Noncoherent Communications with Large Antenna Arrays

Mainak Chowdhury
Joint work with: Alexandros Manolakos, Andrea Goldsmith, Felipe Gomez-Cuba and Elza Erkip

Stanford University

September 29, 2016
Wireless propagation
The promise of large antenna arrays

What is an antenna array?

Group of antennas receiving/transmitting signals simultaneously

- Signal becomes stronger and less uncertain
- Fading and noise vanish [Marzetta, 2010]
- Beams are directed and steerable
- Used for precise imaging and tracking

Note large gain along a particular direction
The promise of large antenna arrays

What is an antenna array?

Group of antennas receiving/transmitting signals simultaneously

Benefits

- Signal becomes stronger and less uncertain
  - Fading and noise vanish [Marzetta, 2010]
- Beams are directed and steerable
- Used for precise imaging and tracking

Note large gain along a particular direction
The promise with large antenna arrays

Better cellular network with heterogenous demand (objects and people)
Challenges with large arrays in communication systems

- Space constraints
- Number of radio frequency (RF) front ends is large
- Channel state information (CSI) needs to be acquired fast
- CSI needs to go through rate limited (fronthaul) links
Challenges with large arrays in communication systems

- Space constraints
- Number of radio frequency (RF) front ends is large
- Channel state information (CSI) needs to be acquired fast
- CSI needs to go through rate limited (fronthaul) links

Number of antennas doesn’t scale easily!
Our contributions

In large antenna arrays:

- Fading channel models need to be rethought
  ⇒ Coherence time and antenna correlation are different
Our contributions

In large antenna arrays:

- Fading channel models need to be rethought
  ⇒ Coherence time and antenna correlation are different

- Many benefits also possible with noncoherent schemes
  ⇒ Schemes with slowly changing statistics of the channel
Our contributions

In large antenna arrays:

- Fading channel models need to be rethought
  \(\Rightarrow\) Coherence time and antenna correlation are different

- Many benefits also possible with noncoherent schemes
  \(\Rightarrow\) Schemes with slowly changing statistics of the channel

- Simple transmission and detection works well
  \(\Rightarrow\) ON-OFF keying, one-shot detectors
Plan

Why large antenna arrays?

Channel models for MIMO large antenna arrays

Noncoherent schemes

Transmission and detection schemes

Conclusions
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
- Ray tracing models
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
- Ray tracing models
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
- Ray tracing models

Models for multi antenna systems

- Independent and identically distributed (IID) fading models
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
- Ray tracing models

Models for multi antenna systems

- Independent and identically distributed (IID) fading models
- Actual antennas not IID $\Rightarrow$ modeling antenna correlation
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
- Ray tracing models

Models for multi antenna systems

- Independent and identically distributed (IID) fading models
- Actual antennas not IID $\Rightarrow$ modeling antenna correlation
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
- Ray tracing models

Models for multi antenna systems

- Independent and identically distributed (IID) fading models
- Actual antennas not IID $\Rightarrow$ modeling antenna correlation

Models for a single time varying channel

- Characterize joint distribution of $h_m(t), h_m(t + \tau)$
Fading models

Channel from user equipment (UE) to $m^{th}$ antenna element is $h_m$

Models for single antenna systems

- Statistical fading models (distribution on $h_m$)
- Ray tracing models

Models for multi antenna systems

- Independent and identically distributed (IID) fading models
- Actual antennas not IID $\Rightarrow$ modeling antenna correlation

Models for a single time varying channel

- Characterize joint distribution of $h_m(t), h_m(t + \tau)$
- Coherence time $T_c$ implies $h_m(t) \approx h_m(t + \tau)$ for $\tau < T_c$
Main assumptions

- Number of antennas $N$ colocated
- UE $b$ leads to $L$ rays (or paths)
- Each ray has gain $G_{b,p} = \frac{1}{\sqrt{L}}$ and angle of arrival $\theta_{b,p}$

$$h_m = \frac{1}{\sqrt{L}} \sum_{p=1}^{L} e^{j\phi_{b,p}} + j2\pi m \frac{\sin(\theta_{b,p})}{\lambda}$$
Our analysis: starting point

Main assumptions

- Number of antennas $N$ colocated
- UE $b$ leads to $L$ rays (or paths)
- Each ray has gain $G_{b,p} = \frac{1}{\sqrt{L}}$ and angle of arrival $\theta_{b,p}$

\[
h_m = \frac{1}{\sqrt{L}} \sum_{p=1}^{L} e^{j \phi_{b,p}} + j2\pi m \frac{\sin(\theta_{b,p})}{\lambda}
\]

Implications

- For a finite $L$, there exists antenna correlations
- As $L \to \infty$, $h \sim \mathcal{CN}(0, \textbf{I})$
Our analysis: starting point

Main assumptions

- Number of antennas $N$ colocated
- UE $b$ leads to $L$ rays (or paths)
- Each ray has gain $G_{b,p} = \frac{1}{\sqrt{L}}$ and angle of arrival $\theta_{b,p}$

$$h_m = \frac{1}{\sqrt{L}} \sum_{p=1}^{L} e^{j \phi_{b,p}} + j 2\pi m \frac{\sin(\theta_{b,p})}{\lambda}$$

Implications

- For a finite $L$, there exists antenna correlations
- As $L \to \infty$, $\mathbf{h} \sim \mathcal{CN}(0, \mathbf{I}) \Rightarrow$ IID Rayleigh fading
Fact

$\phi_{b,p}[t]$ changes much faster than $G_{b,p}[t]$ or $\theta_{b,p}[t]$ ($10 \times$ to $1000 \times$)
Coherence time as a function of $N$

Fact

$\phi_{b,p}[t]$ changes much faster than $G_{b,p}[t]$ or $\theta_{b,p}[t]$ (10× to 1000×)

$\theta_{1,1}$ variation with time

$\phi_{1,1}$ variation with time
Coherence time as a function of $N$

Fact

$\phi_{b,p}[t]$ changes much faster than $G_{b,p}[t]$ or $\theta_{b,p}[t]$ ($10\times$ to $1000\times$)

| $h_1$ | and | $h_2$ | variation

Phase variation
Coherence time as a function of $N$

Fact

$\phi_{b,p}[t]$ changes much faster than $G_{b,p}[t]$ or $\theta_{b,p}[t]$ ($10 \times$ to $1000 \times$)

$|h_1|$ and $|h_2|$ variation

Phase variation

$h_m$ changes much faster than spatial parameters $\theta$ or $G$
Quantifying coherence time

Idea
When $N \to \infty$, becomes possible to track $\theta, G$

$$\text{DFT}(h) = \left\{ \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} h_n e^{-j2\pi kn/N} \right\}_{k=0}^{N-1}$$

Coherence time of $r = |\sum_n h_n \omega_n|^2$ can be much larger
Quantifying coherence time

Idea
When $N \to \infty$, becomes possible to track $\theta, G$

\[
\text{DFT}(h) = \left\{ \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} h_n e^{-j2\pi kn/N} \right\}_{k=0}^{N-1}
\]

Coherence time of $r = \frac{1}{N} \sum_i |h_n|^2$ can be much larger
Quantifying coherence time $T_c$

Idea
For any time series $r(t)$, look at $A_r(t, \tau) = E[r(t)r(t + \tau)]$
Autocorrelation function plots for different values of $N$

Note that coherence time increases for increasing $N$
Thus ...

Coherence times can be large for large antenna arrays
Thus ...

Coherence times can be large for large antenna arrays

Rest of the talk ...
How do we exploit that?
Plan

Why large antenna arrays?

Channel models for MIMO large antenna arrays

Noncoherent schemes

Transmission and detection schemes

Conclusions
Main assumptions

Fast changing channel unknown, slow fading channel known

Slow changing information depends on the number of multipath components $L$

- When $L$ large, statistics of the channel known
  - energy $\sigma_h^2$ for Rayleigh fading
  - line of sight (LOS) and non-LOS energy for Rician fading
- For small $L$, the spatial parameters ($\theta$ and $G$) are known

![Diagram showing user 1, user 2, base station, beam path, and user 2 connected](image-url)
Related (prior) work

Noncoherent capacity
Capacity when the channel realization is unknown at the transmitter and the receiver

- Capacity/bounds with knowledge only of channel statistics
- Characterization of noncoherent capacity achieving distributions
- Error-exponent optimal distributions for channels
Related (prior) work

**Noncoherent capacity**
Capacity when the channel realization is *unknown* at the transmitter and the receiver

- Capacity/bounds with knowledge only of channel statistics
- Characterization of noncoherent capacity achieving distributions
- Error-exponent optimal distributions for channels

We focus instead on the *one shot* communication problem
The one-shot problem

- Base station has an $N$ element linear antenna array
- User equipments (UEs) have 1 antenna each

Uplink (single antenna UEs to base station)

$$y = \sum_i h_i x_i + \nu$$

Downlink (base station to UEs)

$$y = h_i^T x + \nu$$

Objective

How do we choose transmissions, and how do we detect them?
The one-shot problem

- Base station has an $N$ element linear antenna array
- User equipments (UEs) have 1 antenna each

**Uplink** (single antenna UEs to base station)

$$y = \sum_{i} h_{i}x_{i} + \nu$$

**Downlink** (base station to UEs)

$$y = h_{i}^{T} x + \nu$$

**Short answer**

Depends heavily on channel characteristics and $N$
Channel models considered in the next few slides

• Narrowband IID Rayleigh fading models
  • Uplink
  • Downlink

• Wideband IID fading model
  • Uplink
  • Downlink

• Ray tracing models (uplink and downlink)
Narrowband IID Rayleigh fading uplink

Observation
In the uplink,

\[ y \sim f(\|y\|^2, x) \]

Discussions

- One-shot detection performance depends only on the received energy \( \|y\|^2 \)

- As \( N \) increases, \( \frac{\|y\|^2}{N} \) converges to \( \|x\|^2 + \sigma^2 \) almost surely

- At detector use threshold energy detection
Optimizing transmit constellations

Idea

The width of the histogram can be derived from the rate function $I_x(d)$ which satisfies

$$\text{Prob} \left( \left| \frac{\|y\|^2}{N} - \|x\|^2 - \sigma^2 \right| < d \right) = e^{-NI_x(d)}$$

Design criterion

Choose $\{x\}$ and thresholds to minimize probability of error or maximize $\min_x I_x(d)$ (detection threshold width for $x$)
Optimizing transmit constellations

Idea
The width of the histogram can be derived from the rate function $I_x(d)$ which satisfies

$$\text{Prob} \left( \left| \frac{\|y\|^2}{N} - \|x\|^2 - \sigma^2 \right| < d \right) = e^{-NI_x(d)}$$

Design criterion
Choose $\{x\}$ and thresholds to minimize probability of error or maximize $\min_x I_x$ (detection threshold width for $x$)

Capacity scaling result
For a large $N$, and a fixed number of users, the rate scales as $\Theta(\log(N))$ with a vanishing probability of error
Optimizing transmit constellations

Idea
The width of the histogram can be derived from the rate function $I_x(d)$ which satisfies

$$\text{Prob} \left( \left| \frac{\|y\|^2}{N} - \|x\|^2 - \sigma^2 \right| < d \right) = e^{-NI_x(d)}$$

Design criterion
Choose \{x\} and thresholds to minimize probability of error or maximize $\min_x I_x$ (detection threshold width for x)

Capacity scaling result
For a large $N$, and a fixed number of users, the rate scales as $\Theta(\log(N))$ with a vanishing probability of error

Knowing the channel $h$ does not improve capacity scaling
Observation

\[ y = h^T x + \nu \sim \mathcal{CN}(0, \sigma^2 + P), \text{ independent of } N \]

Result

Under an average transmit power constraint \( \| x \|^2 \leq P \), increasing \( N \) does not improve detection performance.
Narrowband IID Rayleigh fading downlink

**Observation**
\[ y = h^T x + \nu \sim \mathcal{C}N(0, \sigma^2 + P), \text{ independent of } N \]

**Result**
Under an average transmit power constraint \( \|x\|^2 \leq P \), increasing \( N \) does not improve detection performance

With IID fading, adding antennas does not help in the downlink
Wideband IID Rayleigh fading uplink

Setting

$B$ parallel narrowband channels (frequency bins) under total power constraint $P$

\[
\begin{array}{ccccccc}
\text{bin 0} & \text{bin 1} & \text{bin 3} & \ldots & \text{bin } B
\end{array}
\]

Results

- If $B$ is fixed and $N \to \infty$, rates $\Theta(\log_2(N))$ achievable
- If $B \to \infty$ and $N$ fixed, rates proportional to $\Theta(1)$ achievable
- Joint scaling: if $B = N^\epsilon$,
  - for $0 < \epsilon < 0.5$, rates close to $\Theta(B \log_2(N))$ achievable
  - for $\epsilon > 0.5$, rates cannot be better than $\Theta(N^{0.5})$
Wideband IID Rayleigh fading uplink

Setting

$B$ parallel narrowband channels (frequency bins) under total power constraint $P$

\[
\begin{array}{cccccc}
\text{bin 0} & \text{bin 1} & \text{bin 3} & \ldots & \text{bin } B
\end{array}
\]

Results

- If $B$ is fixed and $N \to \infty$, rates $\Theta(\log_2(N))$ achievable
- If $B \to \infty$ and $N$ fixed, rates proportional to $\Theta(1)$ achievable
- Joint scaling: if $B = N^\epsilon$,
  - for $0 < \epsilon < 0.5$, rates close to $\Theta(B \log_2(N))$ achievable
  - for $\epsilon > 0.5$, rates cannot be better than $\Theta(N^{0.5})$

Too much channel uncertainty for $\epsilon > 0.5$, bandwidth limited below $\epsilon < 0.5$
Beyond IID: ray tracing models

Ray tracing with finite number of multipath components

- For a large enough $N$, no noise or interuser interference
- For a finite $N$, may experience significant interference
Beyond IID: ray tracing models

Ray tracing with finite number of multipath components

- For a large enough $N$, no noise or interuser interference
- For a finite $N$, may experience significant interference

Reasons for performance degradation

- **Outage**: Two rays closer than $\Delta$ from each other cause destructive or constructive interference
- **Inter-path interference (IPI)**: Rays or paths not in outage also interfere
- **Additive receiver noise**

Objective

Study number of users $B$ supported with a vanishing outage and good detection error performance
Our result

Theorem
For both uplink and downlink, if the number of users $B = o(\sqrt{N})$ then one can choose $\Delta$ with a vanishing outage probability and with a non-outage diversity gain same as the one with a single user.

Diversity gain definition

$$\lim_{N \to \infty} \frac{-\log(\text{Probability of error})}{N}$$
Our result

Theorem
For both uplink and downlink, if the number of users $B = o(\sqrt{N})$ then one can choose $\Delta$ with a vanishing outage probability and with a non-outage diversity gain same as the one with a single user.

Diversity gain definition

$$\lim_{N \to \infty} - \frac{\log(\text{Probability of error})}{N}$$

With a slowly growing number of users, multiuser interference vanishes even if fast fading not known.
Plan

Why large antenna arrays?

Channel models for MIMO large antenna arrays

Noncoherent schemes

Transmission and detection schemes

Conclusions
Noncoherent transceivers

Why noncoherent?

- Do not have to track instantaneous phase/fast fading
- Simpler design
- Energy efficient
- Lower channel feedback rate requirements
Noncoherent transceivers

Why noncoherent?

- Do not have to track instantaneous phase/fast fading
- Simpler design
- Energy efficient
- Lower channel feedback rate requirements

Examples of noncoherent architectures
Noncoherent transceivers

Why noncoherent?

- Do not have to track instantaneous phase/fast fading
- Simpler design
- Energy efficient
- Lower channel feedback rate requirements

Examples of noncoherent architectures

- Energy detectors
Noncoherent transceivers

Why noncoherent?

- Do not have to track instantaneous phase/fast fading
- Simpler design
- Energy efficient
- Lower channel feedback rate requirements

Examples of noncoherent architectures

- Energy detectors
- Fourier transform
Noncoherent transceivers

Why noncoherent?

- Do not have to track instantaneous phase/fast fading
- Simpler design
- Energy efficient
- Lower channel feedback rate requirements

Examples of noncoherent architectures

- Energy detectors
- Fourier transform
- Other unitary precoders/receiver shaping transforms not dependent on fast fading
Benefits from noncoherent transceiver

• Bit rate requirements
  • Fronthaul limits the number of antennas that can be supported
    \[ N \leq \frac{9 \text{ Gbps}}{2 \times \text{bits per sample} \times \text{bandwidth in hertz} \times \text{no. of sectors}} \approx 5 \]
  • Noncoherent architectures are independent of the fronthaul rate limits

• Energy consumption
  • RF front ends for each antenna infeasible
  • Energy consumption for analog-to-digital converters high
  • Efficient architectures for energy detection or fixed precoders
Plan

Why large antenna arrays?

Channel models for MIMO large antenna arrays

Noncoherent schemes

Transmission and detection schemes

Conclusions
Noncoherent architectures promise benefits for large antenna arrays. We discussed how:

- channel fading is different
- noncoherent schemes are optimal according to various metrics
- noncoherent transceiver architectures are practical
Conclusions

Noncoherent architectures promise benefits for large antenna arrays. We discussed how:

- channel fading is different
- noncoherent schemes are optimal according to various metrics
- noncoherent transceiver architectures are practical

TODO...
Test theory on real implementations
The end

Thank you for your attention!