Distributed Nullforming for Distributed MIMO Communications

Soura Dasgupta

The University of Iowa
Background

- MIMO Communications
- Promise Much

- Centralized Antennae
  - 802.11n, 802.11ac, LTE, WiMAX, IMT-Advanced

- Limited in wireless networks by:
  - Form factor
  - Antenna Size
  - Number of Antennae
Distributed MIMO

- D-MIMO attractive alternative
- Transmitters form a virtual antenna
  - Cover and El Gamal, Gastpar and Vetterli
- Carry Separate Oscillators that drift
- Uncertain Geometries
- Extolled by Theoreticians
- Dismissed by practitioners
Major Tools

• Distributed Beamforming
  • Constructive interference at a target
  • $N^2$ gain

• Distributed Nullforming
  • Destructive interference at a target
  • interference avoidance for increased spatial spectrum reuse
  • cognitive radio
  • physical-layer security

• Both require tight synchronization
Concept of Distributed Beam/Nullforming

- Many radios with single-element antennas
- Together act like large antenna array
- Focus transmission in direction of receiver
- Constructive/Destructive interference
- Spatial Multiplexing
Beamforming v. non-coherent cooperation

- SNR increases as $N^2$
  - e.g. 5 element array gives 25x higher SNR than individual transmitter
  - contrast with “amplify & forward relaying”, or “cooperative diversity”
  - Synchronization crucial
Synchronization

- Frequency lock between cooperating nodes

- Phase lock needed at the receiver
Stringent synchronization requirements

Some numbers to illustrate

- Carrier Frequency 2.4 GHz
- 10 nodes, beamforming: Received SNR 20 dB
- Typical clock drift stdev 2.5 ns/sec
- At $t=50$ millisecond clock offset: 125 pico seconds
- Expected received SNR at $t=50$ ms: SNR 11 dB
- Incoherence $\geq 10$ dB
Unpredictable Clock Dynamics: An Example

Clocks are synchronized here (same frequency and phase)
Considerable recent progress on beamforming

- A menu of synchronization techniques have been developed
  - featuring different sets of tradeoffs between complexity, overheads and performance
Work on Beamforming

- Receiver aided feedback
- Tu and Pottie 2002
  - Separate feedback to each node
  - Not scalable
- Coordinated multipoint (CoMP)” for 4G-LTE cellular systems
  - Cooperating Base stations
  - Complex High Speed Backhaul
  - Ubiquitous GPS
- Scalable Feedback
  - One-bit algorithm (Mudumbai et. al.)
  - DARPA project with BBN-Raytheon-Beamforming at 1 km
1-bit algorithm classical version

• Assumes frequency synchronization
• Used for phase synchronization
• Each node perturbs its phase randomly
• Receiver compares new received power to old
• Sends 1-bit information
  • Power increased or decreased?
• If increased nodes retain perturbation
• If decreased discard perturbation
• Guaranteed convergence
1-bit feedback control algorithm

If GOOD keep.
Repeat.
If BAD discard.
And try again.

Really neat: no calibration, channel-estimation
Nullforming Much More Challenging

• Much more sensitive to phase errors

• Beamforming
  • Align phases

• Nullforming
  • Phases and magnitudes must be carefully chosen
  • More intricate than mere phase alignment
Past Work on Null Forming

- Brown et. Al. (CISS 12, SSP 12)

- Feedback based

- Every node knows every other node’s complex channel gain

- Scalability dented
Distributed nullforming

State-of-the-art

Each node needs to know \( h_1, h_2, h_3, h_4 \).

1-bit feedback algorithm does not work.

Our algorithm

Node \( i \) needs to know \( h_i \) only.
Goals of this talk

- **MISO nullforming with phase only adjustments**
  - Transmitting at full power
  - Protecting a cooperating transmitter
  - Non-convex problem

- **Joint beam and nullforming with phase and gain adjustments**
  - Null at multiple locations
  - Beam at one location
  - Convex problem

- **Aggregate Feedback**
- **Distributed**
New Algorithm

- Receiver broadcasts received total baseband signal to all the nodes
  - Aggregate feedback signal

- Adjusts broadcast in response

- Each node only knows its complex channel gain

- More distributed and scalable

- Gradient based
Framework

• First assume all nodes frequency synchronized

• Each node knows its complex channel gains and equalizes the phase but not the magnitude
  • Assumes that the actual compensated channel is $r_D^i$

• Baseband signal sent by $i$-th node: $e^{j\theta_D^i[k]}$

• Signal fed back: $s[k]=\sum_{i=1}^{N} r_D^i[k] e^{j(\theta_D^i[k]+\phi_D^i[k])} + \text{noise}$

• $\phi_D^i[k]$ small uncompensated channel phase error
Algorithm

• For conceptual ease ignore noise
• Call $\theta$ the vector of phases
• Received power: $J(\theta[k]) = |s[k]|^2 = |\sum_{i=1}^{N} r_i[k] e^{j(\theta_i[k] + \phi_i[k])}|^2$
• Gradient descent: $\theta[k+1] = \theta[k] - \mu \frac{\partial J(\theta)}{\partial \theta} \mid \theta = \theta[k]$
• Cannot implement true gradient if $\phi_i[k]$ $r_i[k]$ unknown
• Implement:
  $\theta_i[k+1] = \theta_i[k] - \mu 1 r_i \{ \cos(\theta_i[k]) \Im[s[k]] - \sin(\theta_i[k]) \Re[s[k]] \}$
Observations

• $\theta_{i[k+1]} = \theta_{i[k]} - \mu_1 r_{i[k]} \{ \cos(\theta_{i[k]}) \Im[s[k]] - \sin(\theta_{i[k]}) \Re[s[k]] \}$

• True gradient if $\varphi_{i} = 0$ and $r_{i[k]} = r_{i}$

• Each node needs its current phase, channel gain and $s[k]$ to implement
Similarity to Kuramoto

\[ \theta_{\downarrow i}[k+1] = \theta_{\downarrow i}[k] - \mu_{\downarrow 1} r_{\downarrow i} \{ \cos(\theta_{\downarrow i}[k]) \text{Im}[s[k]] - \sin(\theta_{\downarrow i}[k]) \text{Re}[s[k]] \} \]

\[ = \theta_{\downarrow i}[k] + \sum_{l=1}^{N} r_{\downarrow l} \sin(\theta_{\downarrow i}[k] - \theta_{\downarrow l}[k] - \phi_{\downarrow l}[k]) \]
Stability Analysis

- Will show practical uniform convergence with zero phase offsets $\varphi \downarrow i$ and $r \downarrow i [k] \equiv r \downarrow i$
- Convergence to a critical point
- Only locally stable critical points are global minima
- Rest unstable
  - Unattainable
  - Cannot be maintained
- Uniform convergence guarantees robustness
Convergence Behavior with $\varphi \downarrow i = 0$

- For small enough $\mu > 0$:

$$J(\theta[k+1]) \leq J(\theta[k])$$

$$\lim_{k \to \infty} \frac{\partial J(\theta)}{\partial \theta} \bigg|_{\theta = \theta[k]} = 0$$

- Convergence to a critical point

- Need to characterize critical points

- Examine Hessian
  - If it has a negative eigenvalue then unstable
Critical Points

- \( r \downarrow i \{ \cos(\theta \downarrow i [k]) \text{Im}[s[k]] - \sin(\theta \downarrow i [k]) \text{Re}[s[k]] \} = 0 \)

- Two possibilities:
  - \( s[k] = 0 \quad \Longleftrightarrow \quad \text{Null manifold} \)
  - \( \tan(\theta \downarrow i [k]) = \tan(\theta \downarrow j [k]) \)

- Manifolds:
  \( r \downarrow i \{ \cos(\theta \downarrow i [k]) \text{Im}[s[k]] - \sin(\theta \downarrow i [k]) \text{Re}[s[k]] \} = \sum_{j \neq i} r \downarrow j \cos(\theta \downarrow i [k] - \theta \downarrow j [k]) \)
The Hessian

\[ [H(\theta)]_{ij} = \frac{\partial^2 J(\theta)}{\partial \theta_i \partial \theta_j} = \begin{cases} r_{ij} \cos(\theta_i - \theta_j) & i \neq j \\ -\sum_{l \neq i} r_{il} r_{lj} \cos(\theta_i - \theta_l) & i = j \end{cases} \]

- A critical point is locally unstable if \( H(\theta) \) has a negative eigenvalue

- Two settings to consider
  - A null is possible: \( J^* = 0 \)
  - No null possible: For some \( i, r_{ij} > \sum_{j \neq i} r_{lj} \) \( \Rightarrow \) Global minimum:
    \[ J^* = (r_{ij} - \sum_{j \neq i} r_{lj})^2 \]
Local Behavior

- At any critical point that is not a global minimum the Hessian has a negative eigenvalue
  - Unstable

- At a global minimum
  - Hessian positive semidefinite
  - Not positive definite
  - Global minima are manifolds not isolated points
  - Hessian is singular

\[
[H(\theta)]_{ij} = \partial^2 J(\theta)/\partial \theta_i \partial \theta_j = \begin{cases} r_i r_j \cos(\theta_i - \theta_j) & i \neq j \\ - \sum_{l \neq i} r_i r_l \cos(\theta_i - \theta_l) & i = j \end{cases}
\]
Local Stability of Global Minima

- Let $m$ be the smallest cost at a critical point that is not a global minimum

  \[ J(\theta[0]) < m \]
  \[ J(\theta[k+1]) \leq J(\theta[k]) \]
  \[ \lim_{k \to \infty} \frac{\partial J(\theta)}{\partial \theta} \big|_{\theta = \theta[k]} = 0 \]

  - Then
    \[ \lim_{k \to \infty} J(\theta[k]) = J^{*} \]
Intriguing results for constant phase offsets and equal $r_i$

- The elements of the gradient equalize

- May not go to zero!

- Consensus of sorts

- Includes the possibility of attaining nulls
  - Always seems to happen when phase offsets are less than $\pi/2$
Simulation 1

- $\phi_i$ constant, uniformly distributed in $[0, \pi/2]$ 
- 10 nodes 
- Unit channel gains 
- SNR: Per node SNR
Simulation 2

- $\phi_{i}$ constant, uniformly distributed in $[0, \pi/2]$
- $r_{i}[k]$ constant uniformly distributed between $[1, 2]$
- 10 nodes
- Unit channel gains
Simulation 3

- $\varphi \downarrow i [0]$ uniformly distributed in $[0, \pi/4]$
- 10 nodes
- Unit channel gains
- Channel changes every $C$ iterations
  - $C$: coherence
  - Change in phase uniformly distributed in $[-1.5, 1.5]$ degrees
  - Change in gain by a factor uniformly distributed in $[.99, 1.01]$
Theoretical limit

• \( E\{|\sum_{i=1}^{N} r_{i} e^{j(\theta_{i} + \varphi_{i})}|^{2}\} \)

• \( r_{i} \sim U[0.99, 1.01] \) and \( \varphi_{i} \sim U[-1.5, 0] \)

• Subject to

• \(|\sum_{i=1}^{N} e^{j\theta_{i}}|^{2} = 0\)

• Optimistic performance limit
  • Ignores initial offsets
• Constant phase offset uniformly distributed between $[0,2\pi]$
• Steady State Gradient:
  
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
<tr>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
<td>2.1402</td>
</tr>
</tbody>
</table>
Drifting oscillators: Brownian motion

Power at target (dB)

Mean Brownian motion phase drift between iterations (degrees)

SNR=30dB
SNR=60dB
SNR=100dB
Convergence Speed: Scalability

- **Blue line**: \(N=10, mu=0.2\)
- **Red line**: \(N=30, mu=0.06\)
- **Green line**: \(N=50, mu=0.04\)

Graph showing the convergence speed with different values of \(N\) and \(mu\).
Effect on a Beam: SDMA?

![Graph showing the effect on beam power with iterations.]
Goals of this talk

- MISO null forming with phase only adjustments
  - Transmitting at full power
  - Protecting a cooperating transmitter
  - Non-convex problem

- Joint beam and null forming with phase and gain adjustments
  - Null at multiple locations
  - Beam at one location
  - Convex problem
Technical Difficulties

• $J_i$ cost function at $i$-th receiver

• Naïve approach: Minimize: $\sum_{i=2}^{M} J_i - c_J 1$
  • Nonconvex
  • Likely to have weird behavior
  • Particularly if we allow gain adjustment

• Adjustment of gains while doing just nullforming has problems
  • Could drive gains to zero

• Key observation: Specify desired power at node 1 rather than maximizing it
Framework

• At startup:
  • $l$-th receiver sends to the $i$-th transmitter the channel gain:
    $$h_{il} \downarrow \uparrow \ast = \alpha_{il} e^{j \varphi_{il}}$$
  
• During operation in the $n$-th time slot:
  • $i$-th transmitter broadcasts the baseband signal $x_{il}[n]$
  • $l$-th receiver broadcasts its total received signal:
    $$s_{il}[n] = \sum_{i=1}^{M} h_{il} \uparrow \downarrow x_{il}[n] + w_{il}[n]$$

• Feedback needs TDM
• Adjust $x_{il}$ to minimize:
  $$\sum_{l=1}^{M \downarrow 1} \|s_{il}\|^{2} + \sum_{l=M \downarrow 1 + 1}^{M} \|s_{il} - b_{il}\|^{2}$$
TDM for aggregate feedback

- All N transmitters broadcast to all M receivers
- Receiver 1 broadcasts to all N transmitters
- Receiver M broadcasts to all N transmitters

(time-slot k)
Matrix Framework

\[ s_{\downarrow l}[n] = \sum_{i=1}^{\uparrow M} h_{\downarrow i l} \ast x_{\downarrow i}[n] + w_{\downarrow l}[n] \]

- \( s[n] = H^{\uparrow H} x[n] + w[n] \) ← white zero mean Gaussian
- Find \( x \) to minimize

- \( J = \mathbb{E}_{\downarrow w} [\|H^{\uparrow H} x + w[n] - b\|_{\uparrow 2}] \)

- \( H: N \times M, N > M \): Optimum solutions lie on an \((N-M)\)-dimensional subspace: \( X \)

- \( \text{arg min} \ J(x) = \text{arg min} \ \|H^{\uparrow H} x - b\|_{\uparrow 2} \)
Issues

- Decentralized Iterative Solution
  - $i$-th node uses aggregate feedback and knowledge of its channel
  - Pseudo-inversion is centralized

- How does $x[0]$ affect steady state solution

- How to achieve minimum power solution $x_{\downarrow \text{min}}$?

- Effect of noise
  - Is there drift?

- $\lim_{N \to \infty} ?$
  - Speed
  - $x_{\downarrow \text{min}}$
Gradient descent

- \( \text{arg min } J(x) = \text{arg min } \| H^T H x - b \| \uparrow^2 \)

- Gradient descent:
  \[
  x[n+1] = x[n] - \mu \frac{\partial J(x)}{\partial x} \bigg| \downarrow \theta = x[n]
  \]

  \[
  x[n+1] = x[n] - \mu H \omega (s[n] - b) + H^T H x[n] - b + w[n]
  \]
Zero noise convergence

• Assured if $\mu \downarrow \max (H^H H) < 1$

• $x[0] \rightarrow$ projection of $x[0]$ on $X$

• Exponential

• $x[0] = 0 \rightarrow x \downarrow \text{min}$

• $\lim_{N \to \infty} \|x \downarrow \text{min} (N)\|_2 = 0$ if channel coefficients are iid CN(0, 1)
Convergence rate: Scalability

- Suppose the \( i \)-th column of \( H \) is \( h\downarrow i \) and \( H \) has full column rank.

- As \( N \to \infty \) convergence rate is bounded from below if the condition number of \( H^\top H \) is bounded.

- As \( N \to \infty \) convergence rate is bounded from below if there is a \( K \) and \( \alpha \downarrow i > 0 \) such that for all \( n \),

\[
\alpha \downarrow 1 \ I \leq \sum_{i=n^\top n + K} h\downarrow i \ h\downarrow i \ ^\top H \leq \alpha \downarrow 2 \ I
\]

- Successive channel vectors are persistently spanning.

- As \( N \to \infty \) convergence rate is bounded from below if the channel coefficients are in \( \text{CN}(0, 1) \).
Noise performance

- $H^*H x^* = b$,  \quad $\Delta[n] = x[n] - x^*$

- Error model:
  $\Delta[n+1] = (I - \mu HH^*H)\Delta[n] - \mu Hw[n]$

- $T\Delta = (z\downarrow_1, z\downarrow_2)$: $z\downarrow_1$ in the orthogonal complement of $X$
  and $z\downarrow_2$ on $X$

- $(z\downarrow_1[n+1], z\downarrow_2[n+1]) = ((I - \mu A)z\downarrow_1[n], z\downarrow_2[n]) + \mu (v\downarrow_1[n], v\downarrow_2[n])$: $v\downarrow_2[n] \neq 0 \implies$ Brownian motion along the minimizing subspace
Noise performance

• Error model:

\[ \Delta[n+1] = (I - \mu HH^{\top} H) \Delta[n] - \mu Hw[n] \]

• \( T \Delta = (z \downarrow 1 \mid z \downarrow 2) \): \( z \downarrow 1 \) in the orthogonal complement of \( X \) and \( z \downarrow 2 \) on \( X \)

• \( (z \downarrow 1 [n+1] \mid z \downarrow 2 [n+1]) = ((I - \mu A)z \downarrow 1 [n] \mid z \downarrow 2 [n]) + \mu (v \downarrow 1 [n] \mid v \downarrow 2 [n]) \): \( v \downarrow 2 [n] \neq 0 \) → Brownian motion along the minimizing subspace

• \( v \downarrow 2 [n] = 0 \)
Asymptotic noise performance

• Suppose $x[0] = 0$

• Channel coefficients are iid CN(0,1)

• $\lim_{N \to \infty} \lim_{n \to \infty} \|x(n)\|_2 = 0$
Simulation 4

- N=20
- M=5
- 2 beam targets: 1
- SNR=-40dB
Constant $\mu$

**Iterations used**($k$) for convergence of $J$ to -60dB

$\log_2(\text{Iterations})$ vs. $\log_2(\text{Index})$
Phase drift, SNR=-40dB T=50ms, 0 to 5 degrees between iterations

Graph: Objective function in dB vs. phase drift in deg/sqrt(sec)
Conclusion

• For Phase only adaptation
  • Always converges to a stationary point
  • All local minima global minima
  • Practically guaranteed convergence to a global minimum
  • Scalable
  • Fast convergence
  • Insensitive to channel estimation errors

• For joint null and beamforming
  • Scalable
  • Zero initial power $\Rightarrow$ Power Efficient Solution$\Rightarrow$0 as $N\Rightarrow0$
  • No drift along minimizing subspace