Rate Region of Correlated MIMO Multiple Access and Broadcast Channels

Romain Couillet, PhD Student, ST-Ericsson, Supélec
Mérouane Debbah, PhD Advisor, Supélec
Jack Silverstein, Co-Author, North Carolina State University

Supélec

SSP 2009
Outline

1. Introduction

2. Mathematical Background

3. Rate Region of Broadcast Channels
   - Reminders
   - Random Matrix Theoretical Analysis
   - Power Allocation Algorithm

4. Application: channel link-dependent correlations

5. Conclusion
1. Introduction

2. Mathematical Background

3. Rate Region of Broadcast Channels
   - Reminders
   - Random Matrix Theoretical Analysis
   - Power Allocation Algorithm

4. Application: channel link-dependent correlations

5. Conclusion
Figure: Downlink scenario in multi-user broadcast channel
Motivation

Complex Parametrization

MAC/BC MIMO practical channels involve
- multiple transmit/receive antennas
- multiple users
- multiple path-loss exponents
- user-specific transmit/receive antenna correlations

Involved Analysis

This entails complex performance analysis, especially that of MAC/BC rate region.
Motivation

Complex Parametrization

MAC/BC MIMO practical channels involve
- multiple transmit/receive antennas
- multiple users
- multiple path-loss exponents
- user-specific transmit/receive antenna correlations

Involved Analysis

This entails complex performance analysis, especially that of MAC/BC rate region.
Random Matrix Theory

What is it used for?

- theoretical analysis of large random systems
- provides **deterministic approximates** to metrics of finite-size random systems
- allows to handle involved multi-dimensional large problems

Why does it help here?

- random behaviour of multiple intricate channels
- possibly large numbers of antennas at both transmit/receive sides
- power allocation problem involves in general too many parameters
Random Matrix Theory

**What is it used for?**
- theoretical analysis of large random systems
- provides **deterministic approximates** to metrics of finite-size random systems
- allows to handle involved multi-dimensional large problems

**Why does it help here?**
- random behaviour of multiple intricate channels
- possibly large numbers of antennas at both transmit/receive sides
- power allocation problem involves in general too many parameters
Outline

1 Introduction

2 Mathematical Background

3 Rate Region of Broadcast Channels
   • Reminders
   • Random Matrix Theoretical Analysis
   • Power Allocation Algorithm

4 Application: channel link-dependent correlations

5 Conclusion
Mathematical Background

Stieltjes and Shannon Transforms

**Definition (Stieltjes Transform)**

Let $X$ be an $N \times N$ matrix, then, for $z \in \mathbb{C} \setminus \mathbb{R}^+$, the Stieltjes transform $m_N(z)$ of $X$ is

$$m_N(z) = \text{tr} \left( X - zI \right)^{-1} = \int \frac{1}{\lambda - z} dF^X(\lambda) \quad (1)$$

**Definition (Shannon Transform)**

Let $X$ be an $N \times N$ matrix, then, for $z > 0$, the Stieltjes transform $S(z)$ of $X$ is

$$S(z) = \log \text{det} \left( I + \frac{1}{z} X \right)^{-1} = \int \frac{1}{w} - m_N(-w) dw \quad (2)$$
Mathematical Background

Stieltjes and Shannon Transforms

Definition (Stieltjes Transform)

Let $X$ be an $N \times N$ matrix, then, for $z \in \mathbb{C} \setminus \mathbb{R}^+$, the Stieltjes transform $m_N(z)$ of $X$ is

$$m_N(z) = \text{tr} \ (X - zI)^{-1} = \int \frac{1}{\lambda - z} dF_X(\lambda)$$  \hspace{1cm} (1)

Definition (Shannon Transform)

Let $X$ be an $N \times N$ matrix, then, for $z > 0$, the Stieltjes transform $S(z)$ of $X$ is

$$S(z) = \log \det \left( I + \frac{1}{z} X \right)^{-1} = \int \frac{1}{w} - m_N(-w) dw$$  \hspace{1cm} (2)
Mathematical Background

Stieltjes Transform Deterministic Approximate


Theorem

Denote $B_N = \sum_{k=1}^{K} T_k^{\frac{1}{2}} X_k^H R_k X_k T_k^{\frac{1}{2}}$, an $N \times N$ matrix with $R_k$, $n_k \times n_k$, $T_k$, $N \times N$ Hermitian nonnegative definite, such that $\{F^R_k\}_{n_k \geq 1}$ and $\{F^T_k\}_{N \geq 1}$ are tight. $X_k$ is $n_k \times N$ with entries $\mathcal{N}(0, 1/n_k)$. Let $c_k = n_k/N$. Then, as $n_k, N$ grow large ($K$ fixed), with ratio $c_k$

$$m_N(z) - m^o_N(z) \xrightarrow{a.s.} 0$$ (3)

where

$$m^o_N(z) = \frac{1}{N} \text{tr} \left( \frac{1}{N} \text{tr} \left( \sum_{k=1}^{K} \int \frac{r_k dF^R_k(r_k)}{1 + \frac{r_k}{c_k} e_k(z)} T_k - zI_N \right)^{-1} \right)$$ (4)

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

$$e_i(z) = \frac{1}{N} \text{tr} T_i \left( \sum_{k=1}^{K} \int \frac{r_k dF^R_k(r_k)}{1 + \frac{r_k}{c_k} e_k(z)} T_k - zI_N \right)^{-1}$$ (5)
Theorem

Let $B_N$ be as defined previously with the additional assumption that there exists $M > 0$, such that, for all $N, n_k$, $\max(\|T_k\|, \|R_k\|) < M$, and let $x > 0$. Then, for large $N, n_k$, $V(x) - V^0(x) \xrightarrow{\text{a.s.}} 0$, where

$$V(x) = \int \log \left(1 + \frac{b}{x}\right) dF^{B_N}(b)$$

(6)

$$= \log \left|I + \frac{1}{x}B_N\right|$$

(7)

and

$$V^0(x) = \int_{x}^{+\infty} \left(\frac{1}{w} - m_N^{(0)}(-w)\right) dw$$

(8)
Theorem

Let $B_N$ be as defined previously with the additional assumption that not too many eigenvalues of $T_k, R_k$ grow at a not too large rate. Let $x > 0$. Then, for large $N$, $n_k$, $\mathcal{V}(x) - \mathcal{V}^o(x) \xrightarrow{a.s.} 0$, where

$$
\mathcal{V}(x) = \int \log \left( 1 + \frac{b}{x} \right) dF^{B_N}(b)
$$

and

$$
\mathcal{V}^o(x) = \log \det \left( I_N + \frac{1}{x} \sum_{k=1}^{K} R_k \int \frac{\tau_k}{1 + c_k e_k(-x) \tau_k} dF^{T_k}(\tau_k) \right) \\
+ \sum_{k=1}^{K} \frac{1}{c_k} \int \log \left( 1 + c_k e_k(-x) \tau_k \right) dF^{T_k}(\tau_k) \\
+ x \cdot m_N^{(0)}(-x) - 1
$$
Rate Region of Broadcast Channels

Outline

1. Introduction

2. Mathematical Background

3. Rate Region of Broadcast Channels
   - Reminders
   - Random Matrix Theoretical Analysis
   - Power Allocation Algorithm

4. Application: channel link-dependent correlations

5. Conclusion
1. Introduction

2. Mathematical Background

3. Rate Region of Broadcast Channels
   - Reminders
     - Random Matrix Theoretical Analysis
     - Power Allocation Algorithm

4. Application: channel link-dependent correlations

5. Conclusion
System Model

Scenario

- $K$ users, each equipped with $n_k$ antennas, transmits $s_k \in \mathbb{C}^{n_k}$, such that $E[s_k s_k^H] = P_k$ and receives $y_k \in \mathbb{C}^{n_k}$ including noise $n_k$.
- one base station, equipped with $N$ antennas, transmits $s \in \mathbb{C}^N$, such that $E[ss^H] = P$ and receives $y \in \mathbb{C}^N$ including noise $n$.
- between base station and user $k$, channel is $H_k$

Uplink MAC/Downlink BC

- Uplink MAC model
  \[
  y = \sum_{k=1}^{K} H_k^H s_k + n_k
  \] 
  (11)

- Downlink BC model
  \[
  y_k = H_k s + n
  \] 
  (12)

Channel Model

We consider Gaussian channels with separable variance profile,

\[
H_k = R_k^{\frac{1}{2}} X_k T_k^{\frac{1}{2}}
\] 
(13)
System Model

Scenario

- $K$ users, each equipped with $n_k$ antennas, transmits $s_k \in \mathbb{C}^{n_k}$, such that $\mathbb{E}[s_k s_k^H] = P_k$ and receives $y_k \in \mathbb{C}^{n_k}$ including noise $n_k$.
- one base station, equipped with $N$ antennas, transmits $s \in \mathbb{C}^N$, such that $\mathbb{E}[ss^H] = P$ and receives $y \in \mathbb{C}^N$ including noise $n$.
- between base station and user $k$, channel is $H_k$

Uplink MAC/Downlink BC

- Uplink MAC model
  
  $$y = \sum_{k=1}^{K} H_k^H s_k + n_k$$  \hspace{1cm} (11)

- Downlink BC model
  
  $$y_k = H_k s + n$$  \hspace{1cm} (12)

Channel Model

We consider Gaussian channels with separable variance profile,

$$H_k = R_k^{1/2} X_k T_k^{1/2}$$  \hspace{1cm} (13)
**System Model**

**Scenario**
- \( K \) users, each equipped with \( n_k \) antennas, transmits \( s_k \in \mathbb{C}^{n_k} \), such that \( \mathbb{E}[s_k s_k^H] = P_k \) and receives \( y_k \in \mathbb{C}^{n_k} \) including noise \( n_k \).
- One base station, equipped with \( N \) antennas, transmits \( s \in \mathbb{C}^{N} \), such that \( \mathbb{E}[ss^H] = P \) and receives \( y \in \mathbb{C}^{N} \) including noise \( n \).
- Between base station and user \( k \), channel is \( H_k \).

**Uplink MAC/Downlink BC**

- **Uplink MAC model**
  \[
  y = \sum_{k=1}^{K} H_k^H s_k + n_k \tag{11}
  \]

- **Downlink BC model**
  \[
  y_k = H_k s + n \tag{12}
  \]

**Channel Model**

We consider Gaussian channels with separable variance profile,

\[
H_k = R_k^{\frac{1}{2}} X_k T_k^{\frac{1}{2}} \tag{13}
\]
Theorem (Verdú, 1989)

The rate region $C_{\text{MAC}}(P_1, \ldots, P_K; H^H)$ of the MAC channel $H^H$ under transmit power constraints $P_1, \ldots, P_K$ and compound channel $H^H = [H_1^H \ldots H_K^H]$, reads

$$C_{\text{MAC}}(P_1, \ldots, P_K; H^H) = \bigcup_{\sum_{i=1}^K P_i \leq P_i \geq 0 \quad i=1, \ldots, K} \left\{ \{R_i, 1 \leq i \leq K\} : \sum_{i \in S} R_i \leq \log I + \left| \frac{1}{\sigma^2} \sum_{i \in S} H_i^H P_i H_i \right|, \forall S \subset \{1, \ldots, K\} \right\}$$ (14)

Theorem (Goldsmith, 2003)

The rate region $C_{\text{BC}}(P; H)$ of the broadcast MIMO channel, for a transmit power constraint $P$ over the compound channel $H$, is

$$C_{\text{BC}}(P; H) = \bigcup_{\sum_{k=1}^K P_k \leq P} C_{\text{MAC}}(P_1, \ldots, P_K; H^H)$$ (15)
The rate region $C_{\text{MAC}}(P_1, \ldots, P_K; H^H)$ of the MAC channel $H^H$ under transmit power constraints $P_1, \ldots, P_K$ and compound channel $H^H = [H_1^H \ldots H_K^H]$, reads

$$C_{\text{MAC}}(P_1, \ldots, P_K; H^H) = \bigcup_{\text{tr}(P_i) \leq P_i, \forall i} \left\{ \{R_i, 1 \leq i \leq K\} : \sum_{i \in S} R_i \leq \log \left| I + \frac{1}{\sigma^2} \sum_{i \in S} H_i^H P_i H_i \right|, \forall S \subset \{1, \ldots, K\} \right\}$$  \hspace{1cm} (14)

The rate region $C_{\text{BC}}(P; H)$ of the broadcast MIMO channel, for a transmit power constraint $P$ over the compound channel $H$, is

$$C_{\text{BC}}(P; H) = \bigcup_{\sum_{k=1}^K P_k \leq P} C_{\text{MAC}}(P_1, \ldots, P_K; H^H)$$  \hspace{1cm} (15)
Outline

1 Introduction

2 Mathematical Background

3 Rate Region of Broadcast Channels
   - Reminders
   - Random Matrix Theoretical Analysis
   - Power Allocation Algorithm

4 Application: channel link-dependent correlations

5 Conclusion
Deterministic Approximate of logdet’s

For any set $S \subset \{1, \ldots, K\}$, we have approximately, for $N, n_k$ large,

$$\log \left| I + \frac{1}{\sigma^2} \sum_{i \in S} H_i^H P_i H_i \right| = \int_{\sigma^2}^{\infty} \left( \frac{1}{w} - m^0_S (-w) \right)$$

where

$$m^0_S (z) = \frac{1}{N} \text{tr} \left( \sum_{k \in S} \int \frac{r_k dF^R_{k} P_k (r_k)}{1 + \frac{r_k}{c_k} e_k (z)} T_k - z I_N \right)^{-1}$$

and the $e_i$’s verify

$$e_i (z) = \frac{1}{N} \text{tr} T_i \left( \sum_{k \in S} \int \frac{r_k dF^R_{k} P_k (r_k)}{1 + \frac{r_k}{c_k} e_k (z)} T_k - z I_N \right)^{-1}$$

Removal of stochastic parameters

The stochastic contributions of the $X_k$’s are discarded.

Simplified analysis, only dependent on $T_k$’s and eigenvalues of $R_k P_k$’s
Deterministic Approximate of logdet’s

For any set \( S \subset \{1, \ldots, K\} \), we have approximately, for \( N, n_k \) large,

\[
\log \left| I + \frac{1}{\sigma^2} \sum_{i \in S} H_i^H P_i H_i \right| = \int_{\sigma^2}^{\infty} \left( \frac{1}{w} - m^0_S (-w) \right) \frac{1}{w} \quad (16)
\]

where

\[
m^0_S(z) = \frac{1}{N} \text{tr} \left( \sum_{k \in S} \int \frac{r_k dF^{R_k P_k}(r_k)}{1 + \frac{r_k}{c_k} e_k(z)} T_k - zI_N \right)^{-1} \quad (17)
\]

and the \( e_i \)’s verify

\[
e_i(z) = \frac{1}{N} \text{tr} T_i \left( \sum_{k \in S} \int \frac{r_k dF^{R_k P_k}(r_k)}{1 + \frac{r_k}{c_k} e_k(z)} T_k - zI_N \right)^{-1} \quad (18)
\]

Removal of stochastic parameters

The stochastic contributions of the \( X_k \)’s are discarded.

**Simplified analysis, only dependent on \( T_k \)’s and eigenvalues of \( R_k P_k \)’s**
Outline

1. Introduction

2. Mathematical Background

3. Rate Region of Broadcast Channels
   - Reminders
   - Random Matrix Theoretical Analysis
   - Power Allocation Algorithm

4. Application: channel link-dependent correlations

5. Conclusion
Known Results

- If all $R_k$'s are identical, optimal power allocation known.
- In this case, optimal eigenvalue allocation is known in closed-form.
- If at least two $R_k$'s are different, no result known to this day.

Chosen Power Allocation

We decide here to:

- Align $P_k = U_k Q_k U_k^H$ eigenvectors to $R_k$ eigenvectors.
- Optimize the eigenvalues $q_{kn}$ allocation by convex optimization.

But this is not proven to be optimal!
**Known Results**

- If all $R_k$'s are identical, optimal power allocation known.
- In this case, optimal eigenvalue allocation is known in closed-form.
- If at least two $R_k$'s are different, no result known to this day.

**Chosen Power Allocation**

We decide here to:

- Align $P_k = U_k Q_k U_k^H$ eigenvectors to $R_k$ eigenvectors.
- Optimize the eigenvalues $q_{kn}$ allocation by convex optimization.

**But this is not proven to be optimal!**
Proposition (Convex Property)

For all $N, n_k$, 

$$
\int_{\sigma^2}^{\infty} \left( \frac{1}{w} - m_S^0(-w) \right) \, dw \text{ is a convex function of the } q_{kn} \text{'s, } k \in S, n \in \{1, \ldots, n_k\}.
$$

Methodology

We perform convex optimization under,

- $K$ equality constraints $\sum_{k=1}^{K} \text{tr}Q_k = P$,
- $n_1 + \ldots + n_K$ inequality constraints $q_{kn} \geq 0$.

This is done thanks to classical convex optimization methods which turn the problem into an unconstrained convex optimization problem using

- equality constraints elimination
- barrier method to eliminate the inequality constraints
Proposition (Convex Property)

For all $N, n_k, \int_{\sigma^2}^{\infty} \left( \frac{1}{w} - m_S^0 (-w) \right) \, dw$ is a convex function of the $q_{kn}$'s, $k \in S, n \in \{1, \ldots, n_k\}$.

Methodology

We perform convex optimization under,

- $K$ equality constraints $\sum_{k=1}^{K} \text{tr}Q_k = P$,
- $n_1 + \ldots + n_K$ inequality constraints $q_{kn} \geq 0$.

This is done thanks to classical convex optimization methods which turn the problem into an unconstrained convex optimization problem using

- equality constraints elimination
- barrier method to eliminate the inequality constraints
Proposition (Concavity Property)

If any of the correlation matrices $R_k, \ k \in S$ is invertible, then $V^o(x)$ is a strictly concave function of $P_1, \ldots, P_{|S|}$.

Water-Filling Algorithm

At initialization, for all $k \in S$, $Q_k = \frac{Q_k}{n_k} I_{n_k}$, $\delta_k = 1$, $e_k = 1$.

while the $P_k$’s have not converged do
   for $k \in S$ do
      Set $e_k$ as solution of its fixed-point equation
      for $i = 1 \ldots, n_k$ do
         Set $q_{k,i} = \left( \mu_k - \frac{1}{c_k e_k r_{ki}} \right)^+$, with $\mu_k$ such that $\text{tr}Q_k = P_k$.
      end for
   end for
end while

Asymptotic Optimality

- align $P_k = U_k Q_k U_k^H$ eigenvectors to $R_k$ eigenvectors now proven to maximize $V^o(x)$ for any $N$.
- Water-Filling solution now proven optimal in the $N \to \infty$ limit.
Proposition (Concavity Property)

If any of the correlation matrices $\mathbf{R}_k$, $k \in S$ is invertible, then $\mathcal{V}^\circ(x)$ is a strictly concave function of $P_1, \ldots, P_{|S|}$.

Water-Filling Algorithm

At initialization, for all $k \in S$, $\mathbf{Q}_k = \frac{Q_k}{n_k} \mathbf{I}_{n_k}$, $\delta_k = 1$, $e_k = 1$.

while the $P_k$’s have not converged do

for $k \in S$ do

Set $e_k$ as solution of its fixed-point equation

for $i = 1 \ldots, n_k$ do

Set $q_{k,i} = \left(\mu_k - \frac{1}{c_k e_k r_{ki}}\right)^+$, with $\mu_k$ such that $\text{tr} \mathbf{Q}_k = P_k$.

end for

end for

end while

Asymptotic Optimality

align $\mathbf{P}_k = \mathbf{U}_k \mathbf{Q}_k \mathbf{U}_k^H$ eigenvectors to $\mathbf{R}_k$ eigenvectors now proven to maximize $\mathcal{V}^\circ(x)$ for any $N$.

Water-Filling solution now proven optimal in the $N \to \infty$ limit.
Proposition (Concavity Property)

If any of the correlation matrices $R_k$, $k \in S$ is invertible, then $V^o(x)$ is a strictly concave function of $P_1, \ldots, P_{|S|}$.

Water-Filling Algorithm

At initialization, for all $k \in S$, $Q_k = \frac{Q_k}{n_k}I_{n_k}$, $\delta_k = 1$, $e_k = 1$.

While the $P_k$’s have not converged do

For $k \in S$ do

Set $e_k$ as solution of its fixed-point equation

For $i = 1, \ldots, n_k$ do

Set $q_{k,i} = \left( \mu_k - \frac{1}{c_k e_k r_{ki}} \right)^+$, with $\mu_k$ such that $\text{tr} Q_k = P_k$.

End for

End for

End while

Asymptotic Optimality

- Align $P_k = U_k Q_k U_k^H$ eigenvectors to $R_k$ eigenvectors now proven to maximize $V^o(x)$ for any $N$.

- Water-Filling solution now proven optimal in the $N \to \infty$ limit.
Outline

1. Introduction
2. Mathematical Background
3. Rate Region of Broadcast Channels
   - Reminders
   - Random Matrix Theoretical Analysis
   - Power Allocation Algorithm
4. Application: channel link-dependent correlations
5. Conclusion
Scenario

Directive Transmissions

We assume here

- $K = 2$ users, $n_1 = n_2 = 4$, $N = 8$.
- the existence of transmit/receive solid angles due to
  - not all isotropically transmitted signal directions result in useful energy.
  - not all isotropically received directions contain energy.

$R_k$ and $T_k$ models

We use a generalized Kronecker model. For instance, entry $(a, b)$ of matrix $T_1$ reads

$$T_{1_{ab}} = \int_{\theta_{\min}^{(T_1)}}^{\theta_{\max}^{(T_1)}} \exp \left( 2\pi i |a - b| \frac{d_T}{\lambda} \cos(\theta) \right) d\theta$$ (19)
Directive Transmissions

We assume here
- $K = 2$ users, $n_1 = n_2 = 4$, $N = 8$.
- the existence of transmit/receive solid angles due to
  - not all isotropically transmitted signal directions result in useful energy.
  - not all isotropically received directions contain energy.

$R_k$ and $T_k$ models

We use a generalized Kronecker model. For instance, entry $(a, b)$ of matrix $T_1$ reads

$$T_{1}^{(a,b)} = \int_{\theta_{\min}^{(T_1)}}^{\theta_{\max}^{(T_1)}} \exp \left( 2\pi i |a - b| \frac{d_{T}}{\lambda} \cos(\theta) \right) d\theta$$

(19)
**Figure:** Rate region $C_{BC}$ for $K = 2$ users, theory against simulation, $N = 8$, $n_1 = n_2 = 4$, SNR = 20 dB, random transmit-receive solid angle of aperture $\pi/2$, $d_T/\lambda = 10$, $d_R/\lambda = 1/4$. 

R. Couillet, M. Debbah, J. Silverstein (©)
Figure: Rate region $C_{BC}$ for $K = 2$ users, $N = 8$, $n_1 = n_2 = 4$, SNR = 20 dB, random transmit-receive solid angle of aperture $\pi/2$, $d_T/\lambda = 10$, $d_R/\lambda = 1/4$. In thick line, capacity limit when $E[ss^H] = I_N$. 
Figure: Rate region $C_{BC}$ for $K = 2$ users, $N = 8$, $n_1 = n_2 = 4$, $\text{SNR} = -5 \text{ dB}$, random transmit-receive solid angle of aperture $\pi/2$, $d_T/\lambda = 10$, $d_R/\lambda = 1/4$. In thick line, capacity limit when $E[ss^H] = I_N$. 
We have derived here a formula to analytically compute the rate region of multiple access and broadcast channels for multi-user with large numbers of antennas with transmit and receive correlations.

- A suboptimal power optimization method using convex optimization.

Results:
- Formulas only depend on the matrices $R_k P_k$ and $T_k$'s.
- Deterministic approximate is a good fit even for small dimensions.
- Huge capacity gain are expected with power optimization at low SNR.
We have derived here a formula to analytically compute the rate region of multiple access and broadcast channels for multi-user with large numbers of antennas with transmit and receive correlations.

Results:
- Formulas only depend on the matrices $R_k P_k$ and $T_k$'s.
- Deterministic approximate is a good fit even for small dimensions.
- Huge capacity gain are expected with power optimization at low SNR.