CURRICULUM VITAE

Romain COUILLET
Full Professor at CentraleSupélec
GSTATS DataScience IDEX Chair Holder at Univ. Grenoble-Alpes
LargeDATA MIAI Chair Holder at Univ. Grenoble-Alpes

Born on March 18th, 1983 (36 years old). French

Table of Contents

General Information

Work positions

Université Grenoble–Alpes, Saint-Martin d'Hères, France.

GSTATS IDEX DataScience Chair Holder **April 2018 - Present**

LargeDATA MIAI Chair Holder **September 2019 - Present**

- Research in applied random matrix theory for large dimensional statistics and machine learning.
- Teaching in data science masters.

CentraleSupélec, Université ParisSaclay, Gif sur Yvette, France.

Full Professor **January 2016 - Present**

Assistant Professor **January 2011 - Present**

- Research in probability, statistics, machine learning, signal processing.
- Teaching in undergrad, master (SAR and MVA), PhD levels.

ST-Ericsson, Sophia Antipolis, France.

Development Engineer and PhD student **September 2007 - December 2010**

- Research in random matrix theory.
- Applications to 3GPP-LTE Advanced and MIMO standards.

Languages

French (mother language), English (fluent), German (school level).

University titles

Diplomas

Université d'Orsay, Saclay, France. **January 2011 - February 2015**

HDR diploma in Physics

- Title : Robust estimation methods in the large random matrix regime
- Jury : A. Hero, L. Pastur, J-Y. Tourneret (evaluators), F. Benaych-Georges, P. Bondon, M. McKay, E. Ollila.

CentraleSupélec, Gif sur Yvette, France. **January 2008 - November 2010**

Ph.D. in Physics (Telecommunications), November 2010

- Title : Application of random matrix theory to future wireless flexible networks
- Advisor : Mérouane Debbah
- Jury : P. Loubaton, X. Mestre (evaluators), M. Debbah, P. Duhamel, W. Hachem, A. Moustakas, J. Silverstein.

EURECOM, Sophia Antipolis, France. **September 2005 - June 2007**

M.S., Telecommunication Engineering, September 2007

Mobile communications, embedded systems, computer science.

Telecom ParisTech, Paris, France. **September 2004 - June 2007**

M.S., System of Communications (SiCom), (mention TB), March 2008

Wireless communications, image processing, blind detection techniques.

Education

Lycée Louis le Grand, Paris, France. **September 2001 - June 2004**

Preparation to Engineering School

Teaching Activities

Teaching service

[2011-2018] Full-time (assistant) professor at CentraleSupélec **∼240h ETD**

[2018-2020] Part-time (20%) professor at CentraleSupélec & chair holder at UGA **∼80h ETD**

PhD and master levels

ENS ParisSaclay (Saclay, France) **since 2013**

- Random matrix theory and machine learning applications (Master MVA, lectures, 24hrs ETD)

Université Grenoble-Alpes (Grenoble, France) **since 2018**

- Introduction to Convex Optimization Theory (Master SIGMA, lectures, 20hrs ETD)
- Introduction to Scientific Writing (Master SIGMA, lectures, 9hrs ETD)
- Scientific Writing and Associated Softwares (PhD level, lectures, 24hrs ETD)

CentraleSupélec (Gif sur Yvette, France) **since 2011**

- Techniques of scientific writing (PhD students, lectures, 24hrs ETD)
- Introduction to random matrix theory (Master SAR, seminar lectures, 18hrs ETD)
- Theoretical foundations of flexible radio networks (Master SAR, seminar lectures, 18hrs ETD)

Undergraduate level

CentraleSupélec (Gif sur Yvette, France) **2011-2018**

- Statistical signal modelling (exercises, 2×12hrs ETD)
- Signals and Systems (exercises, 24hrs ETD)
- Digital and analog filtering (practical lectures, 32hrs ETD)
- Introduction to scientific writing (optional class, 18hrs ETD)

Interns

ENS Paris Saclay **since 2011**

- Master internship (4 to 6 months, up to 4 students/year)

CentraleSupélec (Gif sur Yvette, France) **2011-2018**

- Long master project (master SAR, 3-month projects, 2 students/year)
- Telecom major project (3rd-year students, 3-month projects, 3 students/year)
- Projets de conception (1st-year undergraduate students, 2-month projects, 6 students/year)
- Projets de synthèse (2nd-year undergraduate students, 2-month projects, 4 students/year)

Admin

CentraleSupélec (Gif sur Yvette, France) **2015-2018**

- CentraleSupélec engineering program definition
- CentraleSupélec strategical roadmap definition

COURSE ORGANIZATION

Grenoble INP (Grenoble, France) **since 2018**

- Creation of new Convex Optimization course (master level ; lectures, TPs, and exams)
- Creation of Scientific Writing class (Master and PhD levels)

CentraleSupélec (Gif-sur-Yvette, France) **since 2012**

- Organization of master SAR seminars (lectures and exams)
- Organization of the scientific writing classes

ENS ParisSaclay (Saclay, France) **since 2013**

- Organization of master MVA class (lectures, practical works, and exam)

PERIPHERAL ACTIVITIES

Concours CPGE Centrale–Supélec **2016-2018**

- Evaluator at the "Concours CentraleSupélec" entrance examinations (80hrs)

Research Activities

GENERAL INFORMATION

Fields of interest : statistics, signal processing, machine learning, wireless communications, graph theory, random matrix theory, robust statistics, statistical finance, miscellaneous applications (natural language processing, virology, medical data).

Publication record (Google scholar figures as of March 2020)

Overall	1 book, 3 chapters, 50+ journals, 85+ conferences, 6 patents/innovative ideas.
Citations	3400+ (five best : 723, 542, 145, 110, 88)
h-index	27
i10-index	58

AWARDS

CNRS Bronze Medal. **2013**

2013 CNRS Bronze Medal in subsection "science of information and its interaction"

Awarding my work in Signal Processing and Wireless Communications as a young researcher since 2008.

IEEE ComSoc Young Researcher Award. **2013**

2013 IEEE ComSoc Outstanding Young Researcher Award for the EMEA Region

Awarding my work in Communications-related topics as a young researcher since 2008.

Best PhD Thesis Award. **2011**

EEA/GdR ISIS/GRETSI 2011 Award of the Best 2010 Thesis

Price for my PhD thesis "Application of random matrix theory to future wireless flexible networks"

Best Student Paper Award. **2019**

Best Student Paper Award of EUSIPCO 2019 conference

T. Mali, R. Couillet, "Random Matrix-Improved Estimation of the Wasserstein Distance between two Centered Gaussian Distributions"

Best Student Paper Award. **2013**

Second prize of the 2012-2013 IEEE Australia Council Student Paper Contest

G. Geraci, R. Couillet, J. Yuan, M. Debbah, I. B. Collings, "Large System Analysis of Linear Precoding in MISO Broadcast Channels with Confidential Messages"

Best Student Paper Award. **2011**

Best Student Paper Award Final of the 2011 Asilomar Conference

J. Hoydis, R. Couillet, M. Debbah, "Asymptotic Analysis of Double-Scattering Channels"

Best Student Paper Award. **2008**

Best Student Paper Award of the 2008 ValueTools Conference

R. Couillet, S. Wagner, M. Debbah, A. Silva, "The Space Frontier : Physical Limits of Multiple Antenna Information Transfer"

Keynote speaker. **2016**

ACM RACS 2016 conference at Odense, Denmark.

Random matrices and Machine Learning.

Tutorial speaker. **2010–2017**

Ten times tutorial speaker at international conferences.

EUSIPCO'2019, EUSIPCO'2018, IEEE ICASSP'2017, IEEE ICASSP'2015, EUSIPCO'2014, IEEE SPAWC'2013, IEEE ICASSP'2011, CrownCom'2010, European Wireless'2010.

Projects

On-going projects.

Project Name	Contribution	Period
ANR MIAI LargeDATA chair	50% (PI)	2019-2023
IDEX GSTATS chair	100% (PI)	2018-2020
ANR DARLING	50% (co-PI)	2019-2022
HUAWEI RMT4AI	100% (PI)	2019-2021

Previous projects.

Project Name	Contribution	Period
ANR RMT4GRAPH	100% (PI)	2014-2019
ERC MORE	50%	2012-2017
ParisSaclay RMT4ML	100% (PI)	2017-2020
Fondation Supélec DeepRMT	100% (PI)	2017-2020
Singapore MERLION	100% (PI)	2015-2017
Mastodons AGADIR	20%	2017
HUAWEI RMTin5G	100% (PI)	2015-2016
ANR DIONISOS	25% (co-PI)	2012-2016
ANR SESAME	20%	2008-2012
FP7 NEWCOM#	10%	2012-2015
FP7 NEWCOM++	10%	2009-2011

Community Activities

Jury CRCN-INRIA **since 2018**

Jury member for young scientist recruitment (CRCN) at INRIA-GRA (Grenoble center), in 2018, 2019, 2020

European Research Council **since 2018**

Reviewer for ERC-StG grants (phase 1)

Agence Nationale pour la Recherche (ANR, France) **since 2015**

Reviewer for multiple French ANR projects (approx 2/year since 2015)

Fonds de la Recherche Scientifique (FNRS, Belgium) **since 2018**

Reviewer for multiple Belgium FRS-FNRS projects (approx 2/year since 2018)

Academy of Finland **2015**

Reviewer for the Academy of Finland (reviews + 2-day decision in Helsinki)

L'Oréal-UNESCO for Women in Science **since 2020**

Evaluator for the 2021 International Prize

Evaluator for the 2020 Young French Talents Prize

IEEE (Senior member) **since 2007**

IEEE TSP Associate Editor (2015–2019)

IEEE STPM Technical Committee Member (since 2015)

World Scientific **since 2019**

Associate Editor of RMTA (since 2019)

GRETSI (member) **since 2011**

Member of the GRETSI association (since 2011)

Editor for special issue "Matrices aléatoires" of TS magazine (2015)

Other Research Groupes **since 2016**

Member of the GdR MEGA, AMIES main contact

Workshop Organization

Special Session Chair of IEEE CAMSAP 2019.

Technical Area Chair of Asilomar Conference 2016.

Special session and workshop organizer :

- *International* : Special Session "Random Matrices in Signal Processing and Machine Learning" (IEEE SSP, 2016), Special Session "Random Matrix Advances in Signal Processing" (IEEE SSP, 2014), Special Session "Random Matrices and Applications" (IEEE Asilomar, 2013).
- *France* : GdR day "Random matrix advances in large dimensional statistics and machine learning" (2017), GdR day "Estimation et traitement statistique en grande dimension" (2013).

Member of the GDR MEGA (on random matrices)

PhD Students

On-going PhD theses.

Charles Séjourné **2020-2023**

- fully advised by myself
- "Non-convex optimization in large dimensional machine learning"
- Defense expected in September 2023
- Grant : HUAWEI RMT4AI project

Tayeb Zarrouk **2020-2023**

- co-advised (50%) with F. Chatelain and N. LeBihan.
- "Structured models in machine learning and the complexity-performance trade-off"
- Defense expected in September 2023
- Grant : MIAI LargeDATA chair, Grenoble

Bernard Nabet **2019-2025**

- fully advised by myself
- "Large dimensional machine learning methods for statistical finance"
- Defense expected in 2025
- Grant : self-financed, part-time PhD

Cyprien Doz **2019-2022**

- co-advised (33%) with C. Ren and J.-P. Ovarlez
- "Large dimensional machine learning for radars"
- Defense expected in 2022
- Grant : SONDRA-lab at CentraleSupélec

Lorenzo Dall'Amico **2018-2021**

- co-advised (50%) with N. Tremblay and T. Hueber.
- "Statistical physics and graph mining"
- Defense expected in September 2021
- Grant : ED EEATS grant, Grenoble

Malik Tiomoko **2018-2021**

- fully advised by myself
- "Large dimensional statistical learning"
- Defense expected in September 2021
- Grant : ED STIC grant, ParisSaclay

Mohamed Seddik **2017-2020**

- fully advised by myself
- "Random matrices for deep networks"
- Defense expected in September 2020
- Grant : CEA-list

Cosme Louart (thèse 100%) **2017-2018**

- fully advised by myself
- "Concentration of measure in large random matrices"
- Grant : CEA List

Completed PhD theses.

Zhenyu Liao (80%) **2016-2019**

- "Random matrix theory for machine learning and neural networks"
- Grant : Supélec Fondation Grant DeepRMT
- Contribution : 4 articles (AAP,2×IEEE TSP,Journal of Diff.Equ.), 6 conferences (2×ICML,2×ICASSP,CAMSAP,EUSIPCO)

Xiaoyi Mai (100%) **2016-2019**

- "Random matrix methods for large dimensional machine learning"
- Grant : University ParisSaclay Grant RMT4ML
- Contribution : 3 articles (2×JMLR,IEEE TSP), 5 conferences (3×ICASSP,EUSIPCO,Asilomar)

Hafiz Tiomoko Ali (100%) **2015-2018**

- "Community detection in large random graphs"
- Grant : ANR RMT4GRAPH
- Contribution : 2 articles (JMLR), 7 conferences (ICML,3×ICASSP,2×SSP,Asilomar), 1 tutorial (ICASSP)

Gil Katz (33%) **2013-2016**

- "Interactive Communication for Distributed Computing"
- Grant : ERC MORE
- Contribution : 2 articles (IEEE IT,Annals of Stats), 6 conference (Asilomar, Allerton, ISIT)

Azary Abboud (50%) **2012-2015**

- "Distributed optimization for Smart Grids"
- Grant : SUPELEC Foundation
- Contribution : 1 article (IEEE SIPN), 1 conference (ICASSP)
- Today postdoc at INRIA

Julia Vinogradova (50%) **2011-2014**

- "Large random matrices, statistical inference, and future wireless communication networks"
- Grant : DIGITEO
- Contribution : 2 articles (IEEE TSP), 2 conferences (EUSIPCO, ICASSP)
- Today postdoc at Linköping University

Axel Müller (50%) **2011-2014**

- "Random matrix applications to multi-cell wireless communication networks"
- Grant : Intel
- Contribution : 2 articles (TSP,IT), 4 conferences (Asilomar,ICASSP,GLOBECOM)
- Today research engineer at HUAWEI Labs, France.

Postdoctoral researchers

Henrique Goulart **since 2019**

- Research theme : Random tensors.

Franck Iutzeler (now Assistant-Professor at University Grenoble-Alpes) **2014-205**

- Research theme : Distributed asynchronous optimization.

Main Collaborations

Matthew M. McKay, Associated professor at HongKong UST, expert in multivariate analysis, statistics, biostatistics, finance, signal processing.

Jack W. Silverstein, Professor at North Carolina State University, expert in random matrix theory.

Alfred O. Hero, Professor at University of Michigan, expert in statistics, data science and signal processing.

Florent Benaych-Georges, Professor at Université de Paris Descartes and Ecole Polytechnique, expert in random graph and random matrix theory.

Gilles Wainrib, Assistant professor at ENS Paris, expert in machine learning and random neural networks.

Walid Hachem, Research Director at CNRS, expert in random matrix theory and applications.

Mérouane Debbah, Professor at CentraleSupélec, expert in random matrix theory and wireless communications.

Abla Kammoun, Research scientist at KAUST University, expert in random matrix theory, signal processing and statistics.

Visiting Appointments

Hong-Kong University of Science and Technology, Hong-Kong. **June 2014**

Collaboration work with Professor M. McKay,

- Department of Electronic and Computer Engineering
- Project Topic : robust estimation in large financial data

North Carolina State University, North Carolina, USA. **Nov. 2009**

Collaboration work with Professor J. W. Silverstein,

- Department of Mathematics
- Project Topic : random matrix theory for multi-source energy detection

Publications in Journals

J1. C. Louart, R. Couillet "A Concentration of Measure Approach to Large Dimensional Robust Statistics", (submitted to) Annals of Applied Probability, 2020.

J2. K. Elkhalil, A. Kammoun, R. Couillet, T. Al-Naffouri, M-S. Alouini, "A Large Dimensional Analysis of Regularized Discriminant Analysis Classifiers" (to appear) IEEE Transactions on Signal Processing, 2020.

J3. L. Dall'Amico, R. Couillet, N. Tremblay, "A unified framework for spectral clustering in sparse graphs", (submitted to) Journal of Machine Learning Research, 2020.

J4. C. Louart, R. Couillet, "Concentration of Measure and Large Random Matrices with an application to Sample Covariance Matrices", (submitted to) Random Matrix Theory and Applications, 2019.

J5. X. Mai, R. Couillet, "Consistent Semi-Supervised Graph Regularization for High Dimensional Data", (submitted) Journal of Machine Learning Research, 2019.

J6. R. Couillet, M. Tiomoko, S. Zozor, E. Moisan, "Random matrix-improved estimation of covariance matrix distances", Journal of Multivariate Analysis, vol. 174, pp. 104531, 2019.

J7. X. Mai, R. Couillet, "A Random Matrix Analysis and Improvement of Semi-Supervised Learning for Large Dimensional Data", Journal of Machine Learning Research, vol. 19, no. 79, pp. 1-27, 2018.

J8. A. Kammoun, R. Couillet, "Subspace Kernel Clustering of Large Dimensional Data" (submitted to) Annals of Applied Probability, 2017.

J9. L. Yang, M. McKay, R. Couillet, "High-Dimensional MVDR Beamforming: Optimized Solutions based on Spiked Random Matrix Models", IEEE Transactions on Signal Processing, vol. 66, no. 1, pp. 1933-1947, 2018.

J10. A. Karadimitrakis, A. L. Moustakas, R. Couillet, "Gallager Bound for MIMO Channels: Large-N Asymptotics" IEEE Transactions on Wireless Communications, vol. 17, no. 2, pp. 1323-1330, 2018.

J11. N. Auguin, D. Morales, M. McKay, R. Couillet, "Large-dimensional behavior of regularized Maronna's M-estimators of covariance matrices" IEEE Transactions on Signal Processing, vol. 66, no. 13, pp. 3529–3542, 2018.

J12. C. Louart, Z. Liao, R. Couillet, "A Random Matrix Approach to Neural Networks" Annals of Applied Probability, vol. 28, no. 2, pp. 1190–1248, 2018.

J13. Z. Liao, R. Couillet, "A Large Dimensional Analysis of Least Square Support Vector Machines" IEEE Transactions on Signal Processing, vol. 67, no. 4, pp. 1065-1074, 2018.

J14. R. Couillet, H. Tiomoko Ali, "Improved spectral community detection in large heterogeneous networks" Journal of Machine Learning Resarch, vol. 18, no. 225, pp. 1–49, 2018.

J15. R. Couillet, M. McKay, "Optimal block-sparse PCA for high dimensional correlated samples" (submitted to) Journal of Multivariate Analysis, 2016.

J16. R. Couillet, G. Wainrib, H. Sevi, H. Tiomoko Ali, "The asymptotic performance of linear echo state neural networks" Journal of Machine Learning Research, vol. 17, no. 178, pp. 1-35, 2016.

J17. R. Couillet, F. Benaych-Georges, "Kernel Spectral Clustering of Large Dimensional Data" Electronic Journal of Statistics, vol. 10, no. 1, pp. 1393-1454, 2016.

J18. F. Benaych-Georges, R. Couillet, "Spectral Analysis of the Gram Matrix of Mixture Models" ESAIM : Probability and Statistics, DOI http ://dx.doi.org/10.1051/ps/2016007, 2016.

J19. R. Couillet, Estimation robuste et matrices aléatoires, revue Traitement du Signal, vol. 33, no. 2-3, pp. 273-320, 2016.

J20. R. Couillet, G. Wainrib, Perspectives en matrices aléatoires et grands réseaux, revue Traitement du Signal, vol. 33, no. 2-3, pp. 351-376, 2016.

J21. M. Sadeghi, L. Sanguinetti, R. Couillet, Y. Chau, "Large System Analysis of Power Normalization Techniques in Massive MIMO", IEEE Transactions on Vehicular Technologies, vol. 66, no. 10, pp. 9005-9017, 2017.

J22. M. Sadeghi, L. Sanguinetti, R. Couillet, Y. Chau, "Reducing the Computational Complexity of Multicasting in Large-Scale Antenna Systems", IEEE Transactions on Wireless Communications, vol. 16, no. 5, pp. 2963-2975, 2017.

J23. L. Sanguinetti, R. Couillet, M. Debbah, "Large System Analysis of Base Station Cooperation for Power Minimization" IEEE Transactions on Wireless Communications, vol. 15, no. 8, pp. 5480-5496, 2016.

J24. A. Abboud, F. Iutzeler, R. Couillet, H. Siguerdidjane, M. Debbah, "Distributed Production-Sharing Optimization and Application to Power Grid Networks," IEEE Transactions on Signal and Information Processing over Networks, vol. 2, no. 1, pp. 1628, 2016.

J25. A. Kammoun, R. Couillet, F. Pascal, M.-S. Alouini, "Optimal Design of the Adaptive Normalized Matched Filter Detector using Regularized Tyler Estimator" IEEE Transactions on Aerospace and Electronic Systems, vol. 54, no. 2, pp. 755–769, 2018.

J26. A. Kammoun, R. Couillet, F. Pascal, M.-S. Alouini, "Convergence and Fluctuations of Regularized Tyler Estimators" IEEE Transactions on Signal Processing, vol. 64, no. 4, pp. 1048-1060, 2016.

J27. D. Morales-Jimenez, R. Couillet, M. McKay, "Large Dimensional Analysis of Robust M-Estimators of Covariance with Outliers" IEEE Transactions on Signal Processing, vol. 63, no. 21, pp. 5784-5797, 2015.

J28. L. Yang, R. Couillet, M. McKay, "A Robust Statistics Approach to Minimum Variance Portfolio Optimization" IEEE Transactions on Signal Processing, vol. 63, no. 24, pp. 6684–6697, 2015.

J29. R. Couillet, A. Kammoun, F. Pascal, "Second order statistics of robust estimators of scatter. Application to GLRT detection for elliptical signals" Elsevier Journal of Multivariate Analysis, vol. 143, pp. 249-274, 2015.

J30. A. Müller, R. Couillet, E. Bjørnson, S. Wagner, M. Debbah, "Interference-Aware RZF Precoding for Multi-Cell Downlink Systems" IEEE Transactions on Signal Processing, vol. 63, no. 15, pp. 3959-3973 2015.

J31. R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator" Elsevier Journal of Multivariate Analysis, vol. 140, pp. 139-161, 2015.

J32. R. Couillet, M. McKay, "Large Dimensional Analysis and Optimization of Robust Shrinkage Covariance Matrix Estimators" Elsevier Journal of Multivariate Analysis, vol. 131, pp. 99-120, 2014.

J33. Y. Chitour, R. Couillet, F. Pascal "On the convergence of Maronna's M-estimators of scatter" IEEE Signal Processing Letters, vol. 22, no. 6, pp. 709-712, 2014.

J34. R. Couillet, F. Pascal, J. W. Silverstein, "The Random Matrix Regime of Maronna's M-estimator with elliptically distributed samples", vol. 139, pp. 56-78, Elsevier Journal of Multivariate Analysis, 2015.

J35. J. Vinogradova, R. Couillet, W. Hachem, "Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection large", IEEE Transactions on Signal Processing, vol. 63, no. 18, pp. 4903-4913, 2015.

J36. R. Couillet, W. Hachem, "Analysis of the limiting spectral measure of large random matrices of the separable covariance type", Random Matrix Theory and Applications, vol. 3, pp. 1-23, 2014.

J37. J. Hoydis, R. Couillet, P. Piantanida, "The Second-Order Coding Rate of the MIMO Rayleigh Block-Fading Channel," IEEE Transactions on Information Theory, vol. 61, no. 12, pp. 6591-6622, 2015.

J38. J. Vinogradova, R. Couillet, W. Hachem, "Statistical Inference in Large Antenna Arrays under Unknown Noise Pattern," IEEE Transactions on Signal Processing, vol. 61, no. 22, pp. 5633-5645, 2013.

J39. F. Chapon, R. Couillet, W. Hachem, X. Mestre, "The outliers among the singular values of large rectangular random matrices with additive fixed rank deformation," Markov Processes and Related Fields, vol. 20, pp. 183-228, 2014.

J40. R. Couillet, F. Pascal, J. W. Silverstein, "Robust Estimates of Covariance Matrices in the Large Dimensional Regime," IEEE Transactions on Information Theory, vol. 60, no. 11, 2014.

J41. G. Geraci, R. Couillet, J. Yuan, M. Debbah, I. B. Collings, "Large System Analysis of Linear Precoding in MISO Broadcast Channels with Confidential Messages," IEEE Journal on Selected Area in Communications, vol. 31, no. 9, pp. 1660-1671, 2013. **Second prize of the 2012-2013 IEEE Australia Council Student Paper Contest.**

J42. J. Hoydis, R. Couillet, M. Debbah, "Iterative Deterministic Equivalents for the Capacity Analysis of Communication Systems," Technical Report.

J43. R. Couillet, S. Medina Perlaza, H. Tembine, M. Debbah, "Electrical Vehicles in the Smart Grid: A Mean Field Game Analysis," IEEE Journal on Selected Areas in Communications : Smart Grid Communications Series, vol. 30, no. 6, pp. 1086-1096, 2012.

J44. J. Yao, R. Couillet, J. Najim, M. Debbah, "Fluctuations of an Improved Population Eigenvalue Estimator in Sample Covariance Matrix Models," IEEE Transactions on Information Theory, vol. 59, no. 2, pp. 1149-1163, 2013.

J45. R. Couillet, M. Debbah,"Signal Processing in Large Systems: a New Paradigm," IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 24-39, 2013.

J46. R. Couillet, W. Hachem,"Fluctuations of spiked random matrix models and failure diagnosis in sensor networks," IEEE Transactions on Information Theory, vol. 59, no. 1, pp. 509-525, 2013.

J47. A. Kammoun, R. Couillet, J. Najim, M. Debbah,"Performance of capacity inference methods under colored interference," IEEE Transactions on Information Theory, vol. 59, no. 2, pp. 1129-1148, 2013.

J48. R. Couillet, J. Hoydis, M. Debbah, "Random beamforming over quasi-static and fading channels: A deterministic equivalent approach," IEEE Transactions on Information Theory, vol. 58, no. 10, pp. 6392-6425, 2012.

J49. S. Wagner, R. Couillet, M. Debbah, D. T. M. Slock, "Large System Analysis of Linear Precoding in MISO Broadcast Channels with Limited Feedback", IEEE Transactions on Information Theory, vol. 58, no. 7, pp. 4509-4537, 2012.

J50. R. Couillet, J. W. Silverstein, Z. Bai, M. Debbah, "Eigen-Inference for Energy Estimation of Multiple Sources", IEEE Transactions on Information Theory, vol. 57, no. 4, pp. 2420-2439, 2011.

J51. R. Couillet, M. Debbah, J. W. Silverstein, "A Deterministic Equivalent for the Analysis of Correlated MIMO Multiple Access Channels", IEEE Transactions on Information Theory, vol. 57, no. 6, pp. 3493-3514, 2011.

J52. R. Couillet, M. Debbah, "A Bayesian Framework for Collaborative Multi-Source Signal Sensing", IEEE Transactions on Signal Processing, vol. 58, no. 10, pp. 5186-5195, 2010.

J53. R. Couillet, A. Ancora, M. Debbah, "Bayesian Foundations of Channel Estimation for Cognitive Radios", Advances in Electronics and Telecommunications, vol. 1, no. 1, pp. 41-49, 2010.

J54. R. Couillet, M. Debbah,"Le téléphone du futur : plus intelligent pour une exploitation optimale des fréquences" Revue de l'Electricité et de l'Electronique, no. 6, pp. 71-83, 2010.

J55. R. Couillet, M. Debbah, "Mathematical foundations of cognitive radios", Journal of Telecommunications and Information Technologies, no. 4, 2009.

J56. R. Couillet, M. Debbah, "Outage performance of flexible OFDM schemes in packet-switched transmissions", Eurasip Journal on Advances on Signal Processing, Volume 2009, Article ID 698417, 2009.

Conference Publications

C1. M. Seddik, R. Couillet, M. Tamaazousti, "A Random Matrix Analysis of Learning with α-Dropout", International Conference on Machine Learning (ICML'20), Artemiss workshop, Graz, Autria, 2020.

C2. M. Seddik, C. Louart, R. Couillet, M. Tamaazousti, "The Unexpected Deterministic and Universal Behavior of Large Softmax Classifiers", (submitted to) Conference on Neural Information Processing Systems (NeurIPS'20), Vacouver, Canada, 2020.

C3. L. Dall'Amico, R. Couillet, N. Tremblay, "Community detection in sparse time-evolving graphs with a dynamical Bethe-Hessian", (submitted to) Conference on Neural Information Processing Systems (NeurIPS'20), Vacouver, Canada, 2020.

C4. Z. Liao, R. Couillet, M. Mahoney "A random matrix analysis of random Fourier features: beyond the Gaussian kernel, a precise phase transition, and the corresponding double descent", (submitted to) Conference on Neural Information Processing Systems (NeurIPS'20), Vacouver, Canada, 2020.

C5. T. Zarrouk, R. Couillet, F. Chatelain, N. Le Bihan, "Performance-Complexity Trade-Off in Large Dimensional Statistics", International Workshop on Machine Learning for Signal Processing (MLSP'20), Espoo, Finland, 2020.

C6. M. Seddik, R. Couillet, M. Tamaazousti "Random Matrix Theory Proves that Deep Learning Representations of GAN-data Behave as Gaussian Mixtures", International Conference on Machine Learning (ICML'20), Graz, Autria, 2020.

C7. M. Tiomoko, H. Tiomoko, R. Couillet "Deciphering and Optimizing Multi-Task and Transfer Learning: a Random Matrix Approach", (submitted to) International Conference on Machine Learning (ICML'20), Graz, Austria, 2020.

C8. L. Dall'Amico, R. Couillet, N. Tremblay "Optimal Laplacian Regularization for Sparse Spectral Community Detection", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'20), Barcelona, Spain, 2020.

C9. M. Tiomoko, C. Louart, R. Couillet "Large Dimensional Asymptotics of Multi-Task Learning", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'20), Barcelona, Spain, 2020.

C10. L. Dall'Amico, N. Tremblay, R. Couillet "Optimized Deformed Laplacian for Spectrum-based Community Detection in Sparse Heterogeneous Graphs", Neural Information Processing Systems (NeurIPS'19), Vancouver, Canada, 2019.

C11. Z. Liao, R. Couillet, "On Inner-product Kernels of High Dimensional Data", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

C12. R. Couillet, "High Dimensional Robust Classification: A Random Matrix Analysis", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

C13. R. Couillet, "A Random Matrix Analysis and Optimization Framework to Large Dimensional Transfer Learning", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

C14. A. Kadavankandy, R. Couillet, "Asymptotic Gaussian Fluctuations of Spectral Clustering Eigenvectors", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

C15. C. Louart, R. Couillet, "A concentration of measure perspective to robust statistics", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

C16. M. Tiomoko, R. Couillet, "Estimation of Covariance Matrix Distances in the High Dimension Low Sample Size Regime", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

C17. M. Tiomoko, R. Couillet, "Random Matrix-Improved Estimation of the Wasserstein Distance between two Centered Gaussian Distributions", European Signal Processing Conference (EUSIPCO'19), A Coruna, Spain, 2019. **Best student paper award**

C18. M. Seddik, M. Tamaazousti, R. Couillet, "Kernel Random Matrices of Large Concentrated Data: The Example of GAN-Generated Image", .IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

C19. M. Tiomoko, F. Bouchard, G. Ginholac, R. Couillet "Random Matrix Improved Covariance Estimation for a Large Class of Metrics", International Conference on Machine Learning (ICML), Long Beach, USA, 2019.

C20. L. Dall'Amico, R. Couillet "Community Detection in Sparse Realistic Graphs: Improving the Bethe Hessian", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

C21. X. Mai, R. Couillet "Revisiting and Improving Semi-Supervised Learning: A Large Dimensional Approach", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

C22. H. Tiomoko Ali, S. Liu, Y. Yilmaz, R. Couillet, I. Rajapakse, A. Hero, "Latent Heterogeneous Multilayer Community Detection", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

C23. Z. Liao, X. Mai, R. Couillet "A Large n, p Analysis of Logistic Regression: Asymptotic Performance and New Insights", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

C24. M. Tiomoko, R. Couillet, S. Zozor, E. Moisan, "Improved Estimation of the Distance between Covariance Matrices", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

C25. R. Couillet, Z. Liao, X. Mai, "Classification Asymptotics in the Random Matrix Regime", European Signal Processing Conference (EUSIPCO'18), Rome, Italy, 2018.

C26. M. Seddik, M. Tamaazousti, R. Couillet, "A Kernel Random Matrix-Based Approach for Sparse PCA", International Conference on Learning Representations (ICLR'19), New Orleans, USA, 2019.

C27. X. Mai, R. Couillet, "Semi-Supervised Spectral Clustering", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2018.

C28. Z. Liao, R. Couillet, "The Dynamics of Learning: A Random Matrix Approach", International Conference on Machine Learning, Stockholm, Sweden, 2018.

C29. Z. Liao, R. Couillet, "On the Spectrum of Random Features Maps of High Dimensional Data", International Conference on Machine Learning, Stockholm, Sweden, 2018.

C30. H. Tiomoko Ali, A. Kammoun, R. Couillet, "Random matrix-improved kernels for large dimensional spectral clustering", Statistical Signal Processing Workshop (SSP'18), Freiburg, Germany, 2018.

C31. L. Yang, M. R. McKay, R. Couillet, "Random Matrix-Optimized High-Dimensional MVDR Beamforming", Statistical Signal Processing Workshop (SSP'18), Freiburg, Germany, 2018.

C32. C. Louart, R. Couillet, "A Random Matrix and Concentration Inequalities Framework for Neural Networks Analysis", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'18), Calgary, Canada, 2018.

C33. H. Tiomoko Ali, A. Kammoun, R. Couillet, "Random matrix asymptotic of inner product kernel spectral clustering", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'18), Calgary, Canada, 2018.

C34. K. Elkalil, A. Kammoun, R. Couillet, T. Al-Naffouri, M.-S. Alouini, "Asymptotic Performance of Regularized Quadratic Discriminant Analysis Based Classifiers", IEEE International Workshop on Machine Learning for Signal Processing (MLSP'17), Roppongi, Tokyo, Japan, 2017. **Best student paper award finalist**

C35. Z. Liao, R. Couillet, "Random matrices meet machine learning: a large dimensional analysis of LS-SVM", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17), New Orleans, USA, 2017.

C36. X. Mai, R. Couillet, "The counterintuitive mechanism of graph-based semi-supervised learning in the big data regime", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17), New Orleans, USA, 2017.

C37. C. Louart, R. Couillet, "Harnessing neural networks: a random matrix approach", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17), New Orleans, USA, 2017.

C38. H. Tiomoko Ali, R. Couillet, "Random Matrix Improved Community Detection in Heterogeneous Networks", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2016.

C39. R. Couillet, A. Kammoun, "Random Matrix Improved Subspace Clustering", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2016.

C40. R. Couillet, G. Wainrib, H. Sevi, H. Tiomoko Ali, "A Random Matrix Approach to-Recurrent Neural Networks", International Conference on Machine Learning (ICML), New York, USA, 2016.

C41. A. Kammoun, R. Couillet, F. Pascal, M. Slim-Alouini, "Optimal Design of Adaptive Normalized Matched Filter For Large Antenna Arrays", IEEE Statistical Signal Processing Workshop (SSP), Palma de Majorca, Spain, 2016.

C42. N. Auguin, D. Morales, M. R. McKay, R. Couillet, "Robust Shrinkage M-estimators of Large Covariance Matrices", IEEE Statistical Signal Processing Workshop (SSP), Palma de Majorca, Spain, 2016.

C43. R. Couillet, G. Wainrib, H. Sevi, H. Tiomoko Ali, "Training performance of echo state neural networks", IEEE Statistical Signal Processing Workshop (SSP), Palma de Majorca, Spain, 2016.

C44. H. Tiomoko Ali, R. Couillet, "Performance analysis of spectral community detection in realistic graph models", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'16), Shangai, China, 2016.

C45. R. Couillet, F. Benaych-Georges, "Understanding Big Data Spectral Clustering", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, Mexico, 2015.

C46. L. Yang, R. Couillet, M. R. McKay, "Minimum Variance Portfolio Optimization in the Spiked Covariance Model", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, Mexico, 2015.

C47. L. Sanguinetti, R. Couillet, M. Debbah, "Base Station Cooperation for Power Minimization in the Downlink: Large System Analysis", IEEE Global Communications Conference (GLOBECOM'15), San Diego, USA, 2015.

C48. R. Couillet, M. S. Greco, J-P. Ovarlez, F. Pascal, "RMT for Whitening Space Correlation and Applications to Radar Detection", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, Mexico, 2015.

C49. D. Morales-Jimenez, R. Couillet, M. McKay,"Large dimensional analysis of Maronna's M-estimator with outliers", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15), Brisbane, Australia, 2015.

C50. A. Kammoun, R. Couillet, F. Pascal, "Second order statistics of bilinear forms of robust scatter estimators", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15), Brisbane, Australia, 2015.

C51. G. Katz, P. Piantanida, R. Couillet, "Joint Estimation and Detection Against Independence", Fifty-second Allerton Conference on Communication, Control, and Computing, Allerton, IL, USA, 2014.

C52. R. Couillet, M. McKay, "Robust covariance estimation and linear shrinkage in the large dimensional regime", IEEE International Workshop on Machine Learning for Signal Processing (MLSP'14), Reims, France, 2014.

C53. L. Yang, R. Couillet, M. McKay, "Minimum variance portfolio optimization with robust shrinkage covariance estimation", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2014.

C54. P. Vallet, X. Mestre, Ph. Loubaton, R. Couillet, "Asymptotic analysis of Beamspace-MUSIC in the context of large arrays", IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM'14), A Coruna, Spain, 2014.

C55. R. Couillet, A. Kammoun, "Robust G-MUSIC", European Signal Processing Conference (EUSIPCO'14), Lisbon, Portugal, 2014.

C56. J. Vinogradova, R. Couillet, W. Hachem, "Estimation of Large Toeplitz Covariance Matrices and Application to Source Detection", European Signal Processing Conference (EUSIPCO'14), Lisbon, Portugal, 2014.

C57. R. Couillet, F. Pascal, "Robust M-estimator of scatter for large elliptical samples", IEEE Worshop on Statistical Signal Processing (SSP'14), Gold Coast, Australia, 2014.

C58. A. Abboud, R. Couillet, M. Debbah, H. Siguerdidjane, "Asynchronous alternating direction method of multipliers applied to the direct-current optimal power flow problem," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'14), Florence, Italy, 2014.

C59. A. Pelletier, R. Couillet, J. Najim, "Second-Order Analysis of the Joint SINR distribution in Rayleigh Multiple Access and Broadcast Channels," Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2013.

C60. A. Müller, E. Björnson, R. Couillet, M. Debbah, "Analysis and management of heterogeneous user mobility in large-scale downlink systems," Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2013.

C61. J. Hoydis, R. Couillet, P. Piantanida, "Bounds on the Second-Order Coding Rate of the MIMO Rayleigh Block-Fading Channel," IEEE International Symposium on Information Theory, Istanbul, Turkey, 2013.

C62. G. Geraci, R. Couillet, J. Yuan, M. Debbah, I. Collings, "Secrecy Sum-Rates with Regularized Channel Inversion Precoding under Imperfect CSI at the Transmitter," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), Vancouver, Canada, 2013.

C63. R. Couillet, F. Pascal, J. W. Silverstein, "A Joint Robust Estimation and Random Matrix Framework with Application to Array Processing," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), Vancouver, Canada, 2013.

C64. J. Vinogradova, R. Couillet, W. Hachem, "A new method for source detection, power estimation, and localization in large sensor networks under noise with unknown statistics," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), Vancouver, Canada, 2013.

C65. M. de Mari, R. Couillet, M. Debbah, "Concurrent data transmissions in green wireless networks: when best send one's packets?," (Invited paper) IEEE International Symposium on Wireless Communication Systems (ISWCS'12), Paris, France, 2012.

C66. A. Müller, J. Hoydis, R. Couillet, M. Debbah, "Optimal 3D Cell Planning: A Random Matrix Approach," IEEE Global Communications Conference (GLOBECOM'12), Anaheim, Caligornia, USA, 2012.

C67. J. Hoydis, R. Couillet, P. Piantanida, M. Debbah, "A Random Matrix Approach to the Finite Blocklength Regime of MIMO Fading Channels," IEEE International Symposium on Information Theory, Boston, Massachusetts, USA, 2012.

C68. M. Rezaee, R. Couillet, M. Guillaud, G. Matz, "Sum-Rate Optimization for the MIMO IC under Imperfect CSI: a Deterministic Equivalent Approach," IEEE International Workshop on Signal Processing Advances for Wireless Communications, Cesme, Turkey, 2012.

C69. J. Hoydis, A. Müller, R. Couillet, M. Debbah, "Analysis of Multicell Cooperation with Random User Locations Via Deterministic Equivalents," Eighth Workshop on Spatial Stochastic Models for Wireless Networks, Paderborn, Germany, 2012.

C70. R. Couillet, E. Zio, "A subspace approach to fault diagnostics in large power systems" (Invited Paper) IEEE International Symposium on Communications, Control, and Signal Processing (ISCCSP'12), Rome, Italy, 2012.

C71. A. Kammoun, M. Kharouf, R. Couillet, J. Najim, M. Debbah, "On the fluctuations of the SINR at the output of the Wiener filter for non centered channels: the non Gaussian case," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'12), Kyoto, Japan, 2012.

C72. R. Couillet, P. Bianchi, J. Jakubowicz, "Decentralized convex stochastic optimization with few constraints in large networks," IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'11), San Juan, Puerto Rico, 2011.

C73. R. Couillet, S. Medina Perlaza, H. Tembine, M. Debbah, "A mean field game analysis of electric vehicles in the smart grid," IEEE International Conference on Computer Communications (INFOCOM'12), Orlando, FL, USA, 2012.

C74. J. Hoydis, R. Couillet, M. Debbah, "Asymptotic Analysis of Double-Scattering Channels," IEEE Asilomar Conference (ASILOMAR'11), Pacific Grove, CA, USA, 2011. **Best student paper award finalist**

C75. R. Couillet, W. Hachem, "Local Failure Localization in Large Sensor Networks," IEEE Asilomar Conference on Signals, Systems, and Computers (ASILOMAR'11), Pacific Grove, CA, USA, 2011.

C76. R. Couillet, M. Guillaud, "Performance of Statistical Inference Methods for the Energy Estimation of Multiple Sources," (Invited Paper) IEEE Statistical Signal Processing Workshop (SSP'11), Nice, France, 2011.

C77. A. Kammoun, R. Couillet, J. Najim, M. Debbah, "Performance of fast rate adaption techniques in interference-limited networks," IEEE Global Communications Conference (GLOBECOM'11), Houston, TX, USA, 2011.

C78. J. Yao, R. Couillet, J. Najim, E. Moulines, M. Debbah, "CLT for eigen-inference methods in cognitive radios," IEEE International Conference on Acoustics, Speech and Signal Processing, Prague, Czech Republic, 2011.

C79. J. Hoydis, R. Couillet, M. Debbah, "Deterministic Equivalents for the Performance Analysis of Isometric Random Precoded Systems," IEEE International Conference on Communications, Kyoto, Japan, 2011.

C80. J. Hoydis, J. Najim, R. Couillet, M. Debbah, "Fluctuations of the Mutual Information in Large Distributed Antenna Systems with Colored Noise," Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing, Allerton, IL, USA, 2010.

C81. R. Couillet, H. V. Poor, M. Debbah, "Self-organized spectrum sharing in large MIMO multiple-access channels," IEEE International Symposium on Information Theory, Austin TX, USA, 2010.

C82. R. Couillet, J. W. Silverstein, M. Debbah, "Eigen-inference for multi-source power estimation," IEEE International Symposium on Information Theory, Austin TX, USA, 2010.

C83. S. Wagner, R. Couillet, D. T. M. Slock, M. Debbah, "Optimal Training in Large TDD Multi-user Downlink Systems under Zero-forcing and Regularized Zero-forcing Precoding," IEEE Global Communication Conference, Miami, 2010.

C84. S. Wagner, R. Couillet, D. T. M. Slock, M. Debbah, "Large System Analysis of Zero-Forcing Precoding in MISO Broadcast Channels with Limited Feedback"

IEEE International Workshop on Signal Processing Advances for Wireless Communications, Marrakech, Morocco, 2010.

C85. R. Couillet, M. Debbah, "Information theoretic approach to synchronization: the OFDM carrier frequency offset example", Advanced International Conference on Telecommunications, Barcelona, Spain, 2010.

C86. R. Couillet, M. Debbah, "Uplink capacity of self-organizing clustered orthogonal CDMA networks in flat fading channels" IEEE Information Theory Workshop Fall'09, Taormina, Sicily, 2009.

C87. R. Couillet, M. Debbah, J. W. Silverstein, "Asymptotic Capacity of Multi-User MIMO Communications" IEEE Information Theory Workshop Fall'09, Taormina, Sicily, 2009.

C88. R. Couillet, M. Debbah, J. W. Silverstein, "Rate region of correlated MIMO multiple access channel and broadcast channel" IEEE Workshop on Statistical Signal Processing, Cardiff, Wales, UK, 2009.

C89. R. Couillet, M. Debbah, "Mathematical foundations of cognitive radios" U.R.S.I.'09, Warsaw, Poland, 2009.

C90. R. Couillet, M. Debbah, "A maximum entropy approach to OFDM channel estimation", IEEE International Workshop on Signal Processing Advances for Wireless Communications, Perugia, Italy, 2009.

C91. R. Couillet, M. Debbah, "Bayesian inference for multiple antenna cognitive receivers", IEEE Wireless Communications & Networking Conference, Budapest, Hungary, 2009.

C92. R. Couillet, M. Debbah, "Flexible OFDM schemes for bursty transmissions", IEEE Wireless Communications & Networking Conference, Budapest, Hungary, 2009.

C93. R. Couillet, S. Wagner, M. Debbah, "Asymptotic Analysis of Correlated Multi-Antenna Broadcast Channels", IEEE Wireless Communications & Networking Conference, Budapest, Hungary, 2009.

C94. R. Couillet, S. Wagner, M. Debbah, A. Silva, "The Space Frontier: Physical Limits of Multiple Antenna Information Transfer", ValueTools, Inter-Perf Workshop, Athens, Greece, 2008. **Best student paper award**

C95. R. Couillet, M. Debbah, "Free deconvolution for OFDM multicell SNR detection", IEEE Personal, Indoor and Mobile Radio Communications Symposium, Cognitive Radio Workshop, Cannes, France, 2008.

Books and book chapters

B1. R. Couillet, M. Debbah, **Random Matrix Methods for Wireless Communications**, Cambridge University Press, 2011. [book]

Theoretical random matrix tools (finite dimensional analysis, limiting spectral laws, free probability, deterministic equivalents, statistical inference) and applications to wireless communications (SU-MIMO, MU-MIMO, CDMA, detection, estimation, channel modelling).

B2. R. Couillet, M. Debbah, **Mathematical Foundations for Signal Processing, Communications and Networking**, CRC Press, Taylor & Francis Group, 2011 [book chapter]

Chapter "Random matrix theory" on random matrix theory and especially statistical inference methods.

B3. R. Couillet, M. Debbah, **Orthogonal Frequency Division Multiple Access Fundamentals and Applications**, Auerbach Publications, CRC Press, Taylor & Francis Group, 2010 [book chapter]

Chapter "Fundamentals of OFDMA Synchronization" on theoretical considerations and application tools for time offset and frequency offset regulation in OFDM and OFDMA systems.

B4. R. Couillet, M. Debbah, **Radio engineering : From software radio to cognitive radio**, John Wiley & Sons, 2013 [book chapter]

Several chapters on detection and estimation in cognitive radios.

Patents and Innovative Ideas

P1. R. Couillet, M. Debbah, **No. 08368028.0** "Process and apparatus for performing initial carrier frequency offset in an OFDM communication system"

P2. R. Couillet, M. Debbah, **No. 08368023.1** "Method for short-time OFDM transmission and apparatus for performing flexible OFDM modulation"

P3. R. Couillet, S. Wagner, **No. 09368025.4** "Precoding process for a transmitter of a MU-MIMO communication system"

P4. R. Couillet, **No. 09368030.4** "Process for estimating the channel in an OFDM communication system, and receiver for doing the same"

II1. R. Couillet, **Innovative Idea** "Fast Block Diagonalization Precoder"

II2. R. Couillet, **Innovative Idea** "User Subspace Clustering"

Tutorials

T1. R. Couillet, T. Malik, M. A. Seddik, "Random Matrix Advances in Large Dimensional Statistics, Machine Learning and Neural Nets", European Signal Processing Conference (EUSIPCO'19), A Coruna, 2019.

T2. R. Couillet, Z. Liao, X. Mai, "Random Matrix Advances for Machine Learning", European Signal Processing Conference (EUSIPCO'18), Rome, 2018.

T3. R. Couillet, H. Tiomoko Ali, "Random Matrices for Big Data Signal Processing and Machine Learning", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17), New Orleans, USA, 2017.

T4. R. Couillet, "Random Matrices, Robust Estimation, and Applications", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15), Brisbane, Australia, 2015.

T5. R. Couillet, A. Kammoun, "Future Random Matrix Tools for Large Dimensional Signal Processing", European Conference on Signal Processing (EUSIPCO), Lisbon, Portugal, 2014.

T6. R. Couillet, M. Debbah, "Random Matrix Advances in Signal Processing", IEEE International Workshop on Signal Processing Advances in Wireless Communications, Darmstadt, Germany, 2013.

T7. R. Couillet, M. Debbah, "Random Matrix Theory for Signal Processing Applications", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'11), Prague, Czech Republic, 2011.

T8. R. Couillet, M. Debbah, "Random Matrices in Wireless Flexible Networks", International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (Crowncom'10), Cannes, France, 2010.

T9. R. Couillet, M. Debbah, "Eigen-Inference Statistical methods for Cognitive Radio", European Wireless, Lucca, Italy, 2010.

Analytic Presentation of Works

RESEARCH AREAS

Applied mathematics :

- *random matrix theory* : Hermitian models, spiked models, complex models with non independent or non linear entries, kernel random matrices, etc.
- *mathematics* : probability, complex analysis, linear algebra
- *statistics* : robust estimation, kernel methods, principal component analysis, etc.

Graphs and automated learning :

- *graphs* : spectral methods of community detection
- *neural networks* : echo-state networks, extreme learning machines, deep networks, backpropagation
- *unsupervised learning* : kernel spectral methods (Ng–Weiss–Jordan type algorithms), semi-supervised learning, support vector machines

Signal processing :

- *array processing* : detection and estimation, subspace methods, robust statistics methods
- *sensor networks* : change point detection, distributed algorithms

Mobile communications :

- *information theory* : large MIMO system performance, finite blocklength communications
- *wireless communications* : massive MIMO precoder design, multi-cell and realistic systems
- *signal processing for wireless communications* : statistical inference for cognitive radios

SUMMARY OF RESEARCH ACTIVITIES

My research activities focus on the theoretical performance analysis and the algorithms and methods improvement in large dimensional systems and networks. The core tool of my research is the so-called random matrix theory, which more generally exploits the mathematical fields of probability theory, complex analysis, and linear algebra for large dimensional vectors and matrices.

My research started with the performance analysis of complex wireless communication systems (from 2007 to 2012 mainly). Among notable works, (J51) establishes for the first time the ergodic capacity region of multi-user MIMO systems in the uplink, based on new random matrix results. The performance of linear precoders for the same systems but in the downlink is then finely studied under realistic assumptions in (J49) which, in passing, generalizes numerous previous works. Other works, technically more involved, extend those results, notably in (J48,J42). A much more recent work, technically

deeper, (J37) analyzes the asymptotic probability of error in finite blocklength MIMO communication systems. On the 2010–2013 period, my research focus then steered towards signal processing for wireless communications, with in particular articles on the fast estimation of rates and positions of wireless user equipments in a cognitive radio context (J52,J50,J47,J44). All these works, undergone during and after my PhD thesis, were performed in collaboration with Jack Silverstein and Zhidong Bai, professors in mathematics and renown experts in random matrix theory. These led to one textbook (B1), to four patents (P1–P4), and to my being granted the IEEE Outstanding Young Researcher and French CNRS Bronze Medal Awards in 2013.

After this first wave of works, mathematically limited to the analysis of linear functionals of the spectrum of Gram matrices, I shifted my centers of interest to the statistical analysis of large dimensional systems, now for signal processing applications and in particular for array processing. Here the focus is on isolated eigenvalue-eigenvector pairs of given random matrix models (so-called spiked models). My works in this direction have mathematically generalized existing works to more realistic signal processing settings, as in particular with the works (J38,J39,J35) on joint spatial and time correlated models as well as with (J46) in the context of sensor networks.

Nonetheless, all these studies, be they useful and technically involved, were natural follow-ups on previous works. I then strived to tackle more challenging new problems. That led me to engage, on a three-year period, in the purely statistical study of so-called robust estimators of scatter in the large dimensional system regime. The mathematical objects and tools involved in this work significantly differ from conventional random matrix techniques, not adequate here. Among the main breakthroughs, the articles (J40,J34,J32,J29) establish a refined understanding of the behavior of these robust estimators then leading to an improved method for exploiting them in large dimensional settings ; applications specific to array processing (J31) or robust analysis of financial data (J28) then unfolded. A large part of these works was developed jointly with Matthew McKay and Frederic Pascal, both professors specialists in array processing. These results also formed the core of my HDR (qualification to advise research) thesis with professors Leonid Pastur (mathematician, expert in random matrix theory) and Jean-Yves Tourneret and Alfred Hero (both world leading experts in signal processing and statistics) as main reviewers.

Today, the growing interest for automated learning for handling big datasets motivated the deeper understanding and improvement of dedicated methods in the large dimensional regime. In the scope of the ANR Jeunes Chercheurs project (ANR RMT4GRAPH) I lead, within my team (composed of 7 PhD students and postdocs) now study classification algorithms, neural nets-based methods as well as inference on graphs. Our results since 2014 on this topic are very promising and allow in particular for a better understanding of kernel spectral clustering algorithms (J17,J18), semi-supervised learning methods for which we have proposed improvements (C**??**), support vector machine analysis for large datasets (J13), community detection methods on large graphs (J**??**), and more importantly have allowed for a first random matrix analysis of the performance of large random neural networks (J12) which is a promising avenue of research in neural networks. These works have already been set to practice with the development of innovative ideas in the scope of industrial projects (I2).
The medium- to long-term objective of this line of work is to develop improved machine learning methods better adapted to large dimensional datasets and to develop a holistic random matrix framework for big data machine learning.

A detailed description of all articles published since 2007 is listed below.

Publications in Journals

J1. C. Louart, R. Couillet "A Concentration of Measure Approach to Large Dimensional Robust Statistics", (submitted to) Annals of Applied Probability, 2020.

Abstract. This article studies the robust covariance matrix estimation of a data collection $X = (x_1, ..., x_n)$ with $x_i = \tau_i z_i + m$, where $z_i \in \mathbb{R}^p$ is a concentrated vector (e.g., an elliptical random vector), $m \in \mathbb{R}^p$ a deterministic signal and $\tau_i \in \mathbb{R}$ a scalar perturbation of possibly large amplitude, under the assumption where both n and p are large. This estimator is defined as the fixed point of a function which we show is contracting for a so-called stable semi-metric. We exploit this semimetric along with concentration of measure arguments to prove the existence and uniqueness of the robust estimator as well as evaluate its limiting spectral distribution.

J2. K. Elkhalil, A. Kammoun, R. Couillet, T. Al-Naffouri, M-S. Alouini, "A Large Dimensional Analysis of Regularized Discriminant Analysis Classifiers" (to appear) IEEE Transactions on Signal Processing, 2020.

Abstract. Abstract—In this paper, we conduct a large dimensional study of regularized discriminant analysis classifiers with its two popular variants known as regularized LDA and regularized QDA. The analysis is based on the assumption that the data samples are drawn from a Gaussian mixture model with different means and covariances and relies on tools from random matrix theory (RMT). We consider the regime in which both the data dimension and training size within each class tends to infinity with fixed ratio. Under mild assumptions, we show that the probability of misclassification converges to a deterministic quantity that describes in closed form the performance of these classifiers in terms of the class statistics as well as the problem dimension. The result allows for a better understanding of the underlying classification algorithms in terms of their performances in practical large but finite dimensions. Further exploitation of the results permits to optimally tune the regularization parameter with the aim of minimizing the probability of misclassification. The analysis is validated with numerical results involving synthetic as well as real data from the USPS dataset yielding a high accuracy in predicting the performances and hence making an interesting connection between theory and practice.

J3. L. Dall'Amico, R. Couillet, N. Tremblay, "A unified framework for spectral clustering in sparse graphs", (submitted to) Journal of Machine Learning Research, 2020.

Abstract. This article considers spectral community detection in the regime of sparse networks with heterogeneous degree distributions, for which we devise an algorithm to efficiently retrieve communities. Specifically, we demonstrate that a conveniently parametrized form of regularized Laplacian matrix can be used to perform spectral clustering in sparse networks, without suffering from its degree heterogeneity. Besides, we exhibit important connections between this proposed matrix and the now popular non-backtracking matrix, the Bethe-Hessian matrix, as well as the standard Laplacian matrix. Interestingly, as opposed to competitive methods, our proposed improved parametrization inherently accounts for the hardness of the classification problem. These findings are summarized under the form of an algorithm capable of both estimating the number of communities and achieving high-quality community reconstruction.

J4. C. Louart, R. Couillet, "Concentration of Measure and Large Random Matrices with an application to Sample Covariance Matrices", (submitted to) Random Matrix Theory and Applications, 2019.

Abstract. The present work provides an original framework for random matrix analysis based on revisiting the concentration of measure theory from a probabilistic point of view. By providing various notions of vector concentration (q-exponential,

linear, Lipschitz, convex), a set of elementary tools is laid out that allows for the immediate extension of classical results from random matrix theory involving random concentrated vectors in place of vectors with independent entries. These findings are exemplified here in the context of sample covariance matrices but find a large range of applications in statistical learning and beyond, thanks to the broad adaptability of our hypotheses.

J5. X. Mai, R. Couillet, "Consistent Semi-Supervised Graph Regularization for High Dimensional Data", (submitted) Journal of Machine Learning Research, 2019.

Abstract. Semi-supervised Laplacian regularization, a standard graph-based approach for learning from both labelled and unlabelled data, is demonstrated by the recent work of (Mai and Couillet, 2017) to have an insignificant high dimensional learning efficiency with respect to unlabelled data, causing it to be outperformed by its unsupervised counterpart, spectral clustering, given sufficient unlabelled data. Following a detailed discussion on the origin of this inconsistency problem, a novel regularization approach is proposed as solution, which is shown both theoretically and empirically to have a superior performance over Laplacian regularization.

J6. R. Couillet, M. Tiomoko, S. Zozor, E. Moisan, "Random matrix-improved estimation of covariance matrix distances", Journal of Multivariate Analysis, vol. 174, pp. 104531, 2019.

Abstract. Given two sets $x_1^{(1)}, \ldots, x_{n_1}^{(1)}$ and $x_1^{(2)}, \ldots, x_{n_2}^{(2)} \in \mathbb{R}^p$ (or $\mathbb{C}^p$) of random vectors with zero mean and positive definite covariance matrices C_1 and $C_2 \in \mathbb{R}^{p\times p}$ (or $\mathbb{C}^{p\times p}$), respectively, this article provides novel estimators for a wide range of distances between C_1 and C_2 (along with divergences between some zero mean and covariance C_1 or C_2 probability measures) of the form $\frac{1}{p}\sum_{i=1}^n f(\lambda_i(C_1^{-1}C_2))$ (with $\lambda_i(X)$ the eigenvalues of matrix X). These estimators are derived using recent advances in the field of random matrix theory and are asymptotically consistent as $n_1, n_2, p \to \infty$ with non trivial ratios $p/n_1 < 1$ and $p/n_2 < 1$ (the case $p/n_2 > 1$ is also discussed). A first "generic" estimator, valid for a large set of f functions, is provided under the form of a complex integral. Then, for a selected set of f's of practical interest (namely, $f(t) = t$, $f(t) = \log(t)$, $f(t) = \log(1 + st)$ and $f(t) = \log^2(t)$), a closed-form expression is provided. Beside theoretical findings, simulation results suggest an outstanding performance advantage for the proposed estimators when compared to the classical "plug-in" estimator $\frac{1}{p}\sum_{i=1}^n f(\lambda_i(\hat{C}_1^{-1}\hat{C}_2))$ (with $\hat{C}_a = \frac{1}{n_a}\sum_{i=1}^{n_a} x_i^{(a)} x_i^{(a)\mathsf{T}}$), and this even for very small values of n_1, n_2, p.

J7. X. Mai, R. Couillet, "A Random Matrix Analysis and Improvement of Semi-Supervised Learning for Large Dimensional Data", Journal of Machine Learning Research, vol. 19, no. 79, pp. 1-27, 2018.

Abstract. This article provides an original understanding of the behavior of a class of graph-oriented semi-supervised learning algorithms in the limit of large and numerous data. It is demonstrated that the intuition at the root of these methods collapses in this limit and that, as a result, most of them become inconsistent. Corrective measures and a new data-driven parametrization scheme are proposed along with a theoretical analysis of the asymptotic performances of the resulting approach. A surprisingly close behavior between theoretical performances on Gaussian mixture models and on real datasets is also illustrated throughout the article, thereby suggesting the importance of the proposed analysis for dealing with practical data. As a result, significant performance gains are observed on practical data classification using the proposed parametrization.

J8. A. Kammoun, R. Couillet, "Subspace Kernel Clustering of Large Dimensional Data" (submitted to) Annals of Applied Probability, 2017.

Abstract. Let $x_1, \ldots, x_n$ be independent observations of size p, each of them belonging to one of c distinct classes. We assume that observations within class a are characterized by their distribution $\mathcal{N}(0, \frac{1}{p}C_a)$ where here $C_1, \ldots, C_c$ are some non-negative definite $p \times p$ matrices. This paper studies the asymptotic behavior of the symmetric matrix $\tilde{\Phi}_{kl} = \sqrt{p}((x_k^T x_l)^2 \delta_{k \neq l})$ when p and n grow to infinity with $n/p \to c_0$. Particularly, we prove that, if the class covariance matrices are sufficiently close in a certain sense, the matrix $\tilde{\Phi}$ behaves as a low-rank perturbation of a Wigner matrix, presenting possibly some isolated eigenvalues outside the bulk of the semi-circular law. We carry out a careful analysis of some of the isolated eigenvalues and eigenvectors of matrix $\tilde{\Phi}$, and illustrate how these results can help understand spectral clustering methods that use $\tilde{\Phi}$ as a kernel matrix.

J9. L. Yang, M. McKay, R. Couillet, "High-Dimensional MVDR Beamforming: Optimized Solutions based on Spiked Random Matrix Models", IEEE Transactions on Signal Processing, vol. 66, no. 1, pp. 1933-1947, 2018.

Abstract. Minimum variance distortionless response (MVDR) beamforming (or Capon beamforming) is among the most popular adaptive array processing strategies due to its ability to provide noise resilience while nulling out interferers. A practical challenge with this beamformer is that it involves the inverse covariance matrix of the received signals, which must be estimated from data. Under modern high-dimensional applications, it is well-known that classical estimators can be severely affected by sampling noise, which compromises beamformer performance. Here we propose a new approach to MVDR beamforming which is suited to high-dimensional settings. In particular, by drawing an analogy with the MVDR problem and the so-called "spiked models" in random matrix theory, we propose robust beamforming solutions which are shown to outperform classical approaches (e.g., matched filters and sample matrix inversion techniques), as well as more robust solutions, such as methods based on diagonal loading. The key to our method is the design of an optimized inverse covariance estimator which applies eigenvalue clipping and shrinkage functions that are tailored to the MVDR application. Our proposed MVDR solution is simple, in closed form, and easy to implement.

J10. A. Karadimitrakis, A. L. Moustakas, R. Couillet, "Gallager Bound for MIMO Channels: Large-N Asymptotics" IEEE Transactions on Wireless Communications, vol. 17, no. 2, pp. 1323-1330, 2018.

Abstract. The use of multiple antenna arrays in transmission and reception has become an integral part of modern wireless communications. To quantify the performance of such systems, the evaluation of bounds on the error probability of realistic finite length codewords is important. In this paper, we analyze the standard Gallager error bound for both constraints of maximum average power and maximum instantaneous power. Applying techniques from random matrix theory, we obtain analytic expressions of the error exponent when the length of the codeword increases to infinity at a fixed ratio with the antenna array dimensions. Analyzing its behavior at rates close to the ergodic rate, we find that the Gallager error bound becomes asymptotically close to an upper error bound obtained recently by Hoydis et al. 2015. We also obtain an expression for the Gallager exponent in the case when the codelength spans several Rayleigh fading blocks, hence taking into account the situation when the channel varies during each transmission.

J11. N. Auguin, D. Morales, M. McKay, R. Couillet, "Large-dimensional behavior of regularized Maronna's M-estimators of covariance matrices" IEEE Transactions on Signal Processing, vol. 66, no. 13, pp. 3529–3542, 2018.

Abstract. Robust estimators of large covariance matrices are considered, comprising regularized (linear shrinkage) modifications of Maronna's classical M-estimators.

These estimators provide robustness to outliers, while simultaneously being well-defined when the number of samples does not exceed the number of variables. By applying tools from random matrix theory, we characterize the asymptotic performance of such estimators when the numbers of samples and variables grow large together. In particular, our results show that, when outliers are absent, many estimators of the regularized-Maronna type share the same asymptotic performance, and for these estimators we present a data-driven method for choosing the asymptotically optimal regularization parameter with respect to a quadratic loss. Robustness in the presence of outliers is then studied : in the non-regularized case, a large-dimensional robustness metric is proposed, and explicitly computed for two particular types of estimators, exhibiting interesting differences depending on the underlying contamination model. The impact of outliers in regularized estimators is then studied, with remarkable differences with respect to the non-regularized case, leading to new practical insights on the choice of particular estimators.

J12. C. Louart, Z. Liao, R. Couillet, "A Random Matrix Approach to Neural Networks" Annals of Applied Probability, vol. 28, no. 2, pp. 1190–1248, 2018.

Abstract. This article studies the Gram random matrix model $G = \frac{1}{T}\Sigma^{\mathrm{T}}\Sigma$, $\Sigma = \sigma(WX)$, classically found in the analysis of random feature maps and random neural networks, where $X = [x_1, \ldots, x_T] \in \mathbb{R}^{p\times T}$ is a (data) matrix of bounded norm, $W \in \mathbb{R}^{n\times p}$ is a matrix of independent zero-mean unit variance entries, and $\sigma : \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous (activation) function — $\sigma(WX)$ being understood entry-wise. We prove that, as n, p, T grow large at the same rate, the resolvent $Q = (G + \gamma I_T)^{-1}$, for $\gamma > 0$, has a similar behavior as that met in sample covariance matrix models, involving notably the moment $\Phi = \frac{T}{n}\mathrm{E}[G]$, which provides in passing a deterministic equivalent for the empirical spectral measure of G. This result, established by means of concentration of measure arguments, enables the estimation of the asymptotic performance of single-layer random neural networks. This in turn provides practical insights into the underlying mechanisms into play in random neural networks, entailing several unexpected consequences, as well as a fast practical means to tune the network hyperparameters.

J13. Z. Liao, R. Couillet, "A Large Dimensional Analysis of Least Square Support Vector Machines" IEEE Transactions on Signal Processing, vol. 67, no. 4, pp. 1065-1074, 2018.

Abstract. In this article, a large dimensional performance analysis of kernel least squares support vector machines (LS-SVMs) is provided under the assumption of a two-class Gaussian mixture model for the input data. Building upon recent random matrix advances, when both the dimension of data p and their number n grow large at the same rate, we show that the LS-SVM decision function converges to a normal-distributed variable, the mean and variance of which depend explicitly on a local behavior of the kernel function. This theoretical result is then applied to real data sets which, despite their non-Gaussianity, exhibit a surprisingly similar behavior. Our analysis provides a deeper understanding of the mechanism into play in SVM-type methods and in particular of the impact on the choice of the kernel function as well as some of their theoretical limits.

J14. R. Couillet, H. Tiomoko Ali, "Improved spectral community detection in large heterogeneous networks" Journal of Machine Learning Resarch, vol. 18, no. 225, pp. 1–49, 2018.

Abstract. In this article, we study spectral methods for community detection based on α-parametrized normalized modularity matrix hereafter called L_α in heterogeneous graph models. We show, in a regime where community detection is not asymptotically trivial, that L_α can be well approximated by a more tractable random matrix which falls in the family of spiked random matrices. The analysis of

this equivalent spiked random matrix allows us to improve spectral methods for community detection and assess their performances in the regime under study. In particular, we prove the existence of an optimal value α_{opt} of the parameter α for which the detection of communities is best ensured and we provide an on-line estimation of α_{opt} only based on the knowledge of the graph adjacency matrix. Unlike classical spectral methods for community detection where clustering is performed on the eigenvectors associated with extreme eigenvalues, we show through our theoretical analysis that a regularization should instead be performed on those eigenvectors prior to clustering in heterogeneous graphs. Finally, through a deeper study of the regularized eigenvectors used for clustering, we assess the performances of our new algorithm for community detection. Numerical simulations in the course of the article show that our methods outperform state-of-the-art spectral methods on dense heterogeneous graphs.

J15. R. Couillet, M. McKay, "Optimal block-sparse PCA for high dimensional correlated samples" (submitted to) Journal of Multivariate Analysis, 2016.

Abstract. A new principal component analysis (PCA) method is proposed which is performed on a subset of blocks of consecutive entries of the population data vectors. This block-based dimensionality reduction introduces a trade-off by which the accuracy of the dominant eigenvector of the dimension-reduced sample covariance matrix is enhanced while some population entries are discarded. This scheme is particularly suited (but not restricted) to population eigenvectors with localized energy and rather sparse structures. Unlike many sparse PCA algorithms, the originality of our scheme lies in its providing an online selection of the subset of blocks which, in the large dimensional regime where both population and sample sizes grow large, provably ensures optimal alignment between population and sample eigenvectors. Moreover, our method inherently handles (a priori unknown) linear correlation between sample data.

J16. R. Couillet, G. Wainrib, H. Sevi, H. Tiomoko Ali, "The asymptotic performance of linear echo state neural networks" Journal of Machine Learning Research, vol. 17, no. 178, pp. 1-35, 2016.

Abstract. In this article, a study of the mean-square error (MSE) performance of linear echo-state neural networks is performed, both for training and testing tasks. Considering the realistic setting of noise present at the network nodes, we derive deterministic equivalents for the aforementioned MSE in the limit where the number of input data T and network size n both grow large. Specializing then the network connectivity matrix to specific random settings, we further obtain simple formulas that provide new insights on the performance of such networks.

J17. R. Couillet, F. Benaych-Georges, "Kernel Spectral Clustering of Large Dimensional Data" Electronic Journal of Statistics, vol. 10, no. 1, pp. 1393-1454, 2016.

Abstract. This article proposes a first analysis of kernel spectral clustering methods in the regime where the dimension p of the data vectors to be clustered and their number n grow large at the same rate. We demonstrate, under a k-class Gaussian mixture model, that the normalized Laplacian matrix associated with the kernel matrix asymptotically behaves similar to a so-called spiked random matrix. Some of the isolated eigenvalue-eigenvector pairs in this model are shown to carry the clustering information upon a separability condition classical in spiked matrix models. We evaluate precisely the position of these eigenvalues and the content of the eigenvectors, which unveil important properties concerning spectral clustering, in particular in simple toy models. Our results are then compared to the practical clustering of images from the MNIST database, thereby revealing an important match between theory and practice.

J18. F. Benaych-Georges, R. Couillet, "Spectral Analysis of the Gram Matrix of Mixture Models" ESAIM : Probability and Statistics, DOI http ://dx.doi.org/10.1051/ps/2016007, 2016.

Abstract. This text is devoted to the asymptotic study of some spectral properties of the Gram matrix $W^\mathsf{T}W$ built upon a collection $w_1, \ldots, w_n \in \mathbb{R}^p$ of random vectors (the columns of W), as both the number n of observations and the dimension p of the observations tend to infinity and are of similar order of magnitude. The random vectors $w_1, \ldots, w_n$ are independent observations, each of them belonging to one of k classes $\mathcal{C}_1, \ldots, \mathcal{C}_k$. The observations of each class $\mathcal{C}_a$ $(1 \leq a \leq k)$ are characterized by their distribution $\mathcal{N}(0, p^{-1}C_a)$, where $C_1, \ldots, C_k$ are some non negative definite $p \times p$ matrices. The cardinality n_a of class $\mathcal{C}_a$ and the dimension p of the observations are such that n_a/n $(1 \leq a \leq k)$ and p/n stay bounded away from 0 and $+\infty$. We provide deterministic equivalents to the empirical spectral distribution of $W^\mathsf{T}W$ and to the matrix entries of its resolvent (as well as of the resolvent of WW^T). These deterministic equivalents are defined thanks to the solutions of a fixed-point system. Besides, we prove that $W^\mathsf{T}W$ has asymptotically no eigenvalues outside the bulk of its spectrum, defined thanks to these deterministic equivalents. These results are directly used in the companion paper (17), which is devoted to the analysis of the spectral clustering algorithm in large dimensions. They also find applications in various other fields such as wireless communications where functionals of the aforementioned resolvents allow one to assess the communication performance across multi-user multi-antenna channels.

J19. R. Couillet, Estimation robuste et matrices aléatoires, revue Traitement du Signal, vol. 33, no. 2-3, pp. 273-320, 2016.

Abstract. This article provides a technical survey of the recent advances between the fields of robust estimation of scatter and of large dimensional random matrix theory. An exposition of the theoretical results will be made which we shall apply to various contexts in the area of statistics and signal processing at large. The theoretical results essentially show that, while robust estimators of scatter are implicitly defined and thus difficult objects to manipulate, in the large dimensional random matrix regime where both the population size and the number of samples are simultaneously large, these implicit robust estimators tend to behave similar to much simpler random matrix models, amenable to analysis. This induces that many statistical properties of these estimators could be unearthed which we shall discuss. In terms of applications, these robust estimators of scatter are long-standing structural elements to handle both outliers and heavy-tailed behavior in the observed data. These impulsiveness harnessing effects will be precisely documented and shall be instrumental to develop improved robust statistics methods for detection and estimation in antenna arrays, portfolio optimization, etc.

J20. R. Couillet, G. Wainrib, Perspectives en matrices aléatoires et grands réseaux, revue Traitement du Signal, vol. 33, no. 2-3, pp. 351-376, 2016.

Abstract. In this article, several research perspectives in random matrix theory applied to graph theory at large are discussed. Specific focus will be made on the spectrum analysis of the adjacency or Laplacian matrices of large dimensional graphs for community detection in networks, of kernel random matrices for clustering in large datasets, along with applications to neural networks.

J21. M. Sadeghi, L. Sanguinetti, R. Couillet, Y. Chau, "Large System Analysis of Power Normalization Techniques in Massive MIMO", IEEE Transactions on Vehicular Technologies, vol. 66, no. 10, pp. 9005-9017, 2017.

Abstract. Linear precoding has been widely studied in the context of Massive MIMO together with the two common power normalization techniques, namely, matrix normalization (MN) and vector normalization (VN). However, the effect of both on

the system performance has not been thoroughly studied. The aim of this paper is to address this problem using large system analysis. Considering a system model that accounts for channel estimation, pilot contamination, arbitrary pathloss, and per-user channel correlation, we compute tight approximations for the signal-to-interference-plus-noise ratio (SINR) and the rate of each user equipment (UE) in the system while employing maximum ratio transmission (MRT), zero forcing (ZF), and regularized ZF (RZF) precoding under both MN and VN techniques. Exploiting such results, we reveal the effect of power normalization on the performance of MRT and ZF, and determine how it affects noise, interference, pilot contamination, and signal powers of any arbitrary UE. We show that the power normalization can convey a notion of fairness or sum rate maximization for ZF. Numerical results are used to validate the accuracy of the asymptotic analysis and to show that in Massive MIMO, non- coherent interference and noise, rather than pilot contamination, are often the major limiting factors of the considered precoding schemes. .

J22. M. Sadeghi, L. Sanguinetti, R. Couillet, Y. Chau, "Reducing the Computational Complexity of Multicasting in Large-Scale Antenna Systems", IEEE Transactions on Wireless Communications, vol. 16, no. 5, pp. 2963-2975, 2017.

Abstract. In this paper, we study the physical layer multicasting to multiple co-channel groups in large-scale antenna systems. The users within each group are interested in a common message and different groups have distinct messages. In particular, we aim at designing the precoding vectors solving the so-called quality of service (QoS) and weighted max-min fairness (MMF) problems, assuming that the channel state information is available at the base station (BS). To solve both problems, the baseline approach exploits the semidefinite relaxation (SDR) technique. Considering a BS with N antennas, the SDR complexity is more than $O(N^6)$, which prevents its application in large-scale antenna systems. To overcome this issue, we present two new classes of algorithms that, not only have significantly lower computational complexity than existing solutions, but also largely outperform the SDR based methods. Moreover, we present a novel duality between transformed versions of the QoS and the weighted MMF problems. The duality explicitly determines the solution to the weighted MMF problem given the solution to the QoS problem, and vice versa. Numerical results are used to validate the effectiveness of the proposed solutions and to make comparisons with existing alternatives under different operating conditions.

J23. L. Sanguinetti, R. Couillet, M. Debbah, "Large System Analysis of Base Station Cooperation for Power Minimization" IEEE Transactions on Wireless Communications, vol. 15, no. 8, pp. 5480-5496, 2016.

Abstract. Abstract This work focuses on a large-scale multi-cell multi-user MIMO system in which L base stations (BSs) of N antennas each communicate with K single-antenna user equipments. We consider the design of the linear precoder that minimizes the total power consumption while ensuring target user rates. Three configurations with different degrees of cooperation among BSs are considered : the single cell processing scheme (no cooperation between BSs), the coordinated beamforming scheme (only channel state information is shared between BSs) and the coordinated multipoint MIMO processing technology (channel state and data cooperation). The analysis is conducted assuming that N and K grow large with a non trivial ratio K/N and imperfect channel state information is available at the BSs. Tools of random matrix theory are used to compute, in explicit form, deterministic approximations for : (i) the parameters of the optimal precoder ; (ii) the powers needed to ensure target rates ; and (iii) the total transmit power. These results are instrumental to get further insight into the structure of the optimal precoders and also to reduce the complexity of its implementation in large-scale

networks. Numerical results are used to validate the asymptotic analysis in the finite system regime and to make comparisons among the different configurations.

J24. A. Abboud, F. Iutzeler, R. Couillet, H. Siguerdidjane, M. Debbah, "Distributed Production-Sharing Optimization and Application to Power Grid Networks," IEEE Transactions on Signal and Information Processing over Networks, vol. 2, no. 1, pp. 1628, 2016.

Abstract. Based on recent works on asynchronous versions of the distributed Alternating Direction Method of Multipliers (ADMM) algorithm, we develop and prove the convergence of a distributed asynchronous method for Production-Sharing Problems over networks. The asynchronous nature of the algorithm allows both for the relaxation of the synchronization constraint often inherent to distributed ADMM-based methods and distributed optimization methods at large, but also allows for random local failures to occur in fully centralized methods. These two considerations motivate the application of the method to the Direct-Current Optimal Power Flow (DC-OPF) problem in power transmission networks. Applied to the DC-OPF, this method leads to an overall network optimal production obtained through a sequence of local computations in subareas of the network (each area waking up randomly while the rest of the network is non-operational) and neighboring data exchanges. In another scenario, the DC-OPF is performed via iterations of a centralized network-wide ADMM method which may contain disconnected nodes (in general with low probability and for a short duration). In both cases, this method still converges and thus provides additional flexibility to classical DC-OPF algorithms. The proposed algorithm, inherently designed for networks of overlapping subareas, is then extended to networks of non-overlapping areas. Simulations are carried out on the IEEE-30 and IEEE-118 bus test systems which illustrate the convergence, scalability and effectiveness of the proposed algorithms.

J25. A. Kammoun, R. Couillet, F. Pascal, M.-S. Alouini, "Optimal Design of the Adaptive Normalized Matched Filter Detector using Regularized Tyler Estimator" IEEE Transactions on Aerospace and Electronic Systems, vol. 54, no. 2, pp. 755–769, 2018.

Abstract. This article addresses improvements on the design of the adaptive normalized matched filter (ANMF) for radar detection. It is well-acknowledged that the estimation of the noise-clutter covariance matrix is a fundamental step in adaptive radar detection. In this paper, we consider regularized estimation methods which force by construction the eigenvalues of the scatter estimates to be greater than a positive regularization parameter ρ. This makes them more suitable for high dimensional problems with a limited number of secondary data samples than traditional sample covariance estimates. While an increase of ρ seems to improve the conditioning of the estimate, it might however cause it to significantly deviate from the true covariance matrix. The setting of the optimal regularization parameter is a difficult question for which no convincing answers have thus far been provided. This constitutes the major motivation behind our work. More specifically, we consider the design of the ANMF detector for two kinds of regularized estimators, namely the regularized sample covariance matrix (RSCM), appropriate when the clutter follows a Gaussian distribution and the regularized Tyler estimator (RTE) for non-Gaussian spherically invariant distributed clutters. The rationale behind this choice is that the RTE is efficient in mitigating the degradation caused by the presence of impulsive noises while inducing little loss when the noise is Gaussian. Based on recent random matrix theory results studying the asymptotic fluctuations of the statistics of the ANMF detector when the number of samples and their dimension grow together to infinity, we propose a design for the regularization parameter that maximizes the detection probability under constant false alarm rates. Simulation results which support the efficiency of the proposed method are provided in order

to illustrate the gain of the proposed optimal design over conventional settings of the regularization parameter.

J26. A. Kammoun, R. Couillet, F. Pascal, M.-S. Alouini, "Convergence and Fluctuations of Regularized Tyler Estimators" IEEE Transactions on Signal Processing, vol. 64, no. 4, pp. 1048-1060, 2016.

Abstract. This article studies the behavior of regularized Tyler estimators (RTEs) of scatter matrices. The key advantages of these estimators are twofold. First, they guarantee by construction a good conditioning of the estimate and second, being a derivative of robust Tyler estimators, they inherit their robustness properties, notably their resilience to the presence of outliers. Nevertheless, one major problem that poses the use of RTEs in practice is represented by the question of setting the regularization parameter ρ. While a high value of ρ is likely to push all the eigenvalues away from zero, it comes at the cost of a larger bias with respect to the population covariance matrix. A deep understanding of the statistics of RTEs is essential to come up with appropriate choices for the regularization parameter. This is not an easy task and might be out of reach, unless one considers asymptotic regimes wherein the number of observations n and/or their size N increase together. First asymptotic results have recently been obtained under the assumption that N and n are large and commensurable. Interestingly, no results concerning the regime of n going to infinity with N fixed exist, even though the investigation of this assumption has usually predated the analysis of the most difficult N and n large case. This motivates our work. In particular, we prove in the present paper that the RTEs converge to a deterministic matrix when $n \to \infty$ with N fixed, which is expressed as a function of the theoretical covariance matrix. We also derive the fluctuations of the RTEs around this deterministic matrix and establish that these fluctuations converge in distribution to a multivariate Gaussian distribution with zero mean and a covariance depending on the population covariance and the parameter ρ.

J27. D. Morales-Jimenez, R. Couillet, M. McKay, "Large Dimensional Analysis of Robust M-Estimators of Covariance with Outliers" IEEE Transactions on Signal Processing, vol. 63, no. 21, pp. 5784-5797, 2015.

Abstract. A large dimensional characterization of robust M-estimators of covariance (or scatter) is provided under the assumption that the dataset comprises independent (essentially Gaussian) legitimate samples as well as arbitrary deterministic samples, referred to as outliers. Building upon recent random matrix advances in the area of robust statistics, we specifically show that the so-called Maronna M-estimator of scatter asymptotically behaves similar to well-known random matrices when the population and sample sizes grow together to infinity. The introduction of outliers leads the robust estimator to behave asymptotically as the weighted sum of the sample outer products, with a constant weight for all legitimate samples and different weights for the outliers. A fine analysis of this structure reveals importantly that the propensity of the M-estimator to attenuate (or enhance) the impact of outliers is mostly dictated by the alignment of the outliers with the inverse population covariance matrix of the legitimate samples. Thus, robust M-estimators can bring substantial benefits over more simplistic estimators such as the per-sample normalized version of the sample covariance matrix, which is not capable of differentiating the outlying samples. The analysis shows that, within the class of Maronna's estimators of scatter, the Huber estimator is most favorable for rejecting outliers. On the contrary, estimators more similar to Tyler's scale invariant estimator (often preferred in the literature) run the risk of inadvertently enhancing some outliers.

J28. L. Yang, R. Couillet, M. McKay, "A Robust Statistics Approach to Minimum Variance Portfolio Optimization" IEEE Transactions on Signal Processing, vol. 63,

no. 24, pp. 6684–6697, 2015.

Abstract. We study the design of portfolios under a minimum risk criterion. The performance of the optimized portfolio relies on the accuracy of the estimated covariance matrix of the portfolio asset returns. For large portfolios, the number of available market returns is often of similar order to the number of assets, so that the sample covariance matrix performs poorly as a covariance estimator. Additionally, financial market data often contain outliers which, if not correctly handled, may further corrupt the covariance estimation. We address these shortcomings by studying the performance of a hybrid covariance matrix estimator based on Tyler's robust M-estimator and on Ledoit-Wolf's shrinkage estimator while assuming samples with heavy-tailed distribution. Employing recent results from random matrix theory, we develop a consistent estimator of (a scaled version of) the realized portfolio risk, which is minimized by optimizing online the shrinkage intensity. Our portfolio optimization method is shown via simulations to outperform existing methods both for synthetic and real market data.

J29. R. Couillet, A. Kammoun, F. Pascal, "Second order statistics of robust estimators of scatter. Application to GLRT detection for elliptical signals" Elsevier Journal of Multivariate Analysis, vol. 143, pp. 249-274, 2015.

Abstract. A central limit theorem for bilinear forms of the type $a^* \hat{C}_N(\rho)^{-1} b$, where $a, b \in \mathbb{C}^N$ are unit norm deterministic vectors and $\hat{C}_N(\rho)$ a robust-shrinkage estimator of scatter parametrized by ρ and built upon n independent elliptical vector observations, is presented. The fluctuations of $a^* \hat{C}_N(\rho)^{-1} b$ are found to be of order $N^{-\frac{1}{2}}$ and to be the same as those of $a^* \hat{S}_N(\rho)^{-1} b$ for $\hat{S}_N(\rho)$ a matrix of a theoretical tractable form. This result is exploited in a classical signal detection problem to provide an improved detector which is both robust to elliptical data observations (e.g., impulsive noise) and optimized across the shrinkage parameter ρ.

J30. A. Müller, R. Couillet, E. Bjørnson, S. Wagner, M. Debbah, "Interference-Aware RZF Precoding for Multi-Cell Downlink Systems" IEEE Transactions on Signal Processing, vol. 63, no. 15, pp. 3959-3973 2015.

Abstract. Recently, the structure of the optimal linear precoder for multi-cell downlink systems has been described. Other references have used simplified versions of the precoder to obtain promising performance gains. These gains have been hypothesized to stem from providing additional degrees of freedom that allow for interference mitigation through interference relegation to orthogonal subspaces. However, no conclusive or rigorous understanding has yet been proposed. In this paper, we take an interference-aware adaption of the generally optimal precoding structure and analyze the rate performance in multi-cell scenarios. A special emphasis is placed on induced interference mitigation. For example, we will verify the intuitive expectation that the precoder structure can either completely remove induced inter-cell or intra-cell interference. We state new results from large-scale random matrix theory, that make it possible to give more intuitive and insightful explanations of the precoder behavior, also for cases involving imperfect channel state information (CSI). We remark especially that the interference-aware precoder makes use of all available information about interfering channels to improve performance. Even extremely bad CSI can be used to enhance the sum rate. Our obtained insights are then used to propose heuristic precoder parameters for arbitrary systems, whose effectiveness is shown in more involved system scenarios. Furthermore, determining these parameters does not require explicit inter base station cooperation. Using a simple heuristic version of the interference aware precoder, one finds that a sum rate performance, close to the optimally parameterized precoder one, can be achieved.

J31. R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator" Elsevier Journal of Multivariate Analysis, vol. 140, pp. 139-161, 2015.

Abstract. A class of robust estimators of scatter applied to information-plus-impulsive noise samples is studied, where the sample information matrix is assumed of low rank ; this generalizes the study (34) (restricted to a noise only setting) to spiked random matrix models. It is precisely shown that, as opposed to sample covariance matrices which may have asymptotically unbounded (eigen-)spectrum due to the sample impulsiveness, the robust estimator of scatter has bounded spectrum and may contain isolated eigenvalues which we fully characterize. We show that, if found beyond a certain detectability threshold, these eigenvalues allow one to perform statistical inference on the eigenvalues and eigenvectors of the information matrix. We use this result to derive new eigenvalue and eigenvector estimation procedures, which we apply in practice to the popular array processing problem of angle of arrival estimation. This gives birth to an improved algorithm based on the MUSIC method, which we refer to as robust G-MUSIC.

J32. R. Couillet, M. McKay, "Large Dimensional Analysis and Optimization of Robust Shrinkage Covariance Matrix Estimators" Elsevier Journal of Multivariate Analysis, vol. 131, pp. 99-120, 2014.

Abstract. This article studies two regularized robust estimators of scatter matrices proposed (and proved to be well defined) in parallel in (Chen et al., 2011) and (Pascal et al., 2013), based on Tyler's robust M-estimator (Tyler, 1987) and on Ledoit and Wolf's shrinkage covariance matrix estimator (Ledoit and Wolf, 2004). These hybrid estimators have the advantage of conveying (i) robustness to outliers or impulsive samples and (ii) small sample size adequacy to the classical sample covariance matrix estimator. We consider here the case of i.i.d. elliptical zero mean samples in the regime where both sample and population sizes are large. We demonstrate that, under this setting, the estimators under study asymptotically behave similar to well-understood random matrix models. This characterization allows us to derive optimal shrinkage strategies to estimate the population scatter matrix, improving significantly upon the empirical shrinkage method proposed in (Chen et al., 2011).

J33. Y. Chitour, R. Couillet, F. Pascal "On the convergence of Maronna's M-estimators of scatter" IEEE Signal Processing Letters, vol. 22, no. 6, pp. 709-712, 2014.

Abstract. In this paper, we propose an alternative proof for the uniqueness of Maronna's M-estimator of scatter for N vector observations $y_1, \ldots, y_N \in \mathbb{R}^m$ under a mild constraint of linear independence of any subset of m of these vectors. This entails in particular almost sure uniqueness for random vectors y_i with a density as long as $N > m$. This approach allows to establish further relations that demonstrate that a properly normalized Tyler's M-estimator of scatter can be considered as a limit of Maronna's M-estimator. More precisely, the contribution is to show that each M-estimator, verifying some mild conditions, converges towards a particular Tyler's M-estimator. These results find important implications in recent works on the large dimensional (random matrix) regime of robust M-estimation.

J34. R. Couillet, F. Pascal, J. W. Silverstein, "The Random Matrix Regime of Maronna's M-estimator with elliptically distributed samples", vol. 139, pp. 56-78, Elsevier Journal of Multivariate Analysis, 2015.

Abstract. This article demonstrates that the robust scatter matrix estimator $\hat{C}_N \in \mathbb{C}^{N\times N}$ of a multivariate elliptical population $x_1, \ldots, x_n \in \mathbb{C}^N$ originally proposed by Maronna in 1976, and defined as the solution (when existent) of an implicit equation, behaves similar to a well-known random matrix model in the limiting regime where the population N and sample n sizes grow at the same speed. We show precisely that $\hat{C}_N \in \mathbb{C}^{N\times N}$ is defined for all n large with probability one and that, under some light hypotheses, $\|\hat{C}_N - \hat{S}_N\| \to 0$ almost surely in spectral norm, where $\hat{S}_N$ follows a classical random matrix model. As a corollary, the limiting

eigenvalue distribution of $\hat{C}_N$ is derived. This analysis finds applications in the fields of statistical inference and signal processing.

J35. J. Vinogradova, R. Couillet, W. Hachem, "Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection large", IEEE Transactions on Signal Processing, vol. 63, no. 18, pp. 4903-4913, 2015.

Abstract. In this article, we derive concentration inequalities for the spectral norm of two classical sample estimators of large dimensional Toeplitz covariance matrices, demonstrating in particular their asymptotic almost sure consistence. The consistency is then extended to the case where the aggregated matrix of time samples is corrupted by a rank one (or more generally, low rank) matrix. As an application of the latter, the problem of source detection in the context of large dimensional sensor networks within a temporally correlated noise environment is studied. As opposed to standard procedures, this application is performed online, i.e., without the need to possess a learning set of pure noise samples.

J36. R. Couillet, W. Hachem, "Analysis of the limiting spectral measure of large random matrices of the separable covariance type", Random Matrix Theory and Applications, vol. 3, pp. 1-23, 2014.

Abstract. Consider the random matrix $\Sigma = D^{1/2} X \widetilde{D}^{1/2}$ where D and $\widetilde{D}$ are deterministic Hermitian nonnegative matrices with respective dimensions $N \times N$ and $n \times n$, and where X is a random matrix with independent and identically distributed centered elements with variance $1/n$. Assume that the dimensions N and n grow to infinity at the same pace, and that the spectral measures of D and $\widetilde{D}$ converge as $N, n \to \infty$ towards two probability measures. Then it is known that the spectral measure of $\Sigma\Sigma^*$ converges towards a probability measure μ characterized by its Stieltjes Transform. In this paper, it is shown that μ has a density away from zero, this density is analytical wherever it is positive, and it behaves in most cases as $\sqrt{|x-a|}$ near an edge a of its support. In addition, a complete characterization of the support of μ is provided. Aside from its mathematical interest, the analysis underlying these results finds important applications in a certain class of statistical estimation problems.

J37. J. Hoydis, R. Couillet, P. Piantanida, "The Second-Order Coding Rate of the MIMO Rayleigh Block-Fading Channel," IEEE Transactions on Information Theory, vol. 61, no. 12, pp. 6591-6622, 2015.

Abstract. The second-order coding rate of the multiple-input multiple-output (MIMO) quasi-static Rayleigh fading channel is studied. We tackle this problem via an information-spectrum approach and statistical bounds based on recent random matrix theory techniques. We precisely derive a central limit theorem (CLT) to analyze the information density in the regime where the block-length n and the number of transmit and receive antennas K and N, respectively, grow simultaneously large. This result leads to the characterization of closed-form upper and lower bounds on the optimal average error probability when the coding rate is within $O(1/\sqrt{nK})$ of the asymptotic capacity.

J38. J. Vinogradova, R. Couillet, W. Hachem, "Statistical Inference in Large Antenna Arrays under Unknown Noise Pattern," IEEE Transactions on Signal Processing, vol. 61, no. 22, pp. 5633-5645, 2013.

Abstract. In this article, a general information-plus-noise transmission model is assumed, the receiver end of which is composed of a large number of sensors and is unaware of the noise pattern. For this model, and under reasonable assumptions, a set of results is provided for the receiver to perform statistical eigen-inference on the information part. In particular, we introduce new methods for the detection, counting, and the power and subspace estimation of multiple sources composing

the information part of the transmission. The theoretical performance of some of these techniques is also discussed. An exemplary application of these methods to array processing is then studied in greater detail, leading in particular to a novel MUSIC-like algorithm assuming unknown noise covariance.

J39. F. Chapon, R. Couillet, W. Hachem, X. Mestre, "The outliers among the singular values of large rectangular random matrices with additive fixed rank deformation," Markov Processes and Related Fields, vol. 20, pp. 183-228, 2014.

Abstract. Consider the matrix $\Sigma_n = n^{-1/2} X_n D_n^{1/2} + P_n$ where the matrix $X_n \in \mathbb{C}^{N \times n}$ has Gaussian standard independent elements, D_n is a deterministic diagonal nonnegative matrix, and P_n is a deterministic matrix with fixed rank. Under some known conditions, the spectral measures of $\Sigma_n \Sigma_n^*$ and $n^{-1} X_n D_n X_n^*$ both converge towards a compactly supported probability measure μ as $N, n \to \infty$ with $N/n \to c$. In this paper, it is proved that finitely many eigenvalues of $\Sigma_n \Sigma_n^*$ may stay away from the support of μ in the large dimensional regime. The existence and locations of these outliers in any connected component of $\mathbb{R} \setminus \text{supp}(\mu)$ are studied. The fluctuations of the largest outliers of $\Sigma_n \Sigma_n^*$ are also analyzed. The results find applications in the fields of signal processing and radio communications.

J40. R. Couillet, F. Pascal, J. W. Silverstein, "Robust Estimates of Covariance Matrices in the Large Dimensional Regime," IEEE Transactions on Information Theory, vol. 60, no. 11, 2014.

Abstract. This article studies the limiting behavior of a class of robust population covariance matrix estimators, originally due to Maronna in 1976, in the regime where both the number of available samples and the population size grow large. Using tools from random matrix theory, we prove that, for sample vectors made of independent entries having some moment conditions, the difference between the sample covariance matrix and (a scaled version of) such robust estimator tends to zero in spectral norm, almost surely. This result can be applied to various statistical methods arising from random matrix theory that can be made robust without altering their first order behavior.

J41. G. Geraci, R. Couillet, J. Yuan, M. Debbah, I. B. Collings, "Large System Analysis of Linear Precoding in MISO Broadcast Channels with Confidential Messages," IEEE Journal on Selected Area in Communications, vol. 31, no. 9, pp. 1660-1671, 2013. **Second prize of the 2012-2013 IEEE Australia Council Student Paper Contest.**

Abstract. In this paper, we study the performance of regularized channel inversion (RCI) precoding in large MISO broadcast channels with confidential messages (BCC). We obtain a deterministic approximation for the achievable secrecy sum-rate which is almost surely exact as the number of transmit antennas M and the number of users K grow to infinity in a fixed ratio $\beta = K/M$. We derive the optimal regularization parameter ξ and the optimal network load β that maximize the per-antenna secrecy sum-rate. We then propose a linear precoder based on RCI and power reduction (RCI-PR) that significantly increases the high-SNR secrecy sum-rate for $1 < \beta < 2$. Our proposed precoder achieves a per-user secrecy rate which has the same high-SNR scaling factor as both the following upper bounds : (i) the rate of the optimum RCI precoder without secrecy requirements, and (ii) the secrecy capacity of a single-user system without interference. Furthermore, we obtain a deterministic approximation for the secrecy sum-rate achievable by RCI precoding in the presence of channel state information (CSI) error. We also analyze the performance of our proposed RCI-PR precoder with CSI error, and we determine how the error must scale with the SNR in order to maintain a given rate gap to the case with perfect CSI.

J42. J. Hoydis, R. Couillet, M. Debbah, "Iterative Deterministic Equivalents for the Capacity Analysis of Communication Systems," Technical Report.

Abstract. In this report, we introduce the notion of nested deterministic equivalents of functional of random matrices. Nested deterministic equivalents extend classical deterministic equivalents in order to study stochastic problems with multiple independent random variables. In particular, we discuss their applications to wireless communications, and especially to the capacity analysis of doubly-scattering multiple input multiple output (MIMO) channels and of multi-hop relay channels for which we derive novel expressions.

J43. R. Couillet, S. Medina Perlaza, H. Tembine, M. Debbah, "Electrical Vehicles in the Smart Grid: A Mean Field Game Analysis," IEEE Journal on Selected Areas in Communications : Smart Grid Communications Series, vol. 30, no. 6, pp. 1086-1096, 2012.

Abstract. In this article, we investigate the competitive interaction between electrical vehicles or hybrid oil-electricity vehicles in a Cournot market consisting of electricity transactions to or from an underlying electricity distribution network. We provide a mean field game formulation for this competition, and introduce the set of fundamental differential equations ruling the behavior of the vehicles at the feedback Nash equilibrium, referred here to as the mean field equilibrium. This framework allows for a consistent analysis of the evolution of the price of electricity as well as of the instantaneous electricity demand in the power grid. Simulations precisely quantify those parameters and suggest that significant reduction of the daily electricity peak demand can be achieved by appropriate electricity pricing.

J44. J. Yao, R. Couillet, J. Najim, M. Debbah, "Fluctuations of an Improved Population Eigenvalue Estimator in Sample Covariance Matrix Models," IEEE Transactions on Information Theory, vol. 59, no. 2, pp. 1149-1163, 2013.

Abstract. In this article, the joint fluctuations of the extreme eigenvalues and eigenvectors of a large dimensional sample covariance matrix are analyzed when the associated population covariance matrix is a finite-rank perturbation of the identity matrix, corresponding to the so-called spiked model in random matrix theory. The asymptotic fluctuations, as the matrix size grows large, are shown to be intimately linked with matrices from the Gaussian unitary ensemble (GUE). When the spiked population eigenvalues have unit multiplicity, the fluctuations follow a central limit theorem. This result is used to develop an original framework for the detection and diagnosis of local failures in large sensor networks, for known or unknown failure magnitude.

J45. R. Couillet, M. Debbah,"Signal Processing in Large Systems: a New Paradigm," IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 24-39, 2013.

Abstract. For a long time, detection and parameter estimation methods for signal processing have relied on asymptotic statistics as the number n of observations of a population grows large comparatively to the population size N, i.e., $n/N \to \infty$. Modern technological and societal advances now demand the study of sometimes extremely large populations and simultaneously require fast signal processing due to accelerated system dynamics. This results in not-so-large practical ratios n/N, sometimes even smaller than one. A disruptive change in classical signal processing methods has therefore been initiated in the past ten years, mostly spurred by the field of large dimensional random matrix theory. The early works in random matrix theory for signal processing applications are however scarce and highly technical. This tutorial provides an accessible methodological introduction to the modern tools of random matrix theory and to the signal processing methods derived from them, with an emphasis on simple illustrative examples.

J46. R. Couillet, W. Hachem, "Fluctuations of spiked random matrix models and failure diagnosis in sensor networks," IEEE Transactions on Information Theory, vol. 59, no. 1, pp. 509-525, 2013.

Abstract. In this article, the joint fluctuations of the extreme eigenvalues and eigenvectors of a large dimensional sample covariance matrix are analyzed when the associated population covariance matrix is a finite-rank perturbation of the identity matrix, corresponding to the so-called spiked model in random matrix theory. The asymptotic fluctuations, as the matrix size grows large, are shown to be intimately linked with matrices from the Gaussian unitary ensemble (GUE). When the spiked population eigenvalues have unit multiplicity, the fluctuations follow a central limit theorem. This result is used to develop an original framework for the detection and diagnosis of local failures in large sensor networks, for known or unknown failure magnitude.

J47. A. Kammoun, R. Couillet, J. Najim, M. Debbah, "Performance of capacity inference methods under colored interference," IEEE Transactions on Information Theory, vol. 59, no. 2, pp. 1129-1148, 2013.

Abstract. The problem of fast point-to-point MIMO channel mutual information estimation is addressed, in the situation where the receiver undergoes unknown colored interference, whereas the channel with the transmitter is perfectly known. The considered scenario assumes that the estimation is based on a few channel use observations during a short sensing period. Using large dimensional random matrix theory, an estimator referred to as *G-estimator* is derived. This estimator is proved to be consistent as the number of antennas and observations grow large and its asymptotic performance is analyzed. In particular, the G-estimator satisfies a central limit theorem with asymptotic Gaussian fluctuations. Simulations are provided which strongly support the theoretical results, even for small system dimensions.

J48. R. Couillet, J. Hoydis, M. Debbah, "Random beamforming over quasi-static and fading channels: A deterministic equivalent approach," IEEE Transactions on Information Theory, vol. 58, no. 10, pp. 6392-6425, 2012.

Abstract. In this work, we study the performance of random isometric precoders over quasi-static and correlated fading channels. We derive deterministic approximations of the mutual information and the signal-to-interference-plus-noise ratio (SINR) at the output of the minimum-mean-square-error (MMSE) receiver and provide simple provably converging fixed-point algorithms for their computation. Although these approximations are only proven exact in the asymptotic regime with infinitely many antennas at the transmitters and receivers, simulations suggest that they closely match the performance of small-dimensional systems. We exemplarily apply our results to the performance analysis of multi-cellular communication systems, multiple-input multiple-output multiple-access channels (MIMO-MAC), and MIMO interference channels. The mathematical analysis is based on the Stieltjes transform method. This enables the derivation of deterministic equivalents of functionals of large-dimensional random matrices. In contrast to previous works, our analysis does not rely on arguments from free probability theory which enables the consideration of random matrix models for which asymptotic freeness does not hold. Thus, the results of this work are also a novel contribution to the field of random matrix theory and applicable to a wide spectrum of practical systems.

J49. S. Wagner, R. Couillet, M. Debbah, D. T. M. Slock, "Large System Analysis of Linear Precoding in MISO Broadcast Channels with Limited Feedback", IEEE Transactions on Information Theory, vol. 58, no. 7, pp. 4509-4537, 2012.

Abstract. In this paper, we study the sum rate performance of zero-forcing (ZF) and regularized ZF (RZF) precoding in large MISO broadcast systems under the assumptions of imperfect channel state information at the transmitter and per-user

channel transmit correlation. Our analysis assumes that the number of transmit antennas M and the number of single-antenna users K are large while their ratio remains bounded. We derive deterministic approximations of the empirical signal-to-interference plus noise ratio (SINR) at the receivers, which are tight as $M, K \to \infty$. In the course of this derivation, the per-user channel correlation model requires the development of a novel deterministic equivalent of the empirical Stieltjes transform of large dimensional random matrices with generalized variance profile. The deterministic SINR approximations enable us to solve various practical optimization problems. Under sum rate maximization, we derive (i) for RZF the optimal regularization parameter, (ii) for ZF the optimal number of users, (iii) for ZF and RZF the optimal power allocation scheme and (iv) the optimal amount of feedback in large FDD/TDD multi-user systems. Numerical simulations suggest that the deterministic approximations are accurate even for small M, K.

J50. R. Couillet, J. W. Silverstein, Z. Bai, M. Debbah, "Eigen-Inference for Energy Estimation of Multiple Sources", IEEE Transactions on Information Theory, vol. 57, no. 4, pp. 2420-2439, 2011.

Abstract. In this paper, a new method is introduced to blindly estimate the transmit power of multiple signal sources in multi-antenna fading channels, when the number of sensing devices and the number of available samples are sufficiently large compared to the number of sources. Recent advances in the field of large dimensional random matrix theory are used that result in a simple and computationally efficient consistent estimator of the power of each source. A criterion to determine the minimum number of sensors and the minimum number of samples required to achieve source separation is then introduced. Simulations are performed that corroborate the theoretical claims and show that the proposed power estimator largely outperforms alternative power inference techniques.

J51. R. Couillet, M. Debbah, J. W. Silverstein, "A Deterministic Equivalent for the Analysis of Correlated MIMO Multiple Access Channels", IEEE Transactions on Information Theory, vol. 57, no. 6, pp. 3493-3514, 2011.

Abstract. In this article, novel deterministic equivalents for the Stieltjes transform and the Shannon transform of a class of large dimensional random matrices are provided. These results are used to characterise the ergodic rate region of multiple antenna multiple access channels, when each point-to-point propagation channel is modelled according to the Kronecker model. Specifically, an approximation of all rates achieved within the ergodic rate region is derived and an approximation of the linear precoders that achieve the boundary of the rate region as well as an iterative water-filling algorithm to obtain these precoders are provided. An original feature of this work is that the proposed deterministic equivalents are proved valid even for strong correlation patterns at both communication sides. The above results are validated by Monte Carlo simulations.

J52. R. Couillet, M. Debbah, "A Bayesian Framework for Collaborative Multi-Source Signal Sensing", IEEE Transactions on Signal Processing, vol. 58, no. 10, pp. 5186-5195, 2010.

Abstract. This paper introduces a Bayesian framework to detect multiple signals embedded in noisy observations from a sensor array. For various states of knowledge on the communication channel and the noise at the receiving sensors, a marginalization procedure based on recent tools of finite random matrix theory, in conjunction with the maximum entropy principle, is used to compute the hypothesis selection criterion. Quite remarkably, explicit expressions for the Bayesian detector are derived which enable to decide on the presence of signal sources in a noisy wireless environment. The proposed Bayesian detector is shown to outperform the classical power detector when the noise power is known and provides very good performance

for limited knowledge on the noise power. Simulations corroborate the theoretical results and quantify the gain achieved using the proposed Bayesian framework.

J53. R. Couillet, A. Ancora, M. Debbah, "Bayesian Foundations of Channel Estimation for Cognitive Radios", Advances in Electronics and Telecommunications, vol. 1, no. 1, pp. 41-49, 2010.

Abstract. In this paper, we revisit the philosophical foundations of the field of channel estimation. Our main intention is to come up with a partial answer to the question : "given some available sensed signals, how should cognitive radios ideally perform channel estimation ?". We specifically introduce a general framework to provide optimal channel estimates under any prior knowledge at the sensing device. Our discussion is articulated as a top-down approach, introducing successively (i) a discussion on the philosophical foundations of channel estimation as a simplification means for the general problem of wireless detection, (ii) an information theoretically optimal approach to channel detection assuming the sensing device has infinite memory, and (iii) a derived optimal approach when limited memory size is accounted for. The key mathematical tools used in this discussion emerge from Bayesian probability theory and are known as the maximum entropy principle and the minimum update principle. Derivations are carried out for the particular case of channel estimation in orthogonal frequency division multiplexing (OFDM) systems. While some theoretical results will be proven to match already known techniques, such as Kalman filters, another set of novel results will be shown by simulations to perform better than known channel estimation schemes.

J54. R. Couillet, M. Debbah,"Le téléphone du futur : plus intelligent pour une exploitation optimale des fréquences" Revue de l'Electricité et de l'Electronique, no. 6, pp. 71-83, 2010.

Résumé. Au jour où les communications mobiles demandent de plus en plus de vitesse de transfert de larges volumes de données à destination d'utilisateurs de plus en plus nombreux, il apparaît que les limites physiques des protocoles de communication sont bientôt atteintes. Une révolution technologique est ainsi nécessaire et sur le point d'éclore : celle-ci passe par la mise en place de systèmes de communications opportunistes, coopératifs, autonomes et idéalement suffisamment intelligents pour servir au mieux les requêtes de l'utilisateur. Ces différents aspects, certains d'ores et déjà d'actualité, d'autres à l'état embryonnaire sont discutés successivement dans cette étude.

J55. R. Couillet, M. Debbah, "Mathematical foundations of cognitive radios", Journal of Telecommunications and Information Technologies, no. 4, 2009.

Abstract. Recently, much interest has been directed towards software defined radios and embedded intelligence in telecommunication devices. However, no theoretical framework for cognitive radios has ever been proposed. In this paper, we introduce an information theoretic point of view on cognitive radios. Specifically, our motivation in this work is to embed human-like intelligence in mobile wireless devices, following the three century-old work on Bayesian probability theory, the maximum entropy principle and minimal probability update. This allows us to partially answer such questions as "what are the signal detection capabilities of a wireless device ?" or "when facing a situation in which most parameters are missing, how to react ?". As an introduction, we will present two examples from the same authors using the cognitive framework namely multi-antenna channel modelling and signal sensing.

J56. R. Couillet, M. Debbah, "Outage performance of flexible OFDM schemes in packet-switched transmissions", Eurasip Journal on Advances on Signal Processing, Volume 2009, Article ID 698417, 2009.

Abstract. In this paper, α-OFDM, a generalization of the OFDM modulation, is proposed. This new modulation enhances the outage capacity performance of bursty communications. The α-OFDM scheme is easily implementable as it only requires an additional time symbol rotation after the IDFT stage and a subsequent phase rotation of the cyclic prefix. The physical effect of the induced rotation is to slide the DFT window over the frequency spectrum. When successively used with different angles α at the symbol rate, α-OFDM provides frequency diversity in block fading channels. Interestingly, simulation results show a substantial gain in terms of outage capacity and outage BER in comparison with classical OFDM modulation schemes. The framework is extended to multi-antenna and multi-cellular OFDM based standards. Practical simulations, in the context of 3GPP-LTE, called hereafter α-LTE, sustain our theoretical claims.

Conference Publications

C1. M. Seddik, R. Couillet, M. Tamaazousti, "A Random Matrix Analysis of Learning with α-Dropout", International Conference on Machine Learning (ICML'20), Artemiss workshop, Graz, Autria, 2020.

Abstract. This article studies a one hidden layer neural network with generalized Dropout (α-Dropout), where the dropped out features are replaced with an arbitrary value α. Specifically, under a large dimensional data and network regime, we provide the generalization performances for this network on a binary classification problem. We notably demonstrate that a careful choice of $\alpha \neq 0$ can drastically improve the generalization performances of the classifier.

C2. M. Seddik, C. Louart, R. Couillet, M. Tamaazousti, "The Unexpected Deterministic and Universal Behavior of Large Softmax Classifiers", (submitted to) Conference on Neural Information Processing Systems (NeurIPS'20), Vacouver, Canada, 2020.

Abstract. In this paper, we provide a large dimensional analysis of the Softmax classifier, extensively used in modern neural networks. We discover and prove that, when the classifier is trained on data satisfying reasonable concentration assumptions, its weights become deterministic and solely depend on the statistical means and covariances of the data. As a striking consequence, despite the implicit and non-linear nature of the underlying optimization problem, the performance of the Softmax classifier is the same as if performed on a mere Gaussian mixture model, thereby disrupting the intuition that non-linearities inherently extract advanced statistical features from the data. Our findings are in particular theoretically sustained as well as numerically confirmed on CNN representations of images produced by GANs.

C3. L. Dall'Amico, R. Couillet, N. Tremblay, "Community detection in sparse time-evolving graphs with a dynamical Bethe-Hessian", (submitted to) Conference on Neural Information Processing Systems (NeurIPS'20), Vacouver, Canada, 2020.

Abstract. This article considers the problem of community detection in sparse dynamical graphs in which the community structure evolves over time. A fast spectral algorithm based on an extension of the Bethe-Hessian matrix is proposed, which benefits from the positive correlation in the class labels and in their temporal evolution and is designed to be applicable to any dynamical graph with a community structure. Under the dynamical degree-corrected stochastic block model, in the case of two classes of equal size, we demonstrate and support with extensive simulations that our proposed algorithm is capable of making non-trivial commu- nity reconstruction as soon as theoretically possible, thereby reaching the optimal detectability threshold and provably outperforming competing spectral methods.

C4. Z. Liao, R. Couillet, M. Mahoney "A random matrix analysis of random Fourier features: beyond the Gaussian kernel, a precise phase transition, and the correspon-

ding double descent", (submitted to) Conference on Neural Information Processing Systems (NeurIPS'20), Vacouver, Canada, 2020.

Abstract. This article characterizes the exact asymptotics of random Fourier feature (RFF) regression, in the realistic setting where the number of data samples n, their dimension p, and the dimension of feature space N are all large and comparable. In this regime, the random RFF Gram matrix no longer converges to the well-known limiting Gaussian kernel matrix (as it does when $N \to \infty$ alone), but it still has a tractable behavior that is captured by our analysis. This analysis also provides accurate estimates of training and test regression errors for large n, p, N. Based on these estimates, a precise characterization of two qualitatively different phases of learning, including the phase transition between them, is provided ; and the corresponding double descent test error curve is derived from this phase transition behavior. These results do not depend on strong assumptions on the data distribution, and they perfectly match empirical results on real-world data sets.

C5. T. Zarrouk, R. Couillet, F. Chatelain, N. Le Bihan, "Performance-Complexity Trade-Off in Large Dimensional Statistics", International Workshop on Machine Learning for Signal Processing (MLSP'20), Espoo, Finland, 2020.

Abstract. This article introduces a random matrix framework for the analysis of the trade-off between performance and complexity in a class of machine learning algorithms, under a large dimensional data $X = [x_1, \ldots, x_n] \in \mathbb{R}^{p \times n}$ regime. Specifically, we analyze the spectral properties of $K \odot B \in \mathbb{R}^{n \times n}$, for a kernel random matrix $K = \{f(\frac{1}{p}\|x_i - x_j\|^2)\}_{i,j=1}^n$ upon which a sparsity mask $B \in \{0,1\}^{n \times n}$ is applied : this reduces the number of K_{ij} to evaluate, thereby reducing complexity, while weakening the power of statistical inference on K, thereby impeding performance. Assuming the data structured as $X = Z + \sqrt{n}\mu v^\mathsf{T}$ for informative vectors $\mu \in \mathbb{R}^p$, $v \in \mathbb{R}^n$, and white noise Z, we exhibit a phase transition phenomenon below which spectral methods must fail and which is a function of the sparsity structure of B. This finds immediate applications to the fundamental limits of complexity-reduced spectral clustering as well as principal component analysis.

C6. M. Seddik, R. Couillet, M. Tamaazousti "Random Matrix Theory Proves that Deep Learning Representations of GAN-data Behave as Gaussian Mixtures", International Conference on Machine Learning (ICML'20), Graz, Autria, 2020.

Abstract. This paper shows that deep learning (DL) representations of data produced by generative adversarial nets (GANs) are random vectors which fall within the class of so-called concentrated random vectors. Further exploiting the fact that Gram matrices, of the type $G = X'X$ with $X = [x_1, ..., x_n] \in \mathbb{R}^{p \times n}$ and x_i independent concentrated random vectors from a mixture model, behave asymptotically (as $n, p \to \infty$) as if the xi were drawn from a Gaussian mixture, suggests that DL representations of GAN-data can be fully described by their first two statistical moments for a wide range of standard classifiers. Our theoretical findings are validated by generating images with the BigGAN model and across different popular deep representation networks.

C7. M. Tiomoko, H. Tiomoko, R. Couillet "Deciphering and Optimizing Multi-Task and Transfer Learning: a Random Matrix Approach", (submitted to) International Conference on Machine Learning (ICML'20), Graz, Austria, 2020.

Abstract. This article provides theoretical insights into the inner workings of multi-task and transfer learning methods. To this end, we study the particularly tractable (least-square) support vector machine (LS-SVM) extension to multi-task learning (MTL) in the limit of large (p) and numerous data (n). This is achieved by means of a random matrix analysis applied to a Gaussian mixture model, and demonstrates that, as $n, p \to \infty$, the performance of MTL LS-SVM converges to a fully deterministic limit involving basic (small-dimensional) statistics of the data model.

Our major conclusions are that (i) the standard MTL LS-SVM algorithm is in general strongly biased and may dramatically fail (to the point that individual single-task LS-SVMs may outperform the MTL approach, even for quite resembling tasks) : our analysis provides a simple method to correct these biases, and (ii) the sufficient statistics at play in the method are revealed and can be efficiently estimated, even for quite small datasets. The latter aspect is exploited to automatically optimize the hyperparameters without resorting to any cross-validation procedure.

Experiments on popular datasets are provided to further justify the applicability of our proposed approach to actual data. These experiments notably demonstrate that, when hyperparametrized using our theoretical findings, the simple and computationally-efficient as MTL LS-SVM algorithm is largely competitive, and even outperforms, much more elaborate state-of-the-art multitask and transfer learning methods.

C8. L. Dall'Amico, R. Couillet, N. Tremblay "Optimal Laplacian Regularization for Sparse Spectral Community Detection", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'20), Barcelona, Spain, 2020.

Abstract. Regularization of the classical Laplacian matrices was empirically shown to improve spectral clustering in sparse networks. It was ob- served that small regularizations are preferable, but this point was left as a heuristic argument. In this paper we formally determine a proper regularization which is intimately related to alternative state-of-the-art spectral techniques for sparse graphs.

C9. M. Tiomoko, C. Louart, R. Couillet "Large Dimensional Asymptotics of Multi-Task Learning", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'20), Barcelona, Spain, 2020.

Abstract. Inspired by human learning, which transfers knowledge from learned tasks to solve new tasks, multitask learning aims at simultaneously solving multiple tasks by a smart exploitation of their similarities. How to relate the tasks so to optimize their performances is however a largely open problem. Based on a random matrix approach, this article proposes an asymptotic analysis of a support vector machine-inspired multitask learning scheme. The asymptotic performance of the algorithm, validated on both synthetic and real data, sets forth the relation between the statistics of the data in each task and the hyperparameters relating the tasks together. The article, as such, provides first insights on an offline control of multitask learning, which finds natural connections to the currently popular transfer learning paradigm.

C10. L. Dall'Amico, N. Tremblay, R. Couillet "Optimized Deformed Laplacian for Spectrum-based Community Detection in Sparse Heterogeneous Graphs", Neural Information Processing Systems (NeurIPS'19), Vancouver, Canada, 2019.

Abstract. Spectral clustering is one of the most popular, yet still incompletely understood, methods for community detection on graphs. This article studies spectral clustering based on the Bethe-Hessian matrix $H_r = (r^2 - 1)I_n + D - rA$ for sparse heterogeneous graphs (following the degree-corrected stochastic block model) in a two-class setting. For a specific value $r = \zeta$, clustering is shown to be insensitive to the degree heterogeneity. We then study the behavior of the informative eigenvector of H_ζ and, as a result, predict the clustering accuracy. The article concludes with an overview of the generalization to more than two classes along with extensive simulations on synthetic and real networks corroborating our findings.

C11. Z. Liao, R. Couillet, "On Inner-product Kernels of High Dimensional Data", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

Abstract. In this article we investigate the eigenspectrum of inner-product kernel matrices of the type $\sqrt{p}K = \{f(x_i^\mathsf{T} x_j/\sqrt{p})\}_{i,j=1}^n$. Under a two-class mixture modeling of the input data $x_i \in \mathbb{R}^p$, we position ourselves in the regime where the number of data n and their dimension p are both large and comparable, and show, for a wide range of kernel functions f, that the spectrum of K only depends on f via three key parameters, with only two of them useful in extracting the statistical structure from the data. By carefully balancing these two parameters, a huge gain in classification performance is observed on real-world datasets.

C12. R. Couillet, "High Dimensional Robust Classification: A Random Matrix Analysis", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

Abstract. This article proposes a random matrix analysis of the spectral properties of a new robust kernel matrix model adapted to elliptically distributed large dimensional data mixtures. It is shown that these kernel matrices, based on robust estimators of scatter, when finely tuned, can perform asymptotic non-trivial (unsupervised) classification while sample covariance matrices are ineffective. Unlike in conventional robust statistics wisdom though, the "maximally robust estimators (such as Tyler's estimator of scatter) also break asymptotic classification feasibility. This entails the existence of an optimal robustness-classification trade-off which we discuss.

C13. R. Couillet, "A Random Matrix Analysis and Optimization Framework to Large Dimensional Transfer Learning", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

Abstract. This article proposes a first performance analysis and optimization of a simple transfer learning method, extending the standard least squares support vector machine. By means of a random matrix analysis, we prove that, for simultaneously large and numerous data, the correct classification rate of the learning task is asymptotically predictable and the hyperparameters in the problem are easily tuned so to maximize the output performance. Simulations confirm our findings. This preliminary work opens the path to a systematic exploration of transfer learning methods by means of large dimensional statistics.

C14. A. Kadavankandy, R. Couillet, "Asymptotic Gaussian Fluctuations of Spectral Clustering Eigenvectors", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

Abstract. In this article, we analyze the asymptotic distribution of the eigenvectors used in spectral clustering of random graphs and in kernel spectral clustering of high dimensional Gaussian random vectors. For dense random graphs drawn from the Stochastic Block Model (SBM), we prove that the isolated dominant eigenvectors of the modularity matrix behave asymptotically like Gaussian random vectors with independent components. As opposed to previous works on SBM eigenvectors, we deal with a more challenging and practically meaningfulgrowth rate of the edge probabilities. Similarly for kernel clustering of a two-class Gaussian mixture we prove the asymptotic Gaussianity of the finite-dimensional marginals of the single isolated eigenvector. We present two practical applications of our results : predicting the classification accuracy of clustering algorithms, and speeding up the convergence of the final Expectation Maximization (EM) clustering using an improved initialization.

C15. C. Louart, R. Couillet, "A concentration of measure perspective to robust statistics", IEEE International Workshop on Computational Advances in Multi-Sensor

Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

Abstract. We provide promising mathematical considerations for the study of robust scatter matrices in the regime where the data number and dimension are large. Chiefly, we present a new realistic model for data with an assumption inspired from the concentration of measure phenomenon. Our technical contribution is to provide a deterministic equivalent for the robust scatter matrix (i) under relaxed assumptions when compared to the robust statistics literature and (ii) with an original proof based on the introduction of a new semi-metric. This brings simultaneously a new methodological approach to robust statistics analysis and a wider application spectrum to more realistic large dimensional data models.

C16. M. Tiomoko, R. Couillet, "Estimation of Covariance Matrix Distances in the High Dimension Low Sample Size Regime", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'19), Guadeloupe, France, 2019.

Abstract. A broad family of distances between two covariance matrices $C_1, C_2 \in \mathbb{R}^{p\times p}$, among which the Frobenhius, Fisher, Battacharrya distances as well as the Kullback-Leibler, Rényi and Wasserstein divergence for centered Gaussian distributions can be expressed as functionals $\frac{1}{p}\sum_{i=1}^{p} f\left(\lambda_i(C_1^{-1}C_2)\right)$ or $\frac{1}{p}\sum_{i=1}^{p} f\left(\lambda_i(C_1C_2)\right)$ of the eigenvalue distribution of $C_1^{-1}C_2$ or C_1C_2. Consistent estimates of such distances based on few (n_1, n_2) samples $x_i \in \mathbb{R}^p$ having covariance C_1, C_2 have been recently proposed using random matrix tools in the regime where $n_1, n_2 \sim p$. These estimates however demand that $n_1, n_2 > p$ for most functions f. The article proposes to alleviate this limitation using a polynomial approximation approach. The proposed method is supported by simulations in practical applications.

C17. M. Tiomoko, R. Couillet, "Random Matrix-Improved Estimation of the Wasserstein Distance between two Centered Gaussian Distributions", European Signal Processing Conference (EUSIPCO'19), A Coruna, Spain, 2019. **Best student paper award**

Abstract. This article proposes a method to consistently estimate functionals $\frac{1}{p}\sum_{i=1}^{p} f(\lambda_i(C_1C_2))$ of the eigenvalues of the product of two covariance matrices $C_1, C_2 \in \mathbb{R}^{p\times p}$ based on the empirical estimates $\lambda_i(\hat{C}_1\hat{C}_2)$ ($\hat{C}_a = \frac{1}{n_a}\sum_{i=1}^{n_a} x_i^{(a)} x_i^{(a)\mathsf{T}}$), when the size p and number n_a of the (zero mean) samples $x_i^{(a)}$ are similar. As a corollary, a consistent estimate of the Wasserstein distance (related to the case $f(t) = \sqrt{t}$) between centered Gaussian distributions is derived.

The new estimate is shown to largely outperform the classical sample covariance-based "plug-in" estimator. Based on this finding, a practical application to covariance estimation is then devised which demonstrates potentially significant performance gains with respect to state-of-the-art alternatives.

C18. M. Seddik, M. Tamaazousti, R. Couillet, "Kernel Random Matrices of Large Concentrated Data: The Example of GAN-Generated Image", .IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

Abstract. Based on recent random matrix advances in the analysis of kernel methods for classification and clustering, this paper proposes the study of large kernel methods for a wide class of random inputs, i.e., concentrated data, which are more generic than Gaussian mixtures. The concentration assumption is motivated by the fact that one can use generative models to design complex data structures, through Lipschitzally transformed concentrated vectors (e.g., Gaussian) which remain concentrated vectors. Applied to spectral clustering, we demonstrate that our theoretical findings closely match the behavior of large kernel matrices, when

considering the fed- in data as CNN representations of GAN-generated images (i.e., concentrated vectors by design).

C19. M. Tiomoko, F. Bouchard, G. Ginholac, R. Couillet "Random Matrix Improved Covariance Estimation for a Large Class of Metrics", International Conference on Machine Learning (ICML), Long Beach, USA, 2019.

Abstract. Relying on recent advances in statistical estimation of covariance distances based on random matrix theory, this article proposes an improved covariance and precision matrix estimation for a wide family of metrics. The method is shown to largely outperform the sample covariance matrix estimate and to compete with state-of-the-art methods, while at the same time being computationally simpler. Applications to linear and quadratic discriminant analyses also demonstrate significant gains, therefore suggesting practical interest to statistical machine learning.

C20. L. Dall'Amico, R. Couillet "Community Detection in Sparse Realistic Graphs: Improving the Bethe Hessian", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

Abstract. This article improves over the recently proposed Bethe Hessian matrix for community detection on sparse graphs, assuming here a more realistic setting where node degrees are inhomogeneous. We notably show that the parametrization proposed in the seminal work on the Bethe Hessian clustering can be ameliorated with positive consequences on correct classification rates. Extensive simulations support our claims.

C21. X. Mai, R. Couillet "Revisiting and Improving Semi-Supervised Learning: A Large Dimensional Approach", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

Abstract. The recent work (Mai, Couillet'2017) shows that in the big data regime (i.e., numerous high dimensional data), the popular semi-supervised graph regularization, known as semi-supervised Laplacian regularization, fails to effectively extract information from unlabelled data. In response to this problem, we propose in this article an improved approach based on a simple yet fundamental update of the classical method. The effectiveness of the former is supported by both asymptotic results and simulations on finite data samples.

C22. H. Tiomoko Ali, S. Liu, Y. Yilmaz, R. Couillet, I. Rajapakse, A. Hero, "Latent Heterogeneous Multilayer Community Detection", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

Abstract. We propose a method for simultaneously detecting shared and unshared communities in heterogeneous multilayer weighted and undirected networks. The multilayer network is assumed to follow a generative probabilistic model that takes into account the similarities and dissimilarities between the communities. We make use of a variational Bayes approach for jointly inferring the shared and unshared hidden communities from multilayer network observations. We show the robustness of our approach compared to state-of-the art algorithms in detecting disparate (shared and private) communities on synthetic data as well as on real genome-wide fibroblast proliferation dataset.

C23. Z. Liao, X. Mai, R. Couillet "A Large n, p Analysis of Logistic Regression: Asymptotic Performance and New Insights", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

Abstract. Logistic regression, one of the most popular machine learning binary classification methods, has been long believed to be unbiased. In this paper, we consider the "hard" classification problem of separating high dimensional Gaussian

vectors, where the data dimension p and the sample size n are both large. Based on recent advances in random matrix theory (RMT) and high dimensional statistics, we evaluate the asymptotic distribution of the logistic regression classifier and consequently, provide the associated test performance. This brings new insights into the internal mechanism of logistic regression classifier, including a possible bias in the separating hyperplane, as well as on practical issues such as hyper-parameter tuning, thereby opening the door to novel RMT-inspired improvements.

C24. M. Tiomoko, R. Couillet, S. Zozor, E. Moisan, "Improved Estimation of the Distance between Covariance Matrices", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'19), Brighton, UK, 2019.

Abstract. A wide range of machine learning and signal processing applications involve data discrimination through statistical covariances. A broad family of covariance matrix metrics, among which the Frobenius, Fisher, Bhattacharyya distances, as well as the Kullback-Leibler or Rényi divergences, are regularly exploited. However, not being directly accessible, these metrics are assessed through *empirical* sample covariances, which we shall show here may lead to dramatically erroneous estimates in large dimensional data.

C25. R. Couillet, Z. Liao, X. Mai, "Classification Asymptotics in the Random Matrix Regime", European Signal Processing Conference (EUSIPCO'18), Rome, Italy, 2018.

Abstract. This article discusses the asymptotic performance of classical machine learning classification methods (from discriminant analysis to neural networks) for simultaneously large and numerous Gaussian mixture modelled data. We first provide theoretical bounds on the minimally discriminable class means and covariances under an oracle setting, which are then compared to recent theoretical findings on the performance of machine learning. Non-obvious phenomena are discussed, among which surprising phase transitions in the optimal performance rates for specific hyperparameter settings.

C26. M. Seddik, M. Tamaazousti, R. Couillet, "A Kernel Random Matrix-Based Approach for Sparse PCA", International Conference on Learning Representations (ICLR'19), New Orleans, USA, 2019.

Abstract. In this paper, we present a random matrix approach to recover sparse principal components from n p-dimensional vectors. Specifically, considering the large dimensional setting where $n, p \to \infty$ with $p/n \to c \in (0, \infty)$ and under Gaussian vector observations, we study kernel random matrices of the type $f(\hat{C})$, where f is a three-times continuously differentiable function applied entry-wise to the sample covariance matrix $\hat{C}$ of the data. Then, assuming that the principal components are sparse, we show that taking f in such a way that $f'(0) = f''(0) = 0$ allows for powerful recovery of the principal components, thereby generalizing previous ideas involving more specific f functions such as the soft-thresholding function.

C27. X. Mai, R. Couillet, "Semi-Supervised Spectral Clustering", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2018.

Abstract. In this article, we propose a semi-supervised version of spec- tral clustering, a widespread graph-based unsupervised learn- ing method. The semi-supervised spectral clustering has the advantage of producing consistent classification of data with sufficiently large number of labelled or unlabelled data, un- like classical graph-based semi-supervised methods which are only consistent on labelled data. Theoretical arguments are provided to support the proposition of this novel approach, as well as empirical evidence to confirm the theoretical claims and demonstrate its superiority over other graph-based semi- supervised methods.

C28. Z. Liao, R. Couillet, "The Dynamics of Learning: A Random Matrix Approach", International Conference on Machine Learning, Stockholm, Sweden, 2018.

Abstract. Understanding the learning dynamics of neural networks is one of the key issues for the improvement of optimization algorithms as well as for the theoretical comprehension of why deep neural nets work so well today. In this paper, we introduce a random matrix-based framework to analyze the learning dynamics of a single-layer linear network on a binary classification problem, for data of simultaneously large dimension and size, trained by gradient descent. Our results provide rich insights into common questions in neural nets, such as overfitting, early stopping and the initialization of training, thereby opening the door for future studies of more elaborate structures and models appearing in today's neural networks.

C29. Z. Liao, R. Couillet, "On the Spectrum of Random Features Maps of High Dimensional Data", International Conference on Machine Learning, Stockholm, Sweden, 2018.

Abstract. Abstract : Random feature maps are ubiquitous in modern statistical machine learning, where they generalize random projections by means of powerful, yet often difficult to analyze nonlinear operators. In this paper we leverage the "concentration" phenomenon induced by random matrix theory to perform a spectral analysis on the Gram matrix of these random feature maps, here for Gaussian mixture models of simultaneously large dimension and size. Our results are instrumental to a deeper understanding on the interplay of the nonlinearity and the statistics of the data, thereby allowing for a better tuning of random feature-based techniques.

C30. H. Tiomoko Ali, A. Kammoun, R. Couillet, "Random matrix-improved kernels for large dimensional spectral clustering", Statistical Signal Processing Workshop (SSP'18), Freiburg, Germany, 2018.

Abstract. Leveraging on recent random matrix advances in the performance analysis of kernel methods for classification and clustering, this article proposes a new family of kernel functions theoretically largely outperforming standard kernels in the context of asymptotically large and numerous datasets. These kernels are designed to discriminate statistical means and covariances across data classes at a theoretically minimal rate (with respect to data size). Applied to spectral clustering, we demonstrate the validity of our theoretical findings both on synthetic and real-world datasets (here, the popular MNIST database as well as EEG recordings on epileptic patients).

C31. L. Yang, M. R. McKay, R. Couillet, "Random Matrix-Optimized High-Dimensional MVDR Beamforming", Statistical Signal Processing Workshop (SSP'18), Freiburg, Germany, 2018.

Abstract. A new approach to minimum variance distortionless response (MVDR) beamforming is proposed under the assumption of simultaneously large numbers of array sensors and observations. The key to our method is the design of an inverse covariance estimator which is appropriately optimized for the MVDR application. This is obtained by exploiting spectral properties of spiked covariance models in random matrix theory. Our proposed solution is simple to implement and is shown to yield performance improvements over competing approaches.

C32. C. Louart, R. Couillet, "A Random Matrix and Concentration Inequalities Framework for Neural Networks Analysis", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'18), Calgary, Canada, 2018.

Abstract. This article provides a theoretical analysis of the asymptotic performance of a regression or classification task performed by a simple random neural network. This result is obtained by leveraging a new framework at the crossroads between

random matrix theory and the concentration of measure theory. This approach is of utmost interest for neural network analysis at large in that it naturally dismisses the difficulty induced by the non-linear activation functions, so long that these are Lipschitz functions. As an application, we provide formulas for the limiting law of the random neural network output and compare them conclusively to those obtained practically on handwritten digits databases.

C33. H. Tiomoko Ali, A. Kammoun, R. Couillet, "Random matrix asymptotic of inner product kernel spectral clustering", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'18), Calgary, Canada, 2018.

Abstract. We study in this article the asymptotic performance of spectral clustering with inner product kernel for Gaussian mixture models of high dimension with numerous samples. As is now classical in large dimensional spectral analysis, we establish a phase transition phenomenon by which a minimum distance between the class means and covariances is required for clustering to be possible from the dominant eigenvectors. Beyond this phase transition, we evaluate the asymptotic content of the dominant eigenvectors thus allowing for a full characterization of clustering performance. However, a surprising finding is that in some particular scenarios, the phase transition does not occur and clustering can be achieved irrespective of the class means and covariances. This is evidenced here in the case of the mixture of two Gaussian datasets having the same means and arbitrary difference between covariances.

C34. K. Elkalil, A. Kammoun, R. Couillet, T. Al-Naffouri, M.-S. Alouini, "Asymptotic Performance of Regularized Quadratic Discriminant Analysis Based Classifiers", IEEE International Workshop on Machine Learning for Signal Processing (MLSP'17), Roppongi, Tokyo, Japan, 2017. **Best student paper award finalist**

Abstract. This paper carries out a large dimensional analysis of the standard regularized quadratic discriminant analysis (QDA) classifier designed on the assumption that data arise from a Gaussian mixture model. The analysis relies on fundamental results from random matrix theory (RMT) when both the number of features and the cardinality of the training data within each class grow large at the same pace. Under some mild assumptions, we show that the asymptotic classification error converges to a deterministic quantity that depends only on the covariances and means associated with each class as well as the problem dimensions. Such a result permits a better understanding of the performance of regularized QDA and can be used to determine the optimal regularization parameter that minimizes the misclassification error probability. Despite being valid only for Gaussian data, our theoretical findings are shown to yield a high accuracy in predicting the performances achieved with real data sets drawn from popular real data bases, thereby making an interesting connection between theory and practice.

C35. Z. Liao, R. Couillet, "Random matrices meet machine learning: a large dimensional analysis of LS-SVM", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17), New Orleans, USA, 2017.

Abstract. This article proposes a performance analysis of kernel least squares support vector machines (LS-SVMs) based on a random matrix approach, in the regime where both the dimension of data p and their number n grow large at the same rate. Under a two-class Gaussian mixture model for the input data, we prove that the LS-SVM decision function is asymptotically normal with means and covariances shown to depend explicitly on the derivatives of the kernel function. This provides improved understanding along with new insights into the internal workings of SVM-type methods for large datasets.

C36. X. Mai, R. Couillet, "The counterintuitive mechanism of graph-based semi-supervised

learning in the big data regime", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17), New Orleans, USA, 2017.

Abstract. In this article, a new approach is proposed to study the performance of graph-based semi-supervised learning methods, under the assumptions that the dimension of data p and their number n grow large at the same rate and that the data arise from a Gaussian mixture model. Unlike small dimensional systems, the large dimensions allow for a Taylor expansion to linearize the weight (or kernel) matrix W, thereby providing in closed form the limiting performance of semi-supervised learning algorithms. This notably allows to predict the classification error rate as a function of the normalization parameters and of the choice of the kernel function. Despite the Gaussian assumption for the data, the theoretical findings match closely the performance achieved with real datasets, particularly here on the popular MNIST database.

C37. C. Louart, R. Couillet, "Harnessing neural networks: a random matrix approach", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'17), New Orleans, USA, 2017.

Abstract. This article proposes an original approach to the performance understanding of large dimensional neural networks. In this preliminary study, we study a single hidden layer feed-forward network with random input connections (also called extreme learning machine) which performs a simple regression task. By means of a new random matrix result, we prove that, as the size and cardinality of the input data and the number of neurons grow large, the network performance is asymptotically deterministic. This entails a better comprehension of the effects of the hyper-parameters (activation function, number of neurons, etc.) under this simple setting, thereby paving the path to the harnessing of more involved structures.

C38. H. Tiomoko Ali, R. Couillet, "Random Matrix Improved Community Detection in Heterogeneous Networks", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2016.

Abstract. This article proposes a new spectral method for community detection in large dense graphs following the degree-corrected stochastic block model. We theoretically support and analyze an approach based on a novel "α-regularization" of the modularity matrix. We next provide a consistent estimator of the choice of α inducing the most favorable community detection in worst case scenarios. We then prove that spectral clustering should be performed on a $1-\alpha$ regularization of the dominant eigenvectors (rather than the eigenvectors themselves) to compensate for biases due to degree heterogeneity. Although focused on dense graph models, our algorithm is shown to be very promising on real world networks with competitive performances versus the Bethe Hessian spectral method developed for sparse homogeneous networks.

C39. R. Couillet, A. Kammoun, "Random Matrix Improved Subspace Clustering", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2016.

Abstract. This article introduces a spectral method for statistical subspace clustering. The method is built upon standard kernel spectral clustering techniques, however carefully tuned by theoretical understanding arising from random matrix findings. We show in particular that our method provides high clustering performance while standard kernel choices provably fail. An application to user grouping based on vector channel observations in the context of massive MIMO wireless communication networks is provided.

C40. R. Couillet, G. Wainrib, H. Sevi, H. Tiomoko Ali, "A Random Matrix Approach to Recurrent Neural Networks", International Conference on Machine Learning (ICML), New York, USA, 2016.

Abstract. Recurrent neural networks, especially in their linear version, have provided many qualitative insights on their performance under different configurations. This article provides, through a novel random matrix framework, the quantitative counterpart of these performance results, particularly in the case of echo-state networks. Beyond mere insights, our approach conveys a deeper understanding on the core mechanism under play for both training and testing.

C41. A. Kammoun, R. Couillet, F. Pascal, M. Slim-Alouini, "Optimal Design of Adaptive Normalized Matched Filter For Large Antenna Arrays", IEEE Statistical Signal Processing Workshop (SSP), Palma de Majorca, Spain, 2016.

Abstract. This paper focuses on the problem of detecting a target in the presence of a compound Gaussian clutter with unknown statistics. To this end, we focus on the design of the adaptive normalized matched filter (ANMF) detector which uses the regularized Tyler estimator (RTE) built from N-dimensional observations $x_1, \ldots, x_n$ in order to estimate the clutter covariance matrix. The choice for the RTE is motivated by its possessing two major attributes : first its resilience to the presence of outliers, and second its regularization parameter that makes it more suitable to handle the scarcity in observations. In order to facilitate the design of the ANMF detector, we consider the regime in which n and N are both large. This allows us to derive closed-form expressions for the asymptotic false alarm and detection probabilities. Based on these expressions, we propose an asymptotically optimal setting for the regularization parameter of the RTE that maximizes the asymptotic detection probability while keeping the asymptotic false alarm probability below a certain threshold. Numerical results are provided in order to illustrate the gain of the proposed detector over a recently proposed setting of the regularization parameter.

C42. N. Auguin, D. Morales, M. R. McKay, R. Couillet, "Robust Shrinkage M-estimators of Large Covariance Matrices", IEEE Statistical Signal Processing Workshop (SSP), Palma de Majorca, Spain, 2016.

Abstract. Robust high dimensional covariance estimators are considered, comprising regularized (linear shrinkage) modifications of Maronna's classical M-estimators. Such estimators aim to provide robustness to outliers, while simultaneously giving well-defined solutions under high dimensional scenarios where the number of samples does not exceed the number of variables. By applying tools from random matrix theory, we characterize the asymptotic performance of such estimators when the number of samples and variables grow large together. In particular, our results show that, when outliers are absent, many estimators of the shrinkage-Maronna type share the same asymptotic performance, and for such estimators we present a data-driven method for choosing the asymptotically optimal shrinkage parameter. Although our results assume an outlier-free scenario, simulations suggest that certain estimators perform substantially better than others when subjected to outlier samples.

C43. R. Couillet, G. Wainrib, H. Sevi, H. Tiomoko Ali, "Training performance of echo state neural networks", IEEE Statistical Signal Processing Workshop (SSP), Palma de Majorca, Spain, 2016.

Abstract. This article proposes a first theoretical performance analysis of the training phase of large dimensional linear echo-state networks. This analysis is based on advanced methods of random matrix theory. The results provide some new insights on the core features of such networks, thereby helping the practitioner when using them.

C44. H. Tiomoko Ali, R. Couillet, "Performance analysis of spectral community detection in realistic graph models", IEEE International Conference on Acoustics,

Speech and Signal Processing (ICASSP'16), Shangai, China, 2016.

Abstract. This article proposes a spectral analysis of dense random graphs generated by (a modified version of) the degree-corrected stochastic block model, for a setting where the inter block probabilities differ by $O(1/\sqrt{n})$ with n the number of nodes. We study a normalized version of the graph modularity matrix which is shown to be asymptotically well approximated by an analytically tractable (spiked) random matrix. The analysis of the latter allows for the precise evaluation of (i) the transition phase where clustering becomes asymptotically feasible and (ii) the alignment between the dominant eigenvectors and the block-wise canonical basis, thus enabling the estimation of misclassification rates (prior to post-processing) in simple scenarios.

C45. R. Couillet, F. Benaych-Georges, "Understanding Big Data Spectral Clustering", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, Mexico, 2015.

Abstract. This article introduces an original approach to understand the behavior of standard kernel spectral clustering algorithms (such as the Ng–Jordan–Weiss method) for large dimensional datasets. Precisely, using advanced methods from the field of random matrix theory and assuming Gaussian data vectors, we show that the Laplacian of the kernel matrix can asymptotically be well approximated by an analytically tractable equivalent random matrix. The analysis of the former allows one to understand deeply the mechanism into play and in particular the impact of the choice of the kernel function and some theoretical limits of the method. Despite our Gaussian assumption, we also observe that the predicted theoretical behavior is a close match to that experienced on real datasets (taken from the MNIST database).

C46. L. Yang, R. Couillet, M. R. McKay, "Minimum Variance Portfolio Optimization in the Spiked Covariance Model", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, Mexico, 2015.

Abstract. We study the design of minimum variance portfolio when asset returns follow a low rank factor model. Using results from random matrix theory, an optimal shrinkage approach for the isolated eigenvalues of the covariance matrix is developed. The proposed portfolio optimization strategy is shown to have good performance on synthetic data but not always on real data sets. This leads us to refine the data model by considering time correlation between samples. By updating the shrinkage of the isolated eigenvalues accounting for the unknown time correlation, our portfolio optimization method is shown to have improved performance and achieves lower risk values than competing methods on real financial data sets.

C47. L. Sanguinetti, R. Couillet, M. Debbah, "Base Station Cooperation for Power Minimization in the Downlink: Large System Analysis", IEEE Global Communications Conference (GLOBECOM'15), San Diego, USA, 2015.

Abstract. This work focuses on the downlink of a large-scale multi-cell multi-user MIMO system in which L base stations (BSs) of N antennas each communicate with KL single antenna user equipments. We consider the design of the optimal linear precoder for minimizing the total power consumption while ensuring target user rates. Two configurations with different degrees of cooperation among BSs are considered : the coordinated beamforming scheme (only channel state information is shared between BSs) and the network-MIMO technology (channel state and data cooperation). The analysis is conducted in the asymptotic regime where NL and KL grow large with a non trivial ratio K/N. In both configurations, tools of random matrix theory are used to compute, often in closed form, deterministic approximations for : (i) the Lagrange multipliers of the optimal precoder ; (ii) the powers needed to ensure target rates ; and (iii) the total transmit power. These

results are instrumental to get further insight into the structure of the optimal precoder and also to reduce the complexity of its implementation in large-scale networks. Numerical results are used to validate the asymptotic analysis in the finite system regime and to make comparisons among the two different configurations.

C48. R. Couillet, M. S. Greco, J-P. Ovarlez, F. Pascal, "RMT for Whitening Space Correlation and Applications to Radar Detection", IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Cancun, Mexico, 2015.

Abstract. Adaptive radar detection and estimation schemes are often based on the independence of the secondary data used for building estimators and detectors. This paper relaxes this constraint and deals with the non-trivial problem of deriving detection and estimation schemes for joint spatial and temporal correlated radar measurements. Latest results from Random Matrix theory, used for large dimensional regime, allows to build a Toeplitz estimate of the spatial covariance matrix while the temporal covariance matrix is then estimated in a conventional way (Sample Covariance Matrix, M-estimates). These two joint estimates of the spatial and temporal covariance matrices leads to build Adaptive Radar Detectors, like Adaptive Normalized Matched Filter (ANMF). We show that taking care of the spatial covariance matrix may lead to significant performance improvements compared to classical procedures.

C49. D. Morales-Jimenez, R. Couillet, M. McKay,"Large dimensional analysis of Maronna's M-estimator with outliers", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15), Brisbane, Australia, 2015.

Abstract. Building on recent results in the random matrix analysis of robust estimators of scatter, we show that a certain class of such estimators obtained from samples containing outliers behaves similar to a well-known random matrix model in the limiting regime where both the population and sample sizes grow to infinity at the same speed. This result allows us to understand the structure of such estimators when a certain fraction of the samples is corrupted by outliers and, in particular, to derive their asymptotic eigenvalue distributions. This analysis is a first step towards an improved usage of robust estimation methods under the presence of outliers when the number of independent observations is not too large compared to the size of the population.

C50. A. Kammoun, R. Couillet, F. Pascal, "Second order statistics of bilinear forms of robust scatter estimators", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15), Brisbane, Australia, 2015.

Abstract. This paper lies in the lineage of recent works studying the asymptotic behaviour of robust-scatter estimators in the case where the number of observations and the dimension of the population covariance matrix grow at infinity with the same pace. In particular, we analyze the fluctuations of bilinear forms of the robust schrinkage estimator of covariance matrix. We show that this result can be leveraged in order to improve the design of robust detection methods. As an example, we provide an improved generalized likelihood ratio based detector which combines robustness to impulsive observations and optimality across the shrinkage parameter.

C51. G. Katz, P. Piantanida, R. Couillet, "Joint Estimation and Detection Against Independence", Fifty-second Allerton Conference on Communication, Control, and Computing, Allerton, IL, USA, 2014.

Abstract. A receiver in a two-node system is required to make a decision of relevance as to received information, using side information that may or may not be

correlated with the received signal. In case the information is judged to be relevant, the receiver is then required to estimate the source with average distortion D. Focusing on the case of testing against independence, a single-letter expression for the rate-error-distortion region is proposed and proven. The resulting region ports a surprising resemblance to a seemingly non- associated classification problem, known as the information-bottleneck. The optimal region is then calculated for a binary symmetric example. Results demonstrate an interesting tradeoff between the achievable error-exponent for the decision and the distortion at the decoder.

C52. R. Couillet, M. McKay, "Robust covariance estimation and linear shrinkage in the large dimensional regime", IEEE International Workshop on Machine Learning for Signal Processing (MLSP'14), Reims, France, 2014.

Abstract. The article studies two regularized robust estimators of scatter matrices proposed in parallel in (Chen et al., 2011) and (Pascal et al., 2013), based on Tyler's robust M-estimator (Tyler, 1987) and on Ledoit and Wolf's shrinkage covariance matrix estimator (Ledoit and Wolf, 2004). These hybrid estimators convey robustness to outliers or impulsive samples and small sample size adequacy to the classical sample covariance matrix estimator. We consider here the case of i.i.d. elliptical zero mean samples in the regime where both sample and population sizes are large. We prove that the above estimators behave similar to well-understood random matrix models, which allows us to derive optimal shrinkage strategies to estimate the population scatter matrix, largely improving existing methods.

C53. L. Yang, R. Couillet, M. McKay, "Minimum variance portfolio optimization with robust shrinkage covariance estimation", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2014.

Abstract. We study the design of portfolios under a minimum risk criterion. The performance of the optimized portfolio relies on the accuracy to the estimated covariance matrix of portfolio asset returns. For large portfolios, the sample size is often of similar order to the number of assets, and the traditional sample covariance matrix performs poorly. Additionally, financial market data often involve outliers and exhibit heavy-tails, which, if not correctly handled, may further corrupt the covariance estimation. We aim to address these problems by studying the performance of a hybrid covariance matrix estimator based on Tyler's robust M-estimator and on Ledoit-Wolf's shrinkage estimator. Employing recent results from random matrix theory, we develop a consistent estimator of a scaled version of the portfolio risk, based on which, the shrinkage intensity is directly optimized to minimize the risk. Our portfolio optimization method is shown via simulations to outperform existing methods both for synthetic data and for a real market data set from Hang Seng Index.

C54. P. Vallet, X. Mestre, Ph. Loubaton, R. Couillet, "Asymptotic analysis of Beamspace-MUSIC in the context of large arrays", IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM'14), A Coruna, Spain, 2014.

Abstract. It is well-known that the MUSIC method for DoA estimation degrades when the number of samples N and the array dimension M are large and of the same order of magnitude. In this context, several improvements have been proposed, among which the G-MUSIC method, which was shown to be consistent in the asymptotic regime where M, N converge to infinity at the same rate, and under an additional separation condition between noise and signal subspaces of the SCM. Nevertheless, this subspace separation condition is only fulfilled for sufficiently high SNR. Dimension reduction techniques are a classical way to partially circumvent this condition. In this paper, we provide an asymptotic analysis in terms of consistency and MSE in the aforementioned regime, of the Beamspace MUSIC, which is one popular technique to reduce the dimension of the observations.

C55. R. Couillet, A. Kammoun, "Robust G-MUSIC", European Signal Processing Conference (EUSIPCO'14), Lisbon, Portugal, 2014.

Abstract. An improved MUSIC algorithm for direction-of-arrival estimation is introduced that accounts both for large array sizes N comparatively with the number of independent observations n and for the impulsiveness of the background environment (e.g., presence of outliers in the observations). This method derives from the spiked G-MUSIC algorithm proposed in (Vallet et al., 2011) and from the recent works by one of the authors on the random matrix analysis of robust scatter matrix estimators (Couillet et al., 2013). The method is shown to be asymptotically consistent where classical approaches are not. This superiority is corroborated by simulations.

C56. J. Vinogradova, R. Couillet, W. Hachem, "Estimation of Large Toeplitz Covariance Matrices and Application to Source Detection", European Signal Processing Conference (EUSIPCO'14), Lisbon, Portugal, 2014.

Abstract. In this paper, performance results of two types of Toeplitz covariance matrix estimators are provided. Concentration inequalities for the spectral norm for both estimators have been derived showing exponential convergence of the error to zero. It is shown that the same rates of convergence are obtained in the case where the aggregated matrix of time samples is corrupted by a rank one matrix. As an application based on this model, source detection by a large dimensional sensor array with temporally correlated noise is studied.

C57. R. Couillet, F. Pascal, "Robust M-estimator of scatter for large elliptical samples", IEEE Worshop on Statistical Signal Processing (SSP'14), Gold Coast, Australia, 2014.

Abstract. It is shown that a certain family of robust scatter estimators of elliptical samples behaves similar to a well-known random matrix model in the limiting regime where both the population N and sample n sizes grow to infinity at the same speed. This result allows us to understand the structure of such estimators and in particular to derive their limiting eigenvalue distributions. This analysis is a first step towards an improved usage of robust estimation methods when the number of independent observations is not too large compared to the size of the population.

C58. A. Abboud, R. Couillet, M. Debbah, H. Siguerdidjane, "Asynchronous alternating direction method of multipliers applied to the direct-current optimal power flow problem," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'14), Florence, Italy, 2014.

Abstract. We consider a distributed convex optimization problem where each agent has its private convex cost function and controls a set of local variables. We provide an algorithm to carry out a multi-area decentralized optimization in an asynchronous fashion, obtained by applying random Gauss-Seidel iterations on the Douglas-Rachford splitting operator. As an application, a Direct-current linear optimal power flow model is implemented and simulations confirm the convergence of the proposed algorithm.

C59. A. Pelletier, R. Couillet, J. Najim, "Second-Order Analysis of the Joint SINR distribution in Rayleigh Multiple Access and Broadcast Channels," Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2013.

Abstract. This article studies the joint distribution of the signal-to-interference-plus-noise ratios (SINR) of the users in Rayleigh multiple access channels and broadcast channels, using large dimensional random matrix theory. Two models are studied : a multiple access channel (MAC) with minimum mean square error (MMSE) decoding, and a broadcast channel with regularized zero-forcing (RZF) precoding. It

is shown that, in both models, the empirical distribution of the SINRs of the users behaves asymptotically as a Gaussian, with identified mean and variance. The result is applied to the estimation of the proportion of users in outage for a given target rate. This asymptotic Gaussian behavior can be derived from a theoretical approach based on Stein's method in a random matrix theory context.

C60. A. Müller, E. Björnson, R. Couillet, M. Debbah, "Analysis and management of heterogeneous user mobility in large-scale downlink systems," Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2013.

Abstract. Modern cellular networks need to serve user terminals with large disparities in mobility, which incurs different accuracy of the channel state information for each user. The impact of such heterogeneous mobility on the multi-cell downlink is analyzed in this paper. The base stations serve a multitude of users by coordinated beamforming. We derive deterministic equivalents for the user performance in a large scale system where the number of transmit antennas and user terminals grow large at a fixed ratio. We show that low and high mobility users can coexist and be served simultaneously, since the CSI imperfections of a user only harms the performance of this particular user. Simulations are used to verify the applicability of our large scale approximations for systems of practical dimensions. Furthermore, we show that the performance of high mobility users can be improved by explicitly managing the user priorities in the network.

C61. J. Hoydis, R. Couillet, P. Piantanida, "Bounds on the Second-Order Coding Rate of the MIMO Rayleigh Block-Fading Channel," IEEE International Symposium on Information Theory, Istanbul, Turkey, 2013.

Abstract. We study the second-order coding rate of the multiple-input multiple-output (MIMO) Rayleigh block-fading channel via statistical bounds from information spectrum methods and random matrix theory. Based on an asymptotic analysis of the mutual information density which considers the simultaneous growth of the block length n and the number of transmit and receive antennas K and N, we derive closed-form upper and lower bounds on the optimal average error probability when the code rate is within $O(1/\sqrt{nK})$ of the asymptotic capacity. A Gaussian approximation is then used to establish an upper bound on the error probability for arbitrary code rates which is shown by simulations to be accurate for small N, K, and n.

C62. G. Geraci, R. Couillet, J. Yuan, M. Debbah, I. Collings, "Secrecy Sum-Rates with Regularized Channel Inversion Precoding under Imperfect CSI at the Transmitter," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), Vancouver, Canada, 2013.

Abstract. In this paper, we study the performance of regularized channel inversion precoding in MISO broadcast channels with confidential messages under imperfect channel state information at the transmitter (CSIT). We obtain an approximation for the achievable secrecy sum-rate which is almost surely exact as the number of transmit antennas and the number of users grow to infinity in a fixed ratio. Simulations prove this anaylsis accurate even for finite-size systems. For FDD systems, we determine how the CSIT error must scale with the SNR, and we derive the number of feedback bits required to ensure a constant high-SNR rate gap to the case with perfect CSIT. For TDD systems, we study the optimum amount of channel training that maximizes the high-SNR secrecy sum-rate.

C63. R. Couillet, F. Pascal, J. W. Silverstein, "A Joint Robust Estimation and Random Matrix Framework with Application to Array Processing," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), Vancouver, Canada, 2013.

Abstract. An original interface between robust estimation theory and random matrix theory for the estimation of population covariance matrices is proposed. Consider a random vector $x = A_N y \in \mathbb{C}^N$ with $y \in \mathbb{C}^M$ made of $M \geq N$ independent entries, $\mathbb{E}[y] = 0$, and $\mathbb{E}[yy^*] = I_N$. It is shown that a class of robust estimators $\hat{C}_N$ of $C_N = A_N A_N^*$, obtained from n independent copies of x, is (N, n)-consistent with the traditional sample covariance matrix $\hat{S}_N$ in the sense that $\|\hat{C}_N - \alpha \hat{S}_N\| \to 0$ in spectral norm for some $\alpha > 0$, almost surely, as $N, n \to \infty$ with N/n and M/N bounded. This result, in general not valid in the fixed N regime, is used to propose improved subspace estimation techniques, among which an enhanced direction-of-arrival estimator called robust G-MUSIC.

C64. J. Vinogradova, R. Couillet, W. Hachem, "A new method for source detection, power estimation, and localization in large sensor networks under noise with unknown statistics," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), Vancouver, Canada, 2013.

Abstract. Most statistical inference methods for array processing assume an array of size N fixed and a number of snapshots T large. In addition, many works are based on the assumption of a white noise model. These two assumptions are increasingly less realistic in modern systems where N and T are usually both large, and where the noise data can be correlated either across successive observations or across the sensor antennas. In this paper an approach to handle this kind of scenario is presented. New algorithms for source number estimation, power estimation, and localization by a sensor array under noise with unknown correlation model are proposed. The results fundamentally rely on recent advances in small rank perturbations of large dimensional random matrices.

C65. M. de Mari, R. Couillet, M. Debbah, "Concurrent data transmissions in green wireless networks: when best send one's packets?," (Invited paper) IEEE International Symposium on Wireless Communication Systems (ISWCS'12), Paris, France, 2012.

Abstract. In this paper, we consider the scenario of a cellular network where base stations aim to transmit several data packets to a set of users in the downlink, within a predefined time, at minimal energy cost. The base stations are non-cooperating and the instantaneous transmission rate depends on the instantaneous SINR at the receiver. The purpose of this article is to highlight a power-efficient transmit policy. By assuming a large number of homogeneous users, we model the problem as a mean field game, with tractable equations, that allow us to bypass the complexity of analyzing a Nash equilibrium in a L-body dynamic game. The framework we propose yields a consistent analysis of the optimal transmit power strategy, that allows every base station to, selfishly but rationally, satisfy its transmission, at a minimal energy cost.

C66. A. Müller, J. Hoydis, R. Couillet, M. Debbah, "Optimal 3D Cell Planning: A Random Matrix Approach," IEEE Global Communications Conference (GLOBECOM'12), Anaheim, Caligornia, USA, 2012.

Abstract. This article proposes a large system approximation of the ergodic sum-rate (SR) for cellular multi-user multiple-input multiple-output uplink systems. The considered system has various degrees of freedom, such as clusters of base stations (BSs) performing cooperative multi-point processing, randomly distributed user terminals (UTs), and supports arbitrarily configurable antenna gain patterns at the BSs. The approximation is provably tight in the limiting case of a large number of single antenna UTs and antennas at the BSs. Simulation results suggest that the asymptotic analysis is accurate for small system dimensions. Our deterministic SR approximation result is applied to numerically study and optimize the effects of

antenna tilting in an exemplary sectorized 3D small cell network topology. Significant SR gains are observed with optimal tilt angles and we provide new insights on the optimal parameterization of cellular networks, along with a discussion of several non-trivial effects.

C67. J. Hoydis, R. Couillet, P. Piantanida, M. Debbah, "A Random Matrix Approach to the Finite Blocklength Regime of MIMO Fading Channels," IEEE International Symposium on Information Theory, Boston, Massachusetts, USA, 2012.

Abstract. This paper provides a novel central limit theorem (CLT) for the information density of the MIMO Rayleigh fading channel under white Gaussian inputs, when the data blocklength n and the number of transmit and receive antennas K, N are large but of similar order of magnitude. This CLT is used to derive closed-form upper bounds on the error probability via Feinstein's lemma and the second-order approximation of the coding rate. Numerical evaluations suggests that the normal approximation is tight for reasonably small values of n, K, N.

C68. M. Rezaee, R. Couillet, M. Guillaud, G. Matz, "Sum-Rate Optimization for the MIMO IC under Imperfect CSI: a Deterministic Equivalent Approach," IEEE International Workshop on Signal Processing Advances for Wireless Communications, Cesme, Turkey, 2012.

Abstract. A new method is proposed to determine precoding matrices that achieve local maxima of the expected sum rate in a multiple input multiple output interference channel (MIMO IC), in the realistic scenario where only partial channel state information (CSI) is available at the transmitters. Relying on a random matrix analysis of the capacity of large dimensional Ricean channels, the expected sum rate of the K-user MIMO IC is approximated by a deterministic equivalent to which an iterative gradient scheme is applied to find local maxima of the approximated sum rate.

C69. J. Hoydis, A. Müller, R. Couillet, M. Debbah, "Analysis of Multicell Cooperation with Random User Locations Via Deterministic Equivalents," Eighth Workshop on Spatial Stochastic Models for Wireless Networks, Paderborn, Germany, 2012.

Abstract. We consider the uplink of a one-dimensional 2-cell network with fixed base stations (BSs) and randomly distributed user terminals (UTs). Assuming that the number of antennas per BS and the number of UTs grow infinitely large, we derive tight approximations of the ergodic sum rate with and without multicell processing for optimal and sub-optimal detectors. We use these results to find the optimal BS placement to maximize the system capacity. This work can be seen as a first attempt to apply large random matrix theory to the study of networks with random topologies. We demonstrate that such an approach is feasible and leads to analytically tractable expressions of the average system performance. Moreover, these results can be used to optimize certain system parameters for a given distribution of user terminals and to assess the gains of multicell cooperation.

C70. R. Couillet, E. Zio, "A subspace approach to fault diagnostics in large power systems" (Invited Paper) IEEE International Symposium on Communications, Control, and Signal Processing (ISCCSP'12), Rome, Italy, 2012.

Abstract. In this article, a recently proposed subspace approach for diagnosing sudden local changes in large dynamical networks is applied to the detection and localization of link failures in power systems, on the basis of nodal voltage measurements.

C71. A. Kammoun, M. Kharouf, R. Couillet, J. Najim, M. Debbah, "On the fluctuations of the SINR at the output of the Wiener filter for non centered channels: the non

Gaussian case," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'12), Kyoto, Japan, 2012.

Abstract. In the context of multidimensional signals, the linear Wiener receiver is frequently encountered in wireless communication and in array processing ; it is in fact the linear receiver that achieves the lowest level of interference. In this contribution, we focus on the study of the associated Signal-to-interference plus noise ratio (SINR) at its output in the context of Ricean multiple-input multiple-output (MIMO) channels. The case of Ricean channels, which induces non-centered random variables, can be encountered in several practical environments and has not been studied so far, as it raises substantial technical issues. With the help of large random matrix theory, which has shown to be fruitful to successfully address several problems in wireless communications, we study the behaviour of the SINR, together with its fluctuations via a central limit theorem. As realistic models also involve non-Gaussian random variables, we relax the Gaussian assumption. This results in an extra term involving the fourth cumulant in the expression of the variance.

C72. R. Couillet, P. Bianchi, J. Jakubowicz, "Decentralized convex stochastic optimization with few constraints in large networks," IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'11), San Juan, Puerto Rico, 2011.

Abstract. This article introduces a distributed convex opti- mization algorithm in a constrained multi-agent system composed by a large number of nodes. We focus on the case where each agent seeks to optimize its own local parameter under few coupling equality and inequality constraints. The objective function is of the power flow type and can be decoupled as a sum of elementary functions, each of which assumed (imperfectly) known by only one node. Under these assumptions, a cost-efficient decentralized iterative solution based on Lagrangian duality is derived, which is provably converging. This new approach alleviates several limitations of algorithms proposed in the stochastic optimization literature. Applications are proposed to decentralized power flow optimization in smart grids.

C73. R. Couillet, S. Medina Perlaza, H. Tembine, M. Debbah, "A mean field game analysis of electric vehicles in the smart grid," IEEE International Conference on Computer Communications (INFOCOM'12), Orlando, FL, USA, 2012.

Abstract. In this article, we develop a mean field game model for the economical analysis of the integration of purely electrical vehicles (EV) or electrical hybrid oil-electricity vehicles (PHEV) in the smart grid energy market. The framework we develop allows for a consistent analysis of the evolution of the price of electricity, of the timely demand, and possibly of the energy reserves in the grid, when EV or PHEV owners buy and sell electricity from their cars, selfishly but rationally, based on collective price incentives.

C74. J. Hoydis, R. Couillet, M. Debbah, "Asymptotic Analysis of Double-Scattering Channels," IEEE Asilomar Conference (ASILOMAR'11), Pacific Grove, CA, USA, 2011. **Best student paper award finalist**

Abstract. We consider a multiple-input multiple-output (MIMO) multiple access channel (MAC), where the channel between each transmitter and the receiver is modeled by the doubly-scattering channel model. Based on novel techniques from random matrix theory, we derive deterministic approximations of the mutual information, the signal-to-noise-plus-interference-ratio (SINR) at the output of the minimum-mean-square-error (MMSE) detector and the sum-rate with MMSE detection, which are almost surely tight in the large system limit. Moreover, we derive the asymptotically optimal transmit covariance matrices. Our simulation results show that the asymptotic analysis provides very close approximations for realistic

system dimensions.

C75. R. Couillet, W. Hachem, "Local Failure Localization in Large Sensor Networks," IEEE Asilomar Conference on Signals, Systems, and Computers (ASILOMAR'11), Pacific Grove, CA, USA, 2011.

Abstract. In this article, the joint fluctuations of the extreme eigenvalues and eigenvectors of a large dimensional sample covariance matrix are analyzed when the associated population covariance is a finite-rank perturbation of the identity matrix. It is shown that these fluctuations are asymptotically normal with zero mean and a variance which is derived explicitly. This result is used in practice to develop an original framework for local failure localization in large sensor networks, among which sudden parameter changes.

C76. R. Couillet, M. Guillaud, "Performance of Statistical Inference Methods for the Energy Estimation of Multiple Sources," (Invited Paper) IEEE Statistical Signal Processing Workshop (SSP'11), Nice, France, 2011.

Abstract. This article considers statistical inference of the transmit powers of multiple signal sources by a sensor network. Using random matrix methods, an exact expression of the posterior probability of the joint transmit powers is derived. This expression is used to implement the associated ML and MMSE detectors of the joint powers. These are compared for small system sizes against an asymptotically unbiased estimator obtained from large dimensional random matrix theory.

C77. A. Kammoun, R. Couillet, J. Najim, M. Debbah, "Performance of fast rate adaption techniques in interference-limited networks," IEEE Global Communications Conference (GLOBECOM'11), Houston, TX, USA, 2011.

Abstract. In a companion paper, a fast estimator for the capacity of a secondary communication in the context of cognitive radio networks was proposed. It was particularly shown that the proposed method largely outperforms traditional ones when the available number of samples is limited. In this paper, we study the fluctuations of the aforementioned estimators around their deterministic equivalents. We prove that in the asymptotic regime, their behaviors can be approximated by Gaussian random variables for which we derive the variances.

C78. J. Yao, R. Couillet, J. Najim, E. Moulines, M. Debbah, "CLT for eigen-inference methods in cognitive radios," IEEE International Conference on Acoustics, Speech and Signal Processing, Prague, Czech Republic, 2011.

Abstract. This article provides a central limit theorem for a consistent estimator of the population eigenvalues of a class of sample covariance matrices. An exact expression as well as an empirical and asymptotically accurate approximation of the limiting variance is also derived. These results are applied in a cognitive radio context featuring an orthogonal-CDMA primary network and a secondary network whose objective is to maximise the coverage of secondary transmissions under low probability of interference with primary users.

C79. J. Hoydis, R. Couillet, M. Debbah, "Deterministic Equivalents for the Performance Analysis of Isometric Random Precoded Systems," IEEE International Conference on Communications, Kyoto, Japan, 2011.

Abstract. We consider a general wireless channel model for different types of code-division multiple access (CDMA) and space-division multiple-access (SDMA) systems with isometric random signature/precoding matrices over frequency-selective and flat fading channels. We derive deterministic approximations of the Stieltjes transform, the mutual information and the signal-to-interference-plus-noise ratio (SINR) at the output of the minimum-mean-square-error (MMSE) receiver and provide a simple fixed-point algorithm for their computation, which is proved to

converge. The deterministic approximations are asymptotically tight, almost surely, but shown by simulations to be very accurate for even small system dimensions. Our analysis requires neither arguments from free probability theory nor the asymptotic freeness or the convergence of the spectral distribution of the involved matrices. The results presented in this work are, therefore, also a novel contribution to the field of random matrix theory and might be useful to further applications involving isometric random matrices.

C80. J. Hoydis, J. Najim, R. Couillet, M. Debbah, "Fluctuations of the Mutual Information in Large Distributed Antenna Systems with Colored Noise," Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing, Allerton, IL, USA, 2010.

Abstract. This paper studies the fluctuations of the mutual information of a class of multiple-input multiple-output (MIMO) channels with arbitrary correlated noise in the large system limit. Under the assumption that the channel dimensions grow infinitely large at the same rate, we find a deterministic approximation of the ergodic mutual information and study its fluctuations around this value in form of a central limit theorem (CLT). This result can be used to predict the outage probability for slow fading channels. The channel model considered in this contribution has a particular application in the context of distributed antenna or network MIMO systems where the path loss between any pair of transmit and receive antennas has a different value. As shown by simulations, the asymptotic approximations translate well into systems of small dimensions.

C81. R. Couillet, H. V. Poor, M. Debbah, "Self-organized spectrum sharing in large MIMO multiple-access channels," IEEE International Symposium on Information Theory, Austin TX, USA, 2010.

Abstract. In this paper, a deterministic approximation for the rate region of multiple access channels is provided when base station and users have a large number of antennas, and when the transmission bandwidth is divided into several independent subbands. An explicit formulation is also given to the transmit covariance matrices, at each frequency, that reach the boundary of the rate region. From the compact expression of these matrices, suboptimal iterative algorithms emerge that allow the multiple access users to derive autonomously the transmit covariance matrices. This comes at the sole expense of a short signalling overhead, which is constant irrespectively of the number of antennas. Simulations confirm the validity of the theoretical derivations and suggest rather good behaviour obtained by the suboptimal self-organization algorithms.

C82. R. Couillet, J. W. Silverstein, M. Debbah, "Eigen-inference for multi-source power estimation," IEEE International Symposium on Information Theory, Austin TX, USA, 2010.

Abstract. This paper introduces a new method to estimate the power transmitted by multiple signal sources, when the number of sensing devices and the available samples are sufficiently large compared to the number of sources. This work makes use of recent advances in the field of random matrix theory that prove more efficient than previous "moment-based" approaches to the problem of multi-source power detection. Simulations are performed which corroborate the theoretical claims.

C83. S. Wagner, R. Couillet, D. T. M. Slock, M. Debbah, "Optimal Training in Large TDD Multi-user Downlink Systems under Zero-forcing and Regularized Zero-forcing Precoding," IEEE Global Communication Conference, Miami, 2010.

Abstract. This paper considers a large multi-user time-division duplex (TDD) system, where the base station (BS) acquires channel state information via pilot signaling from the users. In the downlink the BS employs zero-forcing (ZF) and regu-

larized zero-forcing (RZF) precoding. We derive the optimal sum rate maximizing amount of channel training using sum rate approximations from the large system analysis of MISO downlink channels under (R)ZF precoding. Moreover, in the regime of high signal-to-noise ratio (SNR), we derive approximate solutions of the optimal amount of training for both schemes that are of closed-form. By comparing the two schemes, we find that RZF requires less training than ZF, but the training interval of both schemes is equal for asymptotically high SNR. Furthermore, simulations are carried out which demonstrate the accuracy of our approximate solutions.

C84. S. Wagner, R. Couillet, D. T. M. Slock, M. Debbah, "Large System Analysis of Zero-Forcing Precoding in MISO Broadcast Channels with Limited Feedback" IEEE International Workshop on Signal Processing Advances for Wireless Communications, Marrakech, Morocco, 2010.

Abstract. In this paper we analyze the sum-rate of zero-forcing (ZF) precoding in MISO broadcast channels with limited feedback, transmit correlation and path loss. Our analysis assumes that the nimber of transmit antennas M and the number of users K are large, while their ratio remains bounded. By applying recent results from random matrix theory we derive a deterministic equivalent of the SINR and compute the sum-rate maximizing number of users as well as the limiting sum-rate for high SNR, as a function of the channel uncertainties and the channel correlation pattern. Simulations show that theoretical and numerical results match well, even for small system dimensions.

C85. R. Couillet, M. Debbah, "Information theoretic approach to synchronization: the OFDM carrier frequency offset example", Advanced International Conference on Telecommunications, Barcelona, Spain, 2010.

Abstract. In practical mobile communication systems, data-aided synchronization is performed before actual data exchanges, and synchronization methods are tailored to the pilot sequence. In this paper, we propose a framework for initial synchronization that works independently of the pilot sequence. We show that classical data-aided and blind techniques are particular cases of this general framework. We thoroughly study the specific problem of OFDM data-aided carrier frequency offset estimation. For the latter, we provide theoretical expressions of performance upper-bounds. Also, a practical thin CFO estimator for OFDM is provided under the form of a novel algorithm which is shown by simulations to perform better than classical pilot based methods.

C86. R. Couillet, M. Debbah, "Uplink capacity of self-organizing clustered orthogonal CDMA networks in flat fading channels" IEEE Information Theory Workshop Fall'09, Taormina, Sicily, 2009.

Abstract. In this paper, we derive a deterministic equivalent of the Shannon-transform of certain type of large unitary random matrices. This approximation is exploited to evaluate the uplink channel capacity of clustered orthogonal CDMA network. When non-uniform power allocation among the users of each cluster is allowed, we derive an explicit iterative waterfilling algorithm which, upon convergence, achieves the multiuser decoding capacity. In particular, we show that, in a self-organizing clustered orthogonal CDMA network, each cluster can optimize its power allocation policy independently of the other clusters at the expense of a small feedback overhead. Simulations corroborate the theoretical derivations.

C87. R. Couillet, M. Debbah, J. W. Silverstein, "Asymptotic Capacity of Multi-User MIMO Communications" IEEE Information Theory Workshop Fall'09, Taormina, Sicily, 2009.

Abstract. This paper introduces a new formula to derive explicit capacity expressions of a class of communication schemes, including single-cell multi-user MIMO and multi-cell point-topoint MIMO, when the wireless channels have separable variance profiles and the system dimensions grow large. As an introductory example, we study point-to-point MIMO channels with multicell interference, in downlink. In this setting, we provide new asymptotic capacity expressions when single-user decoding or MMSE decoding are used. Simulations are shown to corroborate the theoretical claims, even when the number of transmit/receive antennas is not very large.

C88. R. Couillet, M. Debbah, J. W. Silverstein, "Rate region of correlated MIMO multiple access channel and broadcast channel" IEEE Workshop on Statistical Signal Processing, Cardiff, Wales, UK, 2009.

Abstract. In this paper, the rate region of large multi-antenna multiple access channels and broadcast channels are investigated. The propagation channels between transmitters and receivers are modelled as independent Gaussian with separable variance profiles. It is shown in particular that the large antenna rate regions do not depend on the specific channel realization, but only on the channel transmit and receive covariance matrices. The theoretical results are corroborated by simulations.

C89. R. Couillet, M. Debbah, "Mathematical foundations of cognitive radios" U.R.S.I.'09, Warsaw, Poland, 2009.

Abstract. Recently, much interest has been directed towards software defined radios and embedded intelligence in telecommunication devices. However, no theoretical framework for cognitive radios has ever been proposed. In this paper, we introduce an information theoretic point of view on cognitive radios. Specifically, our motivation in this work is to embed humanlike intelligence in mobile wireless devices, following the three century-old work on Bayesian probability theory, the maximum entropy principle and minimal probability update. This allows us to partially answer such questions as "what are the signal detection capabilities of a wireless device ?" or "when facing a situation in which most parameters are missing, how to react ?". As an introduction, we will present two examples from the same authors using the cognitive framework namely multi-antenna channel modelling and signal sensing.

C90. R. Couillet, M. Debbah, "A maximum entropy approach to OFDM channel estimation", IEEE International Workshop on Signal Processing Advances for Wireless Communications, Perugia, Italy, 2009.

Abstract. In this work, a Bayesian framework for OFDM channel estimation is proposed. Using the maximum entropy principle to exploit prior system information at the receiver, we successively derive channel estimates in situations when (i) the channel delay spread and (ii) the channel time correlation statistics are a priori unknown. More generally, this framework allows to derive MMSE channel estimates under any state of knowledge at the receiver. Simulations are provided that confirm the theoretical claims and show the novel results to perform as good or better than classical estimators.

C91. R. Couillet, M. Debbah, "Bayesian inference for multiple antenna cognitive receivers", IEEE Wireless Communications & Networking Conference, Budapest, Hungary, 2009.

Abstract. A Bayesian inference learning process for cognitive receivers is provided in this paper. We focus on the particular case of signal detection as an explanatory example to the learning framework. Under any prior state of knowledge on the communication channel, an information theoretic criterion is presented to decide

on the presence of informative data in a noisy wireless MIMO communication. We detail the particular cases of knowledge, or absence of knowledge at the receiver, of the number of transmit antennas and noise power. The provided method is instrumental to provide intelligence to the receiver and gives birth to a novel Bayesian signal detector. The detector is compared to the classical power detector and provides detection performance upper bounds. Simulations corroborate the theoretical results and quantify the gain achieved using the proposed Bayesian framework.

C92. R. Couillet, M. Debbah, "Flexible OFDM schemes for bursty transmissions", IEEE Wireless Communications & Networking Conference, Budapest, Hungary, 2009.

Abstract. In this paper, -OFDM, a generalization of the OFDM modulation, is proposed to enhance the outage capacity of bursty transmissions. This new flexible modulation scheme is easily implemented and only requires a symbol rotation of angle after the IDFT stage. The induced rotation slides the DFT window and provides frequency diversity in block fading channels. Interestingly, the results show a substantial gain in terms of outage capacity and BER in comparison with classical OFDM modulation schemes. The framework is extended to multiuser/multi-antenna OFDM based standards. Simulations, in the context of 3GPP LTE, called hereafter -LTE, sustain our theoretical claims.

C93. R. Couillet, S. Wagner, M. Debbah, "Asymptotic Analysis of Correlated Multi-Antenna Broadcast Channels", IEEE Wireless Communications & Networking Conference, Budapest, Hungary, 2009.

Abstract. In this paper we consider the MIMO broadcast channel with antenna correlation at the transmitter and receiver. We derive the theoretical sum rate of systems with a large number of antennas for zero-forcing and regularized zero-forcing precoders. Particularly, we apply the results to volume-limited devices where the correlation originates from a dense antenna packing. Throughout this contribution we make extensive use of recent tools from random matrix theory. Simulations confirm the theoretical claims and also indicate that in most scenarios the asymptotic derivations applied to a finite number of users give good approximations of the true ergodic sum rate.

C94. R. Couillet, S. Wagner, M. Debbah, A. Silva, "The Space Frontier: Physical Limits of Multiple Antenna Information Transfer", ValueTools, Inter-Perf Workshop, Athens, Greece, 2008. **Best student paper award**

Abstract. In this paper, we study the capacity limits of dense multiantenna systems. We derive asymptotic capacity expressions for point-to-point and broadcast channels by applying recent tools from random matrix theory. In the case of broadcast channels, we focus on linear precoding techniques. We found that the asymptotic capacity depends only on the ratio between the size of the antenna array and the wavelength. This provides useful guidelines on the achievable limits of information transfer. In particular, it is shown that the total capacity grows unbounded if the transmitter has perfect knowledge on the channel, while the capacity saturates in the absence of channel knowledge at the transmitter. We provide numerical results supporting the theoretical derivations.

C95. R. Couillet, M. Debbah, "Free deconvolution for OFDM multicell SNR detection", IEEE Personal, Indoor and Mobile Radio Communications Symposium, Cognitive Radio Workshop, Cannes, France, 2008.

Abstract. In this paper, a new blind OFDM multicell detection method is proposed to determine the number of base stations in a cellular system. Using recent results of free deconvolution, the algorithm enables the terminal to count the number of surrounding base stations as well as the received power using only a limited number of snapshots. This is in sharp contrast with classical asymptotic blind techniques

and a theoretical analysis is proposed to study the impact of frequency selectivity and the number of receive/transmit antennas. Simulations are provided to sustain the theoretical claims and comparisons are provided with classical techniques.

Books and book chapters

B1. R. Couillet, M. Debbah, **Random Matrix Methods for Wireless Communications**, Cambridge University Press, 2011. [book]

Abstract. Blending theoretical results with practical applications, this book provides an introduction to random matrix theory and shows how it can be used to tackle a variety of problems in wireless communications. The Stieltjes transform method, free probability theory, combinatoric approaches, deterministic equivalents and spectral analysis methods for statistical inference are all covered from a unique engineering perspective. Detailed mathematical derivations are presented throughout, with thorough explanation of the key results and all fundamental lemmas required for the reader to derive similar calculus on their own. These core theoretical concepts are then applied to a wide range of real-world problems in signal processing and wireless communications, including performance analysis of CDMA, MIMO and multi-cell networks, as well as signal detection and estimation in cognitive radio networks. The rigorous yet intuitive style helps demonstrate to students and researchers alike how to choose the correct approach for obtaining mathematically accurate results.

B2. R. Couillet, M. Debbah, **Mathematical Foundations for Signal Processing, Communications and Networking**, CRC Press, Taylor & Francis Group, 2011 [book chapter]

Abstract. From basic transforms to Monte Carlo simulation to linear programming, the text covers a broad range of mathematical techniques essential to understanding the concepts and results in signal processing, telecommunications, and networking. Along with discussing mathematical theory, each self-contained chapter presents examples that illustrate the use of various mathematical concepts to solve different applications. Each chapter also includes a set of homework exercises and readings for additional study.

Chapter "Random matrix theory".

B3. R. Couillet, M. Debbah, **Orthogonal Frequency Division Multiple Access Fundamentals and Applications**, Auerbach Publications, CRC Press, Taylor & Francis Group, 2010 [book chapter]

Abstract. Supported by the expert-level advice of pioneering researchers, Orthogonal Frequency Division Multiple Access Fundamentals and Applications provides a comprehensive and accessible introduction to the foundations and applications of one of the most promising access technologies for current and future wireless networks. It includes authoritative coverage of the history, fundamental principles, key techniques, and critical design issues of OFDM systems.

Chapter "Fundamentals of OFDMA Synchronization".

B4. R. Couillet, M. Debbah, **Radio engineering : From software radio to cognitive radio**, John Wiley & Sons, 2013 [book chapter]

Abstract. This collective work provides engineers, researchers and radio designers with the necessary information from mathematical analysis and hardware architectures to design methodology and tools, running platforms and standardization in order to understand this new cognitive radio domain.

Several chapters.

Patents and Innovative Ideas

P1. R. Couillet, M. Debbah, **No. 08368028.0** "Process and apparatus for performing initial carrier frequency offset in an OFDM communication system"

P2. R. Couillet, M. Debbah, **No. 08368023.1** "Method for short-time OFDM transmission and apparatus for performing flexible OFDM modulation"

P3. R. Couillet, S. Wagner, **No. 09368025.4** "Precoding process for a transmitter of a MU-MIMO communication system"

P4. R. Couillet, **No. 09368030.4** "Process for estimating the channel in an OFDM communication system, and receiver for doing the same"

II1. R. Couillet, **Innovative Idea** "Fast Block Diagonalization Precoder"

II2. R. Couillet, **Innovative Idea** "User Subspace Clustering"