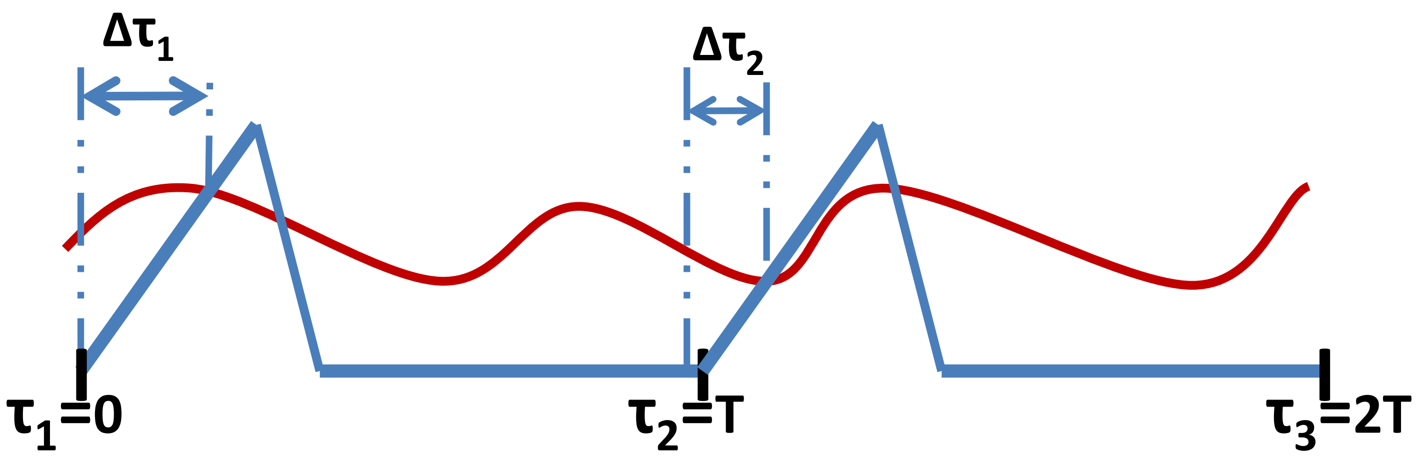
Rand PPM : A low power compressive sampling analog to digital converter

Praveen Yenduri

Anna Gilbert, Mike Flynn, Shahrzad Naraghi

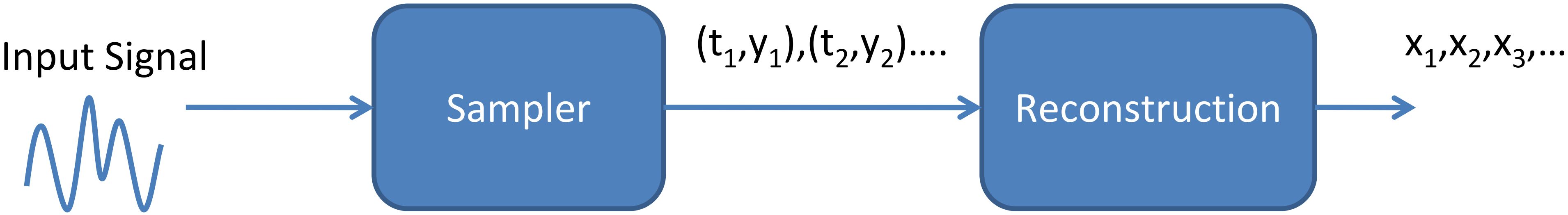
University of Michigan, Ann Arbor

The PPM (Pulse position modulation) ADC



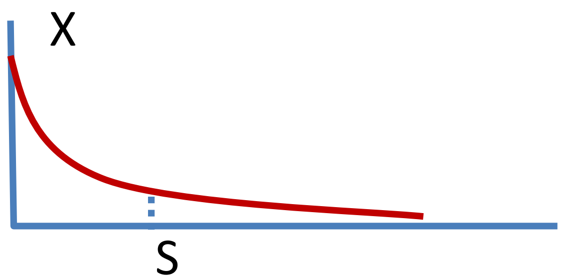
- Features:
  - A periodic reference ramp signal for sampling.
  - Employs a TDC for time delay digitization.
  - Low power : A 9-bit 14µW (2009) Naraghi et al.
  - Replacement of significant analog circuitry by digital components.
  - Non-uniform sampling
  - Reconstruction through a non-linear iterative technique which needs a sampling rate of about 1.2 times the Nyquist.

A Compressed Sensing design

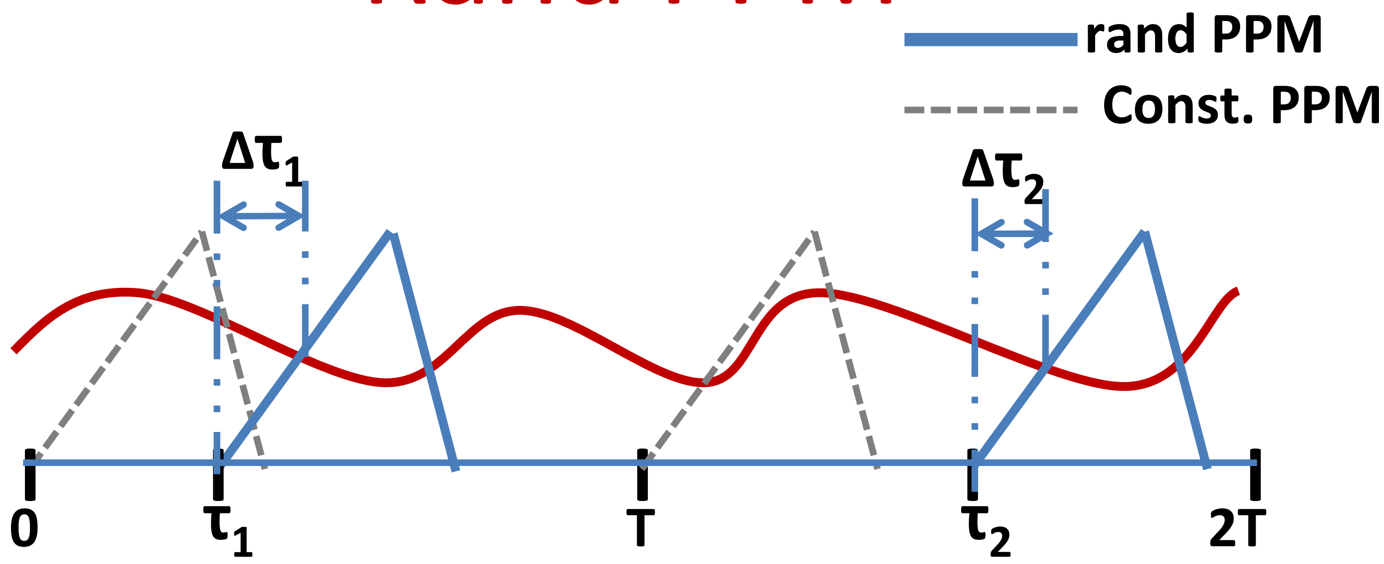


- A Low power Sub-Nyquist Sampler in time domain.
- Fast Reconstruction of the signal in Frequency domain.

- Condition on Input signal :
  - Sparse in the Frequency domain.
    S-sparse => S non-zero frequency coefficients
  - Compressible in the Frequency domain.



Rand PPM



- Features:
  - Ramp starting time in each period is a random variable.
  - Low power TDC with digital components.
  - Non-uniform random sampling
  - Reconstruction through compressive sampling algorithms.
  - Sampling rate is sub-Nyquist.
  - Assumes an s-sparse or compressible input signal.

Measurement Matrix

- Let f be the input signal in time domain.
- In compressed sensing, form $\Phi f = y$.

  Φ => Random linear measurements.
- f is sparse in some Ortho-Normal Basis Ψ.

  $f = \Psi x$, x => sparse. Hence $\Phi\Psi x = y$.
- For random on-grid sampling in time, Φ =

$$\begin{bmatrix} 1\,0\,0\,0 \ldots\ldots\ldots 0\,0 \\ 0\,0\,1\,0 \ldots\ldots\ldots 0\,0 \\ \vdots \\ 0\,0\,0\,0 \ldots\ldots\ldots 1\,0 \end{bmatrix}$$

- Ψ = IDFT Matrix.
- Net Measurement matrix $B = \Phi\Psi$ => sub-IDFT.
- For PPM ADC, Φ is an interpolation matrix.
- Hence, B is not a sub-IDFT matrix.

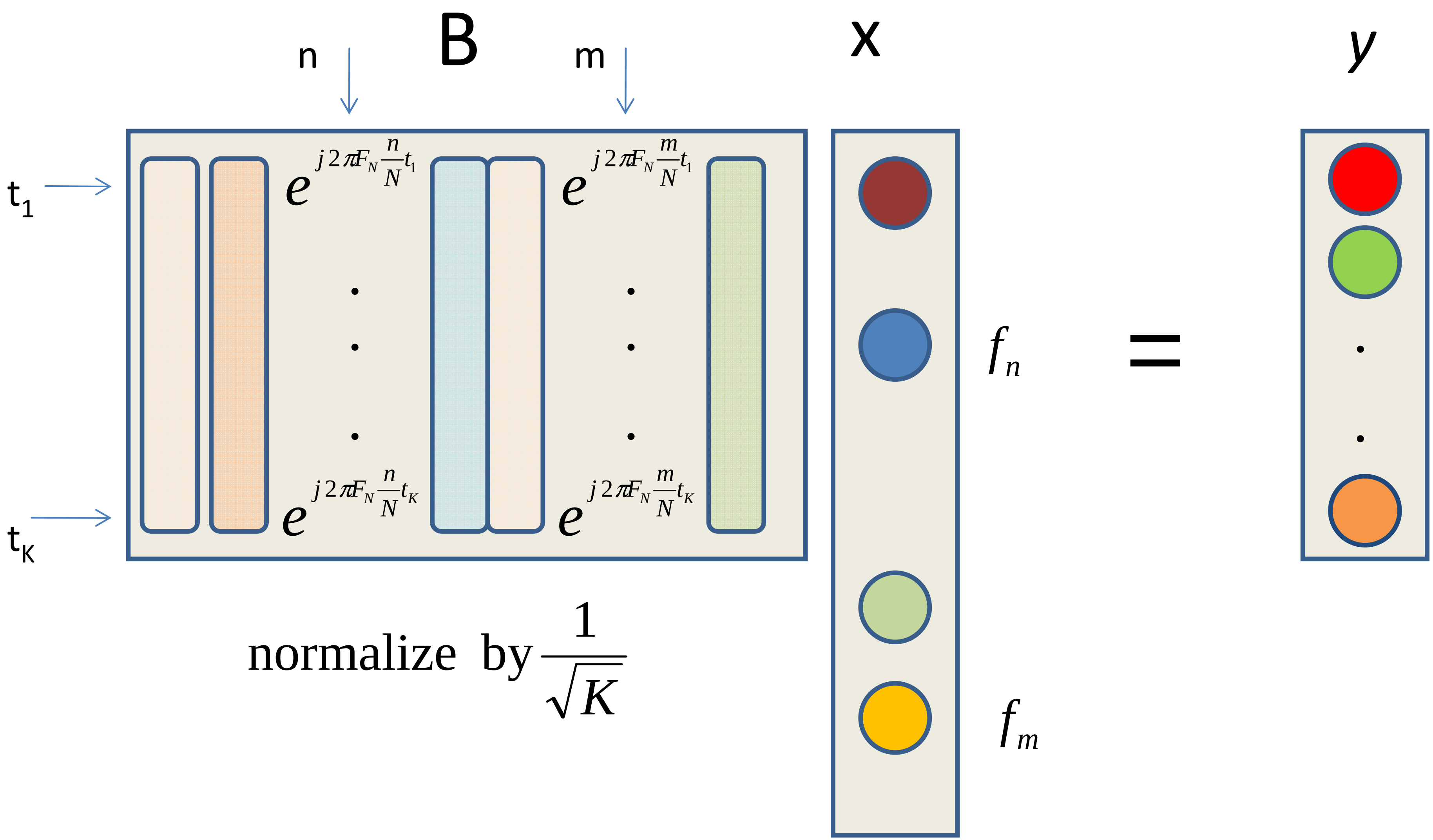
Measurement Matrix

- Let the input signal *f* have a maximum frequency $f_{max} < F_N/2$.
- F_N => Nyquist frequency.
- K => Number of measurements acquired by the PPM Sampler.
- N => Number of measurements if sampled at Nyquist rate, N > K.
- Reconstruction frequencies $f_n = F_N(n/N)$, $n = [-N/2 : N/2 - 1]$.

$$\delta(f - f_0) \Leftrightarrow e^{j2\pi f_0 t}$$

$$\begin{bmatrix} e^{j2\pi f_0 t_1} \\ e^{j2\pi f_0 t_2} \\ . \\ . \\ . \\ e^{j2\pi f_0 t_K} \end{bmatrix}$$

Measurement Matrix



Reconstruction

ALGORITHM 1 :

Signal Model :

Any s-sparse (compressible) signal

input N, sparsity parameter $S, (t_i, y_i), i = 1,2.., K$

output $\underline{\tilde{x}}$ (Signal in frequency domain, length N)

Frequency Identification

For Least Squares : Richardson iteration with Acceleration

Run time : O(SK log(2/δ))

δ => error tolerance in soln.

for $i = 1,2,3...$

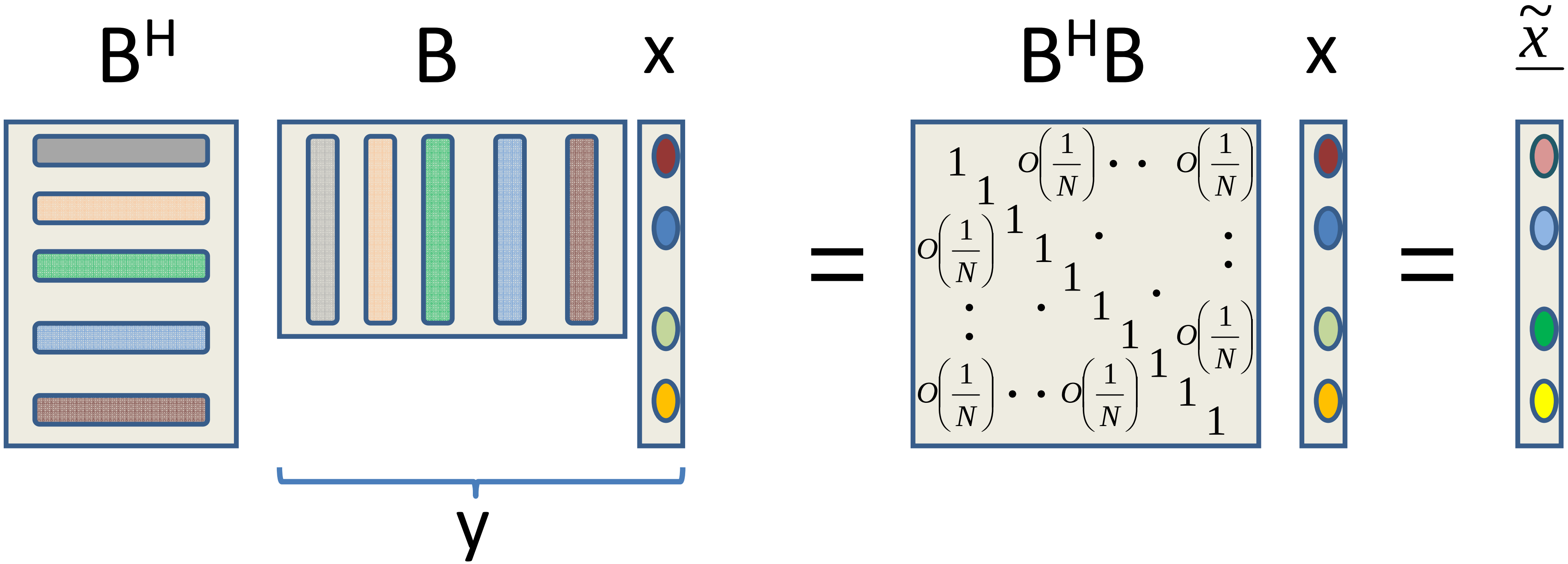
Coefficient estimation and correction

For $B^H r$: Inverse NUFFT (steidl) with Cardinal B-spline Interpolation.

Run time : O(N log N) + O(K)

until $\| \underline{r}_{i+1} \|_2$ is small enough.

Algorithm 1 : Analysis



$$\text{with } K = \mathrm{O}\left(\frac{s}{\varepsilon^2}\right), \qquad \mathrm{E}(\tilde{x}_m) = x_m + \mathrm{O}\left(\frac{1}{N}\right) \| \underline{x} \|_1$$

$$\mathrm{Var}(\tilde{x}_m) \leq \frac{\varepsilon^2}{s} \| \underline{x} \|_2^2$$

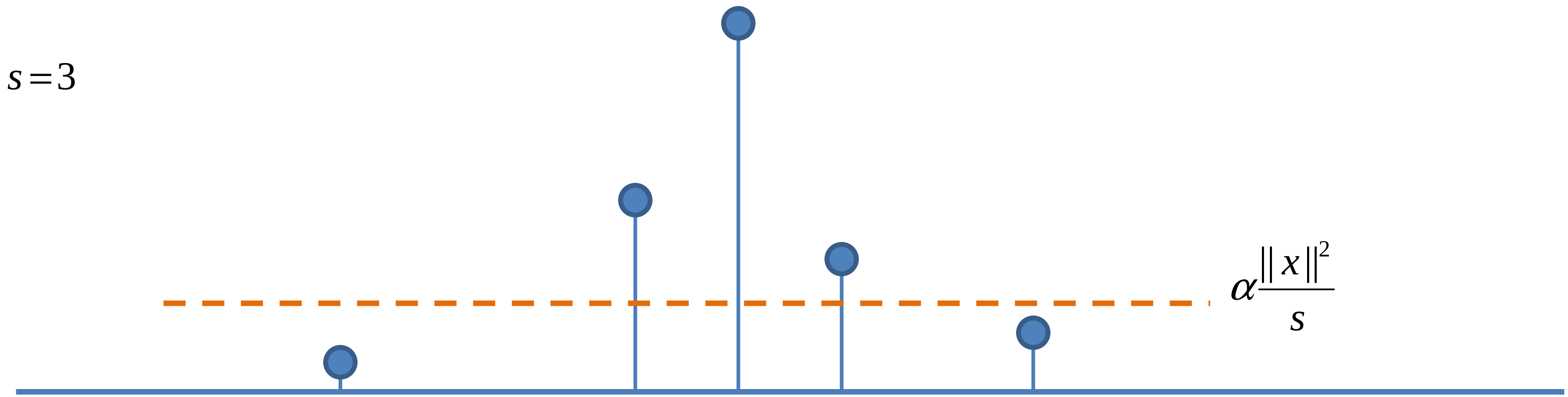
Algorithm 1 : Analysis

- **Theorem :** $y = Bx + \xi$, Given $K = O\left(\frac{s}{\varepsilon^2}\right)$, with probability $1 - O(\varepsilon^2)$,

Runtime $= O(IN \log N)$,

$I =$ # iterations

$$\| x - \tilde{x} \|_2^2 \le \frac{\| x - x_{(s)} \|_2^2}{1-\alpha} + \frac{c(B) \| \xi \|_2^2}{1-\alpha}$$



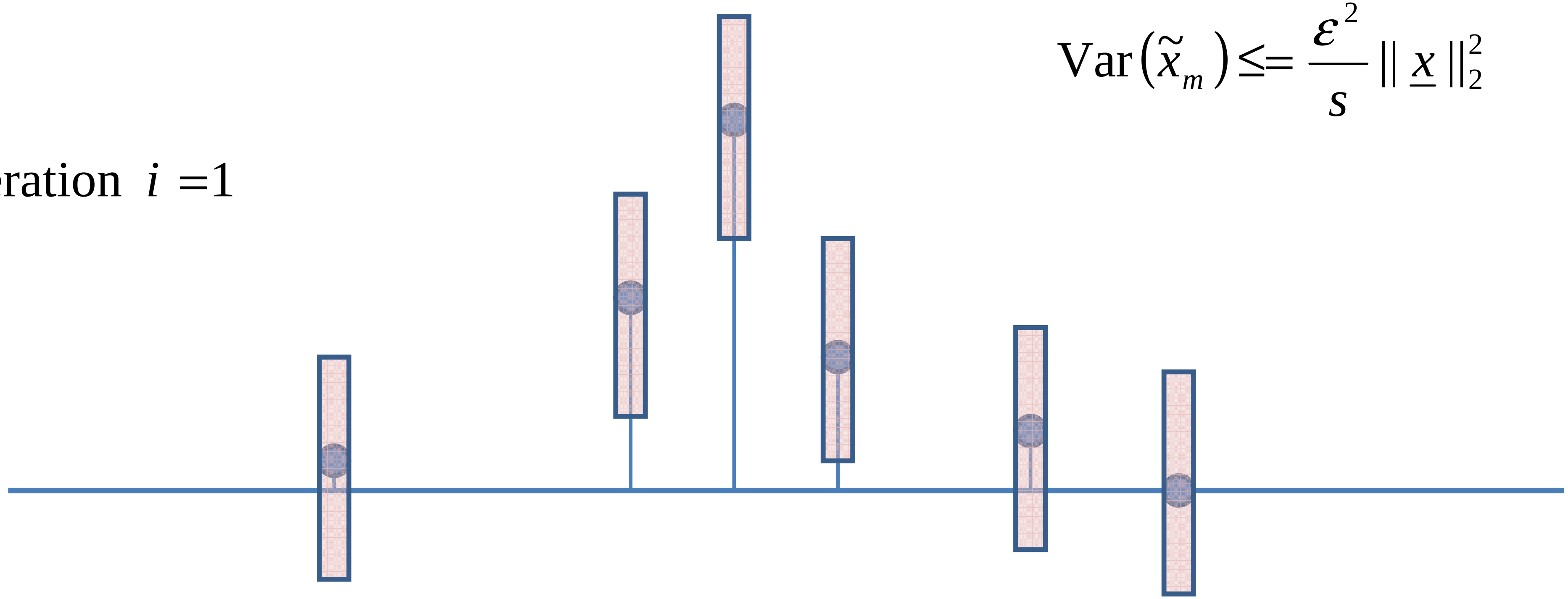
Algorithm 1 : Analysis

- Theorem : $y = Bx + \xi$, Given $K = O\left(\frac{s}{\varepsilon^2}\right)$, with probability $1 - O\left(\varepsilon^2\right)$,

$$\| x - \tilde{x} \|_2^2 \leq \frac{\| x - x_{(s)} \|_2^2}{1-\alpha} + \frac{c(B) \| \xi \|_2^2}{1-\alpha}$$

$$\mathrm{Var}\left(\tilde{x}_m\right) \leq= \frac{\varepsilon^2}{s} \| \underline{x} \|_2^2$$

iteration $i = 1$



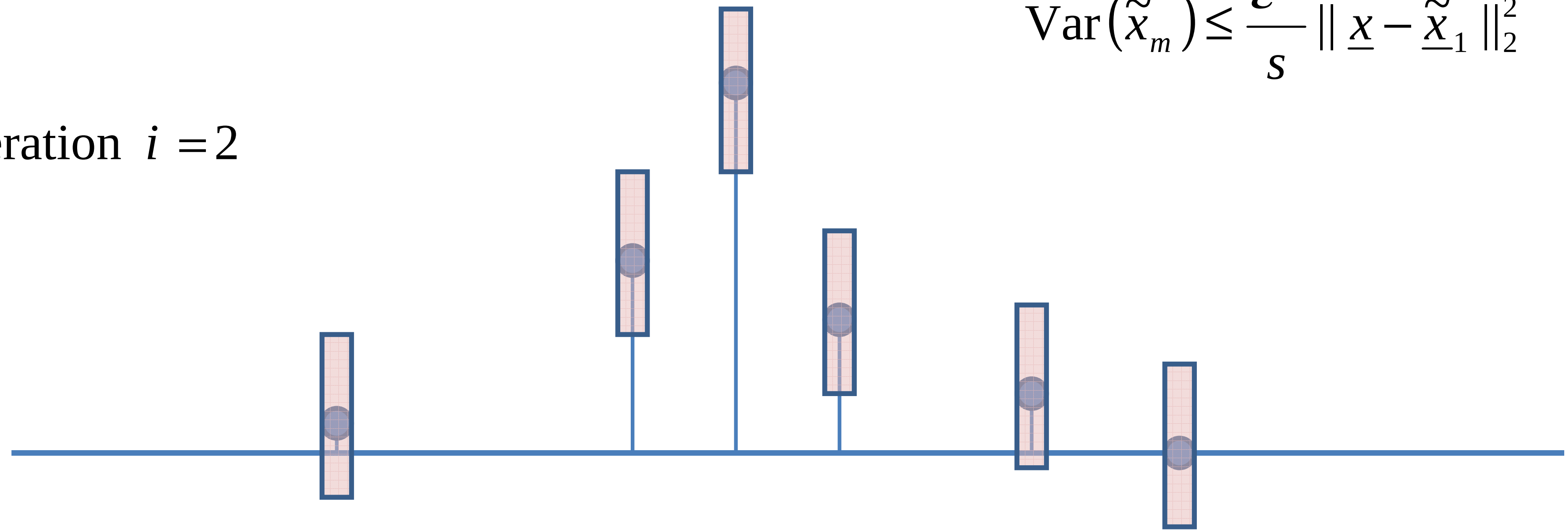
Algorithm 1 : Analysis

- **Theorem :** $y = Bx + \xi$, Given $K = O\left(\dfrac{s}{\varepsilon^2}\right)$, with probability $1 - O\left(\varepsilon^2\right)$,

$$\| x - \tilde{x} \|_2^2 \leq \frac{\| x - x_{(s)} \|_2^2}{1-\alpha} + \frac{c(B) \| \xi \|_2^2}{1-\alpha}$$

$$\mathrm{Var}\left(\tilde{x}_m\right) \leq \frac{\varepsilon^2}{s} \| \underline{x} - \underline{\tilde{x}}_1 \|_2^2$$

iteration $i = 2$



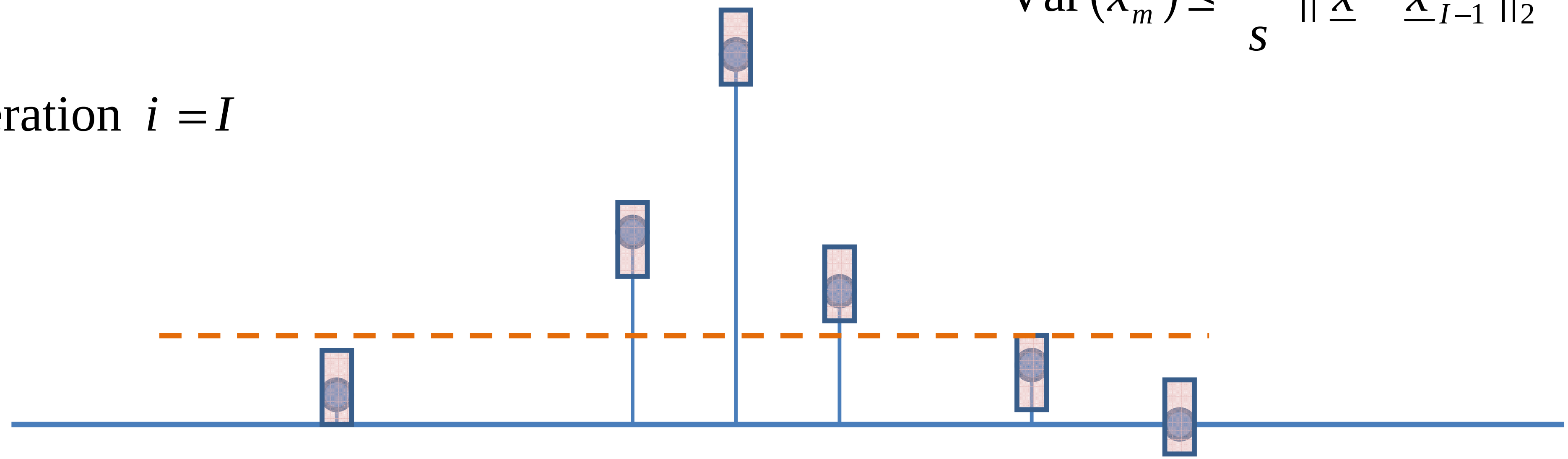
Algorithm 1 : Analysis

- **Theorem :** $y = Bx + \xi$, Given $K = \mathrm{O}\left(\frac{s}{\varepsilon^2}\right)$, with probability $1 - \mathrm{O}\left(\varepsilon^2\right)$,

$$\| x - \tilde{x} \|_2^2 \leq \frac{\| x - x_{(s)} \|_2^2}{1-\alpha} + \frac{c(B) \| \xi \|_2^2}{1-\alpha}$$

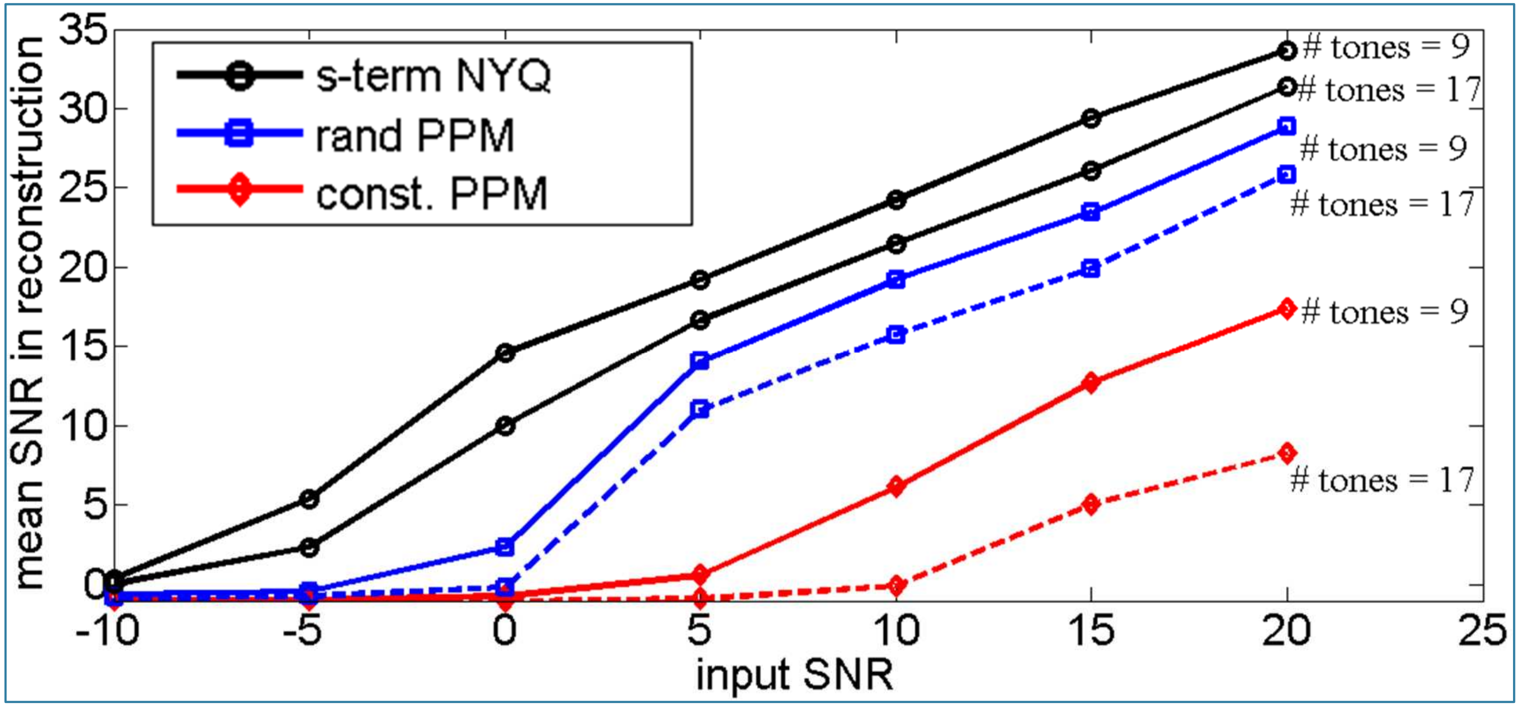
$$\mathrm{Var}\left(\tilde{x}_m\right) \leq \frac{\varepsilon^2}{s} \| \underline{x} - \underline{\tilde{x}}_{I-1} \|_2^2$$

iteration $i = I$



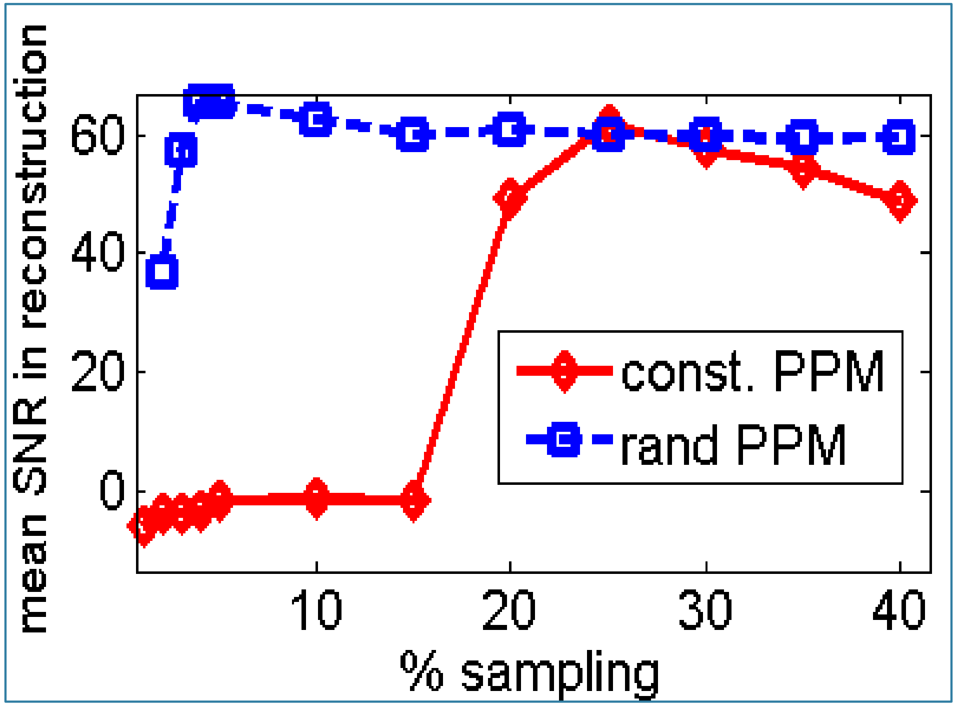
Results : Reconstructing Multitone signals with algorithm 1

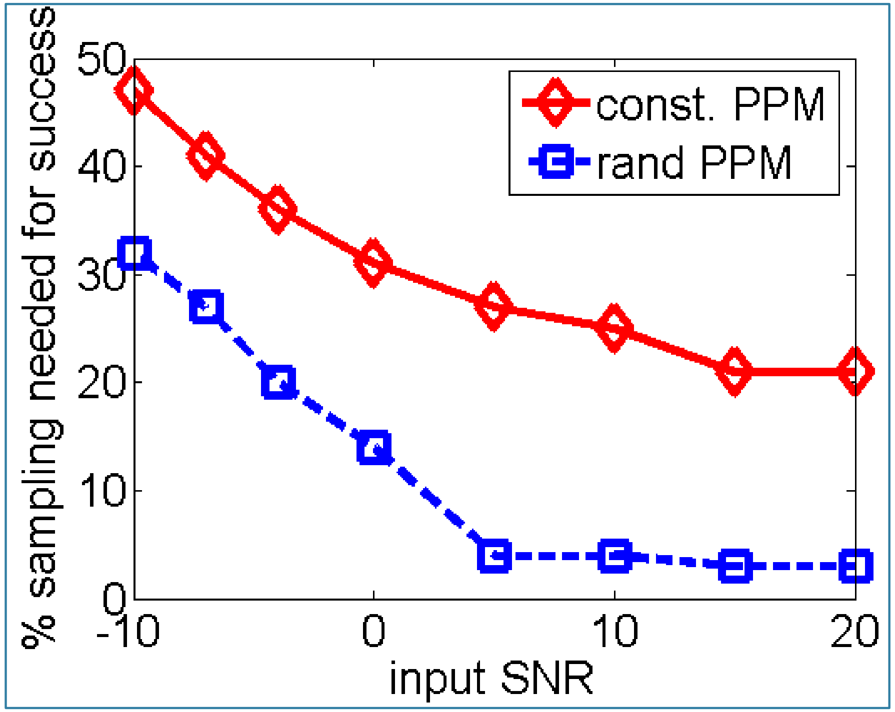
- Input signal is a combination of J sinusoids with frequencies randomly chosen from the Nyquist grid. Sparsity S = 2*J.
- The tones have comparable amplitudes and random phases.
- Sampling frequency = 1 MHz, Nyquist Frequency = 3 MHz.
- S-term Nyquist benchmark : the input signal is sampled at Nyquist rate at the same quantization level of the PPM, and then truncated in frequency domain to retain only the dominant s frequencies



Percentage sampling (compression) results for algorithm 1

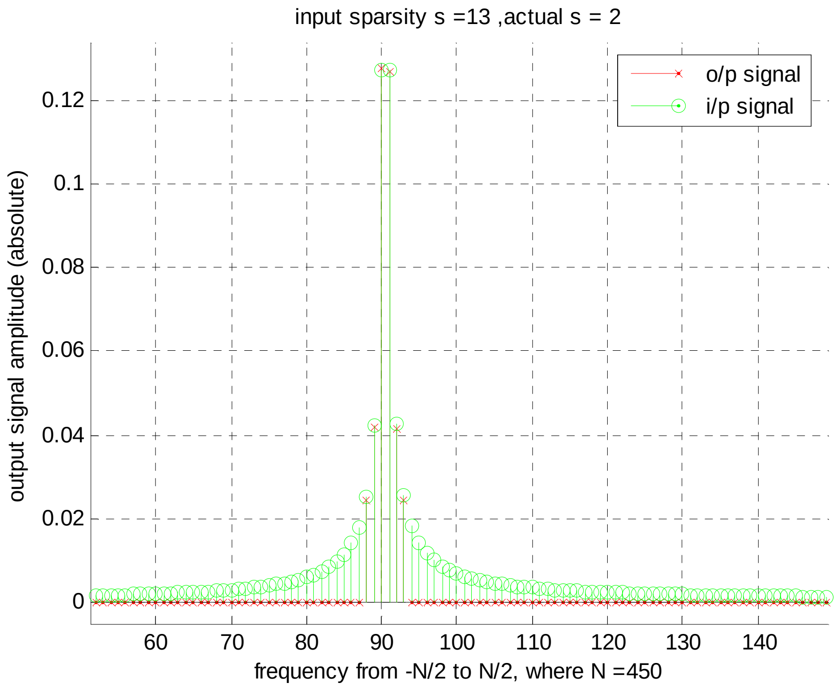
- J = 1 tone, placed at 1.49MHz. Nyquist Frequency = 3 MHz.
- Percentage sampling = 100*K/N
- Success : Output SNR > rect (input SNR) + 10 dB, rect(a) = 0.5*(a+|a|)

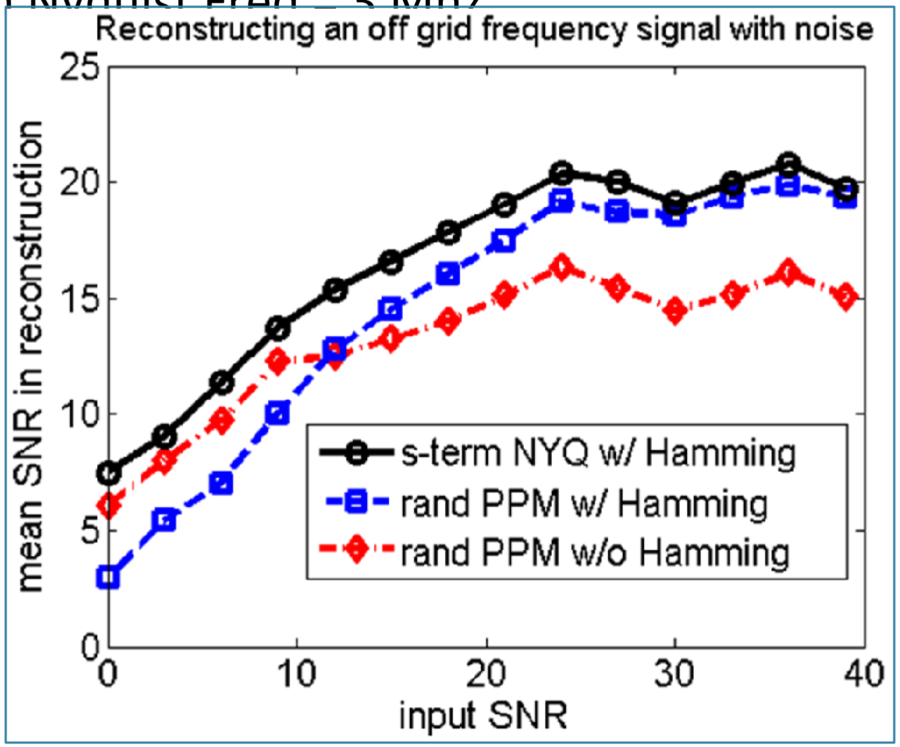




Reconstructing an Off-grid Frequency signal with algorithm 1

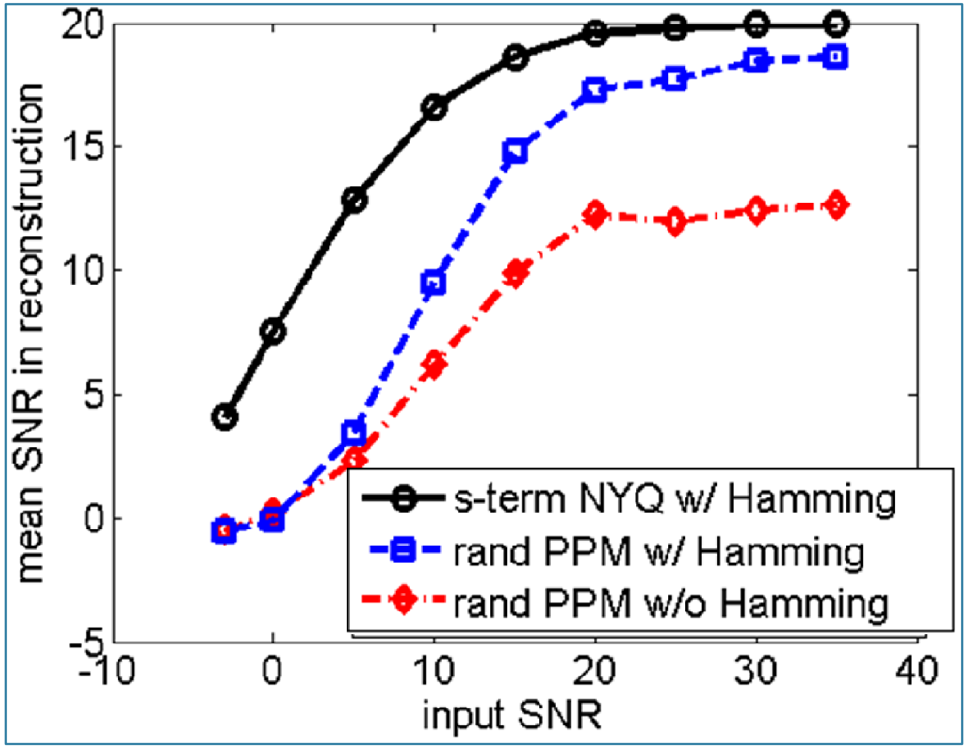
- Reconstructing a two tone signal with the frequencies chosen randomly and off the Nyquist. grid.
- Off-grid frequency causes spectral spread or leakage => adversely affects sparsity
- Reduce spectral leakage through windowing (post sampling , before reconstruction)
- Sampling Freq = 1 MHz, Reconstruction Nyquist Freq = 3 Mhz





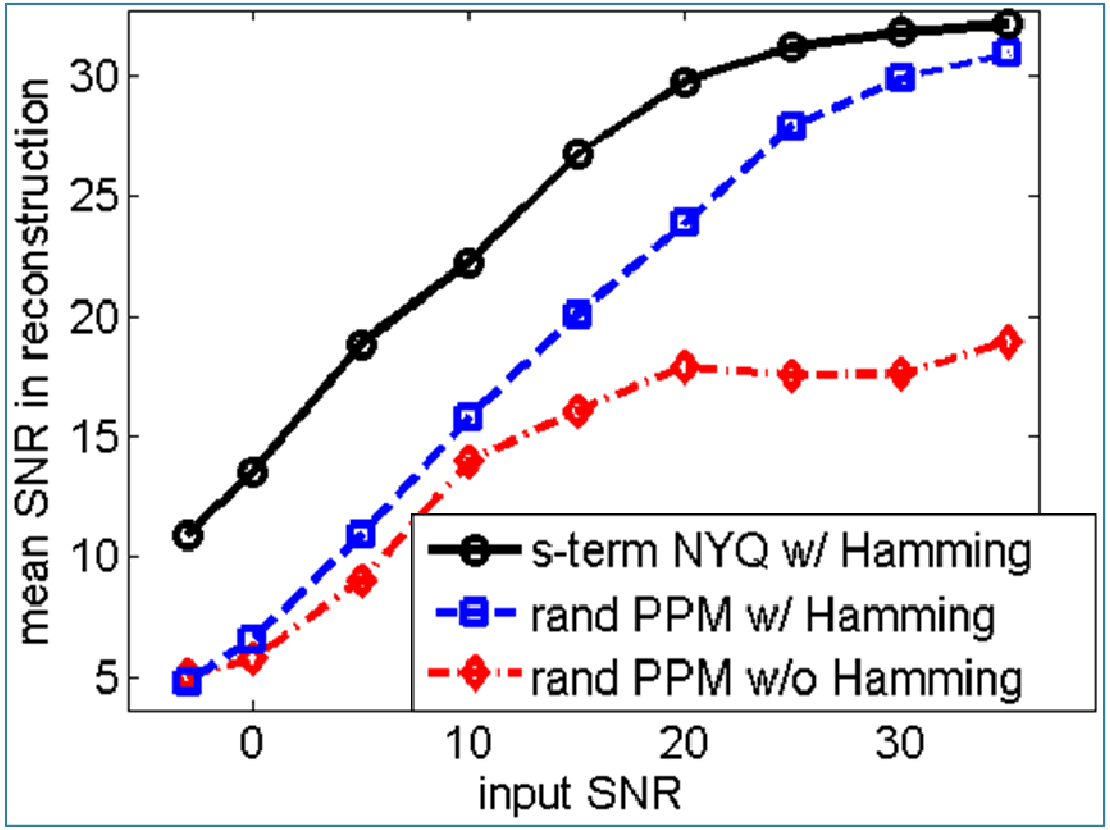
FM signal reconstruction with algorithm 1

- FM signal = $A \cos(2\pi f_c t + \beta \sin(2\pi f_m t))$; Approximate Analytical BW = $2*(\beta+1)*f_m$
- Both Carrier Frequency and message frequency are randomly chosen to be off the Nyquist Grid. $\beta = 2$
- Sampling Freq = 1 MHz, Reconstruction Nyquist Freq = 3 Mhz.
- The sparsity parameter input to the reconstruction algorithm is s = 36 .
- The signal is windowed, reconstructed and demodulated before calculating the output SNR.



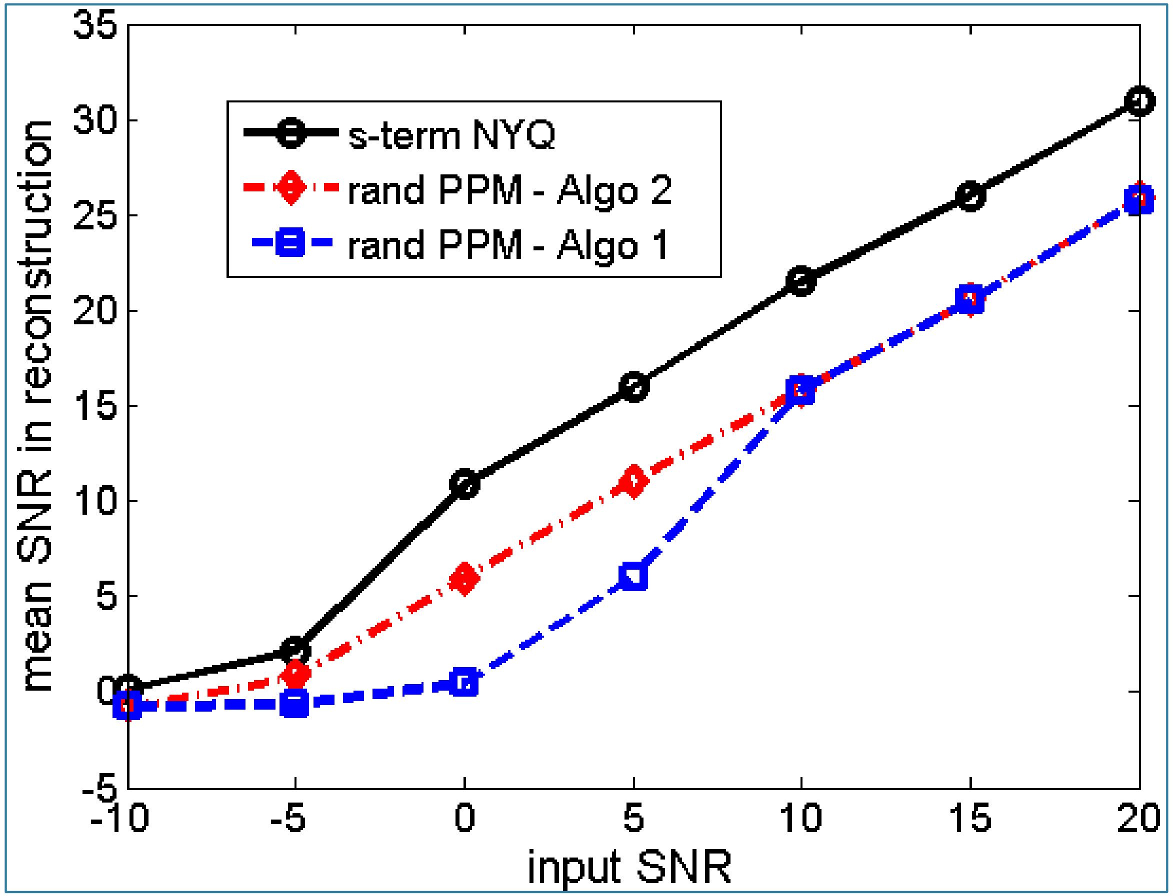
AM signal reconstruction with algorithm 1

- AM signal = A (1+ka*m(t)) cos($2\pi f_c t$); m(t) => message, ka => to keep the envelope of the signal positive.
- AM signal with random off-grid carrier frequency and sawtooth message as modulating signal.
- Sampling Freq = 1 MHz, Reconstruction Nyquist Freq = 3 Mhz.



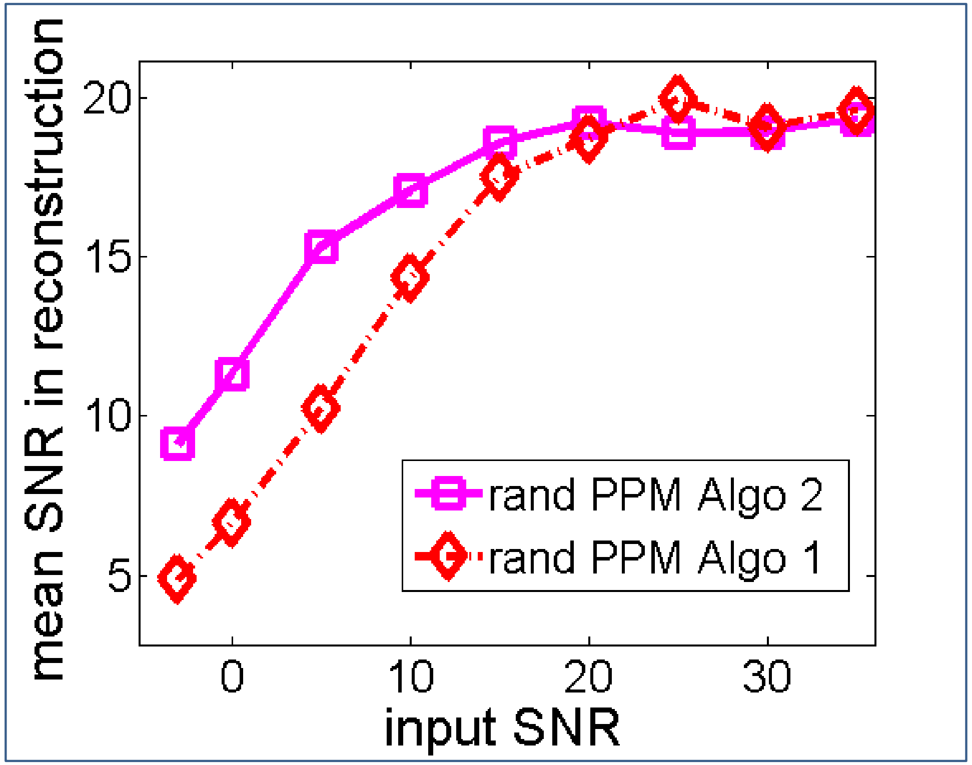
Results : Reconstructing multi-tone signals with algorithm 2

- Input signal is a combination of J sinusoids with frequencies randomly chosen from the Nyquist grid. Sparsity S = 2*J.
- The tones have comparable amplitudes and random phases.
- Sampling frequency = 1 MHz, Nyquist Frequency = 3 MHz.
- Number of blocks m = 5.



AM signal reconstruction

- AM signal = A (1+ka*m(t)) cos(2πf_ct); m(t) => message, ka => to keep the envelope of the signal positive.
- AM signal with random off-grid carrier frequency and sawtooth message as modulating signal.
- Sampling Freq = 1 MHz, Reconstruction Nyquist Freq = 3 Mhz.
- Number of blocks m = 5.
- The plot below compares the reconstruction of First block.



Conclusions

- The Rand PPM design exhibits much better performance than the conventional PPM design in more than one way :
  - better and closer to benchmark performance in terms of the output SNR.
  - Rand PPM can handle less sparse signals better than Const. PPM.
  - can handle signals with off-grid frequencies.
  - enables the reductions of sampling rate to levels much below Nyquist.
  - enables the use of reconstruction algorithms 1 & 2 that are faster and also feasible for a hardware implementation.
  - Algo 1 can be used with Hamming window to boost the performance when the input SNR levels are high enough.
  - Algo 2 can be used to improve computational efficiency whenever the input signal satisfies the required assumptions.

Thank you !!

Reconstruction

- Basis Pursuit (L_1 Minimization) :

$$\underset{\underline{x}}{Min} \| \underline{x} \|_1 \quad s.t. \quad \underline{y} = B\underline{x}$$

or

$$\underset{\underline{x}}{Min} \| \underline{x} \|_1 \quad s.t. \quad \| \underline{y} - B\underline{x} \|_2 < \varepsilon$$

- Limitations :
- B needs a lot of Rows ($O(K\log^4 N)$) to satisfy RIP properties.
- Relatively slow reconstruction, relatively difficult for FPGA implementation.

- Greedy Matching Pursuit Algorithms :

$$\underset{\underline{z}}{Min} \| \underline{x} - \underline{z} \|_1 \quad \text{subject to} \quad \| z \|_0 \leq s$$

- CoSaMP : B needs to be an RIP matrix with $\delta_{2s} < 0.1$
- IT : Assumes that $\| B \|_2 < 1$

- Rand PPM ADC measurement matrix B does not need to satisfy any of the above properties !

=> Matching pursuit with different theoretical analysis.

Algorithm 1 : Analysis

- Theorem :

Let $y = Bx + \xi$, where $\xi = \text{error}$. Suppose $|_s x|^2 \geq \alpha \frac{\|x\|^2}{s} \geq |_{(s+1)} x|^2$

where $|_i x|$ is the magnitude of i^{th} largest term in x. Given $K = O\left(\frac{s}{\varepsilon^2}\right)$,

with probability $1 - O\left(\varepsilon^2\right)$, the s-term estimate $\tilde{x}$ satisfies,

$$\|x - \tilde{x}\|_2^2 \leq \frac{\|x - x_{(s)}\|_2^2}{1-\alpha} + \frac{c(B)\|\xi\|_2^2}{1-\alpha}.$$

$x_{(s)}$ => Best s-term approximation to x.

Runtime $= O(IN \log N)$, $I = \#$ iterations.

Reconstruction

ALGORITHM 2 : Signal Model :

s-sparse (compressible) in each block, support set of top s coefficients constant over m blocks, but the coefficients can vary.

input $N, s, K,$ samples over m blocks

$(\underline{t}(i), \underline{y}(i)), i = 1,2.., m$

output $\underline{\tilde{x}}(i)$ (length $N, i = 1,.., m$)

for $j = 1,..N$

$\hat{x}_j = \text{median } \{|\underline{B}(1)_j^H \underline{y}(1)|,..,|\underline{B}(m)_j^H \underline{y}(m)|\}$

$T = \text{supp}\{ [\underline{\hat{x}}]_S \}$

for $i = 1,..m$

$[\underline{\tilde{x}}(i)]_T = (B(i)_T^H B(i)_T)^{-1} B(i)_T^H \underline{y}(\text{i})$

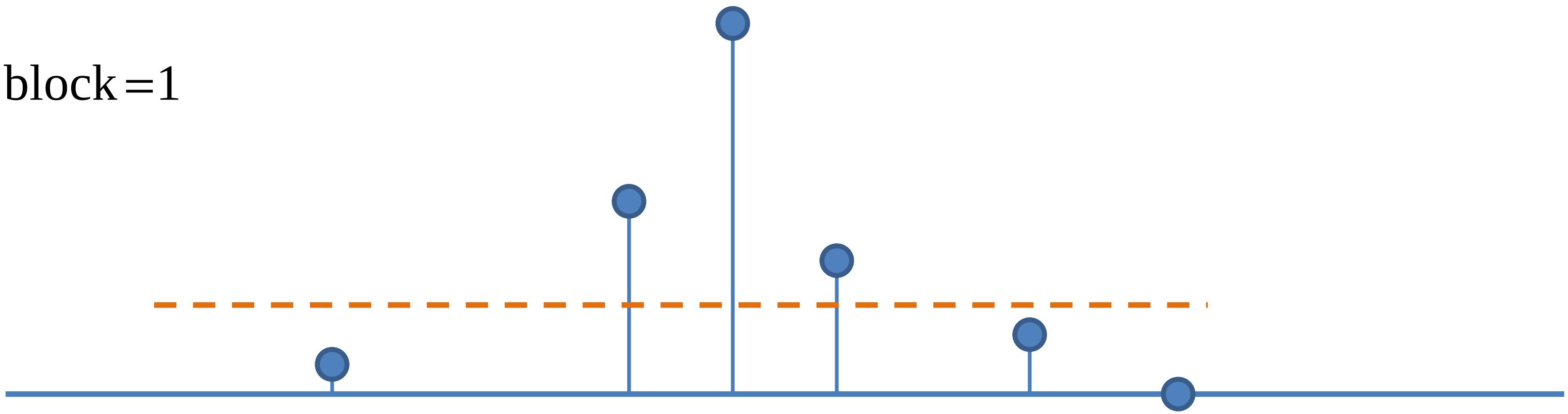
Inverse NUFFT (steidl) with Cardinal B-spline Interpolation.
Run time : O(N log N) + O(K)

Richardson iteration with Acceleration
Run time : O(SK log(2/δ))
δ => error tolerance in soln.

Algorithm 2 : Analysis

- **Theorem :** Given $m = O\left[\ln\left(\frac{N}{\delta}\right)\right]$, with probability $1-\delta$, the top s frequencies are correctly identified with a runtime of $O(N \log N)$ per block.

$\text{block} = 1$



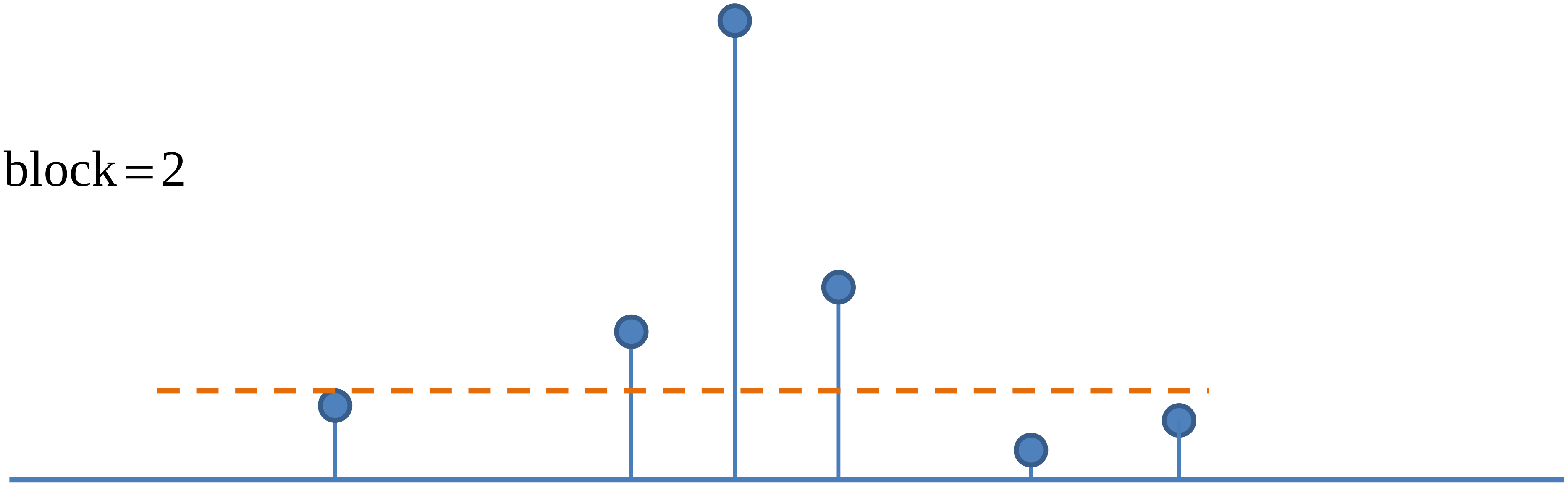
Algorithm 2 : Analysis

- **Theorem :** Given $m = O\left[\ln\left(\frac{N}{\delta}\right)\right]$, with probability $1-\delta$,

  the top s frequencies are correctly identified.

  with a runtime of $O(N \log N)$ per block.

$\text{block} = 2$



Algorithm 2 : Analysis

- **Theorem :** Given $m = O\left[\ln\left(\frac{N}{\delta}\right)\right]$, with probability $1-\delta$, the top s frequencies are correctly identified with a runtime of $O(N \log N)$ per block.

$\text{iteration} = 1$