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

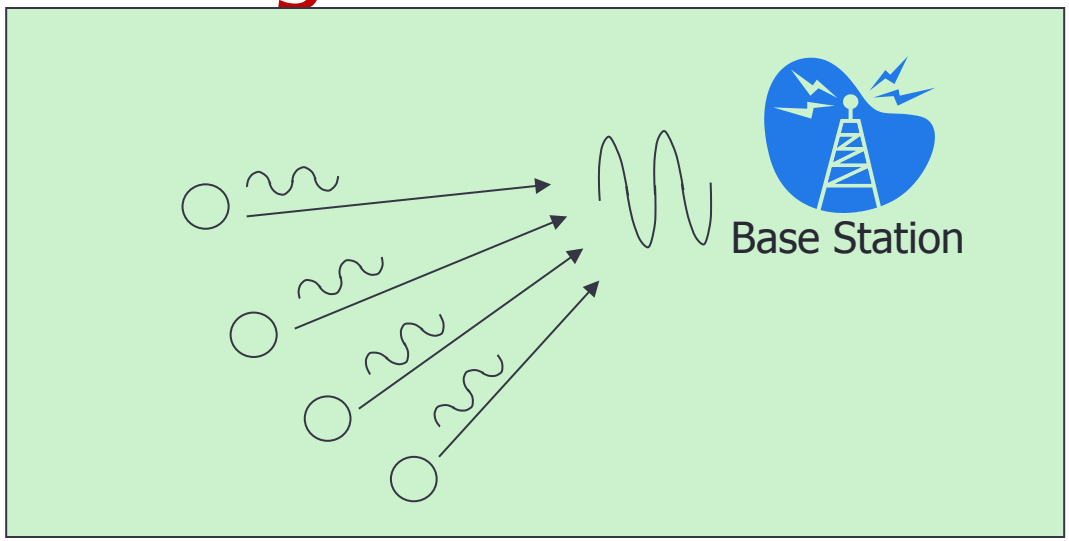
- D-MIMO attractive alternative
- Transmitters form a virtual antenna
  - Cover and El Gamal, Gastpar and Vetterli
- Carry Separate Oscillators that drift
- Uncertain Geometries
- Extolled by Theoreticians
- Dismissed by practitioners

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

THE UNIVERSITY OF IOWA

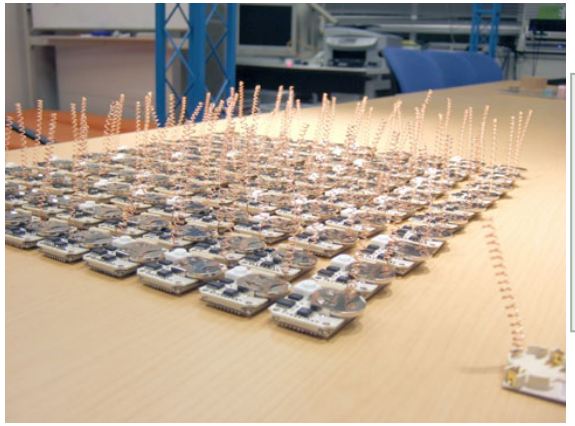
Concept of Distributed Beam/ Nullforming



- 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



INDISTINGUISHABLE IN FUNCTIONALITY

THE UNIVERSITY OF IOWA

Synchronization

- Frequency lock between cooperating nodes

- Phase lock needed at the receiver

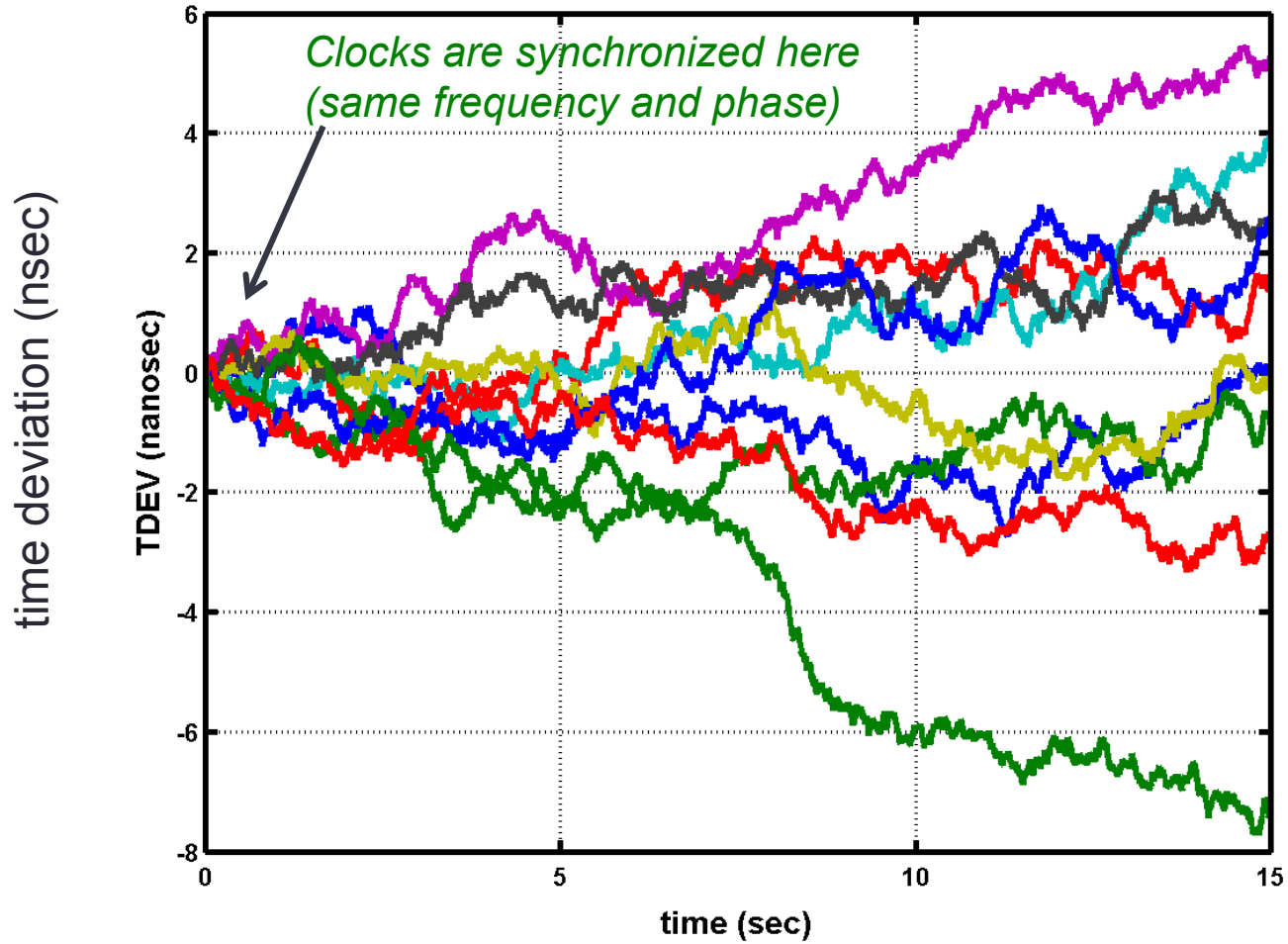
Stringent synchronization requirements

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 millisec clock offset: 125 pico seconds
- Expected received SNR at t=50 ms: SNR 11dB
- Incoherence➔10dB

THE UNIVERSITY OF IOWA

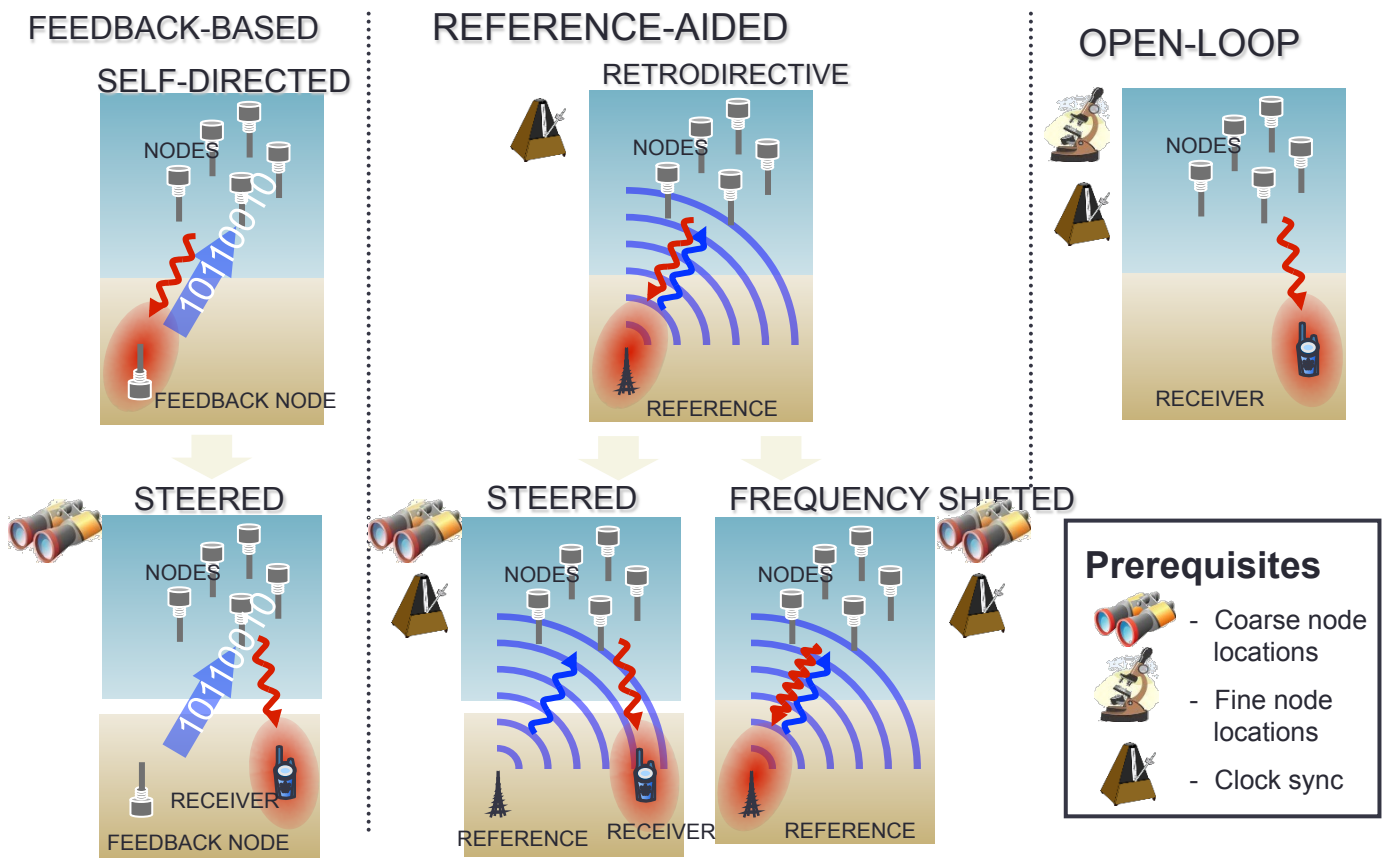
Unpredictable Clock Dynamics: An Example



THE UNIVERSITY OF IOWA

Considerable recent progress on beamforming

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



THE UNIVERSITY OF IOWA

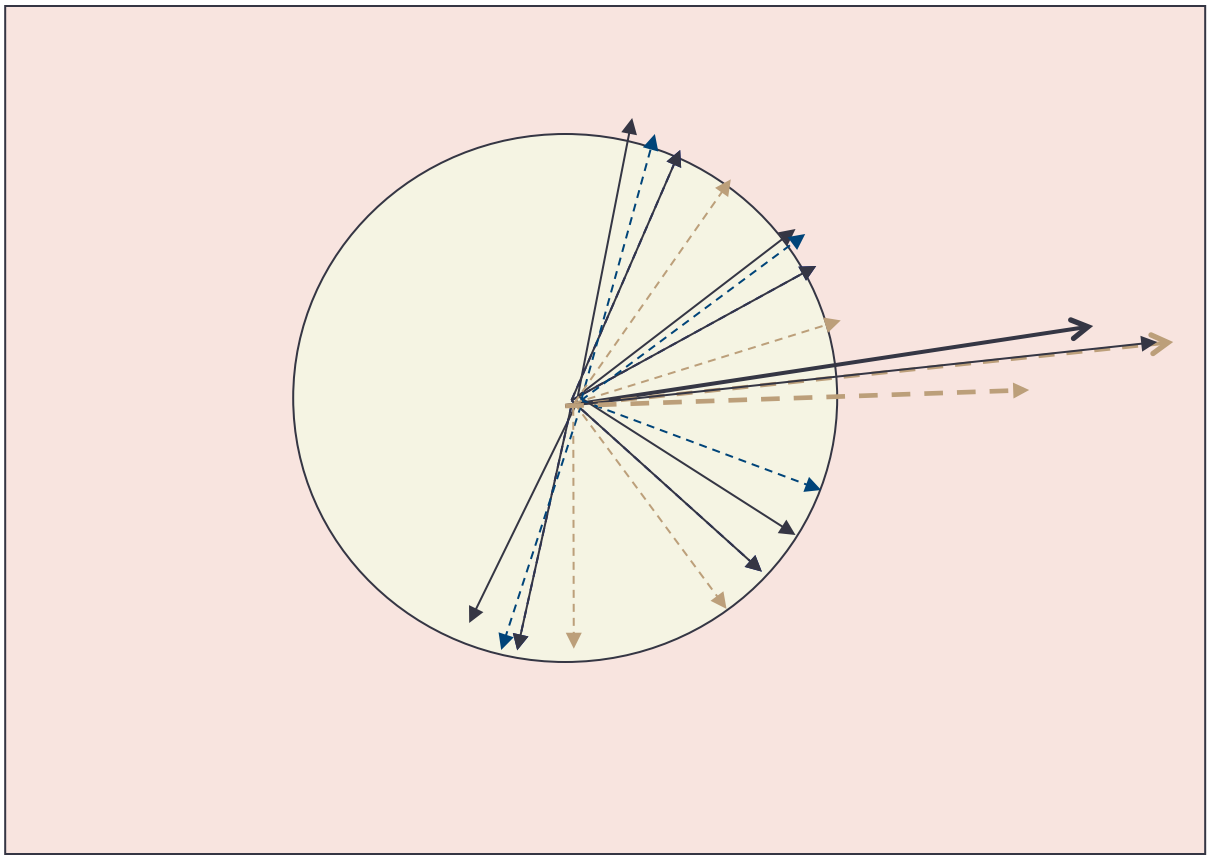
Work on Beamforming

- 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

THE UNIVERSITY OF IOWA

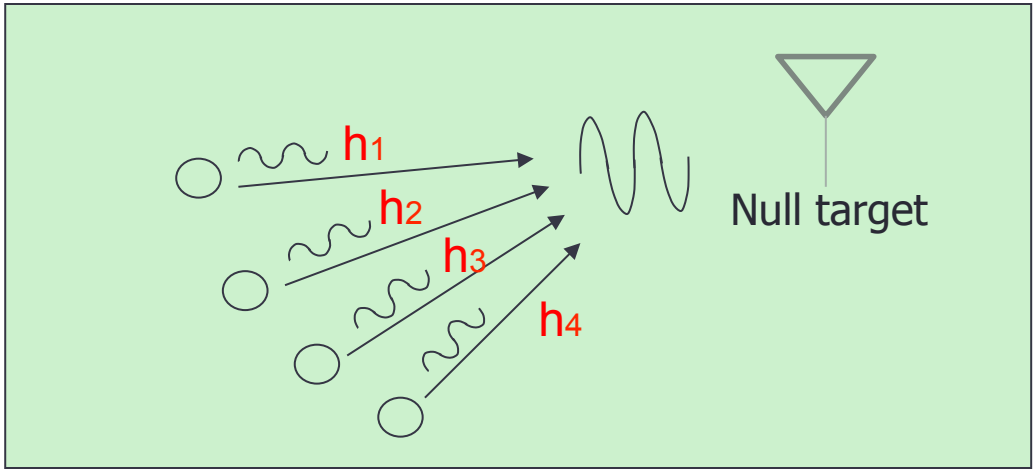
Nullforming Much More Challenging

- 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

1-bit feedback algorithm does not work.

Each node needs to know h_1, h_2, h_3, h_4.

Our algorithm

Node *i* needs to know h_i only.

THE UNIVERSITY OF IOWA

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_i
- Baseband signal sent by i-th node: $e^{j\theta_i}[k]$
- Signal fed back: $s[k]=\sum_{i=1}^{N} r_i[k] e^{j(\theta_i[k]+\varphi_i[k])}$ + noise
- $\varphi_i[k]$ small uncompensated channel phase error

THE UNIVERSITY OF IOWA

Algorithm

- For conceptual ease ignore noise
- Call θ the vector of phases
- Received power: $J(\theta[k])=|s[k]|^2=|\sum_{i=1}^{N} r_i[k] e^{j(\theta_i[k]+\varphi_i[k])}|^2$
- Gradient descent: $\theta[k+1]=\theta[k]-\mu \partial J(\theta)/\partial\theta \,|_{\theta=\theta[k]}$
- Cannot implement true gradient if $\varphi_i[k]$ $r_i[k]$ 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

- $\theta_i[k+1]=\theta_i[k]-\mu_1 r_i\{\cos(\theta_i[k])\, Im[s[k]]-\sin(\theta_i[k])\, Re[s[k]]\}$

- True gradient if $\varphi_i=0$ and $r_i[k]\equiv r_i$

- Each node needs its current phase, channel gain and s[k] to implement

THE UNIVERSITY OF IOWA

Similarity to Kuramoto

- $\theta_i[k+1]=\theta_i[k]-\mu_1 r_i\{\cos(\theta_i[k])\,\mathrm{Im}[s[k]]-\sin(\theta_i[k])\,\mathrm{Re}[s[k]]\}$

$=\theta_i[k]+\boxed{W}\, r_i \sum_{l=1}^{N} r_l \sin(\theta_i[k]-\theta_l[k]-\varphi_l[k])$

THE UNIVERSITY OF IOWA

Stability Analysis

- Will show practical uniform convergence with zero phase offsets φ_i and $r_i[k] \equiv r_i$
- 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 $\varphi_i = 0$

- For small enough $\mu > 0$:

$$J(\theta[k+1]) \leq J(\theta[k])$$
$$\lim_{k \to \infty} \partial J(\theta)/\partial\theta \,|\, \theta = \theta[k] = 0$$

- Convergence to a critical point
- Need to characterize critical points
- Examine Hessian
  - If it has a negative eigenvalue then unstable

Critical Points

- $r_i \{\cos(\theta_i[k])\, Im[s[k]] - \sin(\theta_i[k])\, Re[s[k]]\} = 0$

- Two possibilities:
  - s[k]=0 ←→ Null manifold
  - $\tan(\theta_i[k]) = \tan(\theta_j[k])$
- Manifolds:

$r_i \{\cos(\theta_i[k])\, Im[s[k]] - \sin(\theta_i[k])\, Re[s[k]]\} = \sum_{j \neq i} r_i r_j \cos(\theta_i[k] - \theta_j[k])$

The Hessian

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

- 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, $r_i > \sum_{j \neq i} r_j$ ➔ Global minimum:

$J^* = (r_i - \sum_{j \neq i} r_j)^2$

THE UNIVERSITY OF IOWA

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

- Choose $J(\theta[0])<m$

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

$\lim_{k\to\infty} \partial J(\theta)/\partial\theta \mid_{\theta=\theta[k]} =0$

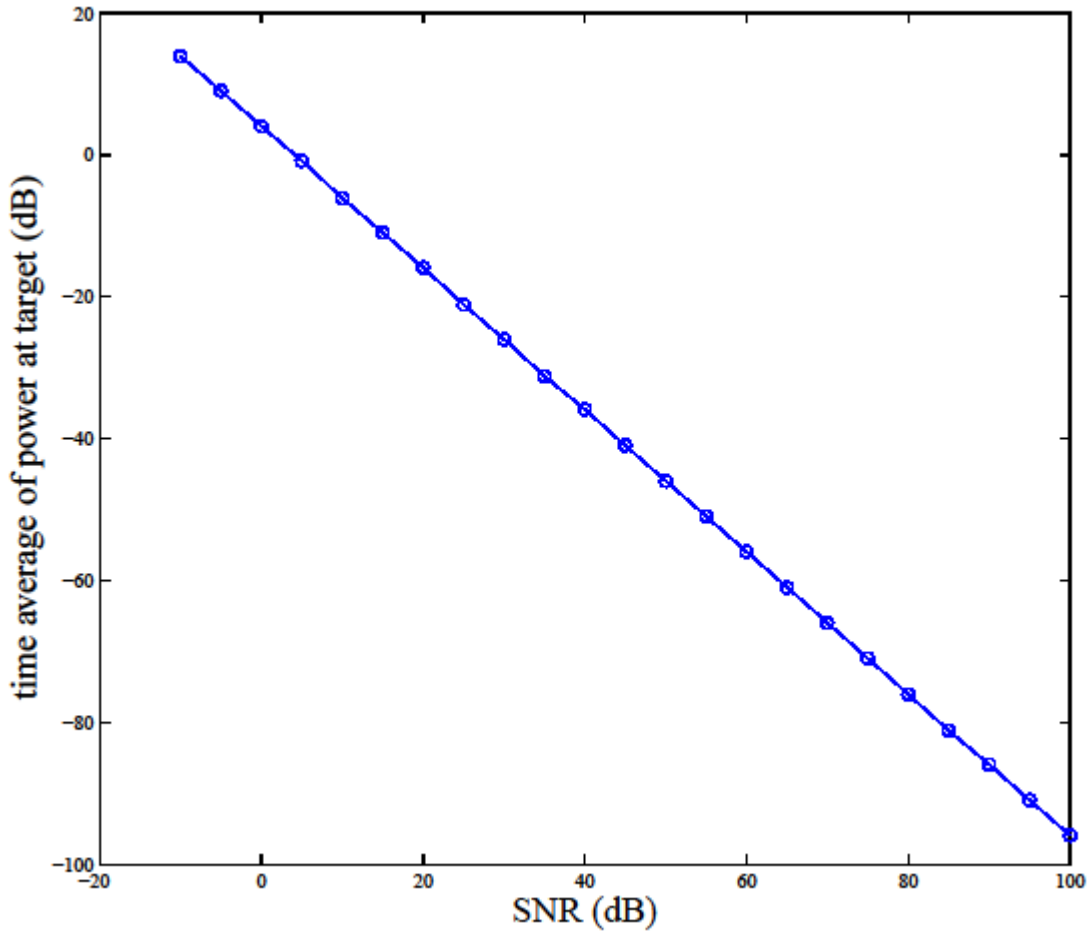
- Then
- $\lim_{k\to\infty} J(\theta[k]) = J^*$

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

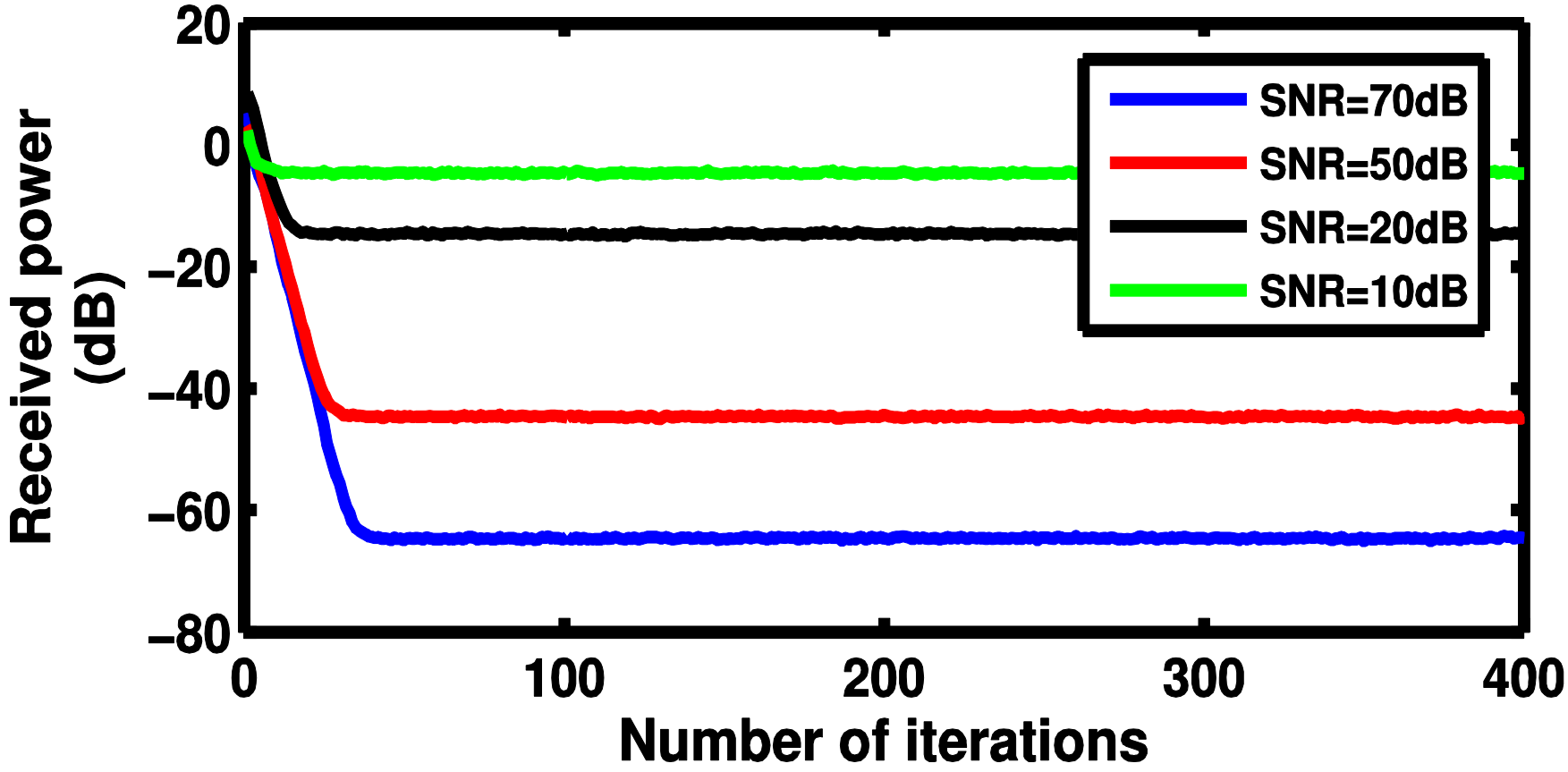
- φ_{i} constant, uniformly distributed in [0,π/2]
- 10 nodes
- Unit channel gains
- SNR: Per node SNR



THE UNIVERSITY OF IOWA

Simulation 2

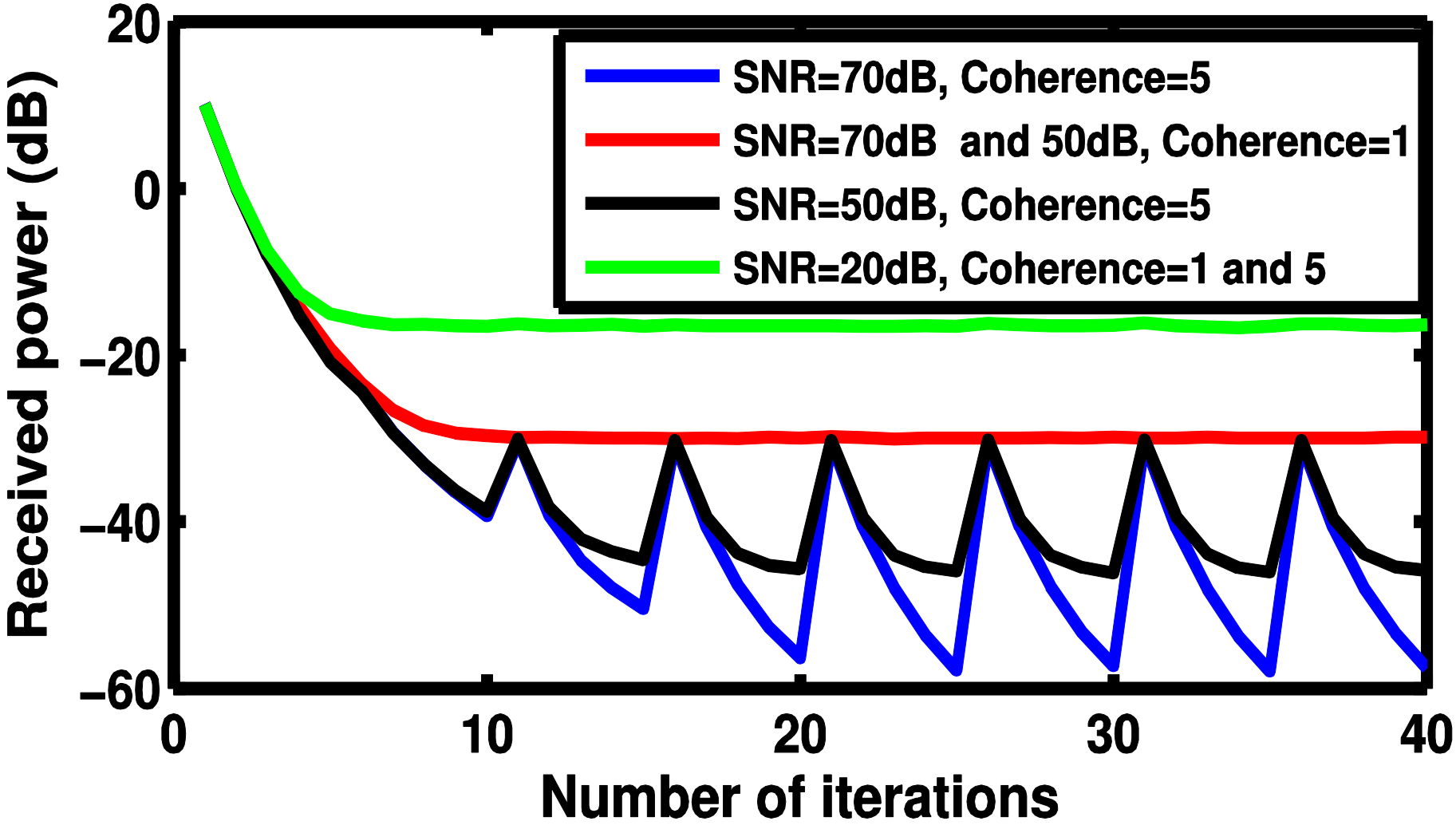
- φ_{i} constant, uniformly distributed in [0,π/2]
- $r_{i}[k]$ constant uniformly distributed between [1, 2]
- 10 nodes
- Unit channel gains



THE UNIVERSITY OF IOWA

Simulation 3

- $\varphi_{i}[0]$ uniformly distributed in [0,π/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]

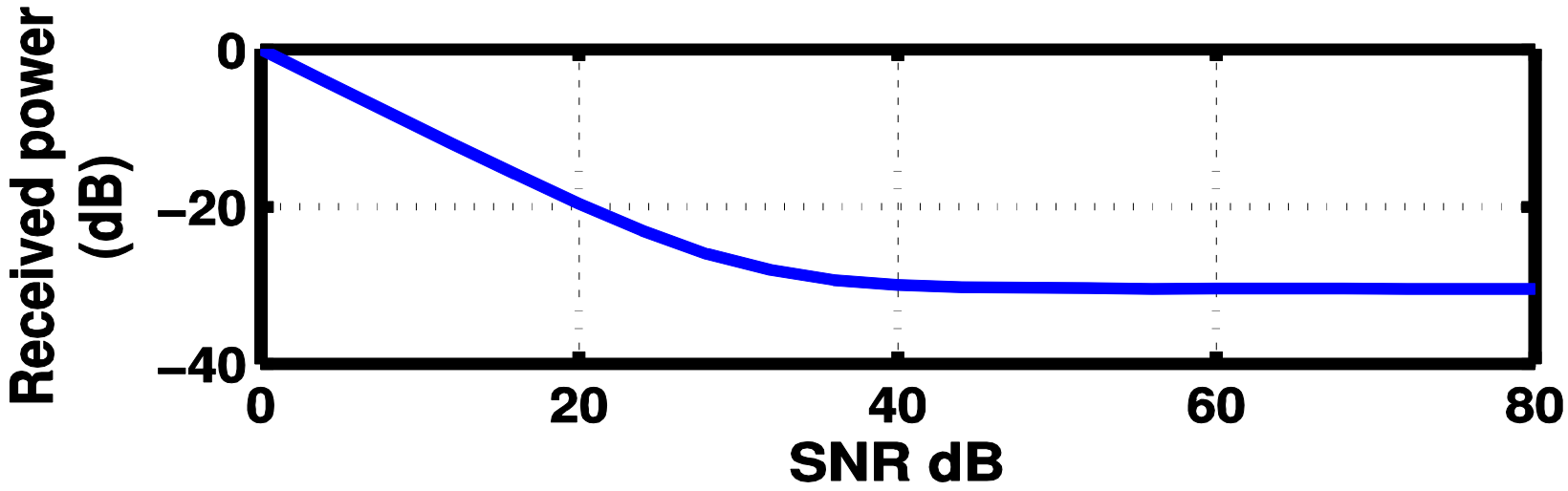


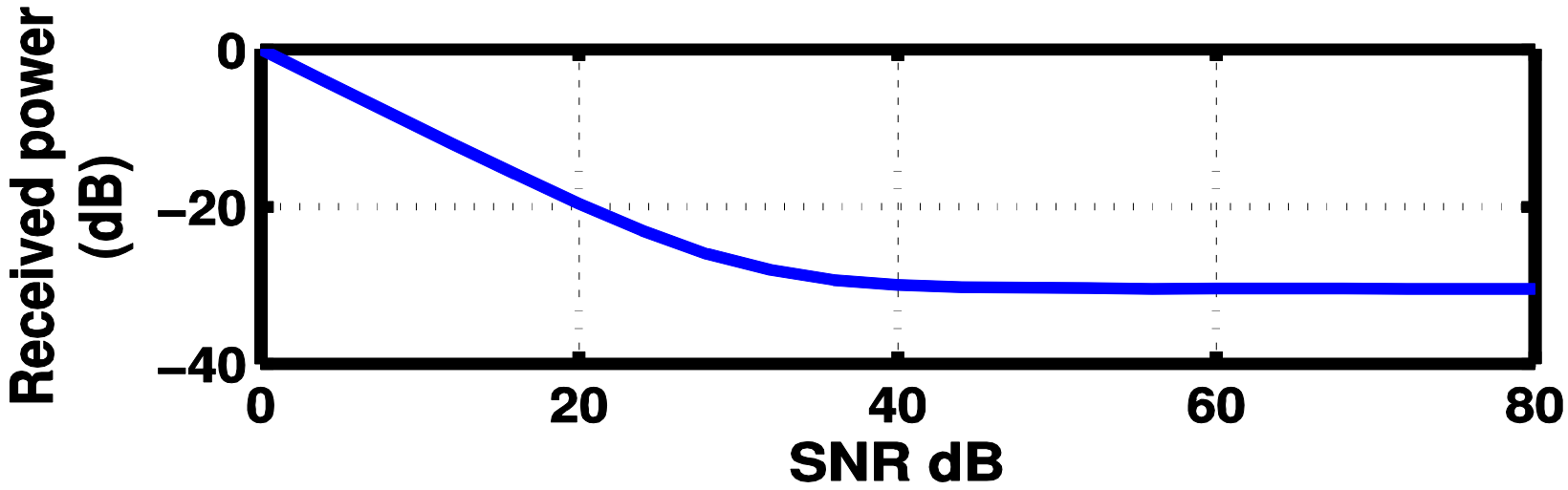
THE UNIVERSITY OF IOWA

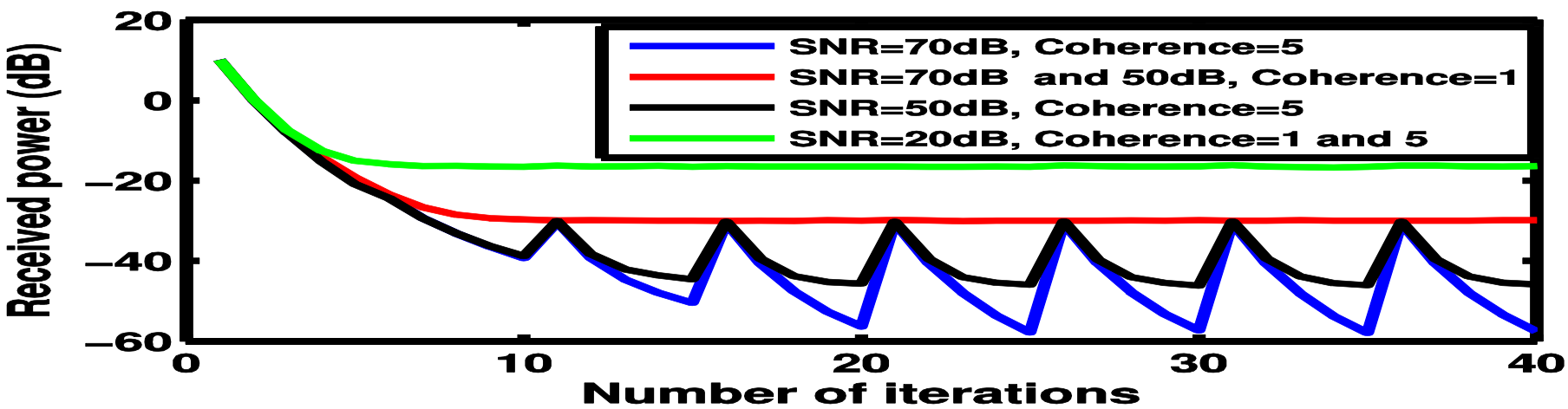
Theoretical limit

- $E\{|\sum_{i=1}^{N} r_i e^{j(\theta_i+\varphi_i)}|^2\}$
- $r_i \sim U[.99, 1.01]$ and $\varphi_i \sim U[-1.5^0, 1.5^0]$
- Subject to
- $|\sum_{i=1}^{N} e^{j\theta_i}|^2 = 0$
- Optimistic performance limit
- Ignores initial offsets

THE UNIVERSITY OF IOWA


Received power
(dB)
0
−20
−40
0
20
40
60
80
SNR dB



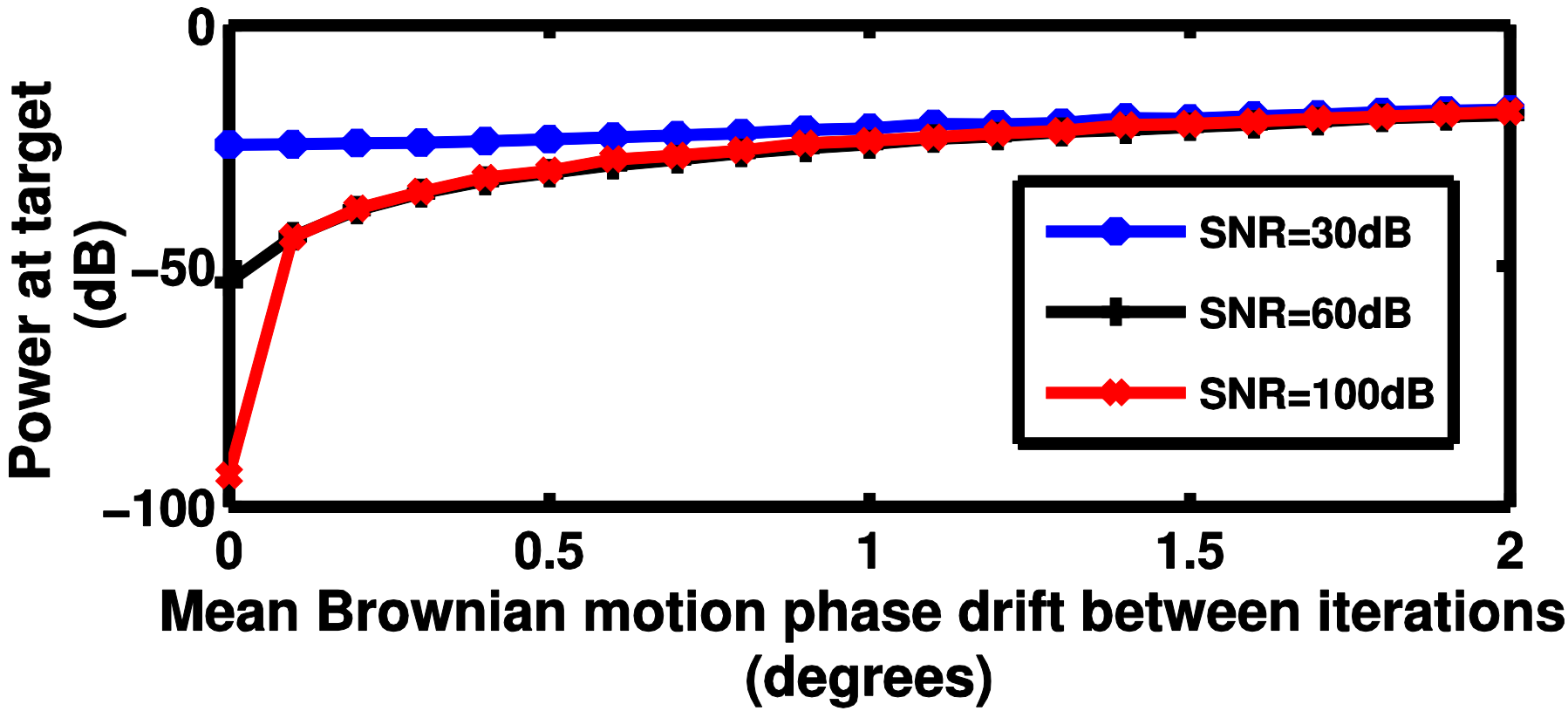


THE UNIVERSITY OF IOWA

- Constant phase offset uniformly distributed between $[0,2\pi]$
- Steady State Gradient:

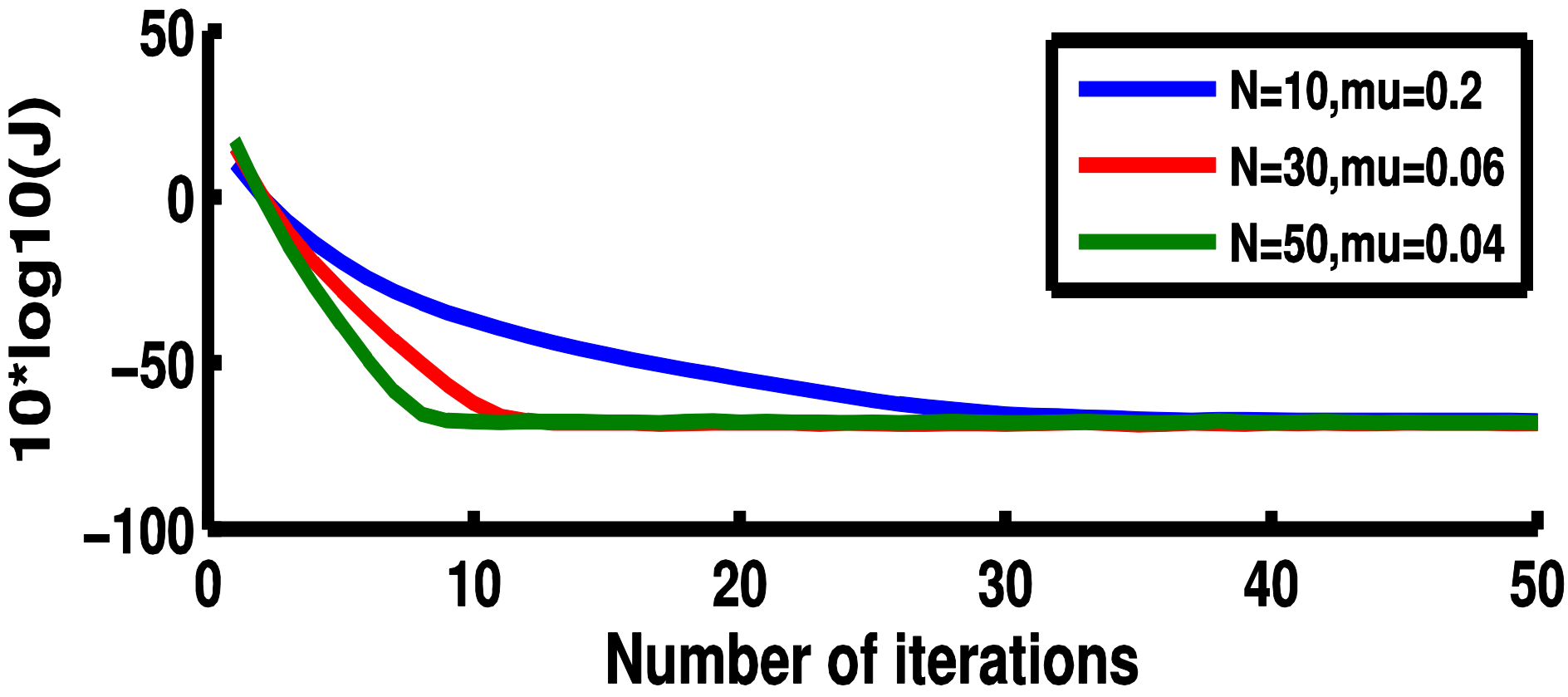
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402
2.1402	2.1402	2.1402	2.1402	2.1402	2.1402

Drifting oscillators: Brownian motion

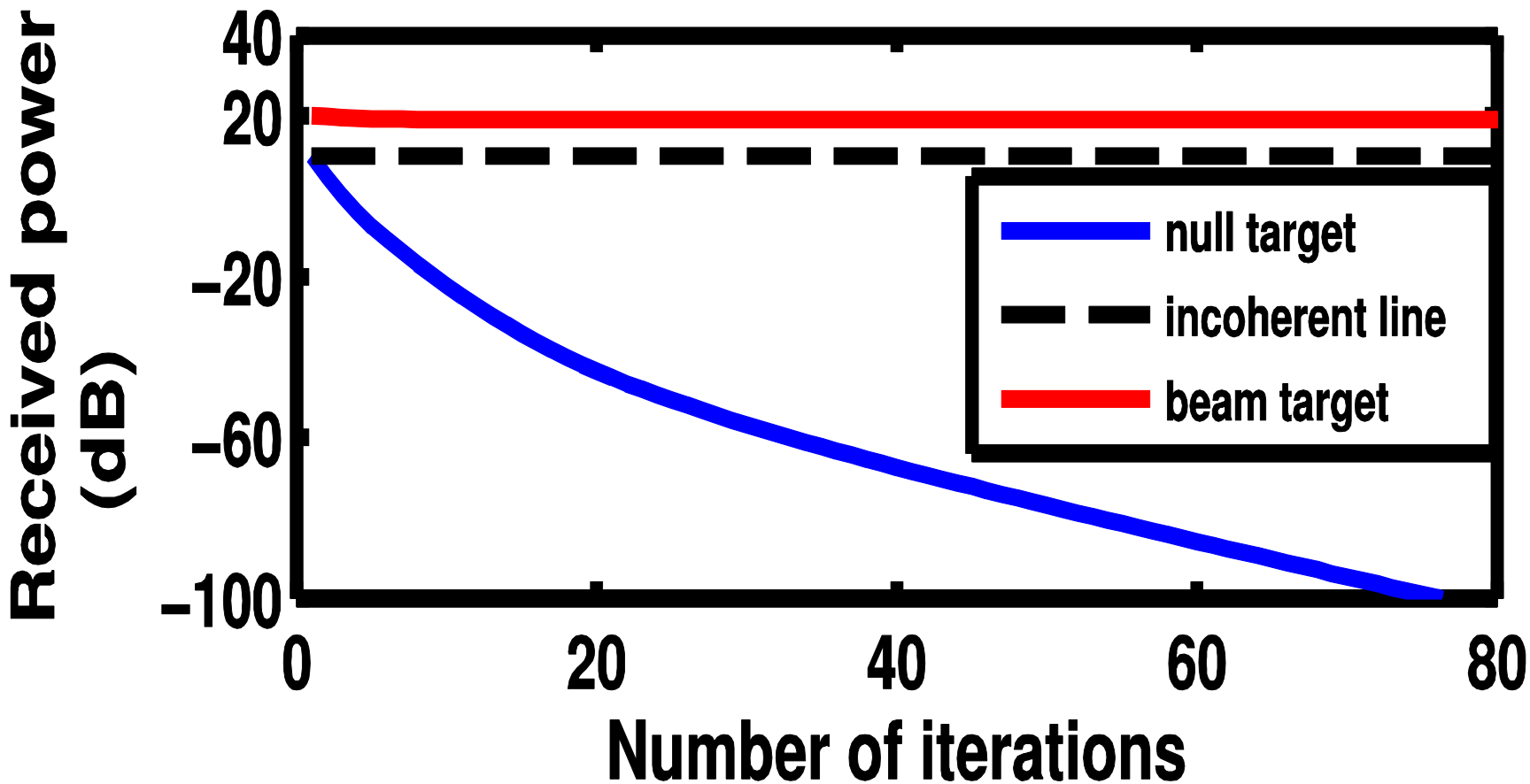


THE UNIVERSITY OF IOWA

Convergence Speed: Scalability



Effect on a Beam: SDMA?



THE UNIVERSITY OF IOWA

Goals of this talk

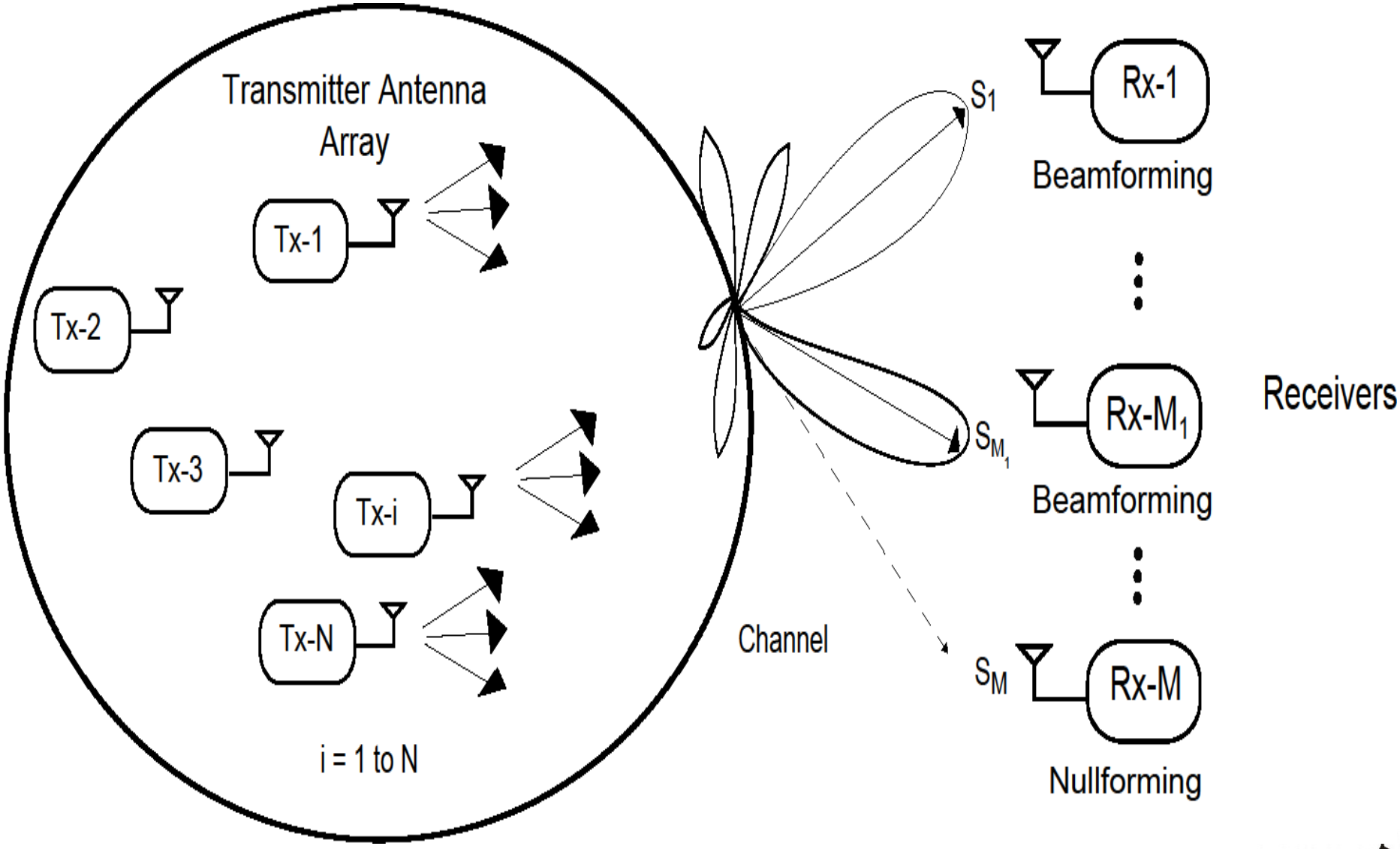
- MISO null forming with phase only adjustments
  - Transmitting at full power
  - Protecting a cooperating transmitter
  - Non-convex problem
- Joint beam and null forming with phase and gain adjustments
  - Null at multiple locations
  - Beam at one location
  - Convex problem

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

The University of Iowa


Transmitter Antenna Array
Tx-1
Tx-2
Tx-3
Tx-i
Tx-N
i = 1 to N
Channel
S1
SM1
SM
Rx-1
Beamforming
Rx-M1
Beamforming
Rx-M
Nullforming
Receivers

Framework

- At startup:
  - l-th receiver sends to the i-th transmitter the channel gain

$$h_{il}^{*} = \alpha_{il} e^{j\varphi_{il}}$$

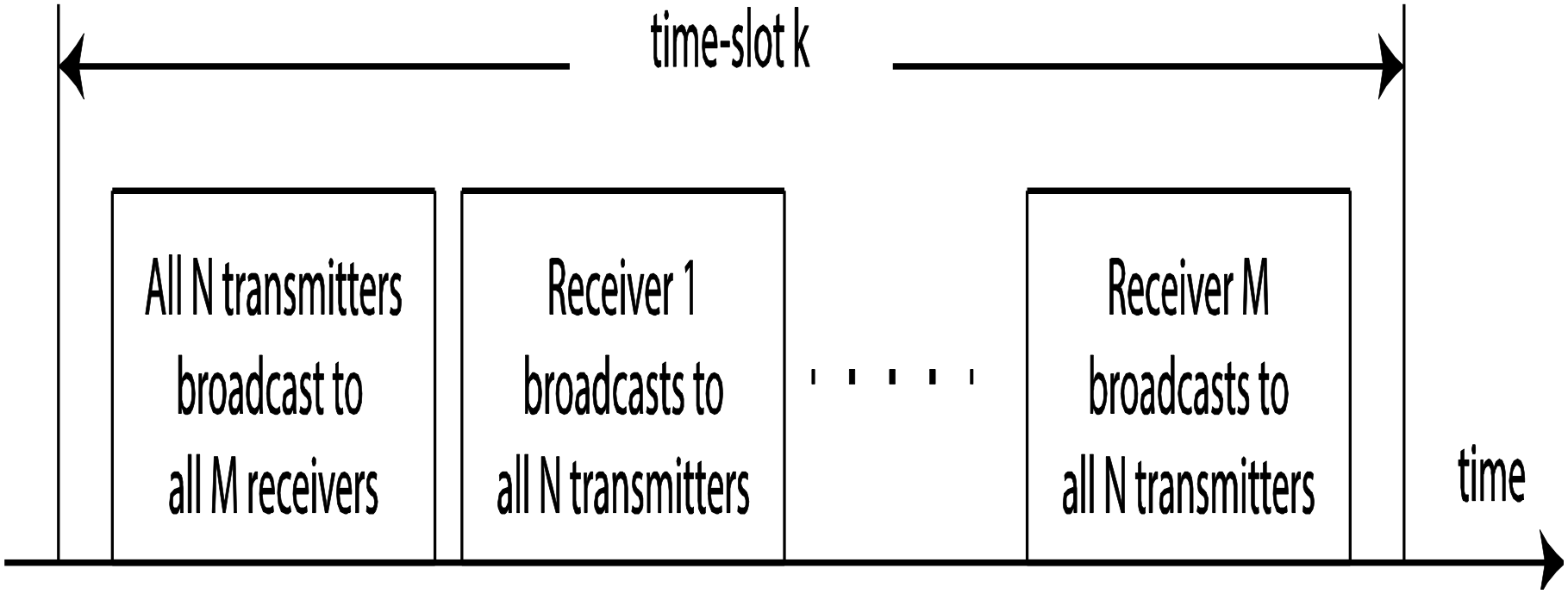
- During operation in the n-th time slot
  - i-th transmitter broadcasts the baseband signal $x_i[n]$
  - l-th receiver broadcasts its total received signal:

$$s_l[n] = \sum_{i=1}^{M} h_{il}^{*} x_i[n] + w_l[n]$$

- Feedback needs TDM
- Adjust x_i to minimize:

$$\sum_{l=1}^{M_1} |s_l|^2 + \sum_{l=M_1+1}^{M} |s_l - b_l|^2$$

TDM for aggregate feedback



Matrix Framework

$$s_l[n]=\sum_{i=1}^{M} h_{il} * x_i[n]+w_l[n]$$

- $s[n]=H^H x[n]+w[n]$ ← white zero mean Gaussian
- Find x to minimize
- $J=E_w[\|H^H x+w[n]-b\|^2]$
- $H:N\times M,\ N>M$: Optimum solutions lie on an $(N-M)$-dimensional subspace: X
- $\arg\min J(x)=\arg\min \|H^H x-b\|^2$

THE UNIVERSITY OF IOWA

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 $x\downarrow min$?
- Effect of noise
  - Is there drift?
- $\lim_{T,N\to\infty}$?
  - Speed
  - $x\downarrow min$

THE UNIVERSITY OF IOWA

Gradient descent

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

- Gradient descent:

$$x[n+1]=x[n]-\mu\partial J(x)/\partial x\,|_{\downarrow\theta=x[n]}$$

$$x[n+1]=x[n]-\mu H_{\smile}(s[n]-b)\ {}_{\top}H^{\uparrow}H\,x[n]-b+w[n]$$

Zero noise convergence

- Assured if $\mu \boxed{W} \downarrow max\,(H\uparrow H\,H)<1$
- $x[0]\rightarrow projection\ of\ x[0]\ on\ X$
- Exponential
- $x[0]=0\ \rightarrow x\downarrow min$
- $\lim_{\top N\rightarrow\infty} \|x\downarrow min\,(N)\|\downarrow 2 = 0$ if channel coefficients are iid CN(0,1)

THE UNIVERSITY OF IOWA

Convergence rate: Scalability

- Suppose the i-th column of H is h_i and H has full column rank.

- As $N \to \infty$ convergence rate is bounded from below if the condition number of $H^H H$ is bounded

- As $N \to \infty$ convergence rate is bounded from below if there is a K and $\alpha_i > 0$ such that for all n,

$$\alpha_1 I \leq \sum_{i=n}^{n+K} h_i h_i^H \leq \alpha_2 I$$

  - Successive channel vectors are persistently spanning

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

THE UNIVERSITY OF IOWA

Noise performance

- $H^{\dagger}H\, x^{*} = b,$ $\Delta[n] = x[n] - x^{*}$

- Error model:

$$\Delta[n+1] = (I - \mu H H^{\dagger} H)\Delta[n] - \mu H w[n]$$

- $T\Delta = (z_1 \,|\, z_2)$: z_1 in the orthogonal complement of X and z_2 on X

- $(z_1[n+1] \,|\, z_2[n+1]) = ((I - \mu A) z_1[n] \,|\, z_2[n]) + \mu(v_1[n] \,|\, v_2[n])$: $v_2[n] \neq 0$ ➔ Brownian motion along the minimizing subspace

Noise performance

- Error model:

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

- $T\Delta=(z_1 \mid z_2)$: z_1 in the orthogonal complement of X and z_2 on X

- $(z_1[n+1] \mid z_2[n+1])=((I-\mu A)z_1[n] \mid z_2[n])+\mu(v_1[n] \mid v_2[n])$: $v_2[n]\neq 0$ ➔ Brownian motion along the minimizing subspace

- $v_2[n]=0$

THE UNIVERSITY OF IOWA

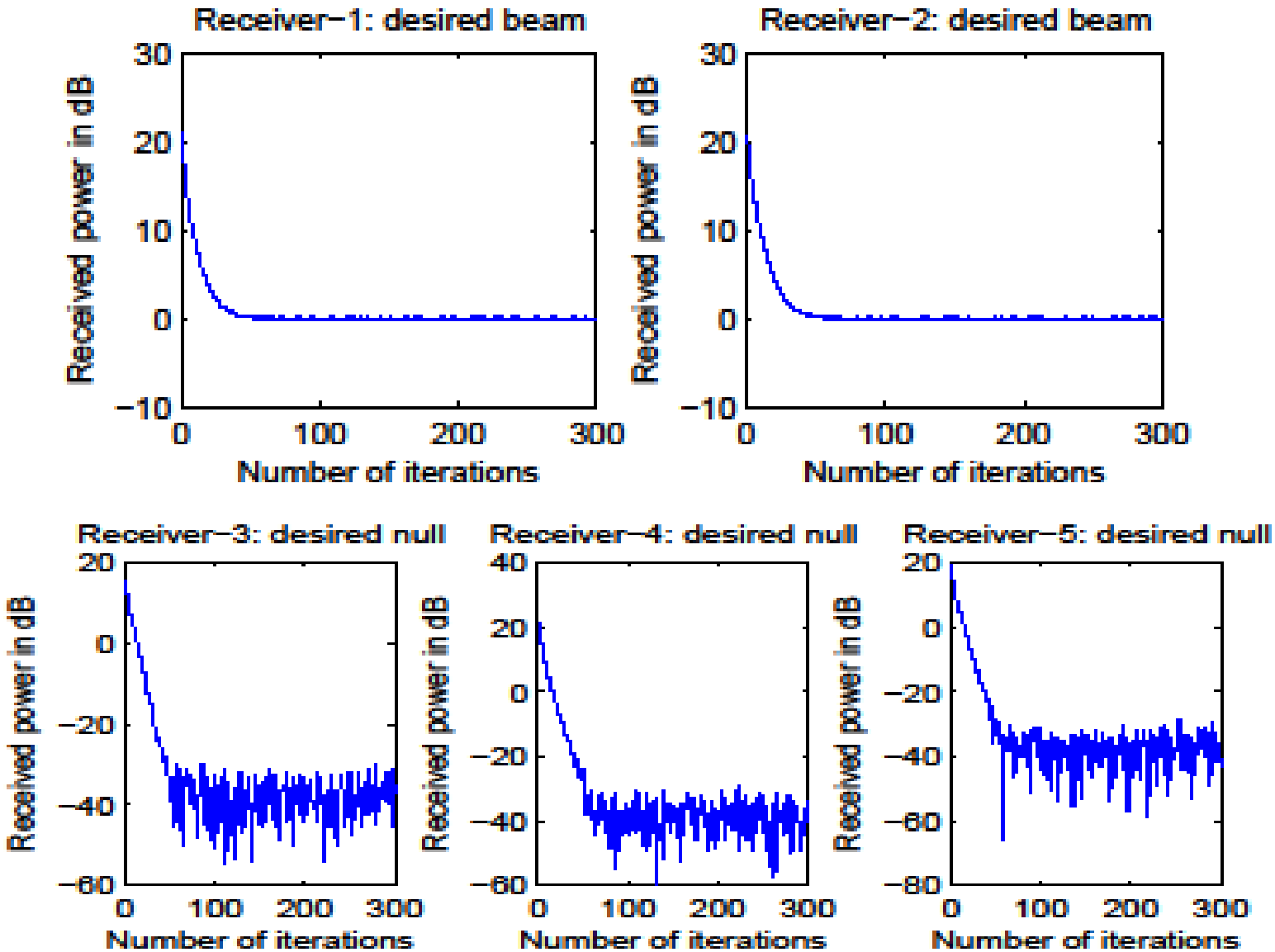
Asymptotic noise performance

- Suppose $x[0]=0$
- Channel coefficients are iid CN(0,1)
- $\lim_{N\to\infty} \lim_{n\to\infty} \| x(n) \|_2 = 0$

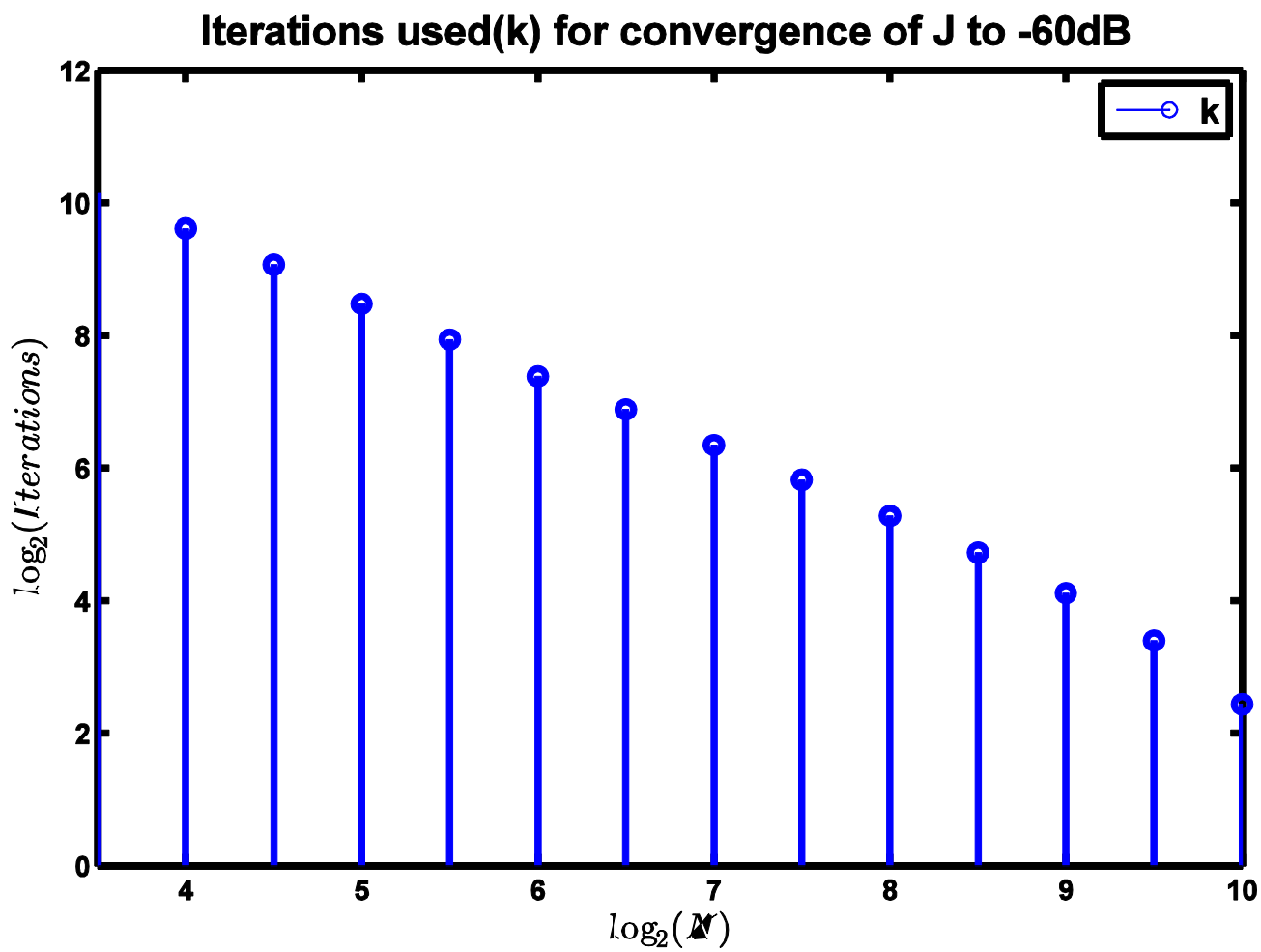
THE UNIVERSITY OF IOWA

Simulation 4

- N=20
- M=5
- 2 beam targets: 1
- SNR=-40dB

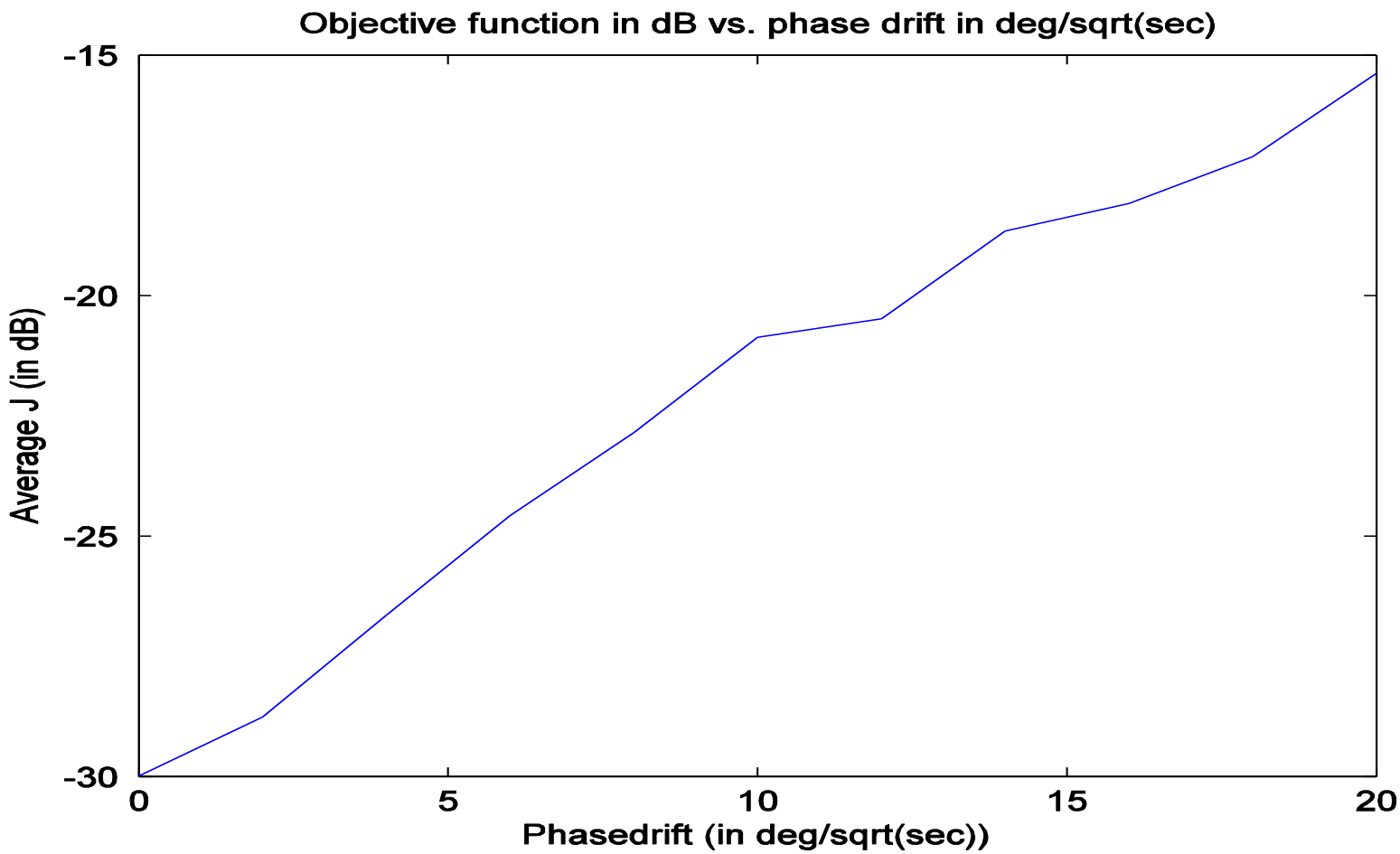

Receiver−1: desired beam
Receiver−2: desired beam
Receiver−3: desired null
Receiver−4: desired null
Receiver−5: desired null
Received power in dB
Number of iterations

Constant μ



THE UNIVERSITY OF IOWA

Phase drift, SNR=-40dB T=50ms, 0 to 5 degrees between iterations



THE UNIVERSITY OF IOWA

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 ➔ Power Efficient Solution➔0 as N➔0
  - No drift along minimizing subspace

THE UNIVERSITY OF IOWA