Random Matrices in Wireless Communications
Course 2: System performance analysis: capacity and rate regions

Romain Couillet
ST-Ericsson, Supélec, FRANCE
romain.couillet@supelec.fr

Supélec

Outline

1. Stieltjes transform methods for more elaborate models
2. Kronecker models and Variance Profiles
3. Capacity expressions, Rate Regions
4. Touching the boundary: optimal power allocation
5. Case study: exchanging relevant data in large self-organized networks
   - Orthogonal CDMA networks
   - Spectrum sharing in multiple access channels

Outline

Reminder and scope

- In Part 1 of this course,
  - we defined the Stieltjes transform:

Definition

Let F be a distribution function, and $z \in \mathbb{C}^+$. Then the Stieltjes transform $m_F(z)$ of F is defined as

$$m_F(z) = \int \frac{1}{\lambda - z} dF(\lambda)$$

For F the spectral distribution of an Hermitian matrix $\mathbf{X} \in \mathbb{C}^{N \times N}$,

$$m_F(z) = \frac{1}{N} \operatorname{tr}(\mathbf{X} - z\mathbf{I}_N)^{-1}$$

  - We gave limiting distribution results for some matrix models.
  - We gave a sketch of the proof of the Marčenko-Pastur law.
- In Part 2, we will
  - extend the notion of limit distributions to deterministic equivalents
  - provide sound mathematical techniques to prove convergence/existence/uniqueness of large N results.
  - provide first wireless communication results
  - apply the results proven above to self-organized networks

Reminder and scope

- In Part 1 of this course,
  - we defined the Stieltjes transform:

Definition

Let F be a distribution function, and $z \in \mathbb{C}^+$. Then the Stieltjes transform $m_F(z)$ of F is defined as

$$m_F(z) = \int \frac{1}{\lambda - z} dF(\lambda)$$

For F the spectral distribution of an Hermitian matrix $\mathbf{X} \in \mathbb{C}^{N \times N}$,

$$m_F(z) = \frac{1}{N} \operatorname{tr}(\mathbf{X} - z\mathbf{I}_N)^{-1}$$

  - We gave limiting distribution results for some matrix models.
  - We gave a sketch of the proof of the Marčenko-Pastur law.
- In Part 2, we will
  - extend the notion of limit distributions to deterministic equivalents
  - provide sound mathematical techniques to prove convergence/existence/uniqueness of large N results.
  - provide first wireless communication results
  - apply the results proven above to self-organized networks

Limiting results against deterministic equivalents

- previously, we showed results of the type:

  "let $\mathbf{X}_N$ be random, $\mathbf{T}_N$ deterministic with $F^{\mathbf{T}_N} \Rightarrow F^T$, etc. Then, when $N \to \infty$, the e.s.d. of $\mathbf{X}_N$ tends to F such that m_F is solution of a fixed-point equation,

  $$m_{\mathbf{X}_N}(z) \to m_F(z) \text{ "}$$

- this has major drawbacks
  - this assumes $\mathbf{T}_N$ *has* a limiting distribution
  - if it does, $m_{\mathbf{X}_N \mathbf{X}_N^H}$ can at best be approximated by m_F which is a function of the limiting F^T. For finite N, $F^{\mathbf{T}_N}$ may be very different from F^T.
  - any sequence $\mathbf{T}_N$ with l.s.d. F^T engenders the same l.s.d. F.
- instead, we shall use results of the type

  "let $\mathbf{X}_N$ be random, $\mathbf{T}_N$ deterministic ~~with $F^{\mathbf{T}_N} \Rightarrow F^T$~~, etc. Then the e.s.d. of $\mathbf{X}_N$ ~~tends to F such that m_F is solution of a fixed-point equation~~ has Stieltjes transform $m_{\mathbf{X}_N}$ well approximated by the deterministic m_N°, which is the unique solution of a fixed-point equation and such that

  $$m_{\mathbf{X}_N}(z) - m_N^\circ(z) \xrightarrow{\text{a.s.}} 0 \text{ "}$$

  In this case, m_N° is a function of $\mathbf{T}_N$, for fixed N and does not require any convergence of $F^{\mathbf{T}_N}$.

Limiting results against deterministic equivalents

- previously, we showed results of the type:

  "let $\mathbf{X}_N$ be random, $\mathbf{T}_N$ deterministic with $F^{\mathbf{T}_N} \Rightarrow F^T$, etc. Then, when $N \to \infty$, the e.s.d. of $\mathbf{X}_N$ tends to F such that m_F is solution of a fixed-point equation,

  $$m_{\mathbf{X}_N}(z) \to m_F(z)\text{ "}$$

- this has major drawbacks
  - this assumes $\mathbf{T}_N$ *has* a limiting distribution
  - if it does, $m_{\mathbf{X}_N\mathbf{X}_N^{\mathsf{H}}}$ can at best be approximated by m_F which is a function of the limiting F^T. For finite N, $F^{\mathbf{T}_N}$ may be very different from F^T.
  - any sequence $\mathbf{T}_N$ with l.s.d. F^T engenders the same l.s.d. F.
- instead, we shall use results of the type

  "let $\mathbf{X}_N$ be random, $\mathbf{T}_N$ deterministic ~~with $F^{\mathbf{T}_N} \Rightarrow F^T$~~, etc. Then the e.s.d. of $\mathbf{X}_N$ ~~tends to F such that m_F is solution of a fixed-point equation~~ has Stieltjes transform $m_{\mathbf{X}_N}$ well approximated by the deterministic m_N°, which is the unique solution of a fixed-point equation and such that

  $$m_{\mathbf{X}_N}(z) - m_N^\circ(z) \xrightarrow{\text{a.s.}} 0\text{ "}$$

  In this case, m_N° is a function of $\mathbf{T}_N$, for fixed N and does not require any convergence of $F^{\mathbf{T}_N}$.

Limiting results against deterministic equivalents

- previously, we showed results of the type:

  "let $\mathbf{X}_N$ be random, $\mathbf{T}_N$ deterministic with $F^{\mathbf{T}_N} \Rightarrow F^T$, etc. Then, when $N \to \infty$, the e.s.d. of $\mathbf{X}_N$ tends to F such that m_F is solution of a fixed-point equation,

$$m_{\mathbf{X}_N}(z) \to m_F(z) \text{ ''}$$

- this has major drawbacks
  - this assumes $\mathbf{T}_N$ *has* a limiting distribution
  - if it does, $m_{\mathbf{X}_N\mathbf{X}_N^{\mathsf{H}}}$ can at best be approximated by m_F which is a function of the limiting F^T. For finite N, $F^{\mathbf{T}_N}$ may be very different from F^T.
  - any sequence $\mathbf{T}_N$ with l.s.d. F^T engenders the same l.s.d. F.
- instead, we shall use results of the type

  "let $\mathbf{X}_N$ be random, $\mathbf{T}_N$ deterministic ~~with $F^{\mathbf{T}_N} \Rightarrow F^T$~~, etc. Then the e.s.d. of $\mathbf{X}_N$ ~~tends to F such that m_F is solution of a fixed-point equation~~ has Stieltjes transform $m_{\mathbf{X}_N}$ well approximated by the deterministic m_N°, which is the unique solution of a fixed-point equation and such that

$$m_{\mathbf{X}_N}(z) - m_N^\circ(z) \xrightarrow{\text{a.s.}} 0 \text{ ''}$$

In this case, m_N° is a function of $\mathbf{T}_N$, for fixed N and does not require any convergence of $F^{\mathbf{T}_N}$.

Outline of the proofs

It will often be the case that the deterministic equivalent $m_N^\circ(z)$ satisfies an implicit equation. The steps are then:

1. find a suitable function f, such that the *true* Stieltjes transform $m_{\mathbf{X}_N}(z)$ satisfies, for fixed $z \in \mathbb{C}^+$,
$$m_{\mathbf{X}_N}(z) = f(m_{\mathbf{X}_N}(z)) + \varepsilon_N$$
where $\varepsilon_N \xrightarrow{\text{a.s.}} 0$ as $N \to \infty$.
This can be done
   - using Pastur's method (see proof of Marčenko-Pastur law in Part 1)
   - using guess-work (see Bai and Silverstein's proofs)

   Remark: This is as far as we went in Part 1.
2. For fixed N, prove the existence of a solution to
$$m = f(m)$$
This is often based on extracting a converging subsequence of $m_N, m_{2N}, \ldots$ such that m_{jN} "has the same properties as $m_{\mathbf{X}_N}(z)$ for all j".
3. For this fixed N, prove the uniqueness of the solution. This involves finding a contradiction if two solutions exist.
4. Calling $m_N^\circ(z)$ the solution, prove finally that
$$m_{\mathbf{X}_N}(z) - m_N^\circ(z) \xrightarrow{\text{a.s.}} 0$$

Outline of the proofs

It will often be the case that the deterministic equivalent $m_N^\circ(z)$ satisfies an implicit equation. The steps are then:

1. find a suitable function f, such that the *true* Stieltjes transform $m_{\mathbf{X}_N}(z)$ satisfies, for fixed $z \in \mathbb{C}^+$,
$$m_{\mathbf{X}_N}(z) = f(m_{\mathbf{X}_N}(z)) + \varepsilon_N$$
where $\varepsilon_N \xrightarrow{\text{a.s.}} 0$ as $N \to \infty$.
This can be done
    - using Pastur's method (see proof of Marčenko-Pastur law in Part 1)
    - using guess-work (see Bai and Silverstein's proofs)

   Remark: This is as far as we went in Part 1.
2. For fixed N, prove the existence of a solution to
$$m = f(m)$$
This is often based on extracting a converging subsequence of $m_N, m_{2N}, \ldots$ such that m_{jN} "has the same properties as $m_{\mathbf{X}_N}(z)$ for all j".
3. For this fixed N, prove the uniqueness of the solution. This involves finding a contradiction if two solutions exist.
4. Calling $m_N^\circ(z)$ the solution, prove finally that
$$m_{\mathbf{X}_N}(z) - m_N^\circ(z) \xrightarrow{\text{a.s.}} 0$$

Outline of the proofs

It will often be the case that the deterministic equivalent $m_N^\circ(z)$ satisfies an implicit equation. The steps are then:

1. find a suitable function f, such that the *true* Stieltjes transform $m_{\mathbf{X}_N}(z)$ satisfies, for fixed $z \in \mathbb{C}^+$,
$$m_{\mathbf{X}_N}(z) = f(m_{\mathbf{X}_N}(z)) + \varepsilon_N$$
where $\varepsilon_N \xrightarrow{\text{a.s.}} 0$ as $N \to \infty$.
This can be done
    - using Pastur's method (see proof of Marčenko-Pastur law in Part 1)
    - using guess-work (see Bai and Silverstein's proofs)

    Remark: This is as far as we went in Part 1.
2. For fixed N, prove the existence of a solution to
$$m = f(m)$$
This is often based on extracting a converging subsequence of $m_N, m_{2N}, \ldots$ such that m_{jN} "has the same properties as $m_{\mathbf{X}_N}(z)$ for all j".
3. For this fixed N, prove the uniqueness of the solution. This involves finding a contradiction if two solutions exist.
4. Calling $m_N^\circ(z)$ the solution, prove finally that
$$m_{\mathbf{X}_N}(z) - m_N^\circ(z) \xrightarrow{\text{a.s.}} 0$$

Outline of the proofs

It will often be the case that the deterministic equivalent $m_N^\circ(z)$ satisfies an implicit equation. The steps are then:

1. find a suitable function f, such that the *true* Stieltjes transform $m_{\mathbf{X}_N}(z)$ satisfies, for fixed $z \in \mathbb{C}^+$,

$$m_{\mathbf{X}_N}(z) = f(m_{\mathbf{X}_N}(z)) + \varepsilon_N$$

   where $\varepsilon_N \xrightarrow{\text{a.s.}} 0$ as $N \to \infty$.
   This can be done
   - using Pastur's method (see proof of Marčenko-Pastur law in Part 1)
   - using guess-work (see Bai and Silverstein's proofs)

   Remark: This is as far as we went in Part 1.

2. For fixed N, prove the existence of a solution to

$$m = f(m)$$

   This is often based on extracting a converging subsequence of $m_N, m_{2N}, \ldots$ such that m_{jN} "has the same properties as $m_{\mathbf{X}_N}(z)$ for all j".

3. For this fixed N, prove the uniqueness of the solution. This involves finding a contradiction if two solutions exist.

4. Calling $m_N^\circ(z)$ the solution, prove finally that

$$m_{\mathbf{X}_N}(z) - m_N^\circ(z) \xrightarrow{\text{a.s.}} 0$$

Outline

Stieltjes transform of a sum of doubly-correlated matrices

R. Couillet, M. Debbah, J. W. Silverstein, “A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels,” *submitted to IEEE Trans. on Information Theory*.

We will give here the method of proof of the following result

Theorem

For $K, N \in \mathbb{N}$, let

$$\mathbf{B}_N = \sum_{k=1}^{K} \mathbf{R}_k^{\frac{1}{2}} \mathbf{X}_k \mathbf{T}_k \mathbf{X}_k^{\mathsf{H}} \mathbf{R}_k^{\frac{1}{2}} + \mathbf{A} \in \mathbb{C}^{N \times N}$$

where $\mathbf{X}_k \in \mathbb{C}^{N \times n_k}$ i.i.d. of zero mean, variance $1/n_k$; $\mathbf{R}_k \in \mathbb{C}^{N \times N}$ Hermitian nonnegative definite; $\mathbf{T}_k = \mathrm{diag}(\tau_1, \ldots, \tau_{n_k}) \in \mathbb{R}^{n_k \times n_k}$, diagonal with $\tau_i \geq 0$; the sequences $\{F^{\mathbf{T}_k}\}_{n_k \geq 1}$ and $\{F^{\mathbf{R}_k}\}_{N \geq 1}$ are tight; $\mathbf{A} \in \mathbb{C}^{N \times N}$ Hermitian positive definite; $0 < a \leq \liminf_N c_k \leq \limsup_N c_N \leq b < \infty$ with $c_k = N/n_k$. Then

$$m_{\mathbf{B}_N}(z) - m_N^{\circ}(z) \xrightarrow{\text{a.s.}} 0$$

where

$$m_N^{\circ}(z) = \frac{1}{N} \operatorname{tr} \left(\mathbf{A} + \sum_{k=1}^{K} \int \frac{\tau_k dF^{\mathbf{T}_k}(\tau_k)}{1 + c_k \tau_k e_k(z)} \mathbf{R}_k - z \mathbf{I}_N \right)^{-1}$$

and the scalars $\{e_i(z)\}$, $i \in \{1, \ldots, K\}$, form the unique solution to

$$e_i(z) = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\mathbf{A} + \sum_{k=1}^{K} \int \frac{\tau_k dF^{\mathbf{T}_k}(\tau_k)}{1 + c_k \tau_k e_k(z)} \mathbf{R}_k - z \mathbf{I}_N \right)^{-1}$$

such that $\mathrm{sgn}(\Im[e_i(z)]) = \mathrm{sgn}(\Im[z])$.

A “telecom-oriented” version of the same result

R. Couillet, M. Debbah, J. W. Silverstein, “A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels,” *submitted to IEEE Trans. on Information Theory*.

$$\mathbf{B}_N = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{H}_k^{\mathsf{H}}, \text{ with } \mathbf{H}_k = \mathbf{R}_k^{\frac{1}{2}} \mathbf{X}_k \mathbf{T}_k^{\frac{1}{2}}$$

with $\mathbf{X}_k \in \mathbb{C}^{N \times n_k}$ with i.i.d. entries of zero mean, variance $1/n_k$, $\mathbf{R}_k$ Hermitian nonnegative definite, $\mathbf{T}_k$ diagonal. Denote $c_k = N/n_k$. Then, as all N and n_k grow large, with ratio c_k,

$$m_{F^{\mathbf{B}_N}}(z) - m_N^{\circ}(z) \xrightarrow{\text{a.s.}} 0$$

where

$$m_N^{\circ}(z) = \frac{1}{N} \operatorname{tr} \left(-z \left[\mathbf{I}_N + \sum_{k=1}^{K} \bar{e}_k(z) \mathbf{R}_k \right] \right)^{-1}$$

and the set of functions $\{e_i(z)\}$ form the unique solution to the K equations

$$e_i(z) = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(-z \left[\mathbf{I}_N + \sum_{k=1}^{K} \bar{e}_k(z) \mathbf{R}_k \right] \right)^{-1}$$

$$\bar{e}_i(z) = \frac{1}{n_i} \operatorname{tr} \mathbf{T}_i \left(-z \left[\mathbf{I}_{n_i} + c_i e_i(z) \mathbf{T}_i \right] \right)^{-1}$$

Pastur's method

Pastur's method is *not* applicable here, unless all $\mathbf{R}_k$'s are diagonal.

Consider $K = 2$, $\mathbf{A} = 0$ and denote $\mathbf{H}_k = \mathbf{R}_k^{\frac{1}{2}} \mathbf{X}_k \mathbf{T}_k^{\frac{1}{2}}$, with diagonal $\mathbf{R}_k$. By block-matrix inversion, we have

$$\left(\mathbf{H}_1\mathbf{H}_1^{\mathsf{H}} + \mathbf{H}_2\mathbf{H}_2^{\mathsf{H}} - z\mathbf{I}_N\right)^{-1}_{11} = \left(\begin{bmatrix} \mathbf{h}_1^{\mathsf{H}} & \mathbf{h}_2^{\mathsf{H}} \\ \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 & \mathbf{U}_1^{\mathsf{H}} \\ \mathbf{h}_2 & \mathbf{U}_2^{\mathsf{H}} \end{bmatrix} - z\mathbf{I}_N\right)^{-1}_{11}$$

$$= \left[-z - z[\mathbf{h}_1^{\mathsf{H}}\mathbf{h}_2^{\mathsf{H}}]\left(\begin{bmatrix} \mathbf{U}_1^{\mathsf{H}} \\ \mathbf{U}_2^{\mathsf{H}} \end{bmatrix}[\mathbf{U}_1\mathbf{U}_2] - z\mathbf{I}_{n_1+n_2}\right)^{-1}\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}\right]^{-1}$$

with the definition $\mathbf{H}_i^{\mathsf{H}} = [\mathbf{h}_i\mathbf{U}_i^{\mathsf{H}}]$.
The inner inverse matrix is

$$\left(\begin{bmatrix} \mathbf{U}_1^{\mathsf{H}} \\ \mathbf{U}_2^{\mathsf{H}} \end{bmatrix}[\mathbf{U}_1\mathbf{U}_2] - z\mathbf{I}_{n_1+n_2}\right)^{-1} = \begin{bmatrix} \mathbf{U}_1^{\mathsf{H}}\mathbf{U}_1 - z\mathbf{I}_{n_1} & \mathbf{U}_1^{\mathsf{H}}\mathbf{U}_2 \\ \mathbf{U}_2^{\mathsf{H}}\mathbf{U}_1 & \mathbf{U}_2^{\mathsf{H}}\mathbf{U}_2 - z\mathbf{I}_{n_2} \end{bmatrix}^{-1}$$

on which we apply another block matrix inverse lemma. The upper-left $(n_1 \times n_1)$ entry equals

$$\left(-z\mathbf{U}_1^{\mathsf{H}}(\mathbf{U}_2\mathbf{U}_2^{\mathsf{H}} - z\mathbf{I}_{N-1})^{-1}\mathbf{U}_1 - z\mathbf{I}_{n_1}\right)^{-1}$$

For the second block diagonal entry, it suffices to revert all 1's in 2's and vice-versa.

Pastur's method

Pastur's method is *not* applicable here, unless all $\mathbf{R}_k$'s are diagonal.

Consider $K = 2$, $\mathbf{A} = 0$ and denote $\mathbf{H}_k = \mathbf{R}_k^{\frac{1}{2}} \mathbf{X}_k \mathbf{T}_k^{\frac{1}{2}}$, with diagonal $\mathbf{R}_k$. By block-matrix inversion, we have

$$\left(\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}} + \mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N\right)^{-1}_{11} = \left(\begin{bmatrix} \mathbf{h}_1^{\mathrm{H}} & \mathbf{h}_2^{\mathrm{H}} \\ \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 & \mathbf{U}_1^{\mathrm{H}} \\ \mathbf{h}_2 & \mathbf{U}_2^{\mathrm{H}} \end{bmatrix} - z\mathbf{I}_N\right)^{-1}_{11}$$

$$= \left[-z - z[\mathbf{h}_1^{\mathrm{H}}\mathbf{h}_2^{\mathrm{H}}]\left(\begin{bmatrix} \mathbf{U}_1^{\mathrm{H}} \\ \mathbf{U}_2^{\mathrm{H}} \end{bmatrix}[\mathbf{U}_1\mathbf{U}_2] - z\mathbf{I}_{n_1+n_2}\right)^{-1}\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}\right]^{-1}$$

with the definition $\mathbf{H}_i^{\mathrm{H}} = [\mathbf{h}_i\mathbf{U}_i^{\mathrm{H}}]$.
The inner inverse matrix is

$$\left(\begin{bmatrix} \mathbf{U}_1^{\mathrm{H}} \\ \mathbf{U}_2^{\mathrm{H}} \end{bmatrix}[\mathbf{U}_1\mathbf{U}_2] - z\mathbf{I}_{n_1+n_2}\right)^{-1} = \begin{bmatrix} \mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1 - z\mathbf{I}_{n_1} & \mathbf{U}_1^{\mathrm{H}}\mathbf{U}_2 \\ \mathbf{U}_2^{\mathrm{H}}\mathbf{U}_1 & \mathbf{U}_2^{\mathrm{H}}\mathbf{U}_2 - z\mathbf{I}_{n_2} \end{bmatrix}^{-1}$$

on which we apply another block matrix inverse lemma. The upper-left $(n_1 \times n_1)$ entry equals

$$\left(-z\mathbf{U}_1^{\mathrm{H}}(\mathbf{U}_2\mathbf{U}_2^{\mathrm{H}} - z\mathbf{I}_{N-1})^{-1}\mathbf{U}_1 - z\mathbf{I}_{n_1}\right)^{-1}$$

For the second block diagonal entry, it suffices to revert all 1's in 2's and vice-versa.

Pastur's method (2)

$$\left(\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}} + \mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N\right)^{-1}_{11} =$$

$$\left[-z - z[\mathbf{h}_1^{\mathrm{H}}\mathbf{h}_2^{\mathrm{H}}]\begin{bmatrix}\left(-z\mathbf{U}_1^{\mathrm{H}}(\mathbf{U}_2\mathbf{U}_2^{\mathrm{H}} - z\mathbf{I}_{N-1})^{-1}\mathbf{U}_1 - z\mathbf{I}_{n_1}\right)^{-1} & \star \\ \star & \left(-z\mathbf{U}_2^{\mathrm{H}}(\mathbf{U}_1\mathbf{U}_1^{\mathrm{H}} - z\mathbf{I}_{N-1})^{-1}\mathbf{U}_2 - z\mathbf{I}_{n_2}\right)^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{h}_1 \\ \mathbf{h}_2\end{bmatrix}\right]$$

The other two terms do not depend on $\mathbf{h}_1$, $\mathbf{h}_2$. We now use both results,

For $\mathbf{x} \in \mathbb{C}^N$, $\mathbf{y} \in \mathbb{C}^N$ i.i.d. with zero mean, variance $1/N$, $\mathbf{A} \in \mathbb{C}^{N\times N}$ Hermitian with bounded spectral norm,

$$\mathbf{x}^{\mathrm{H}}\mathbf{A}\mathbf{x} - \frac{1}{N}\operatorname{tr}\mathbf{A} \xrightarrow{\text{a.s.}} 0$$

$$\mathbf{x}^{\mathrm{H}}\mathbf{A}\mathbf{y} \xrightarrow{\text{a.s.}} 0$$

Since $\mathbf{R}_1$, $\mathbf{R}_2$ are diagonal, $\mathbf{h}_i = \sqrt{r_{i1}}\mathbf{T}_i^{\frac{1}{2}}\mathbf{x}_i$, with the notation $\mathbf{R}_i = \operatorname{diag}(r_{i1}, \ldots, r_{iN})$. Therefore, using the trace and rank-1 perturbation lemma,

$$\left(\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}} + \mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N\right)^{-1}_{11} \to$$

$$\left[-z - zr_{11}\frac{1}{n_1}\operatorname{tr}\mathbf{T}_1\left(-z\mathbf{H}_1^{\mathrm{H}}(\mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N)^{-1}\mathbf{H}_1 - z\mathbf{I}_{n_1}\right)^{-1} - zr_{21}\frac{1}{n_2}\operatorname{tr}\mathbf{T}_2\left(-z\mathbf{H}_2^{\mathrm{H}}(\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}} - z\mathbf{I}_N)^{-1}\mathbf{H}_1 - z\mathbf{I}_{n_2}\right)^{-1}\right]$$

Pastur's method (3)

Now, denoting $\mathbf{H}_i = [\tilde{\mathbf{h}}_i \tilde{\mathbf{U}}_i]$ (column selection instead of row),

$$\mathbf{T}_1\left(-z\mathbf{H}_1^{\mathrm{H}}(\mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N)^{-1}\mathbf{H}_1 - z\mathbf{I}_{n_1}\right)^{-1}_{11} = \tau_{11}\left[-z - z\tilde{\mathbf{h}}_1^{\mathrm{H}}\left(\tilde{\mathbf{U}}_1\tilde{\mathbf{U}}_1^{\mathrm{H}} + \mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N\right)^{-1}\tilde{\mathbf{h}}_1\right]^{-1}$$

$$\to \tau_{11}\left[-z - zc_1\tau_{11}\frac{1}{N}\operatorname{tr}\mathbf{R}_1\left(\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}} + \mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N\right)^{-1}\right]^{-1}$$

with τ_{ij} the j^{th} diagonal entry of $\mathbf{T}_i$. A similar result holds when changing 1's in 2's. Call now

$$f_i = \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}} + \mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N\right)^{-1}$$

and

$$\bar{f}_i = \frac{1}{n_i}\operatorname{tr}\mathbf{T}_i\left(-z\mathbf{H}_1^{\mathrm{H}}(\mathbf{H}_2\mathbf{H}_2^{\mathrm{H}} - z\mathbf{I}_N)^{-1}\mathbf{H}_1 - z\mathbf{I}_{n_1}\right)^{-1}$$

we have shown

$$f_i = \lim_{N\to\infty}\frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(-z\bar{f}_1\mathbf{R}_1 - z\bar{f}_2\mathbf{R}_2 - z\mathbf{I}_N\right)^{-1}$$

$$\bar{f}_i = \lim_{N\to\infty}\frac{1}{n_i}\operatorname{tr}\mathbf{T}_i\left(-zc_if_i\mathbf{T}_i - z\mathbf{I}_{n_i}\right)^{-1}$$

Deterministic equivalent approach: guess work

We will use here the "guess-work" method to find the deterministic equivalent. Consider the simpler case $K = 1$.
Back to the original notations, we seek a matrix $\mathbf{D}$ such that

$$\frac{1}{N}\operatorname{tr}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{D}^{-1} \xrightarrow{\text{a.s.}} 0$$

as $N \to \infty$.

Resolvent lemma

For invertible $\mathbf{A}$, $\mathbf{B}$ matrices,

$$\mathbf{A}^{-1} - \mathbf{B}^{-1} = -\mathbf{A}^{-1}(\mathbf{A} - \mathbf{B})\mathbf{B}^{-1}$$

Taking the matrix differences,

$$-\mathbf{D}^{-1} + (\mathbf{B}_N - z\mathbf{I}_N)^{-1} = \mathbf{D}^{-1}(\mathbf{A} + \mathbf{R}^{\frac{1}{2}}\mathbf{X}\mathbf{T}\mathbf{X}^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_N - \mathbf{D})(\mathbf{B}_N - z\mathbf{I}_N)^{-1}$$

It seems convenient to take $\mathbf{D} = \mathbf{A} - z\mathbf{I}_N - zp_N\mathbf{R}$ with p_N left to be defined

Deterministic equivalent approach: guess work (2)

"Silverstein's" lemma

Let $\mathbf{A}$ be Hermitian invertible, then for any vector $\mathbf{x}$ and scalar τ such that $\mathbf{A}+\tau\mathbf{x}\mathbf{x}^{\mathsf{H}}$ is invertible

$$\mathbf{x}^{\mathsf{H}}(\mathbf{A}+\tau\mathbf{x}\mathbf{x}^{\mathsf{H}})^{-1} = \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}}{1+\tau\mathbf{x}\mathbf{A}^{-1}\mathbf{x}^{\mathsf{H}}}$$

With $\mathbf{D} = \mathbf{A} - z\mathbf{I}_N - zp_N\mathbf{R}$,

$$\begin{aligned}
-\mathbf{D}^{-1} + (\mathbf{B}_N - z\mathbf{I}_N)^{-1} &= \mathbf{D}^{-1}(\mathbf{A} + \mathbf{R}^{\frac{1}{2}}\mathbf{X}\mathbf{T}\mathbf{X}^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}} - z\mathbf{I}_N - \mathbf{D})(\mathbf{B}_N - z\mathbf{I}_N)^{-1} \\
&= \mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}\left(\mathbf{X}\mathbf{T}\mathbf{X}^{\mathsf{H}}\right)\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} + zp_N\mathbf{D}^{-1}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} \\
&= \mathbf{D}^{-1}\sum_{j=1}^{n}\tau_j\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j\mathbf{x}_j^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} + zp_N\mathbf{D}^{-1}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} \\
&= \sum_{j=1}^{n}\tau_j\frac{\mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j\mathbf{x}_j^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)} - z\mathbf{I}_N)^{-1}}{1+\tau_j\mathbf{x}^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)} - z\mathbf{I}_N)^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j} + zp_N\mathbf{D}^{-1}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}
\end{aligned}$$

Choice of p_N: $p_N = -\frac{1}{z}\sum_{j=1}^{n}\frac{\tau_j}{1+\tau_j c\frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}}$

$$\frac{1}{N}\operatorname{tr}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{D}^{-1} = \sum_{j=1}^{n}\tau_j\left[\frac{\mathbf{x}_j^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)} - z\mathbf{I}_N)^{-1}\mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j}{1+\tau_j\mathbf{x}^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)} - z\mathbf{I}_N)^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j} - \frac{\frac{1}{N}\operatorname{tr}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}\mathbf{R}\mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}}{1+c\tau_j\frac{1}{N}\operatorname{tr}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}}\mathbf{R}^{\frac{1}{2}}\right]$$

Deterministic equivalent approach: guess work (2)

"Silverstein's" lemma

Let $\mathbf{A}$ be Hermitian invertible, then for any vector $\mathbf{x}$ and scalar τ such that $\mathbf{A}+\tau\mathbf{x}\mathbf{x}^{\mathsf{H}}$ is invertible

$$\mathbf{x}^{\mathsf{H}}(\mathbf{A}+\tau\mathbf{x}\mathbf{x}^{\mathsf{H}})^{-1} = \frac{\mathbf{x}^{\mathsf{H}}\mathbf{A}^{-1}}{1+\tau\mathbf{x}\mathbf{A}^{-1}\mathbf{x}^{\mathsf{H}}}$$

With $\mathbf{D}=\mathbf{A}-z\mathbf{I}_N-zp_N\mathbf{R}$,

$$\begin{aligned}
-\mathbf{D}^{-1}+(\mathbf{B}_N-z\mathbf{I}_N)^{-1} &= \mathbf{D}^{-1}(\mathbf{A}+\mathbf{R}^{\frac{1}{2}}\mathbf{X}\mathbf{T}\mathbf{X}^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}-z\mathbf{I}_N-\mathbf{D})(\mathbf{B}_N-z\mathbf{I}_N)^{-1}\\
&= \mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}\left(\mathbf{X}\mathbf{T}\mathbf{X}^{\mathsf{H}}\right)\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}+zp_N\mathbf{D}^{-1}\mathbf{R}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}\\
&= \mathbf{D}^{-1}\sum_{j=1}^{n}\tau_j\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j\mathbf{x}_j^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}+zp_N\mathbf{D}^{-1}\mathbf{R}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}\\
&= \sum_{j=1}^{n}\tau_j\frac{\mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j\mathbf{x}_j^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)}-z\mathbf{I}_N)^{-1}}{1+\tau_j\mathbf{x}^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)}-z\mathbf{I}_N)^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j}+zp_N\mathbf{D}^{-1}\mathbf{R}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}
\end{aligned}$$

Choice of p_N: $p_N=-\frac{1}{z}\sum_{j=1}^{n}\frac{\tau_j}{1+\tau_j c\frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}}$

$$\frac{1}{N}\operatorname{tr}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}-\frac{1}{N}\operatorname{tr}\mathbf{D}^{-1} = \sum_{j=1}^{n}\tau_j\left[\frac{\mathbf{x}_j^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)}-z\mathbf{I}_N)^{-1}\mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j}{1+\tau_j\mathbf{x}^{\mathsf{H}}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_{(j)}-z\mathbf{I}_N)^{-1}\mathbf{R}^{\frac{1}{2}}\mathbf{x}_j}-\frac{\frac{1}{N}\operatorname{tr}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}\mathbf{R}\mathbf{D}^{-1}\mathbf{R}^{\frac{1}{2}}}{1+c\tau_j\frac{1}{N}\operatorname{tr}\mathbf{R}^{\frac{1}{2}}(\mathbf{B}_N-z\mathbf{I}_N)^{-1}}\mathbf{R}^{\frac{1}{2}}\right]$$

Deterministic equivalent approach: guess work (3)

The same can be done for $\frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1}$ and we get

$$\frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{R}\mathbf{D}^{-1} \to 0$$

To show that the convergence is almost sure, we use truncation and centralization.

Truncation and centralization

Replace $\mathbf{X}_N$, $\mathbf{T}_N$ and $\mathbf{R}_N$ by $\bar{\mathbf{X}}_N$, $\bar{\mathbf{T}}_N$ and $\bar{\mathbf{R}}_N$ in the following fashion

$$\left(\bar{\mathbf{X}}_N\right)_{ij} = (\mathbf{X}_N)_{ij} \cdot I_{\{(\mathbf{X}_N)_{ij} < g_N\}}$$

for g_N that grows

- fast enough to ensure $F^{\mathbf{B}_N} - F^{\bar{\mathbf{B}}_N} \Rightarrow 0$
- slow enough to ensure $\frac{1}{N}\operatorname{tr}(\bar{\mathbf{B}}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\bar{\mathbf{R}}\bar{\mathbf{D}}^{-1} \xrightarrow{\text{a.s.}} 0$

Showing that some moment of the terms appearing in the difference is summable, applying Borel-Cantelli lemma, we have almost sure convergence.

Application of the Borel-Cantelli lemma

To complete the proof of almost sure convergence, denote

$$w_N = \frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - \frac{1}{N}\operatorname{tr}\mathbf{R}\mathbf{D}^{-1}$$

We divide w_N is 4 successive differences $w_N = w_N^1 + \ldots + w_N^4$. The strategy is as follows:

- for all i, show that
$$\mathrm{E}|w_N^i|^6 < h_N^i$$
where h_N^i is summable
- for $\varepsilon > 0$, applying Markov's inequality,
$$P(|h_N^i| > \varepsilon) < \frac{1}{\varepsilon^6}\mathrm{E}|w_N^i|^6$$
which is summable.
- from Borel-Cantelli, this implies $P(|h_N^i| > \varepsilon \text{ i.o.}) = 0$
- therefore the set $\{\omega : \lim_N m_{\mathbf{B}_N(\omega)}(z) - m_N^\circ(z) = 0\}^c = \bigcup_\varepsilon \{|m_{\mathbf{B}_N(z)} - m_N^\circ(z)| \geq \varepsilon \text{ i.o.}\}$ is a sum of zero probability sets.
- the union above can be done on rational ε's and then the union has probability zero.
- for the z in question, there therefore exists $\Omega_z \subset \Omega$ for which $\lim_N m_{\mathbf{B}_N(\omega)}(z) - m_N^\circ(z) = 0$. It suffices then to countably sample $\mathbb{C}^+$ to generate a dense set of z's which satisfy convergence with probability 1. By local analyticity of m_N° and $m_{\mathbf{B}_N}$, this is true for all $z \in \mathbb{C}^+$.

Deterministic equivalent approach: existence and uniqueness

Fix now N and consider the implicit equation in e

$$e = \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(\mathbf{A} + \int \frac{\tau dF^{\mathbf{T}}(\tau)}{1+c\tau e}\mathbf{R} - z\mathbf{I}_N\right)^{-1}$$

- Existence: for existence, consider the matrices $\mathbf{T}_{[j]} = \mathbf{T}\otimes\mathbf{I}_j$, $\mathbf{R}_{[j]} = \mathbf{R}\otimes\mathbf{I}_j$, $\mathbf{A}_{[j]} = \mathbf{A}\otimes\mathbf{I}_j$. The value of

$$f(e) = \frac{1}{N}\operatorname{tr}\mathbf{R}\left(\mathbf{A}_{[j]} + \int \frac{\tau dF^{\mathbf{T}_{[j]}}(\tau)}{1+c\tau e}\mathbf{R}_{[j]} - z\mathbf{I}_N\right)^{-1}$$

is constant whatever m. Now, take $\omega\in\Omega$ such that $w_N(\omega)\to 0$. For this ω, write

$$\tilde{e}(z) = \frac{1}{N}\operatorname{tr}(\mathbf{B}_N(\omega) - z\mathbf{I}_N)^{-1}$$

Showing that $\tilde{e}(z)$ and $\frac{\tau}{1+c\tau e}$ are uniformly bounded over j, we can take a subsequence of $\tilde{e}(z)$ that goes to, say e. For this e, $w_N = 0$ and then it's a solution.

- Uniqueness: Uniqueness is shown by taking a second solution $\underline{e}$ and by proving that

$$e - \underline{e} = \gamma(e - \underline{e})$$

with $\gamma < 1$.

Deterministic equivalent approach: existence and uniqueness

Fix now N and consider the implicit equation in e

$$e = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\mathbf{A} + \int \frac{\tau dF^{\mathbf{T}}(\tau)}{1 + c\tau e} \mathbf{R} - z\mathbf{I}_N \right)^{-1}$$

- Existence: for existence, consider the matrices $\mathbf{T}_{[j]} = \mathbf{T} \otimes \mathbf{I}_j$, $\mathbf{R}_{[j]} = \mathbf{R} \otimes \mathbf{I}_j$, $\mathbf{A}_{[j]} = \mathbf{A} \otimes \mathbf{I}_j$. The value of

$$f(e) = \frac{1}{N} \operatorname{tr} \mathbf{R} \left(\mathbf{A}_{[j]} + \int \frac{\tau dF^{\mathbf{T}_{[j]}}(\tau)}{1 + c\tau e} \mathbf{R}_{[j]} - z\mathbf{I}_N \right)^{-1}$$

  is constant whatever m. Now, take $\omega \in \Omega$ such that $w_N(\omega) \to 0$. For this ω, write

$$\tilde{e}(z) = \frac{1}{N} \operatorname{tr}(\mathbf{B}_N(\omega) - z\mathbf{I}_N)^{-1}$$

  Showing that $\tilde{e}(z)$ and $\frac{\tau}{1+c\tau e}$ are uniformly bounded over j, we can take a subsequence of $\tilde{e}(z)$ that goes to, say e. For this e, $w_N = 0$ and then it's a solution.
- Uniqueness: Uniqueness is shown by taking a second solution $\underline{e}$ and by proving that

$$e - \underline{e} = \gamma(e - \underline{e})$$

  with $\gamma < 1$.

Deterministic equivalent approach: existence and uniqueness

Fix now N and consider the implicit equation in e

$$e = \frac{1}{N} \operatorname{tr} \mathbf{R}_i \left(\mathbf{A} + \int \frac{\tau dF^{\mathbf{T}}(\tau)}{1 + c\tau e} \mathbf{R} - z\mathbf{I}_N \right)^{-1}$$

- Existence: for existence, consider the matrices $\mathbf{T}_{[j]} = \mathbf{T} \otimes \mathbf{I}_j$, $\mathbf{R}_{[j]} = \mathbf{R} \otimes \mathbf{I}_j$, $\mathbf{A}_{[j]} = \mathbf{A} \otimes \mathbf{I}_j$. The value of

$$f(e) = \frac{1}{N} \operatorname{tr} \mathbf{R} \left(\mathbf{A}_{[j]} + \int \frac{\tau dF^{\mathbf{T}_{[j]}}(\tau)}{1 + c\tau e} \mathbf{R}_{[j]} - z\mathbf{I}_N \right)^{-1}$$

  is constant whatever m. Now, take $\omega \in \Omega$ such that $w_N(\omega) \to 0$. For this ω, write

$$\tilde{e}(z) = \frac{1}{N} \operatorname{tr}(\mathbf{B}_N(\omega) - z\mathbf{I}_N)^{-1}$$

  Showing that $\tilde{e}(z)$ and $\frac{\tau}{1+c\tau e}$ are uniformly bounded over j, we can take a subsequence of $\tilde{e}(z)$ that goes to, say e. For this e, $w_N = 0$ and then it's a solution.
- Uniqueness: Uniqueness is shown by taking a second solution $\underline{e}$ and by proving that

$$e - \underline{e} = \gamma(e - \underline{e})$$

  with $\gamma < 1$.

Deterministic equivalent approach: termination of the proof

- It then suffices to show that $\frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N - z\mathbf{I}_N)^{-1} - e \xrightarrow{\text{a.s.}} 0$
  This exploits the fact that, for some ω in a probability one space, $\frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N(\omega) - z\mathbf{I}_N)^{-1}$ is w_N away from $\frac{1}{N}\mathbf{D}^{-1}\mathbf{R}$. Using the same argument as for uniqueness, we have

$$e - \frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N(\omega) - z\mathbf{I}_N)^{-1} = \gamma(e - \frac{1}{N}\operatorname{tr}\mathbf{R}(\mathbf{B}_N(\omega) - z\mathbf{I}_N)^{-1}) + w_N$$

  for $\gamma < 1$.
- The same argument applies to $m_N(z) - m_N^\circ(z)$.

Result on the Shannon transform of $\mathbf{B}_N$

Remember now that

$$\int \log(1+xt)dF(t) = \int_{1/x}^{\infty} \left(\frac{1}{t} - m_F(-t)\right) dt$$

R. Couillet, M. Debbah, J. W. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," *submitted to IEEE Trans. on Information Theory*.

Theorem

Under the previous model for $\mathbf{B}_N$, as N, n_k grow large,

$$\frac{1}{N}\log\det(\mathbf{B}_N + x\mathbf{I}_N) - \left[\frac{1}{N}\log\det\left(\mathbf{I}_N + \sum_{k=1}^{K}\bar{e}_k(-1/x)\mathbf{R}_k\right) + \sum_{k=1}^{K}\frac{1}{N}\log\det\left(\mathbf{I}_{n_k} + c_k e_k(-1/x)\mathbf{T}_k\right) - \frac{1}{x}\sum_{k=1}^{K}\bar{e}_k(-1/x)e_k(-1/x)\right] \xrightarrow{\text{a.s.}} 0$$

Result on the Shannon transform of $\mathbf{B}_N$

Remember now that

$$\int \log(1+xt)dF(t) = \int_{1/x}^{\infty} \left(\frac{1}{t} - m_F(-t)\right) dt$$

R. Couillet, M. Debbah, J. W. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," *submitted to IEEE Trans. on Information Theory*.

Theorem

Under the previous model for $\mathbf{B}_N$, as N, n_k grow large,

$$\frac{1}{N}\log\det(\mathbf{B}_N + x\mathbf{I}_N) - \left[\frac{1}{N}\log\det\left(\mathbf{I}_N + \sum_{k=1}^{K}\bar{e}_k(-1/x)\mathbf{R}_k\right)\right.$$
$$+\sum_{k=1}^{K}\frac{1}{N}\log\det\left(\mathbf{I}_{n_k} + c_k e_k(-1/x)\mathbf{T}_k\right)$$
$$\left.-\frac{1}{x}\sum_{k=1}^{K}\bar{e}_k(-1/x)e_k(-1/x)\right] \xrightarrow{\text{a.s.}} 0$$

Result on the Shannon transform of $\mathbf{B}_N$

Remember now that

$$\int \log(1+xt)dF(t) = \int_{1/x}^{\infty} \left(\frac{1}{t} - m_F(-t) \right) dt$$

R. Couillet, M. Debbah, J. W. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," *submitted to IEEE Trans. on Information Theory*.

Theorem

Under the previous model for $\mathbf{B}_N$, as N, n_k grow large,

$$\frac{1}{N}\log\det(\mathbf{B}_N + x\mathbf{I}_N) - \left[\frac{1}{N}\log\det\left(\mathbf{I}_N + \sum_{k=1}^{K} \bar{e}_k(-1/x)\mathbf{R}_k \right) \right.$$
$$+ \sum_{k=1}^{K} \frac{1}{N}\log\det\left(\mathbf{I}_{n_k} + c_k e_k(-1/x)\mathbf{T}_k^{\frac{1}{2}}\mathbf{P}_k\mathbf{T}_k^{\frac{1}{2}} \right)$$
$$\left. -\frac{1}{x}\sum_{k=1}^{K} \bar{e}_k(-1/x)e_k(-1/x) \right] \xrightarrow{\text{a.s.}} 0$$

Variance profile

W. Hachem, Ph. Loubaton, J. Najim, "Deterministic equivalents for certain functionals of large random matrices," Annals of Applied Probability, vol. 17, no. 3, pp. 875-930, 2007.

Theorem

Let $\mathbf{X}_N \in \mathbb{C}^{N\times n}$ *have independent entries with* $(i,j)^{th}$ *entry of zero mean and variance* $\frac{1}{n}\sigma_{ij}^2$*. Let* $\mathbf{A}_N \in \mathbb{R}^{N\times n}$ *be deterministic with uniformly bounded column norm. Then*

$$\frac{1}{N}\operatorname{tr}\left((\mathbf{X}_N+\mathbf{A}_N)(\mathbf{X}_N+\mathbf{A}_N)^{\mathsf{H}}-z\mathbf{I}_N\right)^{-1}-\frac{1}{N}\operatorname{tr}\mathbf{T}_N(z)\xrightarrow{\text{a.s.}}0$$

where $\mathbf{T}_N(z)$ *is the unique function that solves*

$$\mathbf{T}_N(z)=\left(\Psi^{-1}(z)-z\mathbf{A}_N\tilde{\Psi}(z)\mathbf{A}_N^{\mathsf{T}}\right)^{-1},\quad \tilde{\mathbf{T}}_N(z)=\left(\tilde{\Psi}^{-1}(z)-z\mathbf{A}_N^{\mathsf{T}}\Psi(z)\mathbf{A}_N\right)^{-1}$$

with $\Psi(z)=\operatorname{diag}(\psi_i(z))$, $\tilde{\Psi}(z)=\operatorname{diag}(\tilde{\psi}_i(z))$*, with entries defined as*

$$\psi_i(z)=-\left(z(1+\frac{1}{n}\operatorname{tr}\tilde{\mathbf{D}}_i\tilde{\mathbf{T}}(z))\right)^{-1},\quad \tilde{\psi}_j(z)=-\left(z(1+\frac{1}{n}\operatorname{tr}\mathbf{D}_j\mathbf{T}(z))\right)^{-1}$$

and $\mathbf{D}_j=\operatorname{diag}(\sigma_{ij}^2, 1\leq i\leq N)$, $\tilde{\mathbf{D}}_i=\operatorname{diag}(\sigma_{ij}^2, 1\leq j\leq n)$

Variance profile

W. Hachem, Ph. Loubaton, J. Najim, "Deterministic equivalents for certain functionals of large random matrices," Annals of Applied Probability, vol. 17, no. 3, pp. 875-930, 2007.

Theorem

For the previous model, we also have that

$$\frac{1}{N}\mathrm{E}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}(\mathbf{X}_N + \mathbf{A}_N)(\mathbf{X}_N + \mathbf{A}_N)^{\mathsf{H}}\right)$$

has deterministic equivalent

$$\frac{1}{N}\log\det\left[\frac{1}{\sigma^2}\Psi(-\sigma^2)^{-1} + \mathbf{A}_N\tilde{\Psi}(-\sigma^2)\mathbf{A}_N^{\mathsf{T}}\right] + \frac{1}{N}\log\det\frac{1}{\sigma^2}\Psi(-\sigma^2)^{-1} - \frac{\sigma^2}{nN}\sum_{i,j}\sigma_{ij}^2\mathbf{T}_{ii}(-\sigma^2)\tilde{\mathbf{T}}_{jj}(-\sigma^2)$$

Alternative strategies

There exists alternative proof strategies for specific models.

- The Gaussian method:
  - this technique is meant for random Gaussian **X** matrices
  - based on two ingredients: a Gaussian integration by parts formula, and the Nash-Poincaré inequality.
  - *advantages:*
    - sequential method, easy to use
    - give results on convergence speed
    - proves convergence of Gaussian-based models of type $N(\mathrm{E}m_N - m_N^\circ) \to 0$
    - $\Rightarrow$ very convenient to prove *total* capacity convergence, instead of *average* capacity.
  - *drawbacks:*
    - somewhat painful to use, leads to much calculus, less “elegant”
    - proves convergence of Gaussian-based models of type $N(\mathrm{E}m_N - m_N^\circ) \to 0$
    - $\Rightarrow$ less powerful than almost sure results
    - $\Rightarrow$ limited to Gaussian.
- Diagrammatic approaches: moment “drawing”-based approach that uses combinatorics to infer limiting results
- Replica methods: physics-based method, non-mathematically accurate, to *conjecture* limiting results.

Alternative strategies

There exists alternative proof strategies for specific models.

- The Gaussian method:
  - this technique is meant for random Gaussian **X** matrices
  - based on two ingredients: a Gaussian integration by parts formula, and the Nash-Poincaré inequality.
  - *advantages:*
    - sequential method, easy to use
    - give results on convergence speed
    - proves convergence of Gaussian-based models of type $N(\mathrm{E}m_N - m_N^\circ) \to 0$
    - $\Rightarrow$ very convenient to prove *total* capacity convergence, instead of *average* capacity.
  - *drawbacks:*
    - somewhat painful to use, leads to much calculus, less “elegant”
    - proves convergence of Gaussian-based models of type $N(\mathrm{E}m_N - m_N^\circ) \to 0$
    - $\Rightarrow$ less powerful than almost sure results
    - $\Rightarrow$ limited to Gaussian.
- Diagrammatic approaches: moment “drawing”-based approach that uses combinatorics to infer limiting results
- Replica methods: physics-based method, non-mathematically accurate, to *conjecture* limiting results.

Outline

Broadcast channel with Kronecker model

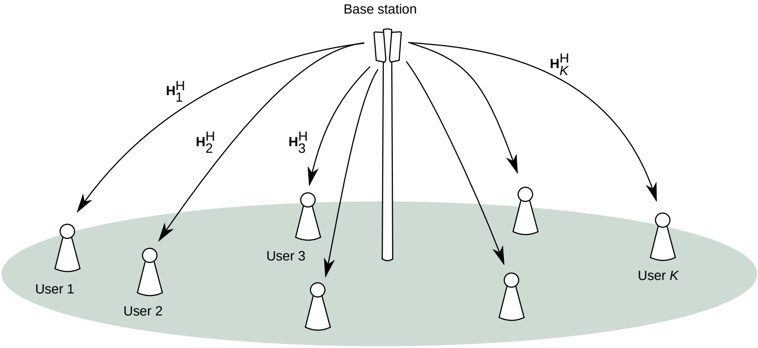

Figure: Downlink scenario in multi-user broadcast channel

Rate region of MAC and BC

S. Vishwanath, N. Jindal and A. Goldsmith, “Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels,” IEEE Trans. on Information Theory, vol. 49, no. 10, 2003.

Assume all channels are modeled as Kronecker; for $k = 1, \ldots, K$

$$\mathbf{H}_k = \mathbf{R}_k^{\frac{1}{2}} \mathbf{X}_k \mathbf{T}_k^{\frac{1}{2}}$$

- Rate region of multiple access channel for K users with channels $\mathbf{H} = [\mathbf{H}_1, \ldots, \mathbf{H}_K]$,

$$\mathbf{C}_{\mathrm{MAC}}(P_1, \ldots, P_K; \mathbf{H}) =$$
$$\bigcup_{\substack{\operatorname{tr}(\mathbf{P}_i) \leq P_i \\ \mathbf{P}_i \geq 0 \\ i=1,\ldots,K}} \left\{ \{R_i, 1 \leq i \leq K\} : \sum_{i \in \mathcal{S}} R_i \leq \frac{1}{N} \log \left| \mathbf{I} + \frac{1}{\sigma^2} \sum_{i \in \mathcal{S}} \mathbf{H}_i \mathbf{P}_i \mathbf{H}_i^{\mathsf{H}} \right|, \forall \mathcal{S} \subset \{1, \ldots, K\} \right\}$$

- Rate region of broadcast channel for $\mathbf{H}^{\mathsf{H}} = [\mathbf{H}_1, \ldots, \mathbf{H}_K]^{\mathsf{H}}$ with total transmit power P,

$$\mathbf{C}_{\mathrm{BC}}(P; \mathbf{H}^{\mathsf{H}}) = \bigcup_{\sum_{k=1}^{K} P_k \leq P} \mathbf{C}_{\mathrm{MAC}}(P_1, \ldots, P_K; \mathbf{H})$$

Reminder: deterministic equivalent for multi-user channel

Under the previous model for $\mathbf{B}_N$, as N, n_k grow large,

$$\frac{1}{N}\log\left|\mathbf{I}+\frac{1}{\sigma^2}\sum_{k\in\mathcal{S}}\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^{\mathsf{H}}\right| - \left[\frac{1}{N}\log\det\left(\mathbf{I}_N+\sum_{k\in\mathcal{S}}\bar{e}_k(-1/x)\mathbf{R}_k\right)\right.$$
$$+\sum_{k\in\mathcal{S}}\frac{1}{N}\log\det\left(\mathbf{I}_{n_k}+c_k e_k(-1/x)\mathbf{T}_k^{\frac{1}{2}}\mathbf{P}_k\mathbf{T}_k^{\frac{1}{2}}\right)$$
$$\left.-\frac{1}{x}\sum_{k=1}^{K}\bar{e}_k(-1/x)e_k(-1/x)\right]\xrightarrow{\text{a.s.}}0$$

Outline

Rate region boundary

- it is desirable to determine the boundary of the rate region
  - for *theoretical purposes*: to fully determine the rate region and alleviate the $\bigcup_{\mathbf{P}_1,\ldots,\mathbf{P}_k}$ sign.
  - for *practical purposes*: to allow users/the base station to transmit at optimal rate.
- it is also desirable to identify the key parameters of the system
  - *in theory*: to extract physical meanings
  - *in theory*: to identify the minimum feedback requirements
  - *in practice*: to minimize information feedback
  - *in practice*: to ease power allocation processing

Rate region boundary

- it is desirable to determine the boundary of the rate region
  - for *theoretical purposes*: to fully determine the rate region and alleviate the $\bigcup_{\mathbf{P}_1,\ldots,\mathbf{P}_k}$ sign.
  - for *practical purposes*: to allow users/the base station to transmit at optimal rate.
- it is also desirable to identify the key parameters of the system
  - *in theory*: to extract physical meanings
  - *in theory*: to identify the minimum feedback requirements
  - *in practice*: to minimize information feedback
  - *in practice*: to ease power allocation processing

Capacity maximizing power allocation

R. Couillet, M. Debbah, J. W. Silverstein, “A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels,” *submitted to IEEE Trans. on Information Theory*.

Consider a subset $\mathcal{S} = \{i_1, \ldots, i_{|\mathcal{S}|}\} \subset \{1, \ldots, K\}$.

- With $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^H$, $\mathbf{D}_k = \operatorname{diag}(\tau_{k1}, \ldots, \tau_{kn_k})$ diagonal, the capacity-achieving matrices $\mathbf{P}_{i_1}^\star, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}^\star$ satisfy
  1. $\mathbf{P}_k^\star = \mathbf{U}_k \mathbf{Q}_k^\star \mathbf{U}_k^H$, with $\mathbf{Q}_k^\star$ diagonal; i.e. the eigenspace of $\mathbf{P}_k^\star$ is the same as the eigenspace of $\mathbf{T}_k$.
  2. with $\bar{e}_k^\star = \bar{e}_k(-\sigma^2, \mathbf{P}_k^\star)$, the i^{th} entry $q_{ki}^\star$ of $\mathbf{Q}_k^\star$ satisfies

$$q_{ki}^\star = \left(\mu_k - \frac{1}{c_k e_k^\star \tau_{ki}}\right)^+$$

  where the μ_k's are evaluated such that $\operatorname{tr}(\mathbf{Q}_k) = P_k$.
- an iterative water-filling method allows to retrieve the $q_{ki}^\star$'s by successively
  - for a given set $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$, evaluating $e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}$
  - updating the new optimal solution $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$ for this system

Iterative water-filling

Upon convergence, the iterative water-filling algorithm leads to the optimal solution.

Capacity maximizing power allocation

R. Couillet, M. Debbah, J. W. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," *submitted to IEEE Trans. on Information Theory*.

Consider a subset $\mathcal{S} = \{i_1, \ldots, i_{|\mathcal{S}|}\} \subset \{1, \ldots, K\}$.

- With $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^{\mathrm{H}}$, $\mathbf{D}_k = \operatorname{diag}(\tau_{k1}, \ldots, \tau_{kn_k})$ diagonal, the capacity-achieving matrices $\mathbf{P}_{i_1}^\star, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}^\star$ satisfy
  1. $\mathbf{P}_k^\star = \mathbf{U}_k \mathbf{Q}_k^\star \mathbf{U}_k^{\mathrm{H}}$, with $\mathbf{Q}_k^\star$ diagonal; i.e. the eigenspace of $\mathbf{P}_k^\star$ is the same as the eigenspace of $\mathbf{T}_k$.
  2. with $\bar{e}_k^\star = \bar{e}_k(-\sigma^2, \mathbf{P}_k^\star)$, the i^{th} entry $q_{ki}^\star$ of $\mathbf{Q}_k^\star$ satisfies
  $$q_{ki}^\star = \left(\mu_k - \frac{1}{c_k e_k^\star \tau_{ki}} \right)^+$$
  where the μ_k's are evaluated such that $\operatorname{tr}(\mathbf{Q}_k) = P_k$.
- an iterative water-filling method allows to retrieve the $q_{ki}^\star$'s by successively
  - for a given set $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$, evaluating $e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}$
  - updating the new optimal solution $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$ for this system

Iterative water-filling

Upon convergence, the iterative water-filling algorithm leads to the optimal solution.

Capacity maximizing power allocation

R. Couillet, M. Debbah, J. W. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," *submitted to IEEE Trans. on Information Theory*.

Consider a subset $\mathcal{S} = \{i_1, \ldots, i_{|\mathcal{S}|}\} \subset \{1, \ldots, K\}$.

- With $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^{\mathsf{H}}$, $\mathbf{D}_k = \mathrm{diag}(\tau_{k1}, \ldots, \tau_{kn_k})$ diagonal, the capacity-achieving matrices $\mathbf{P}_{i_1}^\star, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}^\star$ satisfy
  1. $\mathbf{P}_k^\star = \mathbf{U}_k \mathbf{Q}_k^\star \mathbf{U}_k^{\mathsf{H}}$, with $\mathbf{Q}_k^\star$ diagonal; i.e. the eigenspace of $\mathbf{P}_k^\star$ is the same as the eigenspace of $\mathbf{T}_k$.
  2. with $\bar{e}_k^\star = \bar{e}_k(-\sigma^2, \mathbf{P}_k^\star)$, the i^{th} entry $q_{ki}^\star$ of $\mathbf{Q}_k^\star$ satisfies

$$q_{ki}^\star = \left(\mu_k - \frac{1}{c_k e_k^\star \tau_{ki}} \right)^+$$

  where the μ_k's are evaluated such that $\mathrm{tr}(\mathbf{Q}_k) = P_k$.
- an iterative water-filling method allows to retrieve the $q_{ki}^\star$'s by successively
  - for a given set $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$, evaluating $e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}$
  - updating the new optimal solution $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$ for this system

Iterative water-filling

Upon convergence, the iterative water-filling algorithm leads to the optimal solution.

Capacity maximizing power allocation

R. Couillet, M. Debbah, J. W. Silverstein, “A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels,” *submitted to IEEE Trans. on Information Theory*.

Consider a subset $\mathcal{S} = \{i_1, \ldots, i_{|\mathcal{S}|}\} \subset \{1, \ldots, K\}$.

- With $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^H$, $\mathbf{D}_k = \operatorname{diag}(\tau_{k1}, \ldots, \tau_{kn_k})$ diagonal, the capacity-achieving matrices $\mathbf{P}_{i_1}^\star, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}^\star$ satisfy
  1. $\mathbf{P}_k^\star = \mathbf{U}_k \mathbf{Q}_k^\star \mathbf{U}_k^H$, with $\mathbf{Q}_k^\star$ diagonal; i.e. the eigenspace of $\mathbf{P}_k^\star$ is the same as the eigenspace of $\mathbf{T}_k$.
  2. with $\bar{e}_k^\star = \bar{e}_k(-\sigma^2, \mathbf{P}_k^\star)$, the i^{th} entry $q_{ki}^\star$ of $\mathbf{Q}_k^\star$ satisfies

$$q_{ki}^\star = \left(\mu_k - \frac{1}{c_k e_k^\star \tau_{ki}} \right)^+$$

  where the μ_k's are evaluated such that $\operatorname{tr}(\mathbf{Q}_k) = P_k$.
- an iterative water-filling method allows to retrieve the $q_{ki}^\star$'s by successively
  - for a given set $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$, evaluating $e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}$
  - updating the new optimal solution $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$ for this system

Iterative water-filling

Upon convergence, the iterative water-filling algorithm leads to the optimal solution.

Capacity maximizing power allocation

R. Couillet, M. Debbah, J. W. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," *submitted to IEEE Trans. on Information Theory*.

Consider a subset $\mathcal{S} = \{i_1, \ldots, i_{|\mathcal{S}|}\} \subset \{1, \ldots, K\}$.

- With $\mathbf{T}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{U}_k^{\mathrm{H}}$, $\mathbf{D}_k = \mathrm{diag}(\tau_{k1}, \ldots, \tau_{kn_k})$ diagonal, the capacity-achieving matrices $\mathbf{P}_{i_1}^\star, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}^\star$ satisfy
  1. $\mathbf{P}_k^\star = \mathbf{U}_k \mathbf{Q}_k^\star \mathbf{U}_k^{\mathrm{H}}$, with $\mathbf{Q}_k^\star$ diagonal; i.e. the eigenspace of $\mathbf{P}_k^\star$ is the same as the eigenspace of $\mathbf{T}_k$.
  2. with $\bar{e}_k^\star = \bar{e}_k(-\sigma^2, \mathbf{P}_k^\star)$, the i^{th} entry $q_{ki}^\star$ of $\mathbf{Q}_k^\star$ satisfies

$$q_{ki}^\star = \left(\mu_k - \frac{1}{c_k e_k^\star \tau_{ki}} \right)^+$$

  where the μ_k's are evaluated such that $\mathrm{tr}(\mathbf{Q}_k) = P_k$.
- an iterative water-filling method allows to retrieve the $q_{ki}^\star$'s by successively
  - for a given set $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$, evaluating $e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}$
  - updating the new optimal solution $\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}$ for this system

Iterative water-filling

Upon convergence, the iterative water-filling algorithm leads to the optimal solution.

Proof of water-filling optimality

- Consider the functions

$$C(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}}) =$$

$$\sum_{k\in\mathcal{S}} \frac{1}{N}\log\det\left(\mathbf{I}_{n_k} + c_k e_k \mathbf{T}_k \mathbf{P}_k\right) + \frac{1}{N}\log\det\left(\mathbf{I}_N + \sum_{k\in\mathcal{S}} \bar{e}_k \mathbf{R}_k\right) - \sigma^2 \sum_{k\in\mathcal{S}} \bar{e}_k(-\sigma^2) e_k(-\sigma^2)$$

where

$$e_i = e_i(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}}) = \frac{1}{N}\operatorname{tr}\mathbf{T}_i\left(\sigma^2\left[\mathbf{I}_N + \sum_{k\in\mathcal{S}} \bar{e}_k \mathbf{T}_k\right]\right)^{-1}$$

$$\bar{e}_i = \bar{e}_i(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}}) = \frac{1}{n_i}\operatorname{tr}\mathbf{R}_i\mathbf{P}_i\left(\sigma^2\left[\mathbf{I}_{n_i} + c_i e_i(z)\mathbf{R}_i\mathbf{P}_i\right]\right)^{-1}$$

and $V : (\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}}, \bar{e}_{i_1},\ldots,\bar{e}_{i_{|\mathcal{S}|}}, e_{i_1},\ldots,e_{i_{|\mathcal{S}|}}) \mapsto C(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}})$.

- From chain rule,

$$\frac{\partial}{\partial \mathbf{P}_i} C = \sum_{k\in\mathcal{S}} \frac{\partial V}{\partial e_k}\frac{\partial e_k}{\partial \mathbf{P}_i} + \frac{\partial V}{\partial \bar{e}_k}\frac{\partial \bar{e}_k}{\partial \mathbf{P}_i} + \frac{\partial V}{\partial \mathbf{P}_i}$$

Proof of water-filling optimality

- Consider the functions

$$C(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}}) =$$

$$\sum_{k\in\mathcal{S}} \frac{1}{N}\log\det\left(\mathbf{I}_{n_k} + c_k e_k \mathbf{T}_k \mathbf{P}_k\right) + \frac{1}{N}\log\det\left(\mathbf{I}_N + \sum_{k\in\mathcal{S}} \bar{e}_k \mathbf{R}_k\right) - \sigma^2 \sum_{k\in\mathcal{S}} \bar{e}_k(-\sigma^2) e_k(-\sigma^2)$$

where

$$e_i = e_i(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}}) = \frac{1}{N}\operatorname{tr}\mathbf{T}_i\left(\sigma^2\left[\mathbf{I}_N + \sum_{k\in\mathcal{S}} \bar{e}_k \mathbf{T}_k\right]\right)^{-1}$$

$$\bar{e}_i = \bar{e}_i(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}}) = \frac{1}{n_i}\operatorname{tr}\mathbf{R}_i\mathbf{P}_i\left(\sigma^2\left[\mathbf{I}_{n_i} + c_i e_i(z)\mathbf{R}_i\mathbf{P}_i\right]\right)^{-1}$$

and $V : (\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}},\bar{e}_{i_1},\ldots,\bar{e}_{i_{|\mathcal{S}|}},e_{i_1},\ldots,e_{i_{|\mathcal{S}|}}) \mapsto C(\mathbf{P}_{i_1},\ldots,\mathbf{P}_{i_{|\mathcal{S}|}})$.

- From chain rule,

$$\frac{\partial}{\partial \mathbf{P}_i} C = \sum_{k\in\mathcal{S}} \frac{\partial V}{\partial e_k}\frac{\partial e_k}{\partial \mathbf{P}_i} + \frac{\partial V}{\partial \bar{e}_k}\frac{\partial \bar{e}_k}{\partial \mathbf{P}_i} + \frac{\partial V}{\partial \mathbf{P}_i}$$

Proof of water-filling optimality (2)

- Remark that

$$\frac{\partial}{\partial \bar{e}_k} V(\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}, \bar{e}_{i_1}, \ldots, \bar{e}_{i_{|\mathcal{S}|}}, e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}) = \frac{1}{N} \operatorname{tr}\left[\left(\mathbf{I} + \sum_{i \in \mathcal{S}} \bar{e}_i \mathbf{R}_i\right)^{-1} \mathbf{R}_k\right] - \sigma^2 e_k$$

$$\frac{\partial}{\partial e_k} V(\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}, \bar{e}_{i_1}, \ldots, \bar{e}_{i_{|\mathcal{S}|}}, e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}) = c_k \frac{1}{N} \operatorname{tr}\left[(\mathbf{I} + c_k e_k \mathbf{T}_i \mathbf{P}_i)^{-1} \mathbf{T}_k \mathbf{P}_k\right] - \sigma^2 \bar{e}_k$$

  both being null whenever, for all k, $e_k = e_k(-\sigma^2, \mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}})$ and $\bar{e}_k = \bar{e}_k(-\sigma^2, \mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}})$, which is true in particular for the unique power optimal solution $\mathbf{P}^\star_{i_1}, \ldots, \mathbf{P}^\star_{i_{|\mathcal{S}|}}$ whenever $e_k = e_k^\star$ and $\bar{e}_k = \bar{e}_k^\star$.
- When, for all k, $e_k = e_k^\star$, $\bar{e}_k = \bar{e}_k^\star$, the maximum of V over the $\mathbf{P}_k$'s is then obtained by maximizing the expressions $\log\det(\mathbf{I}_{n_k} + c_k e_k^\star \mathbf{T}_k \mathbf{P}_k)$ over $\mathbf{P}_k$.

Proof of water-filling optimality (2)

- Remark that

$$\frac{\partial}{\partial \bar{e}_k} V(\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}, \bar{e}_{i_1}, \ldots, \bar{e}_{i_{|\mathcal{S}|}}, e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}) = \frac{1}{N} \operatorname{tr} \left[\left(\mathbf{I} + \sum_{i \in \mathcal{S}} \bar{e}_i \mathbf{R}_i \right)^{-1} \mathbf{R}_k \right] - \sigma^2 e_k$$

$$\frac{\partial}{\partial e_k} V(\mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}}, \bar{e}_{i_1}, \ldots, \bar{e}_{i_{|\mathcal{S}|}}, e_{i_1}, \ldots, e_{i_{|\mathcal{S}|}}) = c_k \frac{1}{N} \operatorname{tr} \left[(\mathbf{I} + c_k e_k \mathbf{T}_i \mathbf{P}_i)^{-1} \mathbf{T}_k \mathbf{P}_k \right] - \sigma^2 \bar{e}_k$$

  both being null whenever, for all k, $e_k = e_k(-\sigma^2, \mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}})$ and $\bar{e}_k = \bar{e}_k(-\sigma^2, \mathbf{P}_{i_1}, \ldots, \mathbf{P}_{i_{|\mathcal{S}|}})$, which is true in particular for the unique power optimal solution $\mathbf{P}^\star_{i_1}, \ldots, \mathbf{P}^\star_{i_{|\mathcal{S}|}}$ whenever $e_k = e_k^\star$ and $\bar{e}_k = \bar{e}_k^\star$.
- When, for all k, $e_k = e_k^\star$, $\bar{e}_k = \bar{e}_k^\star$, the maximum of V over the $\mathbf{P}_k$'s is then obtained by maximizing the expressions $\log\det(\mathbf{I}_{n_k} + c_k e_k^\star \mathbf{T}_k \mathbf{P}_k)$ over $\mathbf{P}_k$.

About this result

Some consequences of the previous results are worth mentioning

- deterministic equivalents do not impose any underlying convergence
- truncation and centralization lead to stronger convergence results under the form $m_N - m_N^\circ \xrightarrow{\text{a.s.}} 0$ instead of $\mathrm{E}m_N - m_N^\circ \to 0$
- loose hypotheses on the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s: strong antenna correlation allowed
- the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s are general purpose Hermitian nonnegative, no need of a common eigenspace
- no restriction to Gaussian $\mathbf{X}_k$'s for diagonal $\mathbf{T}_k$'s

Compact expressions

Only K scalar parameters (the e_k's) determine the behaviour of the whole system.

About this result

Some consequences of the previous results are worth mentioning

- deterministic equivalents do not impose any underlying convergence
- truncation and centralization lead to stronger convergence results under the form $m_N - m_N^\circ \xrightarrow{\text{a.s.}} 0$ instead of $\mathrm{E}m_N - m_N^\circ \to 0$
- loose hypotheses on the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s: strong antenna correlation allowed
- the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s are general purpose Hermitian nonnegative, no need of a common eigenspace
- no restriction to Gaussian $\mathbf{X}_k$'s for diagonal $\mathbf{T}_k$'s

Compact expressions

Only K scalar parameters (the e_k's) determine the behaviour of the whole system.

About this result

Some consequences of the previous results are worth mentioning

- deterministic equivalents do not impose any underlying convergence
- truncation and centralization lead to stronger convergence results under the form $m_N - m_N^\circ \xrightarrow{\text{a.s.}} 0$ instead of $\mathrm{E}m_N - m_N^\circ \to 0$
- loose hypotheses on the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s: strong antenna correlation allowed
- the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s are general purpose Hermitian nonnegative, no need of a common eigenspace
- no restriction to Gaussian $\mathbf{X}_k$'s for diagonal $\mathbf{T}_k$'s

Compact expressions

Only K scalar parameters (the e_k's) determine the behaviour of the whole system.

About this result

Some consequences of the previous results are worth mentioning

- deterministic equivalents do not impose any underlying convergence
- truncation and centralization lead to stronger convergence results under the form $m_N - m_N^\circ \xrightarrow{\text{a.s.}} 0$ instead of $\mathrm{E} m_N - m_N^\circ \to 0$
- loose hypotheses on the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s: strong antenna correlation allowed
- the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s are general purpose Hermitian nonnegative, no need of a common eigenspace
- no restriction to Gaussian $\mathbf{X}_k$'s for diagonal $\mathbf{T}_k$'s

Compact expressions

Only K scalar parameters (the e_k's) determine the behaviour of the whole system.

About this result

Some consequences of the previous results are worth mentioning

- deterministic equivalents do not impose any underlying convergence
- truncation and centralization lead to stronger convergence results under the form $m_N - m_N^\circ \xrightarrow{\text{a.s.}} 0$ instead of $\mathrm{E} m_N - m_N^\circ \to 0$
- loose hypotheses on the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s: strong antenna correlation allowed
- the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s are general purpose Hermitian nonnegative, no need of a common eigenspace
- no restriction to Gaussian $\mathbf{X}_k$'s for diagonal $\mathbf{T}_k$'s

Compact expressions

Only K scalar parameters (the e_k's) determine the behaviour of the whole system.

About this result

Some consequences of the previous results are worth mentioning

- deterministic equivalents do not impose any underlying convergence
- truncation and centralization lead to stronger convergence results under the form $m_N - m_N^\circ \xrightarrow{\text{a.s.}} 0$ instead of $\mathrm{E} m_N - m_N^\circ \to 0$
- loose hypotheses on the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s: strong antenna correlation allowed
- the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s are general purpose Hermitian nonnegative, no need of a common eigenspace
- no restriction to Gaussian $\mathbf{X}_k$'s for diagonal $\mathbf{T}_k$'s

Compact expressions

Only K scalar parameters (the e_k's) determine the behaviour of the whole system.

About this result

Some consequences of the previous results are worth mentioning

- deterministic equivalents do not impose any underlying convergence
- truncation and centralization lead to stronger convergence results under the form $m_N - m_N^\circ \xrightarrow{\text{a.s.}} 0$ instead of $\mathrm{E}m_N - m_N^\circ \to 0$
- loose hypotheses on the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s: strong antenna correlation allowed
- the $\mathbf{R}_k$'s and $\mathbf{T}_k$'s are general purpose Hermitian nonnegative, no need of a common eigenspace
- no restriction to Gaussian $\mathbf{X}_k$'s for diagonal $\mathbf{T}_k$'s

Compact expressions

Only K scalar parameters (the e_k's) determine the behaviour of the whole system.

Performance of the deterministic equivalent

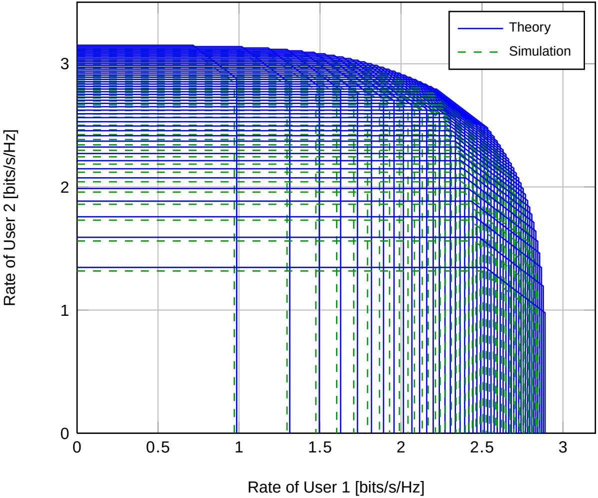

Figure: (Per-antenna) rate region $\mathbf{C}_{\mathrm{BC}}$ for $K = 2$ users, theory against simulation, $N = 8$, $n_1 = n_2 = 4$, $\mathrm{SNR} = 20$ dB, random transmit-receive solid angle of aperture $\pi/2$, $d_{\mathrm{T}}/\lambda = 10$, $d_{\mathrm{R}}/\lambda = 1/4$.

Performance of the deterministic equivalent (2)

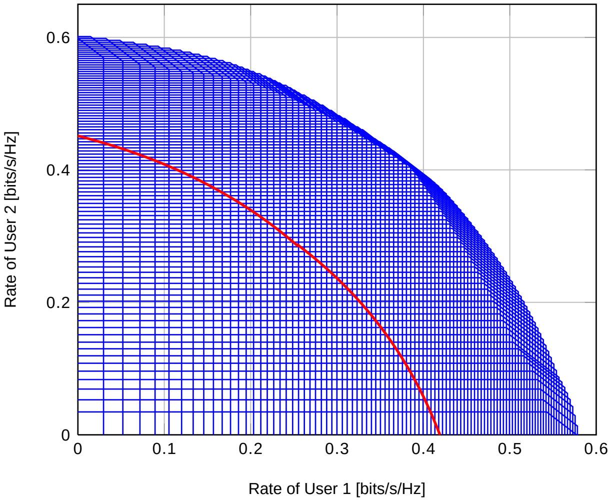

Figure: (Per-antenna) rate region $\mathcal{C}_{\rm BC}$ for $K = 2$ users, $N = 8$, $n_1 = n_2 = 4$, $\mathrm{SNR} = -5$ dB, random transmit-receive solid angle of aperture $\pi/2$, $d_{\rm T}/\lambda = 10$, $d_{\rm R}/\lambda = 1/4$. In thick line, capacity limit when $\mathbf{E}[\mathbf{s}\mathbf{s}^{\rm H}] = \mathbf{I}_N$.

Other results using deterministic equivalents

R. Couillet, S. Wagner, M. Debbah, D. Slock, “Asymptotic analysis of linear precoding in vector broadcast channels with limited feedback”

Deterministic equivalents of sum-rate capacity for linearly precoded broadcast channels,

- accounting for base station antenna correlation, user path losses
- assuming limited channel state information

Results:

- on optimal number of users to serve
- on optimal regularization parameter
- eventually, optimal feedback time
- close behaviour with respect to finite size systems for $N \geq 4$

Other results using deterministic equivalents

R. Couillet, S. Wagner, M. Debbah, D. Slock, “Asymptotic analysis of linear precoding in vector broadcast channels with limited feedback”

Deterministic equivalents of sum-rate capacity for linearly precoded broadcast channels,

- accounting for base station antenna correlation, user path losses
- assuming limited channel state information

Results:

- on optimal number of users to serve
- on optimal regularization parameter
- eventually, optimal feedback time
- close behaviour with respect to finite size systems for $N \geq 4$

Other results using deterministic equivalents

R. Couillet, S. Wagner, M. Debbah, D. Slock, “Asymptotic analysis of linear precoding in vector broadcast channels with limited feedback”

Deterministic equivalents of sum-rate capacity for linearly precoded broadcast channels,

- accounting for base station antenna correlation, user path losses
- assuming limited channel state information

Results:

- on optimal number of users to serve
- on optimal regularization parameter
- eventually, optimal feedback time
- close behaviour with respect to finite size systems for $N \geq 4$

Linearly precoded broadcast channels with imperfect CSI

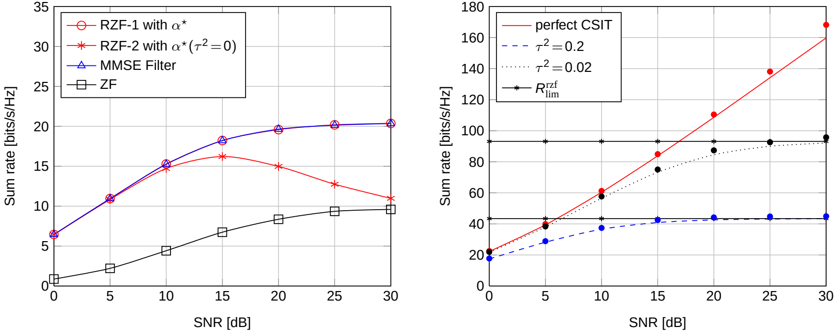

Figure: **Left**: Ergodic sum-rate vs. average SNR with $\mathbf{R}=\mathbf{I}_M$, $\mathbf{L}=\mathbf{I}_K$, $M=10$, $\beta=1$, $\tau^2=0.1$. **Right**: RZF, $\mathbf{R}=\mathbf{I}_M$, $\mathbf{L}=\mathbf{I}_K$, $M=32$, $\beta=1$, simulation results are indicated by circle marks

Outline

Outline

1. Stieltjes transform methods for more elaborate models
2. Kronecker models and Variance Profiles
3. Capacity expressions, Rate Regions
4. Touching the boundary: optimal power allocation
5. Case study: exchanging relevant data in large self-organized networks
   - Orthogonal CDMA networks
   - Spectrum sharing in multiple access channels

Introduction of the self-organized network

Before to apply the previous results, we consider first an alternative, simpler, better adapted model, which

- provides a deterministic equivalent to a model involving Haar (unitary) matrices
- uses R-, S- and η-transforms
- is a striking example of the feedback minimization discussed before.

Introduction of the self-organized network

Before to apply the previous results, we consider first an alternative, simpler, better adapted model, which

- provides a deterministic equivalent to a model involving Haar (unitary) matrices
- uses R-, S- and η-transforms
- is a striking example of the feedback minimization discussed before.

Self-Organized Clustered Networks

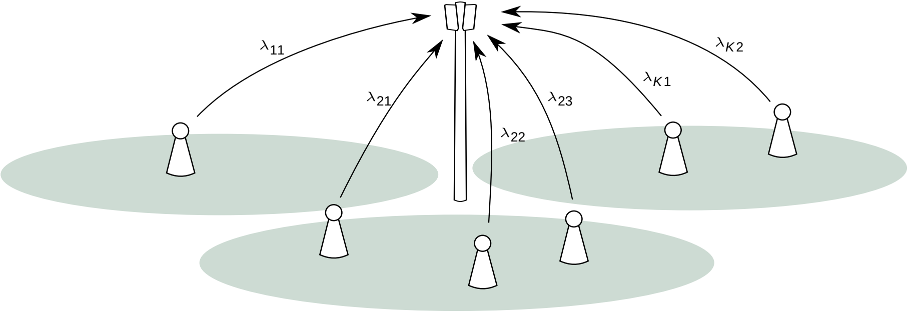

Figure: Self-organizing CDMA network

Self-Organized Clustered Networks

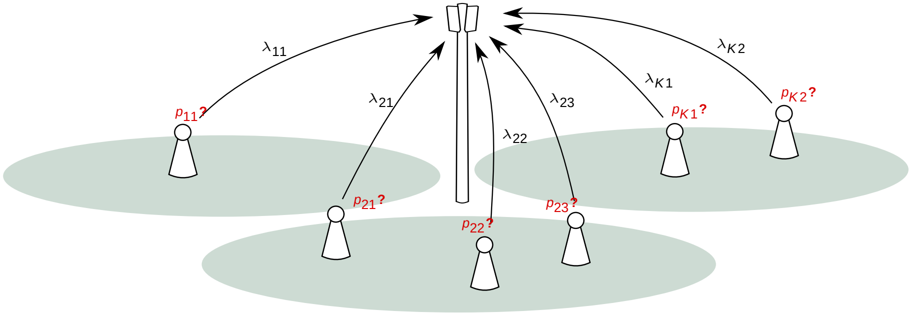

Figure: Self-organizing CDMA network

Uplink clustered CDMA networks

- Consider a set of K clusters, all using independently orthogonal CDMA transmissions. Each cluster is composed of *at most* N users. We wish
  - to obtain a deterministic equivalent for the achievable uplink sum-rate
  - to provide a cheap feedback solution for the network to organize itself to collectively maximize the uplink rate.
- We denote
  - $\mathbf{L}_k = \mathrm{diag}(\lambda_{k1}, \ldots, \lambda_{kN})$ the diagonal of channel gains (inverse path losses).
  - $\mathbf{P}_k = \mathrm{diag}(p_{k1}, \ldots, p_{kN})$ the diagonal of transmit powers from the users in cell k.
  - $\mathbf{W}_k \in \mathbb{C}^{N \times N}$ the unitary CDMA code matrix used in cell k.
  - the received signal $\mathbf{y} \in \mathbb{C}^N$ at the base station reads

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{W}_k \mathbf{L}_k^{\frac{1}{2}} \mathbf{P}_k^{\frac{1}{2}} \mathbf{s}_k + \mathbf{n}$$

  - the sum-rate $C(\sigma^2)$ is

$$C(\sigma^2) = \frac{1}{N} \log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2} \sum_{k=1}^{K} \mathbf{W}_k (\mathbf{P}_k \mathbf{L}_k) \mathbf{W}_k^{\mathsf{H}}\right)$$

Uplink clustered CDMA networks

- Consider a set of K clusters, all using independently orthogonal CDMA transmissions. Each cluster is composed of *at most* N users. We wish
  - to obtain a deterministic equivalent for the achievable uplink sum-rate
  - to provide a cheap feedback solution for the network to organize itself to collectively maximize the uplink rate.
- We denote
  - $\mathbf{L}_k = \operatorname{diag}(\lambda_{k1}, \ldots, \lambda_{kN})$ the diagonal of channel gains (inverse path losses).
  - $\mathbf{P}_k = \operatorname{diag}(p_{k1}, \ldots, p_{kN})$ the diagonal of transmit powers from the users in cell k.
  - $\mathbf{W}_k \in \mathbb{C}^{N \times N}$ the unitary CDMA code matrix used in cell k.
  - the received signal $\mathbf{y} \in \mathbb{C}^N$ at the base station reads

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{W}_k \mathbf{L}_k^{\frac{1}{2}} \mathbf{P}_k^{\frac{1}{2}} \mathbf{s}_k + \mathbf{n}$$

  - the sum-rate $C(\sigma^2)$ is

$$C(\sigma^2) = \frac{1}{N} \log\det \left(\mathbf{I}_N + \frac{1}{\sigma^2} \sum_{k=1}^{K} \mathbf{W}_k (\mathbf{P}_k \mathbf{L}_k) \mathbf{W}_k^{\mathsf{H}} \right)$$

Capacity expression

R. Couillet, M. Debbah, “Uplink capacity of self-organizing clustered orthogonal CDMA networks in flat fading channels”, ITW 2009 Fall, Taormina, Sicily.

Theorem

For large N, we have

$$C_N(\sigma^2) - C_N^\circ(\sigma^2) \to 0$$

with

$$C_N^\circ(\sigma^2) = \log\left(1 + \frac{1}{\sigma^2}\sum_{k=1}^{K}\beta_k\right) + \sum_{k=1}^{K}\frac{1}{N}\log\det\left(\frac{\eta}{\sigma^2}\mathbf{P}_k\mathbf{L}_k + \left[1 - \frac{\eta\beta_k}{\sigma^2}\right]\mathbf{I}_N\right)$$

where β_k and η are defined as

$$\eta = \left(1 + \frac{1}{\sigma^2}\sum_{i=1}^{K}\beta_i\right)^{-1}$$

and $\{\beta_k\}$ are solutions of

$$\beta_k = \frac{1}{N}\operatorname{tr}\mathbf{P}_k\mathbf{L}_k\left(\frac{\eta}{\sigma^2}\mathbf{P}_k\mathbf{L}_k + \left[1 - \frac{\eta\beta_k}{\sigma^2}\right]\mathbf{I}_N\right)^{-1}$$

Capacity expression

R. Couillet, M. Debbah, “Uplink capacity of self-organizing clustered orthogonal CDMA networks in flat fading channels”, ITW 2009 Fall, Taormina, Sicily.

Theorem

For large N, we have

$$C_N(\sigma^2) - C_N^\circ(\sigma^2) \to 0$$

with

$$C_N^\circ(\sigma^2) = \log\left(1 + \frac{1}{\sigma^2}\sum_{k=1}^{K}\beta_k\right) + \sum_{k=1}^{K}\frac{1}{N}\log\det\left(\frac{\eta}{\sigma^2}\mathbf{P}_k\mathbf{L}_k + \left[1 - \frac{\eta\beta_k}{\sigma^2}\right]\mathbf{I}_N\right)$$

where β_k and η are defined as

$$\eta = \left(1 + \frac{1}{\sigma^2}\sum_{i=1}^{K}\beta_i\right)^{-1}$$

and $\{\beta_k\}$ are solutions of

$$\beta_k = \frac{1}{N}\operatorname{tr}\mathbf{P}_k\mathbf{L}_k\left(\frac{\eta}{\sigma^2}\mathbf{P}_k\mathbf{L}_k + \left[1 - \frac{\eta\beta_k}{\sigma^2}\right]\mathbf{I}_N\right)^{-1}$$

Capacity expression

R. Couillet, M. Debbah, "Uplink capacity of self-organizing clustered orthogonal CDMA networks in flat fading channels", ITW 2009 Fall, Taormina, Sicily.

Theorem

For large N, we have

$$C_N(\sigma^2) - C_N^\circ(\sigma^2) \to 0$$

with

$$C_N^\circ(\sigma^2) = \log\left(1 + \frac{1}{\sigma^2}\sum_{k=1}^{K}\beta_k\right) + \sum_{k=1}^{K}\frac{1}{N}\log\det\left(\frac{\eta}{\sigma^2}\mathbf{P}_k\mathbf{L}_k + \left[1 - \frac{\eta\beta_k}{\sigma^2}\right]\mathbf{I}_N\right)$$

where β_k and η are defined as

$$\eta = \left(1 + \frac{1}{\sigma^2}\sum_{i=1}^{K}\beta_i\right)^{-1}$$

and $\{\beta_k\}$ are solutions of

$$\beta_k = \frac{1}{N}\operatorname{tr}\mathbf{P}_k\mathbf{L}_k\left(\frac{\eta}{\sigma^2}\mathbf{P}_k\mathbf{L}_k + \left[1 - \frac{\eta\beta_k}{\sigma^2}\right]\mathbf{I}_N\right)^{-1}$$

Proof

Instead of working with the Stieltjes transform, we use the (totally equivalent) η-transform. We define $\eta_1, \ldots, \eta_K$ as

$$\eta_k(x) = \int \frac{1}{1+xt}\mu_k(dt)$$

with μ_k the probability distribution of $\mathbf{P}_k\mathbf{L}_k$. We will use the R-transform for further development. For each k, denote R_k the R-transform of $\mathbf{W}_k\mathbf{L}_k\mathbf{P}_k\mathbf{W}_k^{\mathrm{H}}$, defined as

$$\eta(-\frac{1}{R(x)+\frac{1}{x}}) = xR(x)+1$$

Since the $\mathbf{W}_k$'s are isometric and independent, they are free random variables. Hence, the R-transform $R(x)$ of the sum of the individual R-transforms $R_1(x), \ldots, R_K(x)$ satisfiesi asymptotically

$$R(x) = \sum_{k=1}^{K} R_k(x)$$

The strategy is then to use the R-transform as a "pivot" in the proof,

- obtain a relation of R_k as a function of the entries of $\mathbf{P}_k\mathbf{L}_k$
- obtain an expression of the eigenvalues of $\sum_{k=1}^{K}\mathbf{W}_k\mathbf{P}_k\mathbf{L}_k\mathbf{W}_k^{\mathrm{H}}$ as a function of R

The first relation is obtained by the definition of the η-transform applied in $-1/(R_k(x)+\frac{1}{x})$

$$xR_k+1 = \int \frac{1}{1-\frac{t}{R_k(x)+\frac{1}{x}}}\mu_k(dt)$$

Proof

Instead of working with the Stieltjes transform, we use the (totally equivalent) η-transform. We define $\eta_1, \ldots, \eta_K$ as

$$\eta_k(x) = \int \frac{1}{1+xt} \mu_k(dt)$$

with μ_k the probability distribution of $\mathbf{P}_k\mathbf{L}_k$. We will use the R-transform for further development. For each k, denote R_k the R-transform of $\mathbf{W}_k\mathbf{L}_k\mathbf{P}_k\mathbf{W}_k^{\mathrm{H}}$, defined as

$$\eta(-\frac{1}{R(x)+\frac{1}{x}}) = xR(x) + 1$$

Since the $\mathbf{W}_k$'s are isometric and independent, they are free random variables. Hence, the R-transform $R(x)$ of the sum of the individual R-transforms $R_1(x), \ldots, R_K(x)$ satisfiesi asymptotically

$$R(x) = \sum_{k=1}^{K} R_k(x)$$

The strategy is then to use the R-transform as a "pivot" in the proof,

- obtain a relation of R_k as a function of the entries of $\mathbf{P}_k\mathbf{L}_k$
- obtain an expression of the eigenvalues of $\sum_{k=1}^{K} \mathbf{W}_k\mathbf{P}_k\mathbf{L}_k\mathbf{W}_k^{\mathrm{H}}$ as a function of R

The first relation is obtained by the definition of the η-transform applied in $-1/(R_k(x)+\frac{1}{x})$

$$xR_k + 1 = \int \frac{1}{1 - \frac{t}{R_k(x)+\frac{1}{x}}} \mu_k(dt)$$

Proof

Instead of working with the Stieltjes transform, we use the (totally equivalent) η-transform. We define $\eta_1, \ldots, \eta_K$ as

$$\eta_k(x) = \int \frac{1}{1+xt}\mu_k(dt)$$

with μ_k the probability distribution of $\mathbf{P}_k\mathbf{L}_k$. We will use the R-transform for further development. For each k, denote R_k the R-transform of $\mathbf{W}_k\mathbf{L}_k\mathbf{P}_k\mathbf{W}_k^{\mathrm{H}}$, defined as

$$\eta(-\frac{1}{R(x)+\frac{1}{x}}) = xR(x)+1$$

Since the $\mathbf{W}_k$'s are isometric and independent, they are free random variables. Hence, the R-transform $R(x)$ of the sum of the individual R-transforms $R_1(x), \ldots, R_K(x)$ satisfiesi asymptotically

$$R(x) = \sum_{k=1}^{K} R_k(x)$$

The strategy is then to use the R-transform as a “pivot” in the proof,

- obtain a relation of R_k as a function of the entries of $\mathbf{P}_k\mathbf{L}_k$
- obtain an expression of the eigenvalues of $\sum_{k=1}^{K}\mathbf{W}_k\mathbf{P}_k\mathbf{L}_k\mathbf{W}_k^{\mathrm{H}}$ as a function of R

The first relation is obtained by the definition of the η-transform applied in $-1/(R_k(x)+\frac{1}{x})$

$$xR_k+1 = \int \frac{1}{1-\frac{t}{R_k(x)+\frac{1}{x}}}\mu_k(dt)$$

Proof (2)

The expression

$$xR_k(x) + 1 = \int \frac{1}{1 - \frac{t}{R_k(x) + \frac{1}{x}}} \mu_k(dt)$$

leads to

$$R_k(x) = \frac{1}{x} \int \frac{t}{R_k(x) + \frac{1}{x} - t} \mu_k(dt)$$

and, in particular, defining $\beta_k(x) = R_k(-x\eta)$, we have

$$\beta_k(x) = \int \frac{t}{1 - x\eta\beta_k + x\eta t} \mu_k(dt)$$

Now, since $R(x) = \sum_{k=1}^{K} R_k(x)$ asymptotically on N, using the reverse definition of the R-transform

$$R(-x\eta(x)) = -\frac{1}{x}(1 - \frac{1}{\eta}(x))$$

we have

$$\eta(x) = \left(1 + x \sum_{k=1}^{K} R_k(-x\eta)\right)^{-1} = \left(1 + x \sum_{k=1}^{K} \beta_k\right)^{-1}$$

which completes the proof.

Optimal power allocation

The power allocation policy $p_{kn} = p^\star_{kn}$ optimizing the deterministic approximation of $C(\sigma^2)$ satisfies, for all k, n,

$$p^\star_{kn} = \left(\alpha_k - \frac{\sigma^2 - \eta^\star \beta^\star_k}{\lambda_{kn}\eta^\star} \right)^+$$

where $\eta^\star$, $\beta^\star_k$ are the respective values of η and β_k when C achieves its maximum, and α_k is such that $\sum_k p^\star_{kn} = P_k$.

Lemma *(Iterative Water-filling)*

Upon convergence, the following algorithm converges to the optimal power allocation policy,

At initialization, for all k, $p_{kn} = \frac{P_k}{N}$, $\eta = 1$, $\beta_k = 1$.
while the p_{kn}'s have not converged **do**
  for $k \in \{1, \ldots, K\}$ **do**
    Solve fixed-point equation for (η, β_k), p_{kn} fixed
    for $n = 1 \ldots, N$ **do**
      Set $p_{kn} = \left(\alpha_k - \frac{\sigma^2 - \eta\beta_k}{\lambda_{kn}\eta} \right)^+$, with α_k such that $\sum_n p_{kn} = P_k$.
    end for
  end for
end while

Optimal power allocation

The power allocation policy $p_{kn} = p_{kn}^\star$ optimizing the deterministic approximation of $C(\sigma^2)$ satisfies, for all k, n,

$$p_{kn}^\star = \left(\alpha_k - \frac{\sigma^2 - \eta^\star \beta_k^\star}{\lambda_{kn} \eta^\star} \right)^+$$

where $\eta^\star$, $\beta_k^\star$ are the respective values of η and β_k when C achieves its maximum, and α_k is such that $\sum_k p_{kn}^\star = P_k$.

Lemma *(Iterative Water-filling)*

Upon convergence, the following algorithm converges to the optimal power allocation policy,

At initialization, for all k, $p_{kn} = \frac{P_k}{N}$, $\eta = 1$, $\beta_k = 1$.
while the p_{kn}'s have not converged **do**
  for $k \in \{1, \ldots, K\}$ **do**
    Solve fixed-point equation for (η, β_k), p_{kn} fixed
    for $n = 1 \ldots, N$ **do**
      Set $p_{kn} = \left(\alpha_k - \frac{\sigma^2 - \eta\beta_k}{\lambda_{kn}\eta} \right)^+$, with α_k such that $\sum_n p_{kn} = P_k$.
    end for
  end for
end while

Sequential feedback in the network

- *Local optimization:* From the formulas of η and β_k, at step (t) of the iterative water-filling, we can write
  - $\eta^{(t)}(x) = \left(\frac{1}{\eta^{(t-1)}(x)} + x(\beta_k^{(t)} - \beta_k^{(t-1)}) \right)^{-1}$
  - $\beta_k^{(t)} = f(\beta_k^{(t)}, \eta^{(t)})$

This is only dependent on k.

$\Rightarrow$ Cluster k does not need to know all λ_{in}, $i \neq k$.

- Iterative self-organized process The preceding algorithm can be rewritten such that,
  - at each time step (t), based on $\eta^{(t-1)}$, cell k performs self-optimization of $\mathbf{P}_k$ and updates $\eta^{(t-1)}$ to $\eta^{(t)}$
  - cell k forwards $\eta^{(t)}$ to next cell $(k+1)$
  - upon convergence (not proven), this proceeds until convergence to the optimal solution (proven)

Sequential feedback in the network

- *Local optimization:* From the formulas of η and β_k, at step (t) of the iterative water-filling, we can write
  - $\eta^{(t)}(x) = \left(\frac{1}{\eta^{(t-1)}(x)} + x(\beta_k^{(t)} - \beta_k^{(t-1)}) \right)^{-1}$
  - $\beta_k^{(t)} = f(\beta_k^{(t)}, \eta^{(t)})$

This is only dependent on k.

$\Rightarrow$ Cluster k does not need to know all λ_{in}, $i \neq k$.

- Iterative self-organized process The preceding algorithm can be rewritten such that,
  - at each time step (t), based on $\eta^{(t-1)}$, cell k performs self-optimization of $\mathbf{P}_k$ and updates $\eta^{(t-1)}$ to $\eta^{(t)}$
  - cell k forwards $\eta^{(t)}$ to next cell $(k+1)$
  - upon convergence (not proven), this proceeds until convergence to the optimal solution (proven)

Sequential feedback in the network

- *Local optimization:* From the formulas of η and β_k, at step (t) of the iterative water-filling, we can write
  - $\eta^{(t)}(x) = \left(\frac{1}{\eta^{(t-1)}(x)} + x(\beta_k^{(t)} - \beta_k^{(t-1)}) \right)^{-1}$
  - $\beta_k^{(t)} = f(\beta_k^{(t)}, \eta^{(t)})$

This is only dependent on k.

$\Rightarrow$ Cluster k does not need to know all λ_{in}, $i \neq k$.

- Iterative self-organized process The preceding algorithm can be rewritten such that,
  - at each time step (t), based on $\eta^{(t-1)}$, cell k performs self-optimization of $\mathbf{P}_k$ and updates $\eta^{(t-1)}$ to $\eta^{(t)}$
  - cell k forwards $\eta^{(t)}$ to next cell $(k+1)$
  - upon convergence (not proven), this proceeds until convergence to the optimal solution (proven)

Self-organization in orthogonal CDMA networks

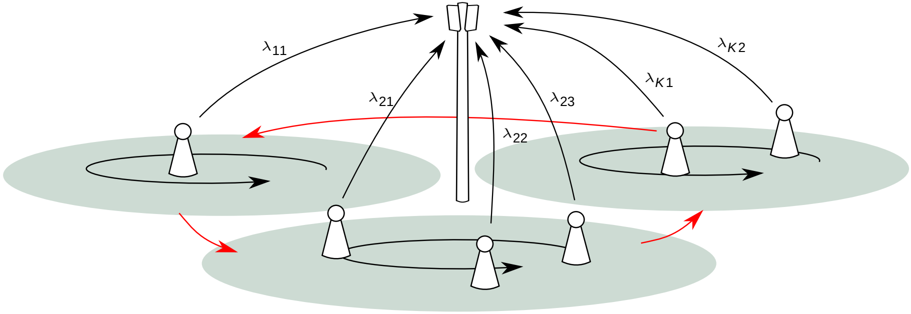

Figure: Self-organization in orthogonal CDMA network

Self-organization in orthogonal CDMA networks

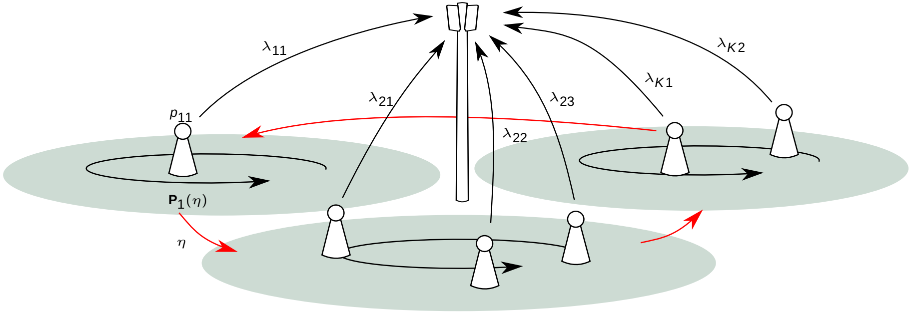

Figure: Self-organization in orthogonal CDMA network

Self-organization in orthogonal CDMA networks

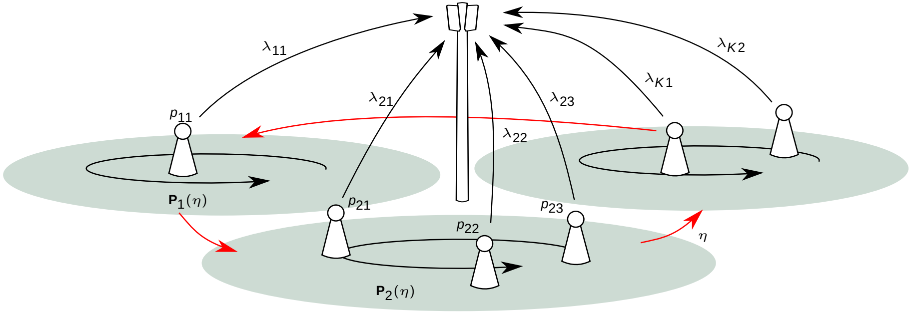

Figure: Self-organization in orthogonal CDMA network

Self-organization in orthogonal CDMA networks

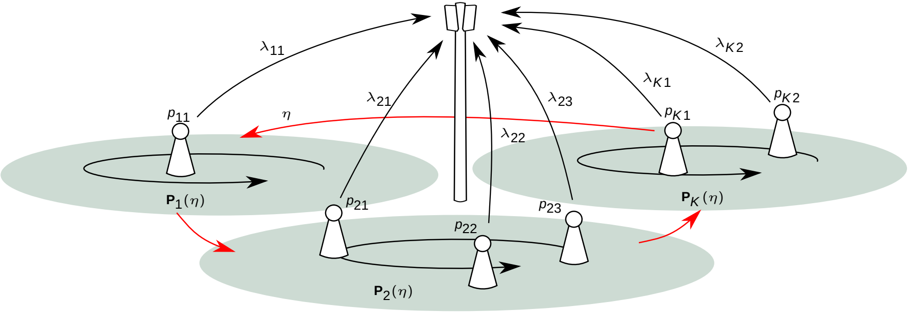

Figure: Self-organization in orthogonal CDMA network

Self-organization in orthogonal CDMA networks

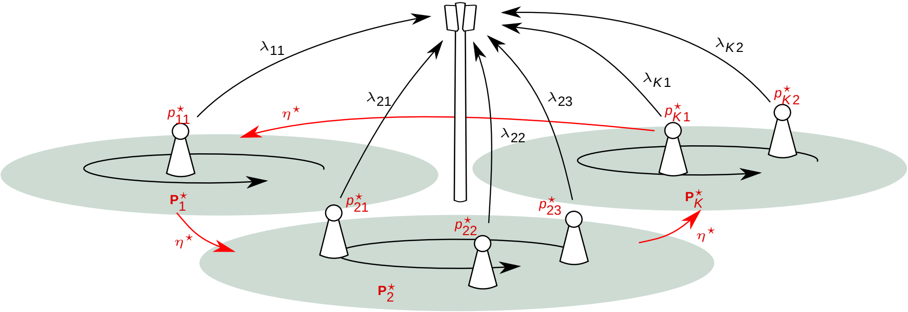

Figure: Self-organization in orthogonal CDMA network

Outline

Spectrum sharing in MIMO-MAC

- Somewhat similarly as η for the clustered CDMA system, user k of a multiple-access channel can find its optimal transmit covariance matrix from the estimation of e_k.
- if F frequency bands are shared among the users, the MAC rate region is the set of rates $R_1, \ldots, R_K$ such that, for any subset $\mathcal{K} \subset \{1, \ldots, K\}$,

$$\sum_{k \in \mathcal{K}} R_k \leq \frac{1}{N} \sum_{f=1}^{F} \log\det \left(\mathbf{I}_N + \frac{1}{\sigma^2} \sum_{k \in \mathcal{K}} \mathbf{H}_{k,f}^{\mathrm{H}} \mathbf{P}_{k,f} \mathbf{H}_{k,f} \right)$$

- the optimal $\mathbf{P}_{k,f}$'s have eigenvectors aligned to the transmit correlation matrix and eigenvectors $q_{k,f,1}, \ldots, q_{k,f,n_k}$ given by

$$q_{k,f,i} = \left(\mu_k - \frac{1}{c_k e_{k,f} t_{k,f,i}} \right)^+$$

with

$$\begin{cases} e_{k,f} &= \frac{1}{N} \operatorname{tr} \mathbf{R}_{k,f} \left(\sigma^2 \left[\mathbf{I}_N + \sum_{k' \in \mathcal{K}} \delta_{k',f} \mathbf{R}_{k',f} \right] \right)^{-1} \\ \bar{e}_{k,f} &= \frac{1}{n_k} \operatorname{tr} \mathbf{T}_{k,f} \left(\sigma^2 \left[\mathbf{I}_{n_k} + c_k e_{k,f} \mathbf{P}_{k,f} \mathbf{T}_{k,f} \right] \right)^{-1} . \end{cases}$$

- Iterative water-filling is still optimal in this case.

Spectrum sharing in MIMO-MAC

- Somewhat similarly as η for the clustered CDMA system, user k of a multiple-access channel can find its optimal transmit covariance matrix from the estimation of e_k.
- if F frequency bands are shared among the users, the MAC rate region is the set of rates $R_1, \ldots, R_K$ such that, for any subset $\mathcal{K} \subset \{1, \ldots, K\}$,

$$\sum_{k \in \mathcal{K}} R_k \leq \frac{1}{N} \sum_{f=1}^{F} \log\det \left(\mathbf{I}_N + \frac{1}{\sigma^2} \sum_{k \in \mathcal{K}} \mathbf{H}_{k,f}^{\mathrm{H}} \mathbf{P}_{k,f} \mathbf{H}_{k,f} \right)$$

- the optimal $\mathbf{P}_{k,f}$'s have eigenvectors aligned to the transmit correlation matrix and eigenvectors $q_{k,f,1}, \ldots, q_{k,f,n_k}$ given by

$$q_{k,f,i} = \left(\mu_k - \frac{1}{c_k e_{k,f} t_{k,f,i}} \right)^+$$

with

$$\begin{cases} e_{k,f} & = \frac{1}{N} \operatorname{tr} \mathbf{R}_{k,f} \left(\sigma^2 \left[\mathbf{I}_N + \sum_{k' \in \mathcal{K}} \delta_{k',f} \mathbf{R}_{k',f} \right] \right)^{-1} \\ \bar{e}_{k,f} & = \frac{1}{n_k} \operatorname{tr} \mathbf{T}_{k,f} \left(\sigma^2 \left[\mathbf{I}_{n_k} + c_k e_{k,f} \mathbf{P}_{k,f} \mathbf{T}_{k,f} \right] \right)^{-1}. \end{cases}$$

- Iterative water-filling is still optimal in this case.

Spectrum sharing, alternative approaches

- Classical ways to share spectrum,
  - via central entity: may be onerous and/or not possible
  - game theoretical considerations: may fall in bad Nash equilibrium
- Through random matrix theory approaches, it seems that the fundamental system parameters naturally appear. In this case,
  - for given $e_{k,1}, \ldots, e_{k,F}$, user k can evaluate $\bar{e}_{k,1}, \ldots, \bar{e}_{k,f}$ and optimize $\mathbf{P}_{k,1}, \ldots, \mathbf{P}_{k,F}$
  - for given $\bar{e}_{k,1}, \ldots, \bar{e}_{k,F}$, the base station can evaluate $e_{k,1}, \ldots, e_{k,f}$

Spectrum sharing, alternative approaches

- Classical ways to share spectrum,
  - via central entity: may be onerous and/or not possible
  - game theoretical considerations: may fall in bad Nash equilibrium
- Through random matrix theory approaches, it seems that the fundamental system parameters naturally appear. In this case,
  - for given $e_{k,1}, \ldots, e_{k,F}$, user k can evaluate $\bar{e}_{k,1}, \ldots, \bar{e}_{k,f}$ and optimize $\mathbf{P}_{k,1}, \ldots, \mathbf{P}_{k,F}$
  - for given $\bar{e}_{k,1}, \ldots, \bar{e}_{k,F}$, the base station can evaluate $e_{k,1}, \ldots, e_{k,f}$

Spectrum access in MIMO-MAC

R. Couillet, H. V. Poor, M. Debbah, "Self-organized spectrum sharing in large MIMO multiple-access channels", *to be submitted to ISIT 2010*.

Depending on the correlation pattern at the base station, we obtain two iterative algorithms,

- Base-station aided algorithm, in case of receive correlation

Initialization: for all k, f, $\bar{e}_{k,f} = 1$. Define convergence threshold $\varepsilon > 0$.
while $\max_{k,f} \|\mathbf{P}_{k,f} - \mathbf{P}^\star_{k,f}\| > \varepsilon$ **do**
  for $k \in \{1, \ldots, K\}$ **do**
    for $f \in \{1, \ldots, F\}$ **do**
      The base station computes $e_{k,f}$
    end for
    The base station transmits $(e_{k,1}, \ldots, e_{k,F})$ to user k
    for $f \in \{1, \ldots, F\}$ **do**
      Based on $e_{k,f}$, user k computes $\mathbf{P}_{k,f}$
      Based on $e_{k,f}$ and $\mathbf{P}_{k,f}$, user k computes $\bar{e}_{k,f}$
    end for
    User k transmits $(\bar{e}_{k,1}, \ldots, \bar{e}_{k,F})$ to the base station
  end for
end while

Spectrum access in MIMO-MAC (2)

- Self-organized iterative water-filling, if no correlation at the base station

Initialization: for all k, f, $\bar{e}_{k,f} = 1$. Define convergence threshold $\varepsilon > 0$.
while $\max_{k,f} \|\mathbf{P}_{k,f} - \mathbf{P}^\star_{k,f}\| > \varepsilon$, **do**
  for $k \in \{1, \ldots, K\}$ **do**
    for $f \in \{1, \ldots, F\}$ **do**
      Based on e_f, user k computes $\mathbf{P}_{k,f}$
      Based on $\{e_f, \mathbf{P}_{k,f}\}$, user k computes $\bar{e}_{k,f}$
      Based on $\bar{e}_{k,f}$, user k updates $e_{k,f}$
    end for
    User k transmits $(e_{k,1}, \ldots, e_{k,F})$ to user $k+1 \pmod K$
  end for
end while

- however, proposed algorithm is sequential, time harvesting. Next step is to work on asynchronous schemes using,
  - gossiping approaches
  - graph theory
  - coding theory
- transmission bands must be uncorrelated. Currently working on frequency selective channels.

Spectrum access in MIMO-MAC (2)

- Self-organized iterative water-filling, if no correlation at the base station

Initialization: for all k, f, $\bar{e}_{k,f} = 1$. Define convergence threshold $\varepsilon > 0$.
while $\max_{k,f} \|\mathbf{P}_{k,f} - \mathbf{P}^\star_{k,f}\| > \varepsilon$, **do**
  for $k \in \{1, \dots, K\}$ **do**
    for $f \in \{1, \dots, F\}$ **do**
      Based on e_f, user k computes $\mathbf{P}_{k,f}$
      Based on $\{e_f, \mathbf{P}_{k,f}\}$, user k computes $\bar{e}_{k,f}$
      Based on $\bar{e}_{k,f}$, user k updates $e_{k,f}$
    end for
    User k transmits $(e_{k,1}, \dots, e_{k,F})$ to user $k+1 \pmod K$
  end for
end while

- however, proposed algorithm is sequential, time harvesting. Next step is to work on asynchronous schemes using,
  - gossiping approaches
  - graph theory
  - coding theory
- transmission bands must be uncorrelated. Currently working on frequency selective channels.

MIMO multi-band multiple access channel

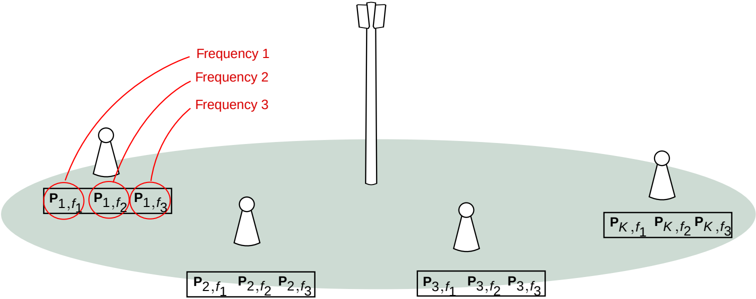

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

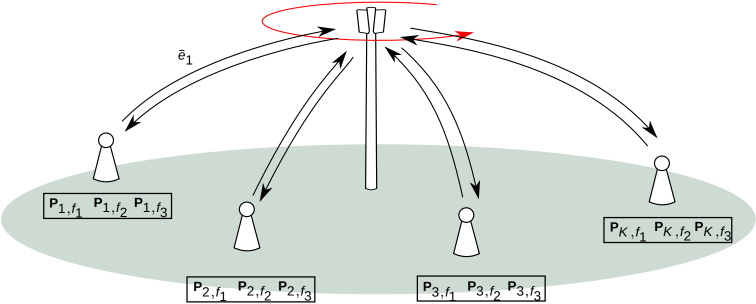

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

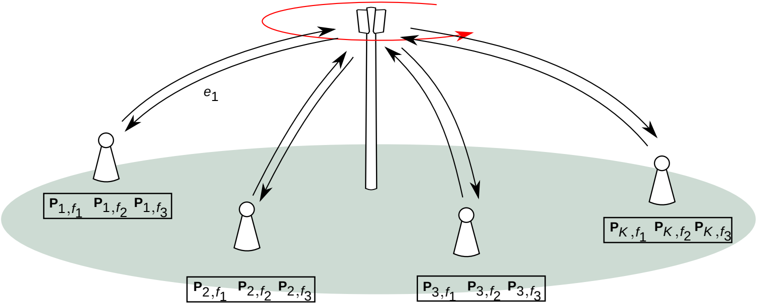

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

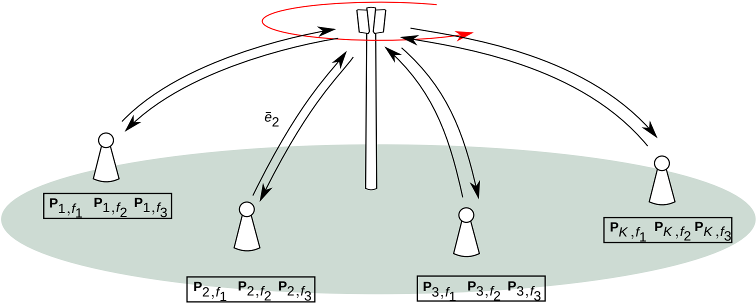

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

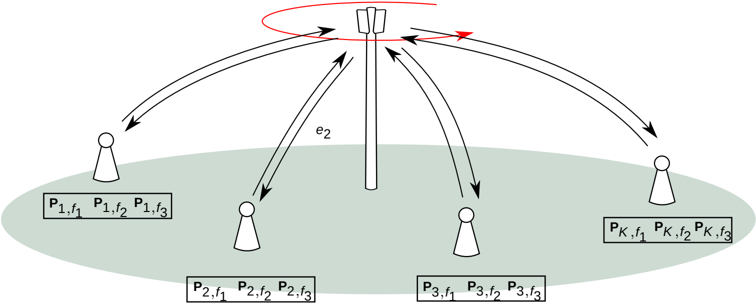

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

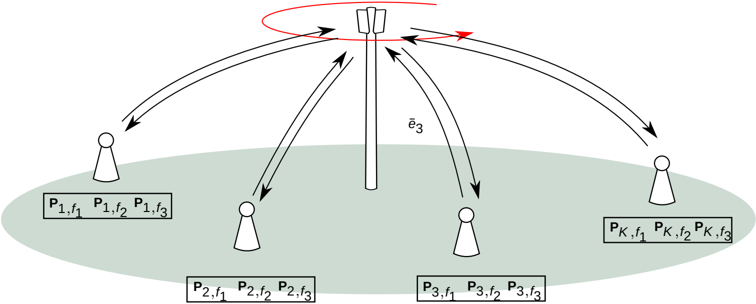

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

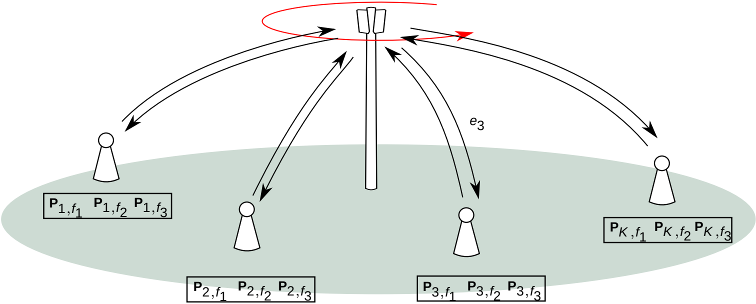

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

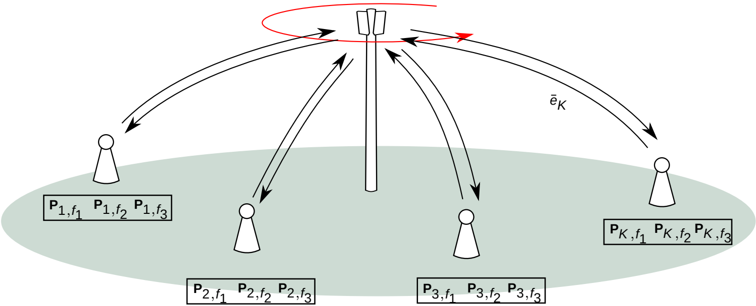

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

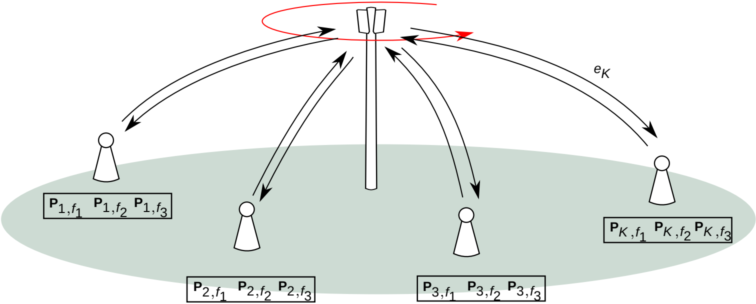

Figure: MIMO multi-band MAC

MIMO multi-band multiple access channel

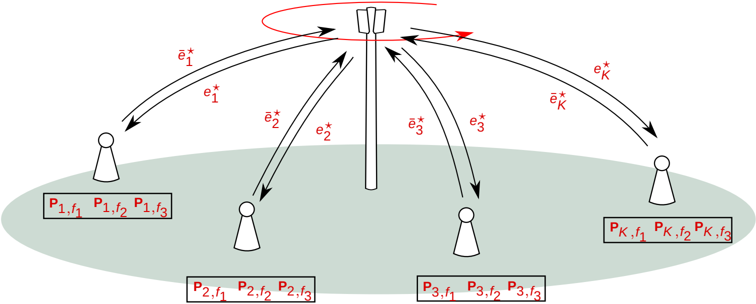

Figure: MIMO multi-band MAC

Work left to be done

- generalize Stieltjes transform approaches to structured matrices
- use random matrix theory to solve open issues
  - optimal Wiener filter in broadcast channels
  - optimal feedback for communications with imperfect CSI
- decentralized network organization using random matrix theory,
  - propose efficient feedback schemes
  - prove convergence or quasi-convergence
  - develop suboptimal schemes

Work left to be done

- generalize Stieltjes transform approaches to structured matrices
- use random matrix theory to solve open issues
  - optimal Wiener filter in broadcast channels
  - optimal feedback for communications with imperfect CSI
- decentralized network organization using random matrix theory,
  - propose efficient feedback schemes
  - prove convergence or quasi-convergence
  - develop suboptimal schemes

Related bibliography

- A. Moustakas, S. Simon and A. Sengupta, "MIMO Capacity Through Correlated Channels in the Presence of Correlated Interferers and Noise: A (Not So) Large *N* Analysis," IEEE Trans. on Information Theory, vol. 49, no. 10, 2003.
- C. K. Wen, Y. N. Lee, J. T. Chen, and P. Ting, "Asymptotic spectral efficiency of MIMO multiple-access wireless systems exploring only channel spatial correlations," IEEE Transactions on Signal Processing, vol. 53, no. 6, pp. 2059-2073, 2005.
- D. Guo and S. Verdú, "Multiuser Detection and Statistical Physics," Communications on Information and Network Security, Kluwer Academic Publishers, 2003.
- B. Zaidel, S. Shamai and S. Verdú, "Multicell Uplink Spectral Efficiency of Coded DS-CDMA With Random Signatures," IEEE Journal on Selected Areas in Communications, vol. 19, no. 8, 2001.
- Z. Bai and J. Silverstein, "On the signal-to-interference-ratio of CDMA systems in wireless communications," Annals of Applied Probability vol. 17 no. 1, pp. 81-101, 2007.
- R. B. Dozier and J. W. Silverstein, "On the Empirical Distribution of Eigenvalues of Large Dimensional Information-Plus-Noise Type Matrices," Journal of Multivariate Analysis 98(4), pp. 678-694, 2007.
- A.M. Tulino, A. Lozano and S. Verdù, "Impact of antenna correlation on the capacity of multiantenna channels," IEEE Trans. on Information Theory, vol. 51, no. 7, pp. 2491-2509, 2005.
- J. Dumont, S. Lasaulce, W. Hachem, Ph. Loubaton and J. Najim, "On the Capacity Achieving Covariance Matrix for Rician MIMO Channels: An Asymptotic Approach", submitted to IEEE Trans. on Information Theory.
- D. N. C. Tse, and S. V. Hanly, "Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Information Theory*, vol. 44, no. 7, pp. 2796-2815, 1998.

Related bibliography

- W. Hachem, Ph. Loubaton, and J. Najim, "A CLT for information theoretic statistics of Gram random matrices with a given variance profile," *Annals of Applied Probability*, vol. 18, no. 6, pp. 2071-2130, 2008.
- S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Information Theory*, vol. 49, no. 10, pp. 2658-2668, 2003.
- A. M. Tulino, and S. Verdù, "Random Matrix Theory and Wireless Communications," Now Publishers Inc., Dordrecht, The Netherlands, 2004.
- R. Couillet, S. Wagner, M. Debbah, "Asymptotic Analysis of Correlated Multi-Antenna Broadcast Channels", WCNC 2009, Budapest, Hungary.
- R. Couillet, M. Debbah, "Bayesian Inference for Multiple Antenna Cognitive Receivers", WCNC 2009, Budapest, Hungary.
- R. Couillet, M. Debbah, "Uplink capacity of self-organizing clustered orthogonal CDMA networks in flat fading channels", ITW 2009 Fall, Taormina, Sicily.
- R. Couillet, S. Wagner, M. Debbah, "Asymptotic Analysis of Linear Precoding Techniques in Correlated Multi-Antenna Broadcast Channels," submitted to IEEE Trans. on Information Theory.
- R. Couillet, M. Debbah, J. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," submitted to IEEE Trans. on Information Theory.
- R. Couillet, J. Silverstein, M. Debbah, "Eigen-inference for multi-source power estimation", *in preparation, to be submitted to ISIT 2010*.
- R. Couillet, M. Debbah, V. Poor, "Self-organized spectrum sharing in large MIMO multiple access channels", *submitted to ISIT 2010*.
- R. Couillet, S. Wagner, M. Debbah, D. Slock, "Asymptotic analysis of linear precoding in vector broadcast channels with limited feedback" *to be submitted*