Multiprocessor Approximate Message Passing with Column-Wise Partitioning

Yanting Ma,* Yue M. Lu,# and Dror Baron*

* Department of Electrical and Computer Engineering, North Carolina State University

\# Paulson School of Engineering and Applied Sciences, Harvard University

Motivation

- **Linear inverse problem**
  - Input signal: $x \in \mathbb{R}^N$
  - Measurement matrix: $A \in \mathbb{R}^{n \times N}$
  - Measurements: $y = Ax + w$, where w is noise
  - Estimate x given A
  - Efficiently solved by approximate message passing (AMP)
- **Column-wise multiprocessor computing**
  - Alleviate storage and computational burden
  - Privacy preservation in data sharing
- **Goal**
  - AMP for column-wise multiprocessor computing
  - Establish performance guarantee for specific model
  - Numerical tests on more general models

C-MP-AMP Algorithm

s (t): outer (inner) iteration index; $\hat{t}$: max #inner iteration

Fusion center: $g^s = \sum_u r_u^{s-1,\hat{t}-1}$

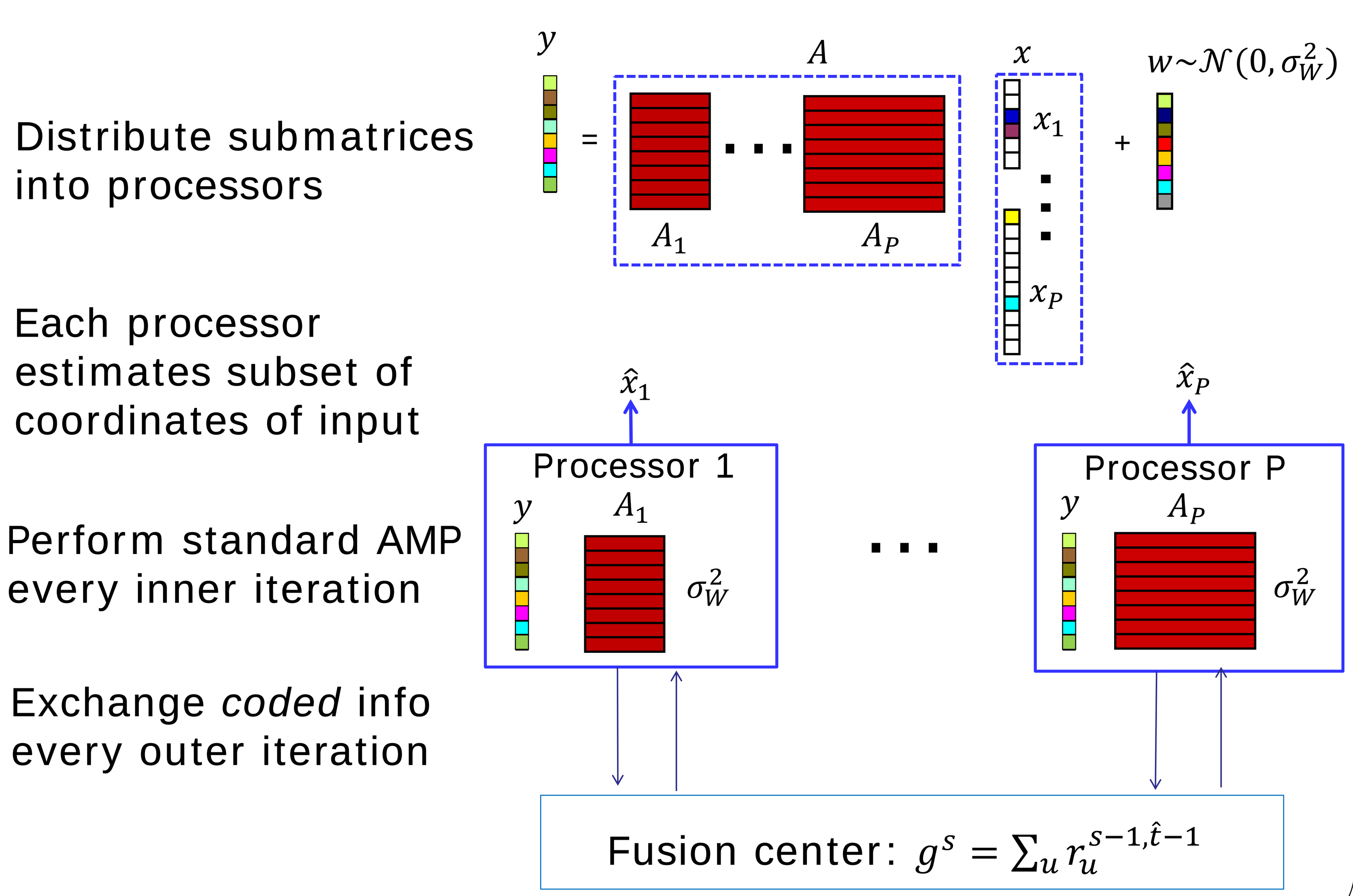
Initialization at each outer iteration: $x_p^{s,0} = x_p^{s-1,\hat{t}-1}, r_p^{s,0} = r_p^{s-1,\hat{t}-1}$

Inner iterations at each processor:

Residual: $z_p^{s,t} = y - g_s - (r_p^{s,t} - r_p^{s,0})$

Denoising: $x_p^{s,t+1} = \eta_{s,t}(x_p^{s,t} + A_p^* z_p^{s,t})$

"Onsager" correction: $r_p^{s,t+1} = A_p x^{s,t+1} - \frac{z_p^{s,t}}{n} \sum_{i=1}^{N_p} \eta'_{s,t}\left(\left[x_p^{s,t} + A_p^* z_p^{s,t}\right]_i\right)$



Performance Guarantees

- **State evolution sequences**

$$\left(\sigma_p^{s,0}\right)^2 = \left(\sigma_p^{s-1,\hat{t}}\right)^2$$

$$\left(\tau_p^{s,t}\right)^2 = \sigma_W^2 + \sum_{u \neq p}\left(\sigma_u^{s,0}\right)^2 + \left(\sigma_p^{s,t}\right)^2$$

$$\left(\sigma_p^{s,t+1}\right)^2 = \delta_p^{-1}\mathbb{E}\left[\left(\eta_{s,t}\left(X + \tau_p^{s,t}Z\right) - X\right)^2\right], \quad \delta_p = \frac{n}{N_p} \in (0,\infty)$$

- **Performance guarantee**

Theorem 1. For models where AMP obeys state evolution, let input x be i.i.d. p_x. Then for all $\epsilon \in (0,1)$, there exist $K_{s,t}, \kappa_{s,t} > 0$ independent of n, ϵ such that

$$P\left(\left|\frac{1}{N_p}\sum_i\left(\left[x_p^{s,t+1}\right]_i - \left[x_p\right]_i\right)^2 - \mathbb{E}\left[\left(\eta_{s,t}\left(X + \tau_p^{s,t}Z\right) - X\right)^2\right]\right| \geq \epsilon\right) \leq K_{s,t}e^{-\kappa_{s,t}n\epsilon},$$

where $X \sim p_x$ and Z is standard normal and independent of X.

- **Remarks**
  - Implies convergence with probability 1 (almost surely) by Borel Cantelli Lemma
  - C-MP-AMP converges to a fixed point no worse than AMP, hence achieves information theoretic minimum mean squared error (MMSE) whenever AMP achieves
  - Asymptotic dynamics of C-MP-AMP are identical to AMP when running one inner iteration per outer iteration

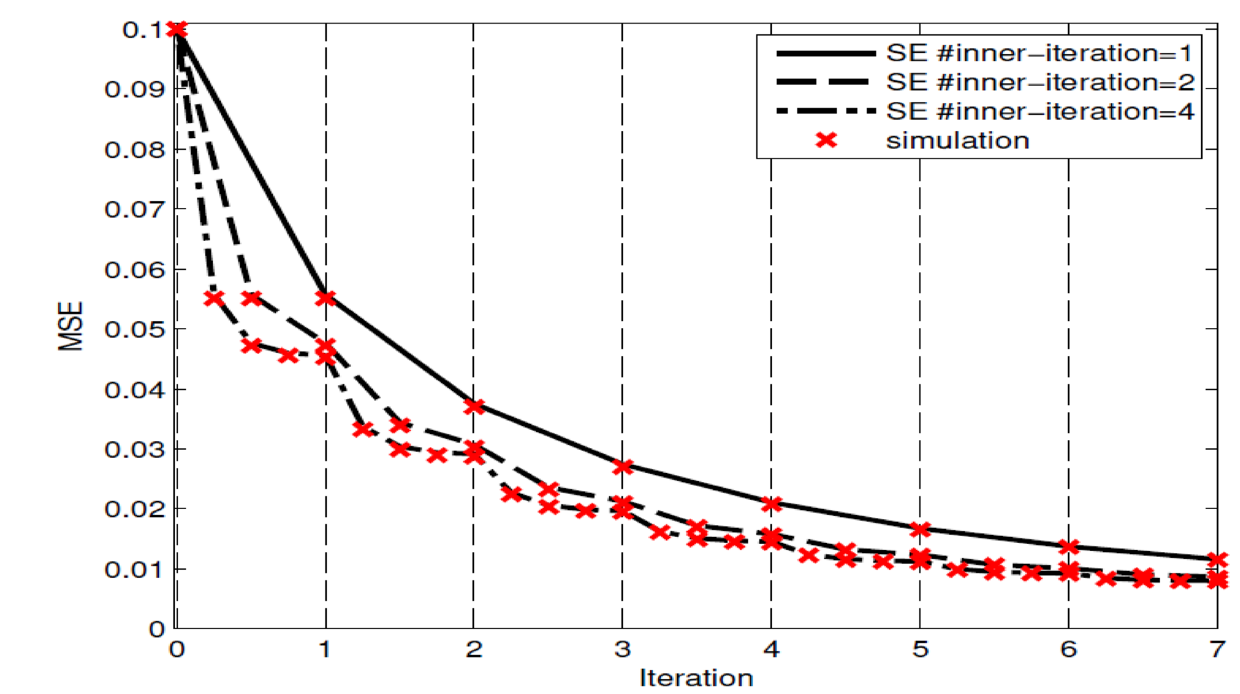
Proof Sketch

- **Compute conditional distribution of submatrices**:
  - $A_p|\{past, x, w\} \stackrel{d}{=} E_p + \mathcal{P}(\tilde{A}_p)$, where $\tilde{A}_p \stackrel{d}{=} A_p$, $\tilde{A}_p$ independent of $\{past, x, w\}$, projector $\mathcal{P}$ and matrix E_p defined by elements in $\{past, x, w\}$
  - $\tilde{A}_p$ independent of $\tilde{A}_q$ for $p \neq q$
- **Characterize conditional distribution of effective noise as i.i.d. Gaussian plus small deviation**
  - $x_p^{s,t} + A_p^* z_p^{s,t} - x_p|\{past, x, w\} \stackrel{d}{=} \tau_p^{s,t} Z_p^{s,t} + \Delta_p^{s,t}$
  - Show derivation $\Delta_p^{s,t}$ is small with high probability (w.h.p.)
- **Characterize conditional distribution of residual as i.i.d. Gaussian plus small deviation**
  - Define $b_p^{s,t} := r_p^{s,t} - A_p x_p$, hence $z_p^{s,t} = w - b_p^{s,t} - \sum_{u \neq p} b_u^{s,0}$
  - Show $< b_p^{s,t}, b_q^{s',t'} >/\mathrm{n}$ small w.h.p. ($\tilde{A}_p$ independent of $\tilde{A}_q$)
  - Show $b_p^{s,t}|\{past, x, w\} \stackrel{d}{=} \sigma_p^{s,t}\, \tilde{Z}_p^{s,t} + \widetilde{\Delta}_p^{s,t}$, and $\widetilde{\Delta}_p^{s,t}$ small w.h.p.
  - Yield $w - z_p^{s,t}|\{past, x, w\} \stackrel{d}{=} \left(\sum_{u \neq p}\left(\sigma_u^{s,0}\right)^2 + \left(\sigma_p^{s,t}\right)^2\right)^{\frac{1}{2}} \hat{Z}_p^{s,t} + \hat{\Delta}_p^{s,t}$

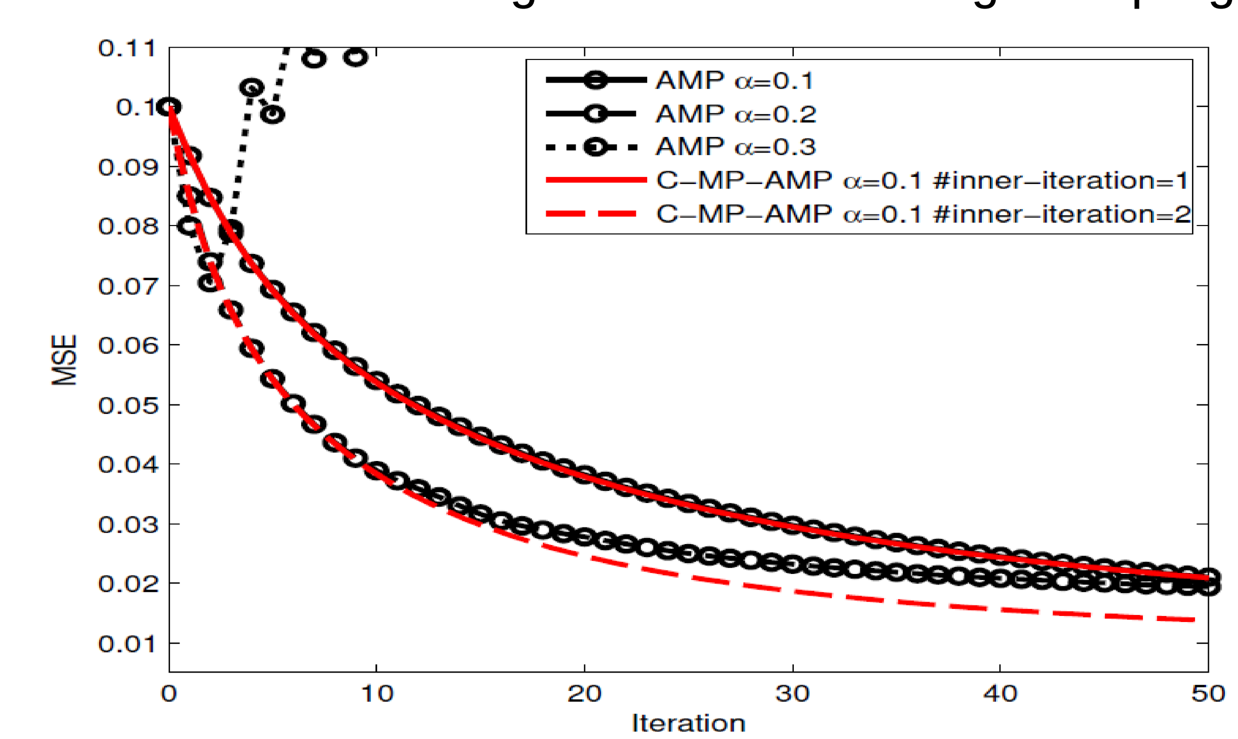
Details online: https://arxiv.org/abs/1701.02578

Numerical Results

- **Gaussian matrix setting**
  - 3 processors, length-30,000 input signal, 9,000 measurements, 15 dB input SNR
  - Verified state evolution prediction



- **Non-Gaussian matrix setting**
  - Models function learning: input signal consists of Taylor coefficients, columns of matrix are 1st, 2nd, and 3rd orders of the variables
  - C-MP-AMP converges well after using damping



Summary

- Designed C-MP-AMP algorithm for linear regression with column-wise portioning of the design matrix (distributed in feature)
- Derived state evolution to track the dynamics of C-MP-AMP
- Proved correctness of state evolution prediction for design matrix with i.i.d. Gaussian coordinates
- Tested C-MP-AMP algorithm for non-i.i.d. matrix; empirical performance was favorable