Channel Estimation and Precoder Design for mmWave Communications: The Sparse Way

Phil Schniter

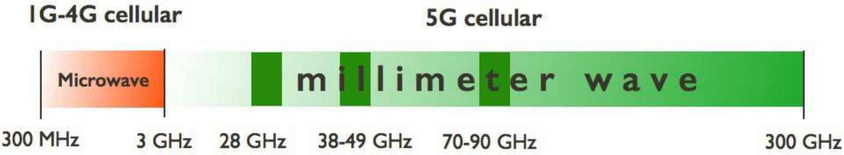
THE OHIO STATE UNIVERSITY

Supported in part by NSF grants CCF-1018368 and CCF-1218754.

(And thanks to Robert Heath, Jr. for lending me a few graphics!)

Asilomar — Nov'14

mmWave Communications



Potential:[1]

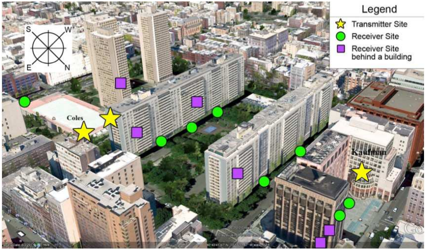
- Huge amount of bandwidth available $\rightarrow$ Huge throughput?
- Many antennas fit in a small form-factor $\rightarrow$ Massive MIMO?

Challenges:

- Path-loss/shadowing are ~ 40dB worse than in microwave bands.
- Huge bandwidth leads to serious implementational issues.

[1] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges" *Proc. IEEE*, Mar. 2014.

mmWave Channel Sparsity



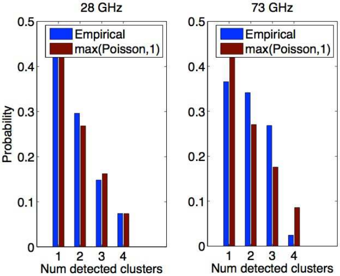
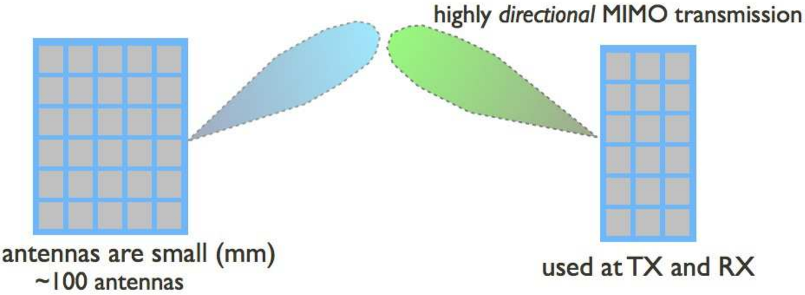


Fig. 4: Distribution of the number of detected clusters at 28 and 73 GHz. The measured distribution is labeled 'Empirical', which matches a Poisson distribution (3) well.

- Physical measurements in dense urban NLOS environments suggest that mmW channels are extremely sparse.[2]
- Can expect at most 3-4 clusters, with very little angle/delay-spread per cluster.

[2] M. Akdeniz, Y. Liu, S. Sun, S. Rangan, T. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE JSAC*, June 2014

mmW uses Massive Arrays



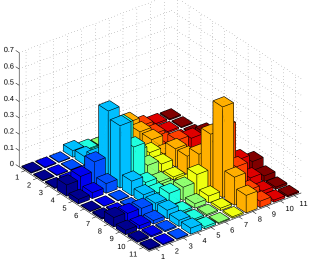
- To counter path-loss, massive arrays are used at both Tx and Rx.
- The goal is beamforming gain, not spatial multiplexing gain. (These systems are power-limited, not bandwidth-limited.)
- Narrow beams also reduce fading, multipath, and interference.

Contributions

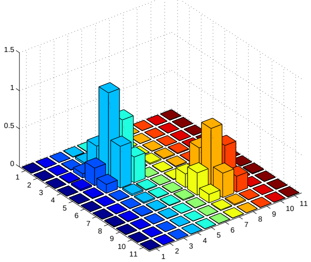
We propose...

- sparsity-exploiting low-complexity space-time channel estimation,
- mutual-information-maximizing beamforming & waterfilling,
- aperture shaping to ensure that physical sparsity manifests as MIMO channel sparsity.

unshaped



shaped



System Model

$N \times N_{\mathsf{r}}$ matrix of received samples at block t:

$$\boldsymbol{Y}_t = \sum_{d=0}^{N_{\mathsf{d}}-1} \boldsymbol{J}_d \boldsymbol{X}_t \boldsymbol{H}_d + \boldsymbol{W}_t$$

where

$\boldsymbol{J}_d$: cyclic d-delay matrix

$\boldsymbol{X}_t$: $N \times N_{\mathsf{t}}$ transmitted signal at block index t

$\boldsymbol{H}_d$: $N_{\mathsf{t}} \times N_{\mathsf{r}}$ MIMO channel at delay d

$\boldsymbol{W}_t$: AWGN of variance ν_{w}

and

N_{d} : channel delay spread

N : length of block transmission (plus N_{d}-length cyclic prefix)

N_{t} : number of transmit antennas

N_{r} : number of receive antennas.

MIMO Channel Models

Per-path channel parameters:

$$(\beta_l, \tau_l, \theta_{\mathsf{t},l}, \theta_{\mathsf{r},l}) = (\text{gain}_l,\ \text{delay}_l,\ \text{transmit-angle}_l,\ \text{receive-angle}_l)$$

MIMO channel matrix at delay d:

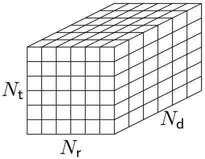
$$\boldsymbol{H}_d = \sum_{l=1}^{L} \beta_l \, p_{\mathsf{srrc}}(d\,T_c - \tau_l)\, \boldsymbol{f}_{N_\mathsf{t}}(\theta_{\mathsf{t},l})\, \boldsymbol{f}_{N_\mathsf{r}}(\theta_{\mathsf{r},l})^{\mathsf{H}}$$

note : $\tau_l, \theta_{\mathsf{t},l}, \theta_{\mathsf{r},l}$ are not discrete!

Virtual[3] MIMO channel matrix at delay d:

$$\boldsymbol{G}_d = \boldsymbol{F}_{N_\mathsf{t}}^{\mathsf{H}} \boldsymbol{H}_d \boldsymbol{F}_{N_\mathsf{r}}$$

$\boldsymbol{F}_{N_\mathsf{t}}, \boldsymbol{F}_{N_\mathsf{r}}$: unitary DFT matrices.

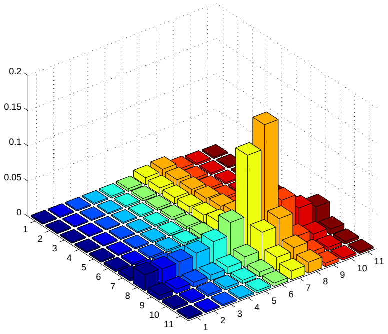


[3] A. M. Sayeed, "Deconstructing multi-antenna fading channels," *IEEE TSP*, Oct. 2002.

Sparsity of the Virtual MIMO Channel

The elements of $\boldsymbol{G}_d$ are the complex channel gains at *discrete* transmit and receive angles and discrete delay d.

Example of virtual MIMO coefficients due to a single path ($N_\mathsf{t} = 11 = N_\mathsf{r}$):



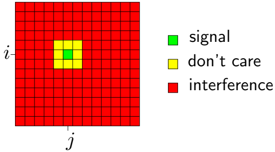
Sparse physical scattering does not yield sparse virtual channel coefs!

The Shaped Virtual MIMO Channel

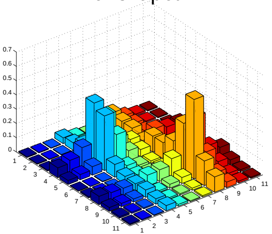
We can restore angle-domain sparsity via aperture shaping,[4] i.e., windowing of the transmit and receive antenna gains:

$$\bar{\boldsymbol{G}}_d = \boldsymbol{F}_{N_\mathsf{t}}^\mathsf{H} \mathrm{diag}(\boldsymbol{w}_\mathsf{t}) \boldsymbol{H}_d \mathrm{diag}(\boldsymbol{w}_\mathsf{r}^*) \boldsymbol{F}_{N_\mathsf{r}}$$

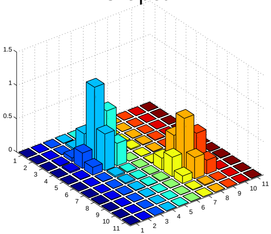
where max-SINR windows $\boldsymbol{w}_\mathsf{t}, \boldsymbol{w}_\mathsf{r}$ are solved via a generalized-eigenvector problem.



unshaped



shaped



[4] P. Schniter and A. M. Sayeed, "A sparseness-preserving virtual MIMO channel model," *Proc. CISS*, 2004.

Training Sequence Design

To facilitate low-complexity channel estimation, we propose to

1. construct the space-time training signal as $\boxed{\boldsymbol{X}_t = \boldsymbol{F}_N^{\mathsf{H}} \boldsymbol{S}_t \boldsymbol{F}_{N_{\mathsf{t}}}}$, where $\boldsymbol{S}_t$ has i.i.d entries in $\{1, j, -1, -j\}$,
2. FFT-process the observations, giving the observation structure

$$\boxed{\begin{aligned} \boldsymbol{F}_N \boldsymbol{Y}_t \boldsymbol{F}_{N_{\mathsf{r}}} &= \mathsf{AWGN}(\nu_{\mathsf{w}}) \\ &+ \sqrt{N} \left[\operatorname{diag}(\boldsymbol{s}_{t,1}) \boldsymbol{F}_{N \times N_{\mathsf{d}}} \cdots \operatorname{diag}(\boldsymbol{s}_{t,N_{\mathsf{t}}}) \boldsymbol{F}_{N \times N_{\mathsf{d}}} \right] \left[\boldsymbol{g}_1 \cdots \boldsymbol{g}_{N_{\mathsf{r}}} \right] \end{aligned}}$$

   where $\boldsymbol{g}_j \in \mathbb{C}^{N_{\mathsf{d}} N_{\mathsf{t}} \times 1}$ contains virtual chan coefs for jth Rx antenna,
3. and, if needed, stack measurements across T blocks, giving a total of NT scalar measurements per $N_{\mathsf{d}} N_{\mathsf{t}}$ scalar unknowns.

Note the near isometry & fast implementation of the training operator.

To Compress or Not To Compress?

Sub-Nyquist regime ($NT < N_t N_d$):

- Low training overhead.
- Requires a sparse reconstruction algorithm.

Super-Nyquist regime ($NT \geq N_t N_d$):

- Higher training overhead.
- Allows classical linear (e.g., LS, LMMSE) estimation.
- Sparse reconstruction can improve performance at very low SNR.

For example, 802.11ad (60GHz) standard uses $N = 512$ and $N_d = 128$.
So $N_t = 64$ Tx antennas $\Rightarrow$ $T \geq 16$ blocks for Nyquist sampling.

Numerical Examples

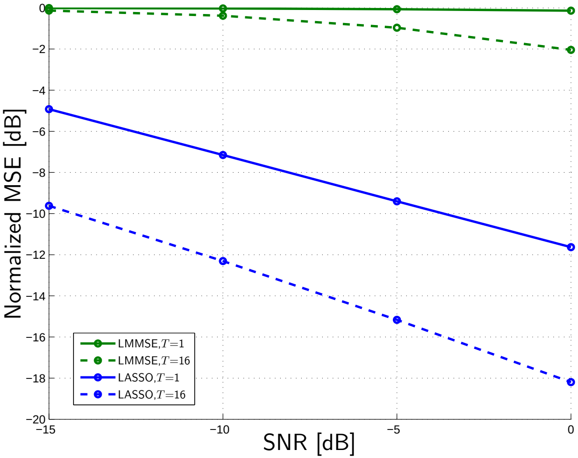
System parameters:

$$
\begin{aligned}
N &= 512 \ \text{ block length} \\
N_\mathsf{d} &= 128 \ \text{ channel delay spread} \\
N_\mathsf{t} &= 64 \ \text{ transmit antennas} \\
N_\mathsf{r} &= 64 \ \text{ receive antennas } (\Rightarrow \text{SNR gain } = 18\text{dB}) \\
\text{SNR} &\sim -8\text{dB} \ \ (\Rightarrow \text{subcarrier SNR} \sim 10\text{dB}) \\
L &= 4 \ \text{ i.i.d Rayleigh paths with uniform delay, Tx angle, Rx angle}
\end{aligned}
$$

Two training lengths considered:

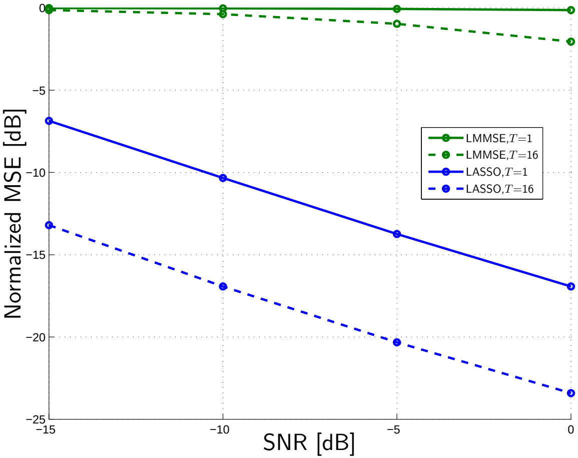
$$
T \in \begin{cases} 1 \text{ block} & \Rightarrow \frac{1}{16} \text{ Nyquist rate} \\ 16 \text{ blocks} & \Rightarrow \text{Nyquist rate} \end{cases}
$$

Numerical Example: Channel Estimation (without shaping)



Sparse reconstruction (via LASSO) shows significant gain over LMMSE at both sub-Nyquist $(T = 1)$ and Nyquist $(T = 16)$ sampling rates.

Numerical Example: Channel Estimation (with shaping)



Aperture shaping yields 2-5dB reduction in LASSO's channel estimation error.

Beamforming and Waterfilling

1. Construct the data matrix as $\boxed{\boldsymbol{X}_t = \boldsymbol{F}_N^{\mathsf{H}} \operatorname{diag}(\sqrt{\boldsymbol{p}}) \boldsymbol{s}_t \boldsymbol{b}^{\mathsf{H}} \boldsymbol{F}_{N_{\mathsf{t}}}}$ with power allocation $\boldsymbol{p} \in \mathbb{C}^{N \times 1}$, QAM $\boldsymbol{s}_t$, and beamformer $\boldsymbol{b} \in \mathbb{C}^{N_{\mathsf{t}} \times 1}$.

2. The observations decouple! A sufficient statistic to estimate $s_{t,n}$ is

$$\boxed{[\boldsymbol{F}_N \boldsymbol{Y}_t \boldsymbol{F}_{N_{\mathsf{r}}}]_{n,:} = s_{t,n} \sqrt{p_n} \boldsymbol{b}^{\mathsf{H}} \boldsymbol{G}_n + \mathsf{AWGN}(\nu_{\mathsf{w}})}$$

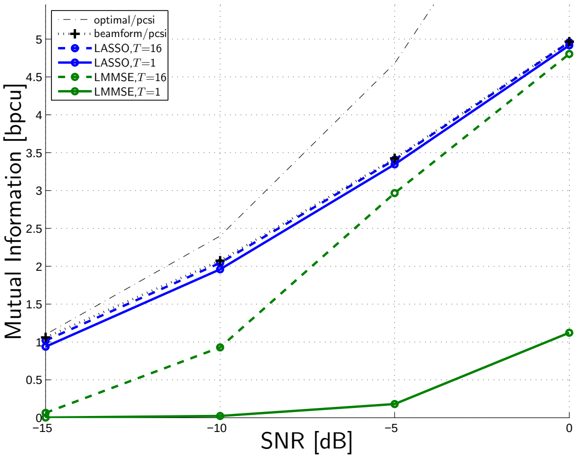
   where $\boldsymbol{G}_n \in \mathbb{C}^{N_{\mathsf{t}} \times N_{\mathsf{r}}}$ is the MIMO channel at *frequency* bin n.

3. Can solve for mutual-information maximal beamformer/powers via

$$\boxed{\arg\max_{\boldsymbol{p},\boldsymbol{b}} \sum_{n=1}^{N} \log_2 \left(1 + p_n \frac{\boldsymbol{b}^{\mathsf{H}} \boldsymbol{G}_n \boldsymbol{G}_n^{\mathsf{H}} \boldsymbol{b}}{\nu_{\mathsf{w}}} \right) \text{ s.t.} \begin{cases} \sum_{n=0}^{N-1} p_n = N, \; p_n \geq 0, \\ \|\boldsymbol{b}\|_2 = 1. \end{cases}}$$

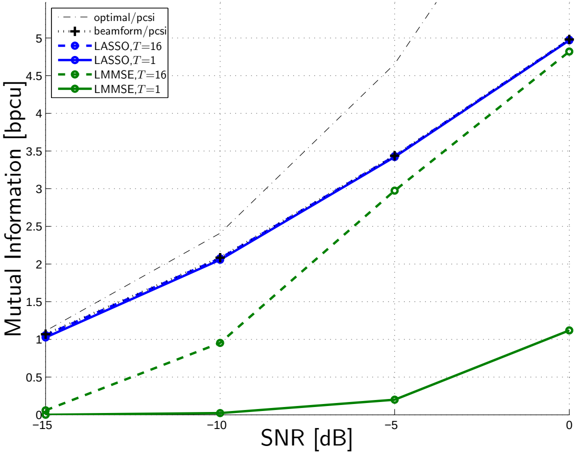
In practice, use *estimated* channel $\hat{\boldsymbol{G}}_n$ for Rx combining and $(\boldsymbol{p}, \boldsymbol{b})$ design.

Numerical Example: Spectral Efficiency (without shaping)



- Beamforming & waterfilling is near-optimal at low SNR.
- LASSO performs nearly as well as perfect CSI, even under compressed pilots (T=1).
- LMMSE is significantly suboptimal except at high SNR and Nyquist-rate pilots (T=16).

Numerical Example: Spectral Efficiency (with shaping)



Aperture shaping yields 0.5dB SNR gain in the $T=1$ case, closing the gap between LASSO and perfect-CSI.

Summary

- Considered mmW systems, which operate at very low SNR using massive antenna arrays at transmitter and receiver.
- Proposed an aperture shaping scheme that promotes sparsity in the virtual MIMO channel coefficients.
- Proposed a low-complexity space-time channel estimation scheme that exploits the extreme sparsity of mmW channels.
- Proposed a beamforming + waterfilling scheme that is near-optimal at low SNR.
- Numerical experiments suggest that LASSO channel estimates yield near-optimal spectral efficiency over a wide SNR range, even under significant pilot compression.