Universal Denoising in Approximate Message Passing

Yanting Ma, Junan Zhu, and Dror Baron

Department of Electrical and Computer Engineering, North Carolina State University

Motivation

- **Linear inverse problem**
  - Input signal: $x \in \mathbb{R}^N$
  - Measurement matrix: $A \in \mathbb{R}^{M \times N}$
  - Measurements: $y = Ax + z$, where z is noise
  - Estimate x given y and A
- **Challenges**
  - 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

- **Linear inverse → universal denoising**
  - Approximate message passing [Donoho *et al.* 2009]

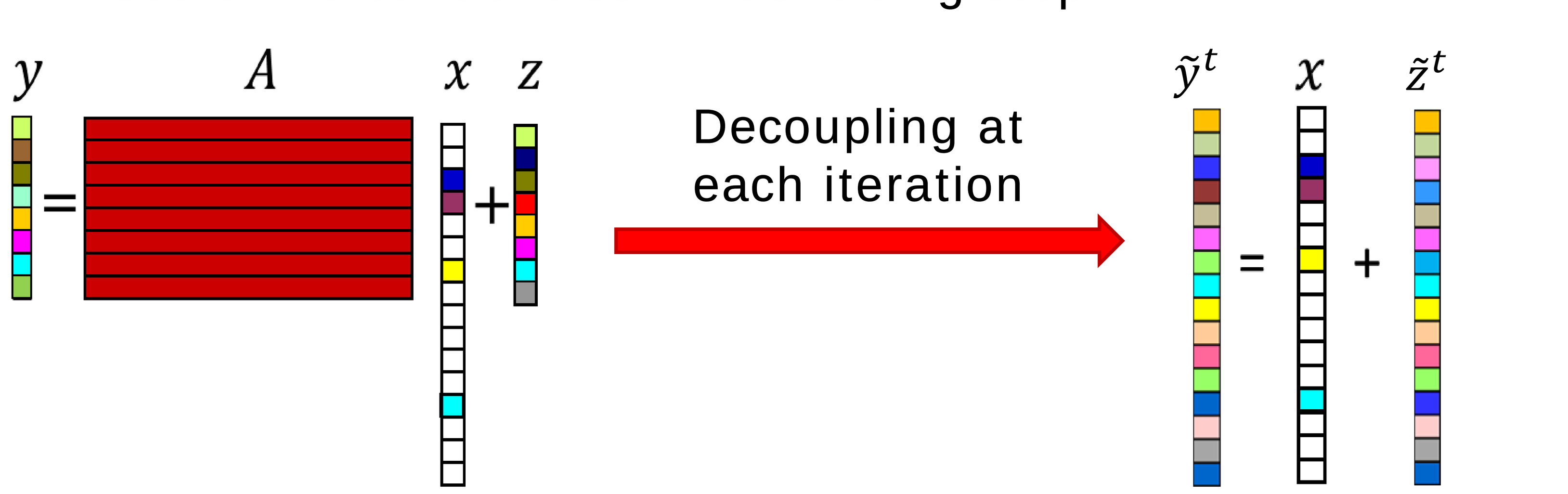
Iterate:

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

Pseudo-data $\quad \tilde{y}^t = x^t + A^T r^t = x + \tilde{z}^t, \ \tilde{z}^t \sim N(0, \sigma_t^2)$

Denoising $\quad x^{t+1} = \eta_t(\tilde{y}^t)$, η_t is denoiser

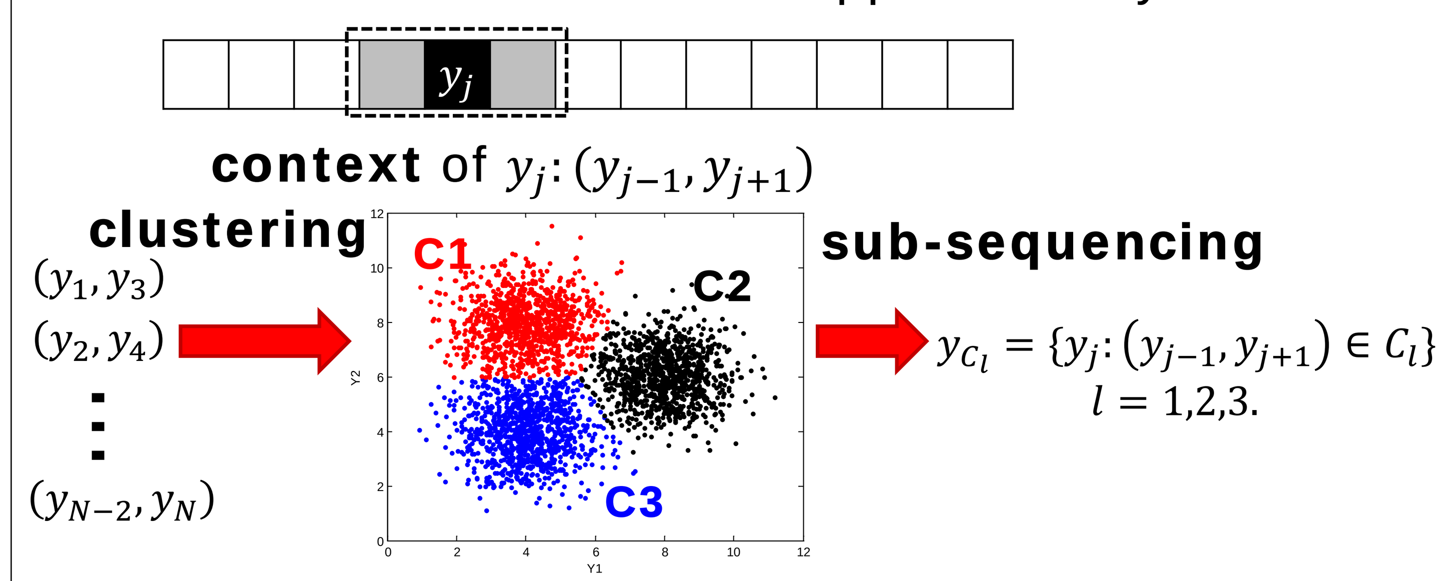
  - Use universal denoiser in denoising step



- **Universal denoising → i.i.d. denoising**
  - Sliding window denoising: estimate an entry from neighboring entries
  - Similar neighbors
    → similar estimation function
    → group entries with similar neighbors and estimate in i.i.d. fashion using MMSE estimator (conditional expectation)

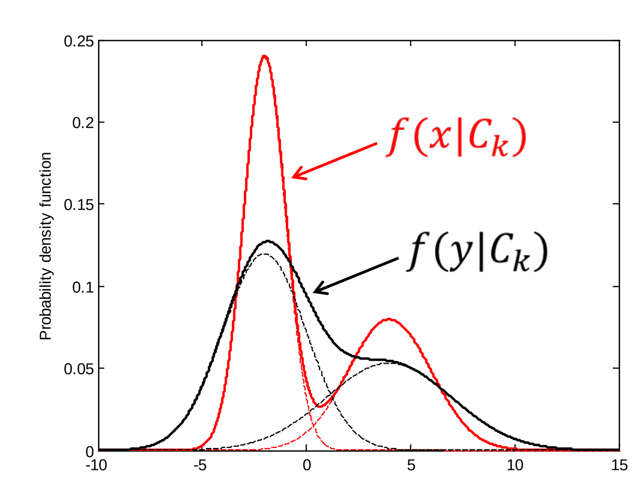
Universal Denoising

- **Context quantization** [Sivaramakrishnan & Weissman 2009]
  - Clustering based on Euclidian distance of context
  - Entries in each cluster are approximately i.i.d.

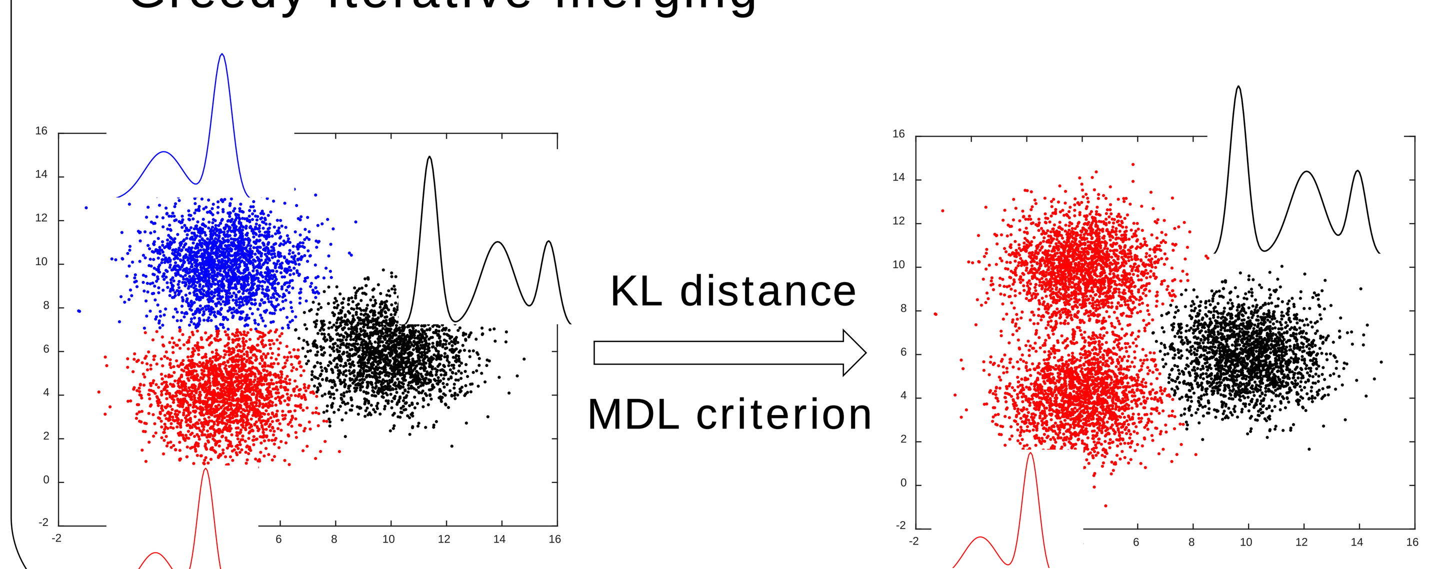


- **Gaussian mixture (GM)**
  - GM approximates many distributions well
  - GM convolved with Gaussian noise is still GM
  - Noise variance can be estimated in AMP

  → learn GM for noisy data, subtract noise variance from each Gaussian component

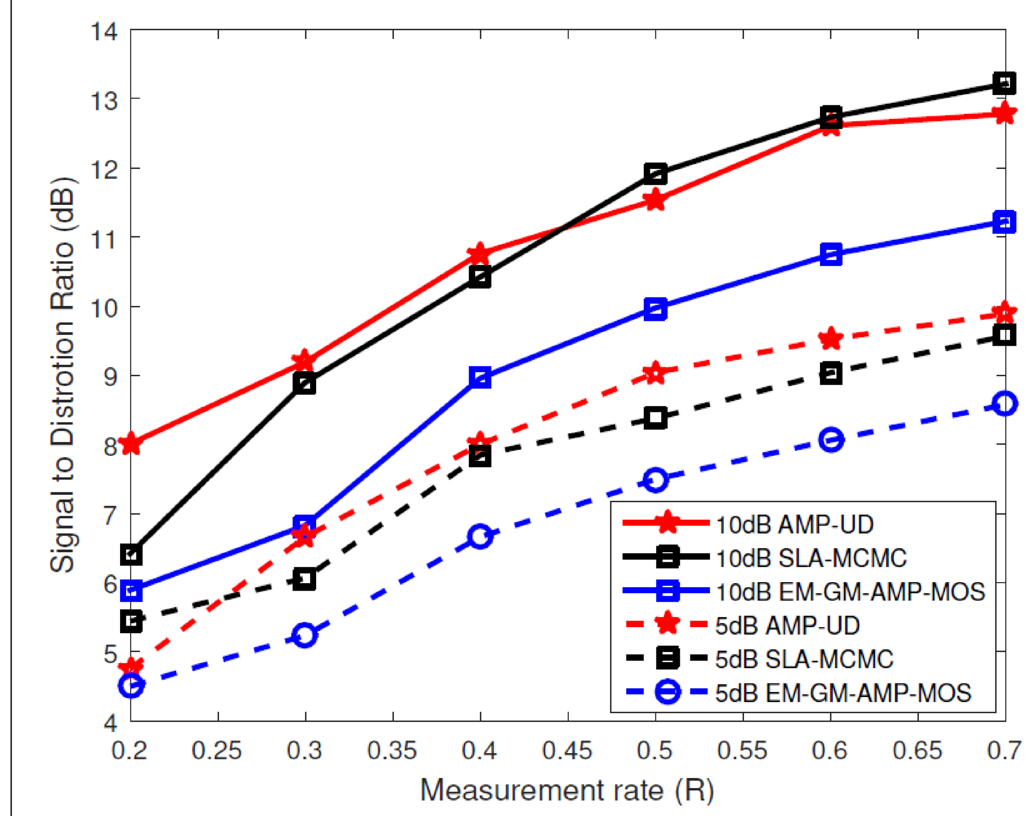


- **Cluster merging**
  - More accurate density estimation for larger clusters
  - Kullback-Leibler (KL) distance measures closeness (which clusters are candidates for merging)
  - MDL as model selection criterion (merge or not)
  - Greedy iterative merging



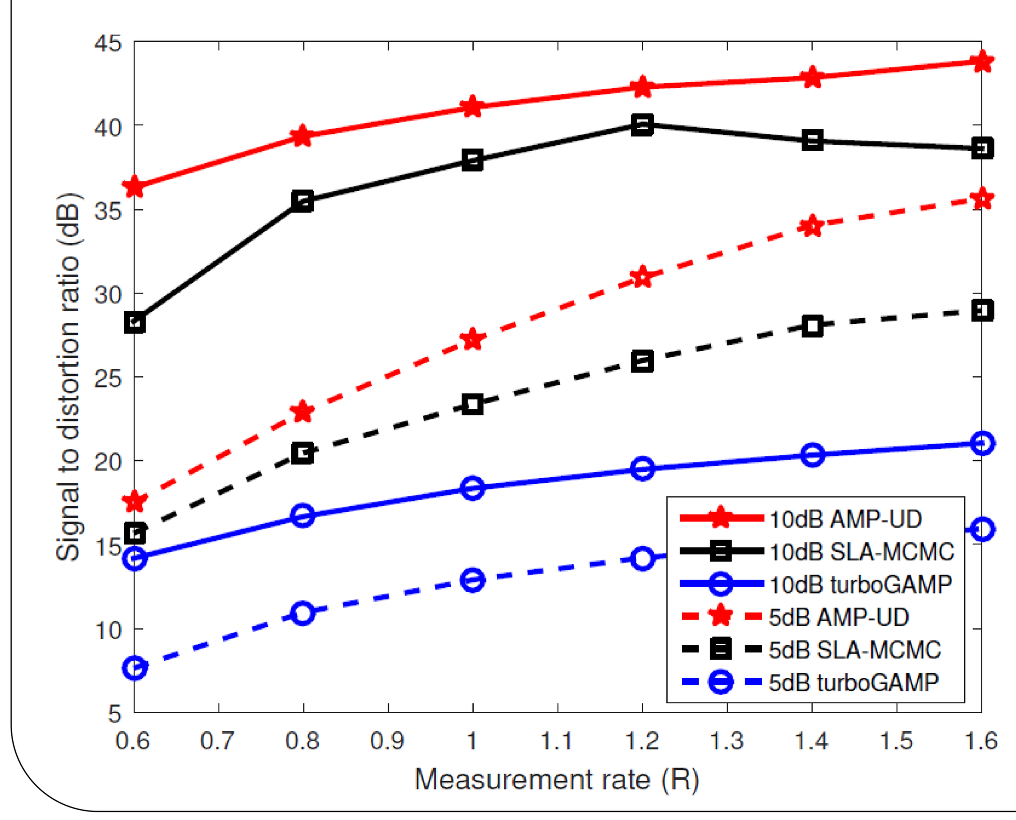
Numerical Results

- **Chirp sound clip**
  - Length 9600 segment of real world signal
  - Short-time discrete cosine transform (DCT)



EM-GM-AMP-MOS uses i.i.d. model. AMP-UD and SLA-MCMC, which use non-i.i.d. models, outperform EM-GM-AMP-MOS, indicating that i.i.d. model is suboptimal for this signal even in the transform domain.

- **Markov Rademacher input**
  - Two-state Markov Machine (zero/nonzero state)
  - +1 and -1 with equal probability in nonzero state
  - Length 10000 with 30% percent nonzero on average



AMP-UD achieves better reconstruction quality and runs faster than SLA-MCMC and turboGAMP. The runtime is 15 minutes for AMP-UD, 30 minutes for turboGAMP, and hours for SLA-MCMC.

Summary

- Designed AMP-UD for solving linear inverse problems with stationary ergodic input
- Merging concepts from AMP, context quantization based universal denoising, Gaussian mixture learning, and MDL model selection criterion
- Numerical results show AMP-UD outperforms the state-of-art algorithms in reconstruction quality and runtime