Sparse Regression Codes for MIMO Detection

Haiwen Cao and Pascal O. Vontobel

Department of Information Engineering
The Chinese University of Hong Kong

haiwen.cao@ieee.org, pascal.vontobel@ieee.org

ITW 2022, Mumbai, India

November 1–2, 2022

Overview

Overview

System Model

- We consider a MIMO system with n_0 antennas at the sender side and m_0 antennas at the receiver side.
- We will consider not only independent MIMO channels, but also correlated MIMO channels.
- Messages are encoded into codewords $\boldsymbol{x}$ of suitable length n. For simplicity, we assume $n = d \cdot n_0$, where d is some positive integer, and then let $m \triangleq d \cdot m_0$.
- In order to recover the transmitted messages at the receiver side, we need to receive d consecutive words, denoted by $\boldsymbol{y}$.

System Model

Definition (MIMO detection)

Mathematically, the problem can be formulated in the following way:

$$\boldsymbol{y} = \widetilde{\boldsymbol{H}}\boldsymbol{x} + \boldsymbol{w}, \tag{1}$$

where

$$\widetilde{\boldsymbol{H}} \triangleq \begin{pmatrix} \boldsymbol{H}_1 & & \\ & \ddots & \\ & & \boldsymbol{H}_d \end{pmatrix}, \quad \boldsymbol{H}_i \in \mathbb{C}^{m_0 \times n_0} \text{ for all } i \in [d],$$

and where the entries of $\boldsymbol{w} = (w_i)_{i \in [m]}$ are i.i.d. according to $\mathcal{CN}(0, \sigma^2)$. We consider quasi-static MIMO channels here, that is, $\boldsymbol{H}_1 = \cdots = \boldsymbol{H}_d = \boldsymbol{H}_0$, where $\boldsymbol{H}_0$ is sampled from some known distribution.

We will consider MIMO matrices with i.i.d. entries and MIMO matrices with correlated entries in this talk.

System Model

Definition (MIMO channel models)

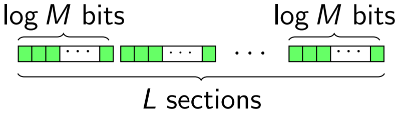
- **IID matrices:** we consider MIMO matrices $\boldsymbol{H}_0$ with i.i.d. entries according to $\mathcal{CN}(0, \frac{1}{m_0})$.

- **Jakes model:** we consider MIMO matrices $\boldsymbol{H}_0$ constructed according to the Jakes model, that is, $\boldsymbol{H}_0 \triangleq \boldsymbol{L}^{1/2} \boldsymbol{H}_{\mathrm{iid}} \boldsymbol{R}^{1/2}$,
  - where $\boldsymbol{H}_{\mathrm{iid}}$ is with entries being i.i.d. according to $\mathcal{CN}(0, \frac{1}{m_0})$,
  - where $\boldsymbol{L}$ with size $m_0 \times m_0$ is given by $L_{i,j} \triangleq J_0(|i-j| \cdot \pi)$,
  - where $\boldsymbol{R}$ with size $n_0 \times n_0$ is given by $R_{i,j} \triangleq J_0(|i-j| \cdot \pi)$,
  - where $J_0(\cdot)$ is the zero-order first-kind Bessel function.

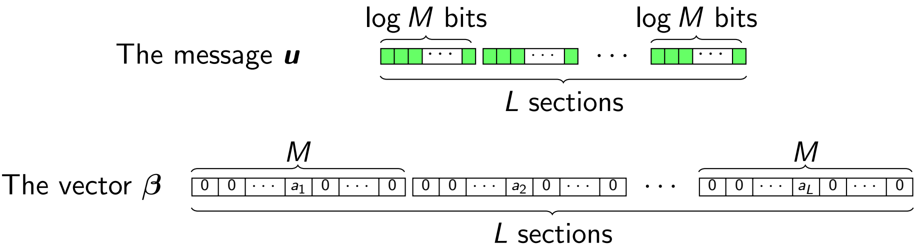
Overview

Sparse Regression Codes

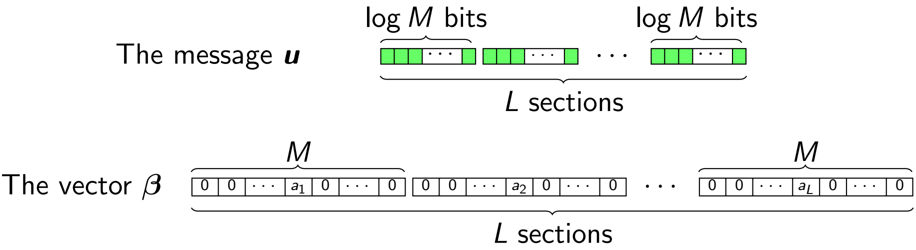
The message $\boldsymbol{u}$



Sparse Regression Codes



Sparse Regression Codes



- The codeword $\boldsymbol{x}$ of length n is of the form $\boldsymbol{A}\beta$, i.e., $\boldsymbol{x} = \boldsymbol{A}\beta$.
- The matrix $\boldsymbol{A}$ of size $n \times ML$ is the so-called design matrix and its entries are i.i.d. Gaussian $\sim \mathcal{N}(0, 1/n)$.
- In an actual implementation, the Gaussian design matrix is usually replaced by a (row-permuated) Hadamard matrix.

NUV Priors

There are a few previous results on normals with unknown variances (NUVs) priors listed chronologically.

- The idea behind NUV priors was commonly used in sparse Bayesian learning (2001).
- Keusch *et al*. (2021) proposed a new method for estimating binary input signals using (binary-enforcing) NUV priors.
- Marti *et al*. (2021) applied Keusch's proposed method to the MIMO detection problem, which only considers uncoded modulation scheme.

Previous works only consider NUV priors for scalar constraint sets. In our work, we generalize it to NUV priors for structured and sparse vector constraint sets and apply our proposed method to the MIMO detection problem with coded data transmission.

NUV Priors

Definition (NUV priors for a single vector)

For simplicity, a binary case, i.e., a constraint set with two candidates $\{\boldsymbol{a}, \boldsymbol{b}\}$ is illustrated as follows. The (improper) NUV prior is constructed as

$$\rho(\beta_0, \boldsymbol{\theta}) \triangleq \mathcal{N}(\beta_0|\boldsymbol{a}, \sigma_{\boldsymbol{a}}^2 \cdot \mathrm{I}) \cdot \mathcal{N}(\beta_0|\boldsymbol{b}, \sigma_{\boldsymbol{b}}^2 \cdot \mathrm{I}),$$

where $\boldsymbol{\theta} \triangleq (\sigma_{\boldsymbol{a}}^2, \sigma_{\boldsymbol{b}}^2)^\mathsf{T}$. Here, the variance vector $\boldsymbol{\theta}$ is unknown.

For an estimation problem, e.g., the estimation of $\beta_0 \in \{\boldsymbol{a}, \boldsymbol{b}\}$ based on the channel observation $\boldsymbol{y} = \beta_0 + \boldsymbol{w}$, the unknown variance vector $\boldsymbol{\theta}$ is obtained via a maximum-a-posterior (MAP) estimator, i.e.,

$$\hat{\boldsymbol{\theta}} \triangleq \operatorname*{argmax}_{\boldsymbol{\theta}} \rho(\boldsymbol{\theta}|\boldsymbol{y}) = \operatorname*{argmax}_{\boldsymbol{\theta}} \int_{-\infty}^{\infty} p(\boldsymbol{y}|\beta_0)\rho(\beta_0, \boldsymbol{\theta}) d\beta_0.$$

Once $\hat{\boldsymbol{\theta}}$ is given, we estimate $\hat{\beta}_0$ via a MAP estimator, i.e.,

$$\hat{\beta}_0(\hat{\boldsymbol{\theta}}) \triangleq \operatorname*{argmax}_{\beta_0} \; p(\boldsymbol{y}|\beta_0)\rho(\beta_0, \hat{\boldsymbol{\theta}}).$$

NUV Priors

Definition (NUV priors for a single vector)

For simplicity, a binary case, i.e., a constraint set with two candidates $\{\boldsymbol{a}, \boldsymbol{b}\}$ is illustrated as follows. The (improper) NUV prior is constructed as

$$\rho\left(\beta_0, \boldsymbol{\theta}\right) \triangleq \mathcal{N}(\beta_0 | \boldsymbol{a}, \sigma_{\boldsymbol{a}}^2 \cdot \mathrm{I}) \cdot \mathcal{N}(\beta_0 | \boldsymbol{b}, \sigma_{\boldsymbol{b}}^2 \cdot \mathrm{I}),$$

where $\boldsymbol{\theta} \triangleq (\sigma_{\boldsymbol{a}}^2, \sigma_{\boldsymbol{b}}^2)^{\mathsf{T}}$. Here, the variance vector $\boldsymbol{\theta}$ is unknown.

Defining $\boldsymbol{\mu}_{\boldsymbol{\theta}} \triangleq \frac{\sigma_{\boldsymbol{a}}^2 \boldsymbol{b} + \sigma_{\boldsymbol{b}}^2 \boldsymbol{a}}{\sigma_{\boldsymbol{a}}^2 + \sigma_{\boldsymbol{b}}^2}$ and $\sigma_{\boldsymbol{\theta}}^2 \triangleq \frac{\sigma_{\boldsymbol{a}}^2 \sigma_{\boldsymbol{b}}^2}{\sigma_{\boldsymbol{a}}^2 + \sigma_{\boldsymbol{b}}^2}$, we obtain

$$\hat{\beta}_0(\hat{\boldsymbol{\theta}}) = \frac{s^2 \boldsymbol{\mu}_{\hat{\boldsymbol{\theta}}} + \sigma_{\hat{\boldsymbol{\theta}}}^2 \boldsymbol{y}}{s^2 + \sigma_{\hat{\boldsymbol{\theta}}}^2}.$$

It is readily seen that $\hat{\beta}_0(\hat{\boldsymbol{\theta}})$ will be either $\boldsymbol{a}$ or $\boldsymbol{b}$ exactly when one of the components of the variance vector $\hat{\boldsymbol{\theta}} = (\sigma_{\boldsymbol{a}}^2, \sigma_{\boldsymbol{b}}^2)$ is zero.

NUV Priors for SPARCs

The decoding of SPARCs is a straightforward extension from the previously discussed case to general linear systems. Concretely,

- the channel observation $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{\beta} + \boldsymbol{w} = \sum_{\ell=1}^{L} \boldsymbol{A}_\ell \boldsymbol{\beta}_\ell + \boldsymbol{w}$, where $\boldsymbol{A} \triangleq (\boldsymbol{A}_1 | \boldsymbol{A}_2 | \cdots | \boldsymbol{A}_L)$, $\boldsymbol{\beta}^\mathsf{T} \triangleq (\boldsymbol{\beta}_1^\mathsf{T} | \boldsymbol{\beta}_2^\mathsf{T} | \cdots | \boldsymbol{\beta}_L^\mathsf{T})$;
- the likelihood function $p(\boldsymbol{y}|\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{A}\boldsymbol{\beta}|\boldsymbol{y}, \sigma^2 \cdot \mathrm{I})$;
- the corresponding NUV prior $\rho(\boldsymbol{\beta}, \boldsymbol{\theta}) \triangleq \prod_{\ell=1}^{L} \rho(\boldsymbol{\beta}_\ell, \boldsymbol{\theta}_\ell)$;
- $\rho(\boldsymbol{\beta}_\ell, \boldsymbol{\theta}_\ell) \triangleq \prod_{k=1}^{M} \mathcal{N}(\boldsymbol{\beta}_\ell | \boldsymbol{a}_\ell^{[k]}, \sigma^2_{\boldsymbol{a}_\ell^{[k]}} \cdot \mathrm{I})$ and $\boldsymbol{\theta}_\ell \triangleq (\sigma^2_{\boldsymbol{a}_\ell^{[1]}}, \cdots, \sigma^2_{\boldsymbol{a}_\ell^{[M]}})^\mathsf{T}$ for $\ell \in [L]$.

NUV Priors for SPARCs

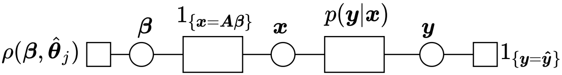
We use the expectation-maximization (EM) algorithm to approximate the parameter estimation, i.e.,

$$\hat{\boldsymbol{\theta}}^{(j+1)} = \sum_{\ell=1}^{L} \sum_{k=1}^{M} \underset{\sigma^2_{\boldsymbol{a}_\ell^{[k]}}}{\operatorname{argmax}} \mathbb{E}_{p(\boldsymbol{\beta}|\boldsymbol{y},\hat{\boldsymbol{\theta}}^{(j)})}[\ln \mathcal{N}(\boldsymbol{\beta}_\ell | \boldsymbol{a}_\ell^{[k]}, \sigma^2_{\boldsymbol{a}_\ell^{[k]}} \cdot \mathrm{I})].$$

It is readily seen that all individual parameters are given by

$$(\sigma^2_{\boldsymbol{a}_\ell^{[k]}})^{(j+1)} = \frac{1}{M}\Bigg(\sum_{\hat{k}=1}^{M} \mathrm{Var}[(\boldsymbol{\beta}_\ell)_{\hat{k}}] + \sum_{\hat{k}=1}^{M} \left(\mathbb{E}[(\boldsymbol{\beta}_\ell)_{\hat{k}}]\right)^2 + nP_\ell - 2\sqrt{nP_\ell}\,\mathbb{E}[(\boldsymbol{\beta}_\ell)_k]\Bigg). \quad \text{(EM1)}$$

NUV Priors for SPARCs



Once the variance vector $\hat{\boldsymbol{\theta}}^{(j)}$ is given, the whole statistical model for estimating $\boldsymbol{\beta}$ is shown as above. The posterior distribution of $\boldsymbol{\beta}$ can be estimated via Gaussian message passing thanks to the NUV priors provided for $\boldsymbol{\beta}$.

Assume the prior distribution of $\boldsymbol{\beta}$ is $\mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})$, then the mean vector $\widetilde{\boldsymbol{\mu}}_{\boldsymbol{\beta}}$ and the covariance matrix $\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}}$ of the posterior distribution are as follows:

$$\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}} = \left(\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1} + \sigma^{-2} \cdot \boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}\right)^{-1}, \quad \text{(EM2)}$$

$$\widetilde{\boldsymbol{\mu}}_{\boldsymbol{\beta}} = \widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}} \cdot \left(\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1}\boldsymbol{\mu}_{\boldsymbol{\beta}} + \sigma^{-2} \cdot \boldsymbol{A}\hat{\boldsymbol{y}}\right). \quad \text{(EM3)}$$

An EM-NUV algorithm for MIMO detection

Computing Eqs.(EM1)-(EM3) together iteratively comprises the so-called EM-NUV algorithm for decoding SPARCs.
Replacing $\boldsymbol{A}$ with $\widetilde{\boldsymbol{H}}\boldsymbol{A}$, the EM-NUV algorithm works for SPARCs-encoded MIMO detection. More specifically, the original MIMO detection problem can be reformulated as decoding SPARCs with a modified design matrix, i.e.,

$$\begin{aligned}\boldsymbol{y} &= \widetilde{\boldsymbol{H}}\boldsymbol{x} + \boldsymbol{w} \\ &= \widetilde{\boldsymbol{H}}(\boldsymbol{A}\boldsymbol{\beta}) + \boldsymbol{w} = (\widetilde{\boldsymbol{H}}\boldsymbol{A})\boldsymbol{\beta} + \boldsymbol{w}.\end{aligned}$$

The most time-consuming part of this algorithm is to compute the covariance matrix via (EM2) since it involves a matrix multiplication and a matrix inversion.

A Hadamard-based Gaussian GAMP algorithm

Computational complexities can be significantly reduced as follows:

- apply the Gaussian generalized approximate message passing (GAMP) algorithm to the considered overall statistical model, which reduces computational complexity from $O(n^3)$ to $O(n^2)$.

A Hadamard-based Gaussian GAMP algorithm

Computational complexities can be significantly reduced as follows:

- apply the Gaussian generalized approximate message passing (GAMP) algorithm to the considered overall statistical model, which reduces computational complexity from $O(n^3)$ to $O(n^2)$.

- replace the Gaussian design matrix with the Hadamard-based design matrix and do some approximations, which reduces computational complexity from $O(n^2)$ to $O(n \log(n))$.

Overview

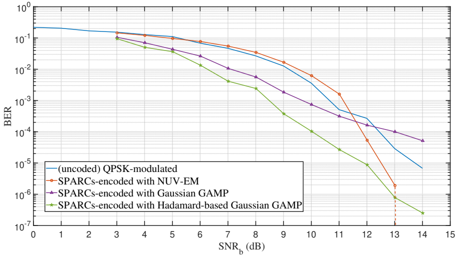
Simulation Results

We consider the following setup for our simulation results.

- we choose the rate of SPARCs R to be 0.5 and the size of each section M to be 4,
- we choose the size of the MIMO system $m_0 \times n_0$ to be 32×32,
- we choose the number of information bits k to be 64,
- we choose the damping factors $\alpha_{\boldsymbol{\beta}}$ and $\alpha_{\boldsymbol{s}}$ to be 0.8 and 0.8, respectively,
- unless otherwise stated, we choose the hyperparameter $\widetilde{\sigma}^2$ to be the noise variance σ^2.

Simulation Results

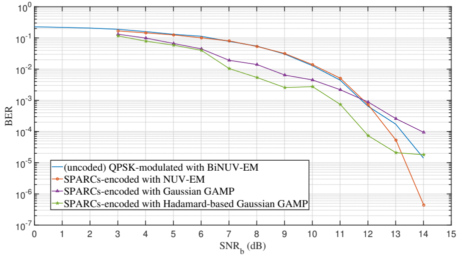
We consider MIMO matrices $\boldsymbol{H}_0$ with i.i.d. entries according to $\mathcal{CN}(0, \frac{1}{m_0})$.



- The figure shows the BER performance comparison among different algorithms and different design matrices.
- From this figure, we know that Gaussian GAMP algorithms work better than the QPSK-modulated scheme in the low-to-medium-SNR range no matter which design matrix we choose.

Simulation Results

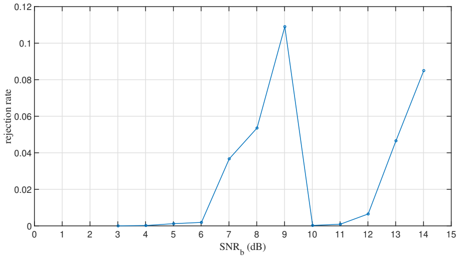
Besides the i.i.d. model, we consider the Jakes model here. A simple rejection technique is needed to tackle some stability issues commonly appearing in AMP-style algorithms.



- The figure shows the BER performance comparison among different algorithms and different design matrices for Jakes models.
- From this figure, algorithms with this simple rejection work well for correlated MIMO matrices of the Jakes model while keeping rejection rates reasonably low.

Simulation Results

Besides the i.i.d. model, we consider the Jakes model here. A simple rejection technique is needed to tackle some stability issues commonly appearing in AMP-style algorithms.



- The figure shows the corresponding rejection rates of our proposed Hadamard-based Gaussian GAMP algorithm.
- From this figure, there is a dramatic decrease of the rejection rate and no performance improvement from 9dB to 10dB since we choose $\widetilde{\sigma}^2 = 4\sigma^2$ in the high-SNR range (i.e., 10dB–14dB).

Overview

- In order to incorporate SPARCs into the NUV framework, we generalize NUV priors for scalar constraint sets to NUV priors for sparse vector constraint sets, and we derive the corresponding NUV-EM algorithm accordingly.

- In order to incorporate SPARCs into the NUV framework, we generalize NUV priors for scalar constraint sets to NUV priors for sparse vector constraint sets, and we derive the corresponding NUV-EM algorithm accordingly.
- In order to avoid matrix inversions and matrix multiplications, a Hadamard-based Gaussian generalized AMP (GAMP) is proposed for the MIMO detection problem.

- In order to incorporate SPARCs into the NUV framework, we generalize NUV priors for scalar constraint sets to NUV priors for sparse vector constraint sets, and we derive the corresponding NUV-EM algorithm accordingly.
- In order to avoid matrix inversions and matrix multiplications, a Hadamard-based Gaussian generalized AMP (GAMP) is proposed for the MIMO detection problem.
- In order to tackle some stability issues, a simple rejection technique is presented for correlated channels.

- In order to incorporate SPARCs into the NUV framework, we generalize NUV priors for scalar constraint sets to NUV priors for sparse vector constraint sets, and we derive the corresponding NUV-EM algorithm accordingly.
- In order to avoid matrix inversions and matrix multiplications, a Hadamard-based Gaussian generalized AMP (GAMP) is proposed for the MIMO detection problem.
- In order to tackle some stability issues, a simple rejection technique is presented for correlated channels.
- Simulation results show that our channel coding scheme (consistently) outperforms the previous state-of-the-art results for independent channels and it also shows excellent performance with reasonable rejection rates for correlated channels.

Thanks!