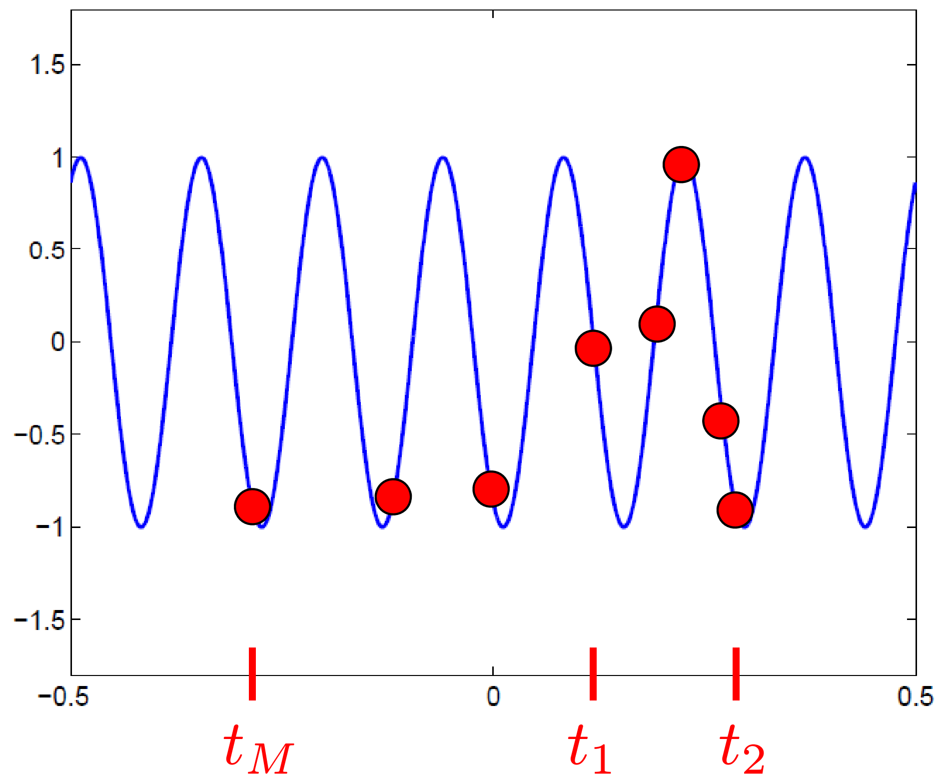
Motivating Scenario

- Analog sinusoid with unknown frequency $\omega_o \in \Omega$

$$e^{j\omega_0 t} = \cos(\omega_0 t) + j\sin(\omega_0 t) \qquad j = \sqrt{-1}$$

- Observe M random samples in time
  - t_m ~ Uniform([-½, ½])

$$y = \begin{bmatrix} e^{j\omega_0 t_1} \\ e^{j\omega_0 t_2} \\ \vdots \\ e^{j\omega_0 t_M} \end{bmatrix}$$



Least-Squares Estimation

- Recall the measurement model

$$y = \begin{bmatrix} e^{j\omega_0 t_1} \\ e^{j\omega_0 t_2} \\ \vdots \\ e^{j\omega_0 t_M} \end{bmatrix}$$

- For every $\omega \in \Omega$, consider the test vector

$$\psi_\omega = \begin{bmatrix} e^{j\omega t_1} \\ e^{j\omega t_2} \\ \vdots \\ e^{j\omega t_M} \end{bmatrix}$$

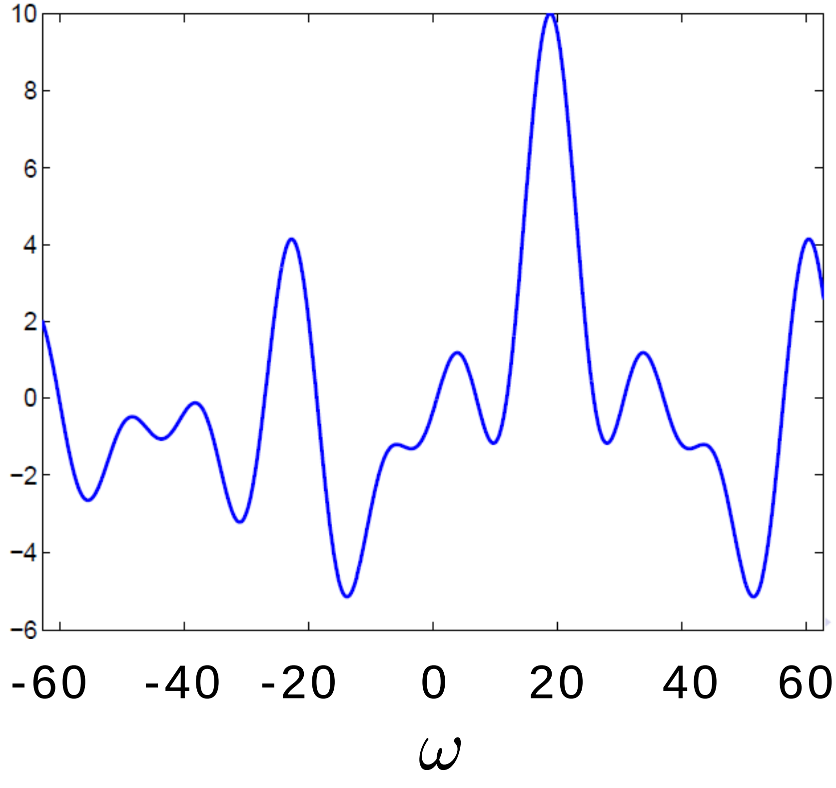
- Compute test statistics $X(\omega) = \langle y, \psi_\omega \rangle$ and let

$$\widehat{\omega}_0 = \arg\max_{\omega \in \Omega} |X(\omega)|$$

Example

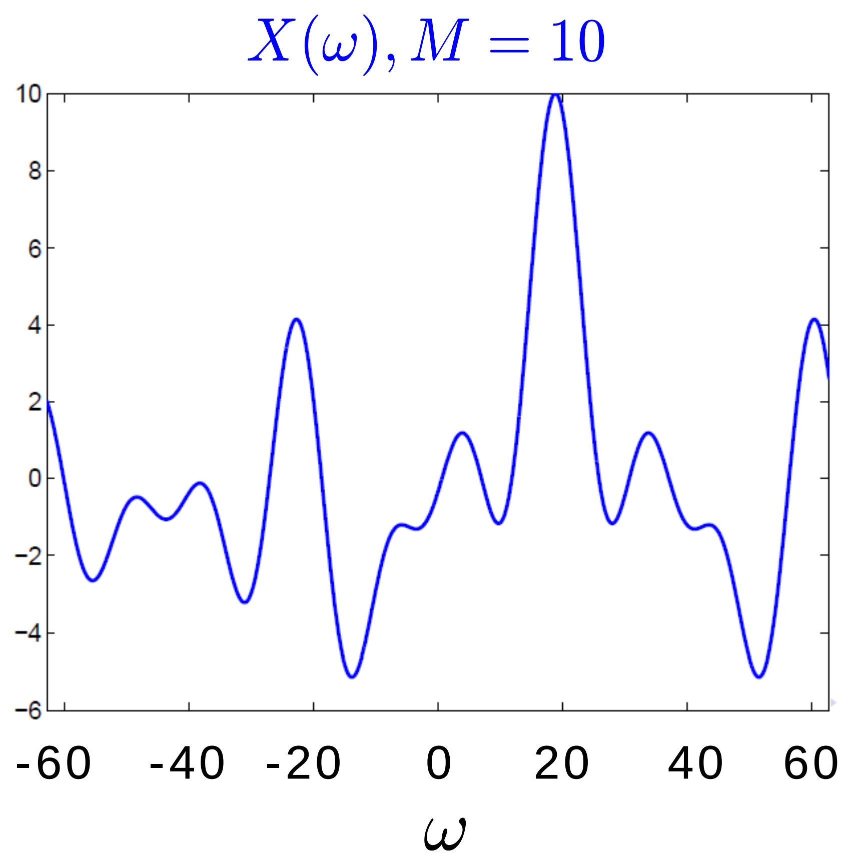
$$X(\omega) = \left\langle \begin{bmatrix} e^{j\omega_0 t_1} \\ e^{j\omega_0 t_2} \\ \vdots \\ e^{j\omega_0 t_M} \end{bmatrix}, \begin{bmatrix} e^{j\omega t_1} \\ e^{j\omega t_2} \\ \vdots \\ e^{j\omega t_M} \end{bmatrix} \right\rangle$$

$X(\omega), M = 10$



Analytical Framework

- $X(\omega) = \langle y, \psi_\omega \rangle$ is a random process indexed by ω.

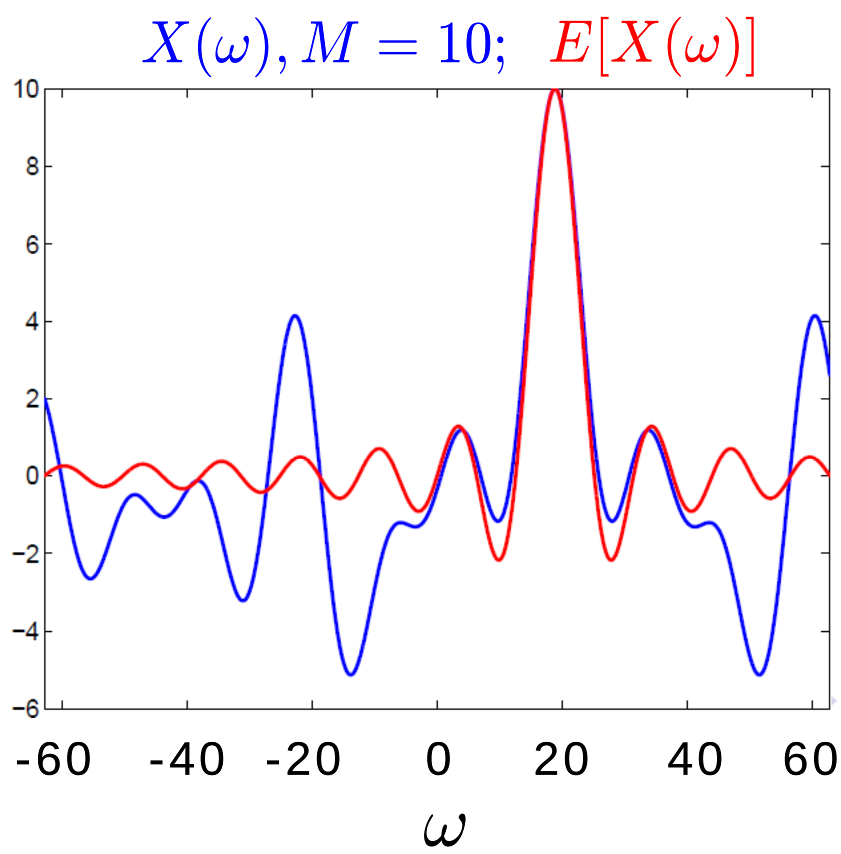


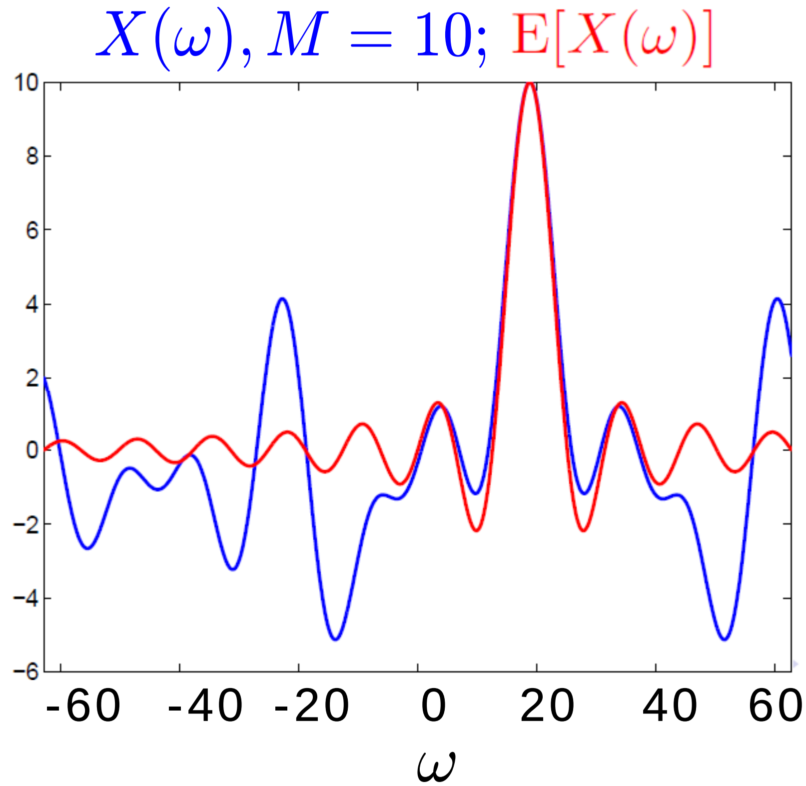
Analytical Framework

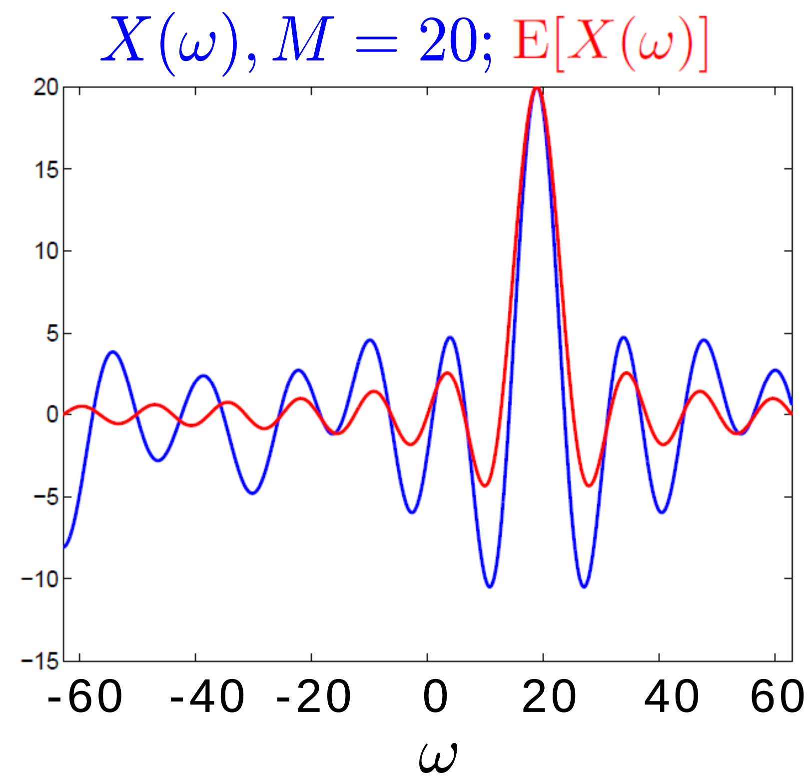
- $X(\omega) = \langle y, \psi_\omega \rangle$ is a random process indexed by ω.
- $X(\omega)$ is an unbiased estimate of the true autocorrelation function:

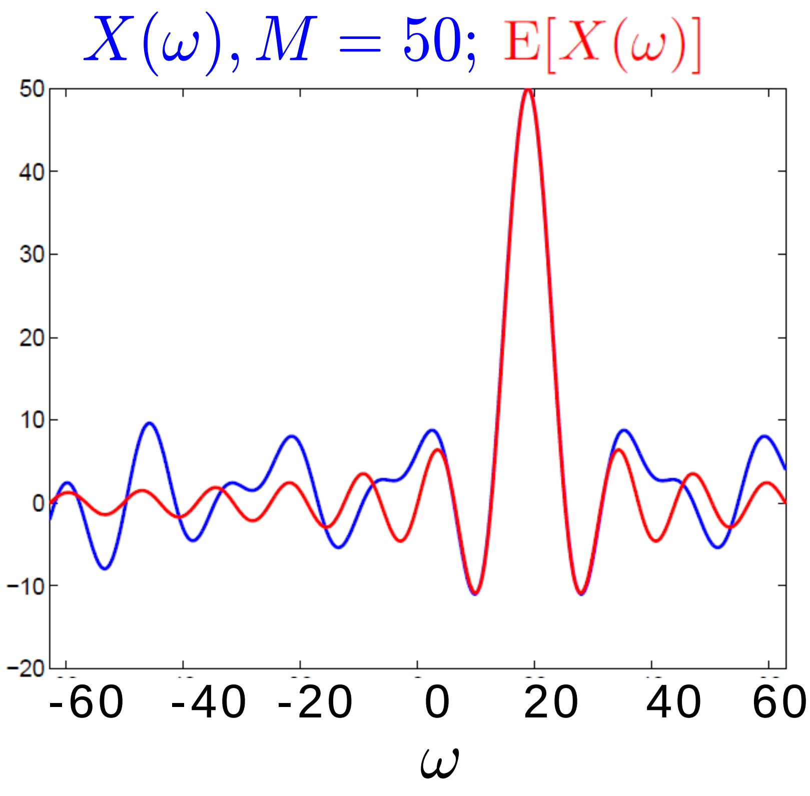
$$\mathrm{E}[X(\omega)] = M \operatorname{sinc}\left(\frac{\omega_0 - \omega}{2}\right)$$

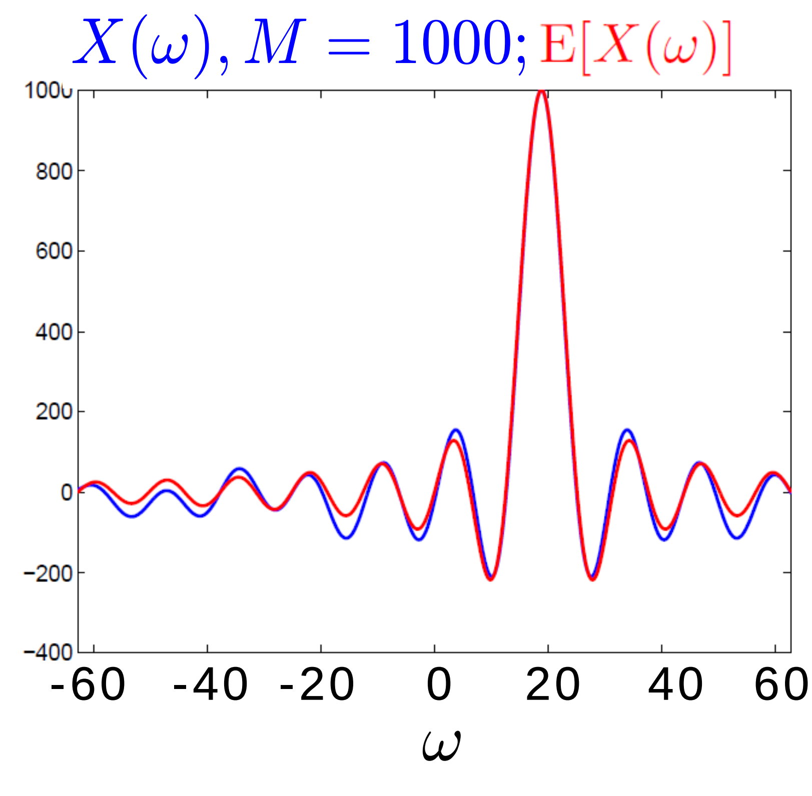
- At each frequency, the variance of the estimate decreases with M.




$X(\omega), M = 10;$ $\mathrm{E}[X(\omega)]$
10
8
6
4
2
0
-2
-4
-6
-60 -40 -20 0 20 40 60
ω


$X(\omega), M = 20;$ $\mathrm{E}[X(\omega)]$
20
15
10
5
0
−5
−10
−15
-60 -40 -20 0 20 40 60
ω


$X(\omega), M = 50;$ $\mathrm{E}[X(\omega)]$
50
40
30
20
10
0
−10
−20
-60 -40 -20 0 20 40 60
ω

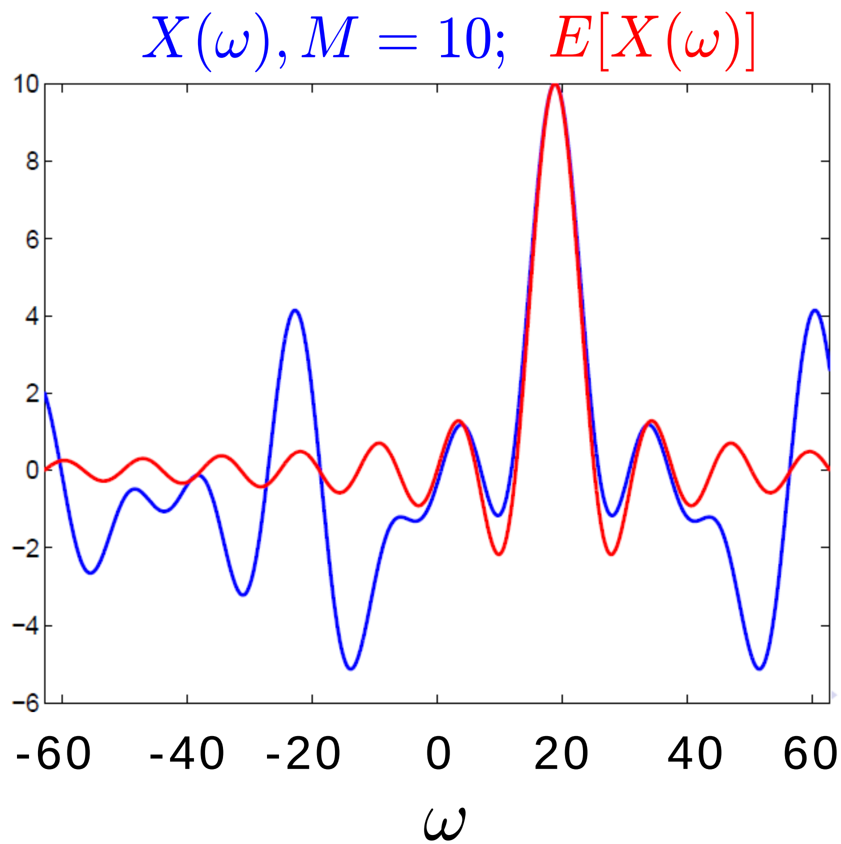

$X(\omega), M = 1000;$ $\mathrm{E}[X(\omega)]$
1000
800
600
400
200
0
−200
−400
-60 -40 -20 0 20 40 60
ω

Analytical Framework

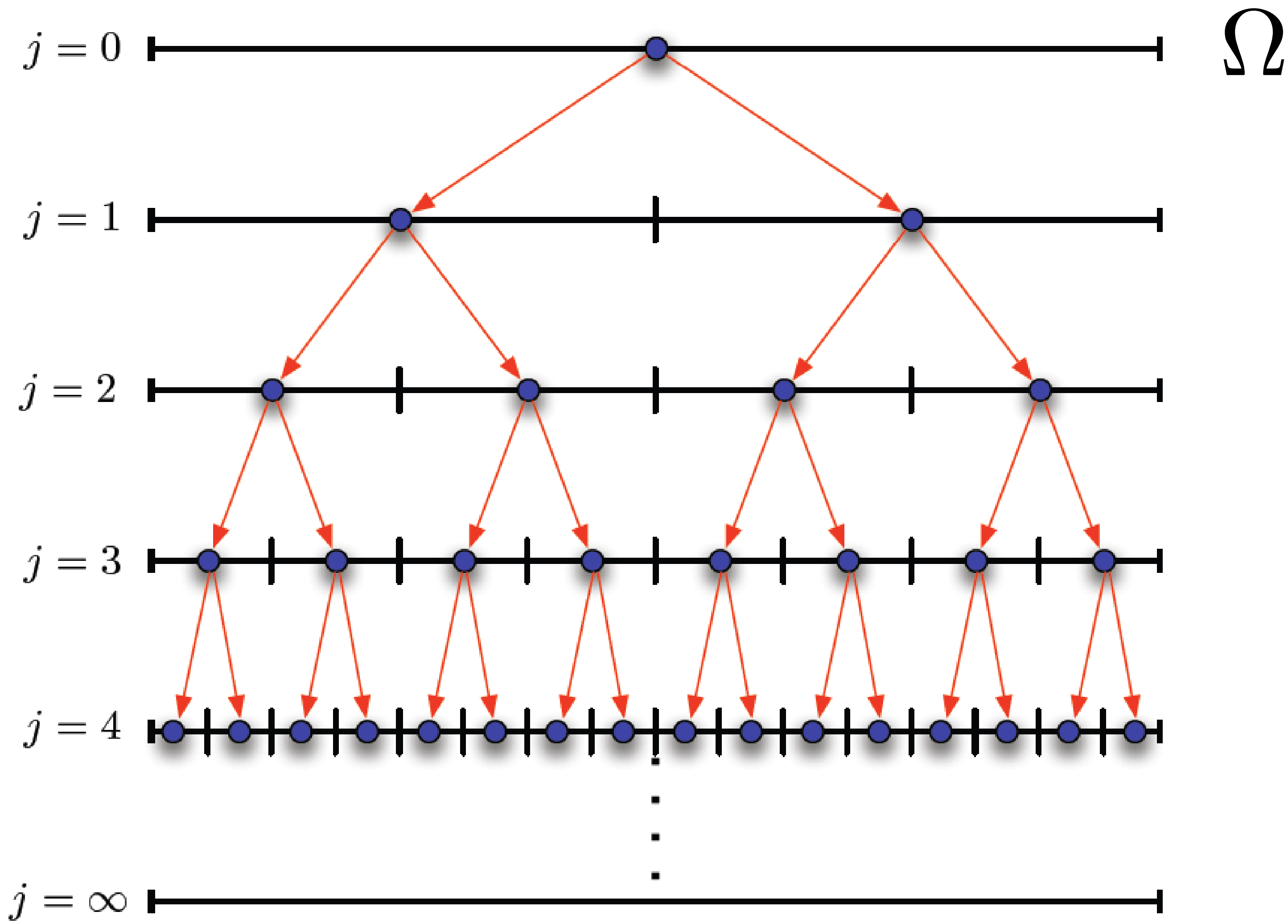
- When will $X(\omega)$ peak at or near the correct ω_o?
- Can we bound the maximum (supremum) of

$$|\, X(\omega) \;-\; E[X(\omega)]\,|$$

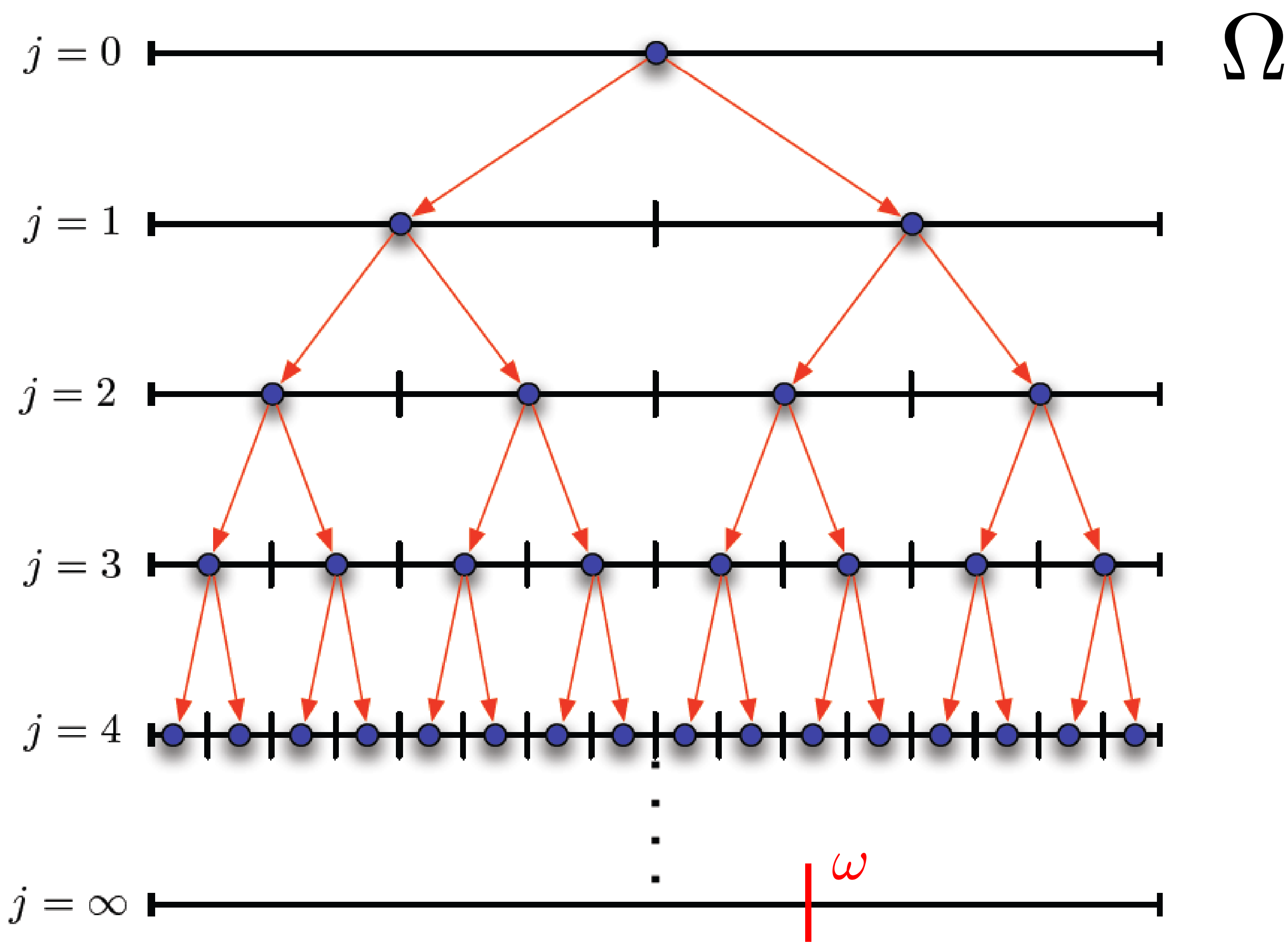
over the infinite set of frequencies $\omega \in \Omega$?



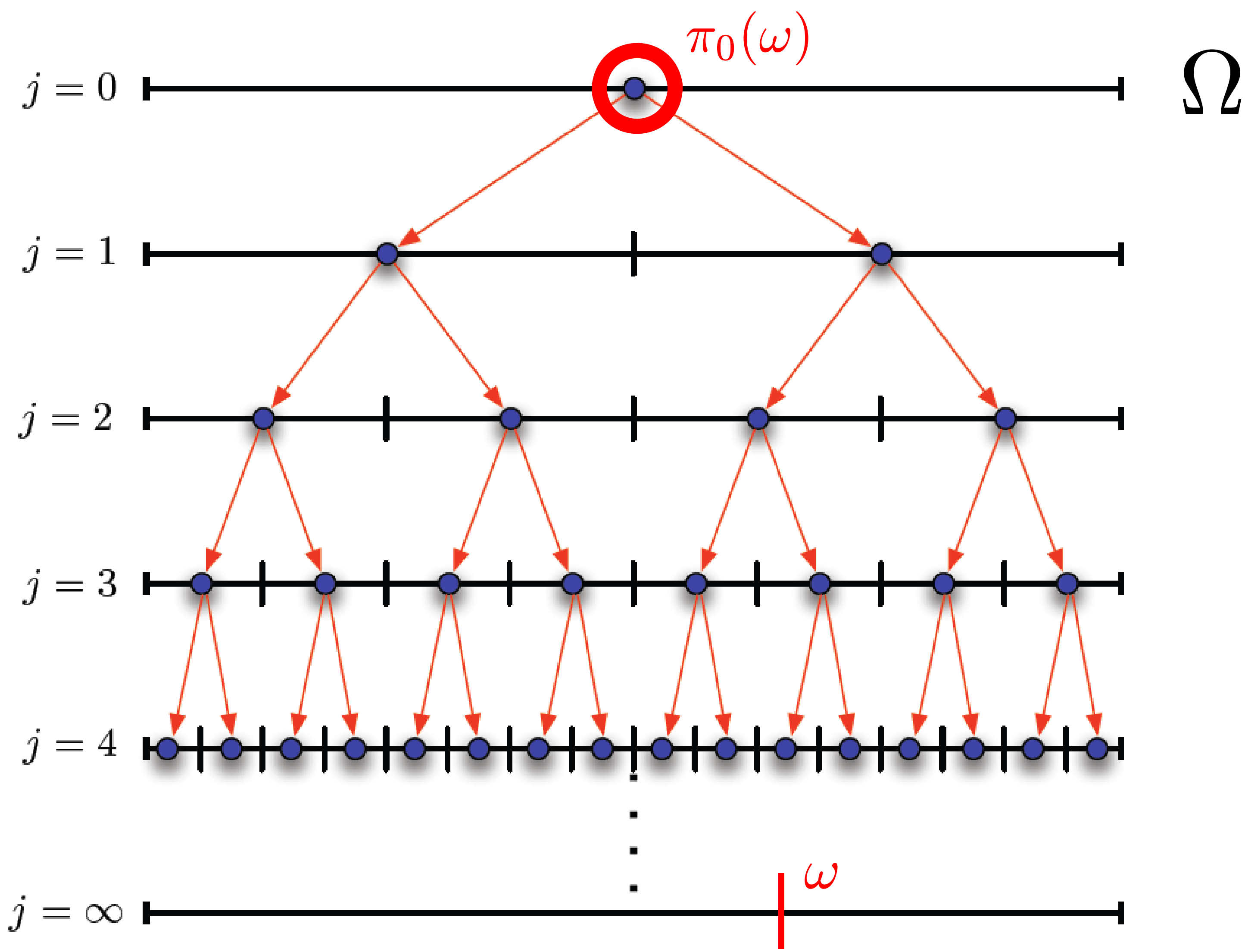
Chaining



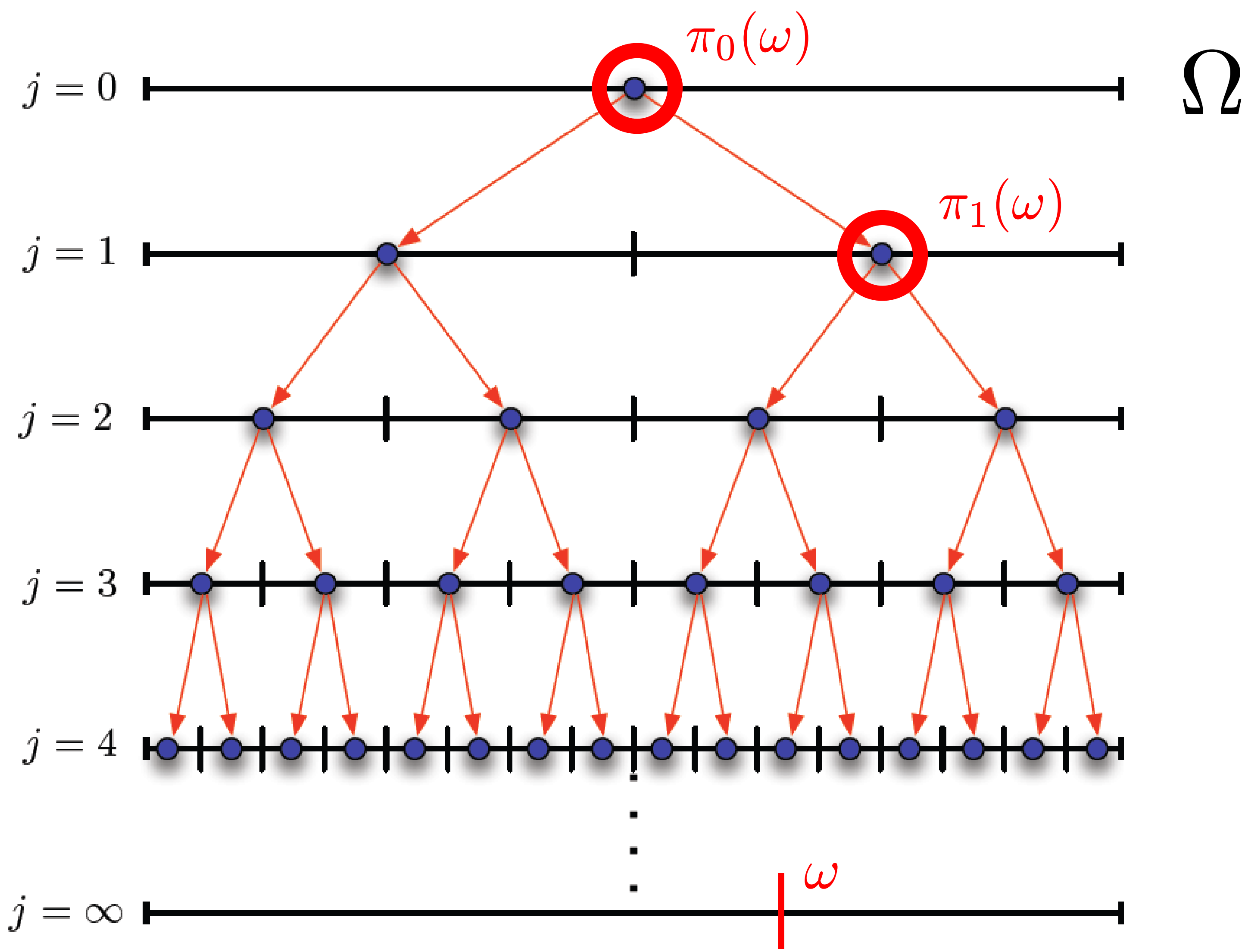
Chaining



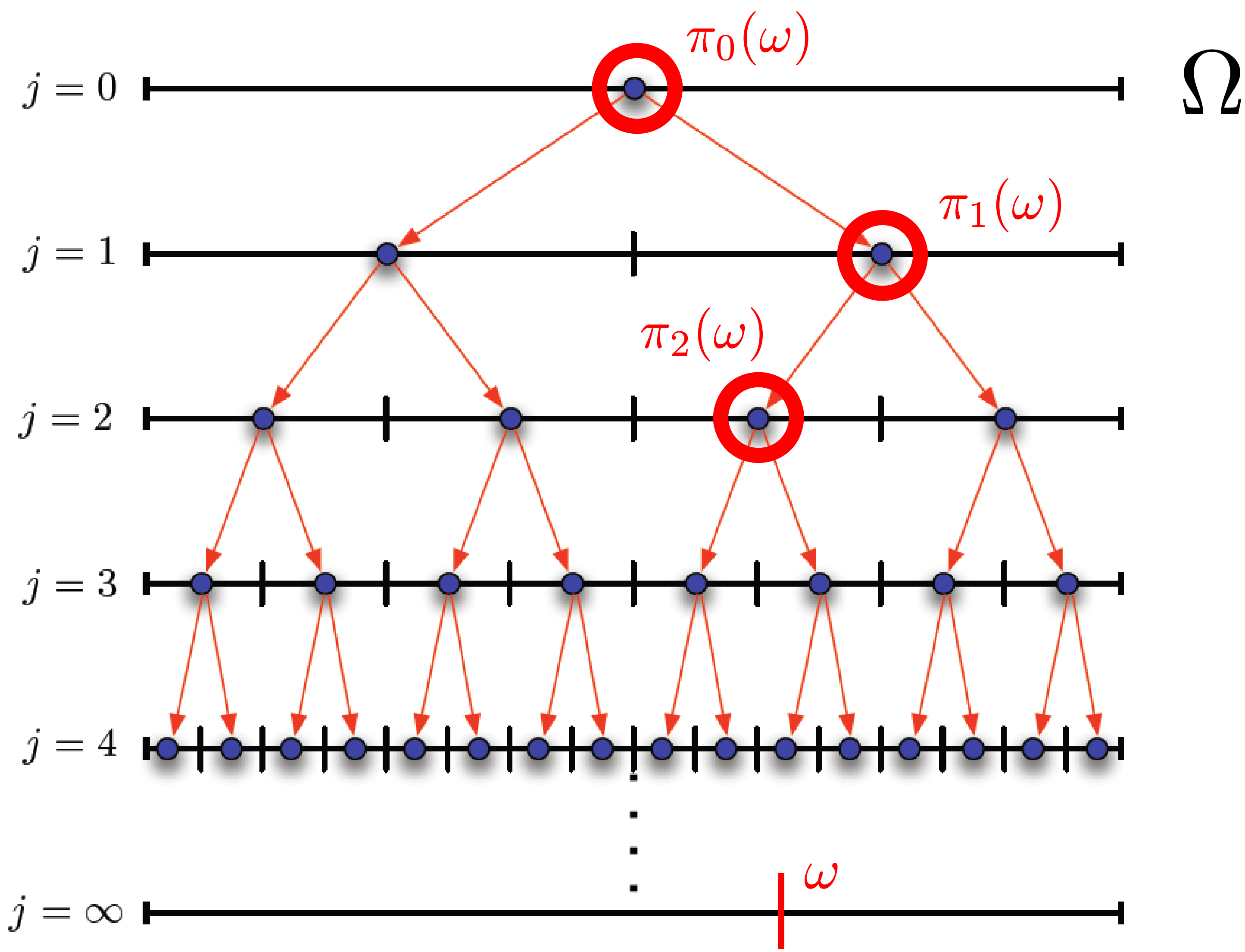
Chaining



Chaining

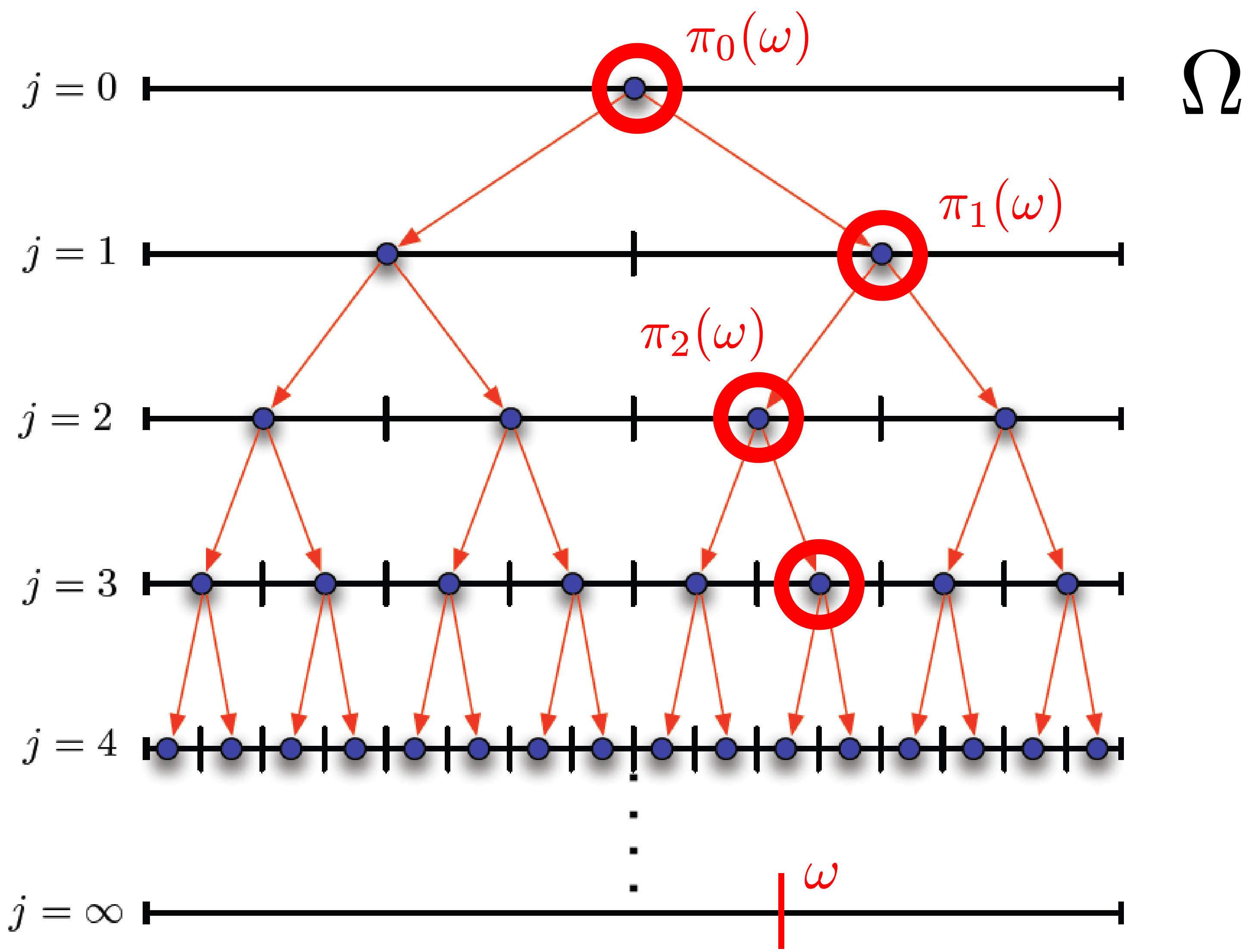


Chaining



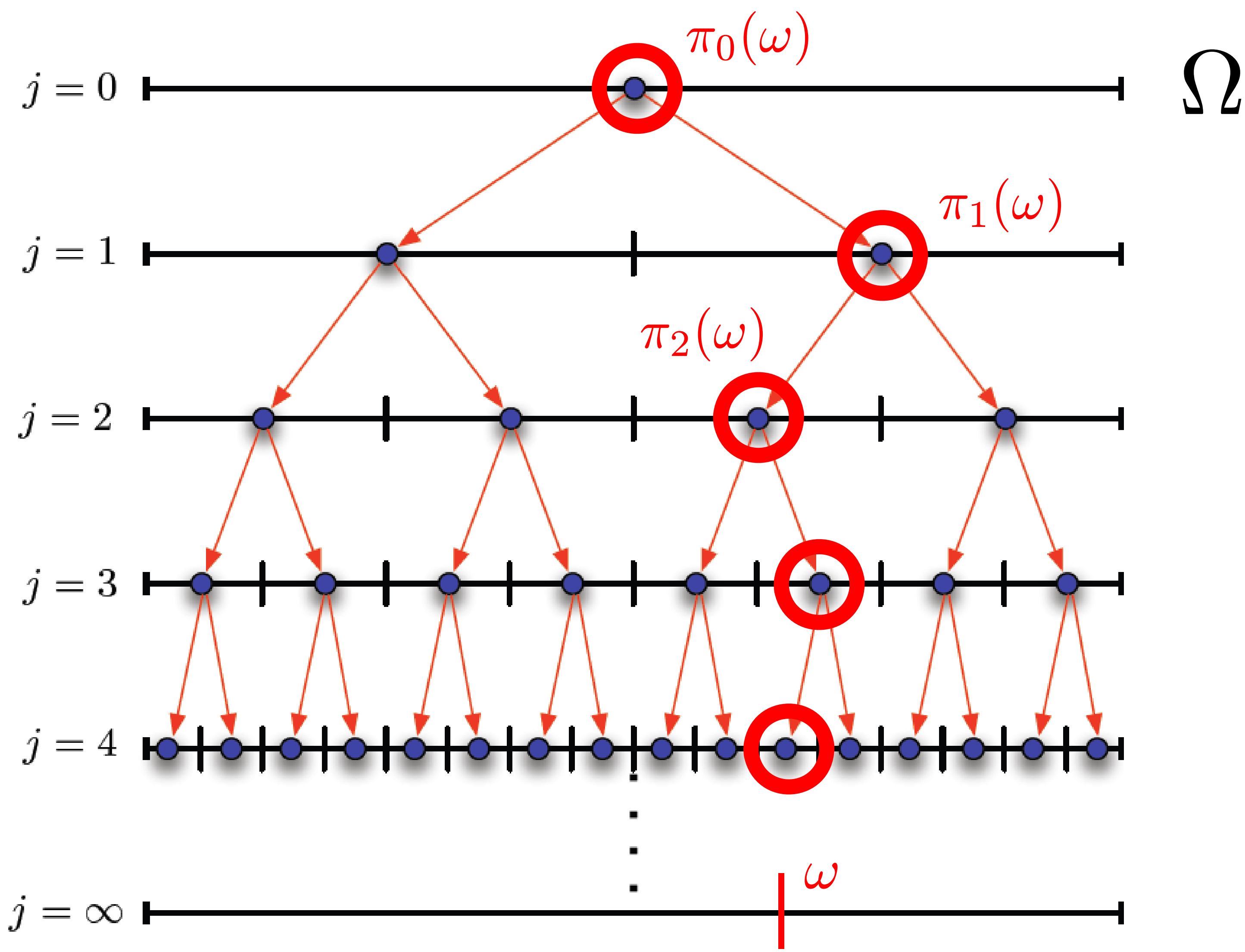
$\pi_0(\omega)$
Ω
$j = 0$
$\pi_1(\omega)$
$j = 1$
$\pi_2(\omega)$
$j = 2$
$j = 3$
$j = 4$
ω
$j = \infty$

Chaining



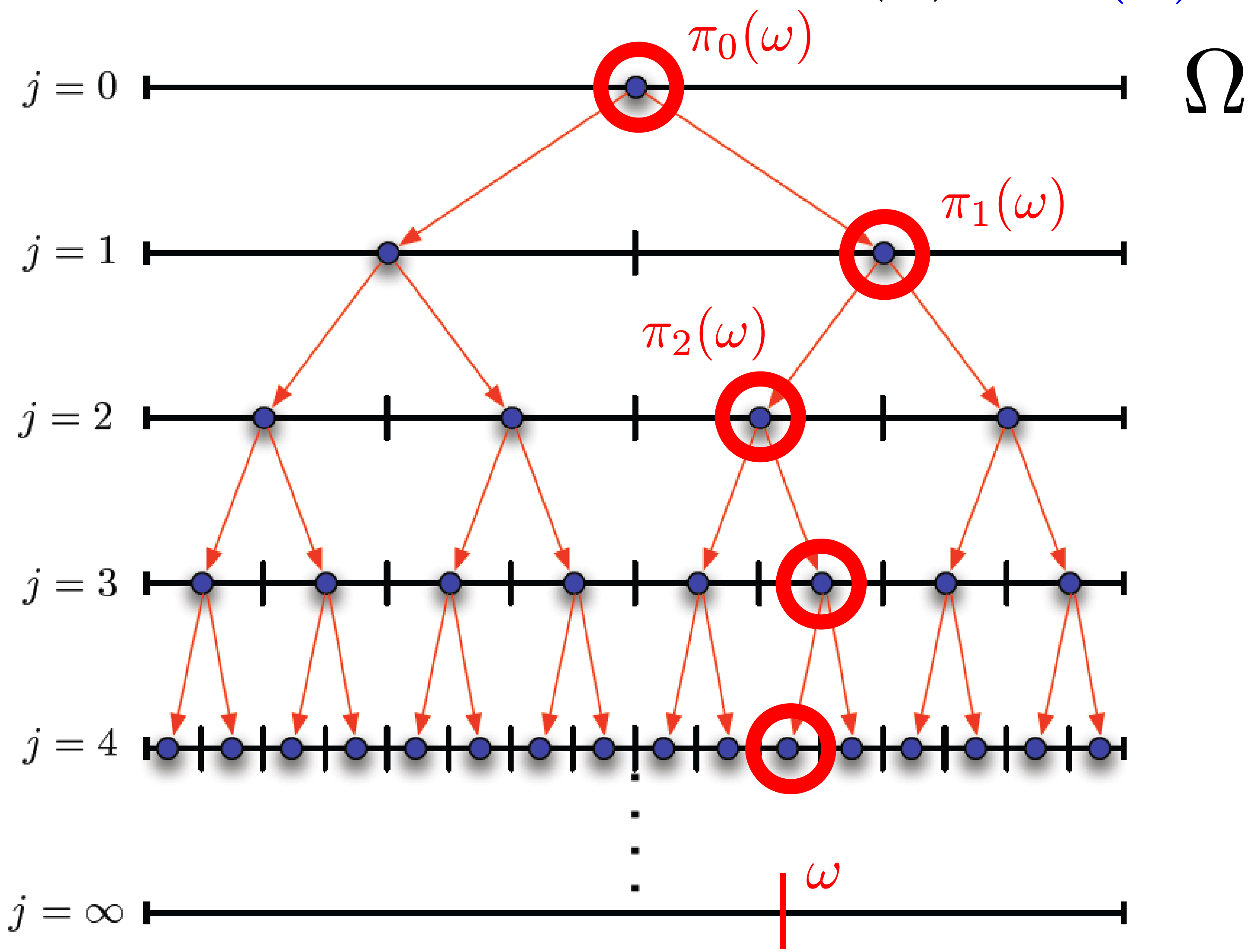
$\pi_0(\omega)$
$\pi_1(\omega)$
$\pi_2(\omega)$
Ω
$j = 0$
$j = 1$
$j = 2$
$j = 3$
$j = 4$
$j = \infty$
ω

Chaining


$\pi_0(\omega)$
Ω
$\pi_1(\omega)$
$\pi_2(\omega)$
$j = 0$
$j = 1$
$j = 2$
$j = 3$
$j = 4$
$j = \infty$
ω

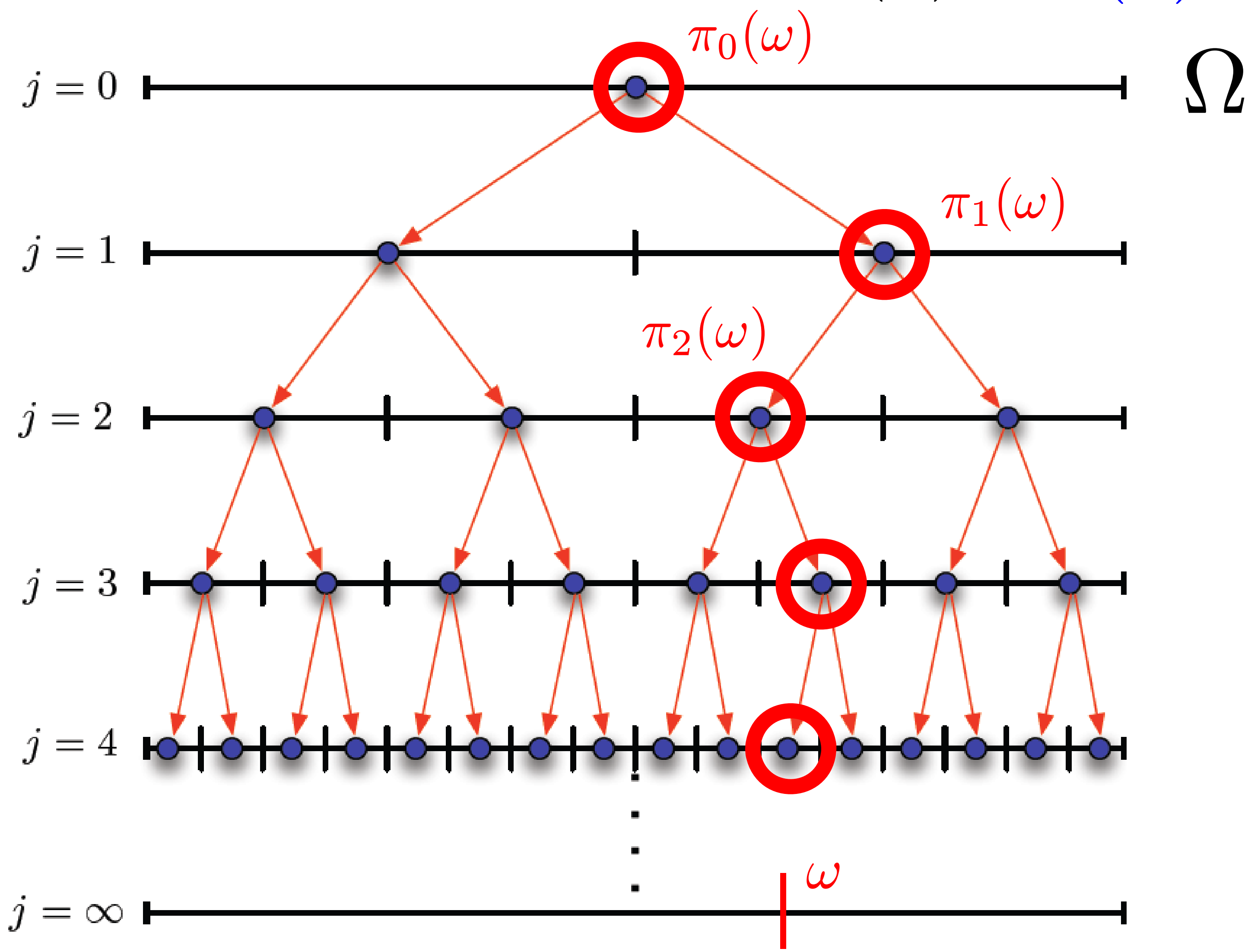
Chaining

- Consider the centered process $Y(\omega) = X(\omega) - E[X(\omega)]$



Chaining

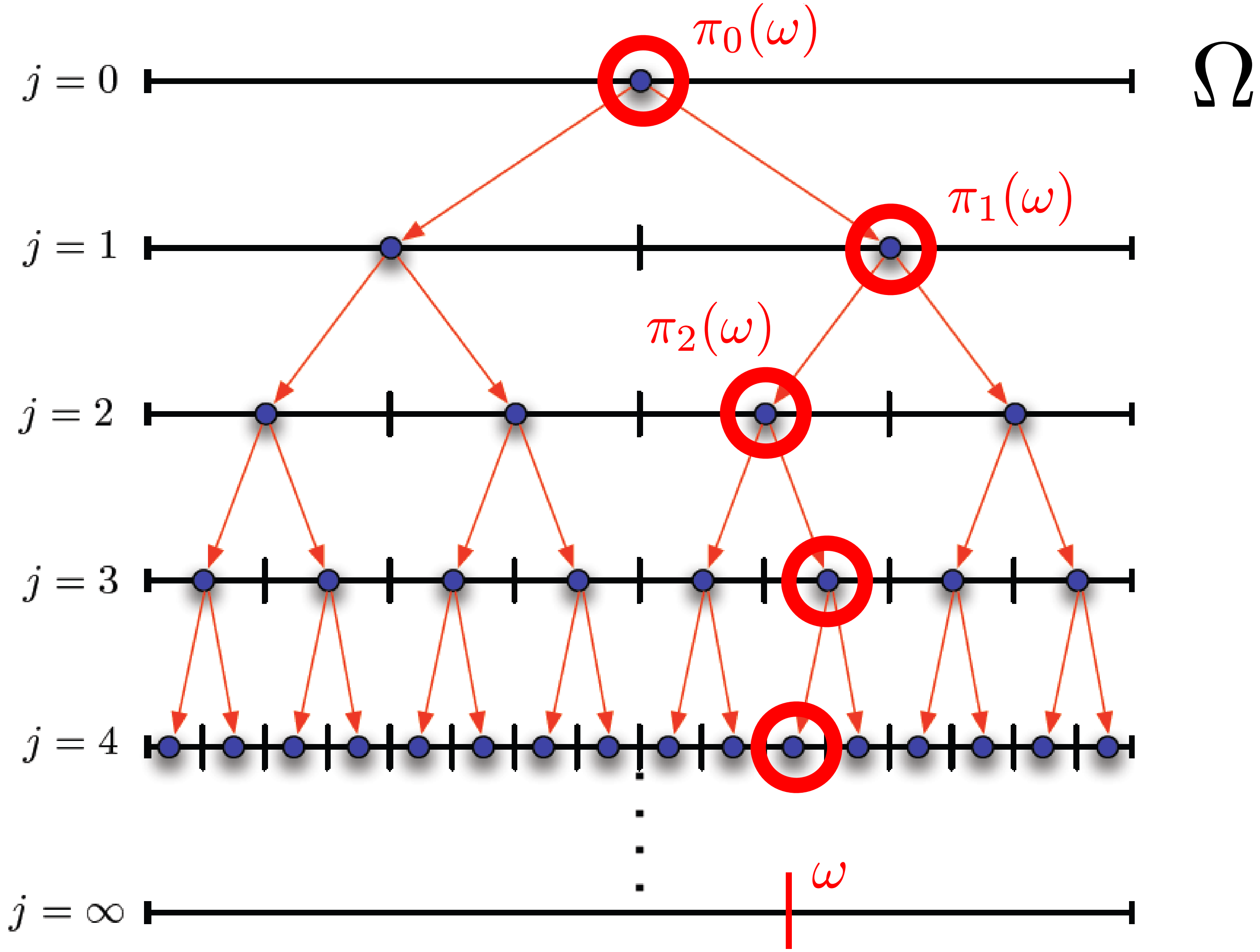
- Consider the centered process $Y(\omega) = X(\omega) - E[X(\omega)]$



$$Y(\omega) = Y(\pi_o(\omega)) + \sum_{j \geq 0} (Y(\pi_{j+1}(\omega)) - Y(\pi_j(\omega)))$$

Chaining

- Consider the centered process $Y(\omega) = X(\omega) - E[X(\omega)]$



$$\sup_{\omega \in \Omega} |Y(\omega)| \leq \max_{p_0 \in \Omega_0} |Y(p_0)| + \sum_{j \geq 0} \max_{(p_j, q_j) \in L_j} |Y(q_j) - Y(p_j)|$$

Modifying the Random Process

- Recall the centered random process

$$Y(\omega) = X(\omega) - E[X(\omega)]$$

- Define an ***independent copy*** called $Y'(\omega)$ with an independent set of random times $\{t'_m\}$
- Define the ***symmetric*** random process

$$Z(\omega) = Y(\omega) - Y'(\omega) = \sum_{m=1}^{M} e^{j(\omega-\omega_0)t_m} - e^{j(\omega-\omega_0)t'_m}$$

- ***Modulate*** with a Rademacher (+/- 1) sequence

$$Z'(\omega) = \sum_{m=1}^{M} \epsilon_m (e^{j(\omega-\omega_0)t_m} - e^{j(\omega-\omega_0)t'_m})$$

Bounding the Random Process

- Conditioned on times $\{t_m\}$ and $\{t'_m\}$, ***Hoeffding's inequality*** bounds $Z'(\omega)$ and its increments:

$$P_{\epsilon_m}\{|Z'(\omega)| > \lambda\} \leq e^{-\frac{C\lambda^2}{M}}$$

$$P_{\epsilon_m}\{|Z'(\omega_1) - Z'(\omega_2)| > \lambda\} \leq e^{-\frac{C\lambda^2}{M|\Omega|^2|\omega_1-\omega_2|^2}}$$

- ***Chaining*** argument bounds supremum of $Z'(\omega)$:

$$\sup_{\omega\in\Omega} |Z'(\omega)| \leq \max_{p_0\in\Omega_0} |Z'(p_0)| + \sum_{j\geq 0} \max_{(p_j,q_j)\in L_j} |Z'(q_j) - Z'(p_j)|$$

- Careful ***union bound*** combines all of this to give:

$$P_{\epsilon_m}\{\sup_{\omega\in\Omega} |Z'(\omega)| > \lambda\} \leq |\Omega| e^{-\frac{C\lambda^2}{M}}$$

Finishing Steps

- After removing the conditioning on times $\{t_m\}$ and $\{t'_m\}$, and relating $Z'(\omega)$ to $Y(\omega)$, we conclude that

$$\mathrm{E}\sup_{\omega\in\Omega}|X(\omega)-E[X(\omega)]| = \mathrm{E}\sup_{\omega\in\Omega}|Y(\omega)| \leq C\cdot\sqrt{M\log|\Omega|},$$

whereas the peak of $\mathrm{E}[X(\omega)]$ scales with M.

- Slightly extending these arguments, we have

$$\sup_{\omega\in\Omega}|X(\omega)-E[X(\omega)]| = \sup_{\omega\in\Omega}|Y(\omega)| \leq C\cdot\sqrt{M\log(|\Omega|/\delta)}$$

with probability at least 1-δ.

Estimation Accuracy

- From our bounds, we conclude that if

$$M \geq C \log(|\Omega|/\delta),$$

then with probability at least 1-δ, the peak of $|X(\omega)|$ will occur within the correct main lobe.

- If our observation interval has length T and we take

$$M \geq C \log(|\Omega||T|/\delta),$$

we are guaranteed a frequency resolution of

$$|\omega_0 - \widehat{\omega}_0| \leq \frac{2\pi}{|T|}$$

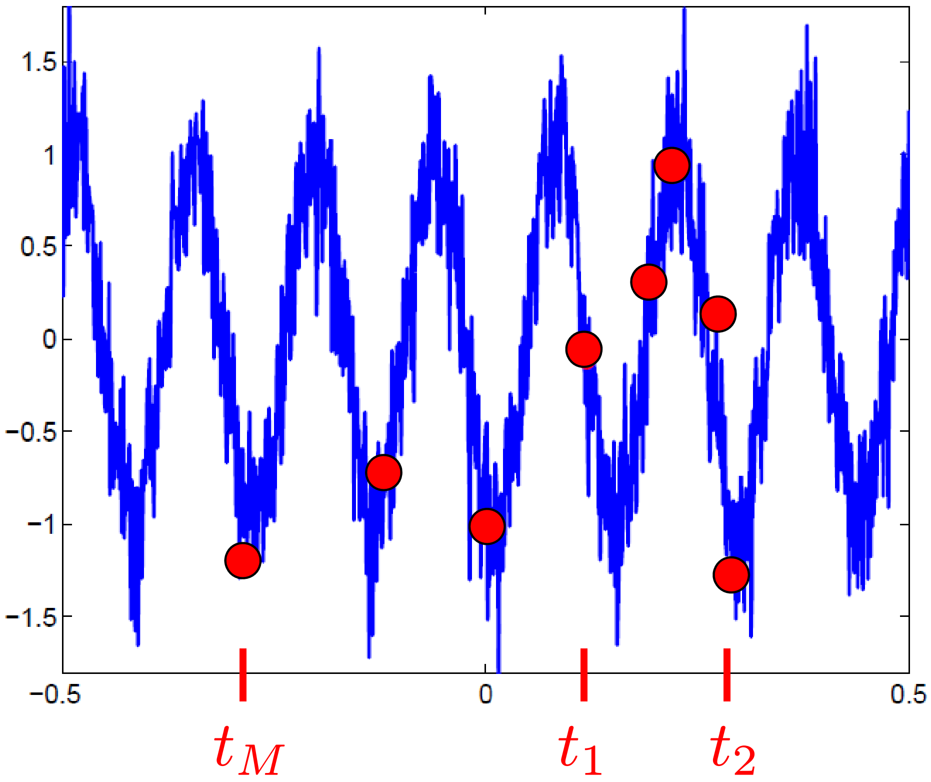
with probability at least 1-δ.

Extensions

- Arbitrary unknown amplitude + Gaussian noise

$$y = \begin{bmatrix} Ae^{j\omega_0 t_1} \\ Ae^{j\omega_0 t_2} \\ \vdots \\ Ae^{j\omega_0 t_M} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}$$

$$n_k \sim \mathcal{N}(0, \sigma_n^2)$$



Extension to Noisy Samples

- Observations

$$y = \begin{bmatrix} Ae^{j\omega_0 t_1} \\ Ae^{j\omega_0 t_2} \\ \vdots \\ Ae^{j\omega_0 t_M} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix}$$

- Random processes

$$\langle y, \psi_\omega \rangle = A \cdot X(\omega) + N(\omega)$$

- Bounds

$$\mathrm{E} \sup_{\omega \in \Omega} |A \cdot X(\omega) - E[A \cdot X(\omega)]| \leq C \cdot A \cdot \sqrt{M \log \Omega}$$

$$\mathrm{E} \sup_{\omega \in \Omega} |N(\omega)| \leq C \cdot \sigma_n \cdot \sqrt{M \log \Omega}$$

Estimation Accuracy

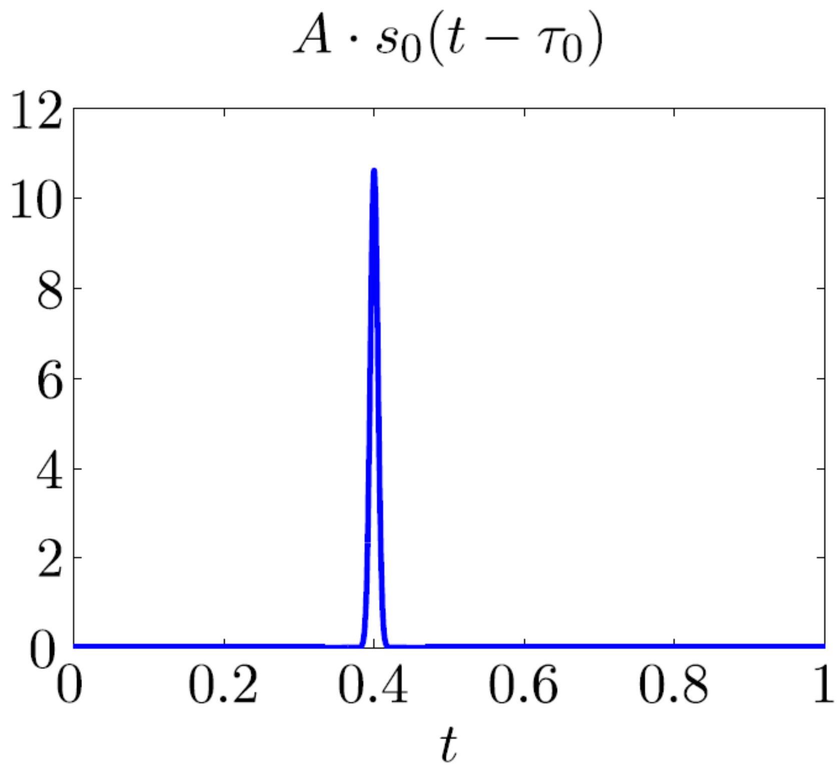
- If
$$M \geq C \cdot \max(\log(|\Omega||T|), \log(2/\delta)) \cdot \frac{\sigma_n^2}{|A|^2},$$
then with probability at least 1-2δ, the peak
$$\widehat{\omega}_0 = \arg\max_{\omega \in \Omega} |A \cdot X(\omega) + N(\omega)|$$
will have a guaranteed a accuracy of
$$|\omega_0 - \widehat{\omega}_0| \leq \frac{2\pi}{|T|}.$$

- The amplitude A can then be accurately estimated via least-squares.

Compressive Matched Filtering

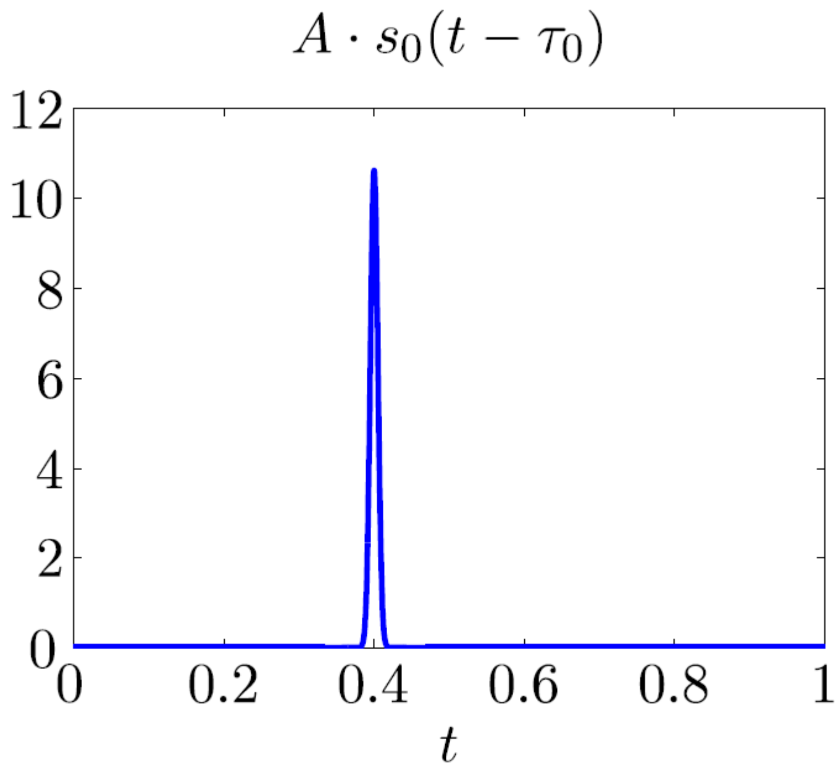
Exchanging Time and Frequency

- Known pulse template $s_0(t)$, unknown delay $\tau_o \in T$



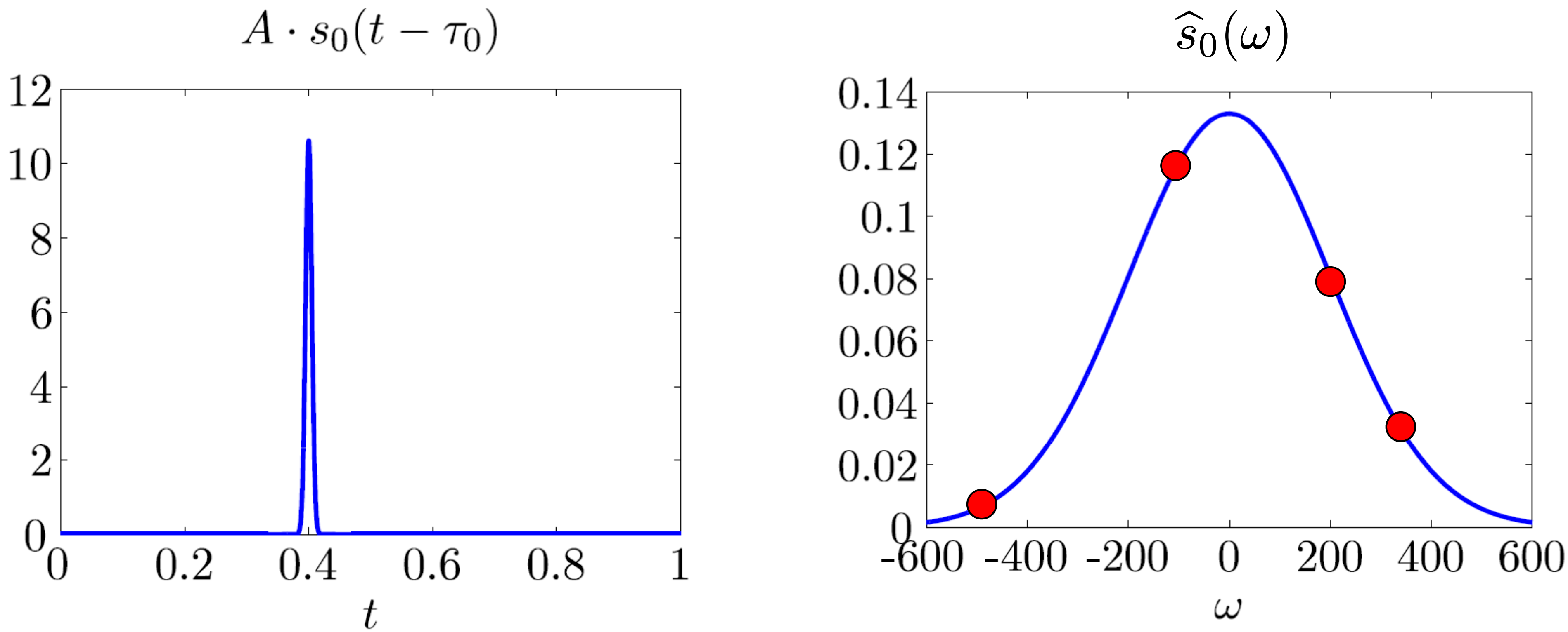
Exchanging Time and Frequency

- Known pulse template $s_0(t)$, unknown delay $\tau_o \in T$
- Random samples in frequency on Ω



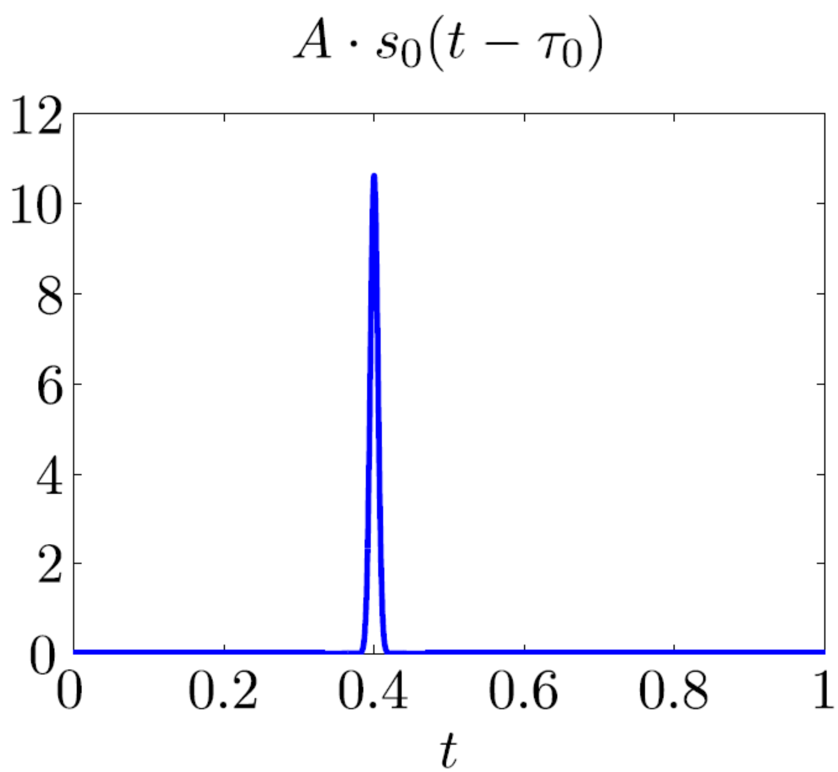
Exchanging Time and Frequency

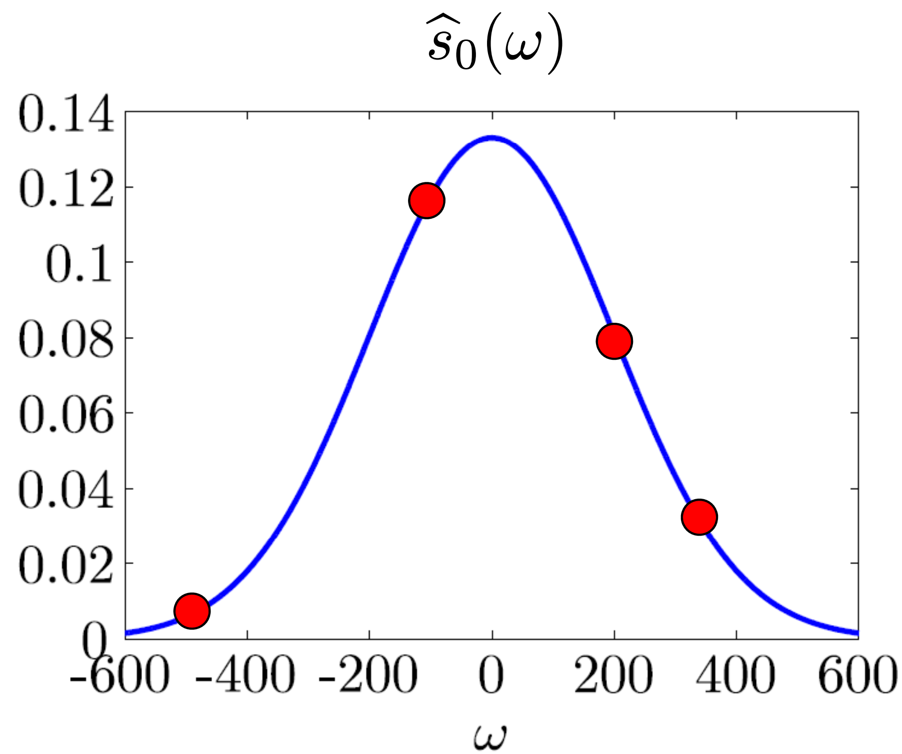
- Known pulse template $s_0(t)$, unknown delay $\tau_o \in T$
- Random samples in frequency on Ω



Exchanging Time and Frequency

- Known pulse template $s_0(t)$, unknown delay $\tau_o \in T$
- Random samples in frequency on Ω

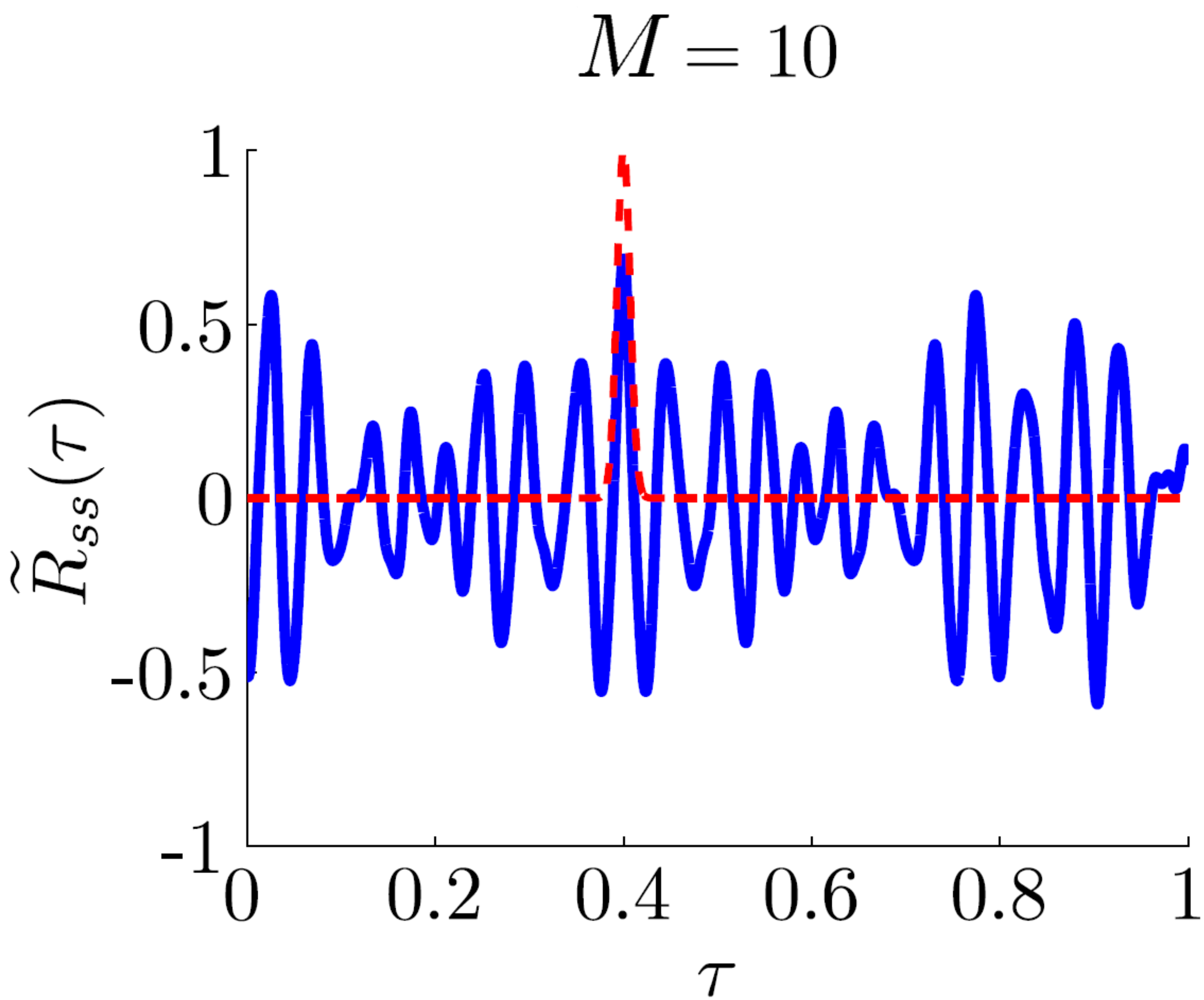




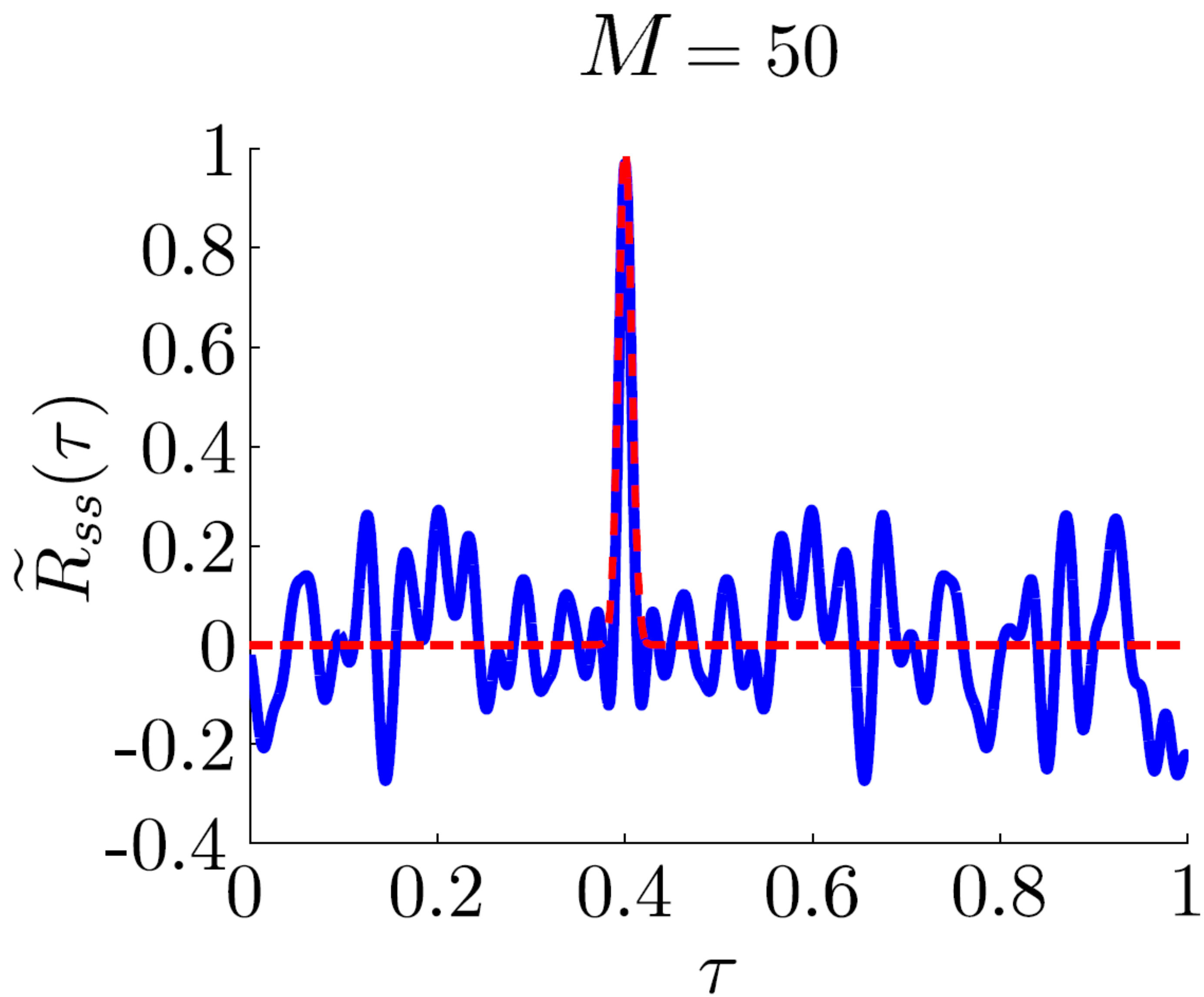
- Compute test statistics $X(\tau) = \langle y, \psi_\tau \rangle$ and let

$$\widehat{\tau}_0 = \arg\max_{\tau \in T} |X(\tau)|$$

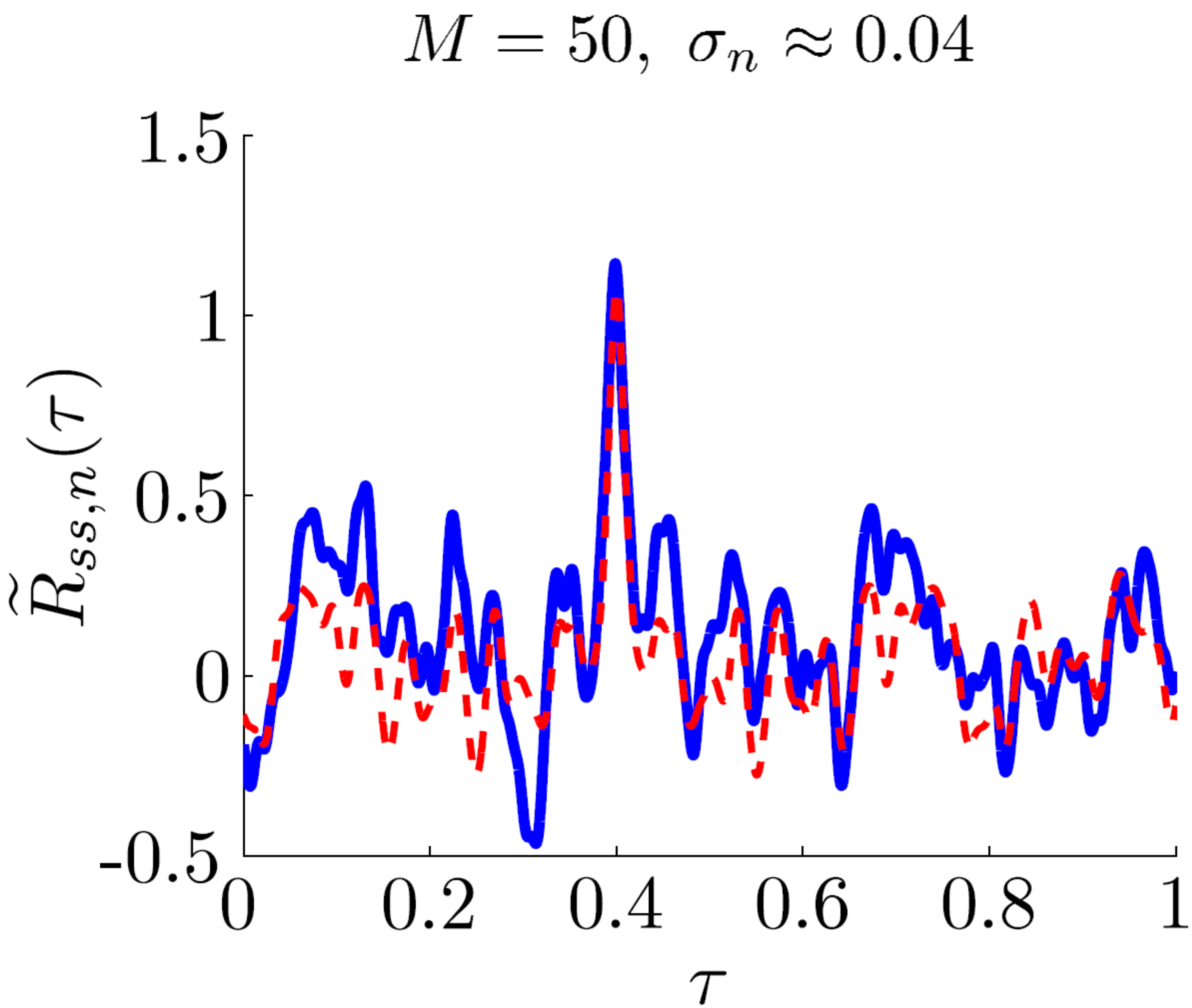
Experiment: Narrow Gaussian Pulse



Experiment: Narrow Gaussian Pulse



Experiment: Narrow Gaussian Pulse

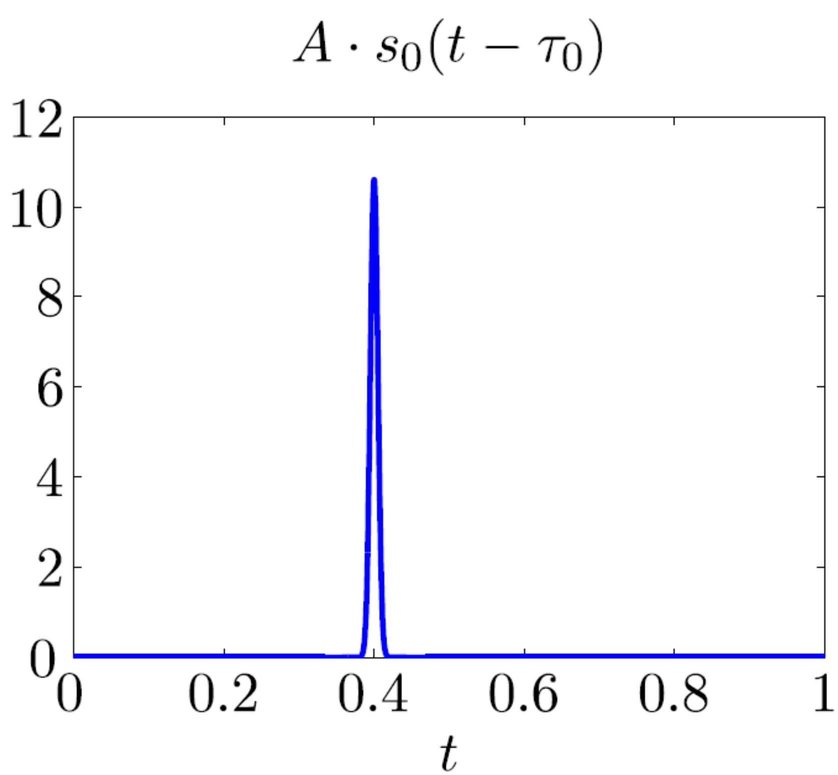


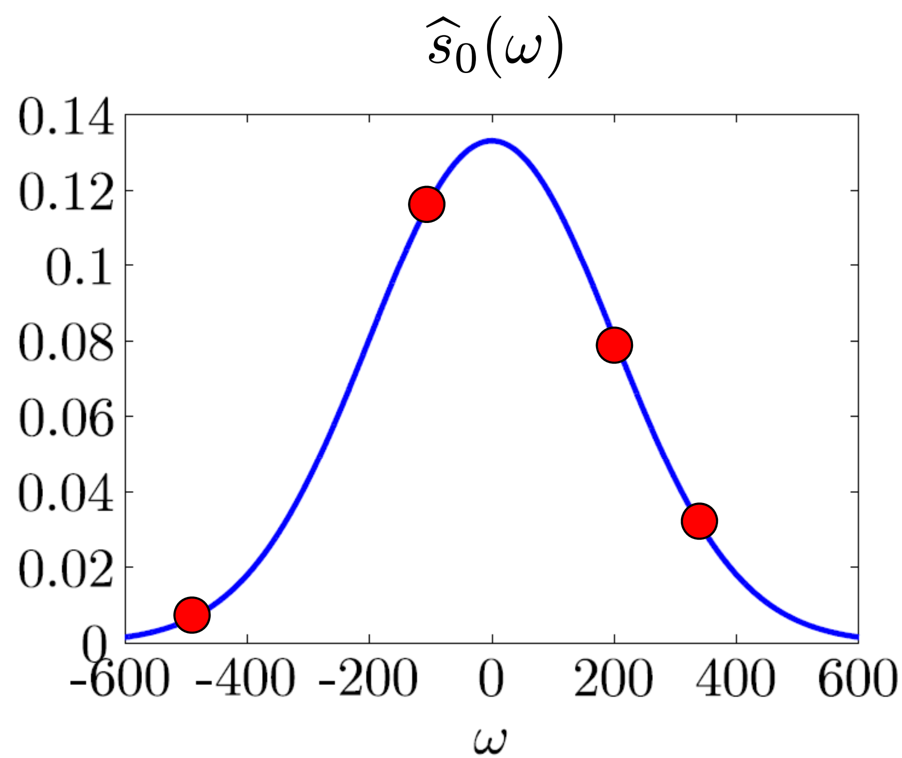
Matched Filter Guarantees

- Measurement bounds again scale with

$$\log(|\Omega||T|) \cdot \mathrm{SNR}^{-1}$$

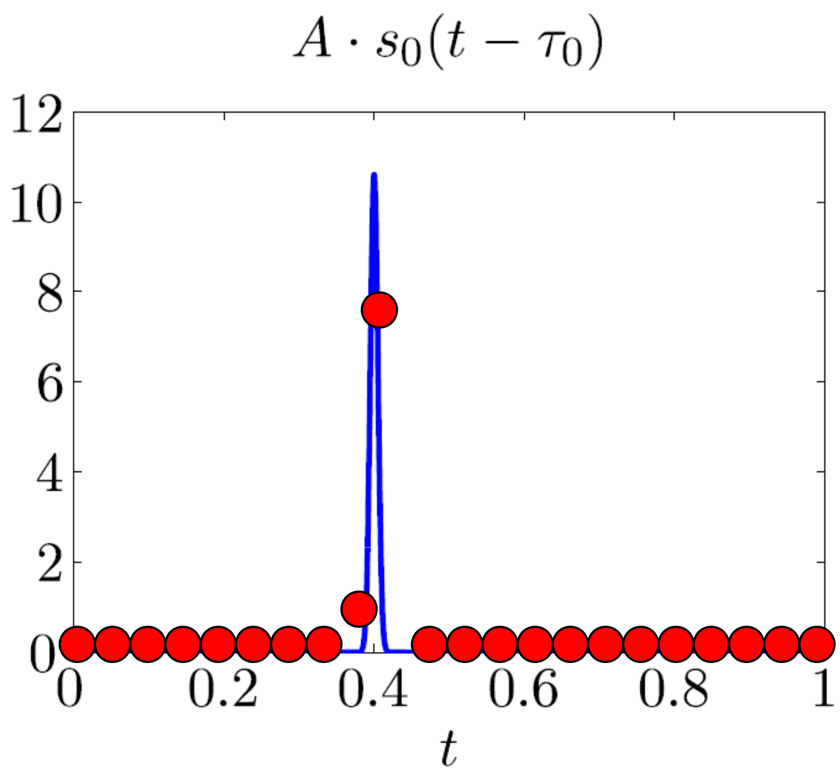
times a factor depending on uniformity of spectrum

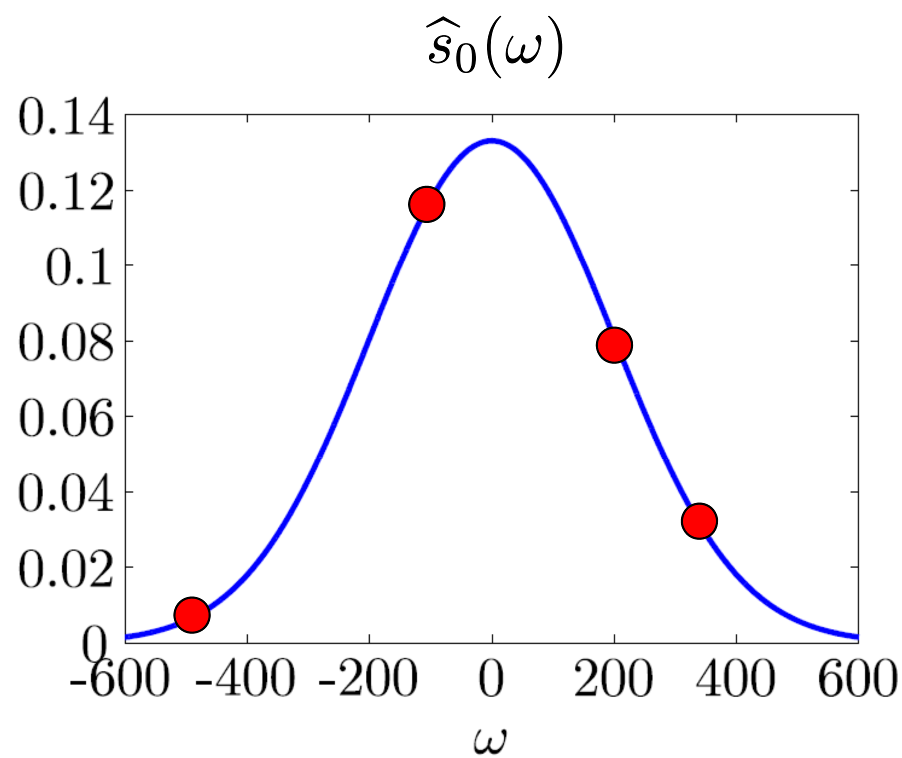




Interpreting the Guarantee

- When $M \sim \Omega$, the compressive matched filter is as robust to noise as traditional Nyquist sampling
- However, when noise is small this gives us a principled way to undersample without the risk of aliasing





Conclusions

- Random measurements
  - recover low-complexity signals
  - answer low-complexity questions

- Compressive matched filter
  - simple least squares estimation
  - analytical framework based on random processes
  - robust performance with sub-Nyquist measurements
  - measurement bounds agnostic to sparsity level
  - could incorporate into larger algorithm

- More is known about these problems
  - spectral compressive sensing [Duarte, Baraniuk]
  - delay estimation using unions of subspaces [Gedalyahu, Eldar]