Universal Denoising and Approximate Message Passing

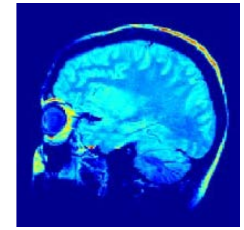
Dror Baron
North Carolina State University

Joint w/ Yanting Ma and Junan Zhu

Information Theory and Applications Workshop
San Diego, CA; February 2015

Supported by NSF CCF-1217749 and ARO W911NF-14-1-0314

Applications of Linear Inverse Problems
(especially compressed sensing)



- Medical imaging (tomography)

- Source and channel coding

- Seismic imaging (oil industry)

- Many more...

Motivation

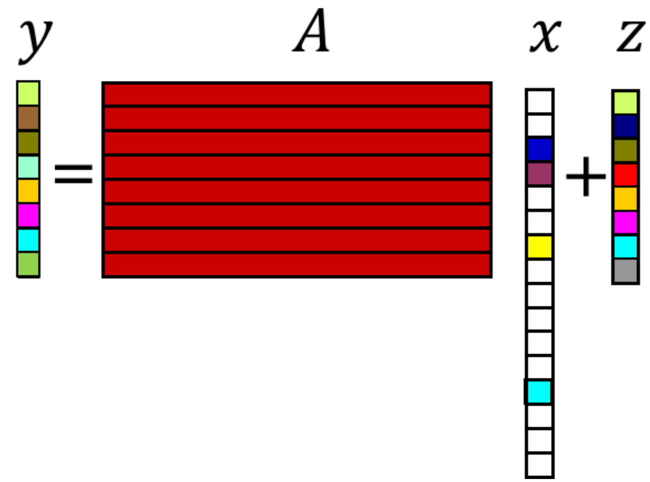
- Input statistics may be *unknown*
- Simple i.i.d. model may be inaccurate

- **Goal**: Approach the minimum mean square error (MMSE) for general *stationary ergodic* input

Universal algorithm

Main Idea

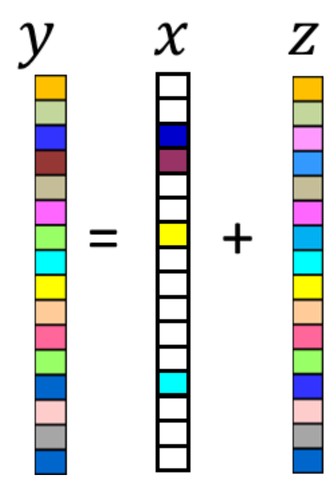
Compressed sensing



approximate message passing

[Donoho et al. 2009]

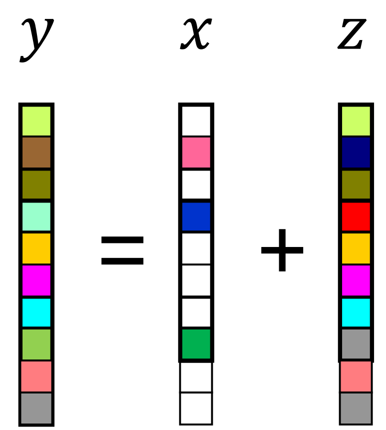
Iterative denoising



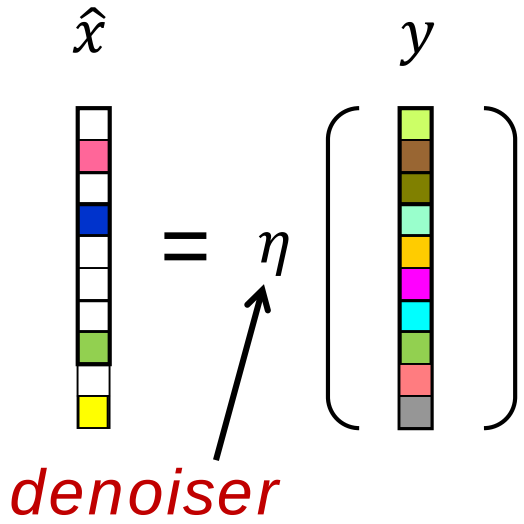
Denoising

Denoising

- Length-N input vector x
- Additive noise z
- Observations $y = x + z$

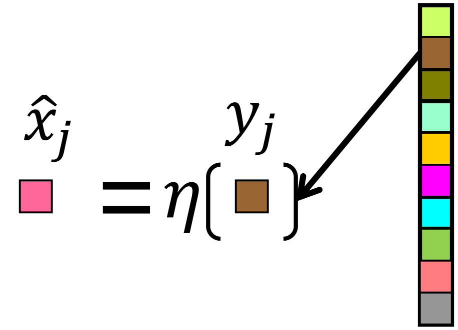


- **Goal**: estimate x from y
- Minimize $\mathbb{E}\left[\frac{1}{N} || x - \hat{x} ||_2^2\right]$



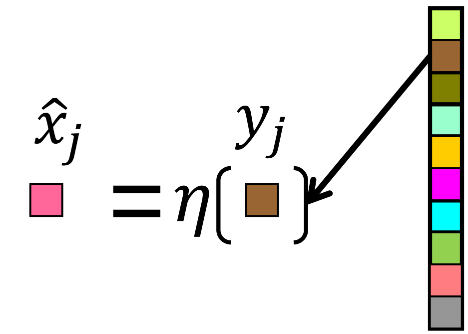
Two Denoising Schemes

- **Symbol-by-symbol scheme**
  - Bayesian denoiser: $\hat{x}_j = \mathbb{E}[x_j | y_j]$
  - Optimal for *memoryless input*

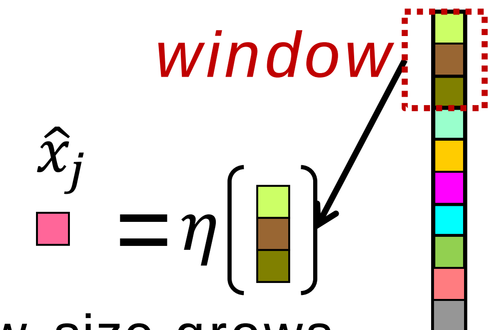


Two Denoising Schemes

- **Symbol-by-symbol scheme**
  - Bayesian denoiser: $\hat{x}_j = \mathbb{E}[x_j | y_j]$
  - Optimal for *memoryless input*



- **Sliding-window scheme**
  - Bayesian denoiser: $\hat{x}_j = \mathbb{E}[x_j | \text{window}]$
  - Optimal for *stationary ergodic input* as window-size grows



Two Denoising Schemes

- **Symbol-by-symbol scheme**
  - Bayesian denoiser: $\hat{x}_j = \boxed{\mathbb{E}[x_j | y_j]}$
  - Optimal for *memoryless input*

Need *input distribution* ***unknown***

- **Sliding-window scheme**
  - Bayesian denoiser: $\hat{x}_j = \boxed{\mathbb{E}[x_j | \text{window}]}$
  - Optimal for *stationary ergodic input* as window-size grows

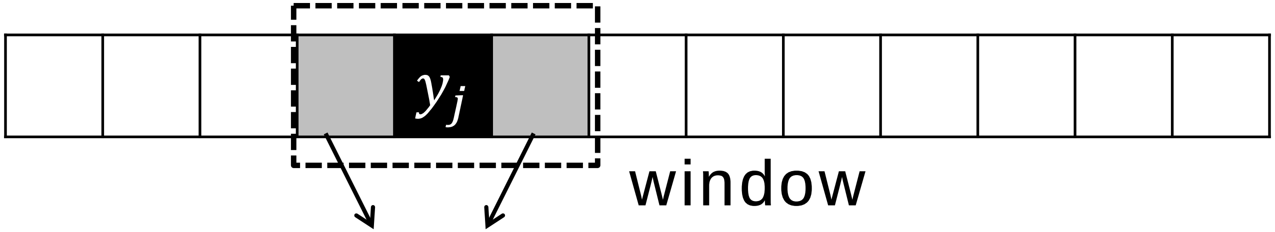
Denoising Unknown Stationary Ergodic Sources

[Universal Denoising]

Inspired by Tsachy Weissman

Universal Denoising
Based on Context Quantization

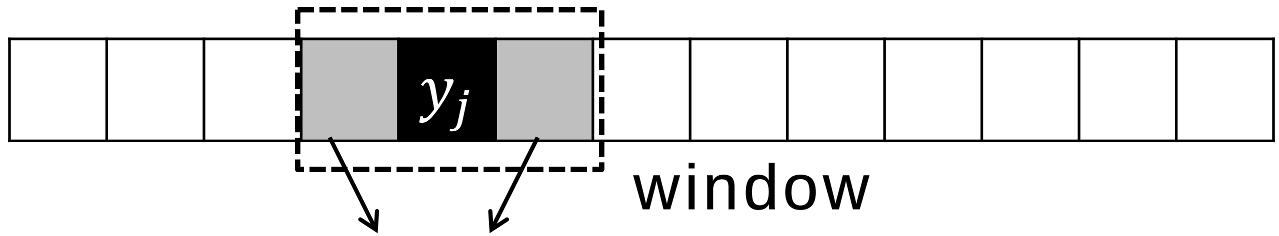
Inspired by [Sivaramakrishnan & Weissman, 2009]



context of y_j: (y_{j-1}, y_{j+1})

Universal Denoising
Based on Context Quantization

Inspired by [Sivaramakrishnan & Weissman, 2009]



context of y_j: (y_{j-1}, y_{j+1})

clustering

(y_1, y_3)

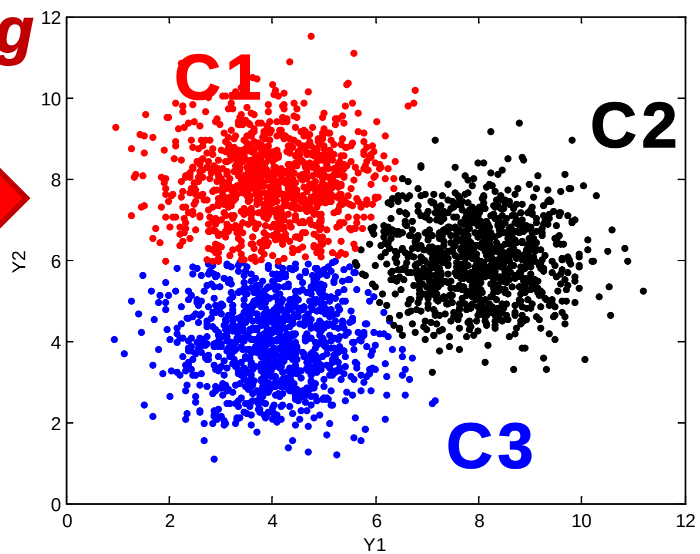
(y_2, y_4)

•

•

•

(y_{N-2}, y_N)

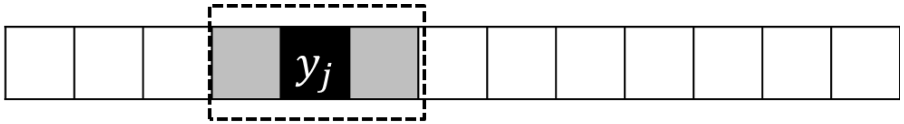


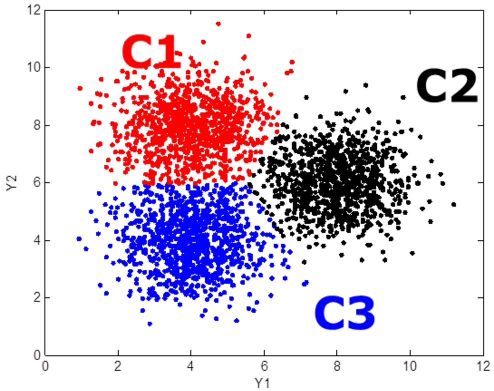
sub-sequencing

$$y_{C_l} = \{y_j : (y_{j-1}, y_{j+1}) \in C_l\}$$
$$l = 1,2,3.$$

Universal Denoising Based on Context Quantization

Inspired by [Sivaramakrishnan & Weissman, 2009]





$$y_{C_l} = \{y_j : (y_{j-1}, y_{j+1}) \in C_l\}$$

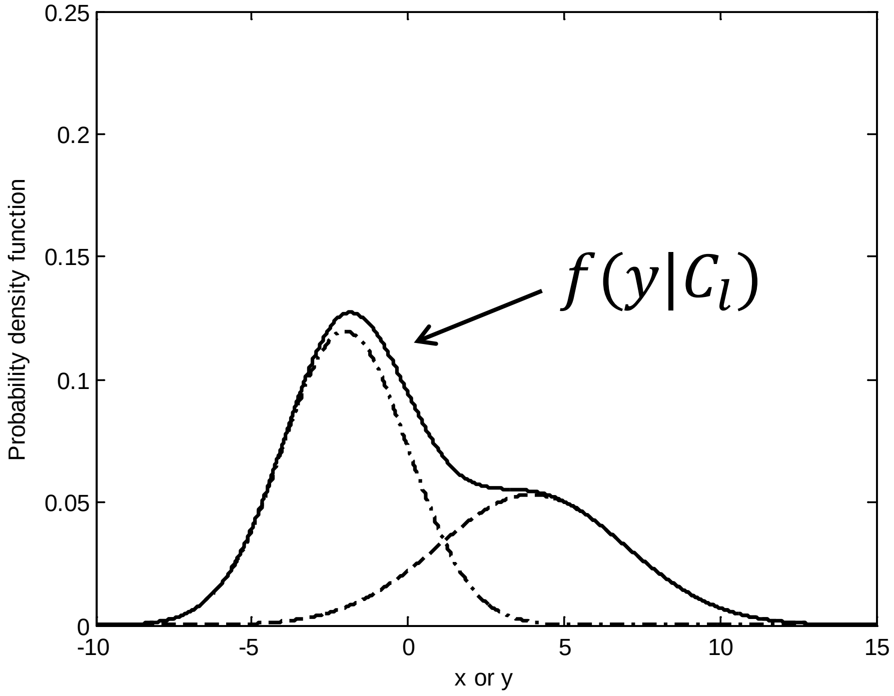
- Symbols in subsequence y_{C_l} *conditionally independent*
- Apply *symbol-by-symbol scheme*: $y_{C_l} = x_{C_l} + z_{C_l}$

Subsequence Denoising [Figueiredo & Jain 2001]

- Consider $y_{C_l} = x_{C_l} + z_{C_l}$, where $z_{C_l} \sim \mathcal{N}(0, \sigma_z^2 I)$
- Fit *Gaussian mixture* model to noisy subsequence y_{C_l}

$$f(y|C_l) = \sum_{s=1}^{S} \alpha_s \mathcal{N}(y; \mu_s, \sigma_s^2)$$

Example:
Gaussian mixture
with two components



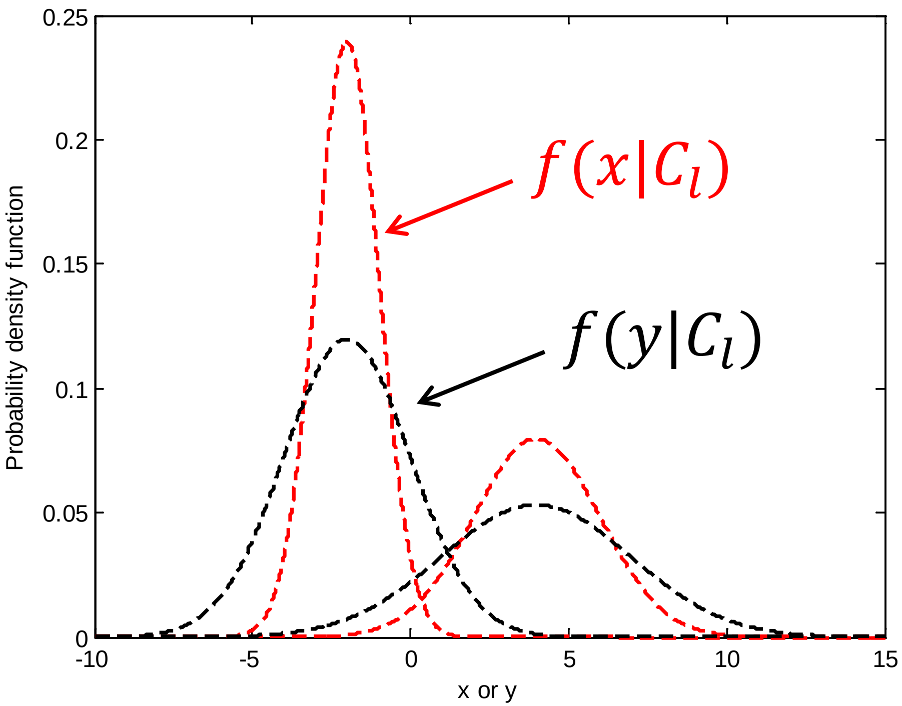
Subsequence Denoising

- Consider $y_{C_l} = x_{C_l} + z_{C_l}$, where $z_{C_l} \sim \mathcal{N}(0, \sigma_z^2 I)$
- Fit *Gaussian mixture* model to noisy subsequence y_{C_l}

$$f(y|C_l) = \sum_{s=1}^{S} \alpha_s \mathcal{N}(y; \mu_s, \sigma_s^2)$$

$$f(x|C_l) = \sum_{s=1}^{S} \alpha_s \mathcal{N}(x; \mu_s, \sigma_s^2 - \sigma_z^2)$$

Example:
Gaussian mixture
with two components



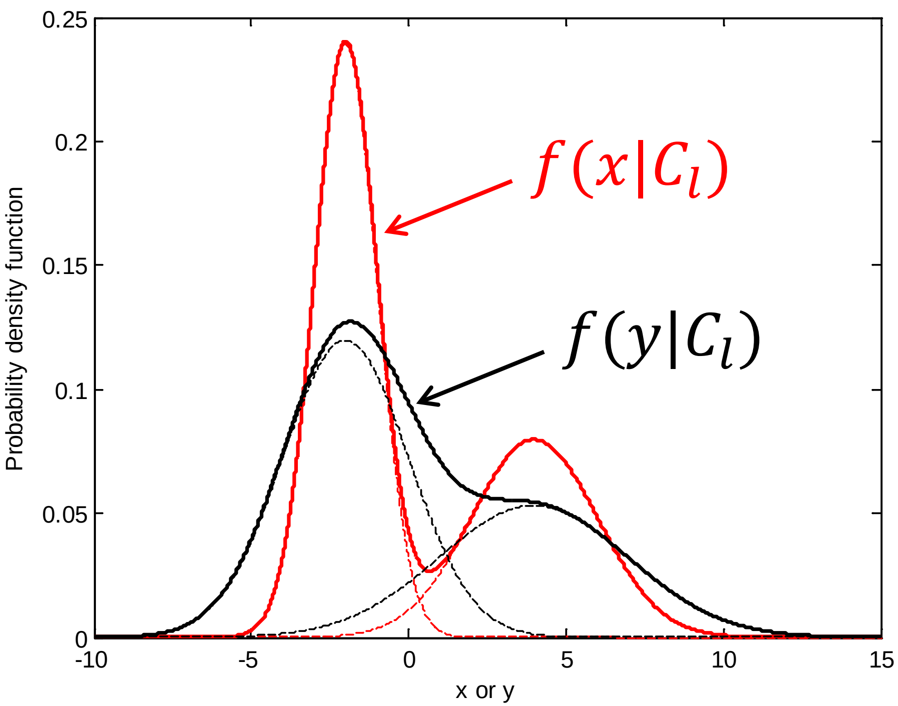
Subsequence Denoising

- Consider $y_{C_l} = x_{C_l} + z_{C_l}$, where $z_{C_l} \sim \mathcal{N}(0, \sigma_z^2 I)$
- Fit *Gaussian mixture* model to noisy subsequence y_{C_l}

$$f(y|C_l) = \sum_{s=1}^{S} \alpha_s \mathcal{N}(y; \mu_s, \sigma_s^2)$$

$$f(x|C_l) = \sum_{s=1}^{S} \alpha_s \mathcal{N}(x; \mu_s, \sigma_s^2 - \sigma_z^2)$$

Example:
Gaussian mixture
with two components



Subsequence Denoising

- Consider $y_{C_l} = x_{C_l} + z_{C_l}$, where $z_{C_l} \sim \mathcal{N}(0, \sigma_z^2 I)$
- Fit *Gaussian mixture* model to noisy subsequence y_{C_l}

$$f(y|C_l) = \sum_{s=1}^{S} \alpha_s \mathcal{N}(y; \mu_s, \sigma_s^2)$$

$$\boxed{f(x|C_l)} = \sum_{s=1}^{S} \alpha_s \mathcal{N}(x; \mu_s, \sigma_s^2 - \sigma_z^2)$$

approximate input distribution

- Approximate *Bayesian sliding-window denoiser*

Universal Denoising
Based on Context Quantization

Theorem [Siviramakrishnan & Wiessman 2009]:
For *bounded* stationary ergodic input, universal denoiser based on context quantization achieves *minimum mean square error* (MMSE) asymptotically

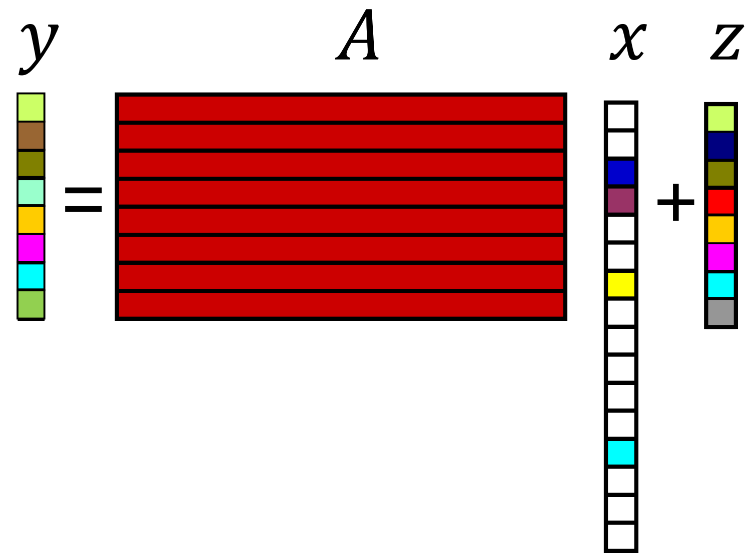
Conjecture: Under some technical conditions

bounded $\longrightarrow$ *unbounded*

Compressed Sensing for Unknown Stationary Ergodic Sources

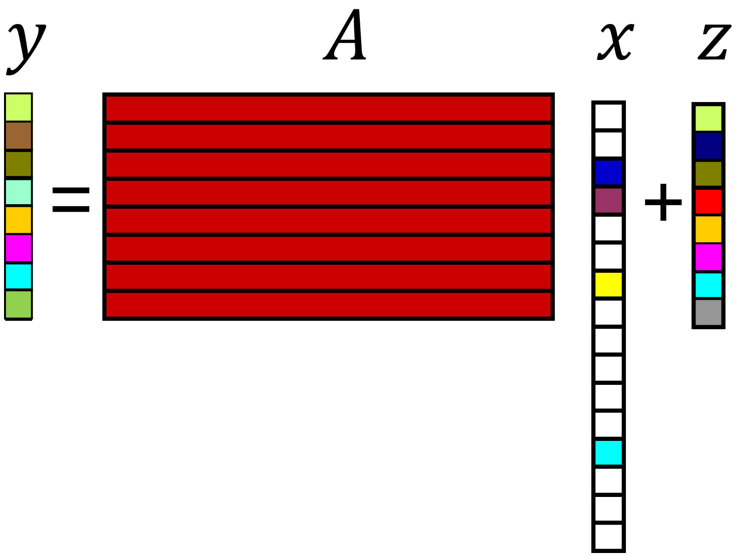
Compressed Sensing

- *Stationary ergodic input* x
- M-by-N measurement matrix A
- Additive white Gaussian noise z
- Observations $y = Ax + z$



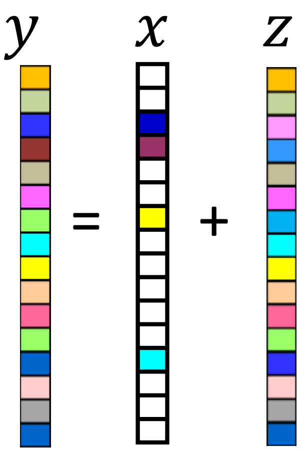
- **Goal**: estimate x from y and A
- Minimize $\mathbb{E}\left[\frac{1}{N} || x - \hat{x} ||_2^2\right]$

Compressed Sensing



- *Stationary ergodic input* x
- M-by-N measurement matrix A
- Additive white Gaussian noise z
- Observations $y = Ax + z$

- **Goal**: estimate x from y and A
- Minimize $\mathbb{E}\left[\frac{1}{N}||x - \hat{x}||_2^2\right]$
- Apply η_{UD} **???**

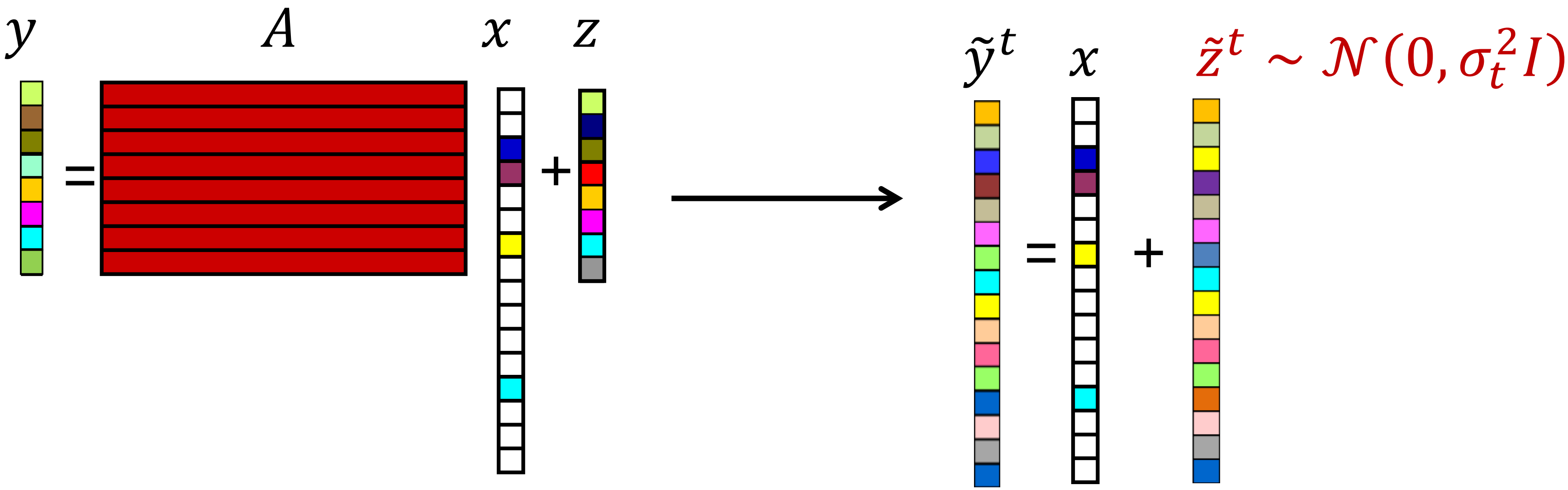


Approximate Message Passing

[Donoho et al. 2009]

Approximate Message Passing (AMP)

[Donoho et al. 2009]



At iteration t:

correction term

Residual $$r^t = y - Ax^t + \frac{r^{t-1}}{M/N} < \eta'_{t-1}\left(x^{t-1} + A^T r^{t-1}\right) >$$

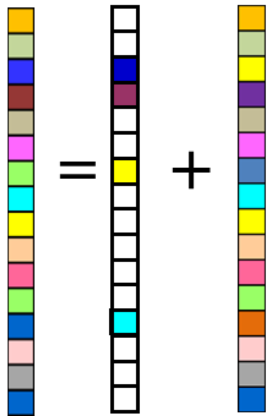
Pseudo-data $$\tilde{y}^t = x^t + A^T r^t = x + \tilde{z}^t$$

Denoising $$x^{t+1} = \eta_t(\tilde{y}^t)$$

State Evolution of AMP

[Donoho et al. 2009, Bayati & Montanari 2011...]

$\tilde{y}^t$ = x + $\tilde{z}^t \sim \mathcal{N}(0, \sigma_t^2 I)$



$$\sigma_{t+1}^2 = \sigma_z^2 + \frac{1}{M/N}\mathbb{E}[(\eta_t(X + \sigma_t W) - X)^2], \quad W \sim \mathcal{N}(0,1)$$

Symbol-by-symbol scheme

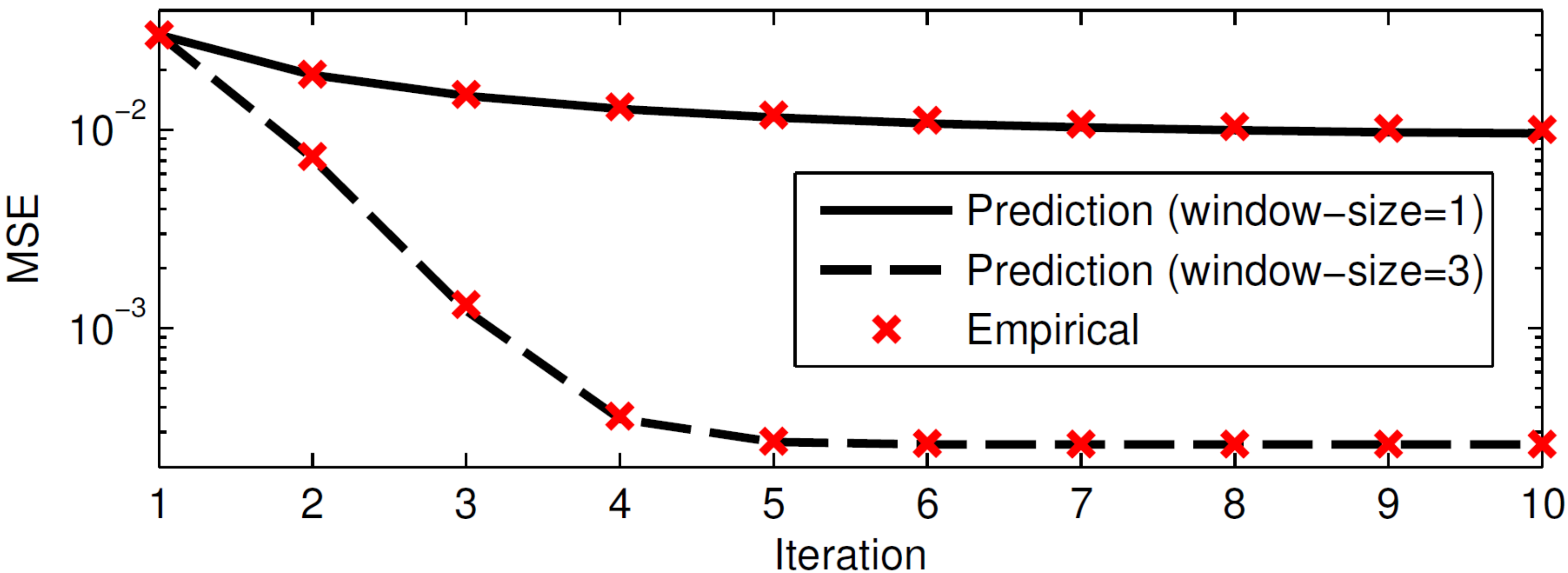
- Separable η_t: rigorously proved [Bayati & Montanari 2011]
- Broader class η_t: conjectured [Donoho et al. 2011]

AMP with Bayesian Sliding-Window Denoiser

Markov Input [N=10,000]

- Bayesian denoiser: $x_j^t = \eta_t(\text{window}) = \mathbb{E}\left[x_j | \text{window}\right]$
- *Non-i.i.d.* input

  $\longrightarrow$ *symbol-by-symbol* denoiser suboptimal



State Evolution of AMP

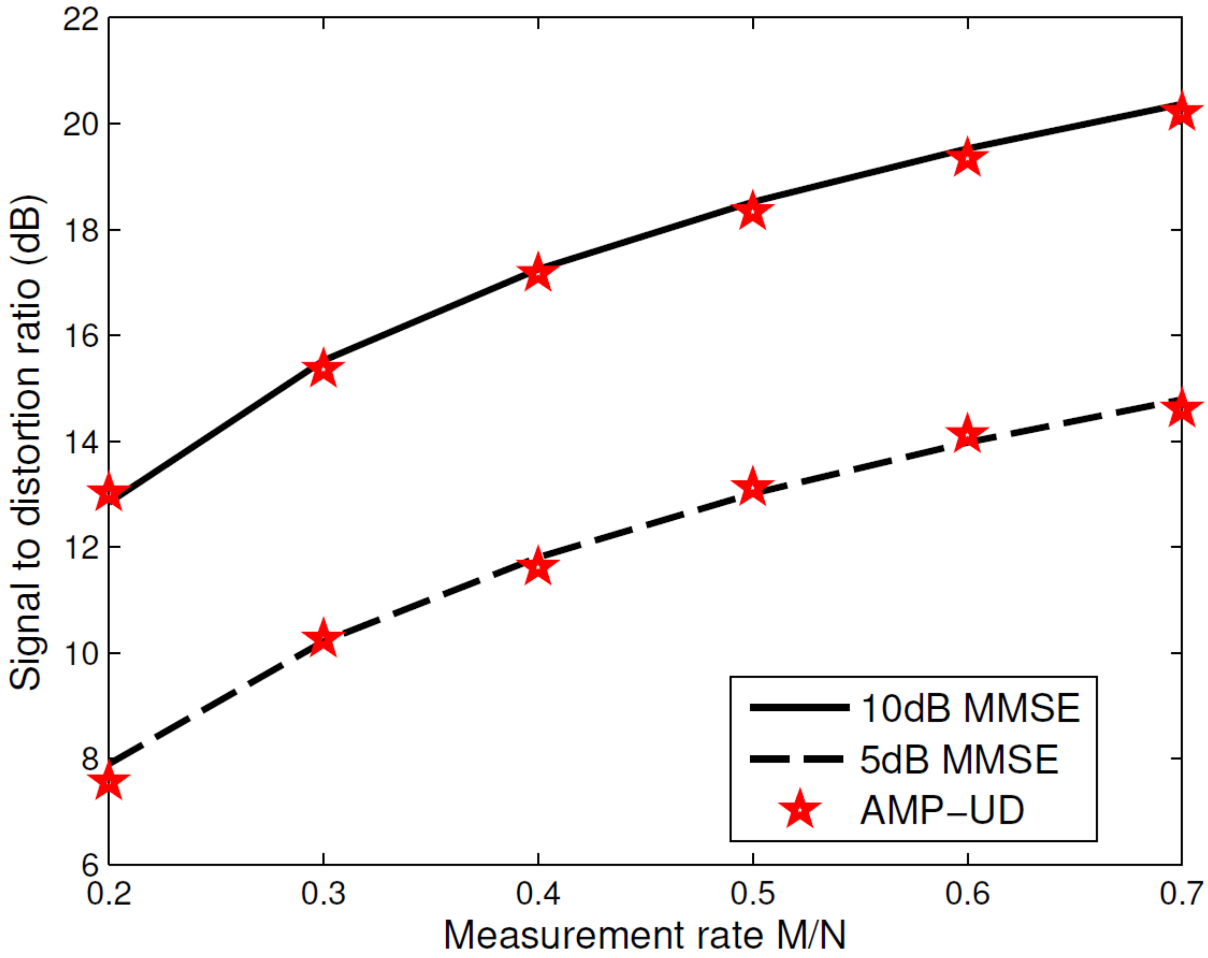
Conjecture: State evolution holds for AMP with Bayesian sliding-window denoisers

AMP with Universal Denoiser

[AMP-UD]

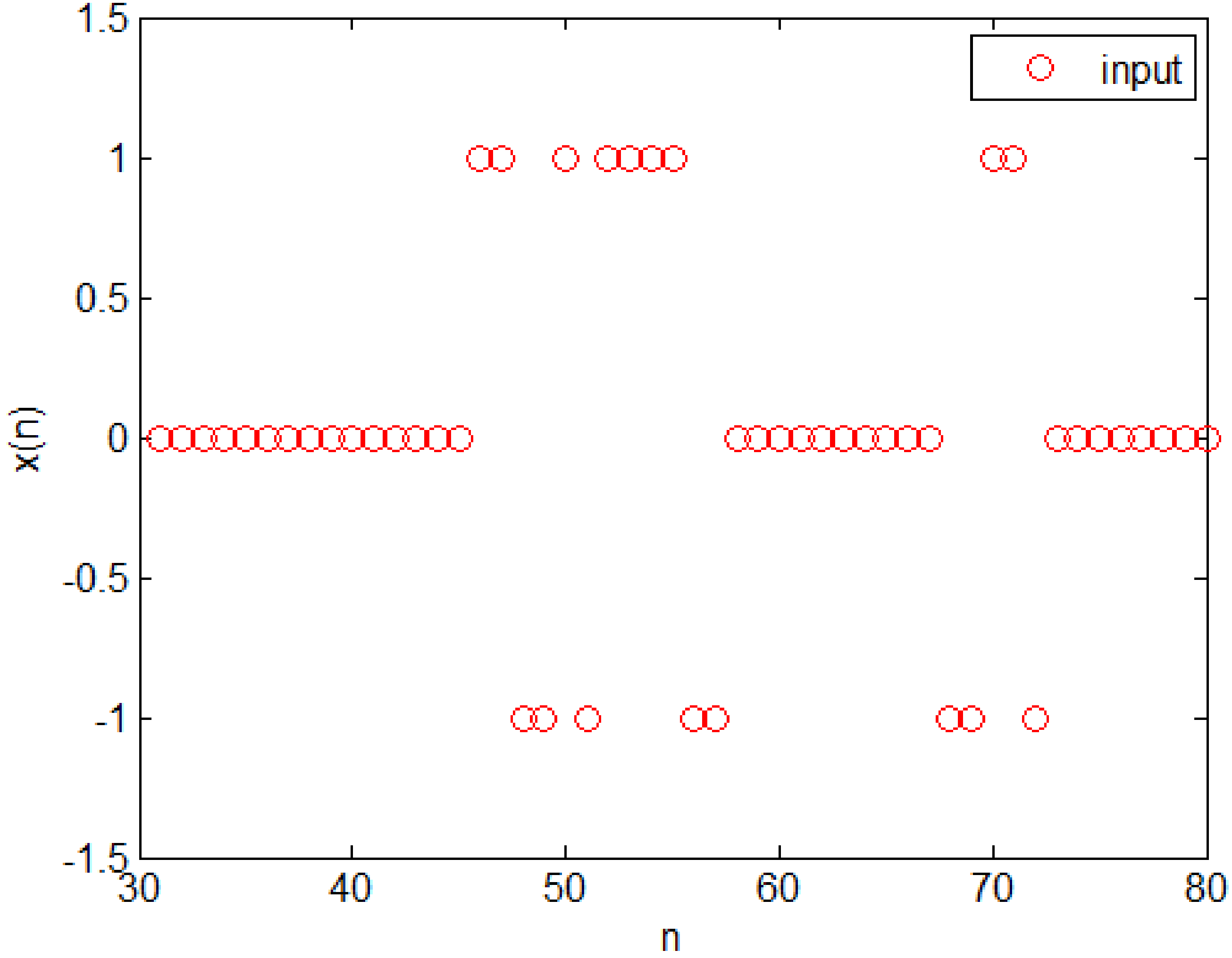
Sparse Laplace Input [N=10,000]

- i.i.d. input
- $f(x) = 0.97\delta(x) + 0.03\mathcal{L}(0,1)$



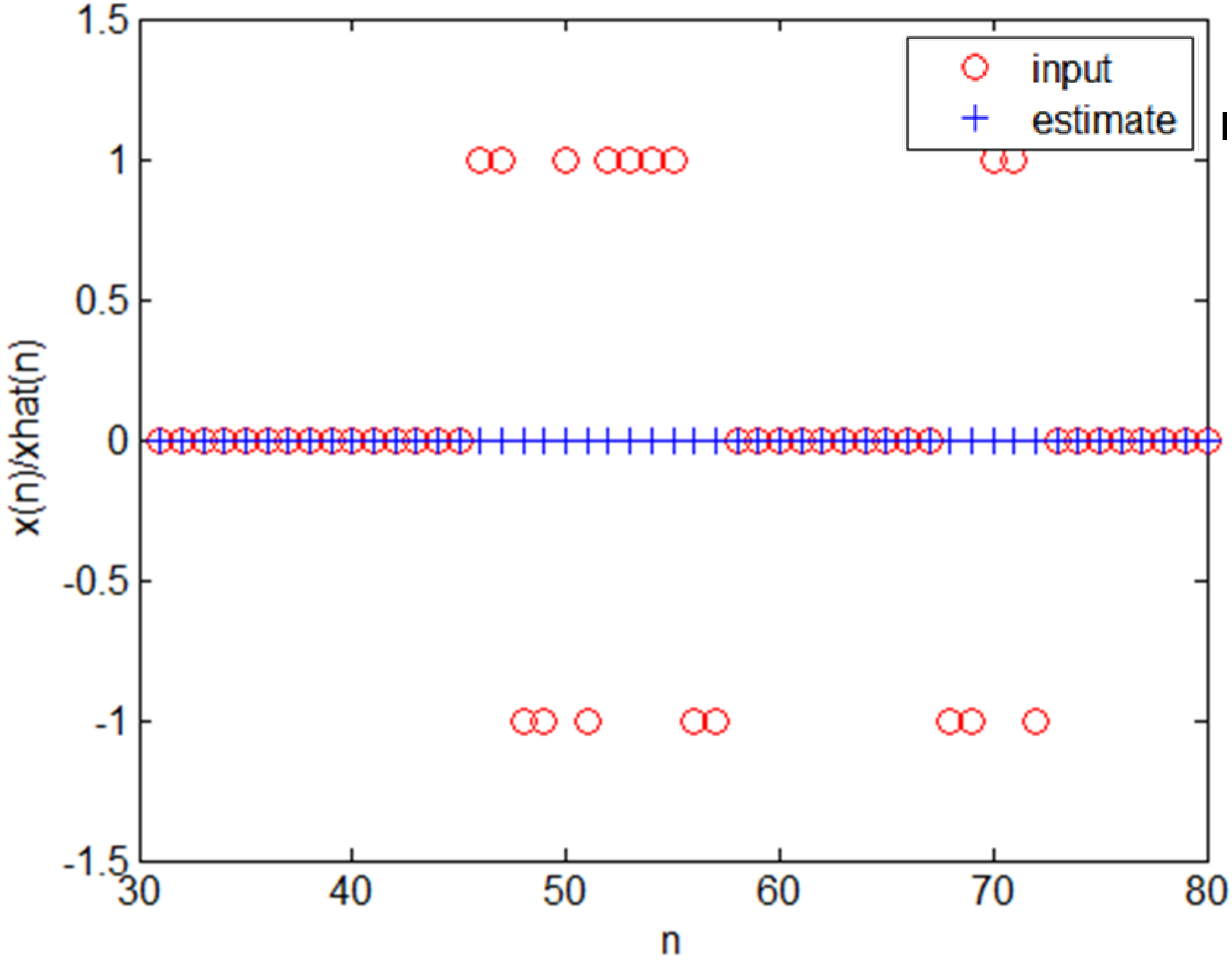
Markov-Rademacher Input

- Markov machine: zero-state & nonzero-state
- Rademacher ± 1 when nonzero
- 30% of nonzeros on average



Markov-Rademacher Input

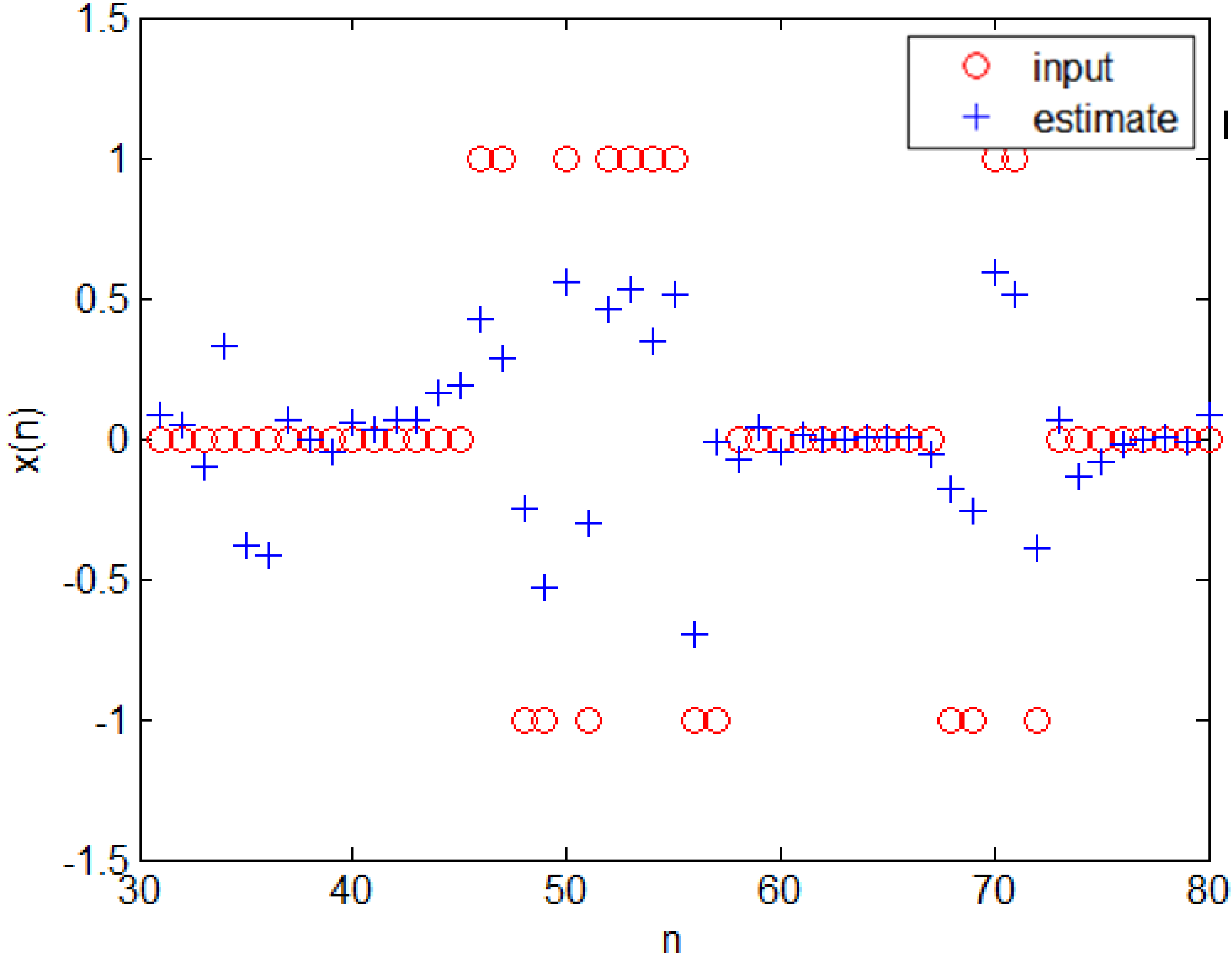
- Markov machine: zero-state & nonzero-state
- Rademacher ±1 when nonzero
- 30% of nonzeros on average



Iteration = 0

Markov-Rademacher Input

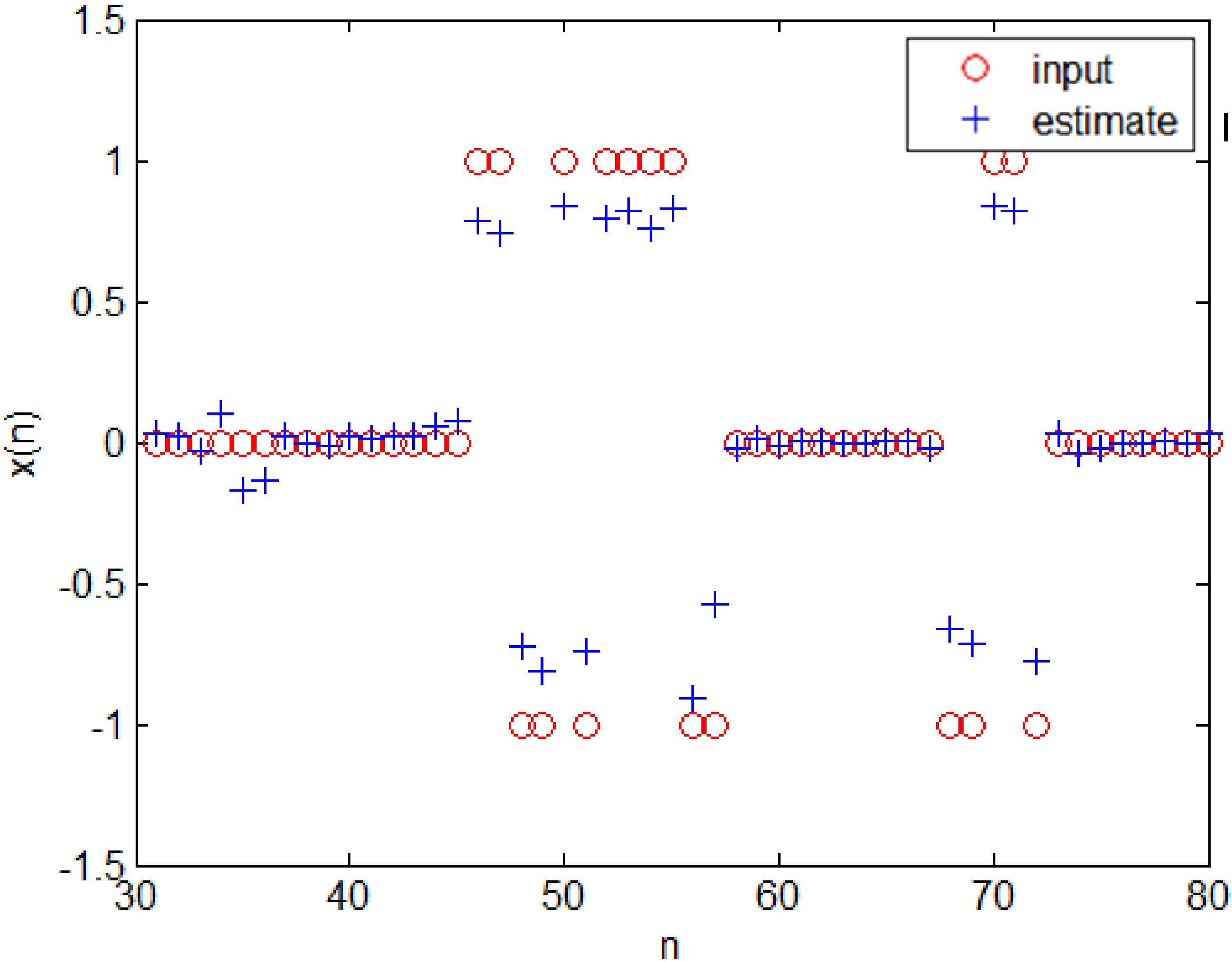
- Markov machine: zero-state & nonzero-state
- Rademacher ±1 when nonzero
- 30% of nonzeros on average



Iteration = 8

Markov-Rademacher Input

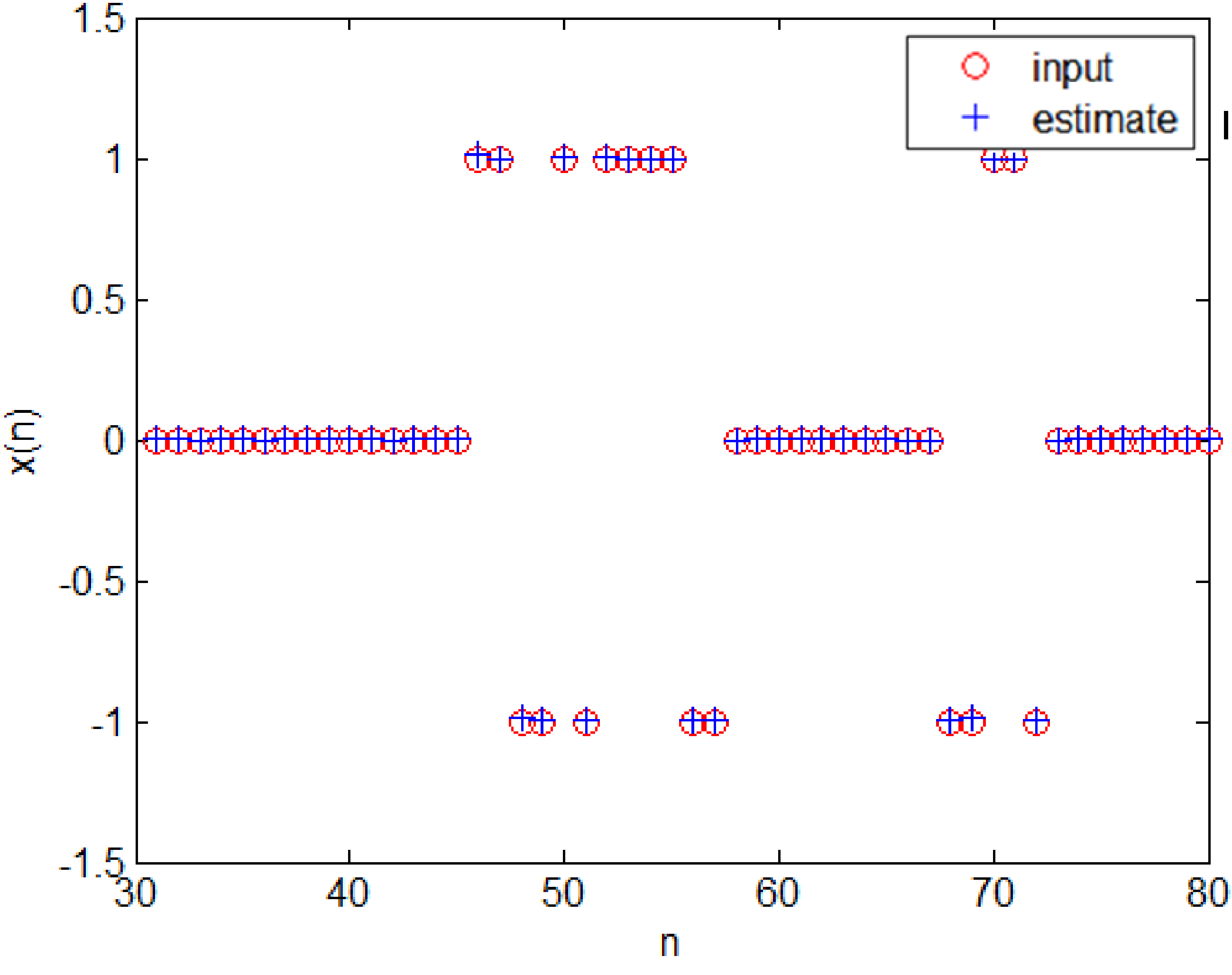
- Markov machine: zero-state & nonzero-state
- Rademacher ±1 when nonzero
- 30% of nonzeros on average



Iteration = 16

Markov-Rademacher Input

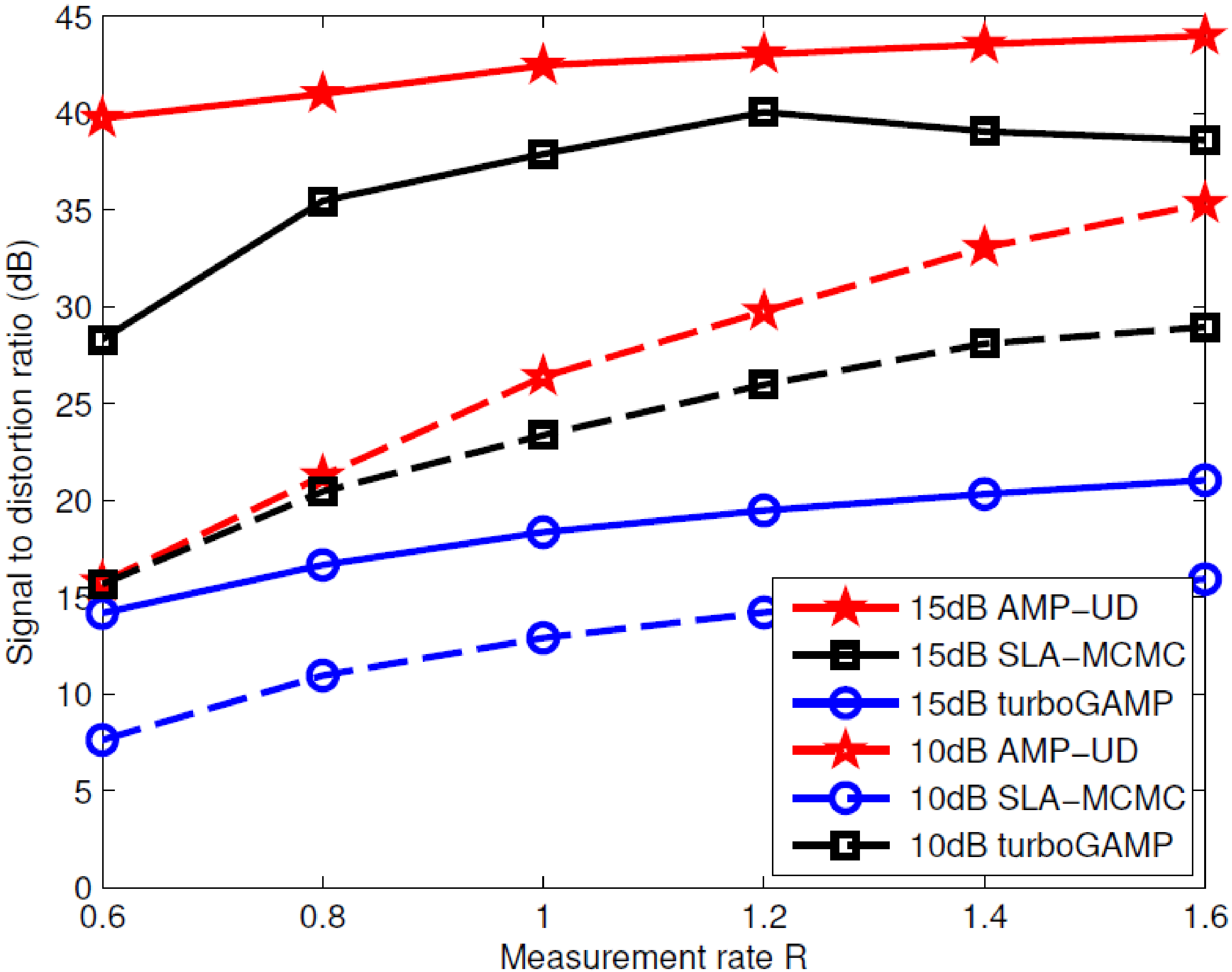
- Markov machine: zero-state & nonzero-state
- Rademacher ±1 when nonzero
- 30% of nonzeros on average



Iteration = 100

Markov-Rademacher Input [N=10,000]

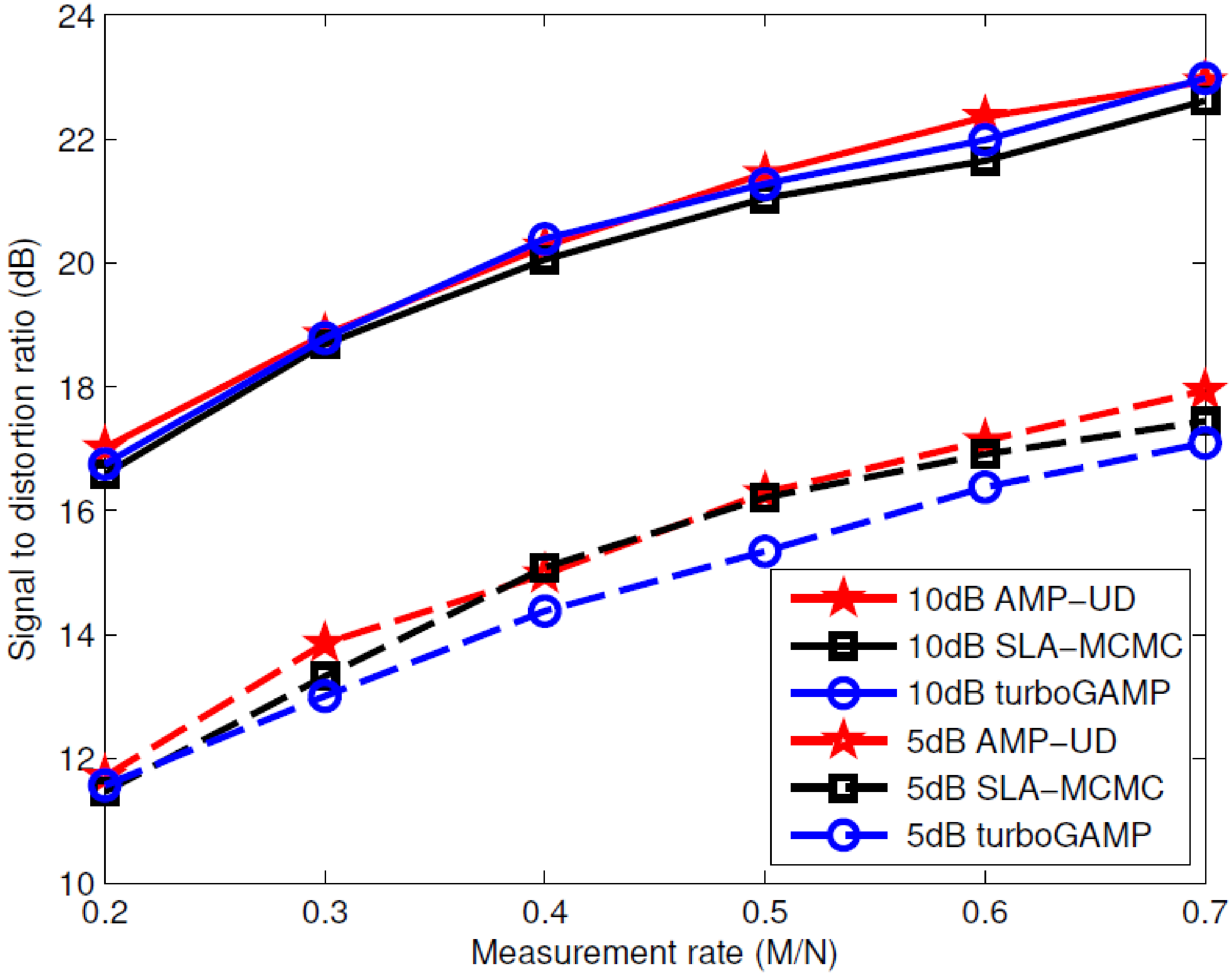
- 30% of nonzeros on average



AMP-UD: 5 minutes
[Zhu et al. 2014] hours
[Ziniel et al. 2012] 0.5 hours

Markov-Uniform Input [N=10,000]

- 10% nonzeros on average
- Nonzeros are uniform U[0,1]



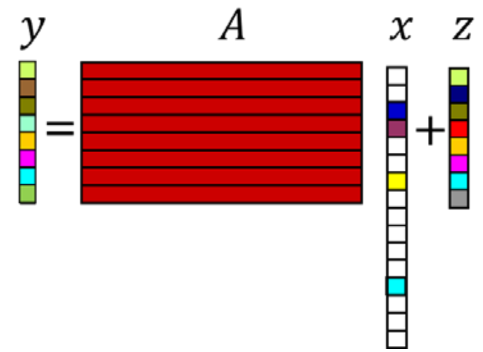
AMP with Universal Denoiser

Conjecture:

- AMP+UD achieves the MMSE for unknown stationary ergodic sources
- Under some technical conditions...

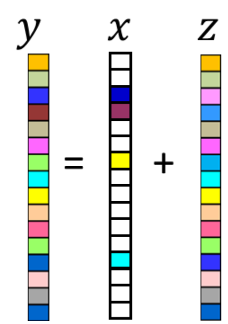
Summary

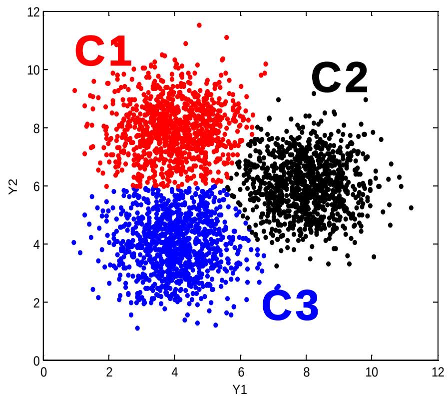
Compressed sensing



Approximate message passing

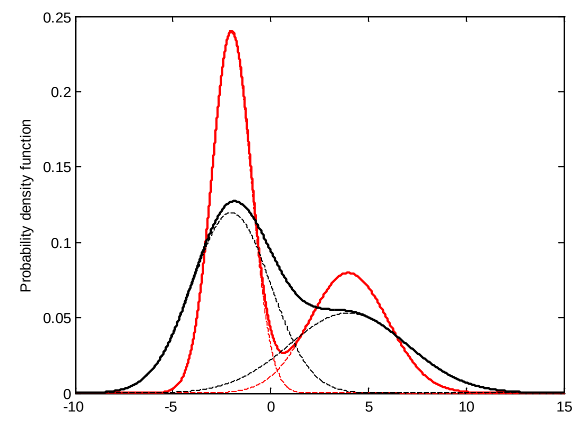
Universal denoising





Context quantization

i.i.d. denoising
(Guassian mixture model)



Future Directions

- Robust subsequence denoising
- Theoretical analysis
- Multi-dimensional input (e.g. images)

Thank you!