Signal processing at the mm wave frontier

Upamanyu Madhow

Dept. of Electrical and Computer Engineering
University of California, Santa Barbara

Plenary lecture at SPAWC 2014, June 23, 2014
Why the buzz?

• “Unlimited” spectrum
  – 30-300 GHz (with strict defn of 10-1 mm wavelength)
  – 60 GHz has received the most recent attention (unlicensed)
  – 71-76 and 81-86 GHz for semi-unlicensed point-to-point
  – 100+ GHz: the wild west of wireless

• Why now?
  – Because we can (mass market RFICs now feasible)
  – Smart phone induced capacity crisis
  – Fits with logic of continued WiFi growth
Agenda today

- How not to fight physics
  - How tiny wavelengths impact us
  - What applications are a natural fit to mm wave
- Case study: from application to theory and back
  - Xtreme spatial reuse via 1000 elt antennas ➔ new theory of compressive estimation
  - Algorithms for attaining CRB
  - Tracking users for mm wave to the mobile
- Challenges and opportunities
  - New MIMO architectures
  - Very high bandwidths

Message: Beyond the hype lie significant intellectual opportunities
How not to fight physics
High directionality is essential

Friis formula for free space propagation

\[
P_{RX} = P_{TX} \cdot G_{TX} \cdot G_{RX} \cdot \frac{\lambda^2}{16\pi^2R^2}, \text{ in terms of antenna gains}
\]

\[
P_{RX} = P_{TX} \cdot \frac{A_{TX}A_{RX}}{\lambda^2R^2}, \text{ in terms of antenna apertures}
\]

Highly directional antennas critical for adequate link budget
High directionality attainable with reasonable form factor
High directionality is attainable

Massive MIMO in your palm
32 x 32 element array fits within 8cm x 8cm
Electrically large, physically small

But how would we steer such large arrays?
Objects look bigger at smaller wavelengths (Huygen’s principle)

- Cannot burn through or diffract around obstacles
- Must steer around them

Again, how do we steer large arrays?
Oxygen absorption and rain can be scary

Oxygen absorption
16 dB/km at 60 GHz

Rain hits all mm wave freqs badly
~ 20 dB/km at 50 mm/hr

What not fighting physics means

- Must use directional TX and RX
  - Hence **electronically steerability is key** if we want flexible usage
- Must steer around, not burn through, obstacles
  - Hence **electronic steerability is key** if we want robust usage
- Should not shoot for kilometers range
  - 16 dB/km (O2) or 20 dB/km (rain) or 36 dB/km (both) are all bad news
- But can certainly go well beyond indoor WPAN
  - Oxygen absorption + heavy rain costs only 3.6 dB at 100m

What applications are consistent with these guidelines?
Example applications

Consistent with the physics
Consistent with mass market economics
Indoor focus over the past few years

WiGig/IEEE 802.11ad
- Up to 7 Gbps
- In room
- Up to 10m

60 GHz CMOS RFICs
- Antenna array in package (32 elt)
- Directional MAC

Showed feasibility of steering around obstacles
But is mm wave comm just nice to have?

- 802.11n is pretty fast already
- Once we upgrade WLAN speeds to a few Gbps, are we done?
- Not quite…
- Millimeter wave communication can play a crucial role in today’s cellular capacity crunch

Figure 2. Global Mobile Data Traffic Forecast by Region

Exabytes per Month

Source: Cisco VNI Mobile, 2012
mm wave for small cells, stage 1

• Increase cellular capacity by drastically increasing spatial reuse
  – Base stations on lampposts, 200 m cell size
  – 4G to mobile, mm wave between base stations

• MultiGigabit wireless mesh backhaul enables dense picocell deployments

Need flexible beamsteering to form mesh
Need five 9s reliability for backhaul
mm wave for small cells, stage 2

• Up the ante on spatial reuse
  – Highly directional mm wave (+LTE) to the mobile
  – 28 GHz being pushed as a possibility
  – Alternative: Downlink 60 GHz with uplink LTE feedback
  – Leverage WiGig radio on mobile device in receive-only model

Need robustness to blockage by user’s body and other obstacles
Focus today: Beamsteering with very large arrays

(The key to “unlimited” spatial reuse)
Beamforming today

DSP-centric, one RF chain per antenna element

Does not scale to 1000 elements!
RF Beamforming with hardware constraints

Coarse phase shifts
\( \phi_i \in \{0^\circ, \pm 90^\circ, 180^\circ\} \)

Much more feasible
But how do we adapt it?
No access to individual elements \( \rightarrow \) least squares does not work
Beam scanning architecture unattractive

- Requires fine control of phases
- Slow adaptation
Can we use the sparsity of the mm wave channel?
Compressive adaptation

Random phases from
$\pm 1, \pm j$

Only coarse phase control
Faster than beam scanning

Feedback from mobiles

Base station estimates channel compressively
Estimation problem

Channel is a sum of a few sinusoids

\[ h = g_1 x(\omega_1) + g_2 x(\omega_2) + g_3 x(\omega_3) \]

\[
\mathbf{x}(\omega) = \begin{pmatrix}
1 \\
e^{j\omega} \\
e^{j2\omega} \\
\vdots \\
e^{j(N-1)\omega}
\end{pmatrix}
\]

\[
\omega_i = \frac{2\pi d}{\lambda} \sin \theta_i
\]

Mobile makes compressive measurements

\[ y_i = \mathbf{a}_i^T \mathbf{h}, \quad i = 1, 2, \ldots, M \]

Estimate gains and spatial frequencies from compressive measurements
Can we use standard compressed sensing?

Observed projections $\rightarrow$ Randomized beamforming weights $\rightarrow$ Fourier Basis $\rightarrow$ Gains of active frequencies

$\mathbf{y} = \sum_{k=1}^{L} g_k \mathbf{A} \mathbf{x}(\omega_k) + \mathbf{n}$

Picture from plenary by Rich Baraniuk, ISIT 2009
Not quite: basis mismatch is the problem

Frequencies come from a continuum, not a grid

With standard CS, off-grid frequencies can have large estimation errors

\[ \frac{2\pi}{n} \]

\[ n = \# \text{ antennas} \]

Need a new theory of compressive estimation!

Sensitivity to Basis Mismatch in Compressed Sensing,
Y. Chi, L. Scharf, A. Pezeshki, R. Calderbank
Compressive estimation in AWGN

Standard parameter estimation

\[ y = s(\theta) + z, \quad z \sim \mathcal{CN}(0, \sigma^2 I_M) \]

\[ \hat{\theta}_{\text{ML}} = \arg \min_{\theta} \| y - s(\theta) \| \]

Performance measures

Cramer-Rao Bound (CRB) when close to truth

Ziv-Zakai bound (ZZB) more generally

(are you in the right bin? How close, once in the right bin?)

ZZB tends to CRB at high SNR (high prob of right bin). This is when estimation can be expected to “work well.”
Performance depends on Euclidean distances

CRB depends on Fisher Information Matrix

\[ F_{m,n}(\theta) = \frac{2}{\sigma^2} \mathbb{R} \left\{ \left( \frac{\partial s(\theta)}{\partial \theta_m} \right)^H \frac{\partial s(\theta)}{\partial \theta_n} \right\} \]

Depends on changes in signal geometry for small changes in parameter

Ziv-Zakai bound is based on an associated detection problem

\[ H_1 : y = s(\theta_1) + z, \quad \Pr(H_1) = \frac{p(\theta_1)}{p(\theta_1) + p(\theta_2)} \]

\[ H_2 : y = s(\theta_2) + z, \quad \Pr(H_2) = \frac{p(\theta_2)}{p(\theta_1) + p(\theta_2)}. \]

Depends on changes in signal geometry for general changes in parameter

\[ d(\theta_1, \theta_2) = \| s(\theta_1) - s(\theta_2) \| \]
Compressive measurements: model

High-dimensional signal space \( x(\theta) \in \mathbb{R}^N \)

(but unknown parameter lies in low-dimensional space)

\( M \) compressive measurements

\[
y_i = \langle w_i, x(\theta_t) + z_i \rangle
\]

\[
A = [w_1 \cdots w_M]^T
\]

\[
y = Ax(\theta_t) + z
\]

\[
z_i \sim \mathcal{N}(0, \sigma^2 I_N)
\]

Noise power is same

\[
z \sim \mathcal{N}(0, \sigma^2 I_M)
\]

When does this provide the “same” performance as standard estimation?
Compressive estimation works well when

1) Signal space geometry is preserved
   (similar to RIP for compressive sensing)

2) “Effective SNR” is high enough
The structure of compressive estimation

GENERAL STRUCTURE

1) Required isometries
   CRB: Preserve distance changes under small perturbations
   ZZB: Preserve distance changes generally

2) SNR penalty (⇒ “effective SNR”)
   Dimension reduction from $N$ to $M$ ⇒ SNR reduction by $M/N$

3) Definition of “working well”
   ZZB tends to CRB (coarse errors highly unlikely)

PROBLEM-SPECIFIC ANALYSIS

How many observations needed to preserve isometries?
Isometries needed

Tangent plane isometry (for CRB)

\[
\sqrt{\frac{M}{N}} (1 - \epsilon) \leq \frac{\|A \sum a_m (\partial x(\theta) / \partial \theta_m)\|}{\|\sum a_m (\partial x(\theta) / \partial \theta_m)\|} \leq \sqrt{\frac{M}{N}} (1 + \epsilon)
\]

\[\forall [a_1, a_2, \ldots, a_K]^T \in \mathbb{R}^K \setminus \{0\}, \forall \theta \in \Theta\]

Pairwise \(\epsilon\)-isometry (for ZZB)

\[
\sqrt{\frac{M}{N}} (1 - \epsilon) \leq \frac{\|A x(\theta_1) - A x(\theta_2)\|}{\|x(\theta_1) - x(\theta_2)\|} \leq \sqrt{\frac{M}{N}} (1 + \epsilon)
\]

\[\forall \theta_1, \theta_2 \in \Theta.\]
What geometry preservation looks like

All measurements

\[ y = x(\theta_t) + z \quad z \sim \mathcal{N}(0, \sigma^2 \mathbb{I}_N) \]

Compressive measurements

\[ y = Ax(\theta_t) + z \quad z \sim \mathcal{N}(0, \sigma^2 \mathbb{I}_M) \]

\[ \hat{\theta} = \arg\min_{\theta_i} \|y - x(\theta_i)\|^2 \]

\[ \hat{\theta} = \arg\min_{\theta_i} \|y - Ax(\theta_i)\|^2 \]
Why we can hope for geometry preservation

\[ \mathbf{v} = \mathbf{x}(\theta_i) - \mathbf{x}(\theta_j) \]

- Random projections must preserve norm of

\[ \frac{1}{M} \| A\mathbf{v} \|^2 = \frac{1}{M} \sum_{i=1}^{M} |\mathbf{w}_i^T \mathbf{v}|^2 \]

\( M \) large enough

concentrates

Mean \( (1/N) \| \mathbf{v} \|^2 \)

i.i.d. with mean \( (1/N) \| \mathbf{v} \|^2 \)

- Chernoff bound on deviations from the mean (with tolerance \( \epsilon \)) + Union bound (for all pairwise differences)

**Johnson-Lindenstrauss (JL) Lemma**

How many measurements?

**Johnson-Lindenstrauss (JL) lemma:**
Pairwise \( \varepsilon \)-isometry for **finite** signal model \( \mathcal{H} = \{ x(\theta_i) \} \) when the number of random projections:

\[
M = O \left( \varepsilon^{-2} \log |\mathcal{H}| \right)
\]

- \( K \) signals, \( M \) measurements
- Chernoff bound + Union bound \( \sim K^2 e^{-\alpha M} \)
- \( \Rightarrow M = O(\log K) \)
Continuous signal model

Parameters come from a continuum \( \theta \in \mathbb{R}^K \)

Need pairwise isometries for all \((\theta_1, \theta_2)\) pairs

Cannot directly use JL lemma
But discretization, JL lemma, and smoothness can be used to do the trick
How many measurements for good performance?

- If pairwise isometry holds, then both CRLB and ZZB go through
  - Only effect of compressive measurements is SNR reduction

- Number of measurements must satisfy two criteria for good performance
  - Should be enough to provide pairwise isometry
  - Effective SNR should be such that ZZB tends to CRLB

\[ \text{SNR in dB} \]

\[ \frac{K_{\text{min}}}{\sigma^2} \]

\[ \frac{N}{\sigma^2} \]
Attaining the CRB for a sinusoid

Problem-specific analysis ➔ Pairwise isometry requires

$$K = O \left( \epsilon^{-2} \log (N\epsilon^{-1}) \right)$$

More random projections
Better isometry constants

RMSE performance for 40+ measurements closely follows that for all N=256 measurements
Isometry constants good for 40+ measurements

$$K = \min(40, \text{ZZB threshold SNR} \times N\sigma^2)$$
Back to the application at hand

How to estimate a 1000-dimensional spatial channel?
Compressive adaptation

Random phases from
\pm 1, \pm j

Base station estimates channel compressively

Only coarse phase control
Faster than beam scanning

Feedback from mobiles
Estimation problem

Channel is a sum of a few sinusoids

\[ h = g_1 x(\omega_1) + g_2 x(\omega_2) + g_3 x(\omega_3) \]

\[ x(\omega) = \begin{pmatrix} 1 \\ e^{j\omega} \\ e^{j2\omega} \\ \vdots \\ e^{j(N-1)\omega} \end{pmatrix} \]

\[ \omega_i = \frac{2\pi d}{\lambda} \sin \theta_i \]

Mobile makes compressive measurements

\[ y_i = a_i^T h, \quad i = 1, 2, \ldots, M \]

Estimate gains and spatial frequencies from compressive measurements
Algorithm

• Acquisition
  – No knowledge of spatial frequencies whatsoever

• Tracking
  – Leverage frequency estimate from previous round
  – Refine based on new measurements
Acquisition: Coarse Estimate

\[ \text{maximize } F(\omega) = \left| \langle Ax(\omega), y \rangle \right|^2 \]

\[ \omega = 0, \frac{2\pi}{2N}, 2 \left( \frac{2\pi}{2N} \right), \ldots, (2N - 1) \left( \frac{2\pi}{2N} \right) \]
Acquisition: Coarse Estimate

maximize $F(\omega) = \left| \langle Ax(\omega), y \rangle \right|^2$

$\hat{\omega}_1 = \frac{\langle Ax(\hat{\omega}_1), y \rangle}{\|Ax(\hat{\omega}_1)\|^2}$
Acquisition: Refine Estimate

Given \( \hat{g}_1 \)
Refine freq: \( \hat{\omega}_1 \leftarrow \hat{\omega}_1 + \Delta_1 \)

\[
y = \hat{g}_1 A x (\hat{\omega}_1 + \Delta_1) + n
\]

\[
[y - \hat{g}_1 A x (\hat{\omega}_1)] = [\hat{g}_1 A \frac{dx(\hat{\omega}_1)}{d\omega}] \Delta_1 + n
\]

\( \Delta_1 \in \mathbb{R} \)

Given \( \hat{\omega}_1 \)
Re-estimate gain

\[
y = g_1 A x (\hat{\omega}_1) + n
\]
Multiple Frequencies

**Given**
- Gains: $\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K$
- Freqs: $\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_K$

Project out contributions from these frequencies

$S = A \begin{bmatrix} x(\hat{\omega}_1) & x(\hat{\omega}_2) & \ldots & x(\hat{\omega}_K) \end{bmatrix}$

$y_r = S^\perp y$

**Coarsely estimate (K+1)th freq**

$max \left| \langle Ax(\omega), y_r \rangle \right|^2$

$\omega = 0, \frac{2\pi}{2N}, 2\left(\frac{2\pi}{2N}\right), \ldots, (2N - 1)\left(\frac{2\pi}{2N}\right)$

$\hat{\omega}_{K+1}, \hat{g}_{K+1}$

**Fix gains:** $\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K, \hat{g}_{K+1}$

**Refine freqs:** $\Delta_1, \Delta_2, \ldots, \Delta_{K+1}$

**Fix freqs:** $\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_{K+1}$

**Estimate g’s:** $\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K, \hat{g}_{K+1}$
Multiple Frequencies

Coarsely estimate \((K+1)\)th freq

\[
\max_{\omega} |\langle Ax(\omega), y_r \rangle|^2
\]

\[
\omega = 0, \frac{2\pi}{2N}, \frac{2\pi}{2N}, \ldots, (2N-1) \left( \frac{2\pi}{2N} \right)
\]

\[\hat{\omega}_{K+1}, \hat{g}_{K+1}\]

Fix gains: \(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K, \hat{g}_{K+1}\)

Refine freqs: \(\Delta_1, \Delta_2, \ldots, \Delta_{K+1}\)

Fix freqs: \(\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_{K+1}\)

Estimate g’s: \(\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_K, \hat{g}_{K+1}\)

Same algorithm works for tracking, just bootstrap with estimate from prior round
Simulation Setup

Array on lamp post
Results

Estimated number of beams

Estimation errors close to CRB
Take-aways from this case study

• Unique challenges of adapting large mm wave arrays
  ➔ Compressive adaptation approach
  ➔ New theory of compressive estimation
• New insight on algorithms attaining CRB
  – Coarse grid, then gradient or Newton based refinement does work
    (If SNR is high enough to get past ZZB threshold)
• Specific motivating application, but leads to rather general techniques
This is only the beginning...

A sampling of new and exciting problems in SPAWC
New MIMO paradigms
MIMO at small carrier wavelengths does not need “rich scattering”

- Degrees of freedom depend on form factor

\[ y_m = \frac{1}{N} \sum_{n=1}^{N} \exp \left( -\frac{i \pi}{\lambda R} (m \lambda - m \lambda) \right) x_n \]

\[ y(p) = \frac{1}{L_T} \int_{-L_T/2}^{L_T/2} \exp \left( -\frac{i \pi}{\lambda R} (q - p)^2 \right) x(q) dq \]

\[ h_{m,n} \]

\[ h(q, p) \]

\[ \lambda_n \phi_n(t) = \int_{-T/2}^{T/2} \frac{\sin 2\pi W(t-s)}{\pi(t-s)} \phi_n(s) ds \]

\[ \frac{L_T^2}{\lambda R} |g_n|^2 \alpha_n(q) = \int_{-L_T/2}^{L_T/2} \frac{\sin 2\pi \frac{L_T}{2\lambda R} (q - q')}{\pi(q - q')} \alpha_n(q') dq' \]

\[ |g_n|^2 \approx 0 \text{ for } n > \frac{L_T L_R}{\lambda R} (1 + \epsilon) \]
Can utilize all degs of freedom with finite # antennas

\[ h_1 = (1, e^{j\phi}, e^{j2^2\phi}, \ldots, e^{j(N-1)^2\phi})^T \]
\[ h_2 = (e^{j\phi}, 1, e^{j\phi}, \ldots, e^{j(N-2)^2\phi})^T \]

\[ |\langle h_1, h_2 \rangle| = \left| \frac{\sin(N\phi)}{\sin \phi} \right| \]

Vectors are orthogonal when \( N\phi = N \frac{\pi d^2}{\lambda R} = \pi \)

\[ d = \sqrt{\frac{\lambda R}{N}} \]

**Example**

\( R=10 \text{ m}, \lambda=5 \text{ mm}, \ N=4 \)
\( d = 11 \text{ cm} \)

Demonstrating LoS MIMO: 4x4 Prototype

- Embedded pilot tones used to identify channels at the receiver
- Decouple receiver functions: channel separation and data demodulation
- Channel separation network implemented with baseband analog circuits

Spatial multiplexing for WiGig

- 14 Gbps on a WiGig channel
- 28 Gbps on a WiGig channel

Small phone form factor
2X spatial multiplexing

Larger tablet form factor
4X spatial multiplexing
Canonical architecture for mm wave MIMO

Rayleigh-spaced arrays: spatial multiplexing
(Shorter spacing: diversity)
Each array is a sub-wavelength spaced subarray: beamforming

RF beamforming per subarray. Mixed signal processing across.
Distributing subarrays to sidestep form factor constraints

The road to long-range wireless fiber: finally a compelling case for relays

Irish, Quitin, Madhow, ITA 2013
Signal processing for multi-GHz signals

How to scale system bandwidth indefinitely?
How to keep riding Moore’s law?
The bandwidth scaling problem

• We like riding Moore’s law
  – Enables economies of scale for cellular and WiFi
  – Keeps going at multiGigabit speeds

• The ADC is the bottleneck
  – High-rate, high-precision ADC costly, power-hungry and/or not available
  – Forces us beyond the OFDM comfort zone

• Clever solutions with low-precision ADC (1-4 bits)?
  – OK if we can keep dynamic range under control

• Time-interleaved ADCs?
  – Each sub-ADC still sees the full bandwidth

Is there a natural successor to OFDM as we scale bandwidth?
Analog Multitone for indefinite scalability

- Off-the-shelf ADC technology determines subchannel speed
- Desired bandwidth determines number of subchannels (much fewer than number of subcarriers in OFDM)
- Analog channelization at transmitter and receiver
- Sophisticated DSP for each subchannel: combat both ISI and ICI
- Promising simulation results for 1 x 2 60 GHz backhaul link

Zhang, Venkateswaran, Madhow, Analog multitone with interference suppression: relieving the ADC bottleneck for wideband 60 GHz systems, IEEE Globecom 2012.
Example: 60 GHz backhaul link

2 GHz bandwidth, 200 m range, two receive antenna arrays for five 9s diversity
4 subchannels require 500 MHz ADCs
8 subchannels require 250 MHz ADCs

\[ \Delta f = \frac{1}{T} \]

Subchannels allowed to overlap

Small number of equalizer taps
No error floor if using two receive arrays

Currently exploring OFDM within subchannels for indoor settings
Parting thoughts on the mm wave frontier

- Getting the most out of 60 GHz indoors
  - Near-LoS MIMO, rapid beam adaptation, handling blockage

- Picocellular backhaul
  - Quasi-deterministic links, highly directional mesh networks

- Mm wave to the mobile
  - Electrically large arrays, rapid adaptation and tracking, network-level coordination

- Wireless data centers
  - 3D beamforming and near-LoS MIMO

- Long-range wireless fiber
  - Distributed architectures for sidestepping geometric constraints

- Signal processing at scale: addressing the ADC bottleneck head on
  - Significant interdisciplinary effort over the next 2 decades
SPAWC focus

• Array of subarrays as a canonical MIMO architecture
  – RF beamforming within subarray
  – Digital, or mixed analog-digital, signal processing across subarrays

• The ADC bottleneck
  – ADC-constrained but DSP-centric design for multiGHz systems
  – Analog multitone as the new OFDM?

• But SP cannot be practiced in a silo
  – Must account for the physics of tiny wavelengths
  – Must account for hardware constraints associated with scaling
  – Must interact with directional networking protocols
Exploring further

Survey

MIMO techniques and channel modeling
Sheldon, Seo, Torkildson, Madhow, Rodwell, A 2.4 Gb/s millimeter-wave link using adaptive spatial multiplexing, APS-URSI 2010.
Zhang, Venkateswaran, Madhow, Channel modeling and MIMO capacity for outdoor millimeter wave links, WCNC 2010. (see also mmCom 2010)

Compressive adaptation
Ramasamy, Venkateswaran, Madhow, Compressive adaptation of large steerable arrays ITA 2012.
Ramasamy, Venkateswaran, Madhow, Compressive tracking with 1000-element arrays..., Allerton 2012.
Ramasamy, Venkateswaran, Madhow, Compressive estimation in AWGN, TSP, April 2014.

ADC Bottleneck
Zhang, Venkateswaran, Madhow, Analog multitone with interference suppression: relieving the ADC bottleneck for wideband 60 GHz systems, IEEE Globecom 2012.
Ponnuru, Seo, Madhow, Rodwell, Joint mismatch and channel compensation for high-speed OFDM receivers with time-interleaved ADCs, IEEE TCOM, August 2010.

Networking with highly directional links
Singh, Mudumbai, Madhow, Distributed coordination with deaf neighbors: efficient medium access for 60 GHz mesh networks, IEEE Infocom 2010.
Singh, Ziliotto, Madhow, Belding, Rodwell, Blockage and directivity in 60 GHz WPANs, IEEE JSAC, October 2009.