Joint Estimation and Decoding for Sparse Channels via Relaxed Belief Propagation

Phil Schniter

THE OHIO STATE UNIVERSITY

(With support from the National Science Foundation via grant CCF-1018368)

Asilomar, Nov. 2010

Problem Statement:

Consider communicating over a channel that is

- Rayleigh *block-fading* with block size N,
- *frequency-selective* with delay spread L (where $L < N$),
- *sparse* with S non-zero taps (where $0 < S < L$),

where *both the channel coefficients and support are unknown* to the receiver.

Important questions:

1. What is the capacity of this channel?
2. How can we build a *practical* system that operates near this capacity?

The Capacity of our Sparse Channel:

For the unknown N-block-fading, L-length, S-sparse channel described earlier, Kannu/Schniter [1] established that

1. In the high-SNR regime, the ergodic capacity obeys

$$C_{\text{sparse}}(\text{SNR}) = \tfrac{N-S}{N}\log(\text{SNR}) + \mathcal{O}(1).$$

2. To achieve the prelog factor $R_{\text{sparse}} = \frac{N-S}{N}$, it suffices to use
   - pilot-aided OFDM (with N subcarriers, of which S are pilots)
   - with (necessarily) *joint* channel estimation and data decoding.

Key points:

- The effect of *unknown channel support* manifests in the $\mathcal{O}(1)$ offset term, not the prelog factor $\frac{N-S}{N}$.
- While [1] uses constructive proofs, the scheme proposed there is impractical.

[1] A. P. Kannu and P. Schniter, "On communication over unknown sparse frequency-selective block-fading channels," arXiv 1006.1548, June 2010.

The Conventional Approach — Compressed Channel Sensing (CCS):

- Motivated by recent advances in the field of "compressed sensing."
- Consider OFDM with P pilot subcarriers, giving $\boxed{\boldsymbol{y}_\mathsf{p} = \boldsymbol{F}_\mathsf{p}\boldsymbol{h} + \boldsymbol{v}_\mathsf{p},}$ where

$$\begin{aligned} \boldsymbol{y}_\mathsf{p} \in \mathbb{C}^P &: \text{ observations on pilot subcarriers} \\ \boldsymbol{F}_\mathsf{p} \in \mathbb{C}^{P\times L} &: \text{ a submatrix of the DFT matrix} \\ \boldsymbol{h} \in \mathbb{C}^L &: \text{ channel impulse response } (S\text{-sparse}) \\ \boldsymbol{v}_\mathsf{p} \in \mathbb{C}^P &: \text{ AWGN with variance } \sigma_v^2. \end{aligned}$$

- CCS-based noncoherent decoding takes the following *decoupled* approach [2]:
  1. Use sparse reconstruction to generate a pilot-aided estimate $\hat{\boldsymbol{h}}$,
  2. Coherently decode assuming $\boldsymbol{h} = \hat{\boldsymbol{h}}$.
- Modern CCS typically employs the LASSO (also known as BPDN):

$$\hat{\boldsymbol{h}}_{\mathsf{LASSO},\lambda} \triangleq \arg\min_{\hat{\boldsymbol{h}}} \|\hat{\boldsymbol{h}}\|_1 \text{ such that } \|\boldsymbol{y}_\mathsf{p} - \boldsymbol{F}_\mathsf{p}\hat{\boldsymbol{h}}\|_2^2 \leq \lambda$$

  which guarantees σ_v^2-proportional ℓ_2-error with $\boxed{P = \mathcal{O}\big(S(\log L)^5\big)}$ pilots.

[2] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, June 2010.

The Main Contribution of this Work:

A new approach to communicating over sparse channels that. . .

- (empirically) achieves the optimal prelog factor $R_{\text{sparse}} = \frac{N-S}{N}$,
- is practical: complexity $\mathcal{O}(NL)$, which supports $L \gtrsim 100$,
- significantly outperforms CCS-based decoding at both low and high SNR.

Our scheme uses. . .

- a conventional transmitter: pilot-aided BICM OFDM,
- a novel receiver: based on loopy *belief propagation* (BP)
  - key enabler: "*relaxed BP*" of Guo/Wang [3] and Rangan [4]

[3] D. Guo and C.-C. Wang, "Random sparse linear systems observed via arbitrary channels: A decoupling principle, in *Proc. ISIT*, June 2007.

[4] S. Rangan, "Estimation with random linear mixing, belief propagation and compressed sensing," arXiv:1001.2228v2, May 2010.

How would we do optimal decoding?:

- To minimize BER, we need to compute the posterior pmfs $\{p(b_q \,|\, \boldsymbol{y}, \boldsymbol{c}_{\text{pt}})\}_{q=1}^{Q}$ where $\boldsymbol{c}_{\text{pt}}$ denotes known pilot/training bits.

- Assuming

  1. bit-interleaved coded modulation (BICM) with OFDM:

$$\begin{array}{c}\text{info}\\ \text{bits } \boldsymbol{b}\end{array} \rightarrow \begin{array}{c}\text{coded/interleaved}\\ \text{bits } \boldsymbol{c}\end{array} \rightarrow \begin{array}{c}M\text{-QAM}\\ \text{symbols } \boldsymbol{s}\end{array} \rightarrow \begin{array}{c}N \text{ OFDM}\\ \text{subcarriers}\end{array},$$

  2. sparse WSSUS Rayleigh-fading channel $\boldsymbol{h}$ and AWGN $\boldsymbol{v}$:

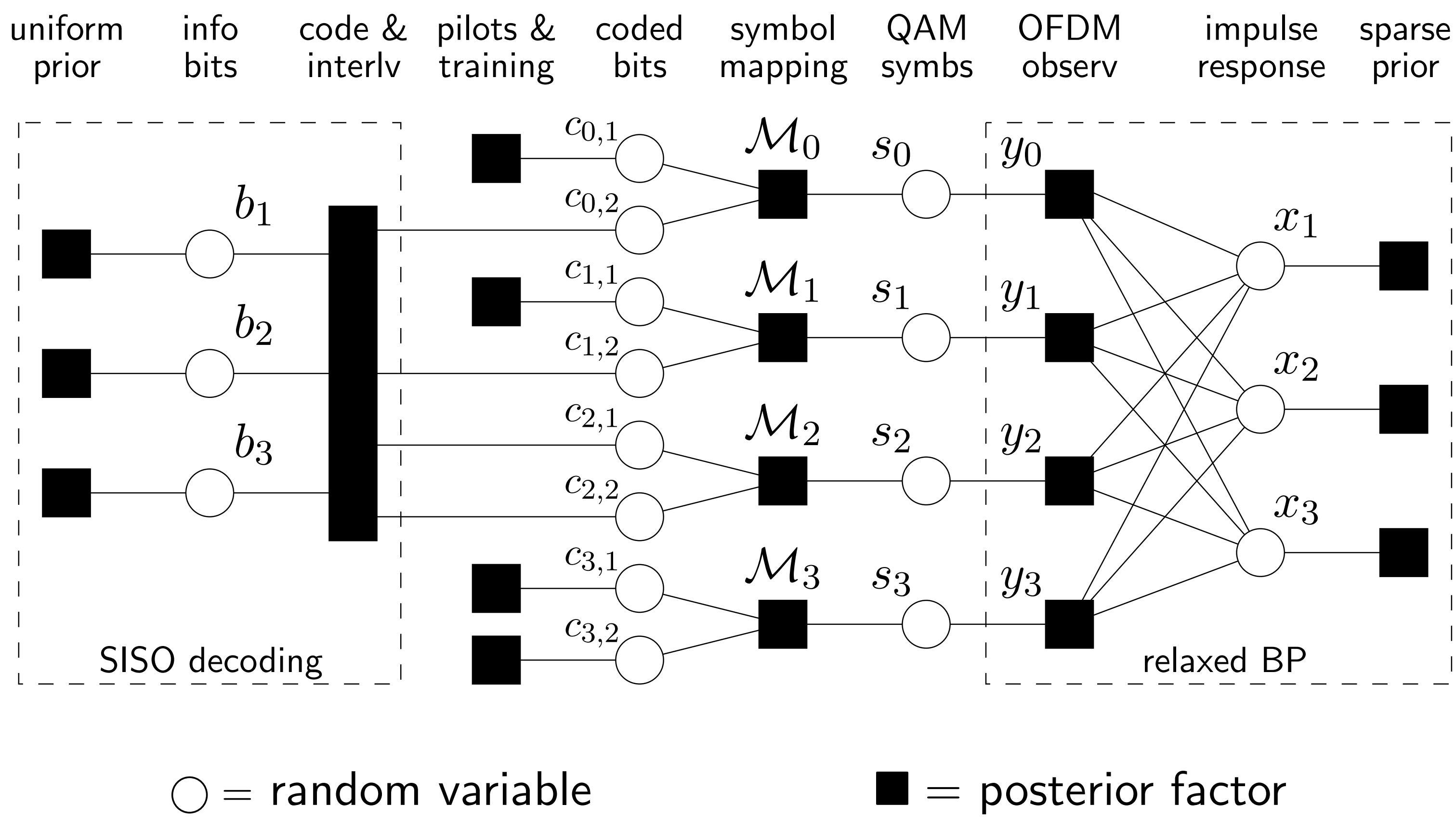
$$\boldsymbol{y} = \mathcal{D}(\boldsymbol{s})\boldsymbol{F}\boldsymbol{h} + \boldsymbol{v},$$

  we can factor the posterior as follows:

$$p(b_q|\boldsymbol{y}, \boldsymbol{c}_{\text{pt}}) \propto \int_{\boldsymbol{x}} \prod_{j=1}^{L} p(x_j) \sum_{\boldsymbol{s}} \prod_{i=1}^{N} p(y_i|s_i, \boldsymbol{x}) \sum_{\boldsymbol{c}} p(s_i|\boldsymbol{c}_i) \sum_{\boldsymbol{b}_{-q}} p(\boldsymbol{c}|\boldsymbol{b}, \boldsymbol{c}_{\text{pt}}) \prod_{q=1}^{Q} p(b_q).$$

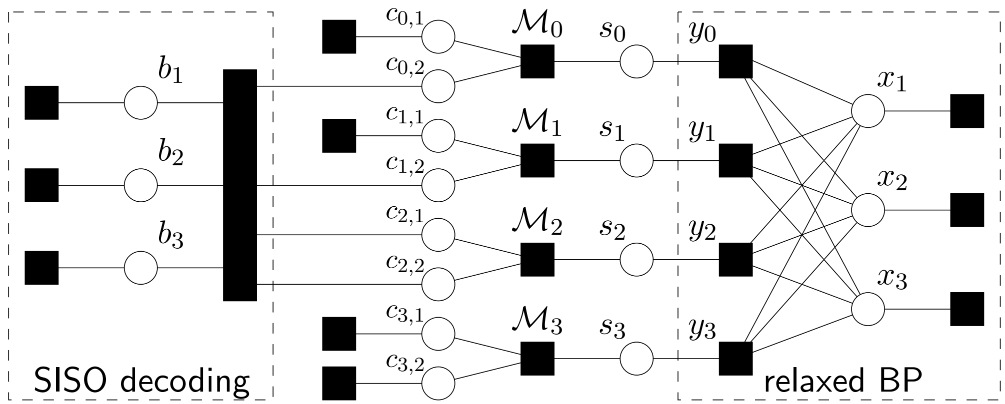
  which can be visualized using a *factor graph*...

The Factor Graph for Noncoherent BICM-OFDM:



To jointly infer all random variables, we perform *belief propagation* (BP) on the factor graph, passing (parameters of) pdfs from node to node.

Approximate BP for Noncoherent BICM-OFDM:



- Since our factor graph (FG) has loops, BP convergence is not guaranteed. However, our simulations suggest that this is not a problem.
- Approximate BP on the left portion of the FG can be efficiently implemented using an off-the-shelf *soft-input/soft-output (SISO) decoder*.
- Approximate BP on the right portion of the FG can be efficiently implemented using a modification of the "*relaxed BP*" algorithm (suitable for $N \gtrsim 100$).
- Repeated forward-backward BP iterations are reminiscent of "noncoherent turbo equalization."

Numerical Results:

Transmitter:

- LDPC codewords with length ≈ 10000 bits.
- M-QAM with $M \in \{4, 16, 64, 256\}$ and multi-level Gray mapping.
- OFDM with $N = 1021$ subcarriers.
- Various choices of P "*pilot subcarriers*" and TM interspersed "*training bits*."

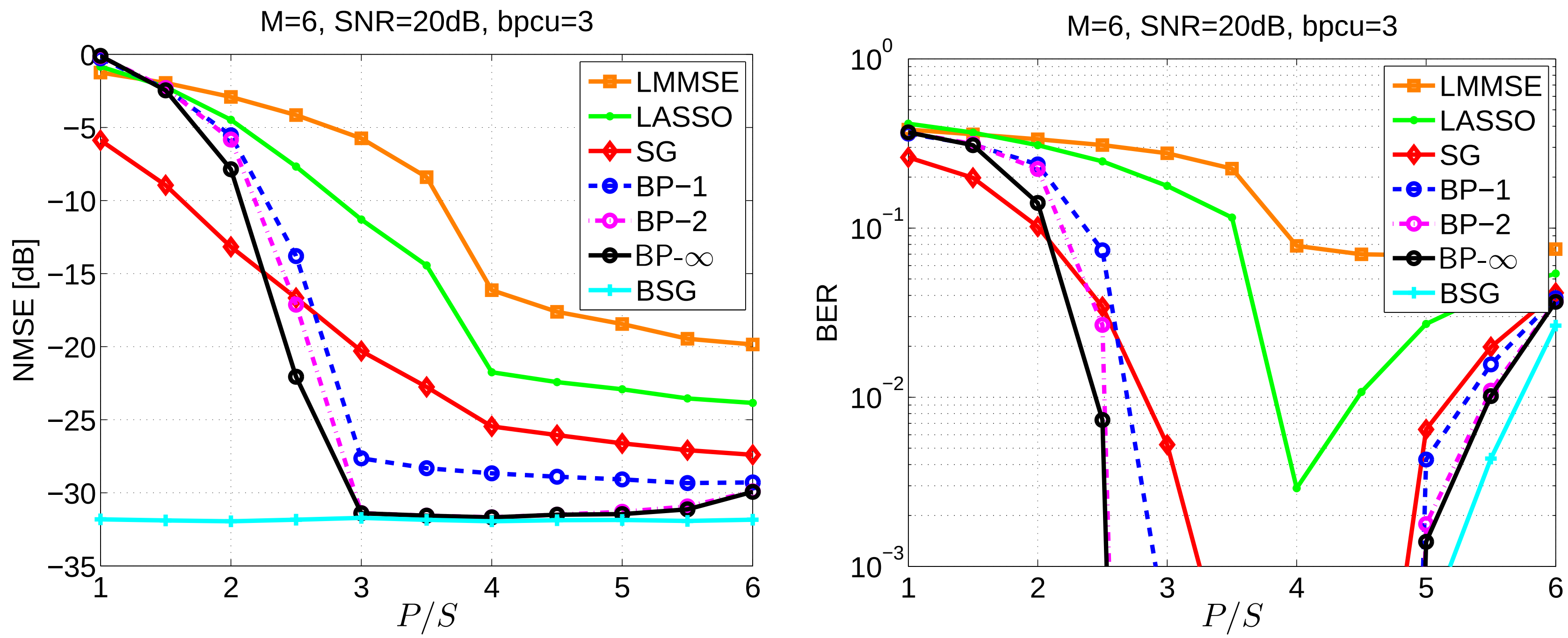
Channel:

- Delay spread $L = 256 \approx N/4$.
- Sparsity $S = 64 = L/4$.

Compressed channel sensing:

- LASSO was implemented using SPGL1 with genie-aided tuning.
- For comparison, we also performed CCS using several reference estimators: LMMSE, support-aware MMSE, and bit+support-aware MMSE.

NMSE & BER versus pilot ratio P/S (at SNR=20dB, $T=0$):

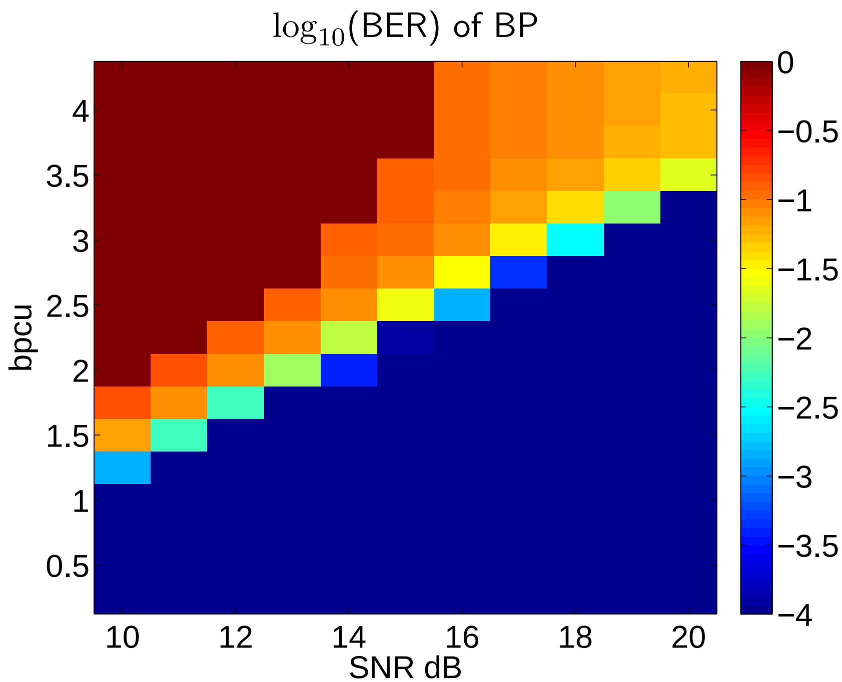


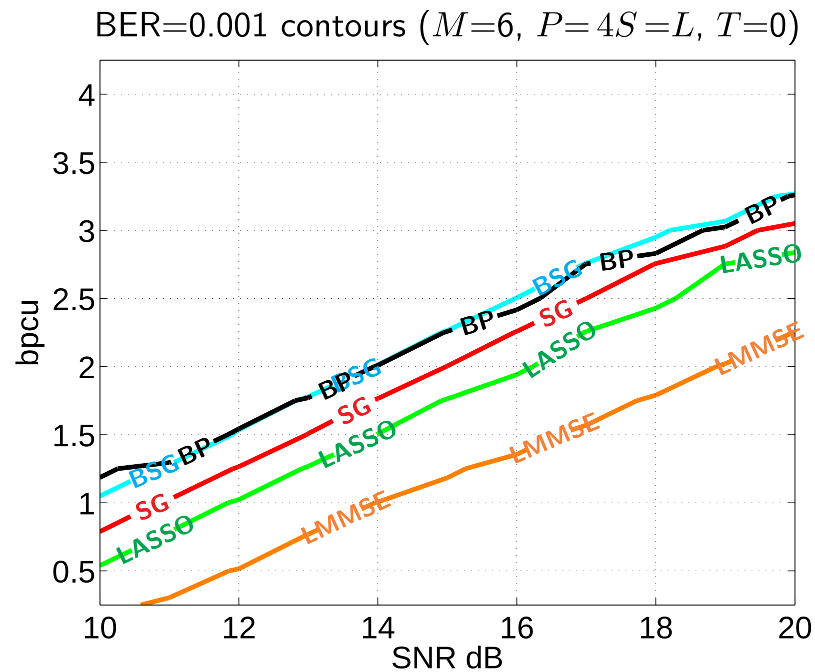
implementable schemes	reference schemes
LMMSE = LMMSE-based CCS LASSO = LASSO-based CCS BP-n = BP after n turbo iterations	SG = support-aware genie BSG = bit- and support-aware genie

Observations:

- For CCS, channel estimation MSE improves monotonically with P.
- As P grows too large, BER suffers due to necessary decrease in LDPC code-rate.
- For CCS, $P = 4S = L$ gives best tradeoff. (No longer "compressed" channel sensing!)

Bit-Rate versus SNR (with $P=4S=L$ pilots and $T=0$ training):

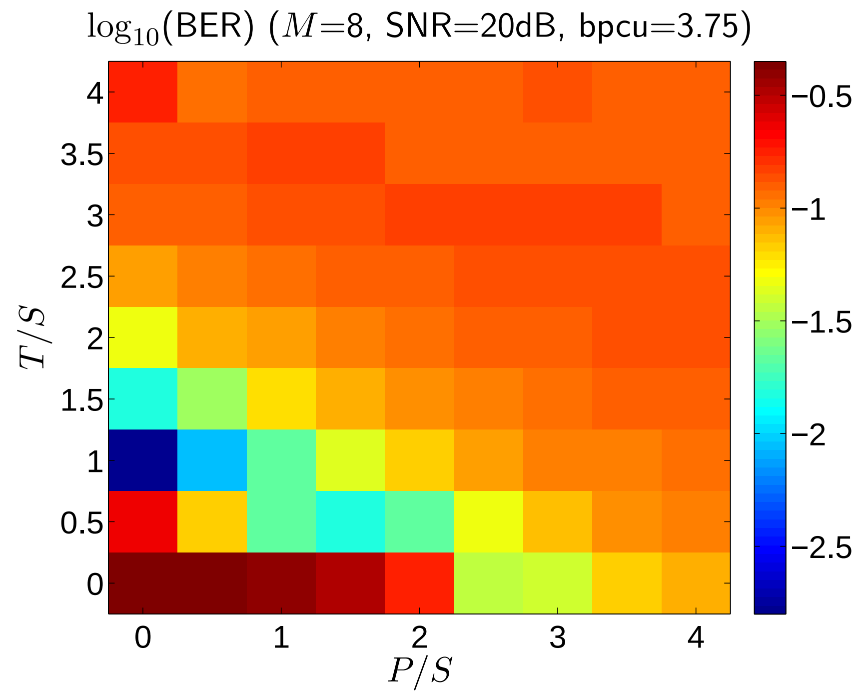


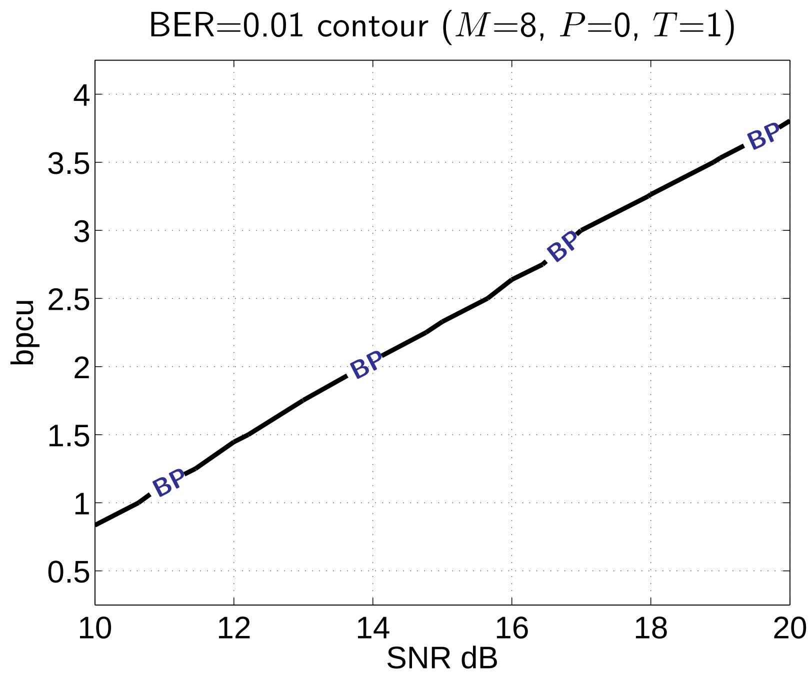


Key points:

- Turbo-BP outperforms not only LASSO, but even the support genie (SG)!
- Turbo-BP performs nearly as well as the bit+support-aware genie (BSG)!
- With $P = L$, all approaches achieve the prelog factor $R \approx \frac{N-L}{N} = \frac{3}{4}$, which falls short of the optimal $R_{\text{sparse}} = \frac{N-S}{N} = \frac{15}{16}$.

Bit-Rate versus SNR (with $P=0$ pilots and $T=S$ training):

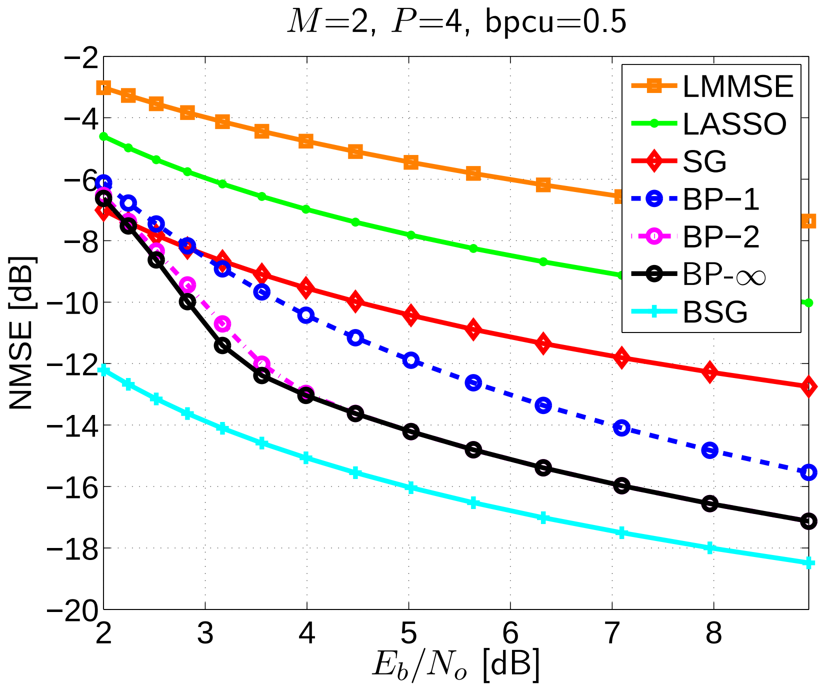


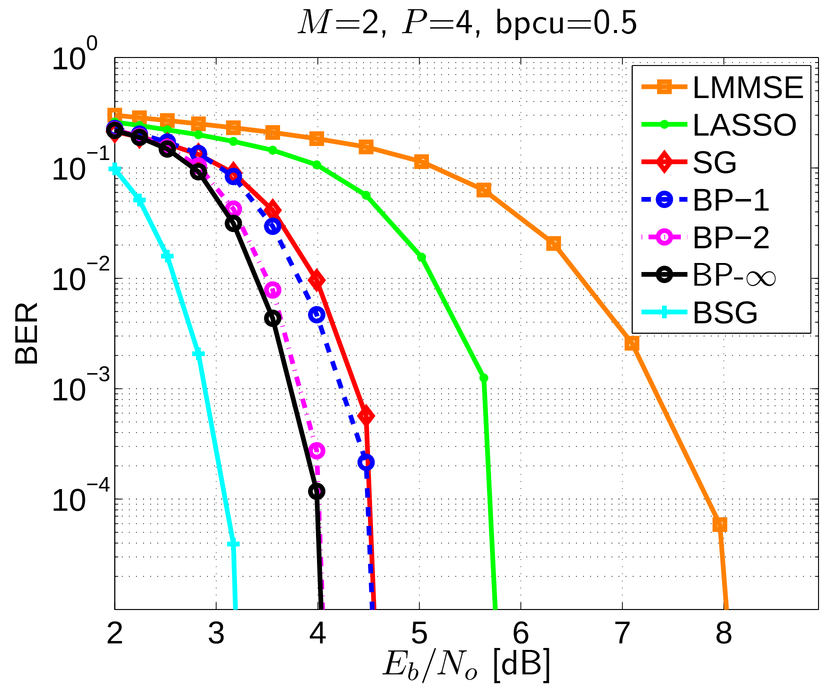


Key points:

- At high-SNR, BP favors the use of $P=0$ pilots and $TM=SM$ training bits.
- With this pilot/training arrangement, BP achieves the channel's capacity prelog factor $R_{\text{sparse}} = \frac{N-S}{N}$.

BER versus SNR (with $P = 4S = L$ pilots and $T = 0$ training):





implementable schemes	reference schemes
LMMSE = LMMSE-based CCS LASSO = LASSO-based CCS BP-n = BP after n turbo iterations	SG = support-aware genie BSG = bit- and support-aware genie

Key points:

- Sparsity can be exploited even at very low SNR. (SNR$=\frac{1}{\text{bpcu}}\frac{E_b}{N_o}$ range is [0,6.5] dB.)
- BP has a 1.8dB advantage over LASSO, which has a 2.2dB advantage over LMMSE.

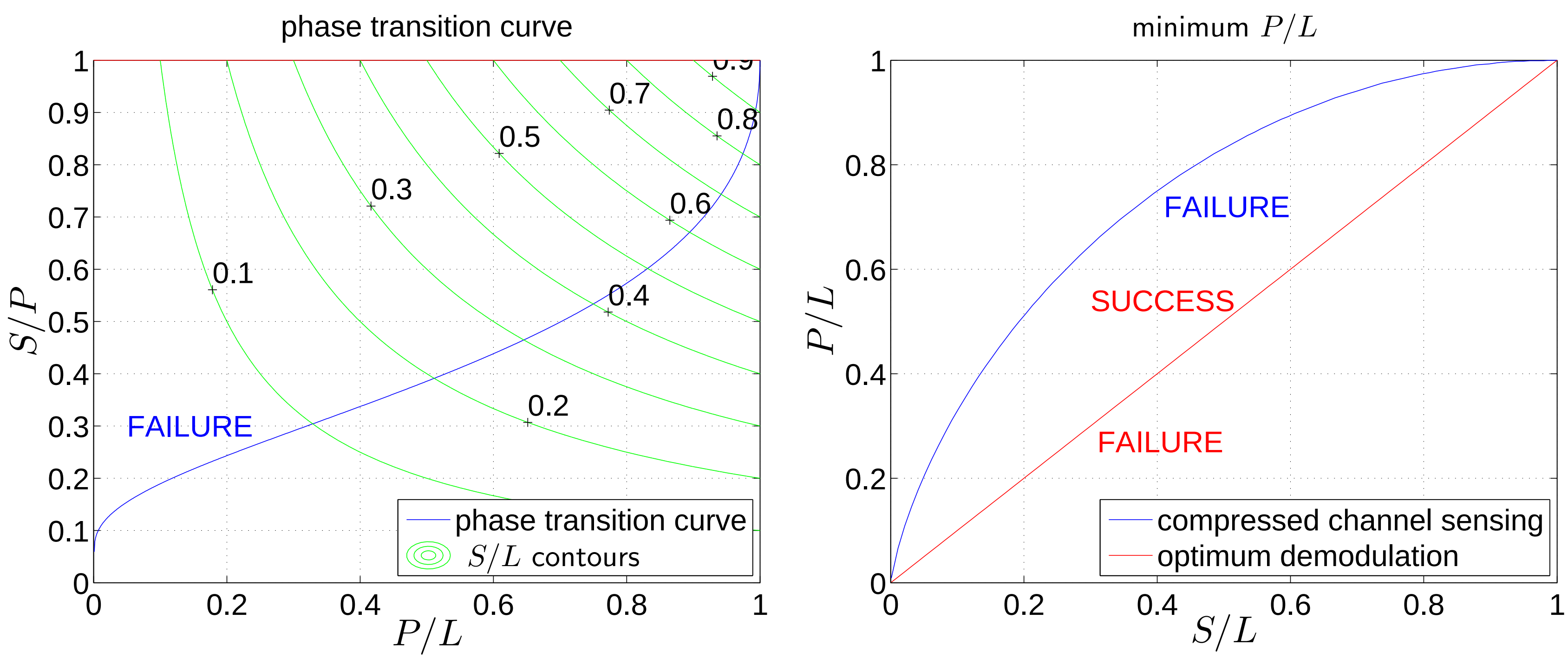
Conclusions:

- We proposed a new noncoherent decoding scheme for N-subcarrier BICM-OFDM transmitted over an S-sparse L-length channel that...
  - is based on approximate belief propagation,
  - is computationally efficient: complexity $\mathcal{O}(NL)$.
- Simulations suggest that our scheme...
  - at high SNR, achieves the channel's capacity prelog factor $R_{\text{sparse}} = \frac{N-S}{N}$,
  - at low SNR, is only 0.8dB worse than bit+support-aware genie,
  - significantly outperforms LASSO-based compressed channel sensing.
- Future work:
  - automatically learn the channel statistics (e.g., SNR, sparsity S),
  - further reduce complexity to $\mathcal{O}(N \log N)$,
  - exploit channel tracking across OFDM symbols,
  - handle channel variation within each OFDM symbol.

Thanks!

Performance Limits of CCS:

In the large system limit (i.e., $L, S, P \to \infty$) with i.i.d $\boldsymbol{F}_{\mathrm{p}}$, the Donoho/Tanner *phase transition curve* (PTC) predicts exactly where noiseless LASSO will fail:



The PTC translates directly to a minimum required P/L for CCS (as SNR$\to\infty$).

[5] D. L. Donoho and J. Tanner, "Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing," *Phil. Trans. Royal Soc. A*, 2009.