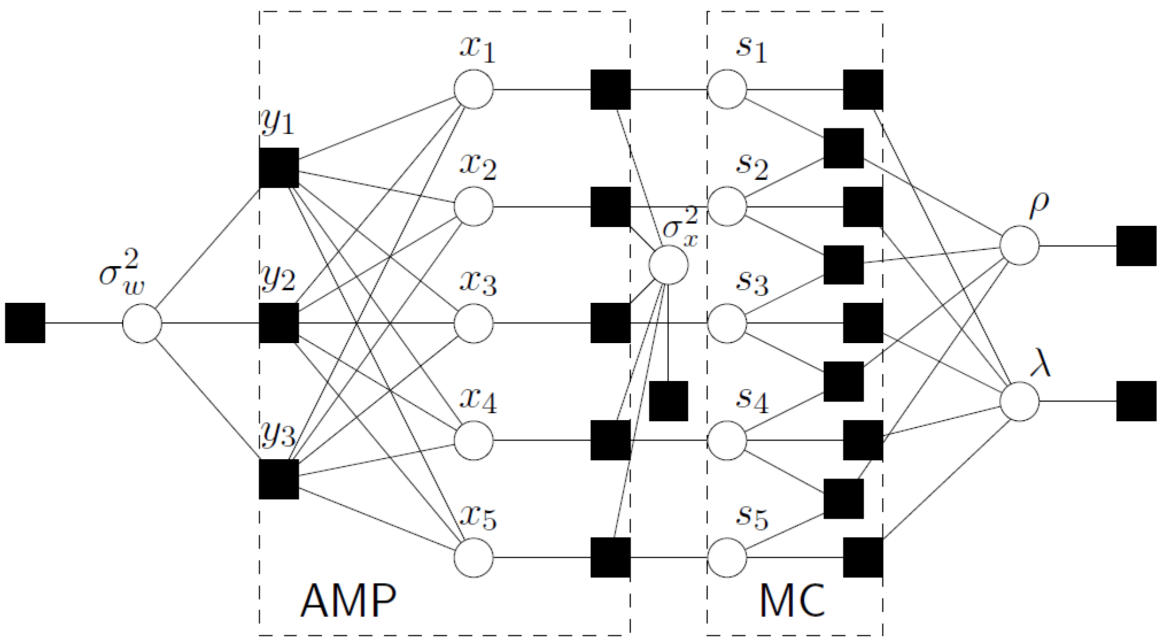
Approximate Message Passing for Bilinear Models



Volkan Cevher

*Laboratory for Information and Inference Systems – **LIONS / EPFL***

http://lions.epfl.ch

& Idiap Research Institute

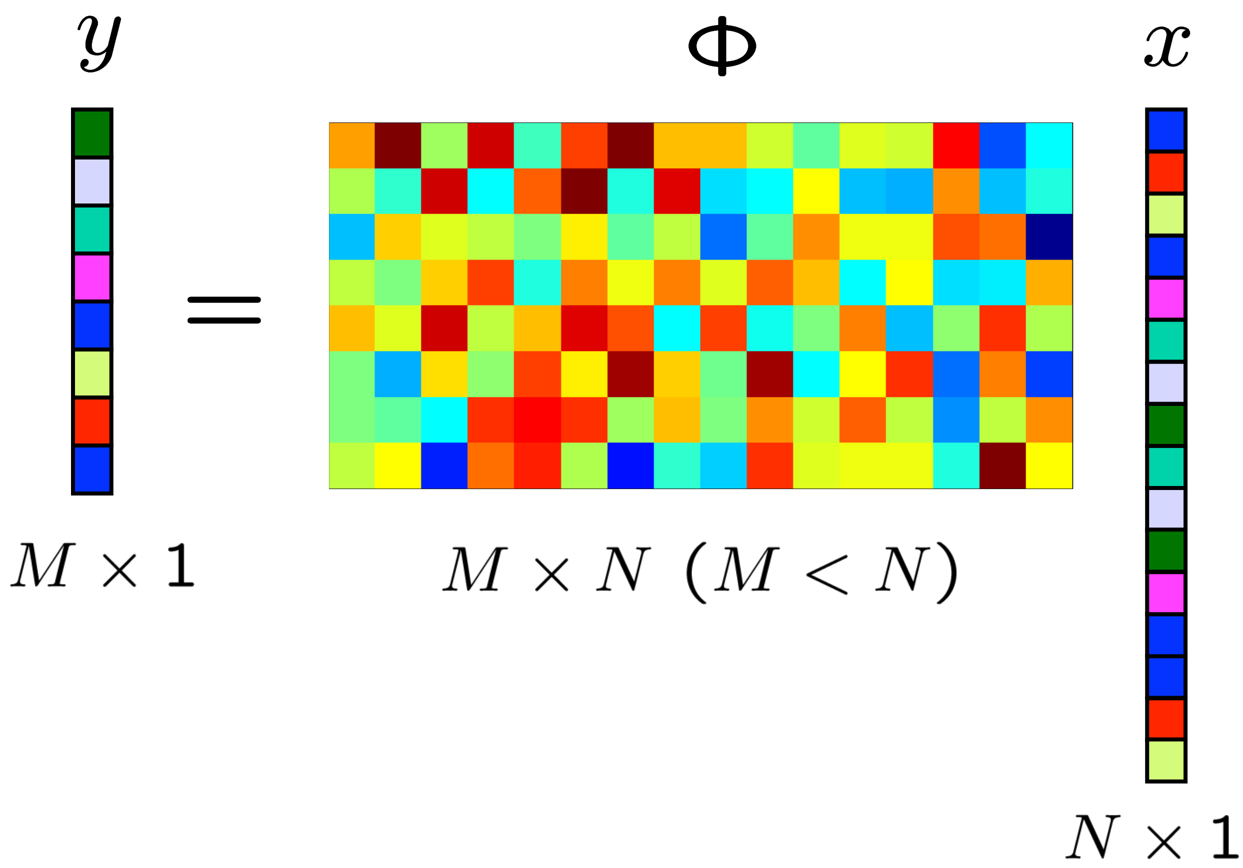
joint work with

Mitra Fatemi ***@Idiap***

Phil Schniter ***@OSU***

Acknowledgements: Sundeep Rangan and J. T. Parker.

Linear Dimensionality Reduction



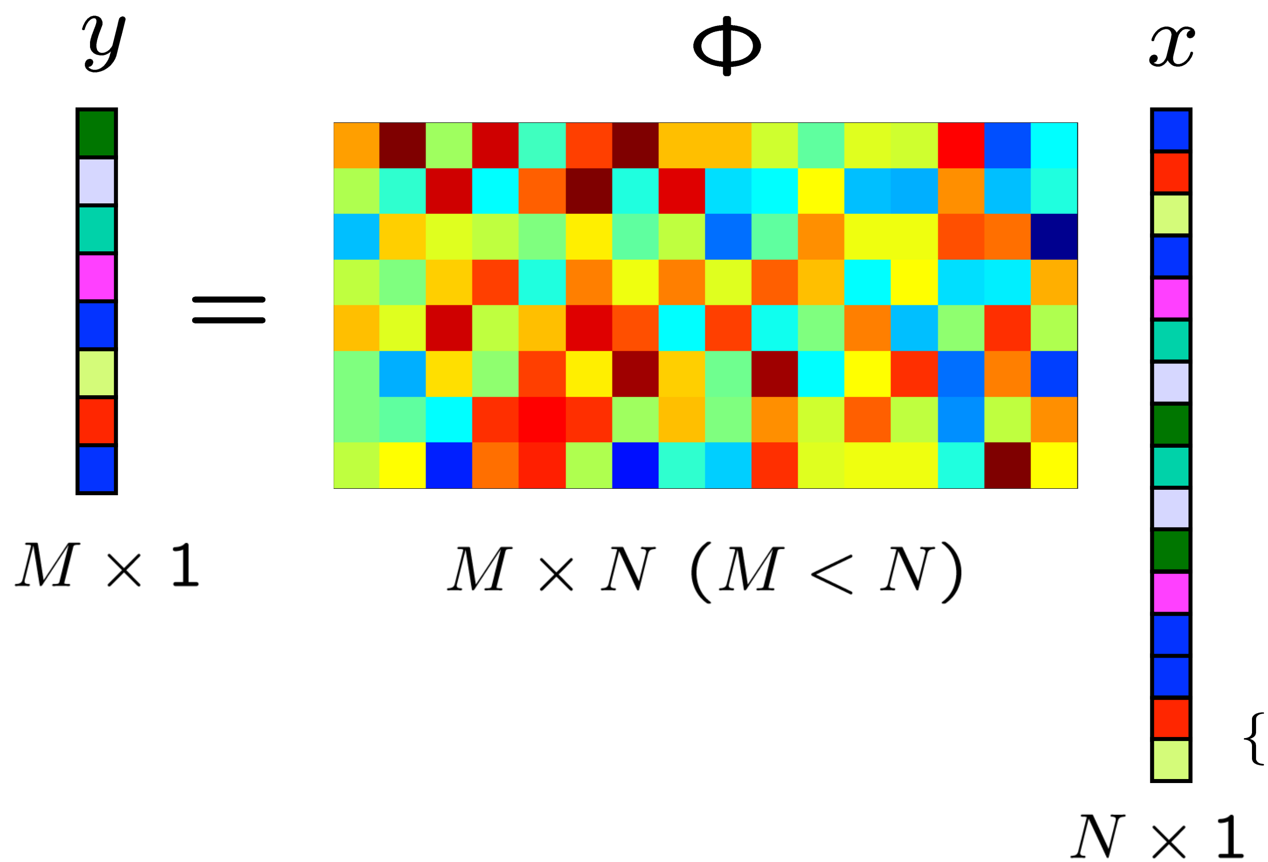
Compressive sensing non-adaptive measurements

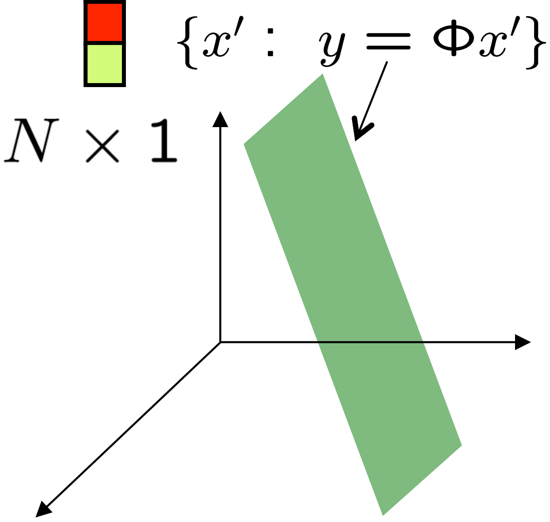
Sparse Bayesian learning dictionary of features

Information theory coding frame

Theoretical computer science sketching matrix / expander

Linear Dimensionality Reduction

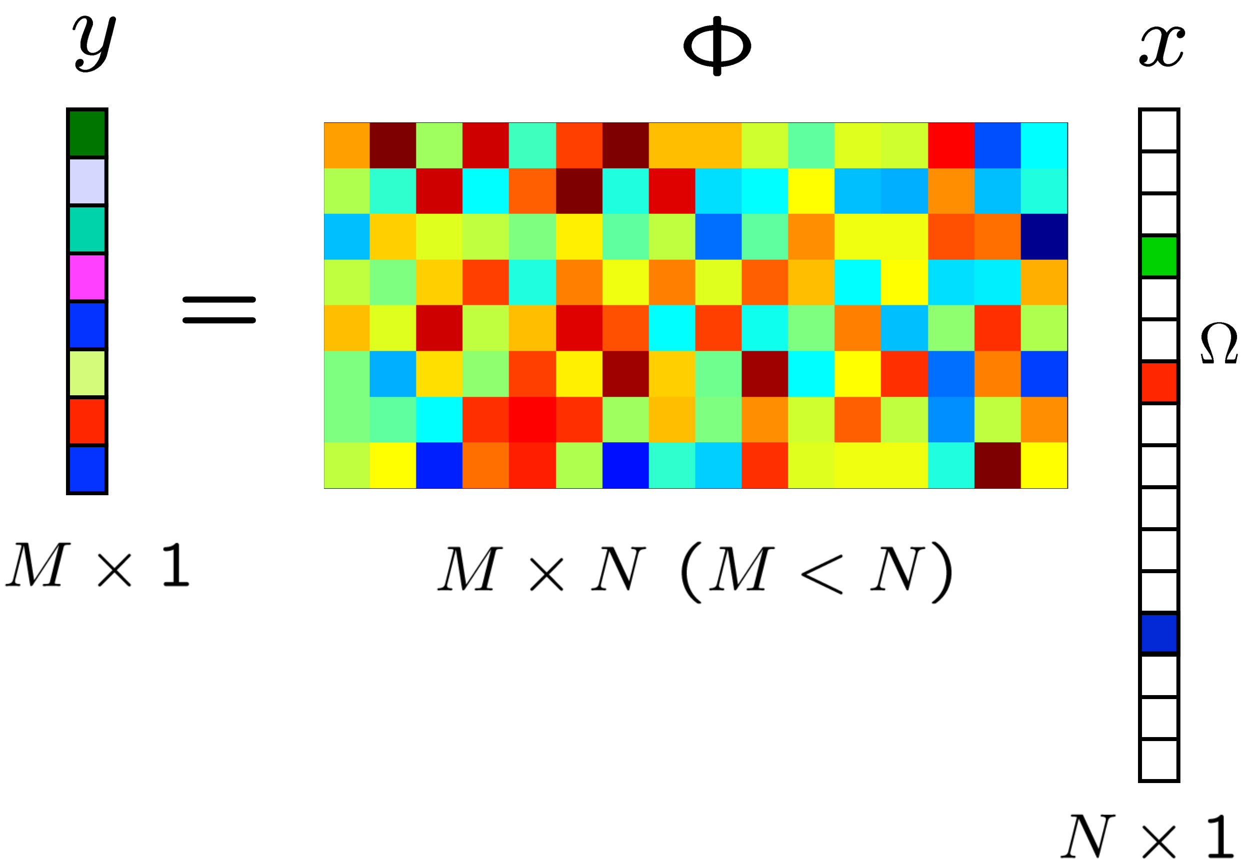




- **Challenge:** Null space of Φ: $\mathcal{N}(\Phi)$

$$x' = x + v \Rightarrow y, \; \forall v \in \mathcal{N}(\Phi)$$

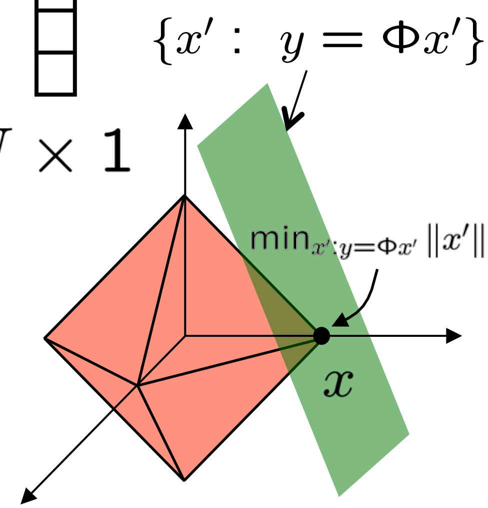
Linear Dimensionality Reduction



support:

$$\Omega = \{i : x_i \neq 0\}$$

$$|\Omega| \leq K$$



- **Key prior: sparsity in a known basis**
  - union-of-subspaces

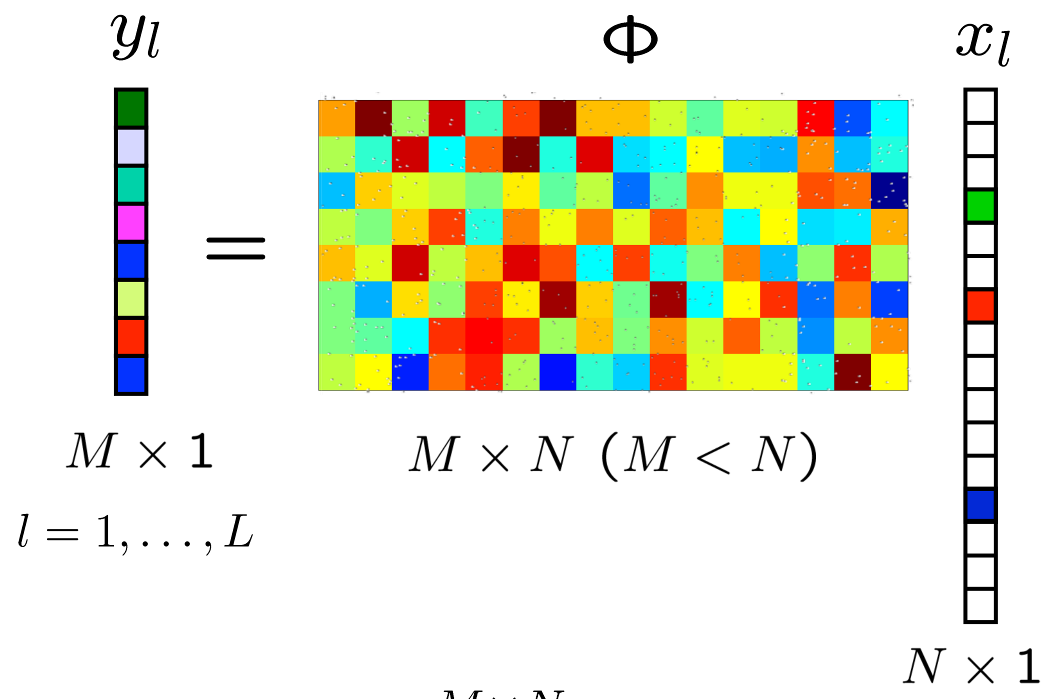
Learning to "Calibrate"

- Suppose we are given

$$\mathbf{Y} = [y_1, \ldots, y_L] \in R^{M \times L}$$

- Calibrate the sensing matrix via the basic bilinear model

$$\mathbf{Y} = \mathbf{\Phi}\mathbf{X} + \mathbf{W}$$



perturbed sensing matrix <> $\mathbf{\Phi} \in R^{M \times N}$

sparse signals <> $\mathbf{X} = [x_1, \ldots, x_L] \in R^{N \times L}$

model perturbations <> $\mathbf{W} \in R^{M \times L}$

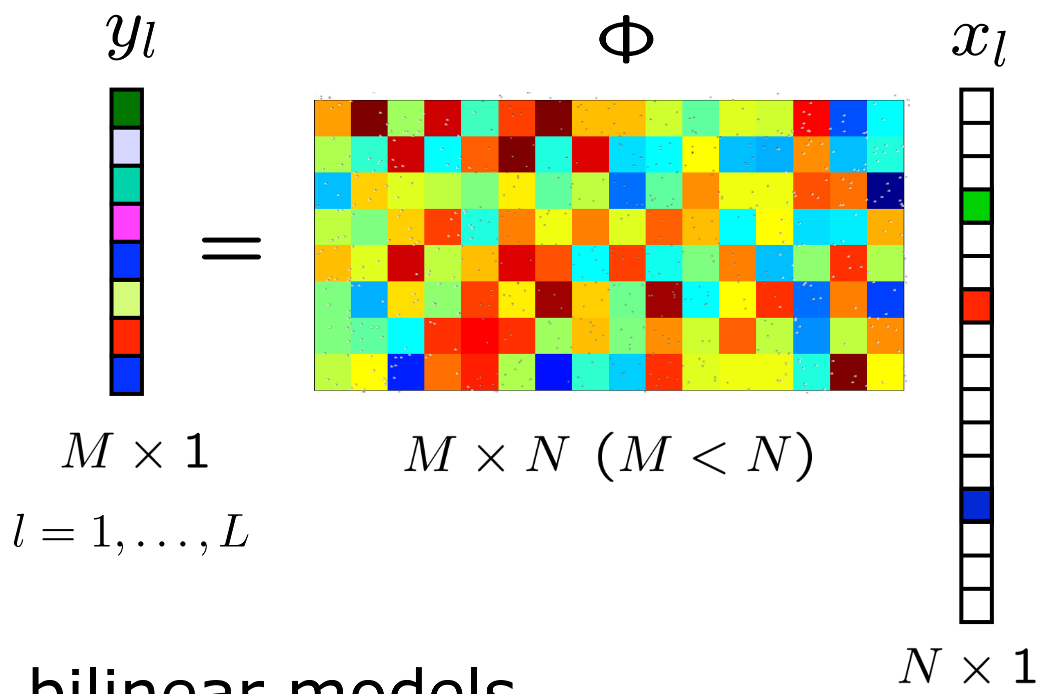
Learning to "Calibrate"

- Suppose we are given

$$\mathbf{Y} = [y_1, \ldots, y_L] \in R^{M \times L}$$

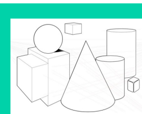
- Calibrate the sensing matrix via the basic bilinear model

$$\mathbf{Y} = \mathbf{\Phi X} + \mathbf{W}$$



- Applications of general bilinear models
  - Dictionary learning
  - Bayesian experimental design
  - Collaborative filtering
  - Matrix factorization / completion

Example Approaches

	Geometric 	**Combinatorial** $\binom{N}{K}$	**Probabilistic** 
Idea	atomic norm / TV regularization	non-convex (sparsity) constraints	Distributions on the space of unknowns
Example	$\min_{\Phi,X} \|\|Y - \Phi X\|\|^2 + \lambda \|\|X\|\|_1$	$\min_{\Phi,X:\|\|X\|\|_0 \leq K} \|\|Y - \Phi X\|\|^2$	Next few slides
Algorithm based on alternating minimization	Nesterov acceleration, Augmented Lagrangian, Bregman distance, DR splitting,…	Hard thresholding (IHT, CoSaMP, SP, ALPS, OMP) + SVD	Variational Bayes, Gibbs sampling, MCMC, Approximate message passing (AMP)
Literature	Mairal et al. 2008; Zhang & Chan 2009, …	Aharon et al. 2006 (KSVD); Rubinstein et al. 2009, …	Zhou et al. 2009; Donoho et al. 2009, …

Probabilistic View

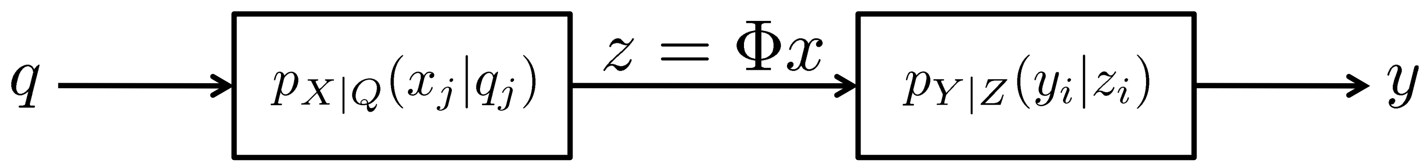
An Introduction to Approximate Message Passing (AMP)

Probabilistic Sparse Recovery: the AMP Way

- **Goal:** given $y = \Phi x + w$

  infer $E\{x|y\}$ and $\mathrm{var}\{x|y\}$

- **Model:**



Example: $p_{X|Q}(x|q = [\lambda, \hat{\theta}, \mu^{\theta}]) = \lambda \mathcal{N}(x; \hat{\theta}, \mu^{\theta}) + (1-\lambda)\delta(x), \quad \lambda \in [0,1]$

$$p_{Y|Z}(y|z) = \mathcal{N}(y; z, \mu^{w})$$

- **Approach:** graphical models / message passing
- **Key Ideas:** **CLT / Gaussian approximations**

[Boutros and Caire 2002; Montanari and Tse 2006; Guo and Wang 2006; Tanaka and Okada 2006; Donoho, Maleki, and Montanari 2009; Rangan 2010]

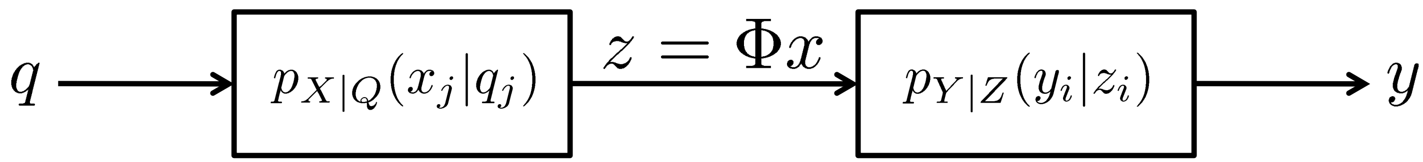
Probabilistic Sparse Recovery: the AMP Way

- **Goal:** given $y = \Phi x + w$

  infer $E\{x|y\}$ and $\mathrm{var}\{x|y\}$

"MAP estimation is so 1996." – Joel T.

- **Model:**



Example: $p_{X|Q}(x|q = [\lambda, \hat{\theta}, \mu^\theta]) = \lambda \mathcal{N}(x; \hat{\theta}, \mu^\theta) + (1-\lambda)\delta(x), \quad \lambda \in [0,1]$

$$p_{Y|Z}(y|z) = \mathcal{N}(y; z, \mu^w)$$

- **Approach:** graphical models / message passing
- **Key Ideas:** **CLT / Gaussian approximations**

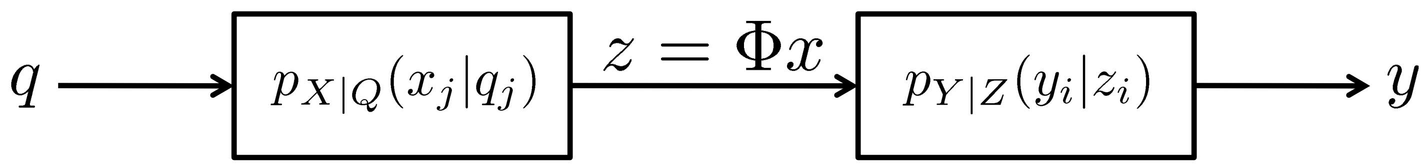
[Boutros and Caire 2002; Montanari and Tse 2006; Guo and Wang 2006; Tanaka and Okada 2006; Donoho, Maleki, and Montanari 2009; Rangan 2010]

Probabilistic Sparse Recovery: the AMP Way

- **Goal:** given $y = \Phi x + w$

  infer $E\{x|y\}$ and $\mathrm{var}\{x|y\}$

- **Model:**



Example: $p_{X|Q}(x|q = [\lambda, \hat{\theta}, \mu^\theta]) = \lambda \mathcal{N}(x; \hat{\theta}, \mu^\theta) + (1-\lambda)\delta(x), \quad \lambda \in [0,1]$

$$p_{Y|Z}(y|z) = \mathcal{N}(y; z, \mu^w)$$

- **Approach:** graphical models / message passing
- **Key Ideas:** **CLT / Gaussian approximations**

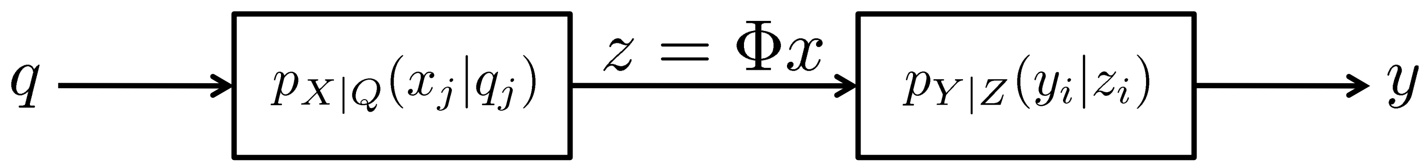
[Boutros and Caire 2002; Montanari and Tse 2006; Guo and Wang 2006; Tanaka and Okada 2006; Donoho, Maleki, and Montanari 2009; Rangan 2010]

Probabilistic Sparse Recovery: the AMP Way

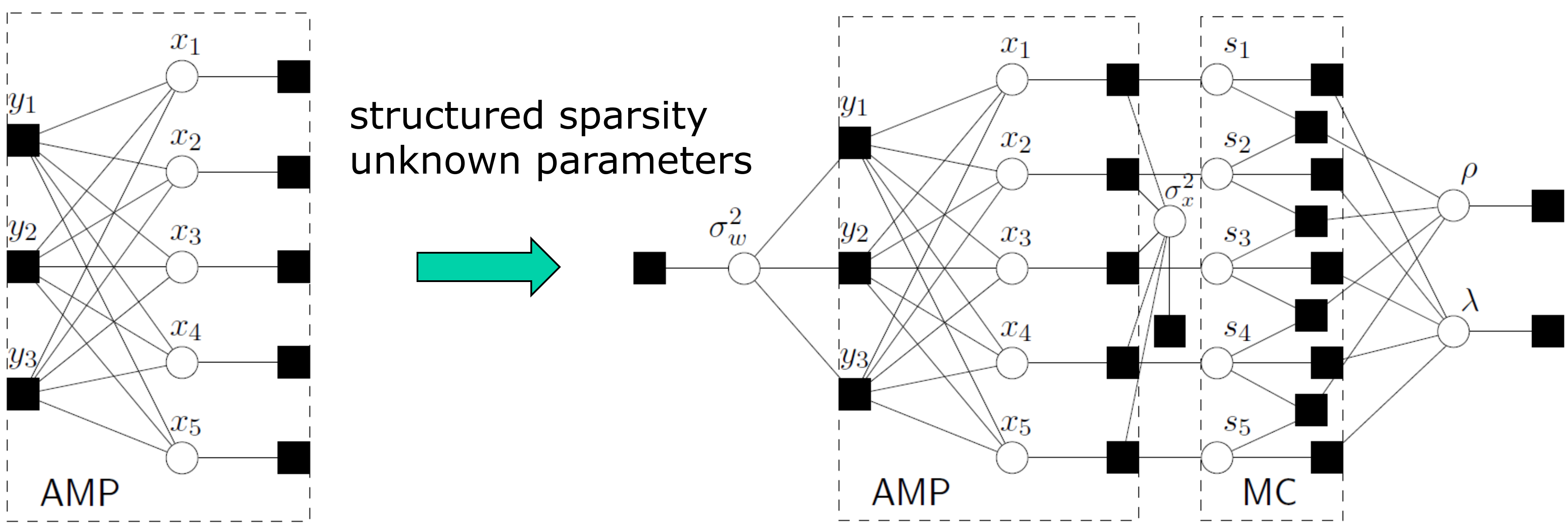
- **Goal:** given $y = \Phi x + w$

  infer $E\{x|y\}$ and $\text{var}\{x|y\}$

- **Model:**



- **Approach:** graphical models are modular!



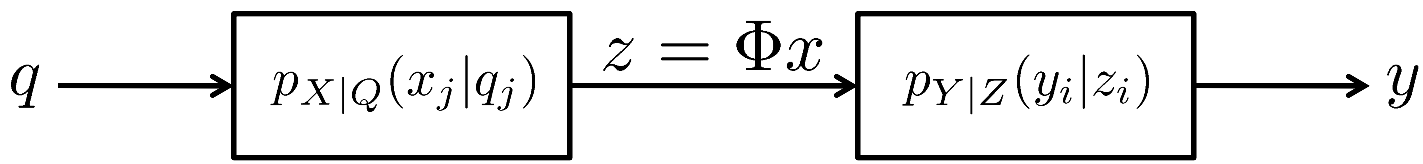
Probabilistic Sparse Recovery: the AMP Way

- **Goal:** given $y = \Phi x + w$

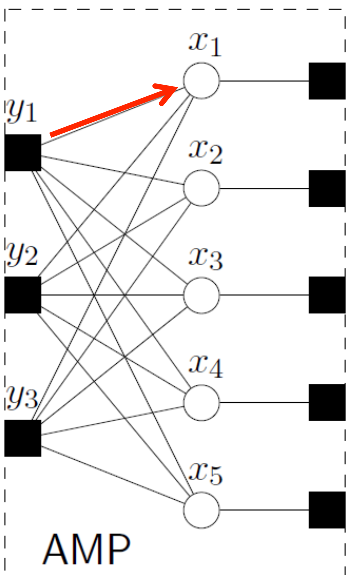
  infer $E\{x|y\}$ and $\mathrm{var}\{x|y\}$

- **Model:**

$p_{Y|Z}(y|z) = \mathcal{N}(y; z, \mu^w)$



- **Approximate message passing:** (sum-product)



$$
\begin{aligned}
m_{i\to j}(x_j) &= E_{x_r:r\neq j}\{p_{Y|Z}(y_i|z_i)|x_j, y_i\} \\
&= E_{x_r:r\neq j}\left\{p_{Y|Z}\left(y_i \middle| [\Phi]_{ij}x_j + \sum_{r\neq j}[\Phi]_{ir}x_r\right)|x_j, y_i\right\} \\
&\approx E_V\left\{p_{Y|Z}\left(y_i \,|[\Phi]_{ij}x_j + V\right)|x_j, y_i\right\} \\
&\approx d_1[\Phi]_{ij}x_j + d_2[\Phi]_{ij}x_j^2
\end{aligned}
$$

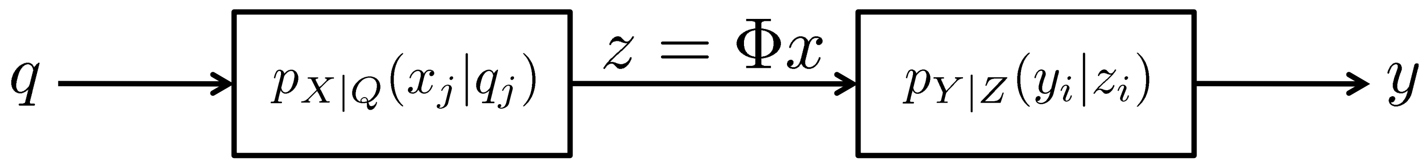
Probabilistic Sparse Recovery: the AMP Way

- **Goal:** given $y = \Phi x + w$

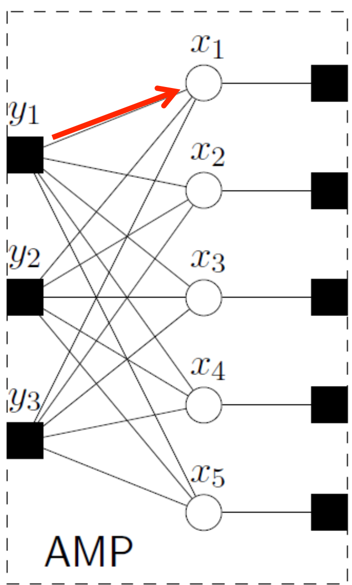
  infer $E\{x|y\}$ and $\mathrm{var}\{x|y\}$

- **Model:**

$p_{Y|Z}(y|z) = \mathcal{N}(y; z, \mu^w)$



- **Approximate message passing:** (sum-product)



$$m_{i\to j}(x_j) = E_{x_r:r\neq j}\{p_{Y|Z}(y_i|z_i)|x_j, y_i\}$$

$$= E_{x_r:r\neq j}\left\{p_{Y|Z}\left(y_i \middle| [\Phi]_{ij}x_j + \sum_{r\neq j}[\Phi]_{ir}x_r\right)|x_j, y_i\right\}$$

$$\approx \mathbf{E_V}\left\{\mathbf{p_{Y|Z}}\left(\mathbf{y_i}\,|[\boldsymbol{\Phi}]_{\mathbf{ij}}\mathbf{x_j} + \mathbf{V}\right)|\mathbf{x_j}, \mathbf{y_i}\right\} \qquad V \sim \mathcal{N}$$

$$\approx d_1[\Phi]_{ij}x_j + d_2[\Phi]_{ij}x_j^2$$

Central limit theorem (blessing-of-dimensionality)

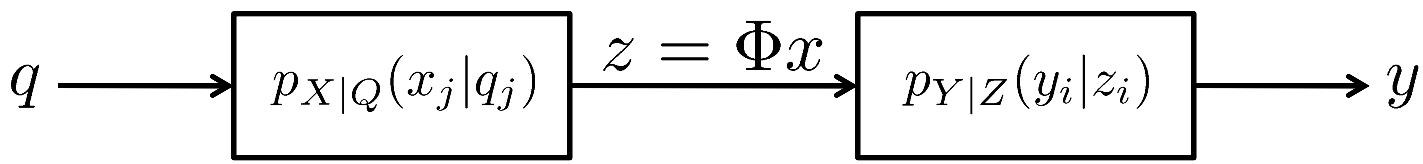
Probabilistic Sparse Recovery: the AMP Way

- **Goal:** given $y = \Phi x + w$

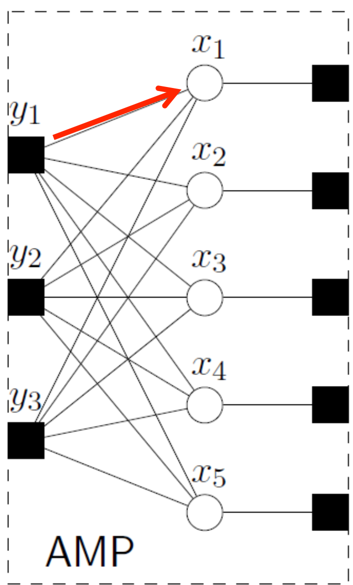
  infer $E\{x|y\}$ and $\mathrm{var}\{x|y\}$

- **Model:**

$p_{Y|Z}(y|z) = \mathcal{N}(y; z, \mu^w)$



- **Approximate message passing:** (sum-product)



$$m_{i\to j}(x_j) = E_{x_r:r\neq j}\{p_{Y|Z}(y_i|z_i)|x_j, y_i\}$$

$$= E_{x_r:r\neq j}\left\{p_{Y|Z}\left(y_i\,\middle|\,[\Phi]_{ij}x_j + \sum_{r\neq j}[\Phi]_{ir}x_r\right)|x_j, y_i\right\}$$

$$\approx E_V\left\{p_{Y|Z}\left(y_i\,|[\Phi]_{ij}x_j + V\right)|x_j, y_i\right\}$$

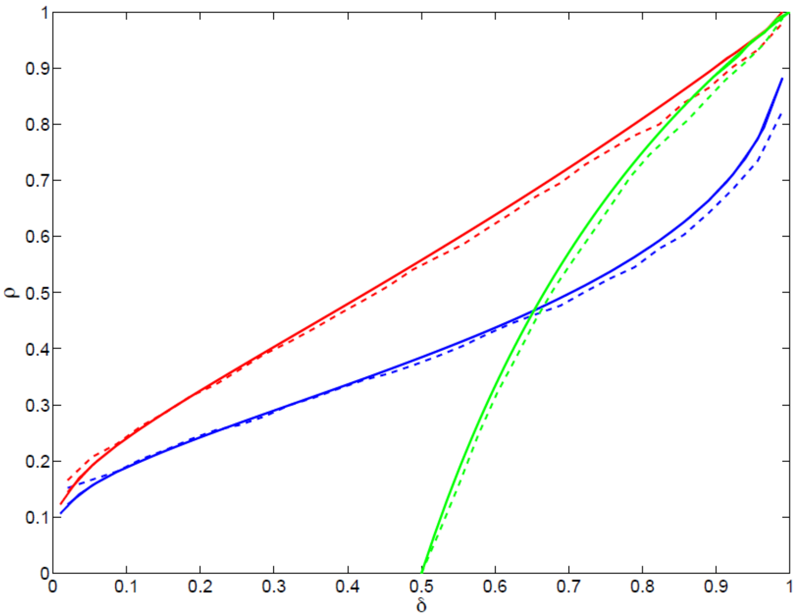
$$\approx \mathbf{d_1[\Phi]_{ij}x_j + d_2[\Phi]_{ij}x_j^2}$$

Taylor series approximation

AMP Performance

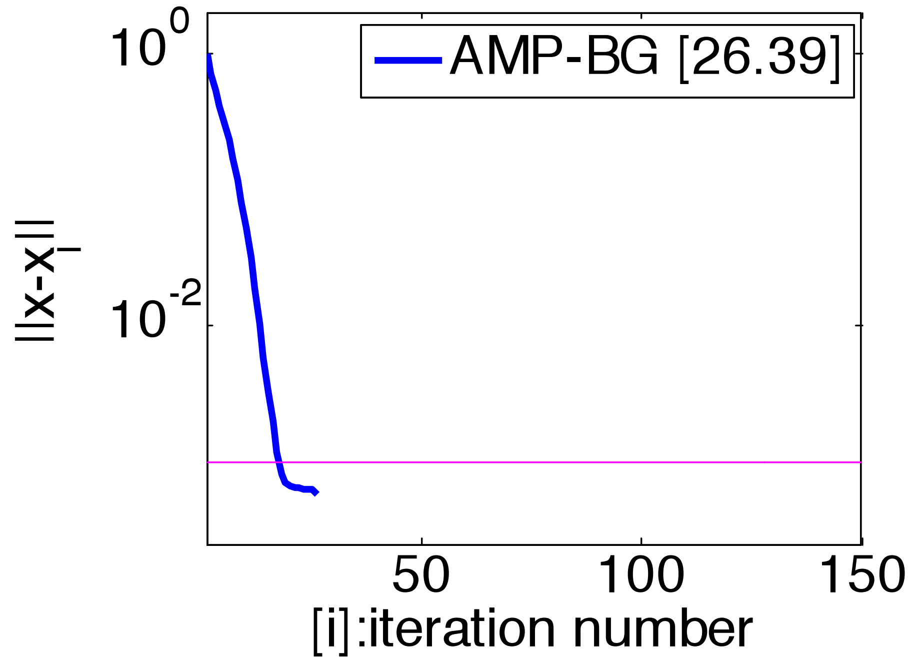
[Donoho, Maleki, and Montanari 2009; Bayati and Montanari 2011; Montanari 2010; Rangan 2010; Schniter 2010]

Phase transition (Laplace prior)



AMP <> state evolution theory
<> fully distributed

Convergence (BG prior)



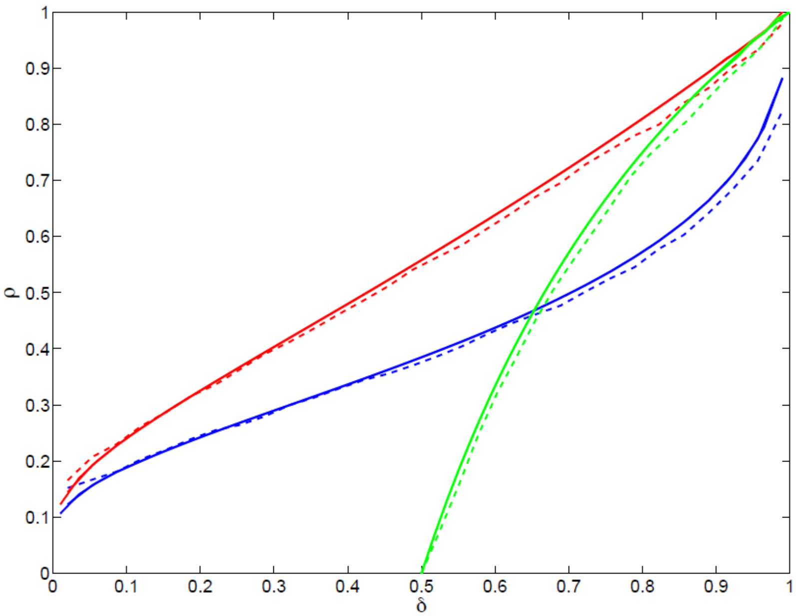
Gaussian matrix

Rangan's GAMP codes are available
http://gampmatlab.sourceforge.net

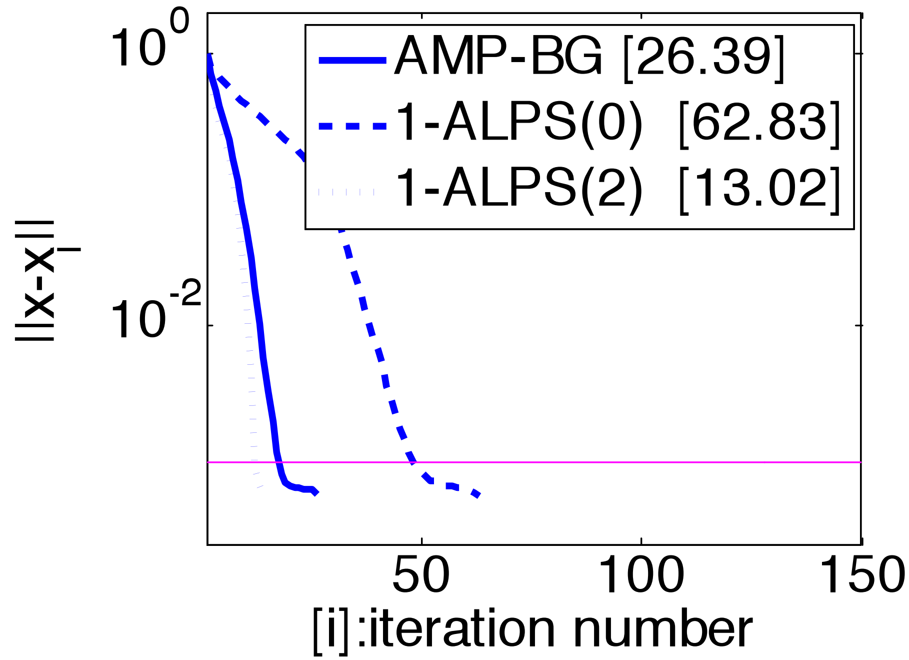
AMP Performance

[Donoho, Maleki, and Montanari 2009; Bayati and Montanari 2011; Montanari 2010; Rangan 2010; Schniter 2010]

Phase transition (Laplace prior)



Convergence (BG prior)



Gaussian matrix

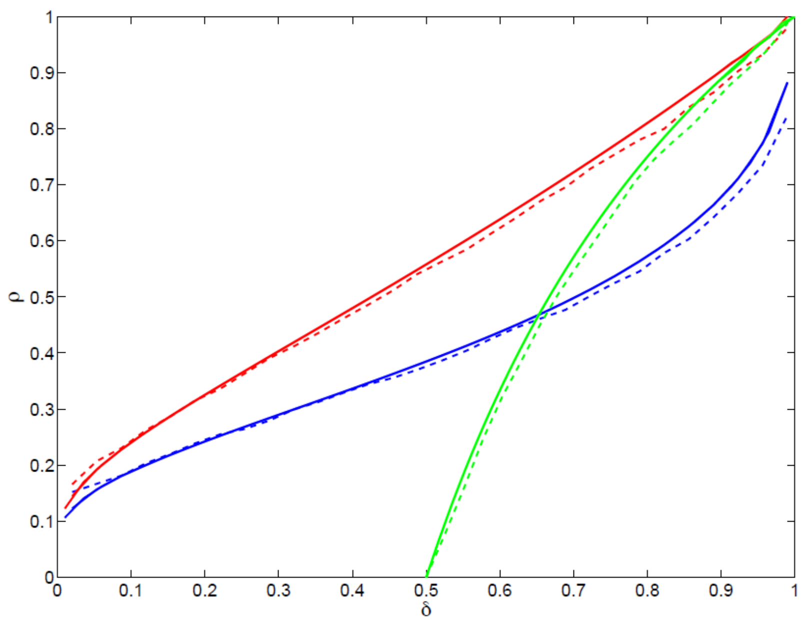
AMP <> state evolution theory
<> fully distributed

http://gampmatlab.sourceforge.net
http://lions.epfl.ch/ALPS/download

AMP Performance

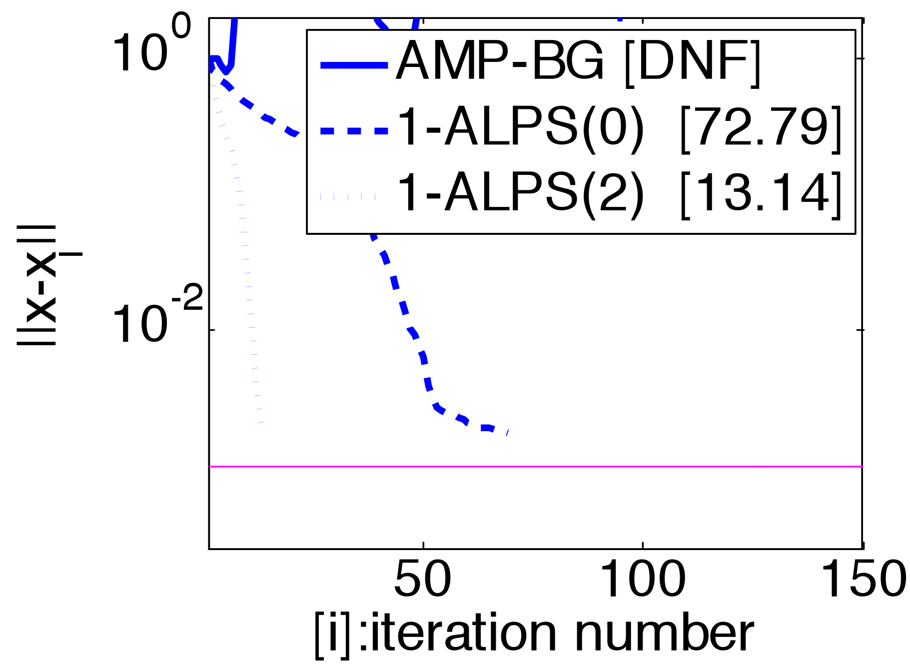
[Donoho, Maleki, and Montanari 2009; Bayati and Montanari 2011; Montanari 2010; Rangan 2010; Schniter 2010]

Phase transition (Laplace prior)



AMP <> state evolution theory
<> fully distributed

Convergence (BG prior)



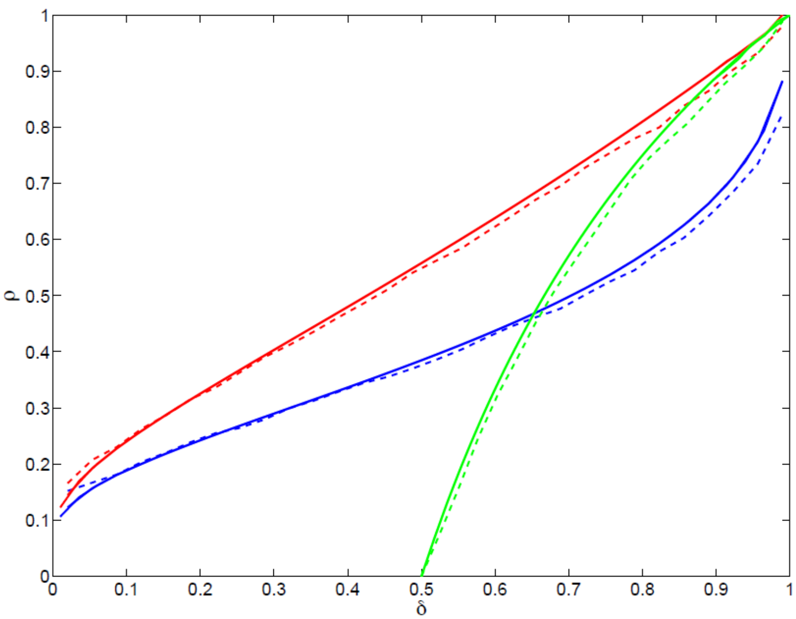
sparse matrix
(model mismatch)

http://gampmatlab.sourceforge.net
http://lions.epfl.ch/ALPS/download

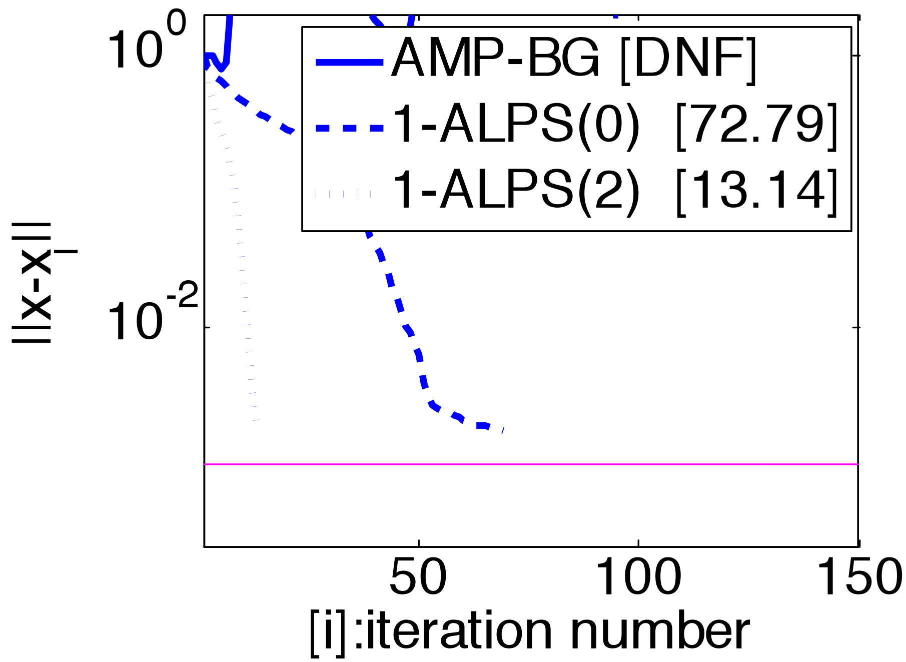
AMP Performance

[Donoho, Maleki, and Montanari 2009; Bayati and Montanari 2011; Montanari 2010; Rangan 2010; Schniter 2010]

Phase transition (Laplace prior)



Convergence (BG prior)



AMP <> state evolution theory
<> fully distributed

sparse matrix
(model mismatch)

Need to switch to normal BP when the matrix is sparse

http://gampmatlab.sourceforge.net
http://lions.epfl.ch/ALPS/download

Bilinear AMP

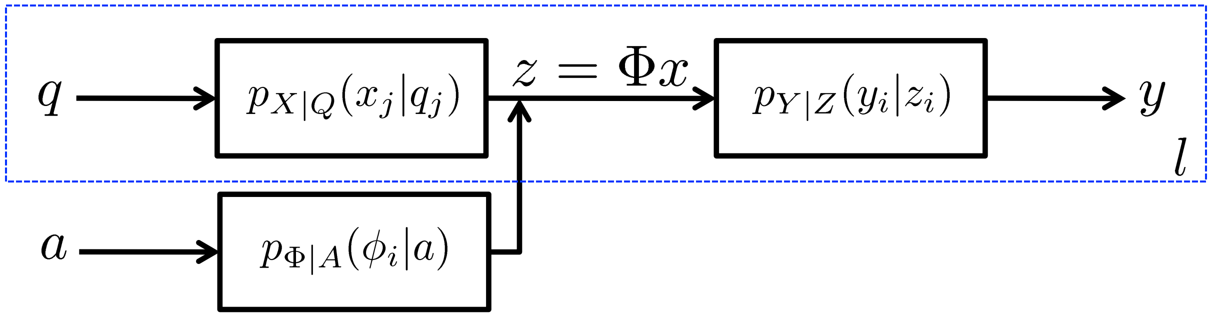
Bilinear AMP

- **Goal:** given $\mathbf{Y} = \mathbf{\Phi}\mathbf{X} + \mathbf{W}$

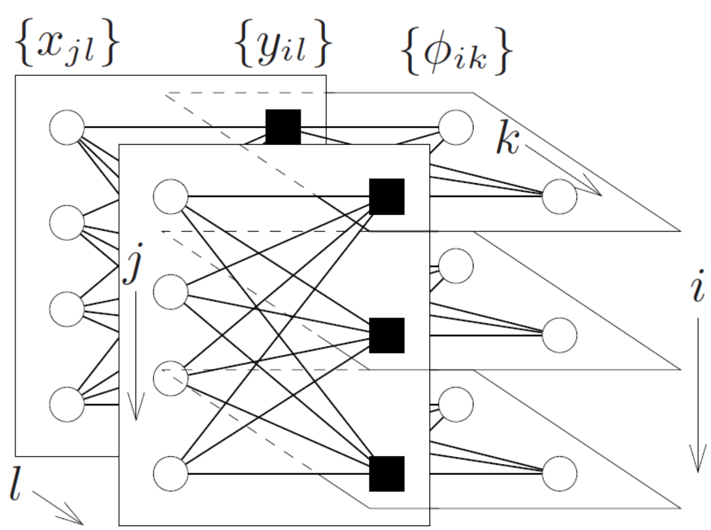
  infer $E\{x_l|\mathbf{Y}\}$ and $\text{var}\{x_{jl}|\mathbf{Y}\}\ \forall l, j$

  $E\{\mathbf{\Phi}|\mathbf{Y}\}$ and $\text{var}\{[\mathbf{\Phi}]_{i,j}|\mathbf{Y}\}\ \forall i, j$

- **Model:**

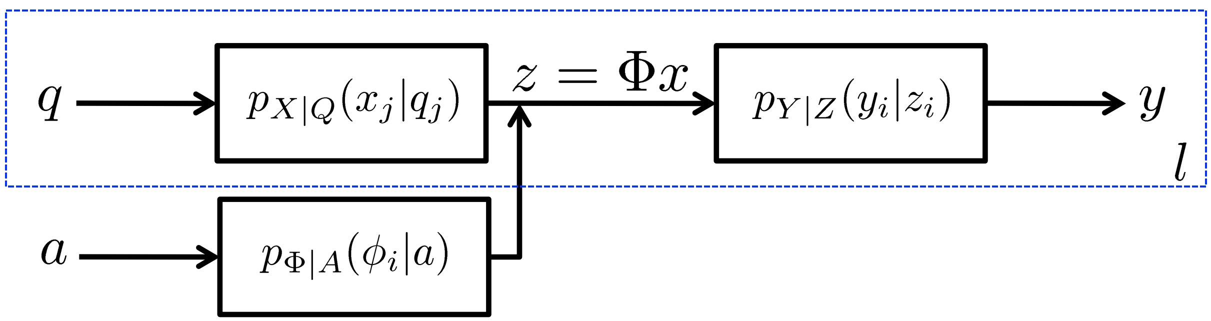


- **Algorithm:** graphical model



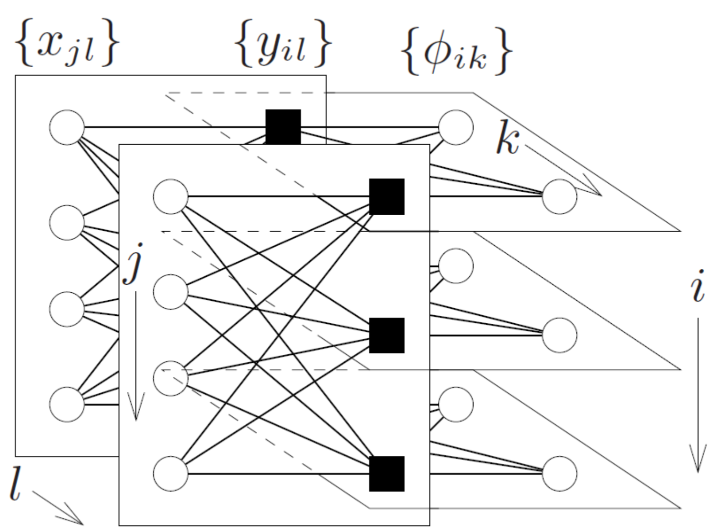
Bilinear AMP

- **Goal:** given $\mathbf{Y} = \mathbf{\Phi}\mathbf{X} + \mathbf{W}$
  infer $E\{x_l|\mathbf{Y}\}$ and $\text{var}\{x_{jl}|\mathbf{Y}\}\ \forall l, j$
  $E\{\mathbf{\Phi}|\mathbf{Y}\}$ and $\text{var}\{[\mathbf{\Phi}]_{i,j}|\mathbf{Y}\}\ \forall i, j$
- **Model:**



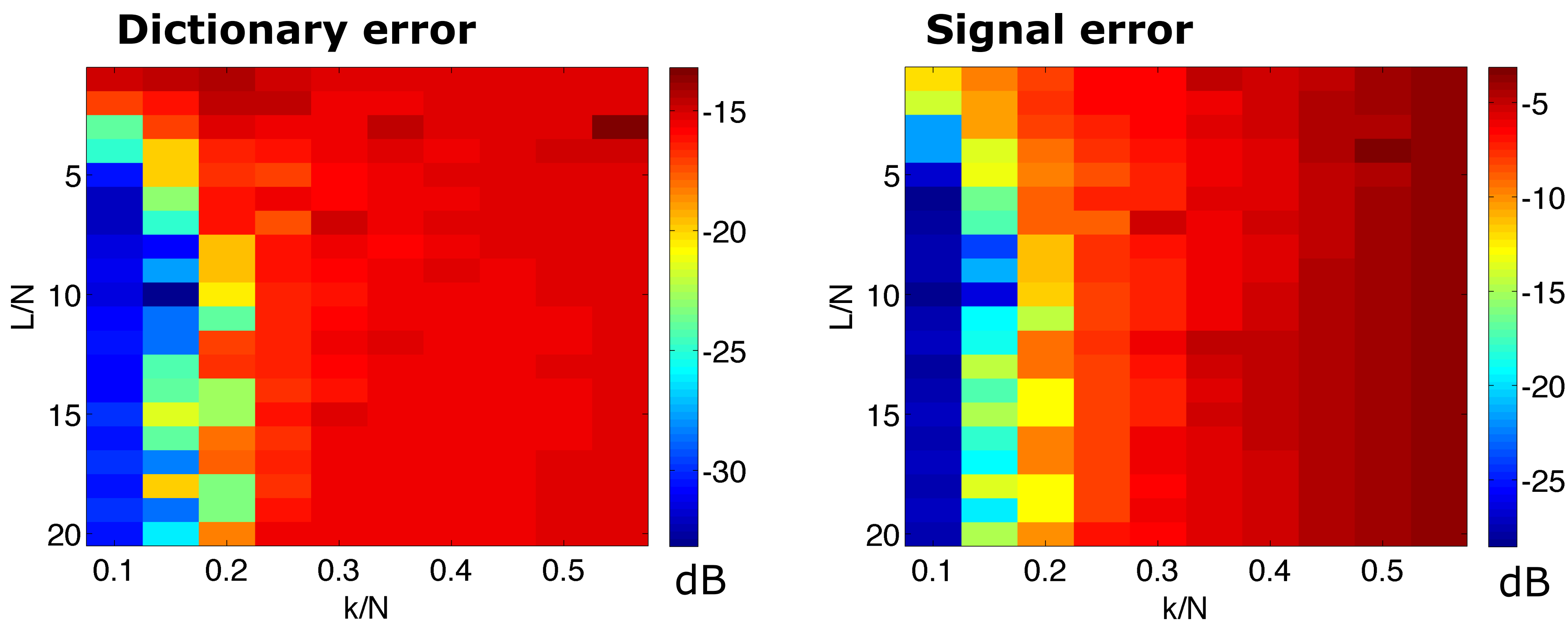
- **Algorithm:** (G)AMP with matrix uncertainty

$$
\begin{aligned}
\mu^r_j(t) &= \left(\sum_{i=1}^{m} |\hat{a}_{ij}|^2 \mu^s_i(t)\right)^{-1} \\
\hat{r}_j(t) &= \hat{x}_j(t) + \mu^r_j(t) \sum_{i=1}^{m} \hat{a}^*_{ij} \hat{s}_i(t) \\
\hat{x}_j(t+1) &= g_{in}(\hat{r}_j(t), q_j, \mu^r_j(t)) \\
\mu^x_j(t+1) &= \mu^r_j(t) g'_{in}(\hat{r}_j(t), q_j, \mu^r_j(t))
\end{aligned}
$$

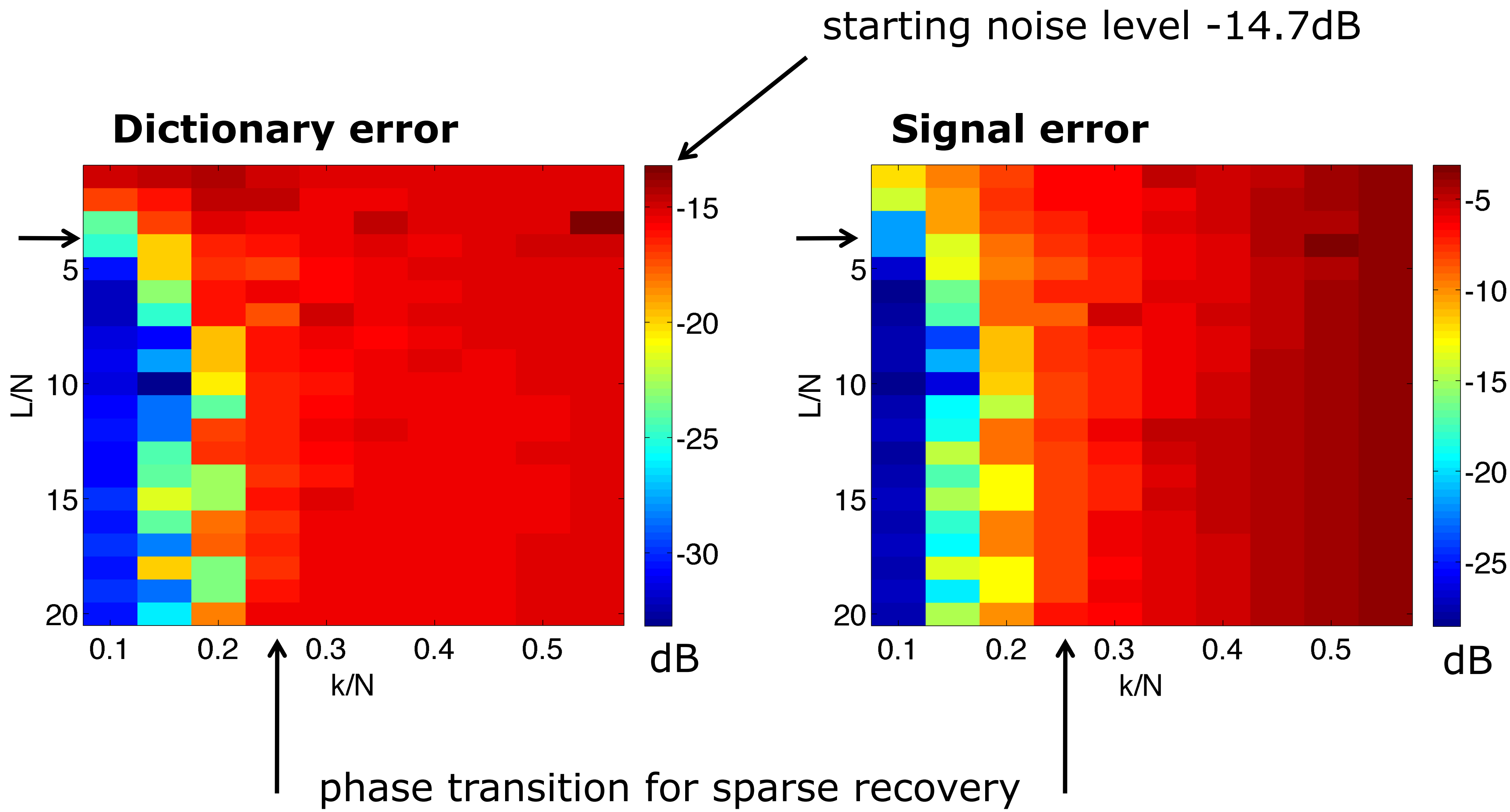


$$
\begin{aligned}
\mu^p_i(t) &= \sum_{j=1}^{n} (|\hat{a}_{ij}|^2 + \mu^a_{ij}) \mu^x_j(t) + \mu^a_{ij} |\hat{x}_j(t)|^2 \\
\hat{z}_i(t) &= \sum_{j=1}^{n} \hat{a}_{ij} \hat{x}_j(t) \\
\hat{p}_i(t) &= \hat{z}_i(t) - \hat{s}_i(t-1) \sum_{j=1}^{n} |\hat{a}_{ij}|^2 \mu^x_j(t) \\
\hat{s}_i(t) &= g_{out}(\hat{p}_i(t), y_i, \mu^p_i(t)) \\
\mu^s_i(t) &= -g'_{out}(\hat{p}_i(t), y_i, \mu^p_i(t))
\end{aligned}
$$

Synthetic Calibration Example



Synthetic Calibration Example



phase transition for sparse recovery

k/N= 0.2632

Conclusions

- (G)AMP <> modular / scalable
  asymptotic analysis
- Extension to bilinear models
  - by products: "robust" compressive sensing
    structured sparse dictionary learning
    double-sparsity dictionary learning
- More to do comparison with other algorithms
- Weakness CLT: dense vs. sparse matrix
- Our codes will be available soon http://lions.epfl.ch

- Postdoc position @ LIONS / EPFL
  contact: volkan.cevher@epfl.ch