Low Resolution Adaptive Compressed Sensing for mmWave MIMO receivers

Cristian Rusu, Nuria González-Prelcic and Robert W. Heath
Motivation
Power consumption at mmWave in a MIMO receiver

Infeasible to dedicate a separate RF chain and ADC for each antenna

Alternative mmWave MIMO architectures are needed*

Low resolution mmWave MIMO receiver

- Output is heavily quantized
  - Treating as Gaussian noise may not be a good approximation
- Reasonable capacity is possible at moderate SNR values*
- May have higher baseband complexity
- Channel estimates are useful at TX/RX for configuring the precoders/combiner

With few bits
\[ y = \mathcal{Q}(Hs + v) \]

With one bit
\[ y = \text{sign}(Hs + v) \]

Threshold in real / imaginary

2b bits per complex dimension

1 bit, 240 Gs/s
Much less at 4 Gs/s

<<10mW

Baseband Processing

RF Chain
b-bit ADC

Transmit Processing

RF Chain
b-bit ADC

See also extensive work by research groups led by U. Madhow, J. Nossek, G. Fettweis, C. Studer, G. Kramer, and O. Dabeer and others
Low resolution channel estimation at mMwave

Channel estimation at mMwave

... and with few-bits ADCs

- Estimation error with few-bit ADCs decreases at best quadratically per bit*
- It also decreases with the sparsity of the channel*


Channel estimation is challenging at mmWave, more with a few-bits RX
Problem formulation
MmWave narrowband channel model

Physical spatial model

Virtual model

\[ H = A_R H_b A_T^* \]

\[ H_b = \text{diag}(\alpha) \]

Path gains

Array steering/response vectors evaluated at the AoAs/AoDs

Spatial resolution

Virtual angles fixed a priori

\[ H \approx \tilde{A}_R \tilde{H}_b \tilde{A}_T^* \]

k-sparse matrix of size GxG

Dictionaries of TX/RX array steering/response vectors with quantized angles

- Few scattering clusters at mmWave – sparsity
- We assume \( N_r = N_t = G \)  

Narrowband received signal model

Training sequence \( T \in \mathbb{C}^{N_t \times p} \)

Received signal

\[
Y = Q \left( HT + N \right)
\]

\[
= Q \left( A_r H_v A_r^* T + N \right)
\]

Choose as DFT matrices \( G=N_t=N_r \)
Channel estimation with low resolution ADCs as a sparse recovery problem

- For a training sequence $\mathbf{T} \in \mathbb{C}^{N_t \times p}$ the vectorized received signal is

$$\text{vec}(\mathbf{Y}) = \mathcal{Q} \left( (\mathbf{T}^T \bar{\mathbf{A}}_t \otimes \mathbf{A}_r) \text{vec}(\mathbf{H}_v) + \text{vec}(\mathbf{N}) \right)$$

- Measurement matrix
- Sparse vector
- iid Gaussian noise
- Quantized measurement

$$\mathbf{y} = \mathcal{Q} (\tilde{\mathbf{A}} \mathbf{x} + \mathbf{n})$$

CS problem with quantized and noisy measurements
Prior work on low resolution mmWave channel estimation

- 1-bit mmWave channel estimation reformulated via conventional one-bit CS*
- Adaptive 1-bit CS based mmWave channel estimation**
- No work on adaptive CS w/ low resolution (>1 bit) noisy measurements
- Adaptive estimation could help to reduce training overhead

What is the potential of adaptive CS?

BOUND ON THE RECONSTRUCTION ERROR

Classical, non-adaptive CS*

- Sparsity level
- Noise variance
- Measurement matrix

\[ E\{\|x - \hat{x}\|_2^2\} \leq C \frac{s\sigma^2}{\|A\|_F^2} N \log N \]

Adaptive CS**

\[ E\{\|x - \hat{x}\|_2^2\} \leq C \frac{s\sigma^2}{\|A\|_F^2} N \]

- Adaptive sensing schemes may not lead to dramatic improvements in MSE
- Is adaptation still worthwhile in a mmWave one-bit/few-bits receiver?


Adative CS with quantized and noisy measurements
Contributions

- An adaptive CS algorithm design from quantized noisy measurements
- Application to channel estimation at mmWave with a low resolution RX
- Training sequence adaptation to reduce MSE in the estimation
- Performance analysis of the algorithm in terms of
  - Resolution of the ADC
  - Training length
  - SNR
  - Sparisty level in the channel
- Comparison to conventional CS approaches
Adaptive CS based on quantized and noisy measurements

\[ y = \mathcal{Q}(Ax + n) \]

- Sparse recovery from **quantized and noisy measurements**
  - General Expectation Maximization (GEM) algorithm based on fixed \( A^* \)
    - Considers noise and quantization together
    - Estimates the noise variance from the data
    - Considers the joint distribution between the observed (quantized) data and missing (unquantized) data
    - Convergence to ML estimator under mild assumptions
  - Non adaptive approaches in prior work
- We propose a new adaptive technique based on GEM

Main ideas for AGEM

Assumptions on A
- Restrict the rows of $A$ to a given set $\mathcal{A}$ of size $N$
- Assume $A$ is made up of orthonormal vectors

Main idea for solving
- Solve using GEM at successive steps
- Increase the number of measurements at every step

Adaptation of A
- After solving at each step,
- Update $A$ to maximize correlation with current estimate
The new adaptive CS algorithm - AGEM

Initialization, $i=1$
1. Set the total measurement budget to $m$
2. Randomly select $A_1$ from $\mathcal{A}$
3. Compute $m/2$ measurements in $y_1$
4. Solve $x_1 = \text{GEM}(A_1,y_1,s)$

Iterate, $i=i+1$
1. Take $\lceil m/2^i \rceil$ new measurement vectors $a^T_j$ such that $|a_j^T \hat{x}_{i-1}|$ is maximized
2. Add them to $A_{i-1}$ to construct $A_i$
3. Compute the new measurements $y_i$
4. Solve $x_i = \text{GEM}(A_i,y_i,s)$
5. Return to step 1 until reaching the total measurement budget
Adaptive channel estimation using AGEM - 1

\[ y = Q(Ax + n), \quad A \in \mathbb{C}^{m \times N}, \quad x \in \mathbb{C}^{N} \]

\[ m = pN_r \]
\[ N = N_t N_r \]

Measurement matrix

\[ \text{vec}(Y) = Q \left( \begin{bmatrix} T^T \bar{A}_t \otimes A_r \end{bmatrix} \text{vec}(H_v) + \text{vec}(N) \right) \]

- \( A_t \) and \( A_r \) are fixed, depend on the antenna array geometry
- An adaptive measurement matrix is obtained adapting the training pilots
  - Each training pilot contributes \( N_r \) new measurements at receiver
  - More training pilots \( p \) → more measurements \( m=pN_r \)
Adaptive channel estimation using AGEM - 2

- Adaptation requires **limited feedback** to the transmitter
  - RX computes the channel estimate at every step
  - At every step, feedbacks the new measurement vectors from $\mathcal{A}$
    - If $\mathcal{A}$ is of size $N_t$, we need $\log N_t$ bits for the feedback
  - For the last step, feedback the channel estimate
Simulation results
Performance analysis of AGEM in terms of SNR

- It is not worthwhile to use more than 3 bits in the ADCs
- Adaptive CS reduces MSE if SNR is not very low

Adaptive CS overcomes conventional CS when SNR > -3 dB
Performance analysis of AGEM in terms of training length

- For AGEM there is a working regime in terms of the number of pilots
- If there are too many measurements AGEM and GEM perform equivalently

Adaptive CS outperforms conventional CS when using a few pilots

SNR=5dB
\( s=8 \)
\( N_t=N_r=16 \)

Too few measurements

Plenty of measurements
Performance analysis of AGEM in terms of sparsity

Adaptive CS is worthwhile for any sparsity level

SNR=5dB
p=8
N_t=N_r=16
How the MSE reduction in the channel estimate impacts on the achievable rate?
Achievables rates in terms of SNR

- Improvement on achievable rate by adaptive CS does not depend on $q$
- Significant improvement on achievable rate if SNR is not very low
Conclusions

- Adaptive CS overcomes conventional CS when SNR is not too low
  - Works various sparsity levels
  - Helps reduces training length and thus overheads

- If many measurements are possible, adaptive does not make sense

Quantization does not dramatically affect performance for channel estimation