

Projet **SESAME**

Programme **MDCO** - Financement **ANR**

Partenaires

Centre National de la Recherche Scientifique

Télécom Paristech

Ecole supérieure d'électricité

Eurécom

Université Marne-La-Vallée

Coordonnateur

Jamal Najim, CNRS et Télécom Paristech

<http://www.tsi.enst.fr/~najim/anr-sesame/>

Présentation du projet

Présentation générale

Présentation scientifique

Objectifs scientifiques

Activité scientifique

Zoom sur 2 work packages

Présentation générale

Projet

- ▶ à **vocation académique**
- ▶ avec **quatre intervenants institutionnels**

CNRS et Télécom ParisTech
Supélec
Université Marne La Vallée
Eurécom

P. Bianchi, W. Hachem, J. Najim
M. Debbah
P. Loubaton
L. Cottatellucci, D. Slock

Présentation générale

Projet

- ▶ à **vocation académique**
- ▶ avec **quatre intervenants institutionnels**

CNRS et Télécom ParisTech
Supélec
Université Marne La Vallée
Eurécom

P. Bianchi, W. Hachem, J. Najim
M. Debbah
P. Loubaton
L. Cottatellucci, D. Slock

Se situe à l'interface des:

- ▶ communications numériques
- ▶ traitement statistique du signal
- ▶ probabilités & statistiques

Présentation générale

Projet

- ▶ à **vocation académique**
- ▶ avec **quatre intervenants institutionnels**

CNRS et Télécom ParisTech
Supélec
Université Marne La Vallée
Eurécom

P. Bianchi, W. Hachem, J. Najim
M. Debbah
P. Loubaton
L. Cottatellucci, D. Slock

Se situe à l'interface des:

- ▶ communications numériques
- ▶ traitement statistique du signal
- ▶ probabilités & statistiques

Durée de **48 mois** + extension **6 mois** - Budget de **470 kE**

Présentation scientifique

SESAME: conSistent **ES**timation and lAрге random **MatricEs**.

Objectif: aborder des problèmes d'estimation motivés par des applications relevant des **communications sans fil**:

- ▶ Réseaux sans fil CDMA
- ▶ Radio cognitive et opportuniste
- ▶ Communications MIMO avec grand nombre d'antennes

Contexte spécifique: Séries temporelles multivariées de grande taille où **nombre** de données et **taille** des données sont du même ordre.

Environnement inadapté aux estimateurs empiriques traditionnels.

Communications numériques et matrices aléatoires

Article fondateur de **Telatar** en 1995 qui applique les résultats de Marčenko et Pastur (1967) à l'étude de la capacité d'un canal **MIMO** particulier.

Communications numériques et matrices aléatoires

Article fondateur de **Telatar** en 1995 qui applique les résultats de Marčenko et Pastur (1967) à l'étude de la capacité d'un canal **MIMO** particulier.

Depuis, engouement majeur et intérêt constant en communications numériques pour la théorie des matrices aléatoires:

- ▶ application de résultats mathématiques antérieurs aux communications numériques
- ▶ élaboration de nouveaux résultats mathématiques pilotés par des applications pertinentes en communications numériques

Communications numériques et matrices aléatoires

Article fondateur de **Telatar** en 1995 qui applique les résultats de Marčenko et Pastur (1967) à l'étude de la capacité d'un canal **MIMO** particulier.

Depuis, engouement majeur et intérêt constant en communications numériques pour la théorie des matrices aléatoires:

- ▶ application de résultats mathématiques antérieurs aux communications numériques
- ▶ élaboration de nouveaux résultats mathématiques pilotés par des applications pertinentes en communications numériques

Notre groupe a développé une expertise sur le sujet depuis début 2000:

- ▶ cf. le projet d'**ACI MALCOM** 2004-2007

Matrices aléatoires et estimation statistique

Le problème:

- ▶ Une des spécificités de l'estimation statistique en communications numériques provient de la rencontre de séries temporelles multivariées de grande taille où nombre de données et taille des données sont du même ordre.

Matrices aléatoires et estimation statistique

Le problème:

- ▶ Une des spécificités de l'estimation statistique en communications numériques provient de la rencontre de séries temporelles multivariées de grande taille où nombre de données et taille des données sont du même ordre.

L'exemple de séries temporelles multivariées

Pour une série temporelle multivariée $(\mathbf{x}_1, \dots, \mathbf{x}_n)$, un problème important est l'estimation de fonctions de la matrice de covariance \mathbf{R} , par exemple $\log \det \mathbf{R}$.

- ▶ Si la dimension K de chaque observation \mathbf{x}_i est très inférieure à n , alors on remplace \mathbf{R} par l'estimateur empirique
$$\hat{\mathbf{R}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \mathbf{x}_i^T.$$
- ▶ Si la dimension de chaque observation est du même ordre que le nombre d'observations, ce n'est plus le cas.

Objectifs du présent projet

- ▶ Promouvoir et développer les techniques mathématiques issues de la théorie des matrices aléatoires pour l'étude des systèmes de communications mobiles.
- ▶ Développer des techniques statistiques efficaces dans le cadre spécifique où nombre de données et taille des données sont du même ordre.

Work packages

Work package 1: Synthèse des résultats disponibles.

- ▶ délivrables: rapport à $t_0 + 12$ (réalisé)

Work package 2: Estimation pour les modèles information + bruit

- ▶ délivrables: rapports à $t_0 + 24$ (réalisé)

Work package 3: Estimation & détection pour des observations multivariées en grande dimension

- ▶ Estimation sous-espace: rapport à $t_0 + 24$ (réalisé)
- ▶ Estimation de débit en radio cognitive: rapport à $t_0 + 48$ (réalisé)
- ▶ Détection en radio cognitive: rapport à $t_0 + 48$ (réalisé)

Work package 4: Estimation dans le cadre des séries temporelles

- ▶ délivrables: rapport à $t_0 + 24$ (réalisé)

Présentation du projet

Activité scientifique

Animation scientifique

Production scientifique

Zoom sur 2 work packages

Réunions plénières

Depuis le lancement du projet (janvier 2008), il y a eu **sept réunions plénières** rassemblant tous les partenaires du projet.

Groupes de travail

Trois groupes de travail ont été organisés dans le cadre du présent projet:

- ▶ Groupe de travail sur la **G-estimation**, impliquant 3 partenaires (Télécom, Supélec, Univ. MLV) ,animé par P. Loubaton. A donné lieu à environ **9 réunions**.
- ▶ Groupe de travail sur le **Collaborative Sensing**, impliquant 2 partenaires (Télécom, Supélec), animé par P. bianchi. A donné lieu à **12 réunions**.
- ▶ Groupe de travail sur l'estimation sous-espace pour les données multivariées de grande dimension, animé par W. Hachem. A donné lieu à **10 réunions**.

Formation à la recherche

Formation à la recherche. Le projet ANR SESAME finance directement:

- ▶ Doctorants:
 - ▶ Antonia Masucci, (Supélec)
 - ▶ Samir Mohammad Omar (Eurécom)
 - ▶ Lei Xiao (Eurécom)
- ▶ Postdoc
 - ▶ Abla Kammoum (Télécom Paristech)

Animation scientifique

Animation scientifique.

Le projet ANR SESAME joue un rôle important d'animation scientifique par l'ensemble des chercheurs qu'il fédère ou a rassemblé à travers ses réunions:

- ▶ **Docs et Postdocs:** M. Kharouf, A. Kammoun, P. Vallet, G. Alfano, R. Couillet, C. Artigues, J. Hoydis, J-F. Yao.
- ▶ **Chercheurs:** J. Jakubowitz (Télécom), S. Yang (Supélec), M. Maïda (Orsay), J. Yao (Rennes).

Invitation d'experts.

- ▶ Alan Edelman, (Massachusetts Institute of Technology - USA),
 - ▷ **matrices aléatoires et aspects computationnels**
- ▶ Xavier Mestre, (Centre Tecnològic de Telecomunicacions de Catalunya),
 - ▷ **traitement statistique du signal**
- ▶ Boaz Nadler, (Weizmann Institute - Israël),
 - ▷ **traitement statistique du signal**
- ▶ Jack Silverstein, (North Carolina State University - USA),
 - ▷ **théorie des grandes matrices aléatoires**

Production scientifique

Le projet SESAME a donné lieu à la production scientifique suivante:

- ▶ **15 Articles de revue** soumis ou acceptés à des revues internationales à comité de lecture.
- ▶ **22 articles** parus ou à paraître dans les actes de conférences internationales à comité de lecture.
- ▶ **6 rapports** réalisés dans le cadre du projet SESAME.
- ▶ **1 colloque international** organisé à Télécom Paristech [3 jours - 40 participants - octobre 2010]

<http://www.tsi.enst.fr/~najim/anr-sesame>

Faits marquants

- ▶ Tutoriel à la conférence EUSIPCO (European Signal Processing Conference), Barcelone 2011 [Hachem, Loubaton, Najim]
- ▶ Tutoriel à la conférence ICASSP (International Conference on acoustics, Speech, and Signal Processing), Prague 2011 [Couillet, Debbah]
- ▶ Livre “Random Matrix Methods for Wireless Communications”, par M. Debbah et R. Couillet qui s’inscrit dans la thématique de l’ANR
- ▶ Keynote speaker “Random Matrices for Cooperative Spectrum Sensing: Some recent results”, par J. Najim, Conférence Physcomnet, Seoul 2009.

Présentation du projet

Activité scientifique

Zoom sur 2 work packages

Etude coopérative pour la radio cognitive (WP2)

Localisation de source bande étroite (WP3)

Travaux en cours au sein de l'ANR SESAME

On décrit ici en détail deux projets en cours au sein du projet SESAME, en essayant de faire l'articulation entre les aspects méthodologiques mis en œuvre et les applications aux communications numériques.

- ▶ Ecoute coopérative de spectre pour la radio cognitive.
- ▶ Avancées en G-estimation.

Présentation du projet

Activité scientifique

Zoom sur 2 work packages

Etude coopérative pour la radio cognitive (WP2)

Localisation de source bande étroite (WP3)

Etude coopérative de spectre pour la radio cognitive

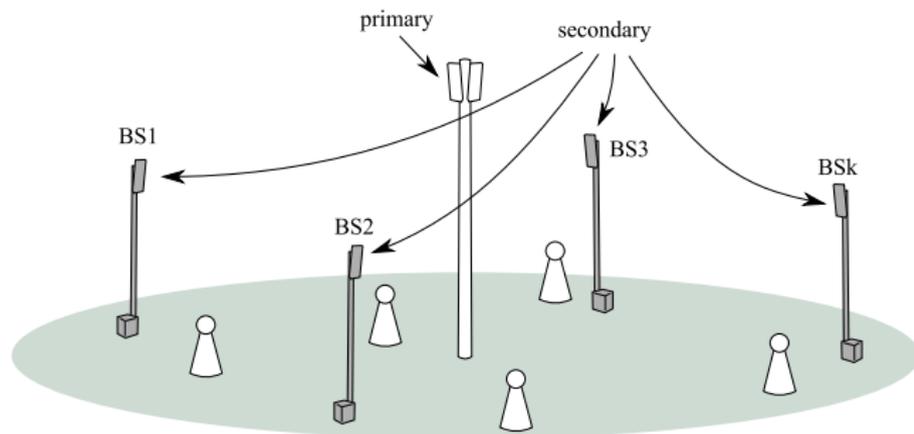


Figure: Scénario considéré

Caractéristiques de la communication

- ▶ Un réseau sans fil **prioritaire** utilise une certaine bande de fréquence,
- ▶ L'utilisation de la bande de fréquence est **irrégulière**,
- ▶ Un réseau **secondaire** sans fil souhaite utiliser cette bande de fréquence quand celle-ci est disponible,
- ▶ les différents agents du réseau secondaire vont **coopérer** pour tester la disponibilité de la bande de fréquence,
- ▶ Leur méthode doit être **autodidacte**: aucune information sur le signal primaire émis n'est requise,
- ▶ Le test doit se faire en **temps réel**.

Modélisation

En l'absence de signal (**Hypothèse H0**). Chaque capteur secondaire $k = 1 : N$ va recevoir un signal $y_k(\ell)$ aux instants d'échantillonnage $\ell = 1 : n$ avec

$$y_k(\ell) = \sigma w_k(\ell) ,$$

où $w_k(\ell)$ est un bruit blanc gaussien, et σ est sa variance.

En présence de signal (**Hypothèse H1**). Le signal est maintenant de la forme:

$$y_k(\ell) = h_k s(\ell) + \sigma w_k(\ell) ,$$

où $s(\ell)$ est le signal primaire gaussien à l'instant ℓ et h_k est le coefficient d'atténuation pour le récepteur secondaire k .

Spécificités

Contraintes:

- ▶ Nombre de capteurs N et dimension du signal reçu n par chaque capteur du même ordre.
- ▶ La variance du bruit σ ainsi que les coefficients d'atténuation h_k sont inconnus.
- ▶ Nécessité d'un traitement en temps réel.

Collaboration entre capteurs secondaires:

- ▶ On range les données sous forme matricielle:

$$\mathbf{Y} = \left(\frac{y_k(\ell)}{\sqrt{n}} \right)_{k=1:N, \ell=1:n},$$

tous les signaux étant mutualisés.

Résultats

- ▶ Construction du test statistique du maximum de vraisemblance généralisé:

$$\mathcal{T}_1 = \frac{\lambda_{\max}(\mathbf{Y}\mathbf{Y}^*)}{\frac{1}{N} \text{Trace}(\mathbf{R})}.$$

- ▶ Etude exhaustive de ses erreurs de 1ère et 2ème espèce
- ▶ Comparaison de ce test avec un test moins performant proposé dans la littérature:

$$\mathcal{T}_1 \succ \mathcal{T}_2 \quad \text{où} \quad \mathcal{T}_2 = \frac{\lambda_{\max}(\mathbf{Y}\mathbf{Y}^*)}{\lambda_{\min}(\mathbf{Y}\mathbf{Y}^*)}.$$

Outils méthodologiques

Les résultats précédents reposent sur les outils méthodologiques suivants, dont certains ont été développés spécifiquement dans le cadre du projet SESAME:

- ▶ Fluctuations de la plus grande valeur propre d'une matrice de Gram (Johnstone 2001)
 - ▷ seuil du test, erreur de première espèce
- ▶ Convergence différenciée de la plus grande valeur propre d'une matrice de Gram sous **H0** et **H1** (Baik et al. 2006)
 - ▷ validité du test, erreur de 2ème espèce
- ▶ Etablissement d'un principe de grandes déviations pour la plus grande valeur propre d'une matrice de Gram sous **H1** (Bianchi et al. 2009)
 - ▷ calcul de l'exposant d'erreur, pour description de l'erreur de 2ème espèce

Comparaison des erreurs de 2ème espèce

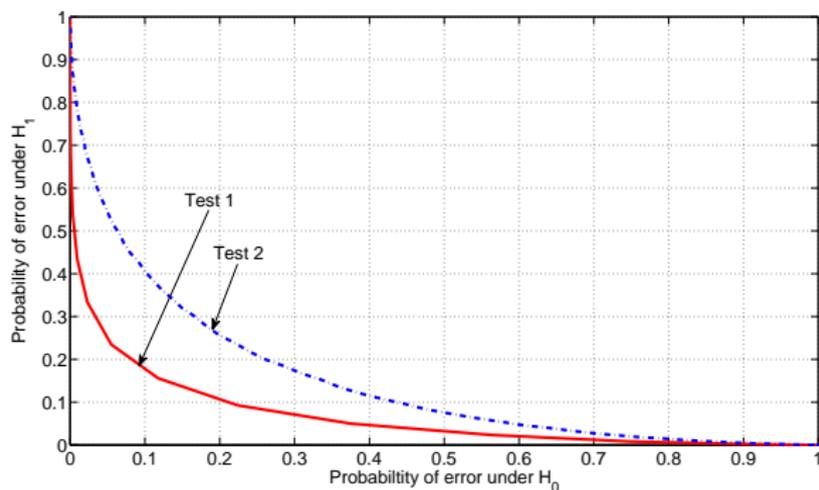


Figure: Erreurs de première et 2ème espèce pour \mathcal{T}_1 et \mathcal{T}_2 dans le cas où $N = 10$, $n = 50$ et $\rho = 1$.

Publications

(Conférences)

1. L. Cardoso, M. Debbah, P. Bianchi, J. Najim. "Cooperative Spectrum Sensing Using Random Matrix Theory" *IEEE ISWPC 2008*, Greece.
2. P. Bianchi, M. Debbah, M. Maïda, J. Najim "Power Analysis of some Hypothesis tests for collaborative sensing" *IEEE Statistical Signal Processing*, Août 2009.
3. G. Alfano, P. Bianchi, M. Debbah, J. Najim "Asymptotics of eigenbased collaborative sensing", *IEEE Information Theory Workshop*, Oct. 2009.

(Revue)

1. A. Kammoun, R. Couillet, J. Najim, M. Debbah "Performance of capacity inference methods under colored interference" *soumis à IEEE Trans. Inf. Th.*
2. P. Bianchi, M. Debbah, J. Najim "Asymptotic Independence in the Spectrum of the Gaussian Unitary Ensemble" *Electronic Communication Probab.*, 2010.
3. P. Bianchi, M. Debbah, M. Maïda, J. Najim "Performance of Statistical tests for Single source detection using Random Matrix theory" *IEEE Inf. Theory*, 2011.

Présentation du projet

Activité scientifique

Zoom sur 2 work packages

Etude coopérative pour la radio cognitive (WP2)

Localisation de source bande étroite (WP3)

Scénario pour la localisation de source

On s'intéresse à l'estimation des **directions d'arrivée en traitement d'antenne**.

Scénario: réception de K signaux sources à l'aide de M antennes ($M > K$) collectant N observations: $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N]$ modélisées par

$$\mathbf{Y} = \mathbf{AS} + \mathbf{V}$$

où

- ▶ $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N]$: matrice $M \times N$ des observations;
- ▶ $\mathbf{A} = [\mathbf{a}(\varphi_1), \dots, \mathbf{a}(\varphi_K)]$: matrice $M \times K$ des vecteurs directionnels;
- ▶ l'application $\varphi \mapsto \mathbf{a}(\varphi)$ est connue; les $(\varphi_k, 1 \leq k \leq K)$ sont inconnus.
- ▶ \mathbf{S} : matrice des signaux transmis (déterministe).

Principe de la localisation de source

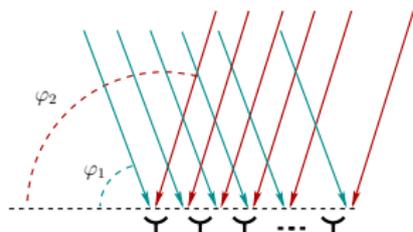


Figure: Estimation des angles d'arrivée φ_1 et φ_2

Principe de la localisation de source

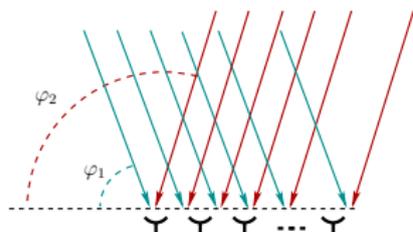


Figure: Estimation des angles d'arrivée φ_1 et φ_2

Principe de la localisation. Si $\mathbf{\Pi}$ est la matrice de projection orthogonale sur les colonnes de $\mathbf{A} = [\mathbf{a}(\varphi_1), \dots, \mathbf{a}(\varphi_K)]$, alors les (φ_k) sont solution de:

$$\mathbf{a}(\varphi)^* \mathbf{\Pi} \mathbf{a}(\varphi) = 0 .$$

Principe de la localisation de source

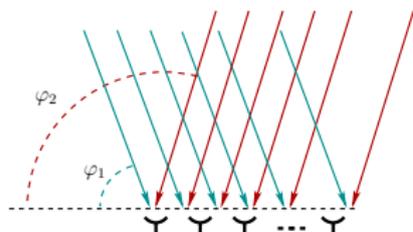


Figure: Estimation des angles d'arrivée φ_1 et φ_2

Principe de la localisation. Si $\mathbf{\Pi}$ est la matrice de projection orthogonale sur les colonnes de $\mathbf{A} = [\mathbf{a}(\varphi_1), \dots, \mathbf{a}(\varphi_K)]$, alors les (φ_k) sont solution de:

$$\mathbf{a}(\varphi)^* \mathbf{\Pi} \mathbf{a}(\varphi) = 0 .$$

L'approche usuelle consiste à remplacer $\mathbf{\Pi}$ par un estimateur $\hat{\mathbf{\Pi}}$ et à estimer les paramètres directionnels par les K minima de

$$\varphi \mapsto \mathbf{a}(\varphi)^* \hat{\mathbf{\Pi}} \mathbf{a}(\varphi)$$

Estimateur sous espace traditionnel

- ▶ N représente la **taille de l'échantillon**,
- ▶ M représente le **nombre d'antennes**.

Si $M \ll N$

alors le projecteur Π est estimé par $\hat{\Pi}_{\text{trad}}$, projecteur sur l'**espace propre associé aux $M - K$ plus petites valeurs propres**

Si $M \sim N$

Alors, ce projecteur **n'est plus consistant!**. La théorie des matrices aléatoires permet de **construire des estimateurs consistants**.

Estimateur sous espace pour $M \sim N$

Deux sous-problèmes selon le nombre des directions d'arrivée K :

- ▶ Le cas où K est **petit** devant M, N . La théorie des grandes matrices aléatoires (petites perturbations et matrices spiked) permet de construire un **estimateur corrigé**:

$$\hat{\mathbf{n}} = \hat{\mathbf{n}}_{\text{trad}} + \sum_{k=1}^K \kappa_k \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^*$$

($\hat{\mathbf{u}}_k$ vecteur propre associé à la valeur propre empirique $\hat{\lambda}_k$)

- ▶ Le cas où K est du même ordre que M, N . Dans ce cas, on construit un **estimateur général**:

$$\hat{\mathbf{n}} = \sum_{k=1}^M \xi_k \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^*$$

- ▶ Les simulations **confirment la validité** de ces résultats.

Publications

(Conférences)

1. P. Vallet, W. Hachem, Ph. Loubaton, X. Mestre, J. Najim. "An Improved MUSIC Algorithm Based on Low Rank Perturbation of Large Random Matrices" *IEEE SSP 2011* Nice.
2. P. Vallet, W. Hachem, Ph. Loubaton, X. Mestre, J. Najim. "On the Consistency of the G-MUSIC DoA Estimator" *IEEE SSP 2011* Nice.

(Journaux)

1. W. Hachem, P. Loubaton, X. Mestre, J. Najim and P. Vallet. "Large information plus noise random matrix models and consistent subspace estimation in large sensor networks". 2011 , *to appear in Random Matrices and Their Applications*.
2. P. Vallet, Ph. Loubaton, X. Mestre, "Improved subspace estimation for multivariate observations of high dimension: the deterministic signals case", *to appear in IEEE Trans. in Inf. Theory, February 2012*.
3. W. Hachem, P. Loubaton, X. Mestre, J. Najim and P. Vallet. "A subspace estimator for fixed rank perturbations of large random matrices" 2011 , *submitted to J. Multivariate Analysis*.