Compressed-Sensing Recovery with Super-Sparse Measurement Matrices

Joint Workshop on Communications and Coding (JWCC)
Kühtai, Austria, March 2019

Norbert Goertz & Stefan Birgmeier

institute of telecommunications

TU WIEN

Overview

- Compressed Sensing Basics and Practical Relevance
- Bayesian Approximate Message Passing (BAMP) for Compressed Sensing Recovery
- Z-BAMP for Matrices with Zero-Elements
- Design of Sparse, Lifted Measurement Matrices
- Specialization of Z-BAMP for use with Lifted Matrices
- Simulation Results

Compressed Sensing Basics

- Given are the linear measurements

$$\mathbf{y}_{[m]} = \mathbf{A}_{[m \times n]} \cdot \mathbf{x}_{[n]} + \mathbf{w}_{[m]} \quad \text{with} \quad m < n \tag{1}$$

  with $\mathbf{A}$ the known $m \times n$ sensing matrix and $\mathbf{w}$ the $m \times 1$ Gaussian measurement noise vector

  ➡ recover vector $\mathbf{x}$ from the measurements $\mathbf{y}$ given $\mathbf{A}$ and knowledge about the structure of $\mathbf{x}$

- Use sampling rate m/n appropriate for the **structure within the signal**
- Typical structure: few components of $\mathbf{x}$ non-zero ➡ sparsity
- More complicated/general "structure" can be exploited by a prior pdf used for recovery
- Classical solution of recovery problem: L_1-norm minimization by convex optimization
- Greedy schemes such as Orthogonal Matching Pursuit and Hard/Soft Thresholding better for larger dimensions

- Much better and much more efficient solution by Bayesian Approximate Message Passing (BAMP), as signal prior pdf can be exploited

Practical Relevance of Compressed Sensing

Systematic approach for many problems in **Signal Processing & Communications**

- source coding of structured (sparse) signals: sampling rates far below the classical theorem
- de-noising (images, audio, speech)
- inpainting (images, audio)
- blind deconvolution, e.g. de-reverberation
- radar signal processing (target detection)
- channel estimation (multipath propagation with few dominant paths)
- signal detection in M2M communication (sparse user activity patterns)
- joint channel estimation and data detection (recent bilinear approaches)
- RFID tag detection
- Single-Pixel Imaging: CS-based THz camera (airport body scanners)

➡ For more applications (and lots of theory), see (long!) list at `http://dsp.rice.edu/cs`

↝ **Key problem: complexity for "big problems" ➡ iterative recovery such as BAMP**

Compressed-Sensing Recovery as an Estimation Problem

- Noisy Compressed Sensing:

$$\mathbf{y}_{[m]} = \mathbf{A}_{[m\times n]} \cdot \mathbf{x}_{[n]} + \mathbf{w}_{[m]} \quad \text{with} \quad m < n$$

  - **known**: observation vector $\mathbf{y}$, measurement matrix $\mathbf{A}$ (full rank m assumed)
  - random measurement noise vector $\mathbf{w}$; variances $\sigma^2_{w_q}$ of noise samples w_q, $q = 1, \ldots, m$ known
  - **to be recovered**: signal vector $\mathbf{x}$; pdf assumed to be known

- Goal: find estimate $\hat{\mathbf{x}}$ that minimizes MSE:

$$\hat{\mathbf{x}} = \arg\min_{\tilde{\mathbf{x}}} \mathbb{E}_{\mathbf{X}}\left\{ \|\mathbf{X} - \tilde{\mathbf{x}}\|_2^2 \middle| \mathbf{Y} = \mathbf{y} \right\}$$

➡ well-known solution from estimation theory: conditional expectation, given observations $\mathbf{y}$:

$$\hat{\mathbf{x}} = \mathbb{E}_{\mathbf{X}}\left\{ \mathbf{X} \middle| \mathbf{Y} = \mathbf{y} \right\}$$

➡ Bayes' rule:

$$\hat{\mathbf{x}} = \int_{\mathbb{R}^n} \mathbf{x}\, p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) d\mathbf{x} = \frac{1}{p_{\mathbf{Y}}(\mathbf{y})} \int_{\mathbb{R}^n} \mathbf{x} \underbrace{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})}_{\text{measurement noise}} \underbrace{p_{\mathbf{X}}(\mathbf{x})}_{\text{signal prior}} d\mathbf{x} \qquad (2)$$

➡ Impossible to realize in practice!

Bayesian Approximate Message Passing (BAMP)

- Approximate solution of (2): "Loopy" Message Passing
  - ➡ based on a graphical model of the measurement process
- Derivation in Donoho *et al.* (2009), Donoho *et al.* (2011), Donoho *et al.* (2010a), Donoho *et al.* (2010b) and Eldar & Kutyniok (2012, Chapter 9), Montanari (2011).
- Formal proof in Bayati & Montanari (2011)
- Derivation is complicated and contains several assumptions, including large m, n
- In what follows Z-BAMP: derived in Birgmeier & Goertz (2018); Goertz & Birgmeier (2019) from intermediate step of derivation of the conventional BAMP scheme
  - Measurement matrix may contain (possibly many) zero-elements
  - Measurement noise variances $\sigma_{w_q}^2$ may be different for all measurements y_q, $q = 1, 2, \ldots, m$
  - Unit Euclidean norms of columns of the measurement matrix $\mathbf{A}$ assumed

Z-BAMP for Measurement Matrices with Zero-Elements

- Sets of indices i of signal components x_i that are connected to the measurements $y_q, q = 1, 2, \ldots, m$ by the components A_{qi} of the measurement matrix $\mathbf{A}$:

$$\mathcal{N}_q \doteq \{i \in \{1, 2, \ldots, n\} : A_{qi} \neq 0\} \tag{3}$$

- Sets of the indices q of measurements y_q that take measurements of the signal components $x_i, i = 1, 2, \ldots, n$:

$$\mathcal{M}_i \doteq \{q \in \{1, 2, \ldots, m\} : A_{qi} \neq 0\} \tag{4}$$

- Integer $m_i = |\mathcal{M}_i|$ is the number of elements in the set $\mathcal{M}_i$.
- Columns normalized to unit Euclidean norm: approximation $A_{qi}^2 \approx \frac{1}{m_i}$ for $q \in \mathcal{M}_i$ and all i.
- Start Z-BAMP at iteration $t = 0$ with the initializations (prior pdf used):

$$\hat{x}_i^0 = \mu_{X_i} = \int_{-\infty}^{\infty} x_i \, p_{X_i}(x_i) dx_i \quad i = 1, 2, \ldots, n \tag{5}$$

$$\sigma_i^{2(0)} = \int_{-\infty}^{\infty} (x_i - \mu_{X_i})^2 \, p_{X_i}(x_i) dx_i \quad i = 1, 2, \ldots, n \tag{6}$$

$$z_q^0 = y_q, \quad q = 1, 2, \ldots, m \tag{7}$$

... Z-BAMP for Measurement Matrices with Zero-Elements: Key equations

- Compute total noise variance at measurement nodes $q = 1, \ldots, m$

$$S_q^{t-1} = \sigma_{w_q}^2 + \sum_{i \in \mathcal{N}_q} \frac{1}{|\mathcal{M}_i|} \sigma_i^{2(t-1)} \quad \mathcal{N}_q\text{: signal components } i \text{ connected to measurement } q \tag{8}$$

- Compute effective noise variance c_i and substitute measurement u_i for variable nodes $i = 1, \ldots, n$

$$\frac{1}{c_i^{t-1}} = \frac{1}{|\mathcal{M}_i|} \sum_{q \in \mathcal{M}_i} \frac{1}{S_q^{t-1}} \qquad \mathcal{M}_i\text{: measurements } q \text{ connected to signal component } i \tag{9}$$

$$u_i^{t-1} = c_i^{t-1} \sum_{q \in \mathcal{M}_i} A_{qi} \frac{1}{S_q^{t-1}} z_q^{t-1} + \hat{x}_i^{t-1} \tag{10}$$

- Compute the signal-component estimates and the error variances

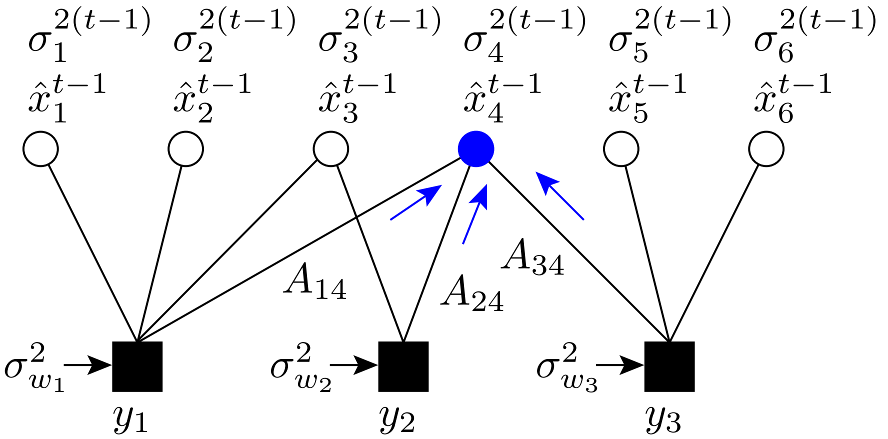
$$\hat{x}_i^t = F_i\left(u_i^{t-1}, c_i^{t-1}\right) \qquad \sigma_i^{2(t)} = c_i^{t-1} F_i'\left(u_i^{t-1}, c_i^{t-1}\right) \tag{11}$$

- Compute residual z_q for the measurement nodes $q = 1, \ldots, m$

$$z_q^t = y_q - \sum_{i \in \mathcal{N}_q} A_{qi} \hat{x}_i^t + \frac{z_q^{t-1}}{S_q^{t-1}} \sum_{i \in \mathcal{N}_q} \frac{1}{|\mathcal{M}_i|} \sigma_i^{2(t)} \tag{12}$$

- Stop, if $\sum_{i=1}^n \left(\hat{x}_i^t - \hat{x}_i^{t-1}\right)^2 < \varepsilon^2 \sum_{i=1}^n \left(\hat{x}_i^t\right)^2$; otherwise, continue with (8) for $t := t + 1$.

factor–to–variable messages



$$S_1^{t-1} = \sigma_{w_1}^2 + \left(\frac{\sigma_1^{2(t-1)}}{m_1} + \frac{\sigma_2^{2(t-1)}}{m_2} + \frac{\sigma_3^{2(t-1)}}{m_3} + \frac{\sigma_4^{2(t-1)}}{m_4}\right)$$

$$S_2^{t-1} = \sigma_{w_2}^2 + \left(\frac{\sigma_3^{2(t-1)}}{m_3} + \frac{\sigma_4^{2(t-1)}}{m_4}\right)$$

$$S_3^{t-1} = \sigma_{w_3}^2 + \left(\frac{\sigma_4^{2(t-1)}}{m_4} + \frac{\sigma_5^{2(t-1)}}{m_5} + \frac{\sigma_6^{2(t-1)}}{m_6}\right)$$

Due to column–normalization:

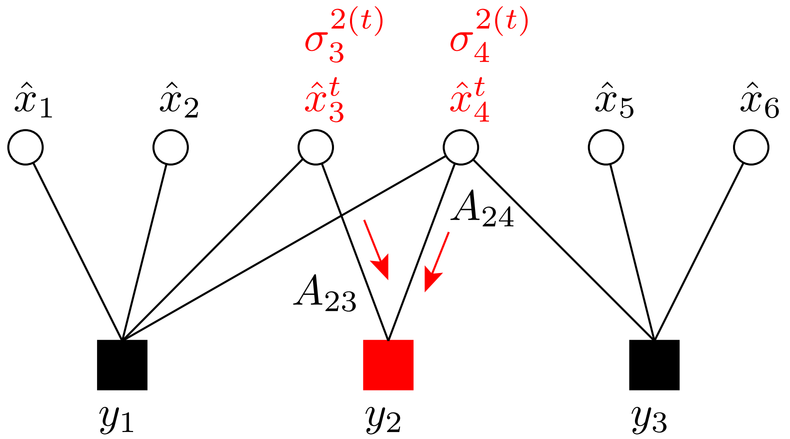
$$|A_{14}| = |A_{24}| = |A_{34}| = \frac{1}{\sqrt{m_4}} = \frac{1}{\sqrt{3}}$$

$$\frac{1}{c_4^{t-1}} = \frac{1}{m_4}\left(\frac{1}{S_1^{t-1}} + \frac{1}{S_2^{t-1}} + \frac{1}{S_3^{t-1}}\right)$$

$$u_4^{t-1} = c_4^{t-1}\left(\frac{A_{14}}{S_1^{t-1}} z_1^{t-1} + \frac{A_{24}}{S_2^{t-1}} z_2^{t-1} + \frac{A_{34}}{S_3^{t-1}} z_3^{t-1}\right) + \hat{x}_4^{t-1}$$

$$\hat{x}_4^t = F_4\left(u_4^{t-1}, c_4^{t-1}\right) \qquad \sigma_4^{2(t)} = c_4^{t-1}\, F_4'\left(u_4^{t-1}, c_4^{t-1}\right)$$

variable–to–factor messages



$$z_2^t = y_2 - (A_{23}\hat{x}_3^t + A_{24}\hat{x}_4^t) + \frac{z_2^{t-1}}{S_2^{t-1}}\left(\frac{\sigma_3^{2(t)}}{m_3} + \frac{\sigma_4^{2(t)}}{m_4}\right)$$

Due to column–normalization:

$$|A_{23}| = \frac{1}{\sqrt{m_3}} = \frac{1}{\sqrt{2}}$$

$$|A_{24}| = \frac{1}{\sqrt{m_4}} = \frac{1}{\sqrt{3}}$$

m_i: number of measurement nodes connected to variable node i

... Z-BAMP for Measurement Matrices with Zero-Elements: Denoisers

- BAMP algorithm contains scalar operators:

$$F_i\left(u_i; c_i\right) = \mathbb{E}_{X_i}\left\{X_i | U_i = u_i\right\} \tag{13}$$

$$F_i'\left(u_i; c_i\right) = \frac{d}{du_i} F_i\left(u_i; c_i\right) \ , \tag{14}$$

- Variance of the estimation error under zero-mean Gaussian noise of effective variance c_i:

$$\sigma_i^2 = \mathbb{E}_{X_i}\left\{\left(X_i - F_i\left(u_i; c_i\right)\right)^2 | U_i = u_i\right\} = c_i \ F_i'\left(u_i; c_i\right) \ .$$

- Operators $F_i()$ and $F_i'()$ computed using posterior pdf

$$p_{X_i|U_i}(x_i|u_i; c_i) = \frac{p_{U_i|X_i}(u_i|x_i; c_i) \ p_{X_i}(x_i)}{p_{U_i}(u_i; c_i)} \quad i = 1, 2, \ldots, n \tag{15}$$

where

$$p_{U_i|X_i}(u_i|x_i; c_i) = \frac{1}{\sqrt{2\pi c_i}} \exp\left(-\frac{1}{2c_i}(x_i - u_i)^2\right) \tag{16}$$

➡ effective variance c_i of Gaussian distribution computed during BAMP iterations in (9)

➡ $p_{X_i}(x_i)$ is the pdf of the signal prior (independent components)

... Z-BAMP for Measurement Matrices with Zero-Elements: Denoiser for BG-Source

- Estimator functions for a Bernoulli-Gaussian (sparse Gaussian) signal prior

$$p_{X_i}(x_i) = \gamma_i \delta(x_i) + (1-\gamma_i)\mathcal{N}(x_i;\ \mu_{X_i}, \sigma^2_{X_i}) \tag{17}$$

  - γ_i: probability of a zero-component; $\sigma^2_{X_i}$: variance of non-zero Gaussian signal part
  - Mean of signal: $\mu_{X_i} = 0$ (used for Z-BAMP-initialization in (5))
  - Variance of the signal X_i: $\sigma_i^{2(0)} = (1-\gamma_i)\sigma^2_{X_i}$ (used for Z-BAMP-initialization in (6))
  - $\mathcal{N}(x_i; \mu_{X_i}, \sigma^2_{X_i})$: value a Gaussian pdf takes when evaluated at x_i

Estimator function (denoiser):

$$F_i(u_i; c_i) = \ u_i\, M(u_i, \gamma_i, c_i, \alpha_i)\ , \tag{18}$$

Variance of the estimation error:

$$\sigma_i^2 = c_i\, M(u_i, \gamma_i, c_i, \alpha_i)\ +\ Q(u_i, \gamma_i, c_i, \alpha_i)\, F_i^2(u_i; c_i) \tag{19}$$

$$M(u_i, \gamma_i, c_i, \alpha_i) = \frac{\alpha_i}{\alpha_i + 1} \frac{1}{1\ +\ Q(u_i, \gamma_i, c_i, \alpha_i)} \tag{20}$$

$$Q(u_i, \gamma_i, c_i, \alpha_i) = \frac{\gamma_i}{1-\gamma_i}\sqrt{\alpha_i + 1}\, e^{-\frac{u_i^2}{2c_i}\frac{\alpha_i}{\alpha_i+1}} \quad \text{where} \quad \alpha_i = \sigma^2_{X_i}/c_i \tag{21}$$

Substitute Decoupled Measurement Model

Remark: Gaussian pdf (16) for the "effective" noise applies for $i = 1, 2, \ldots, n$ but we have only $m < n$ measurements

➡ equivalent *decoupled* measurement model in n dimensions

$$u_i = x_i + \tilde{w}_i \text{ with } i = 1, 2, \ldots, n \text{ and } \tilde{w}_i \sim \mathcal{N}(0, c_i) \,, \tag{22}$$

instead of the coupled model in $m < n$ dimensions given by $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$.

➡ $F()$ in BAMP is a scalar(!) denoiser to reduce the effect of the Gaussian noise $\tilde{w}_j$

- proven in Bayati & Montanari (2011), Donoho *et al.* (2011) and Eldar & Kutyniok (2012, Chapter 9), Montanari (2011) to apply asymptotically for large dimension n
- independence of the n coordinates and the validity of the Gaussian model for $\tilde{w}_j$ in (22) is achieved by the **Onsager-term** in (12) (see Montanari (2011)).

Measurement Matrices

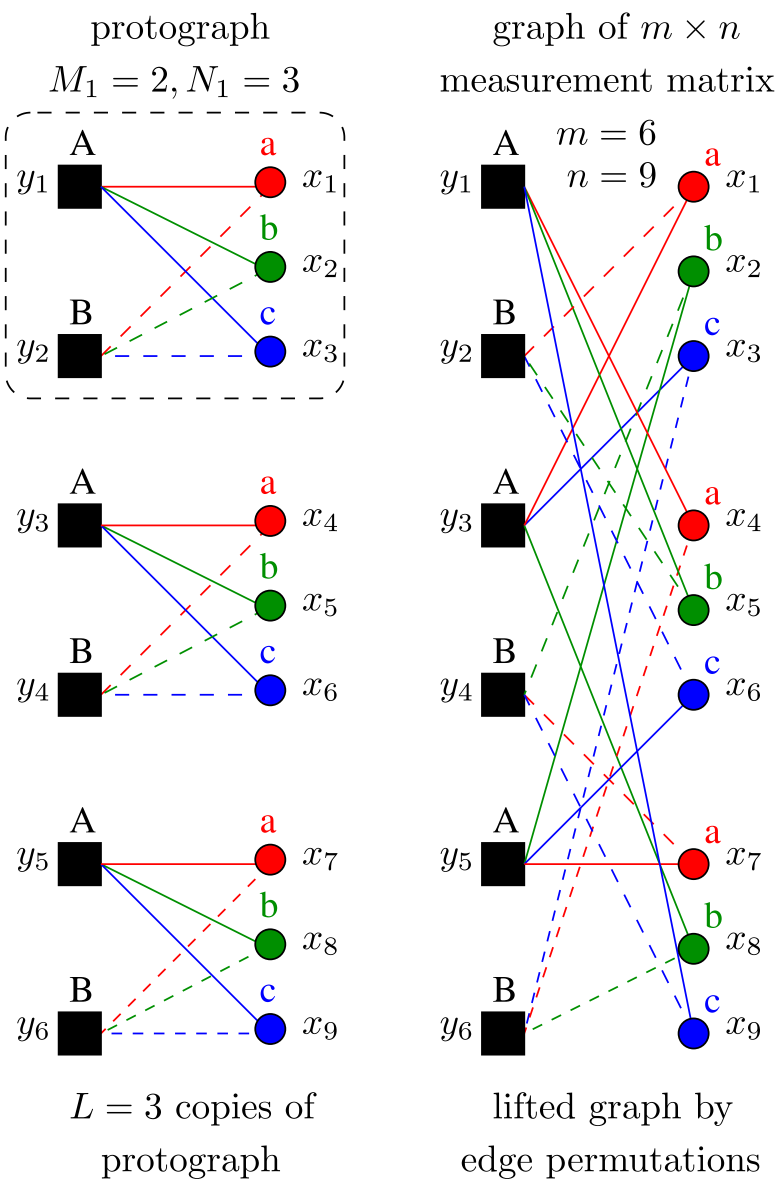
- Random Gaussian or Rademacher (± 1) matrices known to work well for compressed sensing (fulfill RIP with high probability)
- More systematic designs such as randomly sampled DFT matrices also work well
- Problem: for large signal vector dimension n, complexity of matrix-vector multiplications in BAMP as well as storage requirements impractical

  ➡ even worse for other commonly used recovery schemes (OMP or convex optimization)

- Solution proposed here: super-sparse measurement matrices lifted from a protograph

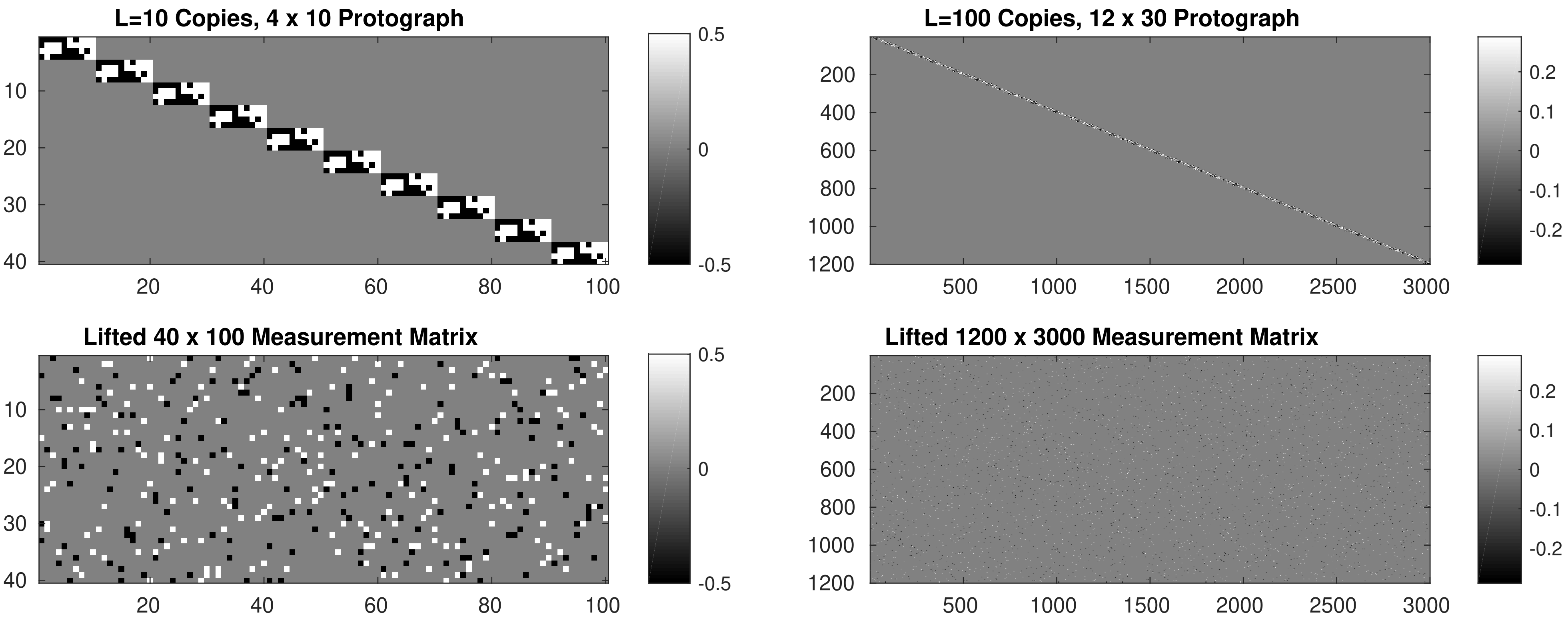
  ➡ design principle borrowed from LDPC convolutional channel codes (e.g. Mitchell *et al.* (2008))

Lifted Measurement Matrices



- start from a Rademacher protograph: random matrix of dimensions $M_1 \times N_1$, with ± 1 elements
- define lifting factor L: number of copies of the protograph
- Generate large $m \times n$ measurement matrix with $m = L \cdot M_1$ and $n = L \cdot N_1$ by random permutations of edges of the same types across the copies of the protograph
- Normalize to unit column norm.
- Note: $A_{qi}^2 = 1/m_i$ for $q \in \mathcal{M}_i$ assumed in the Z-BAMP derivation is exact with $m_i = M_1$ for all i, due to design by lifting

... Lifted Measurement Matrices: Design Examples



… Lifted Measurement Matrices: Storage

- Full matrix:
  - $m \cdot n$ memory locations: $B \cdot m \cdot n$ bits, with $B = 32$ bits for Gaussian (floating point) and $B = 1$ bit for Rademacher matrices
- Lifted matrix: lifting factor L, prototype $M_1 \times N_1$ matrix $\Rightarrow L \cdot N_1 \cdot M_1$ non-zero elements
  - Location of each non-zero element each by $\log_2(m \cdot n)$ bits plus one bit to indicate the sign:
    - ➡ total memory $L \cdot M_1 \cdot N_1 \cdot \log_2(2 \cdot m \cdot n) = \frac{1}{L} \cdot m \cdot n \cdot \log_2(2 \cdot m \cdot n)$ bits
    - ➡ saving-factor due to lifting:

$$s = \frac{B \cdot m \cdot n}{\frac{1}{L} \cdot m \cdot n \cdot \log_2(2 \cdot m \cdot n)} = \frac{L \cdot B}{\log_2(m \cdot n) + 1} = \frac{L \cdot B}{\log_2(L^2 \cdot N_1 \cdot M_1) + 1}$$

$$= \frac{L \cdot B}{2\log_2(L) + 1 + \log_2(N_1 \cdot M_1)} \approx \frac{L \cdot B}{2\log_2(L)} \approx \frac{L}{40} \cdot B \quad \text{(large } L\text{)}$$

- Numerical examples for compression rate of $m/n = M_1/N_1 = 0.4$:

  $L = 100, M_1 = 12, N_1 = 30 \rightarrow s \approx 140$ for $B = 32$ and $s \approx 4.4$ for $B = 1$

  $L = 1000, M_1 = 12, N_1 = 30 \rightarrow s \approx 1080$ for $B = 32$ and $s \approx 34$ for $B = 1$

  $L = 10^6, M_1 = 20, N_1 = 50 \rightarrow s \approx 630000$ for $B = 32$ and $s \approx 20000$ for $B = 1$

➡ Memory saving grows linearly with lifting factor L

... Lifted Measurement Matrices: Complexity

- Full matrix (Gaussian or Rademacher):
  - Within conventional BAMP, per iteration two matrix-vector multiplications, each of complexity-order $\mathcal{O}(m \cdot n)$ (simpler to implement for ± 1 matrix elements)
- Lifted matrices:
  - Same matrix-vector multiplications involve m times $|\mathcal{N}_q| = N_1$ terms in (12) and n times $|\mathcal{M}_i| = M_1$ terms in (10).
  - Complexity order $\mathcal{O}(m \cdot N_1) = \mathcal{O}(L \cdot M_1 \cdot N_1)$ and $\mathcal{O}(n \cdot M_1) = \mathcal{O}(L \cdot N_1 \cdot M_1)$, so that is twice $\mathcal{O}(L \cdot N_1 \cdot M_1) = \mathcal{O}(\frac{1}{L} \cdot m \cdot n)$

➡ Complexity saving in the order of the lifting factor L

Specialization of Z-BAMP for Application with Lifted Matrices

- Sets $\mathcal{N}_q$ of signal components connected to measurements q: $|\mathcal{N}_q| = N_1$ for all $q = 1, \dots, m$
- Sets $\mathcal{M}_i$ of measurements connected to signal components i: $|\mathcal{M}_i| = M_1$ for all $i = 1, \dots, n$
- Assume measurement noise variance σ_w^2 to be the same at all measurement nodes. Then

$$S_q^{t-1} = \sigma_w^2 + \frac{1}{M_1} \sum_{i \in \mathcal{N}_q} \sigma_i^{2(t-1)} \approx \sigma_w^2 + \frac{1}{m} \sum_{i=1}^{n} \sigma_i^{2(t-1)} = S^{t-1} \qquad \text{for all } q \tag{23}$$

with the approximation for the sum:

$$\frac{1}{M_1} \sum_{i \in \mathcal{N}_q} \sigma_i^{2(t-1)} \approx \frac{1}{M_1} \frac{N_1}{n} \sum_{i=1}^{n} \sigma_i^{2(t-1)} = \frac{N_1}{M_1 \cdot L \cdot N_1} \sum_{i=1}^{n} \sigma_i^{2(t-1)} = \frac{1}{m} \sum_{i=1}^{n} \sigma_i^{2(t-1)} \tag{24}$$

- Simplification of effective noise variance:

$$\frac{1}{c_i^{t-1}} = \frac{1}{M_1} \sum_{q \in \mathcal{M}_i} \frac{1}{S_q^{t-1}} = \frac{1}{M_1} \frac{1}{S^{t-1}} \sum_{q \in \mathcal{M}_i} 1 = \frac{1}{S^{t-1}} \;\Rightarrow\; c_i^{t-1} = c^{t-1} = \sigma_w^2 + \frac{1}{m} \sum_{i=1}^{n} \sigma_i^{2(t-1)}$$

➡ Used in Z-BAMP in (10), (12)

➡ obtain conventional BAMP scheme, applied “as is” for lifted matrices

TU WIEN

Conventional BAMP-Recovery / key equations in scalar notation:

For iterations $t = 1, 2, \ldots$:

- Compute the (one) effective noise variance c and the substitute measurements u_i for variable nodes

$$c^{t-1} = \sigma_w^2 + \frac{1}{m} \sum_{i=1}^{n} \sigma_i^{2(t-1)} \tag{25}$$

$$u_i^{t-1} = \sum_{q \in \mathcal{M}_i} A_{qi} \, z_q^{t-1} + \hat{x}_i^{t-1} \quad \text{for} \quad i = 1, 2, \ldots, n \tag{26}$$

- Compute $i = 1, 2, \ldots, n$ the signal-component estimates and the error variances

$$\hat{x}_i^t = F_i\big(u_i^{t-1}, c_i^{t-1}\big) \qquad \sigma_i^{2(t)} = c_i^{t-1} \, F_i'\big(u_i^{t-1}, c_i^{t-1}\big) \tag{27}$$

- Compute for z_q for the measurement nodes $q = 1, \ldots, m$

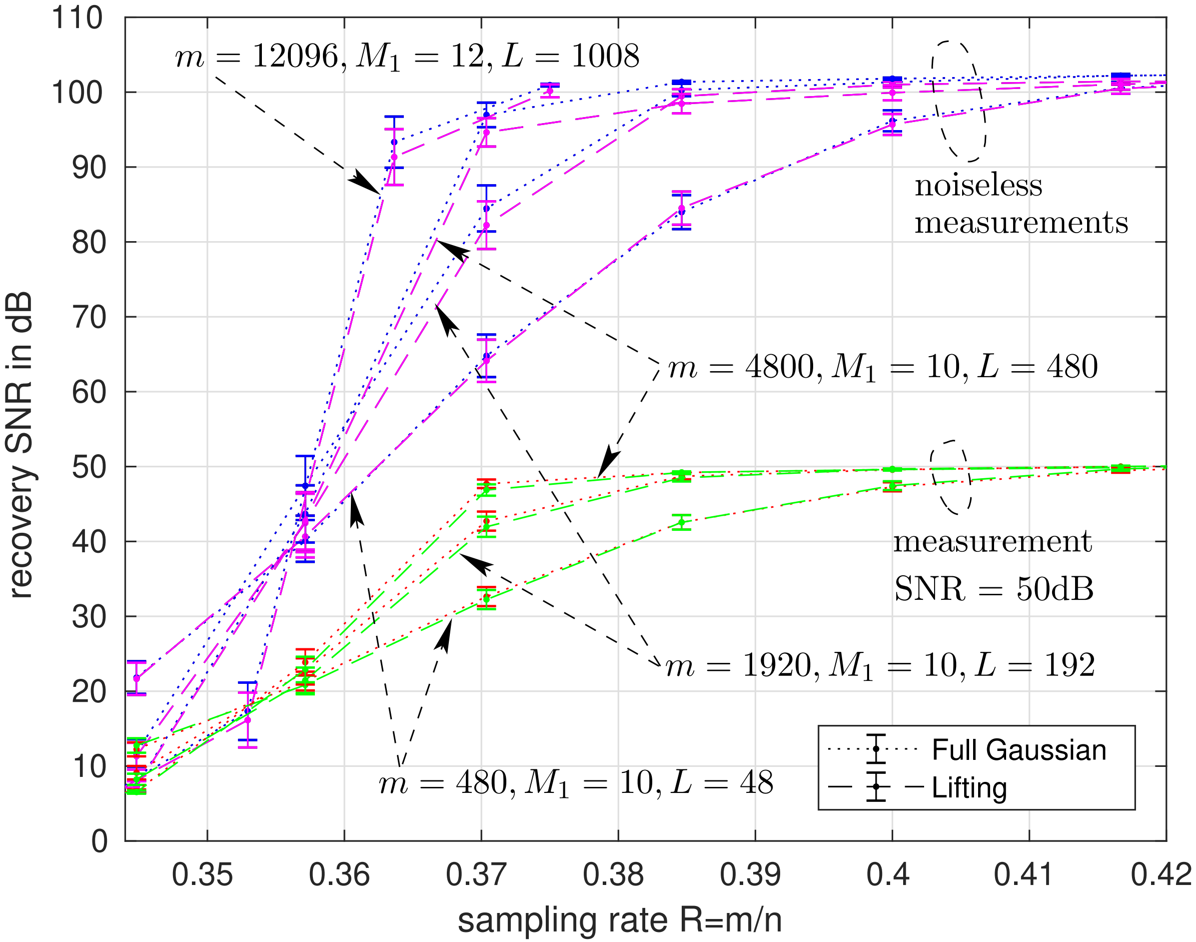
$$z_q^t = y_q - \sum_{i \in \mathcal{N}_q} A_{qi} \, \hat{x}_i^t + z_q^{t-1} \frac{1}{m} \sum_{i=1}^{n} \sigma_i^{2(t)} \tag{28}$$

- Stop, if $\sum_{i=1}^n \big(\hat{x}_i^t - \hat{x}_i^{t-1}\big)^2 < \varepsilon^2 \sum_{i=1}^n \big(\hat{x}_i^t\big)^2$; otherwise, continue with (25) for $t := t + 1$.

Simulation Results

- I.i.d. Bernoulli-Gaussian source, $\gamma_i = 0.8$ (zero-probability) and $\sigma^2_{X_i} = 1$ for all $i = 1, 2, .., n$.
- Sampling rates around $\frac{m}{n} = 0.4$ required for successful recovery
- Protograph with $M_1 = 10$; pick $N_1 \in \{24, ..., 28, 29\}$: leads to sampling rates in the range $0.34 \ldots 0.42$
- pick lifting factors $L \in \{48, 192, 480\}$
  - ➡ Measurement matrices with $m = L \cdot M_1 \in \{480, 1920, 4800\}$ rows and $n \in \{1152 \ldots 1392, 4608 \ldots 5568, 11520 \ldots 13920\}$ columns (signal components)
- $M_1 = 12$ with $N_1 \in \{29, 30, \ldots 35\}$ for lifting factor $L = 1008$
  - ➡ Measurement matrix with $m = 12096$ rows and $n \in \{29232 \ldots 35280\}$ columns
- Compare recovery SNR $= 20 \log_{10} \|\mathbf{x}\|_2 / \|\hat{\mathbf{x}} - \mathbf{x}\|_2$ obtained with full Gaussian random matrices with lifted matrices of the same sizes.
- Compare noiseless measurements and noisy ones with average measurement SNR of 50dB
- BAMP stopping criterion: $\varepsilon = 10^{-5}$ ➡ remaining error converts to SNR/dB $= 20 \log_{10}(10^5) = 100$ dB (saturation of simulations in the noiseless case)

... Simulation Results



- almost no performance difference between lifted and Gaussian matrices of same sizes
- larger lifting factor L: better recovery performance, as block size grows (steeper transition)
- Gaussian *and* lifted matrices: $\approx$ 100 BAMP iterations to converge in the transition region; decrease to fewer than 60 iterations for rates above 0.4

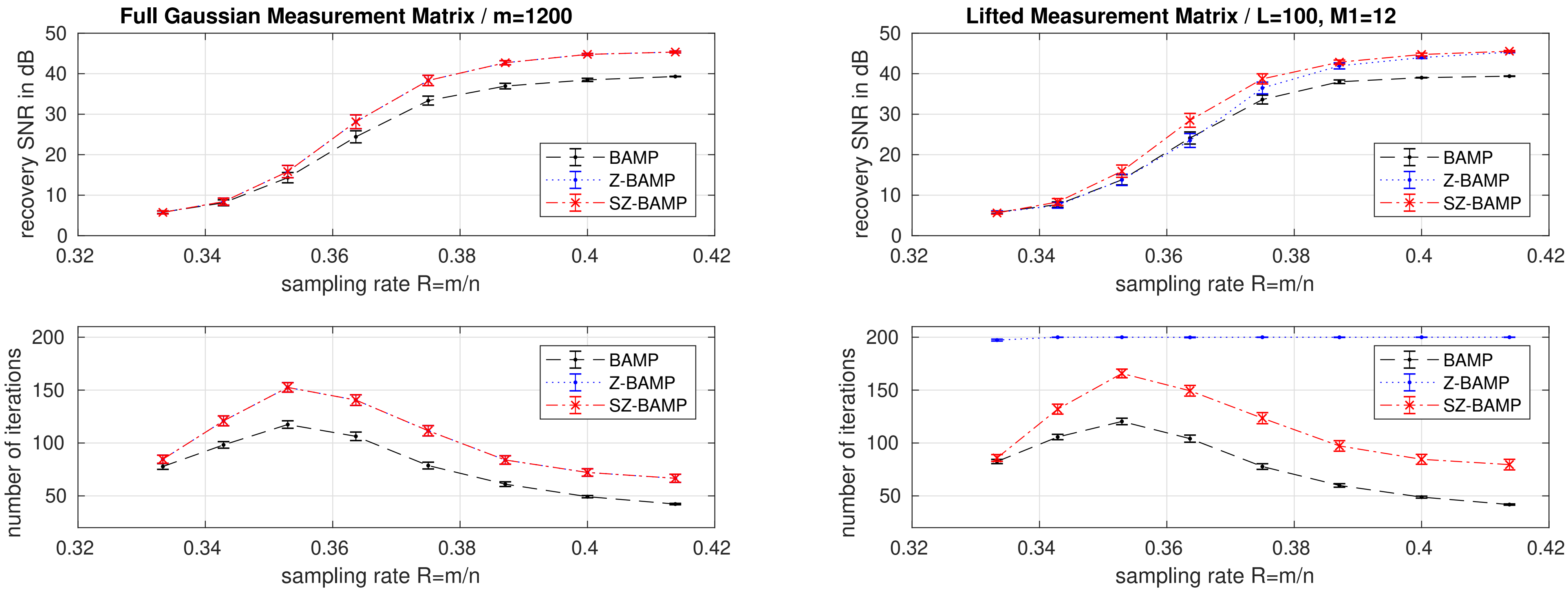
➡ Lifted matrices allow for much larger signal dimension n

➡ in Matlab, full Gaussian matrices require 3 GB storage for $m = 12096$ at sampling rate of 0.4

... Simulation Results

- I.i.d. Bernoulli-Gaussian source, $\gamma_i = 0.8$ (zero-probability) and $\sigma^2_{X_i} = 1$ for all $i = 1, 2, .., n$.
- Sampling rates around $\frac{m}{n} = 0.4$ required for successful recovery
- Protograph with $M_1 = 12$; lifting factor $L = 100$
- **Noise no longer the same for all measurements**: assumption that randomly chosen 50% of the measurements have noise standard deviation 10 times larger than the other 50%
- Average measurement SNR set to 40dB; BAMP uses this *average* knowledge
- Stopping criterion: $\varepsilon = 10^{-3}$ ➡ adapted to the average measurement noise variance that corresponds to 40dB – stopping threshold set to $10 \log_{10} \left((10^3)^2\right) = 60\text{dB}$
- Z-BAMP scheme uses knowledge about quality of *each* of the measurements
- SZ-BAMP: modified Z-BAMP scheme that uses the same sum-approximation as above to get more stable computation of the sum-terms while still exploiting different noise variances
- Compare recovery SNR $= 20 \log_{10} \|\mathbf{x}\|_2 / \|\hat{\mathbf{x}} - \mathbf{x}\|_2$ obtained with full Gaussian random matrices and with lifted matrices of the same sizes.

... Simulation Results



- Z-BAMP and SZ-BAMP: performance gain vs BAMP at larger rates
- Z-BAMP: oscillations around an accurate recovery result when lifted matrices are used: too few terms in the sums over the estimation error variances; stopping criterion doesn't detect convergence
- Modification by SZ-BAMP leads to slightly *better* performance and fewer iterations; still some more iterations than with normal BAMP – but also better performance
- No performance difference between full Gaussian and lifted matrices

Conclusions

- Sparse Lifted Measurement Matrices can replace full Gaussian or Rademacher Matrices with large savings in complexity and storage and without any performance loss
- If the measurement noise variances are different, Z-BAMP and SZ-BAMP schemes can beneficially exploit this for any type of measurement matrix
- Number of required iterations can be kept under control by modifications in the computation of sums over estimation error variances: no performance loss

References

Bayati, M., & Montanari, A. 2011. The Dynamics of Message Passing on Dense Graphs, with Applications to Compressed Sensing. *IEEE Transactions on Information Theory*, **57**(2), 764–785.

Birgmeier, S., & Goertz, N. 2018 (Sept.). Optimizing Approximate Message Passing for Variable Measurement Noise. *Pages 489–493 of: Proceedings 26th European Signal Processing Conference (EUSIPCO)*.

Donoho, D.L., Maleki, A., & Montanari, A. 2009. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, **106**(45), 18914–18919.

Donoho, D.L., Maleki, A., & Montanari, A. 2010a. Message passing algorithms for compressed sensing: I. motivation and construction. *In: Proceedings IEEE Information Theory Workshop (ITW)*.

Donoho, D.L., Maleki, A., & Montanari, A. 2010b. Message passing algorithms for compressed sensing: II. analysis and validation. *In: Proceedings IEEE Information Theory Workshop (ITW)*.

Donoho, D.L., Maleki, A., & Montanari, A. 2011. How to Design Message Passing Algorithms for Compressed Sensing. *unpublished draft, accessed 30 Dec. 2013*, Feb.

Eldar, Y.C., & Kutyniok, G. 2012. *Compressed sensing: theory and applications*. Cambridge University Press.

Goertz, N., & Birgmeier, S. 2019 (July). Sparse Compressed-Sensing Measurement Matrices designed for Signal Recovery with Bayesian Approximate Message Passing. *In: submitted to IEEE International Symposium on Information Theory*.

Mitchell, D. G. M., Pusane, A. E., Goertz, N., & Costello, D. J. 2008 (Sept.). Free Distance Bounds for Protograph-Based Regular LDPC Convolutional Codes. *Pages 408–413 of: Proceedings 5th International Symposium on Turbo Codes and Related Topics*.

Montanari, A. 2011. Graphical Models Concepts in Compressed Sensing. *Arxiv*.