Research Directions of Departure in Multi-Antenna Wireless Communications: a Compressed Sampling View

Dirk T.M. Slock

Mobile Communication Department, EURECOM

ICSPCC, Hong Kong, 13/08/2012

Some Lessons Learned in Wireless Com

- SDMA (Spatial Division Multiple Access) why did it not take off in the early 90s?
  - No cross-layer design (proper scheduling) at that time.
  - Only feedback was for Power Control.
- At the start of 3G+ activities: it was said that no new PHY development was required, only integration of existing systems. What happened? A lot of PHY work! Dimensions of multi-antenna, multi-user and increasing bandwidth (equalization) were underestimated.
- Wireless standardization starts with the PHY layer. Should become more crosslayer though.
- User Selection: $\Rightarrow$ diversity, simplified transceiver designs.
- Channel Feedback: the return of analog transmission?
- Smart phones: location information everywhere.

Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

Outline

fading, diversity

- fading, probability of error, outage probability
- diversity order
- coding gain (SNR offset)
- Diversity-Multiplexing Tradeoff (DMT)
  - MIMO frequency flat
  - MIMO frequency selective
- receivers: ML, DFE, linear equalizers
- cyclic prefix (CP), Cyclic Delay Diversity (CDD), zero-padding (ZP)

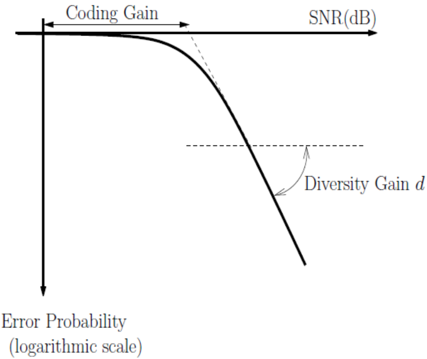
Outage-Rate Tradeoff

- probability of: symbol error, bit error, packet error
- outage probability:
  outage = instantaneous mutual information < rate
- Additive White Gaussian Noise (AWGN) channel:
  $$P_e \sim e^{-\mathrm{SNR}} \qquad (Q(.) \text{ function tail})$$
- fading channel: $P_e = \mathrm{E}_{\,channel}\mathrm{Prob}(error|channel)$
- normalized SINR γ , $\mathrm{SINR} = \rho\,\gamma$
- dominating term in Taylor series of cdf of γ:
  $$\mathrm{Prob}\{\gamma \leq \epsilon\} = c\,\epsilon^k \qquad \text{for small } \epsilon > 0$$
- outage probability $\mathrm{Prob}\{\mathrm{SINR} \leq \alpha\} = c\,\left(\frac{\alpha}{\rho}\right)^k = \left(\frac{\alpha}{g\,\rho}\right)^k$
  k = diversity order, $g = c^{-1/k}$ = coding gain (reduction in SNR required for identical outage probability)
- suboptimal Rx (LE, DFE): channel-equalizer cascade = AWGN channel with Mutual Information $\mathrm{C} = \log(1+\mathrm{SINR})$

Outage-Rate Tradeoff (2)

- at high SNR ρ, $\quad C = \log(1+\text{SINR}) \;\approx\; \log\rho$
- Using Adaptive Modulation and Coding consider taking a fraction r of the capacity:
  rate $R = r\,\log\rho$ (nats), $r \in [0,1]$ is the normalized rate.
- Then the outage probability at high SNR is

$$
\begin{aligned}
P_o &= \text{Prob}\{C \;<\; R\} = \text{Prob}\{\log(1+\text{SINR}) \;<\; \log(\rho^r)\} \\
&= \text{Prob}\{\rho\gamma \;<\; \rho^r - 1\} = \text{Prob}\{\gamma \;<\; \tfrac{1}{\rho^{(1-r)}} - \tfrac{1}{\rho}\} \\
&= \text{Prob}\{\gamma \;<\; \tfrac{1}{\rho^{(1-r)}}\}\,, \text{ for } r > 0 \\
&= c\,\frac{1}{\rho^{(1-r)k}} = \frac{1}{(g\,\rho)^{(1-r)k}}
\end{aligned}
$$



- Hence for $r \in (0, 1]$:

$$d(r) = (1 - r)\,k \quad , \quad g(r)(\text{dB}) = -\frac{10}{(1-r)k} \log_{10} c$$

$d(r)$ = diversity(order)-rate tradeoff,
$g(r)$ = tradeoff dependent coding gain.
The case $r = 0$ (fixed rate) requires separate investigation.

SIMO system

$$\underbrace{\mathbf{y}_k}_{p\times 1} = \underbrace{\mathbf{h}[q]}_{p\times 1}\,\underbrace{a_k}_{1\times 1} + \underbrace{\mathbf{v}_k}_{p\times 1}$$

p is the number of subchannels
the noise power spectral density matrix is $S_{\mathbf{vv}}(z) = \sigma_v^2\, I$
q^{-1} is the unit sample delay operator: $q^{-1}\, a_k = a_{k-1}$
$\mathbf{h}[z] = \sum_{i=0}^{L} \mathbf{h}_i\, z^{-i}$ is the channel transfer function
$L =$ channel delay spread in symbol periods
in the Fourier domain: $\mathbf{h}(f) = \mathbf{h}[e^{j2\pi f}]$

SIMO system (2)

Let $\mathbf{h} = \begin{bmatrix} \mathbf{h}_0 \\ \vdots \\ \mathbf{h}_L \end{bmatrix}$ contain all channel elements.

By default $\mathbf{h} \sim \mathcal{CN}(0, \frac{1}{L+1} I_{p(L+1)})$ i.i.d. channel model

so that spatio-temporal diversity of order $p(L+1)$ is available.

$\rho = \frac{\sigma_a^2}{\sigma_v^2}$ average per subchannel SNR

full CSIR, no CSIT

MFB and SINR of LE and DFE

introduce $\delta = \begin{cases} 0 & \text{, MMSE-ZF design,} \\ 1 & \text{, MMSE design.} \end{cases}$

- $\text{MFB} = \rho \int_{-\frac{1}{2}}^{\frac{1}{2}} \|\mathbf{h}(f)\|^2 \, df = \rho \int_{-\frac{1}{2}}^{\frac{1}{2}} (\|\mathbf{h}(f)\|^2 + \frac{\delta}{\rho}) \, df - \delta$

  arithmetic average
- $\text{SINR}^{\delta}_{DFE} = \rho \, \exp\left[\int_{-\frac{1}{2}}^{\frac{1}{2}} \log(\|\mathbf{h}(f)\|^2 + \frac{\delta}{\rho}) \, df\right] - \delta$

  geometric average
- $\text{SINR}^{\delta}_{LE} = \rho \left[\int_{-\frac{1}{2}}^{\frac{1}{2}} (\|\mathbf{h}(f)\|^2 + \frac{\delta}{\rho})^{-1} \, df\right]^{-1} - \delta$

  harmonic average

$\text{SINR}^{\delta}_{LE} \leq \text{SINR}^{\delta}_{DFE} \leq \text{MFB} \,, \quad \text{SINR}^{0} \leq \text{SINR}^{1}$

Optimal Outage-Rate Tradeoff

- MI with white Gaussian input

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \log(1+\rho\, \|\mathbf{h}(f)\|^2)\, df = \log(1+\mathrm{SINR}_{DFE}^{MMSE})$$

[MedlesSlock:isit04]: $\mathrm{SINR}_{DFE}^{MMSE} \geq \beta_L\, \mathrm{MFB} \sim \|\mathbf{h}\|^2$

- diversity order $k =$ number of constraints need to impose on $\mathbf{h}$ in order to make $\mathrm{SINR} = 0$
- $\Rightarrow\; k = p(L{+}1)$ and

$$d^*(r) = (1-r)\, p\, (L+1)\quad,\quad r \in [0,1]$$

which is valid for Rayleigh $\mathbf{h}$ with non-singular $R_{\mathbf{hh}}$.

LE in SC-CP systems

- after cyclic prefic (CP) insertion, block of N symbols: beq

$$\mathbf{Y} = \mathbf{H}\,A + \mathbf{V} \qquad \mathbf{H} = \begin{bmatrix} \mathbf{h}_0 & & & \mathbf{h}_L & \cdots & \mathbf{h}_1 \\ \vdots & \mathbf{h}_0 & & & \ddots & \vdots \\ \vdots & & \ddots & & & \mathbf{h}_L \\ \mathbf{h}_L & \vdots & & \mathbf{h}_0 & & \\ & \mathbf{h}_L & & & & \\ & & \ddots & \ddots & \ddots & \\ & & & \mathbf{h}_L & \cdots & \mathbf{h}_0 \end{bmatrix}$$

- applying DFT at Rx

$$\underbrace{F_{N,p}\mathbf{Y}}_{\mathbf{U}} = \underbrace{F_{N,p}\mathbf{H}\,F_N^{-1}}_{\mathcal{H}} \;\; \underbrace{F_N\,A}_{X} + \underbrace{F_{N,p}\mathbf{V}}_{\mathbf{W}}$$

where $F_{N,p} = F_N \otimes I_p$,
$\mathcal{H} = \text{blockdiag}\{\mathbf{h}_0, \ldots, \mathbf{h}_{N-1}\}$ with $\mathbf{h}_n = \mathbf{h}(f_n)$, $f_n = \frac{n}{N}$.
At tone n: $\quad \mathbf{u}_n = \mathbf{h}_n\, x_n + \mathbf{w}_n$.

LE in SC-CP systems (2)

- ZF ($\delta = 0$) or MMSE ($\delta = 1$) LE produces per tone $\widehat{x} = (\mathbf{h}^H\mathbf{h} + \frac{\delta}{\rho})^{-1}\mathbf{h}^H\mathbf{u}$ from which $\widehat{a}$ is obtained after IDFT with

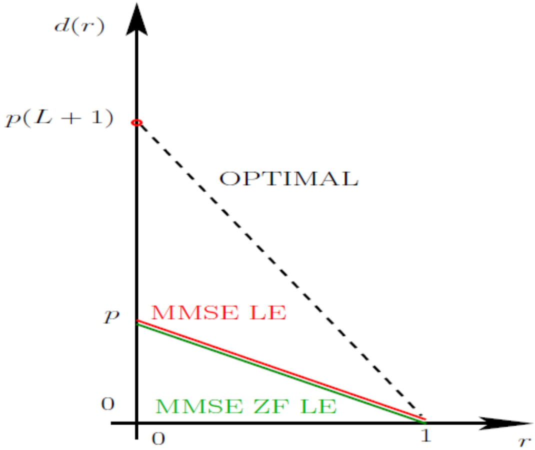
$$\mathrm{SINR}^{\delta}_{CP-LE} = \rho \left(\frac{1}{N} \sum_{n=0}^{N-1} (\|\mathbf{h}_n\|^2 + \frac{\delta}{\rho})^{-1} \right)^{-1} - \delta$$

- For $N \geq L{+}1$, the (ZF) outage manifold is the collection of manifolds for which $\mathbf{h}_n = 0$ for some n ($\mathrm{SINR} = 0$). As a result the codimension is p. Hence

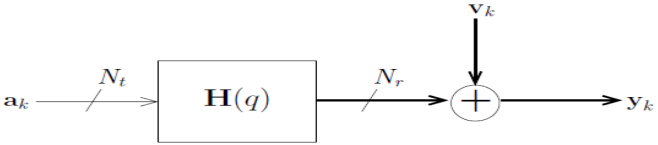
$$d^{ZF}_{CP-LE}(r) = (1 - r)\,p \quad , \quad r \in [0, 1]$$

any frequency diversity is lost! Also for MMSE, except $d^{MMSE}_{CP-LE}(0) = p\,(L + 1)$: the MMSE has full diversity at constant rate R (at finite SNR, some $r > 0$ also).

LE in SC-CP systems (3)



MIMO Channel Model



$$\mathbf{H}(q) = \sum_{l=0}^{L-1} \mathbf{H}_l \, q^{-l}, \quad q^{-1} x_k = x_{k-1}. \qquad \mathbf{H}_l\text{: } N_r \times N_t.$$

L: channel delay spread. SNR $\rho = \frac{P}{N_t \sigma_v^2} = \frac{\sigma_a^2}{\sigma_v^2}$.

$$\mathbf{y}_k = \mathbf{H}(q)\,\mathbf{a}_k + \mathbf{v}_k = \sum_{l=0}^{L-1} \mathbf{H}_l \mathbf{a}_{k-l} + \mathbf{v}_k \,,$$

entries of $\mathbf{H}_l$, $l = 0, \ldots, L-1$: i.i.d. Gaussian $\mathbf{H}_l^{rt} \sim \mathcal{CN}(0,1)$.

Diversity vs Multiplexing Background

[Zheng&Tse'03]

- A coding scheme $\mathcal{C}(\rho)$ is a family of codes of block length T, that supports a bit rate $R(\rho)$.
- Spatial multiplexing r: $\lim\limits_{\rho\to\infty} \dfrac{R(\rho)}{\ln(\rho)} = r$ $\qquad R(\rho) = a + r\ln(\rho)$
- Diversity gain d: $\lim\limits_{\rho\to\infty} \dfrac{\ln P_e(\rho)}{\ln(\rho)} = -d$ $\qquad P_e(\rho) = \dfrac{c}{\rho^d}$

$d^*(r)$ denotes the supremum of the diversity advantage achieved over all possible schemes.

In practice: $\quad P_e(\rho) = P_{out}(\rho)\,,\ \ d^*(r) = d^{out}(r)\ .$

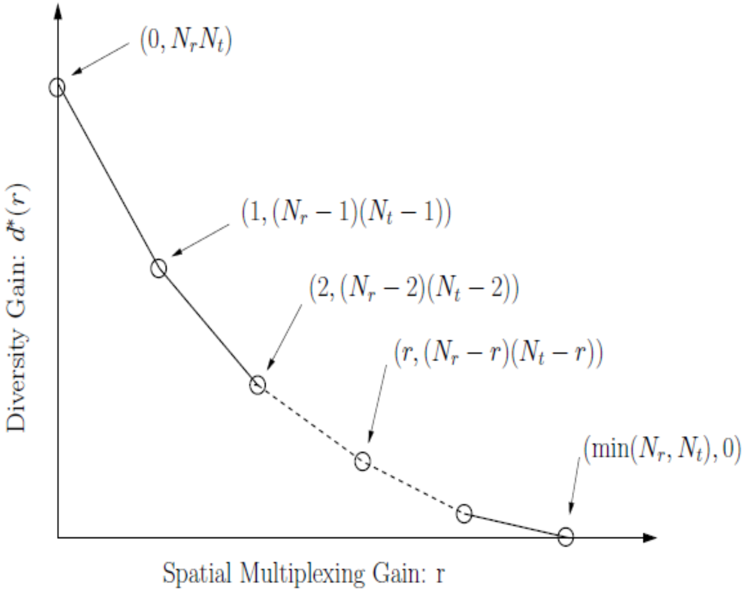
Diversity vs Multiplexing Background(2)

- **Flat MIMO channel ($L = 1$)** For $T \geq N_t$, the optimal trade-off curve $d^*(r)$ is given by the piecewise-linear function connecting the points $(k, d^*(k))$, $k = 0, 1, \ldots, q$, where

$$d^*(k) = (p-k)(q-k),$$
$$q = \min\{N_r, N_t\},$$
$$p = \max\{N_r, N_t\}.$$

  Achieved by the family of codes with non-vanishing determinant [Elia, Pawar *et al*'ALLERTON04].
- **SIMO/MISO frequency selective channel** The optimal trade-off curve is given by the linear function $d^*(r) = L\,p\,(1-r)$ [Grokop& Tse'ISIT04]. For SIMO achieved by using QAM at Tx and MMSE DFE at Rx [Medles& Slock'ISIT04].

Diversity vs Multiplexing Background(3)



DMT for MIMO FS Channel (1)

- The optimal trade-off curve $d^*(r)$ is given by the piecewise-linear function connecting the points $(k, d^*(k))$, $k = 0, 1, \ldots, p$, where

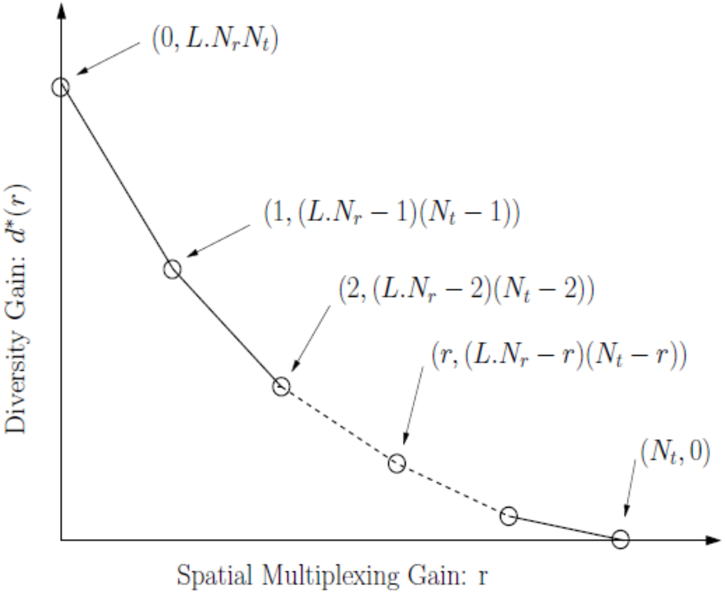
$$\begin{aligned} d^*(k) &= (Lq - k)(p - k), \\ p &= \min\{N_r, N_t\}, \\ q &= \max\{N_r, N_t\}. \end{aligned}$$

- For $N_t \leq N_r$, diversity is the same as for a flat MIMO channel with $N_t{}' = N_t$ and $N_r{}' = LN_r$.
- Coding over L independent OFDM subcarriers (spacing of $\frac{1}{L}$): $d(r) = L(q - r)(p - r) \ \leq \ d^*(r) \rightarrow$ suboptimal.

  Difference $d^*(r) - d(r) = (L-1)\, r\, (p - r)$, peaks at $r = \frac{p}{2}$.

  For large L and $N_r = N_t$, $d^*(p/2) \approx 2\, d(p/2)$.

DMT for MIMO FS Channel (2) $N_r \leq N_t$



Diversity Order Analysis

- parameterization FIR channel of rank $k \leq p = N_t \leq N_r = q$

$$\underbrace{\underbrace{\mathbf{H}(z)}_{q\times p}}_{\text{FIR}-L} = \underbrace{\underbrace{\overline{\mathbf{H}}(z)}_{q\times k}}_{\text{FIR}-L} \; [\, I_k \quad \underbrace{\underbrace{\overline{\overline{\mathbf{H}}}}_{k\times(p-k)}}_{\text{constant}} \,] \quad \underbrace{\mathcal{P}}_{\text{permutation}}$$

degrees of freedom in $q \times p$ rank-k FIR-L manifold:

$$\underbrace{q\,k\,L}_{\overline{\mathbf{H}}} + \underbrace{(p-k)\,k}_{\overline{\overline{\mathbf{H}}}} = q\,p\,L - (q\,L-k)(p-k)$$

- To send at rate k, need to be guaranteed rank k. The diversity degree is the remaining number of degrees of freedom in $\mathbf{H}(z)$:

$$d^*(k) = q\,p\,L - (q\,p\,L - (q\,L-k)(p-k)) = (q\,L-k)(p-k)$$

SISO SC-ZP: [Tepedelenliogu:IT04,globecom05]

- SC-ZP: Single Carrier with Zero Padding (Cyclic Prefix = zeros)

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_0 & 0 & \cdots & 0 \\ \mathbf{h}_1 & \mathbf{h}_0 & & \vdots \\ \vdots & & \ddots & \\ \mathbf{h}_{L-1} & & & \mathbf{h}_0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \mathbf{h}_{L-1} \end{bmatrix} = F_N^{-1} \mathcal{H} F_N \begin{bmatrix} I_{N-L+1} \\ 0 \end{bmatrix}$$

- block MMSE-ZF linear RX: $(\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H$
- $\mathbf{H}^H\mathbf{H}$ only singular if $\mathbf{h} \equiv 0 \quad \Rightarrow \quad$ full diversity.

Extending Tepedelenlioglu to MIMO

- frequency-selective MIMO channel with white Gaussian input:

$$C = \int_{-\frac{1}{2}}^{\frac{1}{2}} \log\det(I_{N_t} + \rho\, \mathbf{h}^H(f)\, \mathbf{h}(f))\, df$$

  equivalent system spectrum $\det(I_{N_t} + \rho\, \mathbf{h}^H(f)\, \mathbf{h}(f)) = \det(I_{N_r} + \rho\, \mathbf{h}(f)\, \mathbf{h}^H(f))$ has memory $\min\{N_r, N_t\}\, L - 1$.
- $\Rightarrow$ if LE diversity to be maximized through ZP $\Rightarrow$ explicit input-output system memory of $\min\{N_r, N_t\}\, L - 1$ is required. Channel memory is only $L-1$, it needs to be increased by linear precoding.

Transmit Diversity Modes

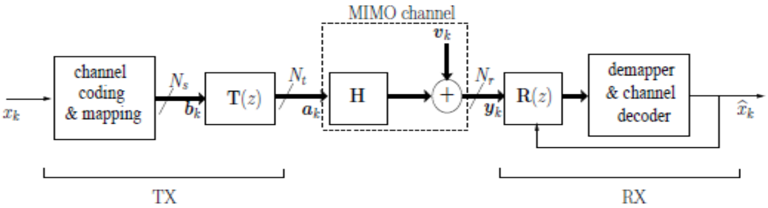
- DTD: Delay Transmit Diversity
  Both Tx antennas transmit the same signal, but with a differential delay D.
  h_i is the channel between Tx antenna i and the mobile and the overall channel is:

$$h(z) = h_1(z) + z^{-D}\, h_2(z) = [h_1(z) \;\; h_2(z)] \begin{bmatrix} 1 & 0 \\ 0 & z^{-D} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

  increases delay spread and requires (more) equalization (only an issue of CP/ZP length in CP/ZP systems)
- STTD: Space-Time Transmit Diversity (Alamouti)

MIMO Linear Convolutive ST Precoding

- general ST precoding setup:



inner code: linear precoding, outer code: channel coding

MIMO Linear Precoder

- This can be accomplished with the convolutive STC introduced in [MedlesSlock:IT06], a structured form of linear dispersion codes. As precoding can only be done at the Tx side, the system memory gets actually increased to $N_t L-1$ for maximal diversity. As at high SNR, $C \approx \min\{N_r, N_t\}\ \log(\rho)$, it suffices (esp. for a LE Rx) to limit the number of symbol streams to $N_s = \min\{N_r, N_t\}$ (instead of N_t).
- paraunitary prefilter $\mathbf{T}[z^L]$

$$\begin{aligned} \mathbf{T}[z] &= \mathbf{D}[z]\, Q\,, \quad Q^H Q = N_t\, I_{N_s}\,, |Q_{ij}| = 1\,, \\ \mathbf{D}[z] &= \mathrm{diag}\{1,\, z^{-1}, \ldots, z^{-(N_t-1)}\} \end{aligned}$$

  where Q is a (constant) tall unitary matrix with equal magnitude elements (e.g. a submatrix of a DFT or a Walsh-Hadamard matrix).
- Corresponds to (MIMO) Cyclic Delay Diversity (CDD) of LTE (circulant filtering).

Cascaded Prefilter-Channel System

- cascaded prefilter-channel system $\mathbf{g}[z] = \mathbf{h}[z]\,\mathbf{T}[z^L] = \sum_{i=1}^{N_t L-1} \mathbf{g}_i z^{-i}$ for which we introduce ZP of length $N_t L{-}1$ symbol periods in a block of length N. This leads to the banded block Toeplitz system matrix

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 & 0 & \cdots & 0 \\ \mathbf{g}_1 & \mathbf{g}_0 & & \vdots \\ \vdots & & \ddots & \\ \mathbf{g}_{N_t L-1} & & & \mathbf{g}_0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & \mathbf{g}_{N_t L-1} \end{bmatrix} = \underbrace{[\mathbf{G}_0 \cdots \mathbf{G}_{N-N_t L}]}_{NN_r \times (N-N_t L+1) N_s}$$

Simulations

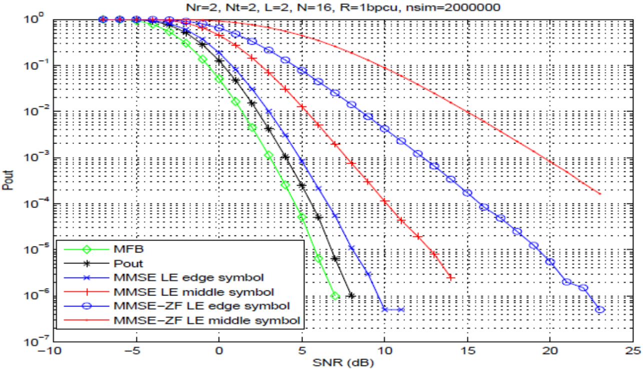


Figure : Outage probability vs. SNR for MFB, block outage, MMSE LE and MMSE-ZF LE for the case $N_t = N_r = 2$, $L = 2$, $N = 16$, $R = 1$bpcu, and 2.10^6 Monte Carlo runs.

Diversity: Issues

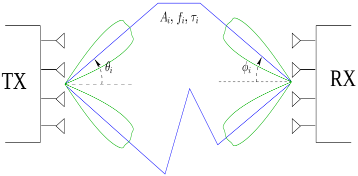
- (no CSIT) systems with equal diversity/DMT: compare coding gain!
  - delay spread L: same diversity order as L receive antennas, but much worse coding gain probably
  - cyclic prefix $\rightarrow$ zero-padding: restores diversity in linear receivers, but with bad coding gain
- partial CSIT: diversity still an issue (space-time coding useful)
- diversity in multi-user problems?
  typically with (full) CSIT: no outage

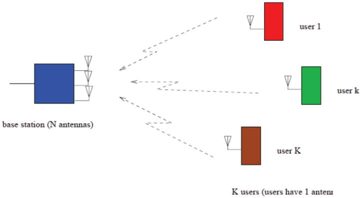
Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
  BC: multi-user (MU) downlink (DL) with multiple BS antennas
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation
  (Wiener/Kalman filtering + parameters)

SDMA considerations

- Whereas single user (SU) MIMO communications represented a big breakthrough and are now integrated in a number of wireless communication standards, the next improvement is indeed multi-user MIMO (MU MIMO).
- This topic is nontrivial as e.g. illustrated by the fact that standardization bodies were not able to get an agreement on the topic until recently to get it included in the LTE-A standard.
- MU MIMO is a further evolution of SDMA, which was THE hot wireless topic throughout the nineties.





MU MIMO key elements

- SDMA is a suboptimal approach to MU MIMO, with transmitter precoding limited to linear beamforming, whereas optimal MU MIMO requires Dirty Paper Coding (DPC).
- Channel feedback has gained much more acceptance, leading to good Channel State Information at the Transmitter (CSIT), a crucial enabler for MU MIMO, whereas SDMA was either limited to TDD systems (channel CSIT through reciprocity) or Covariance CSIT. In the early nineties, the only feedback that existed was for slow power control.
- Since SDMA, the concepts of multiuser diversity and user selection have emerged and their impact on the MU MIMO sum rate is now well understood. Furthermore, it is now known that user scheduling allows much simpler precoding schemes to be close to optimal.

MU MIMO key elements (2)

- Whereas SU MIMO allows to multiply transmission rate by the spatial multiplexing factor, when mobile terminals have multiple antennas, MU MIMO allows to reach this same gain with single antenna terminals.
- Whereas in SU MIMO, various degrees of CSIT only lead to a variation in coding gain (the constant term in the sum rate), in BLUE MU MIMO however CSIT affects the spatial multiplexing factor (= Degrees of Freedom (DoF)) (multiplying the log(SNR) term in the sum rate).

CSIT considerations

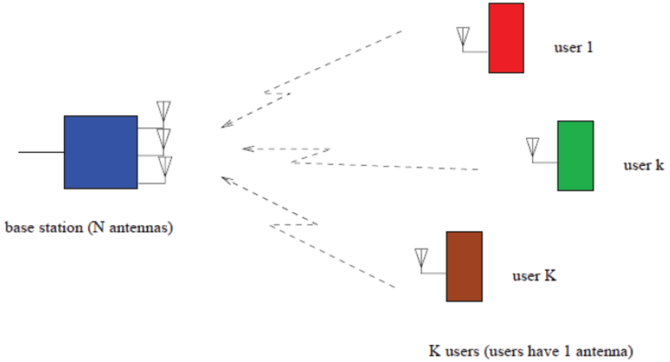
In the process attempting to integrate MU-MIMO into the LTE-A standard, a number of LTE-A contributors had recently become extremely sceptical about the usefulness of the available MU-MIMO proposals. The isue is that they currently do MU-MIMO in the same spirit as SU-MIMO, i.e. with feedback of CSI limited to just a few bits! However, MU-MIMO requires very good CSIT! Some possible solutions:

- Increase CSI feedback enormously (possibly using analog transmission).
- Exploit channel reciprocity in TDD (electronics calibration issue though).
- Limit MU-MIMO to LOS users and extract essential CSIT from DoA or location information.

SDMA system model

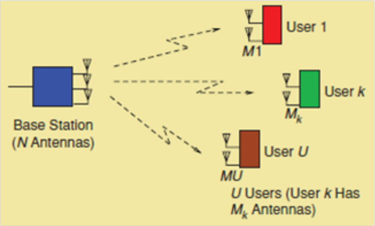
- Rx signal at user k:

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k, \qquad k = 1, 2, \ldots, K$$



MIMO BroadCast

- MIMO BC = Multi-User MIMO Downlink
- N_t transmission antennas.
- K users with N_k receiving antennas.
- Assume perfect CSI
- Possibly multiple streams/user d_k.
- Power constraint P
- Noise variance $\sigma^2 = 1$.
- $\mathbf{H}_k$ the MIMO channel for user k.



$$
\begin{aligned}
\mathbf{F}_k\mathbf{y}_k &= \mathbf{F}_k\mathbf{H}_k\textstyle\sum_{i=1}^{K}\mathbf{G}_i\mathbf{s}_i + \mathbf{F}_k\mathbf{z}_k \\
&= \underbrace{\mathbf{F}_k\mathbf{H}_k\mathbf{G}_k\mathbf{s}_k}_{\text{useful signal}} + \underbrace{\sum_{i=1, i\neq k}^{K}\mathbf{F}_k\mathbf{H}_k\mathbf{G}_i\mathbf{s}_i}_{\text{inter-user interference}} + \underbrace{\mathbf{F}_k\mathbf{z}_k}_{\text{noise}}
\end{aligned}
$$

System Model (2)

- Rx signal: $\mathbf{y}_k = \mathbf{H}_k\mathbf{x} + \mathbf{z}_k = \mathbf{H}_k \sum_{i=1}^{K} \mathbf{G}_i\mathbf{s}_i + \mathbf{z}_k$
- $$\underbrace{\mathbf{F}_k}_{d_k\times N_k}\underbrace{\mathbf{y}_k}_{N_k\times 1} = \underbrace{\mathbf{F}_k}_{d_k\times N_k}\underbrace{\mathbf{H}_k}_{N_k\times N_t}\sum_{i=1}^{K}\underbrace{\mathbf{G}_i}_{N_t\times d_i}\underbrace{\mathbf{s}_i}_{d_i\times 1} + \underbrace{\mathbf{F}_k}_{d_k\times N_k}\underbrace{\mathbf{z}_k}_{N_k\times 1}$$
- [Christensen etal:T-WC08]: use of linear receivers in MIMO BC is not suboptimal (full CSIT, // SU MIMO): can prefilter $\mathbf{G}_k$ with a $d_k \times d_k$ unitary matrix to make interference plus noise prewhitened channel matrix - precoder cascade of user k orthogonal (columns)

User Selection Motivation

- Optimal MIMO BC design requires DPC, which is significantly more complicated than BF.
- User selection allows to
  - improve the rates of DPC
  - bring the rate of BF close to those of DPC
- Optimal user/stream selection requires selection of optimal combination of N_t streams: too complex. Greedy user/stream selection (GUS): select one stream at a time $\Rightarrow$ complexity $\approx$ N_t times the complexity of selecting one stream ($K \gg N_t$).
- Multiple receive antennas cannot improve the sum rate prelog. So what benefit can they bring?
  Of course: cancellation of interference from other transmitters (spatially colored noise): not considered here.

Zero-Forcing (ZF)

- ZF-BF $\mathbf{F}_{1:i}\mathbf{H}_{1:i}\mathbf{G}_{1:i} =$

$$\begin{bmatrix} \mathbf{F}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{F}_2 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{F}_i \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_i \end{bmatrix} [\mathbf{G}_1 \; \mathbf{G}_2 \cdots \mathbf{G}_i] = \begin{bmatrix} \mathbf{F}_1\mathbf{H}_1\mathbf{G}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{F}_2\mathbf{H}_2\mathbf{G}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{F}_i\mathbf{H}_i\mathbf{G}_i \end{bmatrix}$$

- ZF-DPC (modulo reordering issues) $\mathbf{F}_{1:i}\mathbf{H}_{1:i}\mathbf{G}_{1:i} =$

$$\begin{bmatrix} \mathbf{F}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{F}_2 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{F}_i \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_i \end{bmatrix} [\mathbf{G}_1 \; \mathbf{G}_2 \cdots \mathbf{G}_i] = \begin{bmatrix} \mathbf{F}_1\mathbf{H}_1\mathbf{G}_1 & 0 & \cdots & 0 \\ * & \mathbf{F}_2\mathbf{H}_2\mathbf{G}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ * & \cdots & * & \mathbf{F}_i\mathbf{H}_i\mathbf{G}_i \end{bmatrix}$$

- BF-style selection, DPC-style selection: as if it's going to be used in BF/DPC

Stream Selection Criterion from Sum Rate

- At high SNR, both
  - optimized (MMSE style) filters vs. ZF filters
  - optimized vs. uniform power allocation

  only leads to $\frac{1}{\text{SNR}}$ terms in rates.
- At high SNR, the sum rate is of the form

$$\underbrace{N_t}_{\color{red}{\text{DoF}}} \log(\text{SNR}/N_t) + \underbrace{\sum_i \log\det(\mathbf{F}_i\mathbf{H}_i\mathbf{G}_i)}_{\color{blue}{\text{constant}}} + \mathrm{O}(\frac{1}{\text{SNR}}) + \underbrace{\mathrm{O}(\log\log(\text{SNR}))}_{\color{red}{\text{noncoherent Tx}}}$$

for properly normalized ZF Rx $\mathbf{F}_i$ and ZF Tx $\mathbf{G}_i$ (BF or DPC).

MIMO BC Greedy Stream Selection (GUS) Criteria

One can focus on either

(i) (the constant term in) the sum rate at high SNR,

(ii) the sum rate at any SNR of the associated ZF transceiver designs with uniform power loading,

(iii) the sum rate at any SNR of the associated ZF transceiver designs with waterfilling,

(iv) the sum rate at any SNR of optimized transceiver designs.

At high SNR, (i) is the analysis of interest.

More variations could be considered, e.g. regularized ZF as an intermediate between ZF and optimized transceiver designs.

State of the Art: MISO BC

- [TuBlum:COMlet99]: Gram-Schmidt channel orthogonalization with pivoting (DPC-style GUS)
- [DimicSidiropoulos:T-SP05]: introduced proper BF-style GUS, large K analysis DPC-style GUS, simulations. Matrix inversion lemma for bordered matrices, in order to lower complexity of BF-style GUS.
- [YooGoldsmith:06]: analyze BF, but w pseudo-BF-style GUS: SUS (semi-orthogonal) = DPC-style GUS + inner product constraints (limiting size of pool of users for selection). Show that for BF-SUS, as for DPC-US,

$$\lim_{K\to\infty} \frac{SR}{N_t \log(1 + \frac{P}{N_t} \log K)} = 1$$

- [WangLoveZoltowski:08]: small refinement of [YooGoldsmith:06], has more constraints.

State of the Art: MIMO BC

- [SunMcKay:10]: transforms MIMO to MISO. [WangLoveZoltowski:08] type analysis. Pseudo-BF-style GUS (SUS). Analysis for use in DPC and in BF. Analysis only shows effect of antennas in higher-order terms.
- [Jindal:TWC08]: single stream MIMO BC, use of Rx antennas to minimize quantization error (for FB) on resulting virtual channel. Emphasis on partial CSIT (and CSIR) with (G)US.
- [HungerJoham:ciss09]: obtain the high SNR SR offset between BF and DPC, without user selection. They extend the [Jindal:isit05] analysis from MISO to MIMO. [Hassibi] also.
- [Utschick:06] SESAM: proper DPC-style GUS for MIMO case (extension of [TuBlum] from MISO to MIMO).

State of the Art: MIMO BC (2)

- [GuthyUtschick:1209]: propose a BF-style GUS for MIMO-BC-BF. In the style of predecessors, they only adapt the Rx of the new stream to be added. They replace the proper geometric average of the stream channel powers by its harmonic average: $1/\mathrm{tr}\{diag((HH^H)^{-1})\}$, which leads to a generalized eigenvector solution for the Rx filter (minFrob algo). Can be simplified to a classical eigenvector problem: the LISA algo = SESAM algo. No asymptotic analysis.
- [GuthyUtschick:T-SP0410]: same greedy approaches are proposed now for max WSR, without user selection.
- [reference in Yu?]: introducing more than one sweep in GUS.
- [Christensen:TW08]: working per stream is equivalent to working per user.
- many other references on MIMO BF design for max (W)SR

Asymptotic Regimes

Assume for simplicity all $N_k = N_r$. Possible asymptotic regimes for analyzing US:

(0) $K \to \infty$, N_t, N_r finite.

(1) $N_t \to \infty$, $N_r = \alpha N_t$, α, K finite. In this regime, one lets $N_t, N_r \to \infty$ and then one introduces selection mechanisms.

(2) $N_t \to \infty$, $N_r = \alpha N_t$, α, K finite. In this regime, one first introduces selection mechanisms and then one lets $N_t, N_r \to \infty$.

(3) $N_t \to \infty$, $K = \beta N_t$, β, N_r finite.

Regime (0) is the classical asymptotic regime for the analysis of the effect of stream selection. However, the results of this analysis are only relevant when K is extremely large.

The behavior of various stream selection mechanisms is more interesting when there is some redundancy allowing selection, but not too much. In order to benefit from simplified asymptotic analytical results, consider $N_t \to \infty$. Some room for selection is then obtained when $KN_r \to \infty$ also, but at the same rate as N_t with $KN_r > N_t$.

One way for $KN_r \to \infty$ is to keep K finite and let $N_r \to \infty$ as in single-user MIMO asymptotic analysis. The correct selection analysis corresponds to regime (2), whereas regime (1) is a simplified approximation of (2), and is considered in [GuthyUtschickHonig:isit10]. Indeed, in this case all user channels behave identically asymptotically, and hence the selection process becomes very simple.

Another way for $KN_r \to \infty$ is to keep N_r finite and let $K \to \infty$. Such analysis would also encompass the MISO case.

From MISO to MIMO: Rx design

Design of receivers:

(i) Fixed unitary receiver. For i.i.d. channels, any fixed unitary Rx is equivalent, hence can choose an identity matrix, in which case

$$\text{stream selection} = \text{Rx antenna selection.}$$

The analysis of the corresponding algorithms is very simple since in this case K users MIMO with N_r Rx antennas is equivalent to MISO with KN_r users.

(ii) Rx for user k is optimized only in function of its channel $\mathbf{H}_k$. E.g. [SunMcKay:T-SP10]: singular modes of $\mathbf{H}_k$. Rx fixed $\Rightarrow$ K user MIMO $= K\,N_r$ user MISO, but virtual MISO channels no longer i.i.d.

(iii) Optimized receivers (Rx).

MISO DPC-style GUS

- GUS: Greedy User Selection
  Gram-Schmidt orthogonalize $\mathbf{h}_k$ w.r.t. those of already selected users and choose user with maximum residual norm (matched to DPC).
- MISO: $h_k = \mathbf{H}_k^H$, k_i = user selected at stage i, $H_i = h^H_{k_{1:i}}$.
- $$\det(H_i H_i^H) = \prod_{j=1}^{i} \|P^\perp_{h_{k_{1:j-1}}} h_{k_j}\|^2$$
- at stage i: $\quad k_i = \arg\max_k \|P^\perp_{h_{k_{1:i-1}}} h_k\|^2$
- Introduce ϕ_i = angle between h_{k_i} and $h_{k_{1:i-1}}$ $\Rightarrow$ can write $\|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 = \|h_{k_i}\|^2 \sin^2\phi_i$.

MISO BF-style GUS

$$k_i \text{ maximizes } (\det(\text{diag}\{(H_i H_i^H)^{-1}\}))^{-1} =$$
$$\|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 \prod_{j=1}^{i-1} (\|P^\perp_{h_{k_{1:i-1}\setminus k_j}} h_{k_j}\|^2 - \frac{|h_{k_i}^H P^\perp_{h_{k_{1:i-1}\setminus k_j}} h_{k_j}|^2}{\|P^\perp_{h_{k_{1:i-1}\setminus k_j}} h_{k_i}\|^2})$$

For sufficiently large K, the BF-style user selection process will lead to the selection of channel vectors that are close to being mutually orthogonal. As a result we can write up to first order the contribution of stream i to the sum rate offset

$$\begin{aligned} &\|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 \prod_{j=1}^{i-1} \sin^2\phi_{ij} \approx \|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 \sin^2\phi_i \\ &= \|h_{k_i}\|^2 \sin^4\phi_i = \|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^4 / \|h_{k_i}\|^2 . \end{aligned} \tag{1}$$

DPC offset is $\|P^\perp_{h_{k_{1:i-1}}} h_{k_i}\|^2 = \|h_{k_i}\|^2 \sin^2\phi_i$ = certain compromise between max $\|h_{k_i}\|^2$ and min $\cos^2\phi_i$. In the case of BF, $\|h_{k_i}\|^2 \sin^4\phi_i$ leads to a similar compromise, but with more emphasis on orthogonality.

Role of Rx antennas?

- Different distributions of ZF between Tx and Rx give different ZF channel gains! If Rx ZF's k streams, hence Tx only has to ZF $N_t - 1 - k$ streams! So, number of possible solutions (assuming $d_k \equiv 1$):

$$\prod_{k=1}^{N_t}\left(\sum_{i=0}^{N_k-1}\frac{(N_t-1)!}{k!(N_t-1-k)!}\right)$$

  for each user, Rx can ZF k between 0 and $N_k - 1$ streams, to choose among $N_t - 1$.
  Explains non-convexity of MIMO SR at high SNR.
- ZF by Rx can alternatively be interpreted as IA by Tx (Rx adapts Rx-channel cascades to lie in reduced dimension subspace).
- SESAM (and all existing MIMO stream selection algorithms): assumes that all ZF is done by Tx only. Hence, Rx can be a MF, matched to channel-BF cascade.

Concluding Remarks MIMO BC

- Introduced new MISO BC BF-style GUS criterion/interpretation.
- Extension to MIMO BC with receiver design.
- Plenty of room for asymptotic analysis of transient regime of stream selection.
- Here, did not touch upon CSIT FB issues, user preselection schemes to reduce pool size etc.
- Joint Tx-Rx ZF (IA) provides more opportunities (but hence also larger search space and complexity).

Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

Past Instances of Non-circular Symbol Constellations

- SAIC for real signal constellations

  application to GSM: GMSK $\approx$ filtered modulated BPSK

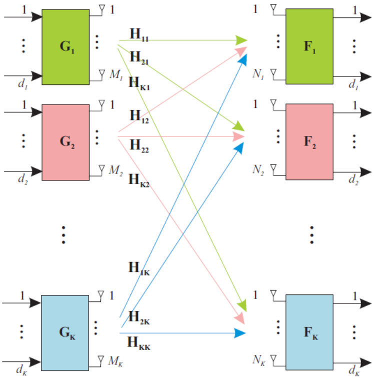
- space-time coding

  Alaouti scheme, = special case of linear dispersion space-time codes

- turbo receivers

  due to channel coding and bit to symbol mapping, reconstructed interfering symbols are typically non-circular (at order 2)

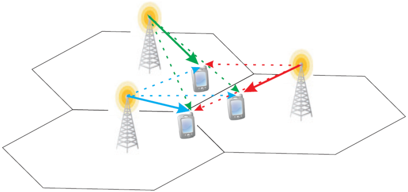
- Interference Alignment (IA) was introduced in [Cadambe,Jafar 2008]
- The objective of IA is to design the Tx beamforming matrices such that the interference at each non intended receiver lies in a common interference subspace
- If alignment is complete at the receiver simple Zero Forcing (ZF) can suppress interference and extract the desired signal
- In [SPAWC2010] we derive a set of interference alignment (IA) feasibility conditions for a K-link frequency-flat MIMO interference channel (IFC)
- $d = \sum_{k=1}^{K} d_k$



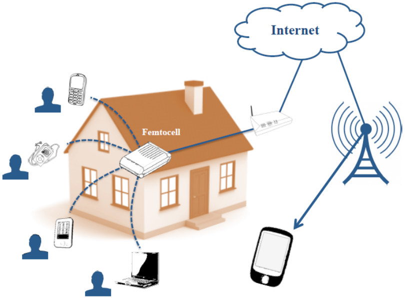
MIMO Interference Channel

Possible Application Scenarios

- Multi-cell cellular systems, modeling intercell interference.
  Difference from Network MIMO: no exchange of signals, "only" of channel impulse responses.



- Coexistence of cellular and femto-cells, especially when femtocells are considered part of the cellular solution.



Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

Why IA?

- The number of streams (degrees of freedom (dof)) appearing in a feasible IA scenario correspond to prelogs of feasible multi-user rate tuples in the multi-user rate region.
  Max Weighted Sum Rate (WSR) becomes IA at high SNR.
- Noisy IFC: interfering signals are not decoded but treated as (Gaussian) noise.
  Apparently enough for dof.
- Lots of recent work more generally on rate prelog regions: involves time sharing, use of fractional power.

Noisy MIMO IFC: Some State of the Art

- IA: alternating ZF algorithm [Jafar etal: globecom08],[Heath etal: icassp09].
- IA feasibility: - $K = 2$ MIMO: [JafarFakhereddin:IT07]
  - [Yetis,Jafar:T10], [Slock etal:eusipco09,ita10, spawc10]
  - 3xNxN, 3xMxN: [BreslerTse:arxiv11]
- max WSR: single stream/link
  - approximately: max SINR [Jafar etal: globecom08]
  - eigenvector interpretation of WSR gradient w.r.t. BF: starting [Honig,Utschick:asilo09]
  - added DA-style approach in [Honig,Utschick:allerton10]
- max WSR: multiple streams/link
  - [Slock etal:ita10] application of [Christensen etal:TW08] from MIMO BC
  - further refined in [Negro etal:allerton10], independently suggested use of DA, developed in [Negro etal:ita11]

IA as a Constrained Compressed SVD

- $F_k^H : \; d_k \times N_k, \;\; H_{ki} : \; N_k \times M_i, \;\; G_i : \; M_i \times d_i \qquad F^H H G =$

$$\begin{bmatrix} F_1^H & 0 & \cdots & 0 \\ 0 & F_2^H & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1K} \\ H_{21} & H_{22} & \cdots & H_{2K} \\ \vdots & & \ddots & \vdots \\ H_{K1} & H_{K2} & \cdots & H_{KK} \end{bmatrix} \begin{bmatrix} G_1 & 0 & \cdots & 0 \\ 0 & G_2 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & G_K \end{bmatrix} = \begin{bmatrix} F_1^H H_{11} G_1 & 0 & \cdots & 0 \\ 0 & F_2^H H_{22} G_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H H_{KK} G_K \end{bmatrix}$$

  F^H, G can be chosen to be unitary for IA

- per user vs per stream approaches:

  IA: can absorb the $d_k \times d_k$ $F_k^H H_{kk} G_k$ in either F_k^H (per stream LMMSE Rx) or G_k or both.

  WSR: can absorb unitary factors of SVD of $F_k^H H_{kk} G_k$ in F_k^H, G_k without loss in rate $\Rightarrow F^H H G =$ diagonal.

- To derive the existence conditions we consider the ZF conditions

$$\underbrace{\mathbf{F}_k^H}_{d_k \times N_k} \underbrace{\mathbf{H}_{kl}}_{N_k \times M_l} \underbrace{\mathbf{G}_l}_{M_l \times d_l} = \mathbf{0} \;, \quad \forall l \neq k$$

$$\text{rank}(\mathbf{F}_k^H \mathbf{H}_{kk} \mathbf{G}_k) = d_k \;, \quad \forall k \in \{1, 2, \ldots, K\}$$

- rank requirement $\Rightarrow$ SU MIMO condition: $\quad d_k \leq \min(M_k, N_k)$
- The total number of variables in $\mathbf{G}_k$ is
  $d_k M_k - d_k^2 = d_k(M_k - d_k)$
  Only the subspace of $\mathbf{G}_k$ counts, it is determined up to a $d_k \times d_k$ mixture matrix.
- The total number of variables in $\mathbf{F}_k^H$ is
  $d_k N_k - d_k^2 = d_k(N_k - d_k)$
  Only the subspace of $\mathbf{F}_k^H$ counts, it is determined up to a $d_k \times d_k$ mixture matrix.

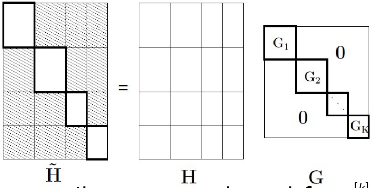
- A solution for the interference alignment problem can only exist if the total number of variables is greater than or equal to the total number of constraints i.e.,

$$\sum_{k=1}^{K} d_k(M_k - d_k) + \sum_{k=1}^{K} d_k(N_k - d_k) \geq \sum_{i\neq j=1}^{K} d_i\, d_j$$
$$\Rightarrow \;\; \sum_{k=1}^{K} d_k(M_k + N_k - 2d_k) \geq (\sum_{k=1}^{K} d_k)^2 - \sum_{k=1}^{K} d_k^2$$
$$\Rightarrow \;\; \sum_{k=1}^{K} d_k(M_k + N_k) \geq (\sum_{k=1}^{K} d_k)^2 + \sum_{k=1}^{K} d_k^2$$

- In the symmetric case: $d_k = d$, $M_k = M$, $N_k = N$:
$d \;\leq\; \frac{M+N}{K+1}$
- For the $K = 3$ user case $(M = N)$: $\quad d \;=\; \frac{M}{2}$.
With 3 parallel MIMO links, half of the (interference-free) resources are available!
However $d \leq \frac{1}{(K+1)/2} M < \frac{1}{2} M$ for $K > 3$.

Interference Alignment: Feasibility Conditions (3)

The main idea of our approach is to convert the alignment requirements at each RX into a rank condition of an associated interference matrix $\mathbf{H}_I^{[k]}=[\mathbf{H}_{k1}\mathbf{G}_1, \ldots \mathbf{H}_{k(k-1)}\mathbf{G}_{(k-1)}, \mathbf{H}_{k(k+1)}\mathbf{G}_{(k+1)}, \ldots \mathbf{H}_{kk}\mathbf{G}_K]$, that spans the interference subspace at the k-th RX (the shaded blocks in each block row).Thus the dimension of the Interference subspace must satisfy $\mathrm{rank}(\mathbf{H}_I^{[k]}) = r_I^{[k]} \leq N_k - d_k$



The equation above prescribes an upperbound for $r_I^{[k]}$ but the nature of the channel matrix (full rank) and the rank requirement of the BF specifies the following lower bound $r_I^{[k]} \geq \max_{l \neq k}(d_l - [M_l - N_k]_+)$. Imposing a rank $r_I^{[k]}$ on $\mathbf{H}_I^{[k]}$ implies imposing $(N_k - r_I^{[k]})(\sum_{\substack{l=1 \\ l \neq k}}^{K} d_l - r_I^{[k]})$ constraints at RX k. Enforcing the minimum number of constraints on the system implies to have maximum rank: $r_I^{[k]} \leq \min(d_{tot}, N_k) - d_k$

Interference Alignment: Feasibility Conditions (4)

- [BreslerTse:arxiv11]: counting equations and variables not the whole story!
- appears in very "rectangular" ($\neq$ square) MIMO systems
- example: $(M, N, d)^K = (4, 8, 3)^3$ MIMO IFC system comparing variables and ZF equations: $d = \frac{M+N}{K+1} = \frac{4+8}{3+1} = \frac{12}{4} = 3$ should be possible
- supportable interference subspace dim. $= N - d =$ 8-3 $= 5$
- however, the 2 interfering 8×4 cross channels generate 4-dimensional subspaces which in an 8-dimensional space do not intersect w.p. 1 !
- hence, the interfering 4×3 transmit filters cannot massage their 6-dimensional joint interference subspace into a 5-dimensional subspace!
- This issue is not captured by # variables vs # equations: $d = \frac{M+N}{K+1}$ only depends on $M + N$: $(5, 7, 3)^3$, $(6, 6, 3)^3$ work.

I and Q components: IA with Real Symbol Streams

- Using real signal constellations in place of complex constellations, transmission over a complex channel of any given dimension can be interpreted as transmission over a real channel of double the original dimensions (by treating the I and Q components as separate channels).
- This doubling of dimensions provides additional flexibility in achieving the total DoF available in the network.
- Split complex quantities in I and Q components:

$$\mathbf{H}_{ij} = \begin{bmatrix} \mathrm{Re}\{\mathbf{H}_{ij}\} & -\mathrm{Im}\{\mathbf{H}_{ij}\} \\ \mathrm{Im}\{\mathbf{H}_{ij}\} & \mathrm{Re}\{\mathbf{H}_{ij}\} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \mathrm{Re}\{\mathbf{x}\} \\ \mathrm{Im}\{\mathbf{x}\} \end{bmatrix}$$

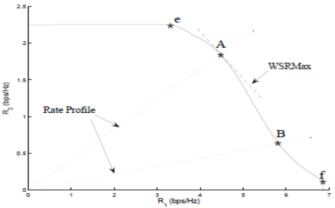
- Example: GMSK in GSM: was considered as wasting half of the resources, but in fact unknowingly anticipated interference treatment: 3 interfering GSM links can each support one GMSK signal without interference by proper joint Tx/Rx design! (SAIC: handles 1 interferer, requires only Rx design).

Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL) and Deterministic Annealing (DA)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

From IA to Optimized IFC's

- from Interference Alignment (=ZF) to max Sum Rate (SR) for the "Noisy IFC".
- to vary the point reached on the rate region boundary: SR $\rightarrow$ Weighted SR (WSR)
- problem: IFC rate region not convex $\Rightarrow$ multiple (local) optima for WSR (multiple boundary points with same tangent direction)
- solution of [CuiZhang:ita10]: WSR $\rightarrow$ max SR under rate profile constraint: $\frac{R_1}{\alpha_1} = \frac{R_2}{\alpha_2} = \cdots = \frac{R_K}{\alpha_K}$: $K-1$ constraints.
  Pro: explores systematically rate region boundary.
  Con: for a fixed rate profile,
  bad links drag down good links.
  $\Rightarrow$ stick to (W)SR
  (monitoring global opt issues).
  Note: multiple WSR solutions
  $\Leftrightarrow$ multiple IA solutions.



The received signal at the $k-$th receiver is:

$$\mathbf{y}_k = \mathbf{H}_{kk}\mathbf{G}_k\mathbf{x}_k + \sum_{\substack{l=1 \\ l\neq k}}^{K} \mathbf{H}_{kl}\mathbf{G}_l\mathbf{x}_l + \mathbf{n}_k$$

Introduce the interference plus noise covariance matrix at receiver k: $\mathbf{R}_{\bar{k}} = \mathbf{R}_{nn} + \sum_{l\neq k} \mathbf{H}_{kl}\mathbf{G}_l\mathbf{G}_l^H\mathbf{H}_{kl}^H$.
The WSR criterion is

$$\mathcal{R} = \sum_{k=1}^{K} \mathrm{u}_k \log\det(\mathbf{I} + \mathbf{G}_k^H\mathbf{H}_{kk}^H\mathbf{R}_{\bar{k}}^{-1}\mathbf{H}_{kk}\mathbf{G}_k) \tag{2}$$

$$\text{s.t.} \quad \mathrm{Tr}\{\mathbf{G}_k^H\mathbf{G}_k\} \leq P_k$$

This criterion is highly non convex in the Tx BFs $\mathbf{G}_k$.

WSR per stream

- Augmented WSR cost function: BFs $\mathbf{g}_{kn}$ plus Rx filters $\mathbf{f}_{kn}$ and weights w_{kn}:

$$\mathcal{O} = -\sum_{k=1}^{K} u_k \sum_{n=1}^{d_k} (-\ln(w_{kn}) + w_{kn}[(1-\mathbf{f}_{kn}^H \mathbf{H}_{kk}\mathbf{g}_{kn})(1-\mathbf{f}_{kn}^H \mathbf{H}_{kk}\mathbf{g}_{kn})^H$$

$$+\mathbf{f}_{kn}^H \underbrace{\Big(\mathbf{R}_{v_k} + \sum_{(im)\neq(kn)} \mathbf{H}_{ki}\mathbf{g}_{im}\mathbf{g}_{im}^H \mathbf{H}_{ki}^H\Big)}_{\mathbf{R}_{\overline{kn}}} \mathbf{f}_{kn}]) + \sum_{k=1}^{K} \lambda \Big(P_k - \sum_{n=1}^{d_k} \mathbf{g}_{kn}^H \mathbf{g}_{kn}\Big) \quad (3)$$

Alternating optimization $\Rightarrow$ quadratic or convex subproblems:

- Opt w_{kn} given $\mathbf{g}_{kn}$, $\mathbf{f}_{kn}$: $-\ln(w_{kn}) + w_{kn}e_{kn} \Rightarrow w_{kn} = e_{kn}^{-1}$
- Opt $\mathbf{f}_{kn}$ given w_{kn}, $\mathbf{g}_{kn}$: MMSE solution
- Opt $\mathbf{f}_{kn}$ given w_{kn}, $\mathbf{g}_{kn}$: MMSE-style solution (UL-DL duality)

Deterministic Annealing (DA) for global optimization

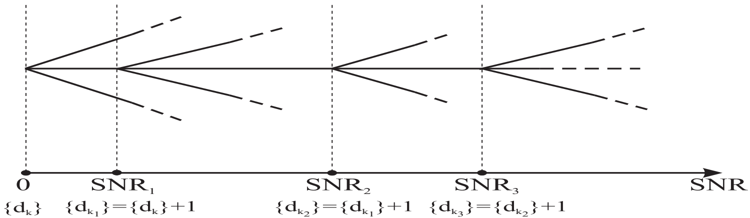
Deterministic Annealing

- In non-convex optimization, in order to avoid getting trapped in local minima/maxima, several heuristic approaches have been proposed.
- In analogy with the physical annealing process, Simulated Annealing (SA) has been proposed in optimization theory for non-convex problems.
- In SA the problem is optimized using a sequence of random moves, the magnitude of which depends on a temperature parameter that gets gradually cooled down.
- Deterministic annealing (DA) is inspired by the same principle but neither the cost function nor the initializations are random.
- The basic principle of DA is that the global optimum of the problem at the next temperature value is in the region of attraction of the solution of the problem at the previous temperature.

Deterministic Annealing (2)

- As in physical systems, also in an optimization problem it can happen that cooling down the temperature leads to phase transitions and hence several possible local optima may appear.
- A phase transition is characterized by a critical temperature that manifests itself by the Hessian of the cost function becoming singular. A stationary point evolves into a non stable point
- In our problem the cost function is the Weighted Sum Rate (WSR), a highly non convex function, and the annealing parameter is the noise variance, $t \propto \sigma^2$.
- Interestingly, in WSR maximization for the K-user MIMO IFC, we can associate phase transitions to the activation of an additional stream for a particular user.

- DA is about optimization of a cost function. In the process, we are tracking what we hope to be the gobal optimum.
- Tracking of extrema, the roots of the KKT conditions, is actually called a homotopy method.
- So DA, in going from one phase transition to the next while tracking the (appropriate) extremum, is a homotopy method.



Homotopy Methods

- Homotopy is used to find the roots of a non linear system of equations $\mathcal{F}(x) = 0$.
- Homotopy transformation is such that it starts from a trivial system $\mathcal{G}(x)$, with known solution, and it evolves towards the target system $\mathcal{F}(x)$ via continuous deformations according to the homotopy parameter t:

$$\mathcal{H}(x,t) = (1-t)\,\mathcal{G}(x) + t\mathcal{F}(x)$$

- Predicting the solution at the next value of $t^{(i+1)} = t^{(i)} + \Delta t$ is called Euler prediction phase
- Once we have a solution at $t^{(i+1)}$ it is possible to refine the estimate using a Newton correction phase for fixed t.
- A property of Homotopy continuation method for the solution of system of equation is that the number of solutions in the target system is at most equal to the number of solution in the trivial system
- The number of solutions along the trajectory remain constant

Homotopy Applied to IA

- Homotopy method can be applied to the IA problem. In particular we can define an homotopy deformation that starts from a trivial system and arrive to the target problem that we want to solve: K-user MIMO IFC: use instead of SNR as temperature, a scale factor for the channel excess singular values: $H_{ji} = \sum_{k=1}^{d} \sigma_{jik} \mathbf{u}_{jik} \mathbf{v}_{jik}^H + t \sum_{k=d+1} \sigma_{jik} \mathbf{u}_{jik} \mathbf{v}_{jik}^H$
- IA Jacobian still full rank if reduce rank(H_{ji}) to $\max(d_j, d_i)$.
- Finding a trivial starting system is easy: e.g. rank 1 channel system.

$$\sigma_{ji1} \mathbf{f}_j^H \mathbf{u}_{ji1} \mathbf{v}_{ji1}^H \mathbf{g}_i = 0$$

- After a coordination phase, where each user decides how is going to suppress a particular stream, the solution of the IA problem is easy.
- Once we define the starting point we describe the homotopy deformation that varies increasing the rank of the channel

Alternative Zero Forcing Approach to IA

- The interpretation of IA as joint transmit-receiver liner zero forcing can be easily understood considering the trivial rank-one MIMO IFC for the case of all $d_k = 1$.
- The channel between Tx i and Rx j can be represented using its SVD decomposition:

$$\mathbf{H}_{ji} = \sigma_{ji}\mathbf{u}_{ji}\mathbf{v}_{ji}^H$$

- The IA (ZF) condition for link $i - j$ can be written as

$$\sigma_{ji}\mathbf{f}_j^H\mathbf{u}_{ji}\mathbf{v}_{ji}^H\mathbf{g}_i = 0$$

- Two configurations are possible:

$$\mathbf{f}_j^H\mathbf{u}_{ji} = 0 \quad \text{or} \quad \mathbf{v}_{ji}^H\mathbf{g}_i = 0$$

- Either the Tx or the Rx suppresses one particular interfering stream
- Homotopy here not suggested for computing IA solutions, but for counting number of solutions.

Homotopy Applied to IA (3)

- Another possible description of the same problem can be done using the expansion of the IA conditions up to the first order:

$$(\mathbf{F}_j^H + d\mathbf{F}_j^H)(\mathbf{H}_{ji} + d\mathbf{H}_{ji})(\mathbf{G}_i + d\mathbf{G}_i) = 0$$

- Expanding the products above and considering only the terms up to first order we get

$$\mathbf{F}_j^H \mathbf{H}_{ji} d\mathbf{G}_i + d\mathbf{F}_j^H \mathbf{H}_{ji} \mathbf{G}_i = -\mathbf{F}_j^H d\mathbf{H}_{ji} \mathbf{G}_i$$

- IA: n joint bilinear equations. Overall number of solutions upper bounded by 2^n (again: $2 =$ either Tx or Rx side). However, equations structured $\Rightarrow$ number of solutions less. Consider e.g. $K = 3$, all $N_k = M_k = N$: in this case all solutions are known analytically and correspond to selecting $d = N/2$ out of N eigenvectors: $\dfrac{N!}{(\frac{N}{2}!)^2} \ll 2^n = 2^{(1.5\,N)^2 - 1.5\,N}$ solutions.

Homotopy applied to WSR backwards: decreasing SNR

- First consider reduced rank channels and apply DA to WSR for SNR decreasing from infinity (IA).
- IA $\Rightarrow$ every IA solution corresponds to a distribution of interference zeroing between Tx and Rx.
- Homotopy in decreasing SNR $\Rightarrow$ can interpret all WSR extrema as different distributions of Tx and Rx roles. Phase transitions $\Rightarrow$ some stream gets turned off $\Rightarrow$ reduces a number of local maxima.
- For a given SNR, homotopy in channel rank allows for a similar interpretation of WSR extrema in original system at any SNR.
- At high SNR, number of WSR extrema $=$ number of IA solutions.
  All WSR extrema (at any SNR) are local maxima: Hessian negative definite.

Algorithm 1 MWSR Algorithm for MIMO IFC

set (SNR) $t = 0$
Fix an initial set of precoding matrices $\mathbf{G}_k$, $\forall \in k = \{1, 2 \ldots K\}$
repeat
  set $t^{(i+1)} = t^{(i+1)} + \delta t$
  set $n = 0$
  Try augmenting $\mathbf{G}^{(t)}$ for one extra stream in any link.
  repeat
    $n = n + 1$
    Given $\mathbf{G}_k^{(n-1)}$ compute $\mathbf{F}_k^{(n)H}$ and $\mathbf{W}_k^{(n)}$, $\forall k$
    Given $\mathbf{F}_k^{(n)H}$, $\mathbf{W}_k^{(n)}$, compute $\mathbf{G}_k^{(n)}$ $\forall k$
  until convergence
until target SNR is reach

Back to Forward DA for WSR

- MWSR iterations for fixed SNR point are equivalent to Newton correction phase in homotopy methods
- MWSR iterations at increased SNR are equivalent to Euler prediction step in homotopy methods
- To find a good initialization at SNR$= 0$ we should study the WSR expansion up to second order in the SNR. The first order term only depends on the useful signal part hence MF are optimal. The second order contribution depends by both useful signal and interference
- When we iterate MWSR algorithm Tx and Rx filter are initialized as MF, after one iteration we optimize up to second order the WSR. This because at each iteration we find the MMSE filter with the other filter being MF.

Concluding Remarks DA

- Deterministic Annealing = Homotopy + Phase Transitions
- Annealing MIMO channel singularity at high SNR $\Rightarrow$ counting number of IA solutions and interpreting each as a different distribution of ZF roles between Tx's and Rx's
- Annealing down in SNR $\Rightarrow$ all WSR local maxima correspond 1-to-1 to continuations of IA solutions
- Annealing up in SNR for Max WSR:
  - global solution known at low SNR: one stream per link, with MF for Tx and Rx
  - alternating WSR maximization leads to simple subproblems for updates of Tx, Rx and weights
  - at any temperature increase, test for phase splitting = introduction of a new stream
  - its Tx and Rx filters are again (colored noise) MF, introduced Jammer WF algo for optimal power redistribution, alternating max algo tracks correct global maximum since WSR is convex up to second order in power variations
  - resulting DA algorithm is perhaps only known structured solution for finding the global Max WSR

Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

IFC CSI Acquisition Outline

- Centralized CSIT Acquisition
- Distributed CSIT Acquisition
- Channel Feedback & Output Feedback
- DoF optimization as a function of coherence time

State of the Art on MIMO IFC w Partial CSI

- MISO BC (MU-MISO DL) w CSIT acquisition: [KobayashiCaireJindal:IT10]
- TDD MISO BC w CSIT acquisition: [SalimSlock:JWCN11]
- Space-Time Coding for Analog Channel Feedback: [ChenSlock:isit08]
- [NegroShenoySlockGhauri: eusipco09]: TDD MIMO IFC IA iterative design via UL/DL duality and TDD reciprocity
- Interference Alignment with Analog CSI Feedback: [ElAyachHeath:Milcom10]
  Centralized approach: BS's are connected to a central unit gathering all CSI, performing BF computations and redistributing BF's.
- [Jafar:GLOBECOM10] Blind IA
- [MaddahAliTse:allerton10] Delayed CSIT approach for $K = 2$ MISO BC
- [VazeVaranasi:submIT] DoF region for MIMO IFC w FB
- [SuhTse:IT11] GDoF for IFC with feedback

Key Points

- Distributed approach: no other connectivity assumed than the UL/DL IFC. FB over reversed IFC
- "distributed" = "duplicated" (decentralized)
- A distributed approach does not have to be iterative. It can be done with a finite overhead (finite prelog loss) and finite SNR loss compared to full CSI, even as SNR $\rightarrow \infty$. Hence of interest compared to non-coherent (no/outdated CSIT) IFC approaches.
- Distributed ($\mathcal{O}(K^2)$) requires more FB than Centralized ($\mathcal{O}(K)$).
- centralized/decentralized IFC CSIT estimation (only exchange of data at temporal coherence variation rate), vs NW-MIMO/CoMP (exchange of data at symbol/sample rate)
- Multiple Rx antennas $\Rightarrow$ Rx training also crucial!
- TDD vs FDD, depends on distributed/centralized.
- Channel FB vs Output feedback (OFB)
- "Practical" scheme far from unique

Signal Structure w Partial CSI

- Perfect CSI: Rx signal at the $k-$th receiver :

$$\mathbf{y}_k = \sum_{i=1}^{K} \sum_{m=1}^{d_i} \mathbf{H}_{ki}\mathbf{g}_{i,m}\, x_{i,m} + \mathbf{v}_k$$

Estimate stream (k, n):

$$\widehat{x}_{k,n} = \mathbf{f}_{k,n}\mathbf{H}_{kk}\mathbf{g}_{k,n}\, x_{k,n} + \sum_{i=1}^{K} \sum_{m \neq n} \mathbf{f}_{k,n}\mathbf{H}_{ki}\mathbf{g}_{i,m}\, x_{i,m}\mathbf{f}_{k,n}\mathbf{v}_k$$

- Imperfect CSI: $\underbrace{\hat{\hat{\mathbf{f}}}_{k,n}}_{\text{est. at Rx } k} \quad \underbrace{\mathbf{H}_{ki}}_{\text{true}} \quad \underbrace{\hat{\mathbf{g}}_{i,m}}_{\text{est. at Tx } i}$
- signal of interest in direct link:

$$\hat{\hat{\mathbf{f}}}_{k,n}\mathbf{H}_{kk}\,\hat{\mathbf{g}}_{k,n} = \underbrace{\hat{\hat{\mathbf{f}}}_{k,n}\widehat{\mathbf{H}_{kk}\,\hat{\mathbf{g}}_{k,n}}}_{\text{known to Rx}} + \underbrace{\hat{\hat{\mathbf{f}}}_{k,n}\widetilde{\mathbf{H}_{kk}\,\hat{\mathbf{g}}_{k,n}}}_{\text{put in interf.}}$$

High SNR Rate Analysis

- Asymptote $\mathcal{R} = a\log(\rho) + b$ permits meaningful optimization for finite (but high) SNR, and may lead to more than minimal FB.
- At very high SNR ρ, only rate prelog a (dof) counts. Its maximization requires FB to be minimal (channel just identifiable).
- At moderate SNR, finding an optimal compromise between estimation overhead and channel quality will involve a properly adjusted overhead. However, the overhead issue is not the only reason for a possibly diminishing multiplexing gain a as SNR decreases, also reducing the number of streams $\{d_k\}$ may lead to a better compromise (as for full CSI).
- The rate offset b is already a non-trivial rate characteristic even in the full CSI case. b may increase as the number of streams decreases, due to reduced noise enhancement.

Unification Stationary & Block Fading

- Doppler Spectrum is bandlimited to $1/T$ ($1/D$ in figure)
- Nyquist's Theorem : downsampling possible with factor T
- Vectorize channel coefficients over T, matrix spectrum of rank 1, MIMO prediction error of rank 1.
- Hence channel coefficient evolution during current "coherence period" T is along a single basis vector, plus prediction from past.
- Block fading: basis vector = rectangular window and prediction from the past = 0

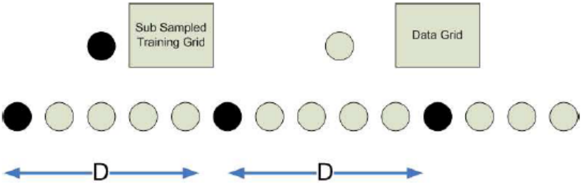
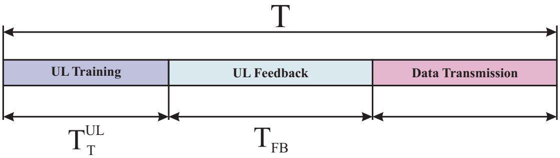
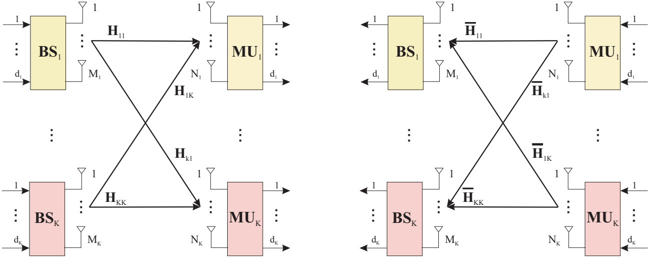


Figure 1: Subsampling Grid.

Centralized Approach

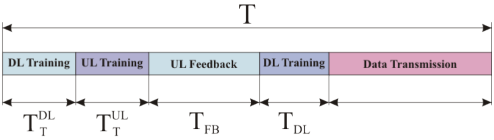


- Proposed by Heath [Milcom10,arxiv]
- The authors extrapolate the single antenna case, where only the estimate of the overall ch-BF gain and associated SINR is required
- In the MIMO IFC Rx not only needs to estimate the ch-BF cascade but also the I+N covariance matrix
- Not trivial. Training length similar as for the BF determination (order K) is required.
- Rate analysis of type 1.



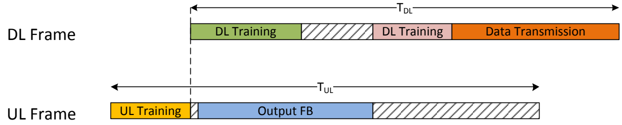
Downlink Uplink

- We Assume FDD transmission scheme
- Downlink channel matrix $\mathbf{H}_{ki}$ from BS_i to MU_k
- Uplink channel matrix $\overline{\mathbf{H}}_{ik}$ from MU_k to BS_i
- Analyze both centralized and distributed approaches.



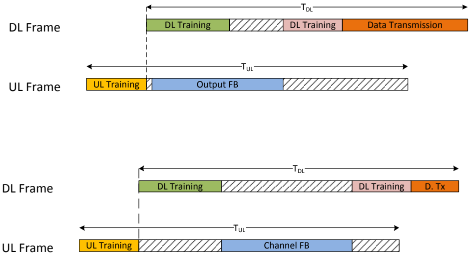
- We consider a block fading channel model with Coherence time interval T
- The general channel matrix $\mathbf{H}_{ik} \sim \mathcal{N}(0, \mathbf{I})$
- To acquire the necessary CSI at BS and MU side several training and feedback phases are necessary
- Hence a total overhead of T_{ovrhd} channel usage is dedicated to BS-MU signaling
- Only part of the time $T_{data} = T - T_{ovrhd}$ is dedicated to real data transmission

Output Feedback



- Each MU feeds back to all BS the noiseless version of its received signal using un-quantized feedback: Output FB (OFB).
- In FDD systems UL and DL transmission can take place at the same time
- T_{UL} represents the UL coherence Time
- T_{DL} represents the DL coherence Time
- OFB phase can start one time instant after the beginning of the DL training phase

Output Feedback



- Output FB allows us to reduce the overhead due to CSI exchange
- In channel FB each MU has to wait the end of the DL training phase before being able to FB DL channel estimates
- For easy of exposition we consider $M_i = N_t \; \forall i$, $N_i = N_r \; \forall i$ where $N_t \geq N_r$

Importance of CSIR

- CSIR is usually neglected
- Some schemes for arbitrary time-varying channels assume that Rxs know all channel matrices at all time: impossible to realize in practice
- An additional DL training phase is required to build the Rx filters

In the end: Sum Rate (high SNR)

- SR

$$\mathcal{R}^{PCSI} = \sum_{k,n} \underbrace{(1 - \frac{\sum T_i}{T})}_{\substack{\text{reduced data} \\ \text{channel uses}}} \ \ln(|\mathbf{f}_{kn}\mathbf{H}_{kk}\mathbf{g}_{kn}|^2 \ \underbrace{\rho/(1 + \sum_i \frac{b_{kni}}{T_i}))}_{\text{SNR loss}} \ ,$$

$T_i \geq T_{i,min}$

Assume $b_{kni} = b_i$ for what follows.

- Fixing $\sum_i T_i = T_{ovrhd}$, optimal $T_i = T_{ovrhd} \sqrt{b_i}/(\sum_i \sqrt{b_i})$.
- Optimizing over T_{ovrhd} now

$$T_{ovrhd} = \frac{\sqrt{T}\,(\sum_i \sqrt{b_i})}{\sqrt{\mathcal{R}^{FCSI}}}$$

Sum Rate at even higher SNR: DoF

- DoF optimization as function of Coherence Time
- full CSI: maximize $d_{tot} = \sum_k d_k$
- with CSI acquisition, larger MIMO systems need more training/FB, hence at short coherence times, the MIMO dimensions and number of streams may need to be reduced
- symmetric systems $(N, N, d)^K$: $\quad N \geq \frac{d}{2}(K+1)$
- DL time overhead for CFB as:

$$T_{ovrhd} = T_T^{DL} + T_{FB} + T_{DL} = \begin{cases} dK(K+2) & \text{(Centr.)} \\ \frac{Kd}{2}((K+1)^2+2) & \text{(Distr.)} \end{cases}$$

- DoF term in Sum Rate:

$$\max_d J(d) = \max_d (1 - \frac{T_{ovrhd}}{T}) Kd \log SNR$$

DoF Optimization (2)

- $\frac{\partial J}{\partial d} = 0 \;\Rightarrow\; d^* = \begin{cases} \frac{T}{2K(K+2)} & \text{(Centr.)} \\ \frac{T}{K[(K+1)^2+2]} & \text{(Distr.)} \end{cases}$
- Hence

$$d \leq \min\left\{d^*, \frac{2N}{K+1}\right\}$$

$$\Rightarrow \quad d = \begin{cases} \min\left\{\frac{T}{2K(K+2)}, \frac{2N}{K+1}\right\} & \text{(Centr.)} \\ \min\left\{\frac{T}{K[(K+1)^2+2]}, \frac{2N}{K+1}\right\} & \text{(Distr.)} \end{cases}$$

DoF Optimization (3)

- if

$$T \geq \frac{4NK(K+2)}{K+1} = 2T_{ovrhd}$$

then number of streams $d = \frac{2N}{K+1}$, kept at its maximum.

- otherwise shrink d and number of active antennas

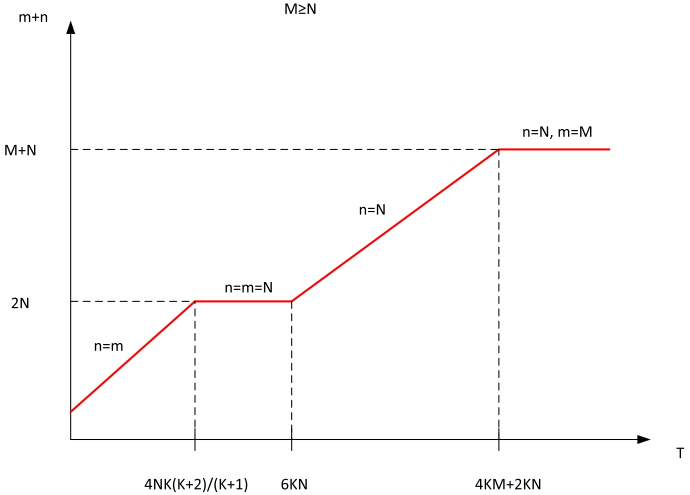
$$n \geq \frac{d^*(K+1)}{2}$$

with a consequent reduction of the time overhead for CSI acquisition.

- similar analysis for output feedback

$(M, N, d)^K$ case

- Assuming $M \geq N$.
- Evolution of number of active transmit m and receive n antennas as a function of coherence time T.



TDD vs FDD

- Usually TDD transmission scheme is used to simplify the DL CSI acquisition at the BS side
- BS_k learns the DL channel $\mathbf{H}_{ik}$, $\forall i$ through reciprocity
- MU_i do not need to feedback $\mathbf{H}_{ik}$ to BS_k but this channel is required at $BS_{j \neq k}$
- In Distributed Processing reciprocity does NOT help in reducing channel feedback overhead $\Longrightarrow$ TDD almost equivalent to FDD
- In Centralized Processing reciprocity makes channel feedback NOT required

Further Optimizing DoF

- data Tx stage (as good as perfect CSI):
  - Can FB increase DoF with perfect CSIT? According to [HuangJafar:IT09] and [VazeVaranasi:ITsubm11] NO for $K = 2$ MIMO IFC; $K > 2$ is OPEN.
  - If not in general, then use of OFB is mainly (only) for CSI acquisition, not for augmenting DoF in presence of CSIT
- CSI acquisition stages:
  - Optimize number of streams/number of active antennas for small T: if less channel to learn then more time to Tx data, even if on reduced number of streams
  - Instead of going from $K = 1$ to full K immediately, could gradually increase number of interfering links (and their CSI acquisition) from 1 to K.
  - When T gets too short: delayed CSIT approaches. Optimal combination: do delayed CSIT during training dead times.
  - A single (the largest) MIMO link can start transmitting right away w/o CSIT (possibly w/o CSIR also).

Perspectives

- When (analog) channel FB is of extended (non-minimal) duration, BF's can get computed and some DL transmission could start while FB is still going on. No need to wait until all CSI is gathered before transmission can get started.
- rate constants: partial CSI Tx/Rx design, diversity issues (optimized IA)
- optimization of training duration/power
- can OFB increase dof w perfect CSIT for $K \geq 3$?
- need to handle CSIR also in delayed CSIT approaches
- users with different coherence times
- full duplex operation (2-way communications)
- minimum reciprocity:
  coherence times equal on UL and DL, feasible dof same on UL and DL
- real IFC system: doubly selective

Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

Recent History of Localization Projects at EURECOM CM

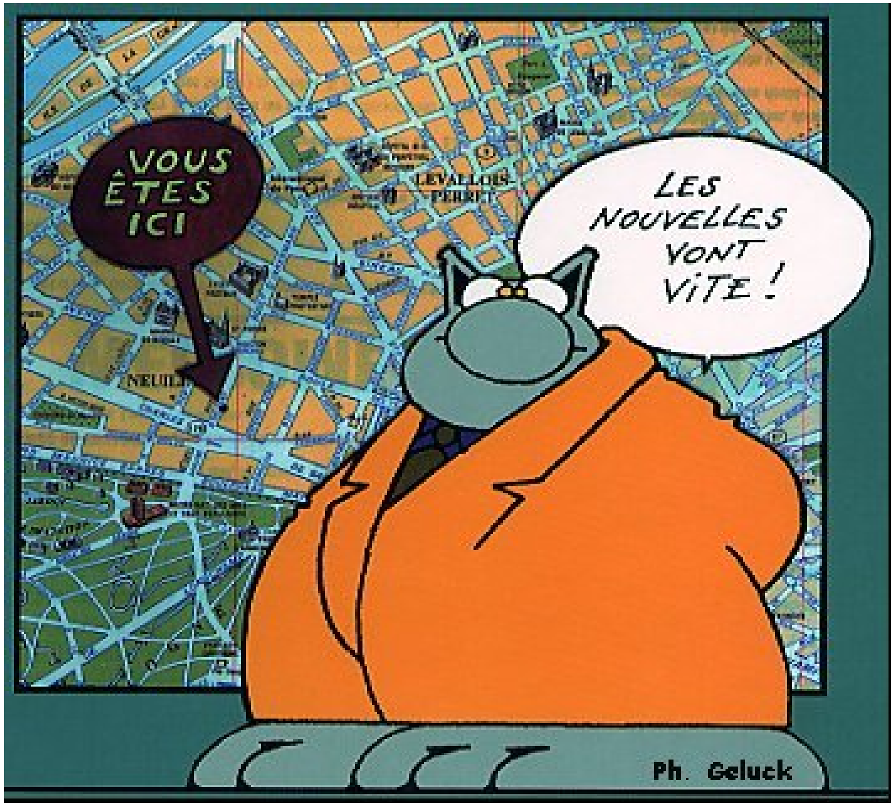
- ANR SEMAFOR
  - Prime: Thales
  - Radio surveillance, identification & localization of (rogue) transmitters

- FP7 WHERE : wireless network based localization
  - Localization: complement GPS for urban canyon & indoor, work on integrating information from heterogeneous sources
  - Initial work on location aided communications

- FP7 WHERE2:
  - Localization: cooperative techniques, accelerometers, map information
  - EU: WHERE2 = demo of WHERE => operators OTE (Greece), PTIN (Portugal), demo of macro-femto handovers etc.

Wireless Network based Localization now in LTE-A

- Enhanced Cell Id
  - Cell Id + RSS (received signal strength)
- O-TDoA (Observed Time Difference of Arrival)
- AoA (Angle of Arrival at BS)

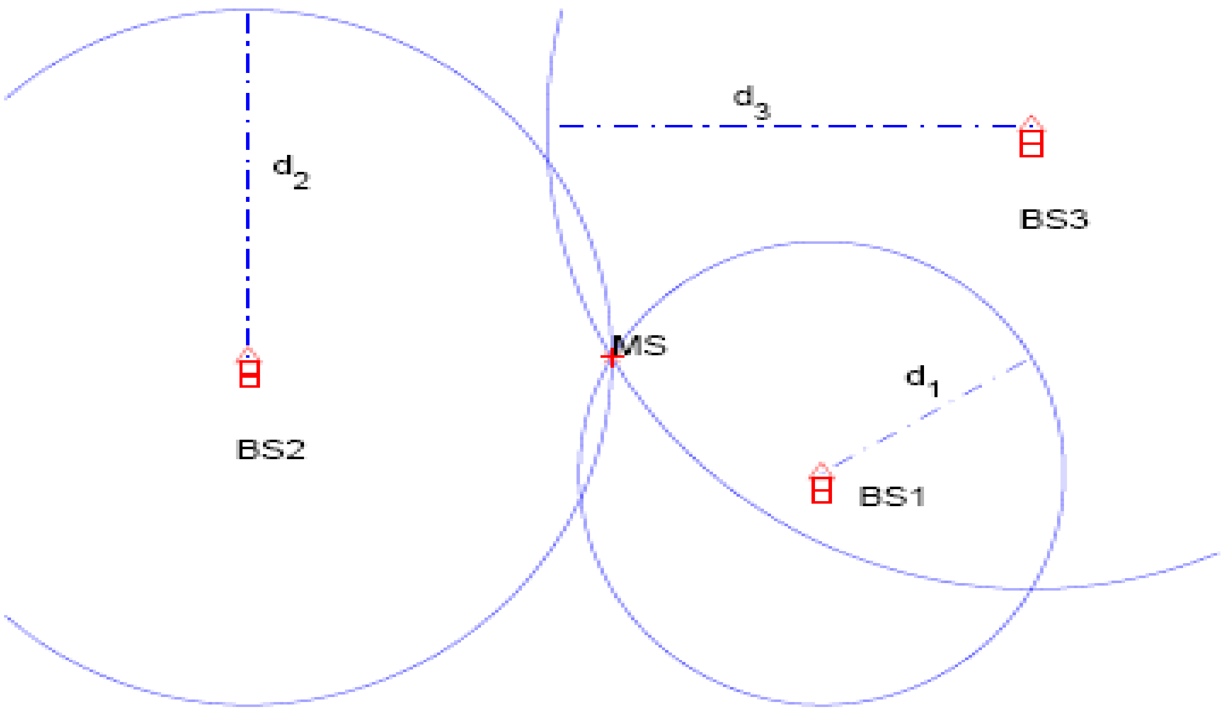


Fingerprinting based Localization Methods

- Fingerprinting based positioning: **fingerprint = RSS usually** (Received Signal Strength)
- Location fingerprinting (LF) (introduced by U.S. Wireless Corp. of San Ramon, Calif.) relies on signal structure characteristics
- LF may exploit the multipath nature of the channel hence the NLOS conditions
- By using multipath propagation pattern, the **LF exploits a signature unique to a given location**
- The position of the mobile is determined by matching measured signal characteristics from the BS-MT link to an entry of the database
- The location corresponding to the highest match of the database entry is considered as the location of the mobile
- For LF, it may be enough to have only one BS-MT link (multiple BSs are not required) to determine the location of the mobile
- Also LF is classified among Direct Location Estimation (DLE) techniques
- Power Delay Profile Fingerprinting (PDP-F) works by matching with the position dependent Power Delay Profiles (PDPs) in a database

Classical ToA based Positioning

LOS: 1 ToA to 1 BS => 1 circle



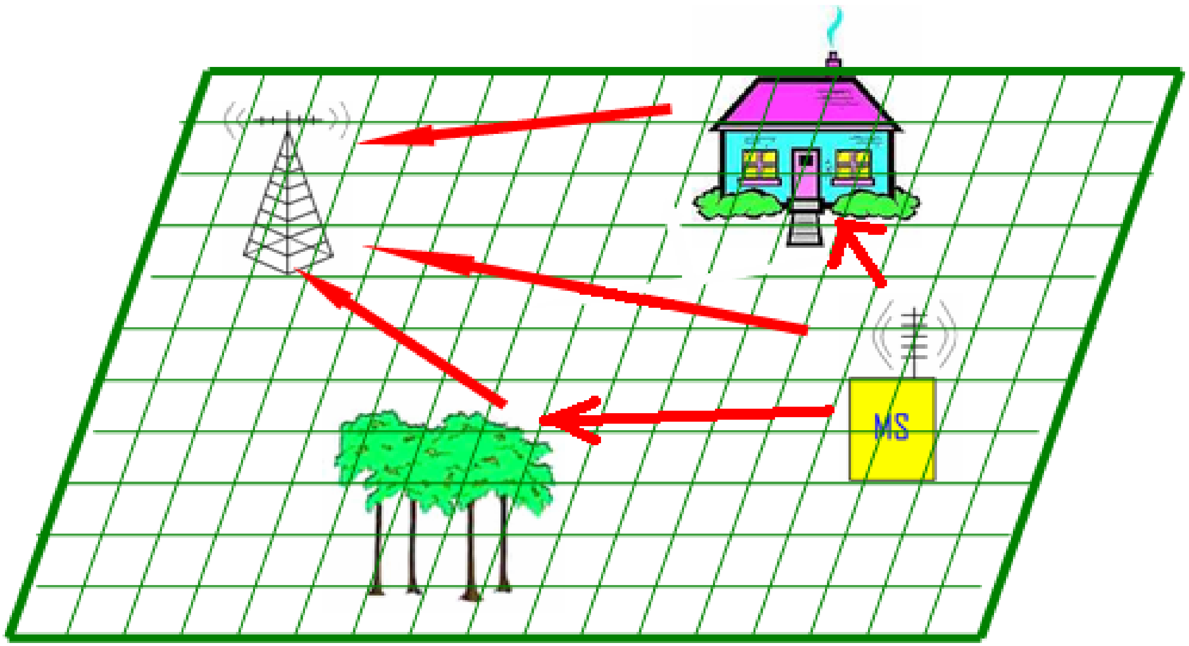
Localization in Multipath Environments

Multipath: Curse or Blessing?

- **Curse:**
  - LOS: hampers estimation of LOS ToA and other parameters
  - NLOS: introduces bias on LOS ToA
- **Blessing:**
  - richer LDP information $\Rightarrow$ may allow single anchor based localization!
  - each path providing as much info as a separate anchor in LOS only case.

Mobile Terminal Localization

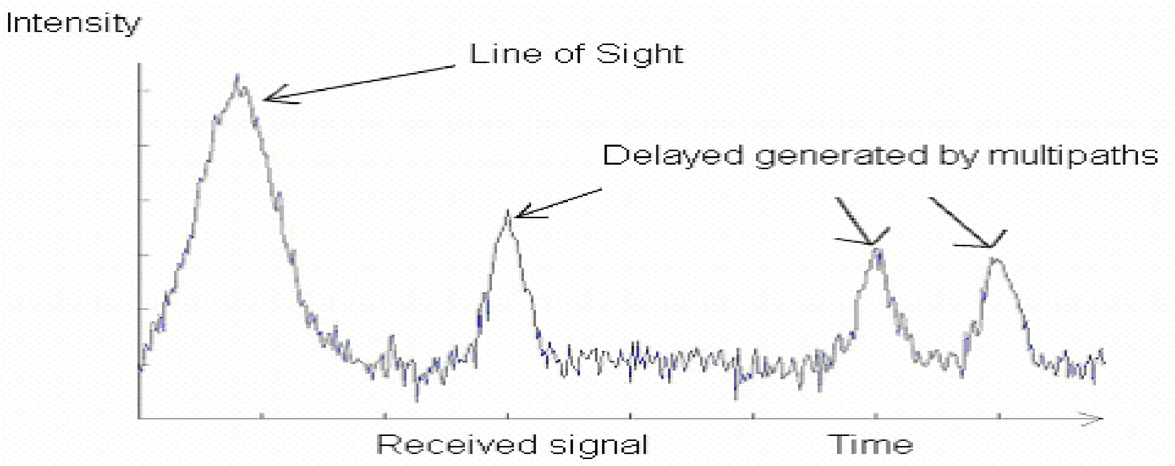
- Exploit (instead of suffer) Multipath Diversity
  & be able to work in NLOS case
  Power Delay Profile Fingerprinting



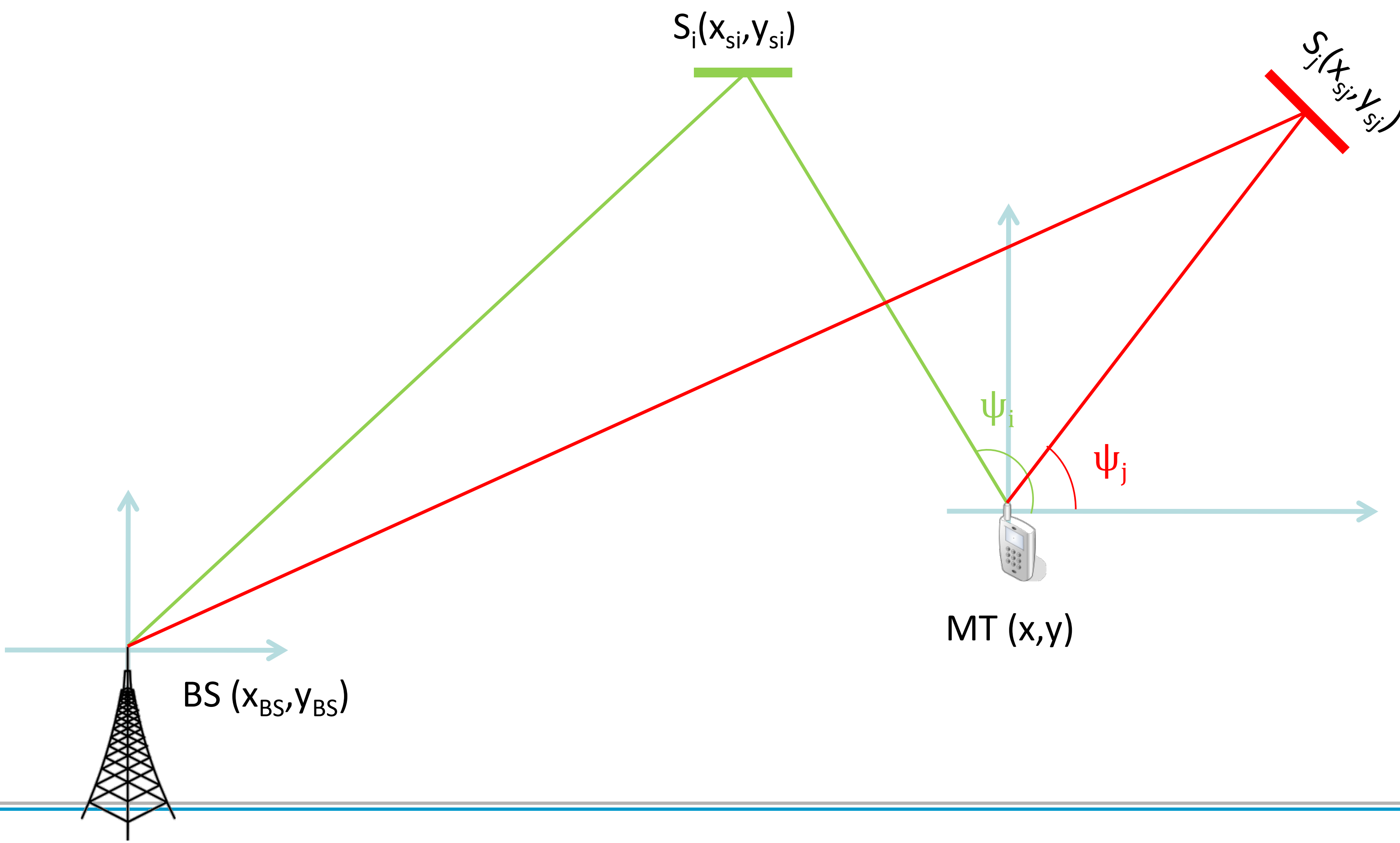
PDP Fingerprinting (with time reference)

- **In the case of a synchronous network,**
  - The first peak of the measured PDP determines the ToA of the line-of-sight path
  - The PDP exploits the ToA of the reflected paths

⇨ Multi-point signal reception is not required



BS-MT Geometry



PEP Analysis

- The Gaussian loglikelihood for T i.i.d. channel estimates $\widehat{\mathbf{h}}_i$ at a given position with channel estimate covariance matrix $\mathbf{C}_{\widehat{\mathbf{h}}_i \widehat{\mathbf{h}}_i}$ is

$$\mathcal{LL} \propto -\ln\left(\det\left(\mathbf{C}_{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}\right)\right) - tr\left(\widehat{\mathbf{C}}\,\mathbf{C}^{-1}_{\widehat{\mathbf{h}}\widehat{\mathbf{h}}}\right)$$

where $\widehat{\mathbf{C}} = \frac{1}{T}\sum_{i=1}^{T} \widehat{\mathbf{h}}_i \widehat{\mathbf{h}}_i^H$ is the sample covariance matrix.

- So we have an error when the loglikelihood for a false position is larger than that for the true position:

$$\text{PEP} = \Pr\{\mathcal{LL}_T < \mathcal{LL}_F\}.$$

Hence,

$$\text{PEP} = \Pr\{tr(\widehat{\mathbf{C}}\,\mathbf{A}) < \ln\det(\mathbf{C}_T \mathbf{C}_F^{-1})\}$$

where $\mathbf{A} = \mathbf{C}_F^{-1} - \mathbf{C}_T^{-1}$.

so that we get at high SNR ρ

$$\text{PEP} = \text{E}_{\mathbf{h}}\, Q\left(\sqrt{\frac{T}{2\sigma_v^2}}\, \|\mathbf{a}_1\|\right) .$$

Now exploiting the Gaussian distribution of $\mathbf{a}_1$, this leads to

$$\text{PEP} \ = \ \frac{c}{\det(T\,\mathbf{D}_1)\ \rho^{N_p-N_c}}$$

(for some constant c) which exhibits the well-known diversity behavior of probability of error for digital communication over fading channels.

- Again, $N_p - N_c$ are the number of path delays in which the mistaken PDP differes from the true PDP. Clearly, the richer the multipath, the smaller the PEP is likely to be, esp. at high SNR.

Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

Some generalities on position based information that can be exploited

- **Slow fading channel characteristics** of various links:
  - LOS/NLOS
  - Attenuation
  - Delay spread, frequency selectivity
  - Angular spreads, MIMO channel characteristics
  - Speed, direction of movement, acceleration (predictibility of movement)

  Some of these aspects may require the use of databases (containing these characteristics as a function of position), compatible with cognitive radio setting.

  Compared to feedback based approaches: some of these characteristics can not easily be determined from isolated channel estimates, or not predicted at all (e.g. slow fading prediction).

- Can **not** be inferred on basis of position: **fast fading** state, instantaneous channel response

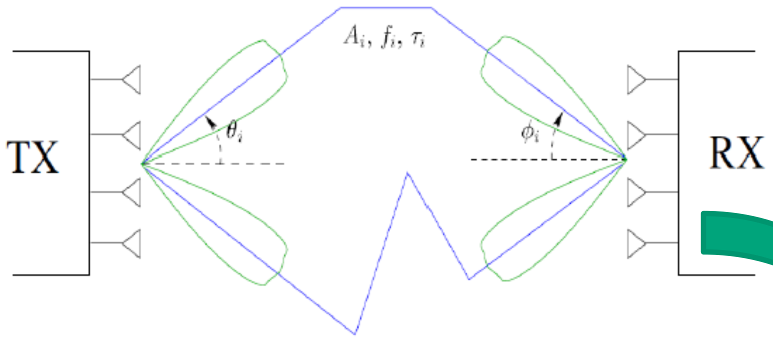
location aided communications : thoughts at the end of WHERE

Impact of position information on channel estimation.

- On the receiver side:
  - position information can lead to information about the channel statistics via a database, which can be used to improve channel estimation. This could be compared to learning of the channel statistics from previous channel estimates (which is not possible though in short packet mode!)

- On the transmitter side:
  - one could consider adapting the (OFDM) Cyclic Prefix (CP) and pilot structure on the basis of environment parameters. This would lead to minimized overhead and would avoid to design for the worst case.

- MCSS’11 Artist4G session, Nokia-Siemens presentation: database of channel impulse responses themselves!! (stable over 40’ in some measurements), to overcome problem of delay in channel feedback (FB). Considered combined FB + location aided approaches as realistic.



Location and database aided channel estimation/prediction

$$\underline{\mathbf{A}}[k] = \mathbf{F}[\mathbf{p}[k],k]\,\underline{\mathbf{A}}[k-1] + \mathbf{Q}^{\frac{1}{2}}(\mathbf{p}[k],k)\,\mathbf{w}[k]$$

$$\hat{\underline{\mathbf{h}}}[k] = \mathbf{H}[\mathbf{p}[k],k]\,\underline{\mathbf{A}}[k] + \tilde{\underline{\mathbf{h}}}[k]$$

Integrated
position tracking (PKF) and
channel tracking (KF)

$\underline{\mathbf{A}}[k]$: fast fading complex path gains

$\mathbf{H}[\mathbf{p}[k],k], \mathbf{F}[\mathbf{p}[k],k], \mathbf{Q}[\mathbf{p}[k],k]$: slow fading position-dependent parameters

For fast fading channel estimation and short-term prediction, the channel dynamics components H,F,Q can be

(1) learned from consecutive channel estimates, but knowledge will often come a bit late in this way, or

(2) determined from position information + database, leading to instantaneous knowledge & extended (short-term) channel prediction range.

(2) Allows furthermore longer-term prediction, but of channel statistics only.

Future work: what if database content is limited?

location aided communications : thoughts at the end of WHERE

Position based Adaptive Modulation and Coding (AMC).

- (Position information leads to) Environment information which in turn leads to information on the channel diversity structure, on the channel frequency selectivity and would allow to adapt frequency allocation/interleaving.

- Information on the MIMO channel richness allows to adapt the spatial multiplexing and the (linear) space-time coding.

- Information on the mobility provides temporal diversity information, which can be used to adapt interleaving in time.

- All these adaptations can take into account channel non-Rayleigh aspects (e.g. LOS/NLOS, LOS leads to reduced or no fading).

location aided communications : thoughts at the end of WHERE

Location aided Resource Allocation.

- A transversal aspect is also that location tracking can lead to **location prediction**. This leads in turn to **slow fading predictibility** (and not just fast fading prediction, which can in principle be done also from past channel response estimates). Another aspect is that **user selection** (multi-user diversity) potentially leads to an **explosion of CSIT requirements** and associated overhead. Location based covariance CSIT might offer a (partial) solution.
- **Single user aspects**:

  the environment (profile) and mobility information can be exploited to aid (time,frequency) allocation in OFDMA.
- **Multi-user MIMO**:

  Use environment information to preselect users, to limit channel feedback to a reduced set of preselected users. The user preselection can e.g. involve: users with similar RSS, users with rank 1 MIMO channels (close to LOS), ...
- **Multicell aspects** (interference coordination) or for **Cognitive Radio** (interference from secondary systems to primary systems):

  the interference level can be predicted from position based information.

Location-aided Multi-cell Processing

- One key observation: **user selection**:
  - Requires enormous overhead: to know instantaneous channels (without delay) for a large pool of users to schedule from (so from many more users than the ones actually scheduled)
  - How far can we go with partial CSIT, typically covariance CSIT (could be location based)? E.g. spatial subspaces (multi-antenna), or PDP holes (even single antenna, see further).

- **Single antenna**: power coordination, can fairly easily be done location-aided.

- **Multi-antenna** techniques:

  Require downlink channel knowledge, in principle of all channels at all transmitters
  - Multi-Cell coordinated beamforming (MIMO/MISO Interference Channel (IFC))
  - Network MIMO (= CoMP): requires furthermore signal distribution over the BS's

Location based SDMA

Multi-User (MU)-MIMO in LTE-A is struggling: needs much better channel information at the transmitter (CSIT) than single-user MIMO! Some possible solutions are:

- Increase CSI feedback enormously (possibly using analog feedback transmission).
- Exploit channel reciprocity in TDD (electronics calibration issue though).
- Limit MU-MIMO to Narrow Angle-of-Departure Aperture (NADA) users and extract essential CSIT from position information (or from DoA estimation?)

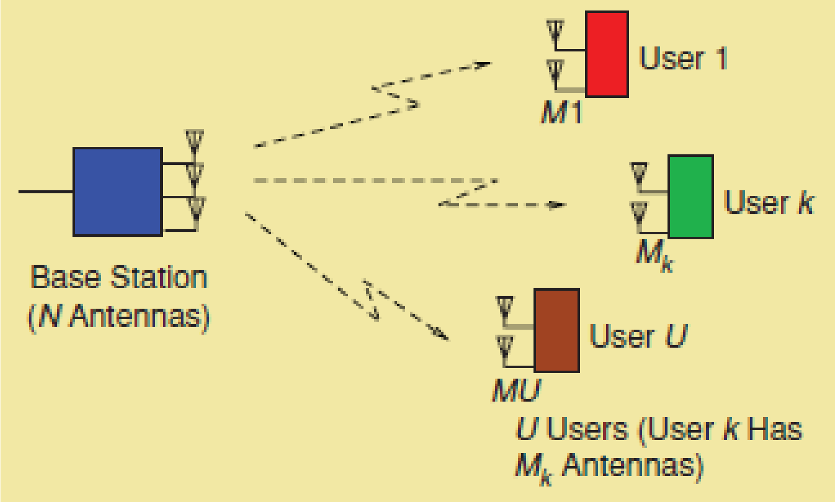
What is new since SDMA of nineties: user selection

- allows to select NADA users => position based approach meaningful
- allows to select quasi-orthogonal users => (ZF) beam-forming close to optimal

$$\mathrm{LB} = \sum_{k \in S} \mathbb{E}_{\mathbf{g}_k} \log\left(1 + \frac{1-\sigma_h^2}{1+p|S|\sigma_h^2}\, p|\mathbf{g}_k^H \bar{\mathbf{v}}_k|^2\right)$$

WHERE: Sum rate (LB) degradation analysis due to imperfect CSIT and applied to:

- Location based: positioning errors and Rice factor (non-LOS)
- LTE-style approach: channel feedback quantization errors



T3.1.2/3.1.3 – Position Aided SDMA (= MU-MIMO) (EUR)

- Context & Problem:
  - Spatial Division Multiple Access = Multi-User MIMO
  - Use location/multipath information to select users and reduce CSIT feedback
  - SDMA (90's) unsuccessful: no user selection (user orthogonality), disregard of multipath

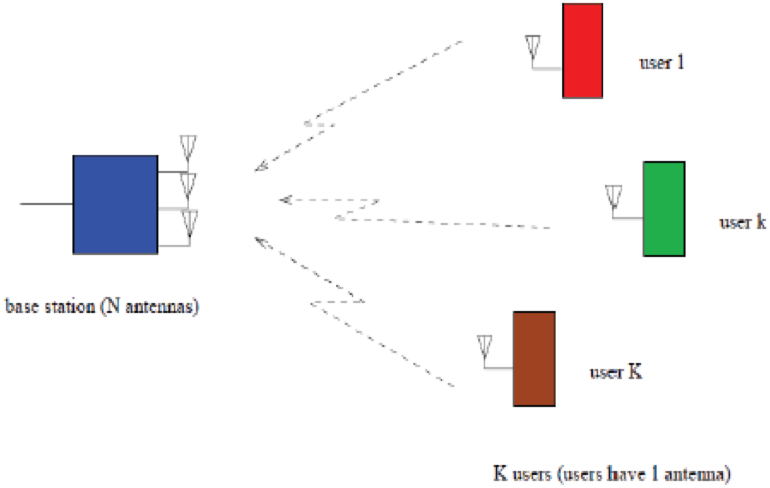
- Principle & Novelty:
  - Only select NADA users (Narrow Angle of Departure Aperture) , that go well together
  - Location aided vs DoA based:
    - More precision / better sensitivity can be obtained if BS antennas are spaced apart further (while satisfying far field assumption, or at least narrowband assumption).
    - Leads to ambiguities in DoA, which can be resolved by location info.

NADA MIMO channel model

$$H = \sum_i h_r(\theta_i) h_t^T(\phi_i) = B\, A^T\ ,\quad A = \left[h_t(\phi)\ \dot{h}_t(\phi) \right]$$

narrow AoD spread: $\phi_i = \phi + \Delta\phi_i$

$$h_t(\phi_i) \approx h_t(\phi) + \Delta\phi_i\, \dot{h}_t(\phi)$$



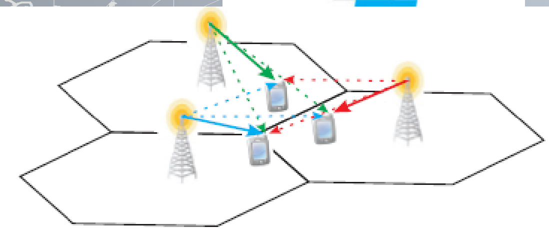
Sequence of Multicell COM paradigms

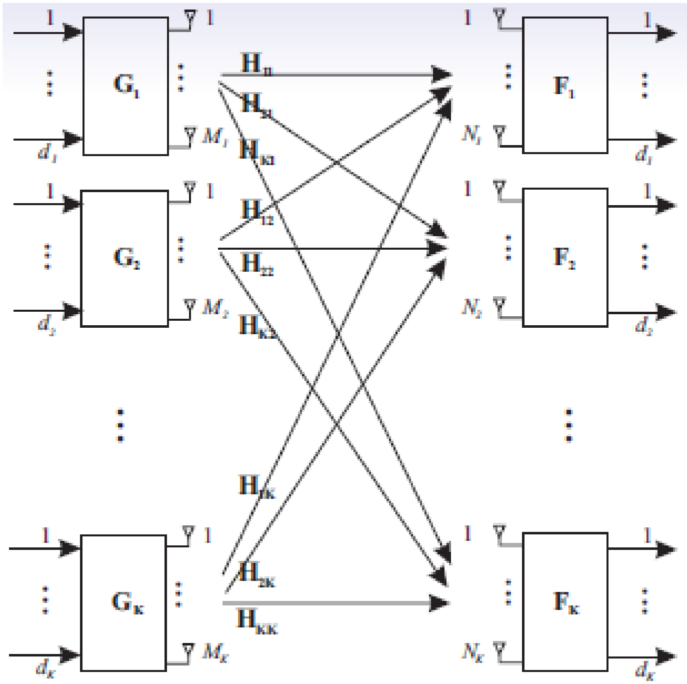
- single-cell MU-MIMO (cell center)
  - This is single-cell, but the opportunities of multiple Rx antennas in MU-MIMO are still not well understood. Here we focus on the use of Rx antennas to handle the multipath difference between NADA (angular spread) and LOS, or general NLOS, whereas the Tx handles the nominal directions (LOS). [WCNC12-W4] (Artist4G/WHERE2 SS)
- MU-MIMO with intercell interference (cell edge)
  - Like MIMO Interference Channel (IFC) but here only the Rx antennas are used to handle intercell interference. Consider LOS and then NADA.
- MIMO IFC
  - MISO IFC: maxmin SINR design via UL/DL duality [pimrc11]
  - MIMO IFC
    - [ita11]: ZF/IA feasibility, globally optimal maximum sum rate design via deterministic annealing
    - [asilomar11,ita12]: centralized and distributed CSI training and feedback (over the reverse IFC)
    - Currently exploring:
      - Spatial aspects: specificities for LOS, extension to NADA (location based (N)LOS info, spatial signature?)
      - Recent approaches // blind IA (BIA): interweave PDP polyphase components

Position Aided Multi-Cell

- Context & Problem:
  - MIMO Interference Channel (IFC): multi-Tx coordination of BF without exchanging signals (=CoMP, NW-MIMO)
  - Model for Multi-Cell, macro-femto, etc.
- Principle & Novelty:
  - Joint Tx/Rx design plagued by numerous local optima (proposed deterministic annealing for guaranteeing global optimum)
- Recent Achievements:
  - In LOS case (all rank-1 MIMO): role of Tx and Rx much clearer: repartition of ZF roles between Tx's and Rx's
  - Furthermore: design of a Tx or Rx only requires CSI of channels connected to it (local).
- Perspectives:
  - LOS -> NADA, mixed CSIT
  - Not all interfering links equally important:

    From “Generalized DoF” or “tier 1 interferers” to location (distance & propagation) dependent interference strengths





LOS: $\mathbf{H}_{ji} = \sigma_{ji}\mathbf{u}_{ji}\mathbf{v}_{ji}^H$

$\mathbf{f}_j^H\mathbf{u}_{ji}\mathbf{v}_{ji}^H\mathbf{g}_i = 0$ => $\mathbf{f}_j^H\mathbf{u}_{ji} = 0$ or $\mathbf{v}_{ji}^H\mathbf{g}_i = 0$

$$\begin{bmatrix} F_1^H & 0 & \cdots & 0 \\ 0 & F_2^H & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} & \cdots & H_{1K} \\ H_{21} & H_{22} & \cdots & H_{2K} \\ \vdots & & \ddots & \vdots \\ H_{K1} & H_{K2} & \cdots & H_{KK} \end{bmatrix} \begin{bmatrix} G_1 & 0 & \cdots & 0 \\ 0 & G_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & G_K \end{bmatrix} = \begin{bmatrix} F_1^H H_{11} G_1 & 0 & \cdots & 0 \\ 0 & F_2^H H_{22} G_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & F_K^H H_{KK} G_K \end{bmatrix}$$

Genuine location aided / Ongoing work

- Location aided beamforming (spatial signature): complementary to DoA based?
- Location aided covariance CSIT vs CSIT feedback (FB) when CSI varies too fast.

  Genuine location aided potential: trajectory based spatial signature prediction. Requires simulation/analytical comparison/combination with
  - DoA estimation and prediction?
  - or feedback based spatial signature estimation plus its prediction over realistic FB delays.
- Location aided covariance CSIT vs CSIT feedback (FB) when too many users to schedule: reducing signaling (FB) overhead.
- NADA (Narrow Angle of Departure Aperture) for IBC (Interfering Broadcast Channel)
  - LOS condition important for location aided spatial signature determination (to a lesser extent also for DoA based)
  - Position info for knowing which interference is strong, which needs to be coordinated and which can be ignored (appears in the noise)?
  - Mixed CSIT: channel feedback based for some users, position based for others.

Location Aided Cognitive Radio Networks

- Single antenna case (primaries)
  - In this case we can only play with the Tx powers.
  - There is an obvious use of location information to determine the secondary2primary attenuation.
  - We elaborate here [pimrc11] on the design of underlay secondary MISO interference channels (interfering secondary MISO links with interference power constraints at single antenna primary receivers).

- Multi-antenna case:
  - We propose a new terminology for spatial underlay/overlay/interweave [CogART11panel]
  - Remark on ASA (Qualcomm-NokiaSiemens): moves strongly to overlay
  - [CogART11] spatial interweave MIMO IFC.
  - CROWN demo: TDD spatial interweave RF calibration

Multiple Antennas in CR: a terminology proposal

- **P = primary, S = secondary**
- **spatial overlay: MIMO Interference Channel**
  - overlay: P and S collaborate
  - exploit multiple antennas and coordinate beamforming to achieve parallel interference-free channels
- **spatial underlay**
  - P Rxs with multiple antennas (P Rx active to suppress interference from S)
  - allows interference subspace of max dimension =

    excess of # P Rx antennas - # P streams
  - consider interference to P Rxs OK as long as interference subspace dimension does not exceed a max; requires active interference ZF by P
- **spatial interweave**
  - # S streams $\leq N^{S}_{Tx} - N^{P}_{Rx}$
  - possible if have excess of S Tx antennas over P Rx antennas
  - S BF nulls to P Rx antennas w/o P cooperation:

    location based (if LOS), reciprocity based in TDD

Underlay Cognitive Radio MISO IFC

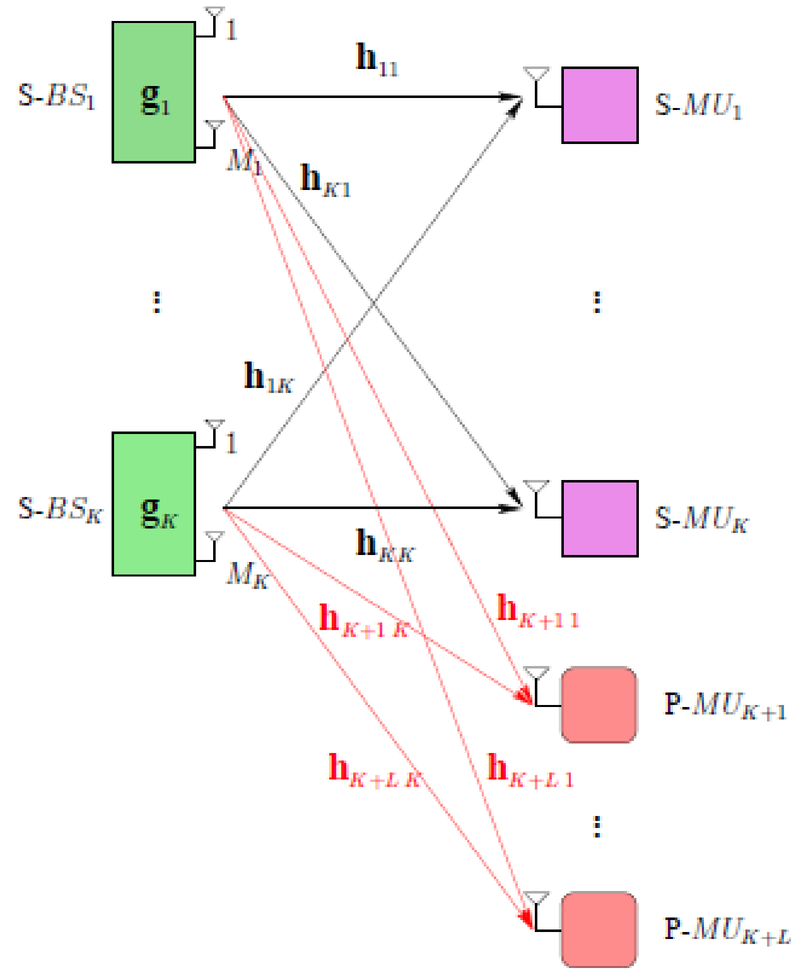
- Context & Problem:

  Allow a secondary MISO network to operate in the presence of a primary system with interference limits

- Principle & Novelty:
  - Popular cognitive radio design problem
  - To make underlay (interference level constraints at primary Rxs) feasible, exploitation of position information to determine attenuations probably only realistic approach

- Recent Achievements:
  - min total Tx power under secondary SINR and primary interference constraints
  - solved max sum rate under secondary Tx power constraints and primary interference level constraints [pimrc11]

- Perspectives:
  - Need a realistic performance evaluation in a complete setup, either analytically or via simulations.

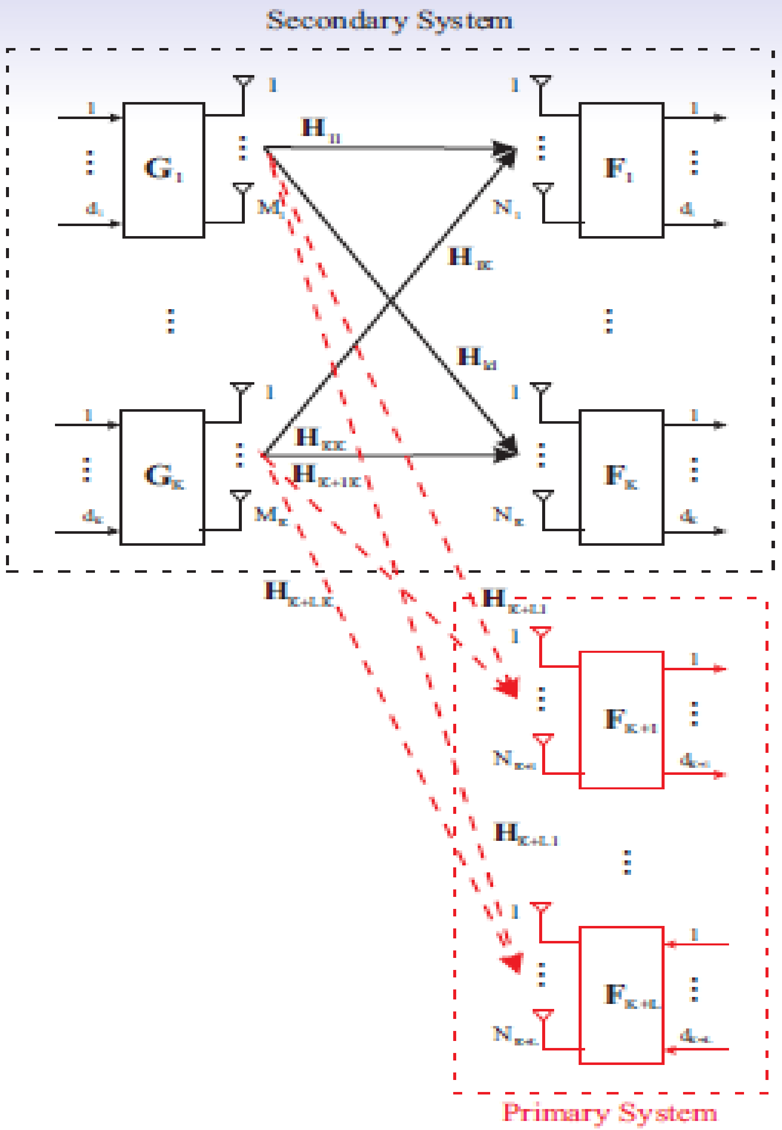


Spatial Interweave Cognitive Radio MIMO IFC

- Context & Problem:

  Allow a secondary MIMO network to operate in the presence of a primary system with multiple antennas also

- Principle & Novelty:
  - Extension of a popular cognitive radio design problem to the multi-antenna case (allows ZF)
  - If secondary2primary link in LOS, number of primary antennas irrelevant, location aided spatial signature
- Recent Achievements:
  - Solved max sum rate under secondary Tx power constraints and primary ZF constraints [CogART11]
  - Addressed secondary RF calibration in a TDD based CSIT approach [JSAC12subm]
- Perspectives:
  - Need a realistic performance evaluation in a complete setup, either analytically or via simulations.



Outline

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)

Kalman Filter

Linear state-space model:

state update equation:

$$\mathbf{x}_{k+1} = \mathbf{F}_k(\theta)\,\mathbf{x}_k + \mathbf{G}_k(\theta)\,\mathbf{w}_k$$

measurement equation:

$$\mathbf{y}_k = \mathbf{H}_k(\theta)\,\mathbf{x}_k + \mathbf{v}_k$$

for $k = 1, 2, \ldots$, with uncorrelated

- initial state $\mathbf{x}_0 \sim \mathcal{N}(\widehat{\mathbf{x}}_0, \mathbf{P}_0)$,
- measurement noise $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k(\theta))$,
- state noise $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k(\theta))$.

State model known up to some parameters θ.

Often $\mathbf{F}_k(\theta)$, $\mathbf{G}_k(\theta)$, $\mathbf{H}_k(\theta)$ linear in θ: bilinear case.

Numerous Applications

- Bayesian adaptive filtering (or wireless channel estimation):

  x_k = FIR filter response,
  θ: Power Delay Profile, AR(1) dynamics

- Position tracking (GPS):

$$\mathbf{x}_{t+1} = \begin{bmatrix} 1 & \Delta t & \frac{1}{2}\Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{x}_t = \begin{bmatrix} x_t + \Delta t \cdot v_t + \frac{1}{2}\Delta t^2 a \\ v_t + \Delta t \cdot a \\ a \end{bmatrix}$$

  θ: acceleration model parameters (e.g. white noise, AR(1))

- Blind Audio Source Separation (BASS):

  x_k = source signals,
  θ: (short+long term) AR parameters, reverb filters

Back to Basics

- $\mathbf{y}$ measurements
  $\mathbf{x}$ random parameters (in some applications (COM, DOA): nuisance parameters, in BASS: parameters of interest)
  θ deterministic parameters
- joint estimation: $f(\mathbf{y},\mathbf{x}|\theta)$, MAP for $\mathbf{x}$, ML for θ
  ML perf $R^J_{\widetilde{\theta}\widetilde{\theta}}$, CRB^J_θ
- marginalized estimation: $f(\mathbf{y}|\theta)$ ML for θ
  ML perf $R^M_{\widetilde{\theta}\widetilde{\theta}}$, CRB^M_θ
- asymptotic perf ordering $C^J_{\widetilde{\theta}\widetilde{\theta}} \overset{(i)}{\geq} C^M_{\widetilde{\theta}\widetilde{\theta}} \overset{(ii)}{=} \mathrm{CRB}^M_\theta \overset{(iii)}{\geq} \mathrm{CRB}^J_\theta$
  (ii): $\widehat{\theta}^M_{ML}$ is consistent
  (i): inconsistent $\widehat{\mathbf{x}}_{MAP}$ prevents $\widehat{\theta}^J_{ML}$ from reaching its CRB

(iii): $$\underbrace{-\frac{\partial^2}{\partial\theta\partial\theta^T}\ln f(\mathbf{y}|\theta)}_{\sim\mathrm{CRB}_\theta^{-M}} = \underbrace{-\mathrm{E}_{\,\mathbf{x}|\mathbf{y},\theta}\frac{\partial^2}{\partial\theta\partial\theta^T}\ln f(\mathbf{y}|\mathbf{x},\theta)}_{\sim\mathrm{CRB}_\theta^{-J}} - Cov_{\mathbf{x}|\mathbf{y},\theta}\left(\tfrac{\partial}{\partial\theta}\ln f(\mathbf{y}|\mathbf{x},\theta)\right)$$

Back to Basics (2)

- EM-KF: converges to joint MAP-ML
  $\widehat{\theta}$ from $\widehat{\mathbf{x}}$ and $\widetilde{\mathbf{x}}$, $\widehat{\mathbf{x}}$ from $\widehat{\theta}$
- alternating MAP-ML: converges to joint MAP-ML
  $\widehat{\theta}$ from $\widehat{\mathbf{x}}$ only, $\widehat{\mathbf{x}}$ from $\widehat{\theta}$
- VB-KF (Variational Bayes): perf?
  $\widehat{\theta}$ from $\widehat{\mathbf{x}}$ and $\widetilde{\mathbf{x}}$, $\widehat{\mathbf{x}}$ from $\widehat{\theta}$ and $\widetilde{\theta}$
  Now also θ random and effect of $\widetilde{\theta}$ on $\mathbf{x}$ acounted for.
  Alternating minimization of Kullback-Leibler divergence of product form of Gaussian posteriors for θ and $\mathbf{x}$ to correct joint posterior: clear approximation sense.
- $\widehat{\theta}^M$ better than $\widehat{\theta}_J$: but in BASS we are interested in $\mathbf{x}$!
  Q: can a reuse of $\widehat{\theta}^M$ in the estimation of $\mathbf{x}$ do better than $\widehat{\mathbf{x}}^J$?
- compared to EKF etc?

Summary

- fading, diversity
- interference single cell: Broadcast Channel (BC)
- interference multi-cell: Interference Channel (IFC)
  - Degrees of Freedom (DoF) and Interference Alignment (IA)
  - Weighted Sum Rate (WSR) maximization and Uplink/Downlink duality (UL/DL)
  - distributed Channel State Information at the Transmitter (CSIT) acquisition
- communications based mobile terminal location finding
- location aided communications
- joint signal and paramater estimation (Wiener/Kalman filtering + parameters)