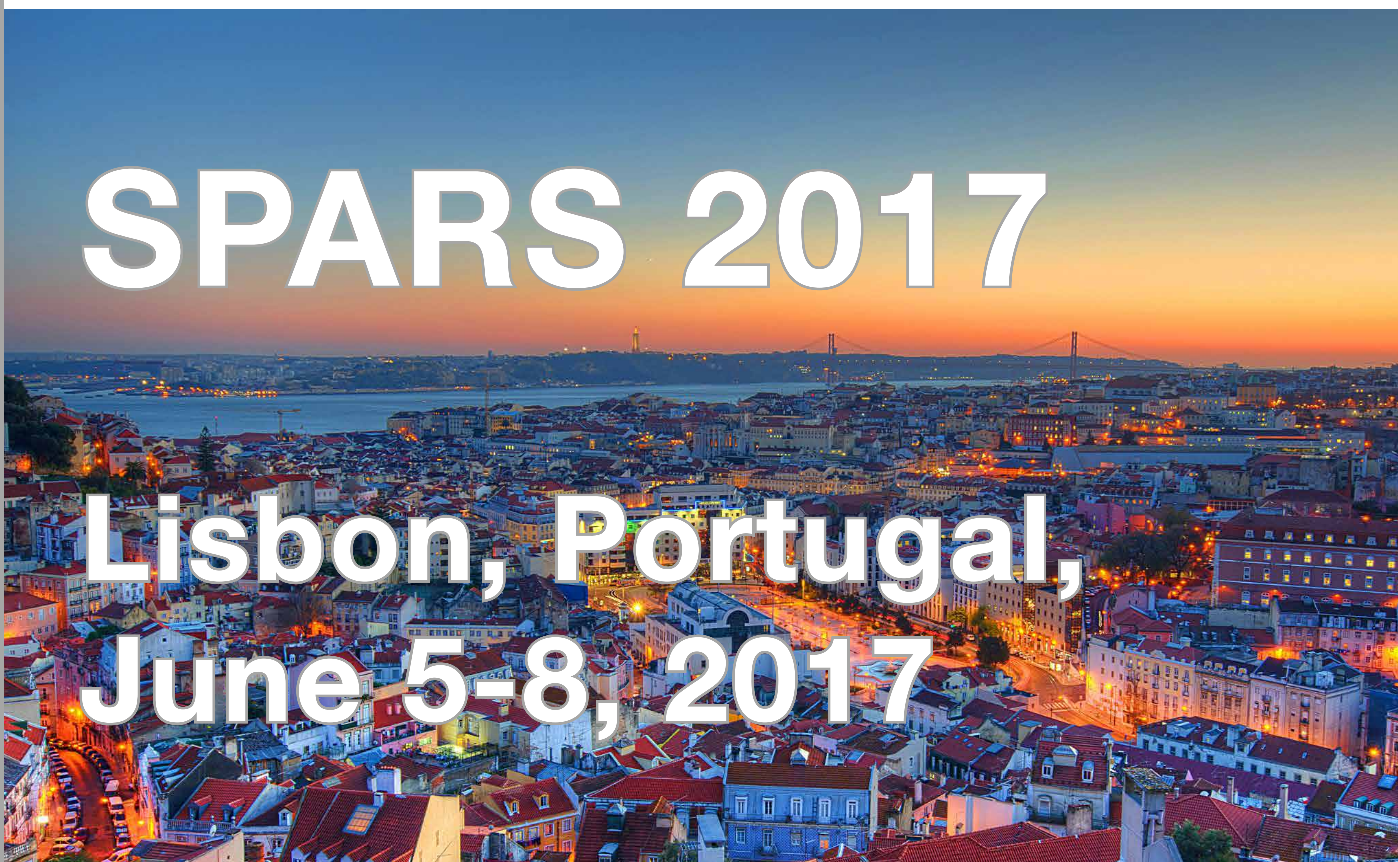
SPARS 2017

Book of Abstracts



SPARS 2017

Lisbon, Portugal, June 5-8, 2017

Program Committee

Francis Bach, *INRIA and École Normale Supérieure, France.*
Laura Balzano, *University of Michigan, USA.*
José Bioucas-Dias, *University of Lisbon, Portugal.*
Laure Blanc-Féraud, *CNRS and INRIA, France.*
Laurent Daudet, *Université Paris Diderot - Paris 7, France.*
Mike Davies, *University of Edinburgh, UK.*
Michael Elad, *Israel Institute of Technology - Technion, Israel.*
Rémi Gribonval, *INRIA & CNRS, France.*
Rodolphe Jenatton, *Amazon, Germany.*
Julien Mairal, *INRIA, France.*
Deanna Needell, *Claremont McKenna College, USA.*
Robert Nowak, *University of Wisconsin, Madison, USA.*
Holger Rauhut, *RWTH Aachen University, Germany.*
Justin Romberg, *Georgia Institute of Technology, USA.*
Ivan Selesnick, *New York University, USA.*
Jared Tanner, *Oxford University, UK.*
Bruno Torresani, *Université d'Aix-Marseille, France.*

Download the full proceedings here

Steering Committee

Volkan Cevher, *École polytechnique Fédérale de Lausanne, Switzerland.*
Mike Davies, *University of Edingurgh, UK.*
Michael Elad, *Technion, Israel*
Jalal Fadili, *École Nationale Supérieure d'Ingénieurs de Caen, France.*
Maryam Fazel, *University of Washington, USA*
Mário A. T. Figueiredo, *Instituto Superior Técnico, Universidade de Lisboa, Portugal.*
Rémi Gribonval, *Centre de Recherche INRIA Rennes, France.*
Nick Kingsbury, *University of Cambridge, UK*
Matthieu Kowalski, *Université Paris-Sud, France.*
Gitta Kutyniok, *Technische Universität Berlin, Germany.*
Mark Plumbley, *University of Surrey, UK.*
Jared Tanner, *University of Oxford, UK.*
Pierre Vandergheynst, *École polytechnique Fédérale de Lausanne, Switzerland.*
Yves Wiaux, *Heriot-Watt University, UK.*

SPARS 2017 is supported by the EU Marie Skłodowska-Curie Innovative Training Networks MacSeNet (www.macsenet.eu) and SpaRTaN (www.spartan-itn.eu).

SpaRTaN
Sparse Representations and Compressed Sensing Training Network

MacSeNet
Machine Sensing Training Network

Welcome from the Chairs

Welcome to SPARS 2017, the 6th edition of the Signal Processing with Adaptive Sparse Structured Representations workshop. As with previous editions, SPARS 2017 brings together statisticians, engineers, mathematicians and computer scientists working in the areas of sparsity-related techniques and computational methods for high dimensional data analysis, signal processing, and related applications.

The program of SPARS 2017 features a set of 8 invited talks by renowned experts: Yoram Bresler (University of Illinois, USA), Volkan Cevher (EPFL, Switzerland), Jalal Fadili (ENSI Caen, France), Anders Hansen (University of Cambridge, UK), Gitta Kutyniok (TU Berlin, Germany), Philip Schniter (Ohio State University, USA), Eero Simoncelli (NYU, USA), and Rebecca Willett (University of Wisconsin, USA).

SPARS 2017 also includes a single track with 34 contributed oral presentations, and three poster sessions (with a total of 111 posters), as well as a special poster session describing work carried out within two Marie Skłodowska-Curie training networks, SpaRTaN and MacSeNet, which provided financial and organizational support to the workshop.

SPARS 2017 is co-organized by Instituto de Telecomunicações and held at Instituto Superior Técnico, the engineering school of the University of Lisbon.

Quoting a recent article by Steve Kin on cnntravel.com:

> *"I've been coming to Lisbon for nearly a quarter of a century. I honestly don't know what took the rest of the world so long. I can't think of another city that more richly deserves the attention."*

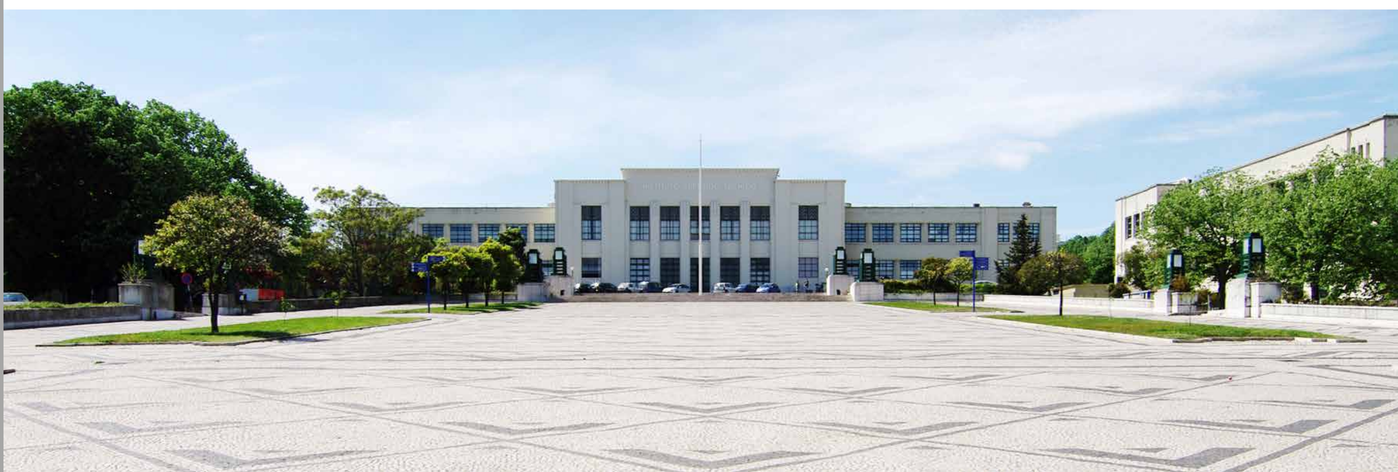
Finally, SPARS has come to Lisbon; we hope you enjoy this great workshop and this beautiful city.

Mário A. T. Figueiredo,
Instituto de Telecomunicações and Instituto Superior Técnico,
Universidade de Lisboa, Portugal.
Mark Plumbley,
University of Surrey, UK.

instituto de telecomunicações

TÉCNICO LISBOA

LISBOA UNIVERSIDADE DE LISBOA



Monday, June 5

09:15 Opening Session

09:25 Plenary I Yoram Bresler (Uni. of Illinois, USA)
Sparse Signal Recovery in Bilinear Inverse Problems

10:25 Oral Session 1

1. **Beyond ℓ_1: Data Driven Sparse Signal Recovery using DeepInverse**
   Mousavi and R. Baraniuk (Rice Uni., USA)
2. **From Sparse Bayesian Learning to Deep Recurrent Nets**
   H. He (Peking Uni.),
   B. Xin and D. Wipf (Microsoft Research Beijing, China)

11:15 Coffee Break

11:45 Oral Session 2

3. **Cosparse Denoising: The Importance of Being Social**
   C. Gaultier (INRIA, France), S. Kitic (Technicolor, France),
   S. Bertin (IRISA, France), R. Gribonval (INRIA, France)
4. **A Matrix Factorization Approach for Learning Semidefinite-Representable Regularizers**
   Y. Sheng Soh and V. Chandrasekaran (Caltech, USA)

12:35 Special Session | Mark Plumbley (Uni. of Surrey, UK)
The SpaRTaN and MacSeNet networks

13:00 Lunch and SpaRTaN-MacSeNet Special Poster Session

14:50 Plenary I Gitta Kutyniok (Technische Universität Berlin, Germany)
Optimal Approximation with Sparse Deep Neural Networks

15:50 Oral Session 3

5. **Analyzing Convolutional Neural Networks Through the Eyes of Sparsity**
   V. Papyan, Y. Romano, and M. Elad (Technion, Israel)
6. **Working Locally Thinking Globally: Guarantees for Convolutional Sparse Coding**
   V. Papyan, J. Sulam, and M. Elad (Technion, Israel)

16:40 Coffee Break

17:10 Oral Session 4

7. **Invariant Multiscale Statistics for Inverse Problems**
   I. Dokmanic (Uni. of Illinois, USA), J. Bruna (New York Uni., USA),
   S. Mallat (ENS, France), M. de Hoop (Rice Uni., USA)
8. **DOA estimation in fluctuating oceans: put your glasses on!**
   C. Herzet (INRIA, France), A. Dremeau (ENSTA Bretagne, France)

18:00 Welcome Reception

Tuesday, June 6

09:00 Plenary I Volkan Cevher (EPFL, Switzerland)
Learning-Based Compressive Subsampling

10:00 Oral Session 5

9. Learning Dictionaries as Sums of Kronecker Products
C. Dantas and R. Gribonval (INRIA, France),
R. Lopes and M. da Costa (Uni. of Campinas, Brazil)

10. Minimax Lower Bounds for Dictionary Learning from Tensor Data
Z. Shakeri, W. Bajwa, and A. Sarwate (Rutgers Uni., USA)

10:50 Coffee Break

11:20 Oral Session 6

11. Scalable Convex Methods for Low-Rank Matrix Problems
A. Yurtsever (EPFL, Switzerland),
Q. Tran-Dinh (Uni. of North Carolina, USA), V. Cevher (EPFL, Switzerland)

12. The Nonconvex Geometry of Low-Rank Matrix Optimizations
Q. Li, Z. Zhu, and G. Tang (Colorado School of Mines, USA)

13. On Computational and Statistical Tradeoffs in Matrix Completion with Graph Information
G. Dasarathy (Rice Uni., USA), N. Rao (Uni. of Texas at Austin, USA),
R. Baraniuk (Rice Uni., USA)

12:35 Lunch and Poster Session 1

14:50 Plenary I Jalal Fadili (ENSICAEN, France)
Exponential Weighted Aggregation vs Penalized Estimation: Guarantees and Algorithms

15:50 Oral Session 7

14. On the Tradeoff between Convergence Speed and Reconstruction Accuracy in Inverse Problems
R. Giryes (Tel Aviv Uni., Israel), Y. Eldar (Technion, Israel),
A. Bronstein (Technion, Israel), G. Sapiro (Duke Uni., USA)

15. Local Linear Convergence of Primal--Dual Splitting Methods for Low Complexity Regularization
J. Liang and J. Fadili (ENSICAEN, France),
G. Peyré (CNRS, DMA, ENS, France)

16:40 Coffee Break

17:10 Oral Session 8

16. Regularized Nonlinear Acceleration
D. Scieur and F. Bach (INRIA, ENS, France),
A. d'Aspremont (INRIA, ENS, CNRS, France)

17. MultiD-AMP: Match Up Accuracy and Fast Computation by Dynamically Denoising Data
A. Perelli and M. Davies (Uni. of Edinburgh, UK)

Wednesday, June 7

09:00 Plenary | Eero Simoncelli (New York Uni., USA)
Cascaded Gain Control Representations for Images

10:00 Oral Session 9

18. Random Moments for Sketched Mixture Learning
N. Keriven and R. Gribonval (INRIA, France),
G. Blanchard (Uni. of Potsdam, Germany), Y. Traonmilin (INRIA, France)

19. Subspace Estimation from Incomplete Observations: A Precise High-Dimensional Analysis
C. Wang (Harvard Uni., USA), Y. Eldar (Technion, Israel),
Y. Lu (Harvard Uni., USA)

10:50 Coffee Break

11:20 Oral Session 10

20. A Guaranteed Poly-Logarithmic Time Relaxation for the Line Spectral Estimation Problem
M. Ferreira Da Costa and W. Dai (Imperial College of London, UK)

21. Spikes super-resolution with random Fourier sampling
Y. Traonmilin, N. Keriven, R. Gribonval (INRIA, France) and
G. Blanchard (Uni. of Potsdam, Germany)

22. Sampling the Fourier Transform along Radial Lines
C. Dossal (Uni. of Bordeaux, France), C. Poon (Uni. of Cambridge, UK),
V. Duval (INRIA, France)

12:35 Lunch and Poster Session 2

14:50 Plenary | Anders Hansen (Uni. of Cambridge, UK)
On Foundational Computational Problems in ℓ_1 and Total Variation Regularisation

15:50 Oral Session 11

23. A Simple Convex Program for Phase Retrieval Using Anchor Vectors
S. Bahmani and J. Romberg (Georgia Institute of Technology, USA)

24. Low Rank Phase Retrieval
N. Vaswani and S. Nayer (Iowa State Uni., USA), and
Y. Eldar (Technion, Israel)

16:40 Coffee Break

17:10 Oral Session 12

25. Convolutional Phase Retrieval via Gradient Descent
Q. Qu and Y. Zhang (Columbia Uni., USA), Y. Eldar (Technion, Israel), J. Wright (Columbia Uni., USA)

26. PhaseMax: Convex Phase Retrieval Without Lifting
T. Goldstein (Stanford Uni., USA), C. Studer (Cornell Uni., USA)

20:00 SPARS Banquet (transportation provided)

Thursday, June 8

09:00 **Plenary | Phil Schniter** (Ohio State Uni., USA)
Recent Advances in Approximate Message Passing

10:00 **Oral Session 13**

27. Robust Outlier Identification for Noisy Data via Randomized Adaptive Compressive Sampling
X. Li and J. Haupt (Uni. of Minnesota, USA)

28. Robustness to Unknown Error in Sparse Regularization
S. Brugiapaglia and B. Adcock (Simon Fraser Uni., USA),
R. Archibald (Oak Ridge National Laboratory, USA)

10:50 **Coffee Break**

11:20 **Oral Session 14**

29. Sparse Recovery From Superimposed Non-Linear Sensor Measurements
M. Genzel and P. Jung (Technical Uni. of Berlin, Germany)

30. Recovery of Nonlinearly Degraded Sparse Signals through Rational Optimization
M. Castella (Telecom SudParis, France),
J.-C. Pesquet (Uni. Paris-Saclay, France)

31. Computing the Spark of a Matrix
Andreas Tillmann (RWTH Aachen Uni., Germany),
M. Pfetsch (Technical Uni. of Darmstadt, Germany)

12:35 **Lunch and Poster Session 3**

14:50 **Plenary | Rebecca Willett** (Uni. of Wisconsin, USA)
Nonlinear Models for Matrix Completion

15:50 **Coffee Break**

16:20 **Oral Session 15**

32. Stabilizing Embedology: Geometry-Preserving Delay-Coordinate Maps
C. Rozell (Georgia Institute of Technology, USA),
M. Wakin (Colorado School of Mines, USA),
H. Lun Yap (Georgia Institute of Technology, USA),
A. Eftekhari (The Alan Turing Institute, USA)

33. Rare Eclipses in Quantised Random Embeddings of Disjoint Convex Sets: a Matter of Consistency?
V. Cambareri, C. Xu, and L. Jacques
(Uni. Catholique de Louvain, Belgium)

34. Multilinear Low-Rank Tensors on Graphs & Applications
N. Shahid, F. Grassi, and P. Vandergheynst (EPFL, Switzerland)

17:35 **Closing Remarks**

Plenary | Yoram Bresler, (Uni. of Illinois, USA)

Sparse Signal Recovery in Bilinear Inverse Problems

Monday June 5th 09:25

Much less is known about the solution of bilinear inverse problems (BLIPs) than about their linear counterparts. In signal processing, BLIPs arise in blind signal recovery, notably in blind deconvolution, with applications in blind image deblurring, blind channel equalization, speech dereverberation, and seismic data analysis. Another important example, is blind gain and phase calibration (BGPC), arising, e.g., in blind albedo estimation in inverse rendering, in sensor array processing with miscalibrated arrays, in multichannel blind deconvolution, and in multichannel MRI.
In certain cases, low dimensionality or sparsity constraints alleviate the ill-posedness of BLIPs. However, until recently, existing theoretical analysis on uniqueness in these problems was rather limited. We present results on sample complexity that guarantee stable recovery in blind deconvolution and in BGPC under minimal requirements. We complement these theoretical results by practical recovery algorithms: (1) a guaranteed sparse blind deconvolution algorithm with optimal (up to log factors) scaling of the sample complexity, both in theory and in numerical experiments; and (2) a BGPC algorithm that in adversarial conditions improves on competing algorithms.
Based on joint work with Yanjun Li, Kiryung Lee, and Marius Junge

---

Yoram Bresler received the B.Sc. (cum laude) and M.Sc. degrees from the Technion, Israel Institute of Technology, in 1974 and 1981 respectively, and the Ph.D degree from Stanford Uni., in 1986, all in Electrical Engineering.
In 1987 he joined the Uni. of Illinois at Urbana-Champaign, where he is currently a Professor at the Departments of Electrical and Computer Engineering and Bioengineering, and Research Professor at the Coordinated Science Laboratory.
His current research interests include compressed sensing and multi-dimensional and statistical signal processing and their applications to inverse problems in imaging, and in particular computed tomography and magnetic resonance imaging.
Dr. Bresler has served on the editorial board of a number of journals, and on various committees of the IEEE. Currently he serves on the editorial boards for the IEEE Journal on Selected Topics in Signal Processing, and the SIAM Journal on Imaging Science. Dr. Bresler is a fellow of the IEEE and of the AIMBE. He received two Senior Paper Awards from the IEEE Signal Processing society, and a paper he coauthored with one of his students received the Young Author Award from the same society in 2002. He is the recipient of a 1991 NSF Presidential Young Investigator Award, the Technion (Israel Inst. of Technology) Fellowship in 1995, and the Xerox Senior Award for Faculty Research in 1998. He was named a Uni. of Illinois Scholar in 1999, appointed as an Associate at the Center for Advanced Study of the Uni. in 2001-2, and Faculty Fellow at the National Center for Super Computing Applications in 2006.
Dr. Bresler is a co-founder (along with Dr. Munson), president, and chief technology officer of InstaRecon, Inc., a software startup that develops and markets breakthrough image reconstruction technology for CT scanners, originating from his research with students and colleagues at the Uni.

Plenary | Gitta Kutyniok, (TU Berlin, Germany)

Optimal Approximation with Sparse Deep Neural Networks

Monday June 5th 14:50

Despite the outstanding success of deep neural networks in real-world applications, most of the related research is empirically driven and a mathematical foundation is almost completely missing. One central task of a neural network is to approximate a function, which for instance encodes a classification task. In this talk, we will be concerned with the question, how well a function can be approximated by a deep neural network with sparse connectivity. We will derive fundamental lower bounds on the connectivity and the memory requirements of deep neural networks guaranteeing uniform approximation rates for arbitrary function classes, also including functions on low-dimensional immersed manifolds. Additionally, we prove that our lower bounds are achievable for a broad family of function classes, thereby deriving an optimality result. Finally, we present numerical experiments demonstrating that the standard stochastic gradient descent algorithm generates deep neural networks providing close-to-optimal approximation rates at minimal connectivity. Moreover, surprisingly, these results show that stochastic gradient descent actually learns approximations that are sparse in the representation systems optimally sparsifying the function class the network is trained on.
This is joint work with H. Bölcskei, P. Grohs, and P. Petersen.

Gitta Kutyniok completed her Diploma in Mathematics and Computer Science in 1996 at the Universitat Paderborn in Germany. She received her Ph.D. degree in the area of time-frequency analysis from the same Uni. in 2000. She completed her Habilitation in Mathematics in 2006 and received her venia legendi. In 2007, she was awarded a Heisenberg Fellowship by the DFG-German Research Foundation.
From 2001 to 2008 she held visiting appointments at several US institutions, including Princeton Uni., Stanford Uni., Yale Uni., Georgia Institute of Technology, and Washington Uni. in St. Louis.
After returning to Germany in October 2008, she became a full professor of mathematics at the Universitat Osnabrueck, and headed the Applied Analysis Group. Since October 2011, she has an Einstein Chair at the Technical Uni. of Berlin and is head of the Applied Functional Analysis Group (AFG).
Her research and teaching have been recognized by various awards, including the von Kaven Prize by the German Research Foundation, awards by the Uni. Paderborn and the Justus-Liebig Uni. Giessen for Excellence in Research, as well as the Weierstrass Prize for Outstanding Teaching. She is an Associate Editor and also Corresponding Editor for several journals in the area of applied mathematics. She is also a board member of the Berlin Mathematical School, a member of the council of the MATHEON Mathematics for key technologies in Berlin, and the chair of the GAMM activity group on Mathematical Signal and Image Processing.

Plenary | Volkan Cevher, (EPFL, Switzerland)

Learning-based compressive subsampling

Tuesday June 6th 09:00

In the past decade, there have been significant advances in methods that directly acquire only the relevant information during data acquisition. For instance, compressive sensing (CS) simultaneously performs data acquisition and compression, seeking to directly sample for the relevant information in signals as opposed to acquiring a full signal only to throw most of it away via compression. Despite the tremendous amount of research in CS, there exist severe limitations in this existing theory and methodology that have prevented its widespread use in practical systems. Indeed, many approaches assume an unrealistic amount of knowledge of the system model (e.g., perfect knowledge of the sparsity basis). Moreover, the strongest theoretical results for CS are based on measurement designs that scale poorly due to the computational challenges associated with dense sensing operators, or that rely on randomization methods that are impractical and perhaps not even possible in applications.

In this talk, we will introduce a new paradigm, partially addressing emerging challenges by unifying compressive sensing with statistical learning theory. Our key idea, which is surprisingly simple in retrospect, is that by using training signals and developing combinatorial training procedures, we can efficiently and effectively learn the structure inherent in the data, and accordingly design measurement matrices that directly acquire only the relevant information during acquisition. As a result, we will show how we can not only outperform the existing state-of-the-art compressive sensing techniques on real-world datasets (including neural signal acquisition and magnetic resonance imaging), but can do so with strong theoretical guarantees.

In particular, we will describe how to optimize the samples for the standard linear acquisition model along with the use of a simple linear decoder, and build towards optimizing the samples for non-linear reconstruction algorithms. Overall, the mathematical premise of our approach represents a re-thinking of the data models and dimensionality reduction using continuous optimization algorithms based on convexity, and combinatorial optimization algorithms based on submodularity.

Volkan Cevher received the B.Sc. (valedictorian) in electrical engineering from Bilkent Uni. in Ankara, Turkey, in 1999 and the Ph.D. in electrical and computer engineering from the Georgia Institute of Technology in Atlanta, GA in 2005. He was a Research Scientist with the Uni. of Maryland, College Park from 2006-2007 and also with Rice Uni. in Houston, TX, from 2008-2009. Currently, he is an Associate Professor at the Swiss Federal Institute of Technology Lausanne and a Faculty Fellow in the Electrical and Computer Engineering Department at Rice Uni.. His research interests include signal processing theory, machine learning, convex optimization, and information theory. Dr. Cevher was the (co-) recipient of the IEEE Signal Processing Society Best Paper Award in 2016, a Best Paper Award at CAMSAP in 2015, a Best Paper Award at SPARS in 2009, and an ERC CG in 2016 as well as an ERC StG in 2011.

Plenary | Jalal Fadili, (ENSICAEN, France)

Exponential Weighted Aggregation vs Penalized Estimation: Guarantees and Algorithms

Tuesday June 6th 14:50

Two approaches are usually used to solve high-dimensional linear inverse problems: 1) the penalized approach which amounts to solving an optimization problem, and 2) the exponential weighted aggregation (EWA) which solves an integration problem. In this talk, I will present a unified analysis of the performance guarantees of both estimators with a general class of priors which encourage objects which conform to some notion of simplicity/complexity. More precisely, we show that these two estimators satisfy (i) sharp oracle inequalities for prediction, and (ii) estimation error bounds ensuring their good theoretical performances. We also highlight the differences between them. When the noise or the design is random, we provide bounds which holds with high probability under mild assumptions on the underlying distribution. Finally, we propose a framework based on proximal splitting to efficiently implement these estimators. The results are then exemplified on several popular instances in signal/image processing and machine learning.

Jalal M. Fadili graduated from the École Nationale Supérieure d'Ingénieurs (ENSI) de Caen, Caen, France, and received the M.Sc. and Ph.D. degrees in signal and image processing from the Uni. of Caen. He was a Research Associate with the Uni. of Cambridge MacDonnel-Pew Fellow), Cambridge, U.K., from 1999 to 2000. He has been an Associate Professor of signal and image processing since September 2001 at ENSI. He also received an Habilitation from Uni. of Caen in 2010. He was a visitor at several universities (QUT-Australia, Stanford Uni., CalTech, MIT, EPFL). He is the co-author of a book entitled Sparse Signal and Image Processing: Wavelets, Curvelets, Morphological Diversity Cambridge Uni. Press. His research interests include statistical approaches in signal and image processing, inverse problems, computational harmonic analysis, optimization and sparse representations. His areas of application include medical and astronomical imaging.

Plenary | Eero Simoncelli, (New York Uni., USA)

Cascaded Gain Control Representations for Images

Wednesday June 7th 09:00

I'll describe some recent work on developing models for visual representation based on a cascaded architecture, in which each stage performs convolutions with a set of kernels, nonlinear rectification, and locally adaptive gain control. All three operations are primarily motivated by known functional properties of neural response throughout the visual system, and their parameters can be set by fitting physiological or perceptual data. Alternatively, they can be optimized to capture statistical relationships found in natural images, or to achieve a specific engineering goal, such as restoration, compression, or recognition. I'll show examples of their use in optimizing the display of photographs (for example, rendering an HDR image on a LDR display), and as the basis of a high-quality image coder that is optimized in terms of its rate-distortion tradeoff.

Dr. Simoncelli is Professor of Neural Science, Mathematics, and Psychology at New York Uni.. He began his higher education as a physics major at Harvard, went to Cambridge Uni. on a Knox Fellowship to study mathematics for a year and a half, and earned a doctorate in electrical engineering and computer science at the Massachusetts Institute of Technology. He then joined the faculty of the Computer and Information Science Department at the Uni. of Pennsylvania. In 1996, he moved to NYU as part of the Sloan Center for Theoretical Visual Neuroscience. In 2000, he became an Investigator of the Howard Hughes Medical Institute. He is a fellow of the IEEE, has received two IEEE Best Paper awards and a Sustained Impact paper award, and received an Engineering Emmy award in 2015.

Plenary | Anders Hansen, (Uni. of Cambridge, UK)

On Foundational Computational Problems in ℓ_1 and Total Variation Regularisation

Wednesday June 7th 14:50

The use of regularisation techniques such as ℓ_1 and Total Variation in Basis Pursuit and Lasso has been a great success in wide areas of mathematics and statistics over the last decades. In this talk we will discuss universal boundaries regarding the existence of algorithms for solving these problems. For example we have the following paradox: it is impossible to design algorithms to solve these general problems accurately when given inaccurate input data, even when the inaccuracies can be made arbitrarily small. As a simple number such as root(2) never comes with an exact numerical representation, inaccurate data input is a daily encounter. The impossibility result implies that for any algorithm designed to solve these problems there will be cases where the algorithm fails in the following way: For fixed dimensions and any small accuracy parameter $\varepsilon > 0$, one can choose an arbitrary large time *T* and find an input such that the algorithm will run for longer than *T* and still not have reached epsilon accuracy. Moreover, it is impossible to determine when the algorithm should halt to achieve an epsilon accurate solution, and hence the algorithm will never be able to produce an output where one knows that the output is at least epsilon accurate. The largest epsilon for which this failure happens is called the Breakdown-ε. For Basis Pursuit and Lasso, the Breakdown-$\varepsilon > 1/3$ even when the absolute value of the input is bounded by one and is well conditioned.
The paradox opens up for a new classification theory to determine the boundaries of what computers can achieve in regularisation, and to explain why empirically many modern algorithms for solving regularisation problem in real-world scenarios perform very well. We will discuss positive classification results showing that sparse problems can be computed accurately. However, this is delicate; e.g. given standard assumptions from sparse recovery, there are algorithms that can compute a solution to Basis Pursuit accurately, however, this is impossible for Lasso and Basis Pursuit with noise parameter $\delta > 0$. However, one can compute a solution accurately up to the Breakdown-ε that tends to zero when δ tends to zero, and coincides with the error bound provided in the theory of sparse recovery. This helps explaining the success of many modern algorithms applied in numerous real-world scenarios, and also explains the cases where algorithms will fail and why.

Anders Hansen leads the Applied Functional and Harmonic Analysis group within the Cambridge Centre for Analysis at DAMTP. He is a Lecturer at the Department of Applied Mathematics and Theoretical Physics, Professor of Mathematics at the Uni. of Oslo, a Royal Society Uni. Research Fellow, and a Fellow of Peterhouse. He received an M.A. from the Uni. of California, Berkeley, and a PhD from the Uni. of Cambridge, both in mathematics. His research interests are: applied functional analysis, operator/spectral theory, complexity theory, foundations of computational mathematics, compressed sensing, mathematical signal processing, computational harmonic analysis, inverse problems, medical imaging, geometric integration, numerical analysis, C algebras. He is on the editorial board of the Proceedings of the Royal Society (Series A).*

Plenary | Phil Schniter, (Ohio State Uni., USA)

Recent Advances in Approximate Message Passing

Thursday June 8th 09:00

The approximate message passing (AMP) algorithm proposed in 2009 by Donoho, Maleki, and Montanari is a computationally efficient iterative approach to sparse reconstruction and related problems. AMP has a remarkable property: for large i.i.d. sub-Gaussian measurement matrices, its per-iteration behavior is rigorously characterized by a scalar state-evolution whose fixed points, when unique, are Bayes optimal. However, AMP is fragile in that even small deviations from the i.i.d. sub-Gaussian model can cause the algorithm to diverge. In this talk, I will describe a vector AMP (VAMP) algorithm, which also has a rigorous scalar state-evolution. VAMP's state-evolution, however, holds under a much broader class of large random measurement matrices, those that are right-rotationally invariant. I will also discuss non-parametric versions of VAMP that can cope with unknown prior and/or likelihood, connections between VAMP and convex optimization algorithms, plug-and-play extensions of VAMP, and connections between VAMP and deep neural networks.

Philip Schniter received the B.S. and M.S. degrees in Electrical Engineering from the Uni. of Illinois at Urbana-Champaign in 1992 and 1993, respectively. From 1993 to 1996 he was employed by Tektronix Inc. in Beaverton, OR as a systems engineer. In 2000, he received the Ph.D. degree in Electrical Engineering from Cornell Uni. in Ithaca, NY. Subsequently, he joined the Department of Electrical and Computer Engineering at The Ohio State Uni. in Columbus, OH, where he is now a Professor and a member of the Information Processing Systems (IPS) Lab. He held visiting professor positions at Eurecom (Sophia Antipolis, France) from October 2008 through February 2009, and at Supelec (Gif sur Yvette, France) from March 2009 through August 2009. He is currently (August 2016 to June 2017) a visiting professor at Duke Uni. (Durham, NC).

Philip Schniter is an IEEE Fellow and a member of the IEEE Signal Processing Society and IEEE Information Theory Society. Since 2013 he has been elected to serve on the IEEE Sensor Array and Multichannel (SAM) Technical Committee, and from 2005-2010 he was elected to serve on the IEEE Signal Processing for Communications and Networking (SPCOM) Technical Committee.

While pursuing his Ph.D. degree, Dr. Schniter received a Schlumberger Fellowship and an Intel Foundation Fellowship. He was awarded the 1999 Prize Paper Award from the IEEE Energy Development and Power Generation Committee for work relating to his M.S. thesis. In 2003, he received the National Science Foundation CAREER Award and, in 2005, the OSU College of Engineering Lumley Research Award. With graduate student Jason Parker and Volkan Cevher, he won the 2016 IEEE Signal Processing Society Best Paper Award.

Dr. Schniter's areas of research include signal processing, machine learning, information theory, communication theory, compressed sensing, and sensor networks. His recent sources of funding include the National Science Foundation, the Defense Advanced Research Projects Agency, the Air Force Research Laboratory, MIT Lincoln Labs, the Office of Naval Research, Sandia National Labs, and Motorola Labs.

Plenary | Rebecca Willett, (Uni. of Wisconsin, USA)

Nonlinear Models for Matrix Completion

Thursday June 8th 14:50

The past decade of research on matrix completion has shown it is possible to leverage linear dependencies to impute missing values in a low-rank matrix. However, the corresponding assumption that the data lies in or near a low-dimensional linear subspace is not always met in practice. Extending matrix completion theory and algorithms to exploit low-dimensional nonlinear structure in data will allow missing data imputation in a far richer class of problems. In this talk, I will describe several models of low-dimensional nonlinear structure and how these models can be used for matrix completion. In particular, we will explore matrix completion in the context of three different nonlinear models: single index models, in which a latent subspace model is transformed by a nonlinear mapping; unions of subspaces, in which data points lie in or near one of several subspaces; and nonlinear algebraic varieties, a polynomial generalization of classical linear subspaces. In these settings, we will explore novel and efficient algorithms for imputing missing values and new bounds on the amount of missing data that can be accurately imputed. The proposed algorithms are able to recover synthetically generated data up to predicted sample complexity bounds and outperform standard low-rank matrix completion in experiments with real recommender system and motion capture data.

Rebecca Willett is an Associate Professor of Electrical and Computer Engineering, Harvey D. Spangler Faculty Scholar, and Fellow of the Wisconsin Institutes for Discovery at the Uni. of Wisconsin-Madison. She completed her PhD in Electrical and Computer Engineering at Rice Uni. in 2005 and was an Assistant then tenured Associate Professor of Electrical and Computer Engineering at Duke Uni. from 2005 to 2013. Willett received the National Science Foundation CAREER Award in 2007, is a member of the DARPA Computer Science Study Group, and received an Air Force Office of Scientific Research Young Investigator Program award in 2010. Willett has also held visiting researcher or faculty positions at the Uni. of Nice in 2015, the Institute for Pure and Applied Mathematics at UCLA in 2004, the Uni. of Wisconsin-Madison 2003-2005, the French National Institute for Research in Computer Science and Control (INRIA) in 2003, and the Applied Science Research and Development Laboratory at GE Healthcare in 2002.

Oral Papers - Abstracts

1. **Beyond ℓ_1: Data Driven Sparse Signal Recovery using DeepInverse**
   Mousavi and R. Baraniuk (Rice Uni., USA)
   We study a novel sparse signal recovery framework called *DeepInverse* that learns the inverse transformation from measurement vectors to signals using a deep convolutional network. We compare DeepInverse with ℓ_1-minimization from the phase transition point of view and demonstrate that it outperforms ℓ_1-minimization in the regions of phase transition plot where ℓ_1-minimization cannot recover the exact solution.

2. **From Sparse Bayesian Learning to Deep Recurrent Nets**
   H. He (Peking Uni.), B. Xin and D. Wipf (Microsoft Research Beijing, China)
   The iterations of many simple sparse estimation algorithms are comprised of a fixed linear filter cascaded with a thresholding nonlinearity, which collectively resemble a typical neural network layer. Consequently, a lengthy sequence of algorithm iterations can be viewed as a deep network with shared, hand-crafted layer weights. In this work we demonstrate that more complex approaches like sparse Bayesian learning mirror the structure of sophisticated long short-term memory (LSTM) networks and, when appropriately learned from training data, can estimate maximally sparse solutions efficiently in regimes where other algorithms fail.

3. **Cosparse Denoising: The Importance of Being Social**
   C. Gaultier (INRIA, France), S. Kitic (Technicolor, France),
   S. Bertin (IRISA, France), R. Gribonval (INRIA, France)
   This work investigates the performance of cosparse vs. social cosparse regularizations in addressing the audio denoising problem. Beyond the cosparse (also known as sparse analysis) model, results show that exploiting structures in the time-frequency domain is beneficial to audio signal restoration for high degradation levels.

4. **A Matrix Factorization Approach for Learning Semidefinite-Representable Regularizers**
   Y. Sheng Soh and V. Chandrasekaran (Caltech, USA)
   Regularization techniques are widely employed in optimization-based approaches for solving ill-posed inverse problems in data analysis and scientific computing. These methods are based on augmenting the objective with a penalty function, which is specified based on prior domain-specific expertise to induce a desired structure in the solution. We consider the problem of learning suitable regularization functions from data in settings in which precise domain knowledge is not directly available. Previous work under the title of 'dictionary learning' or 'sparse coding' may be viewed as learning a regularization function that can be computed via linear programming. We describe generalizations of these methods to learn regularizers that can be computed and optimized via semidefinite programming. Our framework for learning such semidefinite regularizers is based on obtaining structured factorizations of data matrices, and our algorithmic approach for computing these factorizations combines recent techniques for rank minimization problems along with an operator analog of Sinkhorn scaling. Under suitable conditions on the input data, our algorithm provides a locally linearly convergent method for identifying the correct regularizer that promotes the type of structure contained in the data. Our analysis is based on the stability properties of the Operator Sinkhorn scaling and their relation to geometric aspects of determinantal varieties (in particular tangent spaces

with respect to these varieties). The regularizers obtained using our framework can be employed effectively in semidefinite programming relaxations for solving inverse problems.

5. **Analyzing Convolutional Neural Networks Through the Eyes of Sparsity**
V. Papyan, Y. Romano, and M. Elad (Technion, Israel)
It is becoming increasingly difficult to ignore the remarkable results of Convolutional Neural Networks (CNN), and the need for its theoretical analysis. In this work, we aim to alleviate this gap by proposing a novel model - the multi-layer convolutional sparse coding (ML-CSC). This defines a set of signals for which the forward pass of CNN is nothing but a thresholding pursuit. Leveraging this connection, we are able to attribute to the CNN architecture theoretical claims such as uniqueness of the representations (feature maps) throughout the network and their stable estimation, all guaranteed under simple local sparsity conditions. Sitting on these theoretical grounds, we propose a better pursuit that is shown to be theoretically superior to the forward pass.

6. **Working Locally Thinking Globally: Guarantees for Convolutional Sparse Coding**
V. Papyan, J. Sulam, and M. Elad (Technion, Israel)
The acclaimed sparse representation model has led to remarkable results in various signal processing tasks. However, despite its initial purpose of serving as a global prior for entire signals, it has been commonly used for modeling low dimensional patches due to the computational constraints entailed when deployed with learned dictionaries. The emerging convolutional sparse coding (CSC) model comes as a globally-aware alternative. Several works have presented algorithmic solutions to the global pursuit problem under this new model, yet very few truly effective guarantees are known for their success. In this work, we address the theoretical aspects of the CSC model, providing the first meaningful answers to questions of uniqueness of solutions and success of pursuit algorithms. Moreover, we further extend this analysis to the noisy regime, addressing the stability of the sparsest solutions and of the associated algorithms. Finally, we demonstrate practical approaches for solving the global pursuit problem via simple local processing.

7. **Invariant Multiscale Statistics for Inverse Problems**
I. Dokmanic (Uni. of Illinois, USA), J. Bruna (New York Uni., USA), S. Mallat (ENS, France), M. de Hoop (Rice Uni., USA)
We propose a new approach to linear ill-posed inverse problems. Our algorithm stabilizes the inversion by enforcing a new statistical constraint in a suitable feature space. We use the non-linear multiscale scattering transform — a complex convolutional network which discards the phase and thus exposes strong spectral correlations otherwise hidden beneath the phase fluctuations. We apply the algorithm to super-resolution and tomography with synthetic signals, and show that it outperforms regularized methods and stably recovers the missing spectrum. Further, we discuss the choice of the feature transform as a function of the operator and input statistics, and we show convergence of the proposed iterative algorithm.

8. **DOA estimation in fluctuating oceans: put your glasses on!**
C. Herzet (INRIA, France), A. Dremeau (ENSTA Bretagne, France)
In this work, we deal with the problem of estimating the directions of arrival (DOA) of a set of incident plane waves. Most contributions in this field assume that the

received signal is only corrupted by some additive noise. Unfortunately, when the waves travel through highly fluctuating media, as in the case of e.g. atmospheric sound propagation or underwater acoustics, this model does no longer describe accurately the physics underlying the propagation process. In such cases, a multiplicative phase noise typically corrupts the collected signal, making the corresponding DOA estimation problem much more challenging. We propose here a new methodology to address this issue. Our procedure is grounded on a probabilistic model combining a sparsity-enforcing Bernoulli-Gaussian prior on the DOAs and a Markov-Gaussian model on the phase noise. The estimation of the DOA is based on a mean-field approximation of the Minimum Mean Square Error (MMSE) estimate associated to this probabilistic model. Our work also relates to the phase retrieval problem where the phase information on the observations is completely missing: only intensities or amplitudes are acquired. Formally, both the phase retrieval problem and the problem considered here share the same observation model but differ in the prior distribution they enforce on the phase noise, the absence of phase information being modeled by a non-informative prior, such as a uniform law. We show in the present work how to nicely incorporate fine noise-phase models in this framework, extending in this respect, some Bayesian approaches in the literature.

9. **Learning Dictionaries as Sums of Kronecker Products**

   C. Dantas and R. Gribonval (INRIA, France),
   R. Lopes and M. da Costa (Uni. of Campinas, Brazil)

   The choice of an appropriate dictionary is a crucial step in the sparse representation of a given class of signals. Traditional dictionary learning techniques generally lead to unstructured dictionaries which are costly to deploy and do not scale well to higher dimensional signals. In order to overcome such limitation, we propose a learning algorithm that constrains the dictionary to be a sum of Kronecker products of smaller sub-dictionaries. A special case of the proposed structure is the widespread separable dictionary. This approach, named SuKro, is evaluated experimentally on an image denoising application.

10. **Minimax Lower Bounds for Dictionary Learning from Tensor Data**

    Z. Shakeri, W. Bajwa, and A. Sarwate (Rutgers Uni., USA)

    This paper provides fundamental limits on the sample complexity of estimating dictionaries for K-dimensional tensor data, in which case the underlying dictionary can be expressed in terms of K smaller coordinate dictionaries. This work provides lower bounds on the minimax risk of dictionary learning for tensor data in the case of general and sparse-Gaussian coefficient distributions. The results suggest the sample complexity of dictionary learning for tensor data can be significantly lower than that for unstructured data: for unstructured data it scales linearly with the product of the dictionary dimensions, whereas it need only scale linearly with the sum of the product of the dimensions of the coordinate dictionaries for tensor data.

11. **Scalable Convex Methods for Low-Rank Matrix Problems**

    A. Yurtsever (EPFL, Switzerland), Q. Tran-Dinh (Uni. of North Carolina, USA),
    V. Cevher (EPFL, Switzerland)

    We describe our new scalable primal-dual convex optimization framework to solve low-rank matrix recovery problems. Our algorithmic framework (universal primal-dual gradient method, abbreviated as UniPDGrad) extends Nesterov's universal gradient methods for the primal-dual setting in a non-trivial fashion. The main characteristics of our framework is the cheap per-iteration complexity and the low-memory footprint. We demonstrate the flexibility and scalability of our framework by solving matrix completion, quantum tomography and phase retrieval problems.

12. **The Nonconvex Geometry of Low-Rank Matrix Optimizations**
Q. Li, Z. Zhu, and G. Tang (Colorado School of Mines, USA)
This work considers the minimization of a general convex function f(X) either over the cone of positive semi-definite matrices or regularized by the nuclear norm. The optimal solution is assumed to be of low-rank. Standard first-order convex solvers require performing an eigenvalue or singular decomposition in each iteration, severely limiting their scalability. A natural nonconvex reformulation of the problem factors the variable X into the product of two smaller matrices. For a special class of matrix sensing and completion problems with quadratic objective functions, local search algorithms applied to the factored problem have been shown to be much more efficient and, in spite of being nonconvex, to converge to the global optimum. The purpose of this work is to extend this line of study to general convex objective functions f(X) and investigate the geometry of the resulting factored formulations. Specifically, we prove that when f(X) satisfies restricted strong convexity and smoothness, each critical point of the factored problem either corresponds to the optimal solution or is a strict saddle point where the Hessian matrix has a negative eigenvalue. Such a geometric structure of the factored formulation ensures that many local search algorithms can converge to the global optimum with random initializations.

13. **On Computational and Statistical Tradeoffs in Matrix Completion with Graph Information**
G. Dasarathy (Rice Uni., USA), N. Rao (Uni. of Texas at Austin, USA),
R. Baraniuk (Rice Uni., USA)
Matrix completion has generated significant interest owing to the central role it plays in several signal processing and machine learning applications. Modern datasets typically contain additional 'side information', e.g., a similarity graph on the underlying variables. It has been shown recently that incorporating this into the model yields significant statistical advantages. However, such algorithms typically result in higher computational costs. In this work, we propose a technique to gracefully tradeoff computation for statistical power using efficient algorithms for graph sparsification. Our theory and simulations suggest that for slight deterioration in statistical performance, we can obtain significant computational gains.

14. **On the Tradeoff between Convergence Speed and Reconstruction Accuracy in Inverse Problems**
R. Giryes (Tel Aviv Uni., Israel), Y. Eldar (Technion, Israel),
A. Bronstein (Technion, Israel), G. Sapiro (Duke Uni., USA)
Solving inverse problems with iterative algorithms is popular, especially for large data. Due to time constraints, the number of possible iterations is usually limited, potentially limiting the achievable accuracy. Given an error one is willing to tolerate, an important question is whether it is possible to modify the original iterations to obtain faster convergence to a minimizer achieving the allowed error without increasing the computational cost of each iteration considerably. Relying on recent recovery techniques developed for settings in which the desired signal belongs to some low-dimensional set, we show that using a coarse estimate of this set may

lead to a faster convergence at the cost of an additional error in the reconstruction related to the accuracy of the set approximation. Our theory ties to recent advances in sparse recovery, compressed sensing, and deep learning. Particularly, it may provide a possible explanation to the successful approximation of the ℓ_1-minimization solution by neural networks with layers representing iterations, as practiced in the learned iterative shrinkage-thresholding algorithm (LISTA).

15. **Local Linear Convergence of Primal--Dual Splitting Methods for Low Complexity Regularization**
    J. Liang and J. Fadili (ENSICAEN, France), G. Peyré (CNRS, DMA, ENS, France)
    Primal–Dual (PD) splitting method has become very popular for solving sparse recovery problems and beyond (see for instance the review [9]). The goal of this work is to understand the local convergence behaviour of PD which has been observed in practice to exhibit local linear rate of convergence. In this paper, we show that when the involved non-smooth functions are partly smooth, the PD algorithm identifies the associated active manifolds in finite time, and then locally converges linearly with a rate determined by the properties of the primal and dual active manifolds. The result is illustrated by several concrete examples and supported by numerical experiments.

16. **Regularized Nonlinear Acceleration**
    D. Scieur and F. Bach (INRIA, ENS, France),
    A. d'Aspremont (INRIA, ENS, CNRS, France)
    We describe a convergence acceleration technique for generic optimization problems. Our scheme computes estimates of the optimum from a nonlinear average of the iterates produced by any optimization method. The weights in this average are computed via a simple and small linear system, whose solution can be updated online.

17. **MultiD-AMP: Match Up Accuracy and Fast Computation by Dynamically Denoising Data**
    A. Perelli and M. Davies (Uni. of Edinburgh, UK)
    Denoising-AMP (D-AMP) can be viewed as an iterative algorithm where at each iteration a non-linear denoising function is applied to the signal estimate. D-AMP algorithm has been analysed in terms of inferential accuracy without considering computational complexity. This is an important missing aspect since the denoising is often the computational bottleneck in the D-AMP reconstruction. The approach that it is proposed in this work is different; we aim to design a mechanism for leveraging a hierarchy denoising models (MultiD-AMP) in order to minimize the overall complexity given the expected risk, i.e. the estimation error. The intuition comes from the observation that at earlier iteration, when the estimate is far according to some distance to the true signal, the algorithm does not need a complicated denoiser, since the structure of the signal is poor, but faster denoisers and this leads to the idea of defining a family/hierarchy of denoisers of increased complexity. The main challenge is to define a switching scheme which is based on the empirical finding that in MultiD-AMP we can predict exactly, in the large system limit, the evolution of the Mean Square Error. We can exploit the State Evolution, evaluated on a set of training images, to find a proper switching strategy. The proposed framework has been tested on i.i.d. random Gaussian measurements with Gaussian noise and

for deconvolution problem. The results show the effectiveness of the proposed reconstruction algorithm.

18. **Random Moments for Sketched Mixture Learning**
N. Keriven and R. Gribonval (INRIA, France),
G. Blanchard (Uni. of Potsdam, Germany), Y. Traonmilin (INRIA, France)
We present a method to solve large-scale mixture learning tasks from a sketch of the data, formed by random generalized empirical moments. We give empirical and theoretical results on k-means and Gaussian Mixture Model estimation problems.

19. **Subspace Estimation from Incomplete Observations: A Precise High-Dimensional Analysis**
C. Wang (Harvard Uni., USA), Y. Eldar (Technion, Israel), Y. Lu (Harvard Uni., USA)
The problem of estimating and tracking low-rank subspaces from incomplete observations has received a lot of attention recently in the signal processing and learning communities. Popular algorithms, such as GROUSE and PETRELS, are often very effective in practice, but their performance depends on the careful choice of algorithmic parameters. Important questions, such as the global convergence of these algorithms and how the noise level, subsampling ratio, and various other parameters affect the performance, are not fully understood. In this paper, we present a precise analysis of the performance of these algorithms in the asymptotic regime where the ambient dimension tends to infinity. Specifically, we show that the time-varying trajectories of estimation errors converge weakly to a deterministic function of time, which is characterized as the unique solution of a system of ordinary differential equations (ODEs.) Analyzing the limiting ODEs also reveals and characterizes sharp phase transition phenomena associated with these algorithms. Numerical simulations verify the accuracy of our asymptotic predictions, even for moderate signal dimensions.

20. **A Guaranteed Poly-Logarithmic Time Relaxation for the Line Spectral Estimation Problem**
M. Ferreira Da Costa and W. Dai (Imperial College of London, UK)
Line spectral estimation theory aims to estimate the off-the-grid spectral components of a time signal with optimal precision. Recent results have shown that it is possible to recover signals having sparse line spectra from few temporal observations via the use of convex programming. However, the computational cost of such approaches remains the major flaw to their use in practical systems. We present in this work a novel randomized algorithm that is guaranteed to recover the original signal with high probability. This algorithm works in a poly-logarithmic time with respect to the number of initial measurements.

21. **Spikes Super-Resolution with Random Fourier Sampling**
Y. Traonmilin, N. Keriven, R. Gribonval (INRIA, France) and
G. Blanchard (Uni. of Potsdam, Germany)
We leverage recent results from machine learning to show theoretically and practically that it is possible to stably recover a signal made of few spikes (in the gridless setting) from few random weighted Fourier measurements. Given a free choice of frequencies, a number of measurements lower than with the traditional low-pass filter (uniform sampling of low frequencies) guarantees stable recovery.

22. Sampling the Fourier Transform along Radial Lines

C. Dossal (Uni. of Bordeaux, France), C. Poon (Uni. of Cambridge, UK), V. Duval (INRIA, France)

This article considers the use of total variation minimization for the recovery of a superposition of point sources from samples of its Fourier transform along radial lines. We present a theoretical result precising the link between the sampling operator and the recoverability of the point sources.

23. A Simple Convex Program for Phase Retrieval Using Anchor Vectors

S. Bahmani and J. Romberg (Georgia Institute of Technology, USA)

We propose a new convex relaxation for phase retrieval that operates in the natural domain of the signal. Our method considers an enlarged set of solutions by relaxing the quadratic equations describing the problem to inequalities. Then an anchor vector is used in a simple convex program to estimate the ground truth as an (approximate) extreme point of the enlarged solution set. With random measurements, the proposed method is shown to produce accurate solutions. Our proposed method has significantly lower computational cost than the existing convex methods that rely on semidefinite programming; our method competes with the recent non-convex methods.

24. Low Rank Phase Retrieval

N. Vaswani and S. Nayer (Iowa State Uni., USA), and Y. Eldar (Technion, Israel)

We develop two iterative algorithms for solving the low rank phase retrieval (PR) problem. This refers to the problem of recovering a low-rank matrix from phaseless measurements of random linear projections of its columns. Both algorithms consist of a spectral initialization step followed by an iterative algorithm to maximize the observed data likelihood. We obtain sample complexity bounds for our proposed initialization step. When the rank is low enough, these bounds are significantly smaller than what existing single vector PR algorithms need. Via extensive experiments, we show that the same is true for the proposed complete algorithms as well.

25. Convolutional Phase Retrieval via Gradient Descent

Q. Qu and Y. Zhang (Columbia Uni., USA), Y. Eldar (Technion, Israel), J. Wright (Columbia Uni., USA)

We study the convolutional phase retrieval problem, which is to recover a signal from measuring the magnitude of its circulant convolution with a given kernel. This problem is motivated by applications such as channel estimation and underwater acoustic communication, where the phase information is often difficult to acquire. In this work, we assume that the convolutional kernel is random. In this setting, we show that by minimizing a nonconvex and nonsmooth objective using generalized gradient descent, with high probability, it locally converges to the optimal solution with $m \geq \Omega(n \text{ poly} \log n)$ samples. In addition, the method benefits from the benign structure of the convolutional model, which significantly reduces the computation cost by using FFT. Our analysis deals with the entry-wise dependence problem of random convolutional matrix by using ideas from decoupling theory, restricted isometry property of random circulant matrices, and recent analysis for alternating projection methods.

26. PhaseMax: Convex Phase Retrieval Without Lifting
T. Goldstein (Stanford Uni., USA), C. Studer (Cornell Uni., USA)
We consider the recovery of a (real- or complex-valued) signal from magnitude-only measurements, known as phase retrieval. We formulate phase retrieval as a convex optimization problem, which we call PhaseMax. Unlike other convex methods that use semidefinite relaxation and lift the phase retrieval problem to a higher dimension, PhaseMax operates in the original signal dimension. We show that the dual of the PhaseMax formulation is Basis Pursuit, which implies that phase retrieval can be performed using algorithms initially designed for sparse signal recovery. We develop sharp lower bounds on the success probability of PhaseMax for a broad range of random measurement ensembles, and we analyze the impact of measurement noise on the solution accuracy. We use numerical results to demonstrate the accuracy of our recovery guarantees, and we showcase the efficacy and limits of PhaseMax in practice.

27. Robust Outlier Identification for Noisy Data via Randomized Adaptive Compressive Sampling
X. Li and J. Haupt (Uni. of Minnesota, USA)
This paper examines the problem of locating outlier columns in a large, otherwise low-rank matrix, in the setting where the data are noisy. We propose a randomized two-step inference framework, and establish sufficient conditions on the required sample complexities under which these methods succeed (with high probability) in accurately locating the outliers. Numerical experimental results are provided to verify the theoretical bounds and demonstrate the computational efficiency of the proposed algorithm.

28. Robustness to Unknown Error in Sparse Regularization
S. Brugiapaglia and B. Adcock (Simon Fraser Uni., USA),
R. Archibald (Oak Ridge National Laboratory, USA)
From a numerical analysis perspective, assessing the robustness of l1-minimization is a fundamental issue in compressed sensing and sparse regularization. Yet, the recovery guarantees available in the literature usually depend on a priori estimates of the noise, which can be very hard to obtain in practice. In this work, we study the performance of l1-minimization when these estimates are not available, providing robust recovery guarantees for quadratically constrained basis pursuit and random sampling in bounded orthonormal systems. Applications of this work include approximation of high-dimensional functions, infinite-dimensional sparse regularization for inverse problems, and fast algorithms for non-Cartesian Magnetic Resonance Imaging.

29. Sparse Recovery From Superimposed Non-Linear Sensor Measurements
M. Genzel and P. Jung (Technical Uni. of Berlin, Germany)
We study the problem of sparse parameter estimation via wireless sensor networks. Such networks typically encompass a large collection of spatially distributed sensor units, acquiring individual measurements of a sparse source vector which are simultaneously transmitted to a central receiver. Since this sensing process is usually imperfect (e.g., caused by low-quality sensors), the receiver measures a superposition of non-linearly distorted signals. Based on a rigorous mathematical framework, we show that efficient sparse recovery from a very few measurements is still

feasible in this setup, using a simple Lasso estimator. Moreover, we discuss several practical implications and extensions of our approach.

30. Recovery of Nonlinearly Degraded Sparse Signals through Rational Optimization
M. Castella (Telecom SudParis, France), J.-C. Pesquet (Uni. Paris-Saclay, France)
We show the benefit which can be drawn from recent global rational optimization methods for the minimization of a regularized criterion. The regularization term is a rational Geman-McClure like potential, approximating the ℓ_0 norm and the fit term is a least-squares criterion suitable for a wide class of nonlinear degradation models.

31. Computing the Spark of a Matrix
Andreas Tillmann (RWTH Aachen Uni., Germany),
M. Pfetsch (Technical Uni. of Darmstadt, Germany)
The spark of a matrix, i.e., the smallest number of linearly dependent columns, plays an important role in sparse signal recovery and compressed sensing; for instance, it yields characterizations of unique reconstructability. We develop several approaches to tackle the NP-hard problem of computing the spark; in particular, we propose a novel branch & cut scheme based on an integer program arising from a matroid circuit formulation. The potential advantage of our algorithm compared to using general-purpose solvers (applied to a mixed-integer programming model) is demonstrated in some numerical experiments.

32. Stabilizing Embedology: Geometry-Preserving Delay-Coordinate Maps
C. Rozell (Georgia Institute of Technology, USA),
M. Wakin (Colorado School of Mines, USA),
H. Lun Yap (Georgia Institute of Technology, USA),
A. Eftekhari (The Alan Turing Institute, USA)
The seminal Takens' embedding theorem asserts that it is possible to use time-series data to reconstruct an image of the trajectory of a hidden dynamical system using a simple tool known as a delay-coordinate map. Many algorithms for time-series prediction and dimensionality reduction take inspiration from this result. Takens' theorem, however, does not speak to the quality of the reconstruction or the role that the delay-coordinate parameters play. We extend Takens' embedding theorem to stable embeddings using insight from compressed sensing, providing new insight into conditions when delay-coordinate maps can reconstruct a geometrically similar image of a trajectory.

33. Rare Eclipses in Quantised Random Embeddings of Disjoint Convex Sets: a Matter of Consistency?
V. Cambareri, C. Xu, and L. Jacques (Uni. Catholique de Louvain, Belgium)
We study the problem of verifying when two disjoint closed convex sets remain separable after the application of a quantised random embedding, as a means to ensure exact classification from the signatures produced by this non-linear dimensionality reduction. An analysis of the interplay between the embedding, its quantiser resolution and the sets' separation is presented in the form of a convex problem; this is completed by its numerical exploration in a special case, for which the phase transition corresponding to exact classification is easily computed.

34. Multilinear Low-Rank Tensors on Graphs & Applications

N. Shahid, F. Grassi, and P. Vandergheynst (EPFL, Switzerland)

Current low-rank tensor literature lacks development in large scale processing and generalization of the low-rank concepts to graphs. Motivated by the fact that the first few eigenvectors of the knn-nearest neighbors graph provide a smooth basis for the data, we propose a novel framework Multilinear Low-Rank Tensors on Graphs (MLRTG). The applications of our scalable method include approximate and fast methods for tensor compression, robust PCA, tensor completion and clustering. We specifically focus on Tensor Robust PCA on Graphs in this work.

Poster Session 1, Tuesday, June 6th

35. Compressed Learning: A Deep Neural Network Approach
A. Adler, E. Zisselman, and M. Elad (Technion, Israel)
Compressed Learning (CL) is a joint signal processing and machine learning framework for inference from a signal, using a small number of measurements. This paper presents an end-to-end deep learning approach for CL, in which a network composed of fully-connected layers followed by convolutional layers perform the linear sensing and non-linear inference stages. During training, the sensing matrix and the inference operator are *jointly* optimized, leading to a significant advantage compared to existing methods. For example, at a sensing rate of 1% (only 8 measurements of 28 x 28 pixels images), the classification error for the MNIST dataset is 6.46% compared to 41.06% with CL state-of-the-art.

36. Identifying Archetypes by Exploiting Sparsity of Convex Representations
V. Abrol, P. Sharma, and A. Sao (IIT Mandi, India)
This paper presents a computationally efficient greedy archetypal analysis (GAA) algorithm. GAA leverages the underlying sparseness property of AA, and thus is scalable to larger datasets while giving significantly faster convergence as compared to the existing algorithms. Since, extremal points have the sparsest convex representation, archetypes are identified by projecting the data in a linearly transformed/coefficient space involving sparse matrices. Here, appropriate sparse exemplars are selected by employing an iterative fast subset selection approach.

37. Characterizing the maximum parameter of the total-variation denoising through the pseudo-inverse of the divergence
C. Deledalle and N. Papadakis (IMB Bordeaux, France),
J. Salmon (TELECOM ParisTech, France), S. Vaiter (IMB Bordeaux, France)
We focus on the maximum regularization parameter for anisotropic total-variation denoising. It corresponds to the minimum value of the regularization parameter above which the solution remains constant. While this value is well know for the Lasso, such a critical value has not been investigated in details for the total-variation. Though, it is of importance when tuning the regularization parameter as it allows fixing an upper-bound on the grid for which the optimal parameter is sought. We establish a closed form expression for the one-dimensional case, as well as an upper-bound for the two-dimensional case, that appears reasonably tight in practice. This problem is directly linked to the computation of the pseudo-inverse of the divergence, which can be quickly obtained by performing convolutions in the Fourier domain.

38. Learning Fast Orthonormal Sparsifying Transforms
C. Rusu and J. Thompson (Uni. of Edinburgh, UK)
Given a dataset, the task of learning a transform that allows sparse representations of the data bears the name of dictionary learning. In many applications, these learned dictionaries represent the data much better than the static well-known transforms (Fourier, Hadamard etc.). The main downside of learned transforms is that they lack structure and therefore they are not computationally efficient, unlike their classical counterparts. This posses several difficulties especially when using power limited hardware such as mobile devices, therefore discouraging the application of sparsity techniques in such scenarios. In this paper we construct orthonormal dictionaries

that are factorized as a product of a few basic transformations. In the orthonormal case, we solve exactly the dictionary update problem for one basic transformation, which can be viewed as a generalized Givens rotation, and then propose to construct orthonormal dictionaries that are a product of these transformations, guaranteeing their fast manipulation. We show how the proposed transforms can balance very well data representation performance and computational complexity on image data. We also compare with classical fast and learned general and orthonormal transforms.

39. Accelerating the Gradient Projection Iterative Sketch for Large Scale Constrained Least-squares

J. Tang, M. Golbabaee, and M. Davies (Uni. of Edinburgh, UK)

We propose an efficient algorithm for solving the large scale constrained/regularized least-squares, based on the recently introduced iterative sketching programs, projected/proximal gradient descent and Nesterov's acceleration scheme. We observe the computational efficiency of the proposed method over the standard first order methods through numerical experiments.

40. Binary Graph-Signal Recovery from Noisy Samples

G. Eslamlou (Technical Uni. of Vienna, Austria),
N. Goertz (Vienna Uni. of Technology, Austria)

We study the problem of recovering a smooth graph signal from incomplete noisy measurements, using random sampling to choose from a subset of graph nodes. The signal recovery is formulated as a convex optimization problem. We reformulate the optimization problem in a way that the optimality conditions form a system of linear equations which is solvable via Laplacian solvers. We use an incomplete Cholesky factorization conjugate gradient (ICCG) method for graph signal recovery. Numerical experiments validate the performance of the recovery method over real-world blog-data of 2004 US election.

41. Multiview Attenuation Computation and Correction

V. Debarnot (INSA, France), J. Kahn and P. Weiss (CNRS, France)

Measuring attenuation coefficients is a fundamental problem that can be solved with diverse techniques such as X-ray or optical tomography and lidar. We propose a novel technique based on the observation of a sample from a few different angles. In principle, it can be used in existing devices such as lidar or various types of fluorescence microscopes. It is based on the resolution of a nonlinear inverse problem. We propose a specific computational approach to solve it and show the well-foundedness of the approach on simulated data. Some of the tools developed are of independent interest. In particular we propose new robust solvers for the lidar equation and an adaptation of the nonlocal-means to a specific type of heteroscedastic noise. This can be used to correct attenuation defects in images.

42. $\ell 1$-HOUDINI: A New Homotopy Method for $\ell 1$-Minimization

C. Brauer and D. Lorenz (TU Braunschweig, Germany),
A. Tillmann (RWTH Aachen, Germany)

We propose a primal-dual homotopy method for $\ell 1$-minimization problems with infinity norm constraints, where the natural homotopy parameter is the value of the bound for the constraints. Motivated by primal-dual optimality conditions, each iteration of our method decomposes into two relatively small linear programs. The effectiveness and the competitiveness of our method are demonstrated in numerical experiments.

43. ADMM Pursuit for Manifold Regularized Sparse Coding
Y. Yankelevsky and M. Elad (Technion, Israel)
We propose an efficient ADMM-based algorithm for graph regularized sparse coding that explicitly takes into account the local structure of the data. Specifically, the graph Laplacian representing the data manifold structure is used as a regularizer, encouraging the resulting sparse codes to vary smoothly along the geodesics of the manifold. By preserving locality, the obtained representations have more discriminating power and can better facilitate machine learning tasks such as classification and clustering. The experimental results demonstrate the effectiveness of our proposed algorithm over previously suggested approaches, in terms of both lower representation errors and faster, more stable runtimes.

44. Sparsity Regularized Optical Interferometric Imaging
J. Birdi, A. Repetti, and Y. Wiaux (Heriot-Watt Uni., UK)
Optical interferometry involves acquisition of under-sampled data related to the Fourier coefficients of the intensity image of interest, with missing phase information. It poses an ill-posed non-linear inverse problem for image recovery. In this context, for monochromatic imaging, a tri-linear data model was proposed in (Auria, 2014), leading to a non-negative non-linear least squares minimization problem, solved using a Gauss-Seidel method. In the recently published paper (Birdi, 2016), we have developed a new robust method to improve upon the previous approach, by introducing a sparsity prior, imposed either by an ℓ_1 or a reweighted ℓ_1 regularization term. The resulting problem is solved using an alternating forward-backward algorithm, which is applicable to both smooth and non-smooth functions, and provides convergence guarantees in the non-convex context of interest. Moreover, our method presenting a general framework, we have extended it to hyperspectral imaging, where we have promoted a joint sparsity prior by an $\ell_{2,1}$ norm. Here we describe the proposed method and present simulation results to show its performance.

45. Sparse Maximin Aggregation of Neuronal Activity
S. Mogensen, A. Lund, and N. Hansen (Uni. of Copenhagen, Denmark)
When analyzing large and inhomogeneous data sets it is of interest to obtain a robust estimate of an underlying signal. We consider a large data set describing neuronal activity in which systematic noise components are present. We propose the use of (soft) maximin aggregation and L1-penalization to obtain a robust and sparse signal from this noisy data. An approximative computational method and an exact LARS-type method giving the entire solution path are presented.

46. Cloud Dictionary: Sparse Coding and Modeling for Point Clouds
O. Litany, T. Remez, and A. Bronstein (Tel Aviv Uni., Israel)
With the development of range sensors such as LIDAR and TOF camera, 3D point cloud scans have become ubiquitous in computer vision applications, the most prominent ones being gestures recognition and autonomous driving. Parsimony based algorithms have shown great success on images and videos where data points are sampled on a Cartesian regular grid. We propose an adaptation of these techniques to irregularly samples signals by using continuous dictionaries. We present an example application in the form of point cloud denoising.

47. Cover Tree Compressed Sensing

M. Golbabaee and M. Davies (Uni. of Edinburgh, UK)

We adopt data structure in the form of cover trees and iteratively apply approximate nearest neighbour searches for fast compressed sensing reconstruction of signals living on discrete manifolds. We study the convergence of our algorithm within a more general class of inexact gradient projection methods designated to solve constrained least-square problems. We show that the projection cost is mainly determined by the intrinsic dimension of the manifold and grows sub-linearly with its population. We apply our results to quantitative MRI compressed sensing and in particular within the Magnetic Resonance Fingerprinting framework. For a similar reconstruction accuracy, we report 2-3 orders of magnitude reduction in computations compared to the standard iterative method using brute-force searches.

48. Structured Sparse Modelling with Hierarchical GP

D. Kuzin, O. Isupova, L. Mihaylova (Uni. of Sheffield, UK)

In this paper a new Bayesian model for sparse linear regression with a spatio-temporal structure is proposed. It incorporates the structural assumptions based on a hierarchical Gaussian process prior for spike and slab coefficients. We design an inference algorithm based on Expectation Propagation and evaluate the model over the real data.

49. An ODE-based Modeling of Inertial Forward-Backward Algorithms

V. Apidopoulos, J.-F. Aujol, C. Dossal (Uni. de Bordeaux, France)

In this work we are interested in the behaviour of the trajectory of solutions of a differential equation, derived by a discrete scheme which corresponds to a particular inertial Forward-Backward (iFB) algorithm considered in [3]. More precisely, under some appropriate hypotheses, the convergence rates of the values of the functional and the velocity norm of the trajectory are of the same order to the ones obtained in [3] of their “discretized versions”.

50. Convex Optimisation for Partial Volume Estimation in Compressive Quantitative MRI

R. Duarte and Z. Chen (Heriot Watt Uni., UK), S. Gazzola (Uni. of Bath, UK), I. Marshall and M.Davies (Uni. of Edinburgh, UK), Y. Wiaux (Heriot-Watt Uni., UK)

Based on the recently proposed compressive sensing framework for quantitative MRI, a new approach for partial volume reconstruction is developed in this abstract. We first formulate a convex optimisation problem for the recovery of a sparse matrix of coefficients in a dictionary of measured temporal fingerprints associated with specific combinations of quantitative parameters of interest. Each column of the sought matrix represents a voxel in the volume under investigation and the sparsity of this column represents the number of active dictionary elements or partial volumes. In a second step, we employ the weighted k-means algorithm to cluster the recovered coefficient matrix in parameter space and obtain the quantitative parameter maps. The proposed approach was validated through simulations, and its performance is competitive when compared to a state of the art algorithm.

51. Finite-Valued Sparse Signals

S. Keiper (Uni. of Berlin, UK), G. Kutyniok (TU Berlin, Germany), G. Pfander and D. Lee (Uni. of Marbur, Germany)

The need of reconstructing discrete-valued sparse signals from few measurements appears frequently in science and engineering. Whereas classical compressed sensing algorithms do not incorporate the additional knowledge of the discrete nature of the signal, classical lattice decoding approaches do not utilize sparsity constraints. In this talk, we present an approach that incorporates a discrete values prior into basis pursuit. In particular, we address finite-valued sparse signals, i.e., sparse signals with entries in a finite alphabet. We will introduce an equivalent null space characterization and show that phase transition takes place earlier than when using the classical basis pursuit approach. We will further discuss robustness of the algorithm.

52. Infimal Convolution Type Coupling of First and Second Order Differences on Manifold-Valued Images

G. Steidl (Uni. of Kaiserslautern, Germany)

Infimal convolution type functions were successfully applied in regularization terms of variational models for restoring and decomposing images. In this paper, we generalize the infimal convolution of first and second order differences to manifold-valued images in a intrinsic way, i.e. without embedding the manifold into an Euclidean space. We apply a gradient descent reprojection algorithm to find a critical point of the corresponding functional and demonstrate the approach by numerical examples on the 2-sphere and the manifold of positive definite matrices with the affine invariant metric.

53. On the Difficulty of Selecting Ising Models with Approximate Recovery

J. Scarlett and V. Cevher (EPFL, Switzerland)

We consider the problem of estimating the underlying graph associated with an Ising model given a number of independent and identically distributed samples. We adopt an *approximate recovery* criterion that allows for a number of missed edges or incorrectly-included edges, in contrast with the widely-studied exact recovery problem. Our main results provide information-theoretic lower bounds on the sample complexity for graphs having constraints on the number of edges and maximal degree. We identify a range of scenarios where, either up to constant factors or logarithmic factors, our lower bounds match the best known lower bounds for the exact recovery criterion, several of which are known to be tight or near-tight. Hence, in these cases, approximate recovery has a similar difficulty to exact recovery in the minimax sense.

54. Adaptive Orthogonal Basis Pursuit for Volumetric Two-Photon Microscopy

A. Charles (Princeton Uni., USA), A. Song, S. Koay, J. Gauthier, S. Thiberge, D. Tank, J. Pillow (Princeton Neuroscience Institute, USA)

Greedy algorithms have been used extensively in statistical inference and signal processing to provide fast approximate solutions to complex inference problems. One of the most prevalent greedy methods is the Matching Pursuit (MP) family of algorithms (e.g. Orthogonal Matching Pursuit; OMP) for solving sparsity-regularized least-squares optimization programs. Essentially, MP algorithms seek to describe a measured signal in terms of a sparse linear combination of the columns of a matrix A by greedily choosing the columns of A most correlated with the measurements. Variants of the MP algorithm choose the new column(s) to add and construct the residuals in different ways, however they all rely on knowing the matrix A. In some applications, precise information about A is not available, however approximate information is known. In such cases, access to multiple measurement vectors with the same decomposition can compensate for this deficiency. In this work we extend the

MP framework to cover the case where the true A is only approximately modeled. Specifically, we introduce an additional shape-projection step into the OMP algorithm that adapts the idealized templates to the measurement vectors. The SCISM algorithm was motivated by the application of fluorescence two-photon microscopy (TPM) for calcium imaging. In TPM, hundreds of neurons are simultaneously imaged. To extract neural activity, the neural profiles (location and shapes) must be found. While neural shapes can be stereotyped as annuli, the actual shapes may vary widely. We show that the SCISM algorithm can be used with novel optics for stereoscopic TPM in order to image entire neural volumes in-vivo at biologically relevant time-scales and extract the neurla profiles and time traces. The SCISM algorithm permits OMP to adapt stereotypical neural shapes to the actual shapes of neurons in the tissue, providing 3D neural locations and corresponding neural activity for neurons in V1 and CA1 in awake mice.

55. An Iterative Convex Optimization Solver with Side Information for Joint-Sparse Signal Recovery

S.-W. Hu, S.-H. Hsieh, C.-S. Lu (Academia Sinica, Taiwan),
G.-X. Lin (National Cheng-Kung Uni., Taiwan)

We propose an iterative joint-sparse signal recovery algorithm for compressive sensing with multiple measurement vectors. Our algorithm is an iterative procedure to solve $L_{2,1}$-norm minimization with side information that is self-produced from the reconstructed signals at previous iterations instead of being given in advance. Our side information is designed with a theoretical foundation to further reduce the required number of measurements for successful recovery.

56. Image Generation Using a Sparsity Model

G. Vaksman and M. Elad (Technion, Israel)

In this work we present a novel method for generating images using sparse representations. The proposed method learns a dictionary and estimates the probability distribution of the sparse vectors from a given set of images. The probability distribution of the sparse vectors is represented as a product of two functions: one describing the support, and the other for the coefficients, given the support. We model the distribution of the support using a Markov network with pairwise factors, and assume that the atom coefficients, given the support, are normally distributed. We generate new images by drawing sparse vectors from the estimated probability and multiplying them by the pre-learned dictionary. We demonstrate the proposed method on three different sets of images: (i) MNIST, (ii) set of aligned faces, and (iii) set of unaligned faces from the MegaFace database.

57. Berhu Penalty for Matrix and Tensor Estimation

M. Pontil (Istituto Italiano di Tecnologia, Italy; Uni. College London, UK),
G. Denevi (Istituto Italiano di Tecnologia, Italy; Uni. degli studi di Genova, Italy),
Michele Donini (Istituto Italiano di Tecnologia, Italy)

We present a regularizer for learning low rank matrices. It is obtained by applying the Berhu penalty function to the spectrum. We link the regularizer to previous ones for structured sparsity and derive its proximity operator. In numerical experiments the spectral Berhu performs favourably over the standard methods. We discuss how the regularizer can be extended to the tensor setting.

58. Sparsity and Low-Rank Amplitude Based Blind Source Separation
F. Feng (Uni. Paris-Sud, France),
M. Kowalski (SUPELEC, CNRS, Uni. Paris-Sud, France)
We present a regularizer for learning low rank matrices. It is obtained by applying the Berhu penalty function to the spectrum. We link the regularizer to previous ones for structured sparsity and derive its proximity operator. In numerical experiments the spectral Berhu performs favourably over the standard methods. We discuss how the regularizer can be extended to the tensor setting.

59. Support Recovery Guarantees for Group Lasso Estimator
M. K. Elyaderani, S. Jain, J. Haupt, J. Druce,
and S. Gonella (Uni. of Minnesota, USA)
We consider the problem of estimating a high dimensional signal from noisy low-dimensional linear measurements, where the desired unknown signal exhibits a group-sparse structure. Assuming the non-zero groups of the group-sparse signal possess enough strength and are generated according to certain statistical assumptions, we provide conditions to guarantee that the signal support can be exactly recovered via solving the group Lasso problem.

60. A Kaczmarz Method for Low Rank Matrix Recovery
H. Mansour and U. Kamilov (MERL, USA),
O. Yilmaz (Uni. of British Columbia, Canada)
The Kaczmarz method was initially proposed as a row-based technique for reconstructing signals by finding the solutions to overdetermined linear systems. We propose a weighted Kaczmarz method that can recover low rank matrices from linear measurements both in the overdetermined and underdetermined regimes. We demonstrate using numerical simulations the effectiveness of our technique in recovering low rank signals and compare its performance to the linearized Bregman Kaczmarz method as well as the standard Kaczmarz method.

61. Complex-valued Deterministic Matrices with Low Coherence based on Algebraic Geometric Codes
H. Abin and A. Amini (Sharif Uni. of Technology, Iran)
In this work, we introduce new algebraic geometry (AG) curves that can generate extremely fat matrices with low coherence. The previous application of AG codes in matrix design has been limited to binary matrices. Here, we devise a different approach to achieve $m \times n$ complex-valued matrices. As $n > m2$ in our matrices, the Welch bound is no longer achievable; however, the coherence of our matrices surpass the Welch bound only by a O(log m) factor. Moreover, our construction provides flexibility in setting the number of rows and columns.

62. Online Convex Optimization Meets Sparsity
S. Fosson and E. Magli (Politecnico di Torino, Italy), J. Matamoros
and M. Gregori (Centre Tecnologic de Telecomunicacions de Catalunya, Spain)
Tracking time-varying sparse signals is a novel problem, with broad applications. Techniques merging compressed sensing and Kalman filtering have been proposed in the related literature, which typically rely on specific dynamic models. In this work, we propose a new perspective on the problem, based on elements of online convex optimization. In particular, we design a suitable optimization problem and develop algorithms which do not assume any specific dynamic model. For these algorithms,

we analytically evaluate the behaviors of their dynamic regrets that serve as their performance measure.

63. Human Action Attribute Learning Using Low-Rank Representations
T. Wu, P. Gurram, R. Rao, and W. Bajwa (Rudgers Uni., USA)
This paper studies the problem of learning human action attributes based on union-of-subspaces model. It puts forth an extension of the low-rank representation (LRR) model, termed the hierarchical clustering-aware structure-constrained low-rank representation (HCS-LRR) model, for unsupervised learning of human action attributes from video data. The effectiveness of the proposed model is demonstrated through experiments on five human action datasets for action recognition.

64. A Low-Rank Approach to Off-the-Grid Sparse Deconvolution
P. Catala (DMA, ENS, France), V. Duval (INRIA Rocquencourt, France), G. Peyré (CNRS, DMA, ENS, France)
We propose a new solver for the sparse spikes deconvolution problem over the space of Radon measures. A common approach to off-the-grid deconvoluton considers semidefinite relaxations of the total variation minimization problem. The direct resolution of this SDP is however intractable for large scale settings, since the problem size grows as fc^2 where fc is the cutoff frequency of the filter. Our first contribution introduces an unconstrained dualization of this semidefinite lifting, which has low-rank solutions. Our second contribution is a conditional gradient optimization scheme with non-convex updates. This algorithm leverages both the low-rank and the convolutive structure of the involved variable, resulting in an O(fc log(fc)) complexity per iterations. Our numerical simulations are promising and show that this algorithm converges in exactly k steps, where k is the number of Diracs composing the solution.

65. Audio Source Separation with Deep Neural Networks Using the Dropout Algorithm
A. Zermini, Y. Xu, and Q. Kong, M. Plumbley, W. Wang (Uni. of Surrey, UK)
A method based on Deep Neural Networks (DNNs) and time-frequency masking has been recently developed for binaural audio source separation. In this method, the DNNs are used to predict the Direction Of Arrival (DOA) of the audio sources with respect to the listener which is then used to generate soft time-frequency masks for the recovery/estimation of the individual audio sources. In this paper, an algorithm called 'dropout' will be applied to the hidden layers, affecting the sparsity of hidden units activations: randomly selected neurons and their connections are dropped during the training phase, preventing feature co-adaptation. These methods are evaluated on binaural mixtures generated with Binaural Room Impulse Responses (BRIRs), accounting a certain level of room reverberation. The results show that the proposed DNNs system with randomly removed neurons is able to achieve higher SDRs performances compared to the baseline method without the dropout algorithm.

66. Patch-based Interferometric Phase Estimation via Mixture of Gaussian Density Modelling in the Complex Domain
J. Krishnan and J. Bioucas-Dias (Instituto de Telecomunicações, Instituto Superior Técnico, Portugal)
This paper addresses interferometric phase (InPhase) image denoising; that is, the denoising of phase modulo-2π images from sinusoidal 2π-periodic and noisy ob-

servations. The wrapping discontinuities present in the InPhase images, which are to be preserved carefully, make InPhase denoising a challenging inverse problem. We tackle this problem by exploiting the self-similarity of the InPhase images. We propose a novel approach to address the problem by modelling the patches of the phase images using Mixture of Gaussian (MoG) densities in the complex domain. An Expectation Maximization (EM) algorithm is formulated to learn the parameters of the MoG from the noisy data. The learned MoG is used as a prior for estimating the InPhase images from the noisy images using Minimum Mean Square Error (MMSE) estimation. The experiments conducted on simulated and real data of InSAR/InSAS shows results which are competitive with the state-of-the-art techniques

67. Sampling from Binary Measurements - On Reconstructions from Walsh Coefficients

L. Terhaar and A. Hansen (Uni. of Cambridge, UK)

High quality reconstructions from a small amount of measurements offers interesting possibilities in a lot of applications. This is possible due to techniques such as generalized sampling and methods based on data assimilation and additional information modelled by PDEs. The reconstruction quality of both methods depends highly on the subspace angle between the sampling and the reconstruction space. We show that for binary measurements and wavelets, the relation between the amount of data sampled and the coefficients reconstructed has to be only linear to ensure that the angle is bounded from below and hence the reconstruction is accurate and stable.

68. Metric Learning for Tracking a Disease Progress

N. Cvetkovic and T. Conrad (Freie Universitat Berlin, Germany)

A current challenge in medicine is using high-dimensional clinical data to understand the transition of patients between long-lived stages of a disease, e.g. ‘healthy’ or ‘sick’. We apply feature selection and metric learning techniques to such data in order to construct Markov chains on sparse connected networks, where long-lived stages are modeled by clusters of nodes on the network in which the chain spends long times. Using transition path theory, we can characterise the statistics of transitions and identify the most probable transition paths between long-lived stages.

69. Recovery of Sparse Heat Source Signals from Locally Differentially Private Sensor Data via Constrained ℓ1 Minimisation

A. McMillan and A. Gilbert (Uni. of Michigan, USA)

In this paper we consider sparse signal recovery from differentially private measurement data. We produce a differentially private version of the measurement data for the discrete one dimensional heat source problem and explore what information about the sources is captured in this private data. We show it is possible to use the private data to recover an estimate to the initial source locations that is close in the earth mover distance to the true source location by studying the behaviour of the EMD error when constrained ℓ_1 minimisation is used to recover the source locations from the private data. Our work indicates that it is possible to produce locally private sensor measurements that both keep the exact locations of the heat sources private and permit recovery of the “general vicinity” of the sources.

70. Sparse Signal Recovery via Correlated Degradation Model

N. Eslahi (Tampere Uni. of Technology, Finland),
V. Ramakrishnan (Universidad de Concepcion, Chile),
K. Wiik (Uni. of Turku, Finland), A. Foi (Tampere Uni. of Technology, Finland)

Sparse signal recovery aims to recover an unknown signal from few non-adaptive, possibly noisy, linear measurements using a nonlinear sparsity-promoting algorithm, under the assumption that the signal is sparse or compressible with respect to a known basis or frame. Most recovery approaches, such as Approximate Message Passing or Plug & Play-Priors can be regarded as iterative estimation of a signal from a degraded observation, which is commonly modeled as independent and identically distributed (i.i.d.) additive noise that has to be alleviated by a denoising filter at each iteration. We focus on the modeling of such degradations as spatially (or spatiotemporally) correlated noise. Our motivation is that the common assumption of i.i.d. noise is valid only under special conditions that are hardly met in practice. In contrast to i.i.d. noise, correlated noise can lead to disproportions in the magnitude of errors across the spectrum, to an extent that i.i.d. denoisers may not effectively discern between the true signal and noise in regularization via shrinkage, potentially leading to ineffective filtering and distortion of the signal. We model the correlation through the power spectral densities (PSD) with respect to the sparsifying transform used internally by the denoiser: these PSDs modulate the shrinkage thresholds, i.e. allow to compare the magnitude of each transform coefficient against that of the corrupting noise. We perform a robust PSD fitting at every iteration of the recovery algorithm. This PSD (which often features order-of-magnitude differences across the spectrum and noticeable anisotropy) constitutes an adaptive estimate of the noise correlation which the filter can utilize. We demonstrate that our modeling yields a better detection and sharper recovery of fine structures across a wide range of practical sparse recovery applications, including volumetric medical image reconstruction, multi-epoch radio-interferometry, and compressive video encoding.

71. Alternating Group Sparsity for Image Restoration

K. Egiazarian and V. Katkovnik (Tampere Uni. of Technology, Finland)

In this paper, alternating application of specially designed nonlocal collaborative filters in iterative image restoration is proposed. This work is in line with the current trends in image processing, that the accurate formulation of the prior can be omitted in favor of a good denoising algorithm embedded in the iterations. We prove that indeed the used priors results in the efficient filters and, even more, alternative application of these filters leads to the extraordinary good results.

72. Compressed Sensing of FRI Signals using Annihilating Filter-based Low-rank Interpolation

J.-C. Ye and J. M. Kim (KAIST, Korea), K. H. Kim (EPFL, Switzerland),
K. Lee (Georgia Tech, USA)

Here, we propose a novel two-step Fourier compressive sampling framework for the recovery of finite rate of innovations (FRI) signals in a continuous domain. The algorithm can be implemented as measurement domain interpolation, after which signal reconstruction can be carried out by means of classical analytic reconstruction methods. The main idea comes from the fundamental duality between the sparsity in the primary space and the low-rankness of a structured matrix in the spectral domain, which shows that a low-rank interpolator in the spectral domain can enjoy all of the advantages of sparse recovery with performance guarantees. Using a powerful dual certificate and the golfing scheme, we show that the new framework still achieves a near-optimal sampling rate for a general class of FRI signal recovery, while the sampling rate can be further reduced for a class of cardinal splines.

Poster Session 2, Wednesday, June 7

73. Sparse Denoising: Aggregation Versus Global Optimization

D. Carrera and G. Boracchi (Politecnico di Milano, Italy),
A. Foi (Tampere Uni. of Technology, Finland),
B. Wohlberg (Los Alamos National Laboratory, USA)

Denoising is often addressed via sparse coding with respect to an overcomplete translation-invariant dictionary. There are two main approaches for dictionaries composed of translates of an orthonormal basis. The classical approach is cycle spinning, which aggregates partial estimates, each of which is sparse with respect to a different shift of the orthonormal basis. An alternative is offered by convolutional sparse representations, which perform a global optimization over the entire dictionary. It is tempting to view the former approach as providing a suboptimal solution of the latter. Here we compare the two approaches and show that, while the global optimization produces estimates with lower bias than the corresponding aggregation procedure, these are also characterized by a higher variance. In practice, the computationally demanding global optimization outperforms the simpler aggregation of partial estimates only when images admit an extremely sparse representation w.r.t. the dictionary, while they perform similarly on natural images.

74. Identifying Impedances of Walls Using First and Second Order Echoes

H. Peic Tukuljac, H. Lissek, and P. Vandergheynst (EPFL, Switzerland)

Control of the sound field in a room requires detailed characterization of acoustical properties of the room, such as room shape and physical characteristics of the walls. We will focus on the estimation of the acoustic impedances of the walls in a room using room impulse response and image-source model which eliminates the limitation of rectangular room geometry as well as narrow-band low-frequency requirement imposed by many solutions. Previous solutions rely mainly on the finite difference methods that become quite cumbersome when we move to 3D.

75. Reconstructing Signals from a Union of Linear Subspaces Using a Generalized CoSaMP

T. Tirer and R. Giryes (Tel Aviv Uni., Israel)

The idea that signals reside in a union of low dimensional subspaces subsumes many low dimensional models that have been used extensively in the recent decade in many fields and applications. Until recently, the vast majority of works have studied each one of these models on its own. However, a recent approach suggests providing general theory for low dimensional models using their Gaussian mean width, which serves as a measure for the intrinsic low dimensionality of the data. In this work we use this novel approach to study a generalized version of the popular compressive sampling matching pursuit (CoSaMP) algorithm, and to provide general recovery guarantees for signals from a union of low dimensional linear subspaces, under the assumption that the measurement matrix is Gaussian. We discuss the implications of our results for specific models, and use the generalized algorithm as an inspiration for a new greedy method for signal reconstruction in a combined sparse-synthesis and cosparse-analysis model. We perform experiments that demonstrate the usefulness of the proposed strategy.

76. Compressive Spectrometer

C. C. Lu and H.-C. Chen (Industrial Technology Research Institute, Taiwan), H. T. Kung (Harvard Uni., USA)

Spectrometers are widely used for characterizing materials. Recently, filter-based spectrometers have been proposed to lower the manufacturing cost by replacing optical components with low-cost wavelength-selective filters, but at the expense of possibly lowered signal quality. We present compressive spectrometers which, based on the compressive sensing principle, are able to recover signal with improved quality from measurements acquired by a relatively small number of low-cost filters. We achieve high quality recovery by leveraging the fact that spectrometer measurements typically follow the shape of a smooth curve with a few spikes. We validate our method with real-world measurements, and release our dataset to facilitate future research.

77. Gap Safe Screening Rules for Faster Complex-valued Multi-task Group Lasso

M. Massias (Télécom ParisTech, Uni. Paris-Saclay, France), A. Gramfort and J.Salmon (Télécom ParisTech, France)

Linear regression with sparsity-inducing penalties is a popular tool for high-dimensional inverse problems such as source localization or denoising. The Group Lasso is a particular choice of convex penalty that considers the $\ell_{2,1}$ norm to promote group (or block) sparsity patterns. Since no general closed-form solution is available for this problem, iterative solvers are needed. This can lead to very slow convergence especially if the problem is ill-conditioned or if the dimension of the problem is particularly large. Safe screening rules and in particular dynamic ones speed-up the optimization process by progressively discarding regressors identified as irrelevant. In this work, we consider the case where features and observations can be complex-valued, a common case in signal processing when working with time-frequency operators. We derive Gap Safe screening rules in this context and propose a block coordinate descent (BCD) optimization strategy. In practice, we illustrate significant speed-ups in terms of convergence compared to classical solvers on a neuroscience problem, namely the problem of source localization using magneto and electroencephalography (M/EEG).

78. Block-GMCA

C. Kervazo and J. Bobin (CEA Saclay, France)

Blind Source Separation (BSS) is a powerful method to analyze multichannel data in fields that involve processing large-scale data. However, standard methods fail at correctly tackling BSS problems when the number of sources becomes large, especially when the number of available samples is low. Building upon a standard BSS algorithm, namely GMCA (Generalized Morphological Component Analysis), we propose investigating the performances of block- coordinate optimization strategies to tackle sparse BSS problems in the large-scale regime. Preliminary results reveal that the proposed approach, the block-GMCA algorithm, significantly improves the performances of the standard GMCA algorithm.

79. Compressed Dictionary Learning

F. Teixeira and K. Schnass (Uni. of Innsbruck, Austria)

Low complexity models of high-dimensional data lie at heart of many efficient solutions in modern signal processing. One such model is that of sparsity in dictionary. A fundamental question associated with the sparse model is how to find a suitable dictionary providing sparse representations. In this paper we introduce the Iterative Compressed-Thresholding and K-Means (IcTKM) algorithm for fast dictionary learning. The algorithm is based on a fundamental dimensionality-reduction result due to Johnson and Lindenstrauss. We show that the sample complexity of IcTKM is essentially the same as that of ITKM, while the best achievable error is increased and the convergence radius is reduced; However increasing the minimally achievable error is largely negligible for high-dimensional data, since the realistically achievable error is determined by the sample size. The reduction of convergence radius is somewhat more disappointing, but as we show in our numerical experiments on large data sets, in practice this does not affect the good global convergence behavior and reduced computational complexity of IcTKM.

80. Slice Inverse Regression witrh Score Functions
D. Babichev and F. Bach (INRIA, ENS, France)
We consider non-linear regression problems where we assume that the response depends non-linearly on a linear projection of the covariates. We propose score function extensions to sliced inverse regression problems, both for the first- order and second-order score functions. We show that they provably improve estimation in the population case over the non-sliced versions and we study finite sample estimators and their consistency given the exact score functions. We also propose to learn the score function as well, in two steps, i.e., first learning the score function and then learning the effective dimension reduction space, or directly, by solving a convex optimization problem regularized by the nuclear norm. We illustrate our results on a series of experiments.

81. Theoretical Analysis of PCA for Heteroscedastic Data
D. Hong, L. Balzano, and J. Fessler (Uni. of Michigan, USA)
Principal Component Analysis (PCA) is a classical method for estimating a subspace given noisy samples. It is useful in a variety of applications and problems ranging from dimensionality reduction to anomaly detection and the visualization of high dimensional data. Effective use of PCA requires a rigorous understanding of its performance. This work analyzes PCA for samples with heteroscedastic noise, that is, samples that have non-uniform noise variances. In particular, we provide a simple asymptotic prediction of the recovery of a low-dimensional subspace basis from noisy heteroscedastic samples. The prediction enables: (a) easy and efficient calculation of the asymptotic performance, (b) reasoning about the asymptotic performance (with heteroscedasticity such as outliers), and (c) a deeper understanding that PCA has best performance when the noise is homoscedastic (all points share the same noise level).

82. Image Restoration via Successive Compression
Y. Dar, A. Bruckstein, and M. Elad (Technion, Israel)
In this paper we propose a method for solving various imaging inverse problems via complexity regularization that leverages existing image compression techniques. Lossy compression has already been proposed in the past for Gaussian denoising – the simplest inverse problem. However, extending this approach to more complicated inverse problems (e.g., deblurring, inpainting, etc.) seemed to result in intractable optimization tasks. In this work we address this difficulty by decomposing the complicated optimization problem via the Half Quadratic Splitting approach, resulting in a sequential solution of a simpler L2-regularized inverse problem followed

by a rate-distortion optimization, replaced by an efficient compression technique. We demonstrate the proposed scheme for inpainting of corrupted images, using leading image compression techniques such as JPEG2000 and HEVC.

83. Regularized Residual Quantization: a Multi-layer Sparse Dictionary Learning Approach

S. Ferdowsi, S. Voloshynovskiy, D. Kostadinov (Uni. of Geneva, Switzerland)

The Residual Quantization (RQ) framework is revisited where the quantization distortion is being successively reduced in multi-layers. Inspired by the reverse-water-filling paradigm in rate-distortion theory, an efficient regularization on the variances of the codewords is introduced which allows to extend the RQ for very large numbers of layers and also for high dimensional data, without getting over-trained. The proposed Regularized Residual Quantization (RRQ) results in multi-layer dictionaries which are additionally sparse, thanks to the soft-thresholding nature of the regularization when applied to variance-decaying data which can arise from de-correlating transformations applied to correlated data. Furthermore, we also propose a general-purpose pre-processing for natural images which makes them suitable for such quantization. The RRQ framework is first tested on synthetic variance-decaying data to show its efficiency in quantization of high-dimensional data. Next, we use the RRQ in super-resolution of a database of facial images where it is shown that low-resolution facial images from the test set quantized with codebooks trained on high-resolution images from the training set show relevant high-frequency content when reconstructed with those codebooks.

84. Synthesis Sparse Modeling: Application to Image Compression and Image Error Concealment

A. Akbari and M. Trocan (Institut Superieur d'Electronique de Paris, France),
B. Granado (Pierre et Marie Curie Uni., France)

Signal models are a cornerstone of contemporary signal and image processing methodology. Two particular signal modeling methods, called analysis and synthesis sparse representation have been proven to be effective for many signals, such as natural images, and successfully used in a wide range of applications. Both models represent signals in terms of linear combinations of an underlying set, called dictionary, of elementary signals known as atoms. The driving force behind both models is sparsity of the representation coefficients. On the other hands, the dictionary choice determines the success of the entire model. According to these two signal models, there have been two main disciplines of dictionary designing; harmonic analysis approach and machine learning methodology. The former leads to designing the dictionaries with easy and fast implementation, while the latter provides a simple and expressive structure for designing adaptable and efficient dictionaries. The line of research followed in this report is the synthesis-based sparse representation approach in the sense that the dictionary is not fixed and predefined, but learned from training data and adapted to data, yielding a more compact representation. We report recent and novel research results of two particular applications of this signal modeling: image compression and image error concealment.

85. Fusion of Sparse Reconstructions
W. Meiniel (Telecom ParisTech, Uni. Paris-Saclay, France),
J.-C. Christophe Olivo-Marin (Institut Pasteur, France),
E. Angelini (Telecom ParisTech, Uni. Paris-Saclay, France)
In this work, we present several variations of the fusion of sparsity-based reconstructions. The method exploits TV-sparsity to obtain multiple estimators of the original image, that are then aggregated using specific strategies. We tested the technique for denoising of microscopic images.

86. Compression of Multiple Input Streams into Recursive Neural Networks
A. Charles (Princeton Uni., USA), D. Yin (Uni. of California, Berkeley, USA),
C. Rozell (Georgia Institute of Technology, USA)
Recursive neural networks (RNNs) are becoming an increasingly important part of the machine learning toolbox in applications such as video, audio, EEG data. These networks have been used as either stand-alone tools for training classifiers, or as in layers in conventional deep neural networks to expand their use to time-varying data. As the abilities of RNNs to process temporal data are often attributed to RNNs accumulating information from input data over time into the network nodes (the STM), we analyze here the ability of RNNs to store inputs in the network state. In particular, since many real-world signals have low-dimensional representations, we study the STM of RNNs when the inputs are either sparse in a basis (e.g. audio/video), or where the input vectors are correlated such that they form low-rank martix (e.g. two-photon microscopy). This work leverages the tools developed in the compressive sensing literature to develop a theoretical understanding of RNNs, an important tool in machine learning. Our results demonstrate that RNNs are very efficient in compressing long input streams into the network state. As opposed to traditional results that bound the compression rate (how many inputs can be recovered from M nodes) by the number of nodes, we show that for both low-rank and sparse inputs, the compression rate is proportional to the underlying dimension of the inputs (sparsity or rank), and only poly-logarithmic with the total number of inputs (LN). In terms of machine-learning tasks, this means that RNNs operating on structured, dynamic signals have access to long extents of the data history to make classification or prediction decisions.

87. Joint Multichannel Deconvolution and Blind Source Separation
M. Jiang, J. Bobin, and J.-L. Starck (CEA Saclay, France)
In the real world, current Blind Source Separation (BSS) methods are limited since extra instrumental effects like blurring have not been taken into account. Therefore, a more rigorous BSS has to be solved jointly with a deconvolution problem, yielding a new inverse problem: deconvolution BSS (DBSS). We introduce an innovative DBSS approach, called DecGMCA, which is based on sparse signal modeling and an efficient alternative projected least square algorithm. Numerical results demonstrate the performance of DecGMCA and highlight the advantage of jointly solving BSS and deconvolution instead of considering these two problems independently.

88. Joint Multicontrast MRI Reconstruction
L. Weizman (Technion, Israel), J. Mota (Heriot-Watt Uni., UK), P. Song (Uni. College London, UK), Y. Eldar (Technion, Israel), M. Rodrigues (Uni. College London, UK)
Joint reconstruction is relevant for a variety of medical imaging applications, where multiple images are acquired in parallel or within a single scanning procedure. Examples include joint reconstruction of different medical imaging modalities (e.g. CT and PET) and various MRI applications (e.g. different MR imaging contrasts of the same patient). In this paper we present an approach for joint reconstruction of two MR images, based on partial sampling of both. We assume each MR image has a limited number of edges, that is, low total variation, but they are similar in the sense that many of the edges overlap. We examine synthetic phantoms representing T1 and T2 imaging contrasts and realistic T1-weighted and T2-weighted images of the same patient. We show that our joint reconstruction approach outperforms conventional TV-based MRI reconstruction for each image solely. Results are shown both visually and numerically for sampling ratios of 4%-20%, and consist of an improvement of up to 3.6dB.

89. A Compressed Sensing Approach for Ultrasound Imaging
A. Besson (EPFL, Switzerland), R. Carrillo (CSEM, Switzerland), D. Perdios and M. Arditi (EPFL, Switzerland), Y. Wiaux (Heriot-Watt Uni., UK), J.-P. Thiran (EPFL, Switzerland)
Ultrasonography uses multiple piezoelectric element probes to image tissues. Current time-domain beamforming techniques require the signal at each transducer-element to be sampled at a rate higher than the Nyquist criterion, resulting in an extensive amount of data to be received, stored and processed. In this work, we propose to exploit sparsity of the signal received at each transducer-element. The proposed approach uses multiple compressive multiplexers for signal encoding and solves an ℓ_1-minimization in the decoding step, resulting in the reduction of 75 % of the amount of data, the number of cables and the number of analog-to-digital converters required to perform high quality reconstruction.

90. Correlation-Based Super-Resolution Imaging in Microscopy and Ultrasound
O. Dicker, O. Solomon, M. Mutzafi, A. Bar-Zion, M. Segev, Y. Eldar (Technion, Israel)
In this work, we demonstrate super-resolution in two traditionally diffraction limited imaging modalities: fluorescence microscopy and contrast enhanced ultrasound imaging. In both cases we exploit the statistical structure of blinking emitters over time, along side their sparse distribution within the imaged object. Our method enables sub-diffraction resolution with high temporal resolution, dictated by the imaging rate of the camera and ultrasound probe, respectively.

91. Atomic Norm Minimization for Modal Analysis from Compressive Measurements
S. Li, M. Wakin, and G. Tang (Colorado School of Mines, USA)
One technique for structural health monitoring involves estimating a structure's mode shapes and frequencies from vibrational data. To save energy, it is desirable to reduce the dimension of data to be transmitted. We consider scenarios where data is compressed at each sensor via random matrix multiplication. We show that atomic norm minimization can perfectly recover modal parameters and provide new theoretical analysis on the sample complexity of this scheme. In particular, our the-

ory does not require randomness of the mode shapes, and it shows that the sample complexity per sensor will decrease as the the number of sensors increases.

92. SNIPE for Memory-Limited PCA From Incomplete Data: From Failure to Success
A. Eftekhari (The Alan Turing Institute, UK), L. Balzano (Uni. of Michigan, USA), M. Wakin and D. Yang (Colorado School of Mines, USA)
We consider partially observed data vectors drawn from an unknown subspace, presented sequentially to the user who can only store small amounts of data. We are interested in developing a streaming algorithm for PCA from these incomplete measurements. We begin by discussing the drawbacks of a simple averaging of the subspaces spanned by zero-filled data blocks; this is a biased estimator and we bound the amount of bias. We then present SNIPE, a sequential algorithm that uses previous subspace estimates to impute missing data elements. SNIPE converges, globally and linearly, to the true subspace under reasonable requirements.

93. Joint Sparsity and "SPID" Calculation of the Stationary Wavelet Transform for Compressed Sensing Reconstruction in Parallel MRI
E. Shimron (Technion, Israel), A. Webb (Leiden Uni., The Netherlands), H. Azhari (Technion, Israel)
A fast Compressed Sensing (CS) reconstruction scheme for parallel MRI is proposed, utilizing the hybrid "SPID" technique. The proposed method utilizes SPID for calculating the Stationary Wavelet Transform (SWT) of the unknown MR image from the undersampled k-space data; subsequently, it recovers the image through a CS process that promotes joint sparsity of the multi-coils data. In-vivo experiments show that this method eliminates artifacts and expedites the CS convergence.

94. Robust Compressed Sensing with Side Information Based on Laplace Mixtures Models
C. Ravazzi (IEIIT, National Research Council, Italy), Enrico Magli (Politecnico di Torino, Italy)
In this paper, we propose a new method for the recovery of a sparse signal from few linear measurements using a reference signal as side information. Modeling the signal coefficients with a double Laplace mixture model, and assuming that the distribution of the components of the prior information differs slightly from the unknown signal, the problem is formulated as a weighted ℓ_1 minimization problem. We derive sufficient conditions for perfect recovery and we show that our method is able to reduce significantly the number of measurements required for reconstruction. Numerical experiments demonstrate that the proposed approach outperforms the best algorithms for compressed sensing with prior information and is robust in imperfect scenarios.

95. Signal Separation with Magnitude Constraints : a Phase Unmixing Problem
A. Deleforge and Y. Traonmilin (INRIA Rennes, France)
We consider the problem of estimating the phases of K mixed complex signals from a multichannel observation, when the mixing matrix and signal magnitudes are known. We compare three approaches to tackle it: a heuristic method, an alternate minimization method, and a convex relaxation into a semi-definite program. In particular, we show that the convex relaxation approach yields best results, including the potential for exact source separation in under-determined settings.

96. Subspace Regularized Dynamic Time Warping for Spoken Query Detection
D. Ram (Idiap Research Institute and EPFL, Switzerland),
A. Asaei (Idiap Research Institute, Switzerland),
H. Bourlard (Idiap Research Institute and EPFL, Switzerland)
Deep neural network posterior probabilities are the best features for query detection in speech archives. Dynamic time warping (DTW) is the state-of-the-art solution for this task. Posterior features live in low-dimensional subspaces whereas, the current DTW methods do not incorporate this global structure of the data and rely on local feature distances. We exploit the query example as the dictionary for sparse recovery. Local DTW scores are integrated with the sparse reconstruction scores to obtain a subspace regularized distance matrix for DTW. The proposed method yields a substantial performance gain over the baseline system.

97. Graph-based Total Variation for Tomographic Image Reconstruction
F. Mahmood (Okinawa Institute of Science and Technology, Japan),
N. Shahid (EPFL, Switzerland),
U. Skoglund (Okinawa Institute of Science and Technology, Japan)
In this short abstract we introduce a novel method for tomographic reconstructions from low-dose data which is usually noisy and has missing information. Our method is a generalization of sparsity exploiting image reconstruction methods which employ Total Variation (TV) as an additional sparsifying transform. Similar to state-of-the-art Non-local TV (NLTV) method our proposed method goes beyond spatial similarity between different regions of an image being reconstructed by establishing a connection between similar regions in the image regardless of spatial distance. However, it involves updating the graph prior during every iteration and is computationally more efficient as compared to adaptive NLTV.

98. Learning Transforms With a Specified Condition Number
S. Mukherjee and C. S. Seelamantula (Indian Institute of Science, India)
We address the problem of learning data-adaptive square sparsifying transforms with a condition number constraint. We adopt an alternating minimization (Alt. Min.) strategy and propose a projection approach, following the transform-update step within every iteration of Alt. Min., to enforce the condition number constraint by solving a quadratic program. The set of updated singular values of the transform can be expressed as an affine relaxation applied on the current ones. The proposed approach, referred to as singular value relaxation (SVR), is compared with two recently proposed transform learning techniques in terms of signal sparsification performance. The experimental results show that the transform learnt using SVR is in better agreement with the ground-truth and leads to competitive reconstruction performance with the state-of-the-art methods with easier tuning of parameters.

99. Sparsity Order Estimation under Constrained Budget
M. K. Pirbalouti and M. F. Naeini (Shahed Uni., Iran),
A. Amini (Sharif Uni. of Technology, Iran)
In many applications of compressed sensing it is either necessary or beneficial to have a priori information about the sparsity order $||x||_0$ of an unknown sparse vector x. As shown in [1], $||x||_1 / ||x||_2$ provides a more reliable measure of compressibility than the sparsity order. The latter could also be estimated by observing the vector through random linear measurements following Cauchy and Gaussian distributions.

In this paper, we determine the optimal number of Cauchy and Gaussian measurements to minimize the estimation error under limited measurement budget.

100. Multi-Source Image Enhancement via Coupled Dictionary Learning
K. Fotiadou, G. Tsagkatakis, and P. Tsakalides
(Foundation for Research and Technology, Greece)
Hyperspectral remote sensing imagery provides valuable insights regarding the composition of a scene and facilitates tasks such as spectral unmixing, and region clustering. Nevertheless, current remote sensing imaging architectures are unable to concurrently acquire high spatial and spectral resolution imagery. This work proposes a novel post-acquisition technique aiming to enhance the spectral or the spatial dimensionality of imaging systems by exploiting the mathematical frameworks of Sparse Representations and Dictionary Learning. The key contribution of this work is a novel coupled dictionary learning model which considers coupled feature spaces, composed of low and high resolution hypercubes, in order to address the multi-instrument enhancement problem. We formulate our coupled dictionary learning technique within the context of the Alternating Direction Method of Multipliers. Experimental results demonstrate the ability of the proposed approach to synthesize high resolution hypercubes.

101. Learning Convolutional Proximal Filters
U. Kamilov, H. Mansour, and D. Liu (MERL, USA)
In the past decade, sparsity-driven methods have led to substantial improvements in the capabilities of numerous imaging systems. While traditionally such methods relied on analytical models of sparsity, such as total variation (TV) or wavelet regularization, recent methods are increasingly based on data-driven models such as dictionary-learning or convolutional neural networks (CNN). In this work, we propose a new trainable model based on the proximal operator for TV. By interpreting the popular fast iterative shrinkage/thresholding algorithm (FISTA) as a CNN, we train the filters of the algorithm to minimize the error over a training data-set. Experiments on image denoising show that by training the filters, one can substantially boost the performance of the algorithm and make it competitive with other state-of-the-art methods.

102. Uniform Recovery Guarantees for Hadamard Sampling and Wavelet Reconstruction
V. Antun (Uni. of Oslo, Norway), A. Hansen (Uni. of Cambridge, UK),
B. Adcock (Simon Fraser Uni., USA), Ø. Ryan (Uni. of Oslo, Norway)
In most real-world applications of compressed sensing uniform random sampling is suboptimal, however structured sampling is an effective alternative. In order to obtain satisfactory signal reconstructions in such applications one need to incorporate both the signal structure, and the local coherence structure of the change of basis matrix in the choice of sampling patterns. In this text we will estimate the local coherences in a change of basis matrix between Hadamard samples and Daubechies wavelets. These estimates are then combined with newly obtained uniform recovery guarantees, to create concrete guarantees for Hadamard sampling combined with Daubechies wavelets.

103. Sparse Parametric Estimaton of Poisson Processes

M. Moore and M. Davenport (Georgia Institute of Technology, USA)

We present a recovery guarantee for using event observations to estimate the parameters of inhomogeneous Poisson processes with linear representations using maximum-likelihood estimation. Our result allows for an improved bound when the parameterization is known to be sparse and the recovery is appropriately constrained. Not only do we provide a guarantee for a problem that previously lacked one, but we generalize and improve existing bounds for the related "Poisson linear inverse problem" y~Poisson(Ax+b).

104. NUWBS: Non-Uniform Wavelet Bandpass Sampling for Compressive RF Feature Acquisition

M. Pelissier (CEA LETI-MINATEC, France), C. Studer (Cornell Uni., USA)

Feature extraction from wideband radio-frequency (RF) signals, such as spectral activity, interferer energy and type, or direction-of-arrival, finds use in a growing number of applications. Compressive sensing (CS)-based analog-to-information (A2I) converters enable the design of inexpensive and energy-efficient wideband RF sensing solutions for such applications. However, most A2I architectures suffer from a variety of real-world impairments. We propose a novel A2I architecture, referred to as non-uniform wavelet bandpass sampling (NUWBS). Our architecture extracts a carefully-tuned subset of wavelet coefficients directly in the RF domain, which mitigates the main issues of most existing A2I converters. We use simulations to show that NUWBS approaches the performance limits of l1-norm-based sparse signal recovery.

105. Variable Splitting and Cycle Spinning for Sparse Signal Recovery

E. Sakhaee and A. Entezari (Uni. of Florida, USA)

We propose a variable splitting approach for sparse recovery from incomplete Fourier data, which significantly improves conventional wavelet-based compressed sensing/reconstruction, offers the benefits of Shift-invariant Wavelet Transform (SWT), and overcomes the high redundancy factor of SWT. Our method recovers sparse Discrete Wavelet Transform (DWT) coefficients of translated version of the signal in parallel, while enforces consistency between the translated signals via solving the problem in an ADMM formulation. The experiments demonstrate that 4 shifts are sufficient to achieve reconstruction accuracy as high as reconstruction using SWT, hence, significantly reducing the computational cost and redundancy factor of SWT frame.

106. Improved Guarantees for Correlated-PCA (PCA when Data and Noise are Correlated)

N. Vaswani and H. Guo (Iowa State Uni., USA)

We study Principal Component Analysis (PCA) in the setting where a part of the corrupting "noise" is correlated with the true data (correlated-PCA). Such corruption is often called "data-dependent noise". We provide a guarantee for the most commonly used PCA solution, simple eigenvalue decomposition (EVD). To our best knowledge, most existing results that study the simple EVD solution to PCA assume that the true data and the corrupting noise are uncorrelated. This is valid in practice often, but not always. We first studied correlated-PCA in a 2016 NIPS paper. This abstract removes two key limitations of that work.

107. Reweighted L1-norm Minimization with Guarantees: An Incremental Measurement Approach to Sparse Reconstruction

J. Mota (Heriot-Watt Uni., UK), L. Weizman (Technion, Israel), N. Deligiannis (Vrije Uni.it Brussel, Belgium), Y. Eldar (Technion, Israel), M. Rodrigues (Uni. College London, UK)

We propose a sparse reconstruction algorithm that not only provably reconstructs sparse signals, but also automatically decides on the number of measurements to do so. This is achieved via an incremental measurement approach: measurements are acquired in blocks, rather than all at once. Our algorithm is based on the iteratively reweighted L1 minimization (IRL1) method which, however, requires the number of measurements to be selected in advance. Experimental results show that our algorithm automatically selects a number of measurements that is often smaller than the minimum number of measurements that IRL1 requires for perfect reconstruction.

108. Complex Domain Nonlocal Group-Wise Sparsity: Toward Wavelength Super-Resolution Phase Imaging in Coherent Optics

V. Katkovnik and K. Egiazarian (Tampere Uni. of Technology, Finland)

Complex-domain nonlocal block-wise sparsity is discussed. It is demonstrated for optical coherent phase imaging: in-line digital holography configuration with thin lens and coding phase mask in object plane. An iterative super-resolution algorithm is developed as optimal for noisy Poissonian observations. Sparse modeling of amplitude and phase is one of the key instruments of the approach. The high accuracy of sub-wavelength resolution is demonstrated by simulation experiments.

Poster Session 3, Thursday, June 8

109. Convergence Results of GROUSE

D. Zhang and L. Balzano (Uni. of Michigan, USA)

Subspace learning and matrix factorization problems have a great many applications in science and engineering, and efficient algorithms are critical as dataset sizes continue to grow. Many relevant problem formulations are non-convex, and in a variety of contexts it has been observed that solving the non-convex problem directly is not only efficient but reliably accurate. We discuss convergence theory for a particular method: first order incremental gradient descent constrained to the Grassmannian.

110. Low-Rank Tensor Regularization for Improved Dynamic Quantitative Magnetic Resonance Imaging

N. Kargas, S. Weingartner, N. Sidiropoulos, M. Akcakaya (Uni. of Minnesota, USA)

In quantitative MRI, a series of images with different soft-tissue contrast are acquired. For moving organs, such as the heart, this process may be performed across different motion states, leading to ≥4-dimensional images. Furthermore, due to the nature of organ motion and contrast changes, such datasets can be well-represented using low-rank tensors. In this work, we investigate the utility of low-rank tensor regularization for improving the quantification of a state-of-the-art cardiac MRI technique.

111. Sparse Pronunciation Codes for Perceptual Phonetic Information Assessment

A Asaei and M. Cernak (Idiap Research Institute, Switzerland),
H. Bourlard (Idiap Research Institute and EPFL, Switzerland),
D. Ram (Idiap Research Institute, Switzerland)

Speech is a complex signal produced by a highly constrained articulation machinery. Neuro and psycholinguistic theories assert that speech can be decomposed into molecules of structured atoms. Although characterization of the atoms is controversial, the experiments support the notion of invariant speech codes governing speech production and perception. We exploit deep neural network (DNN) invariant representation learning for probabilistic characterization of the phone attributes defined in terms of the phonological classes and known as the smallest-size perceptual categories. We cast speech perception as a channel for phoneme information transmission via the phone attributes. Structured sparse codes are identified from the phonological probabilities for natural speech pronunciation. We exploit the sparse codes in information transmission analysis for assessment of phoneme pronunciation. The linguists define a single binary phonological code per phoneme. In contrast, probabilistic estimation of the phonological classes enables us to capture large variation in structures of speech pronunciation. Hence, speech assessment may not be confined to the single expert knowledge based mapping between phoneme and phonological classes and it may be extended to multiple data-driven mappings observed in natural speech.

112. Dynamic Filtering with Earth Mover's Distance Regularization

A. Charles (Princeton Uni., USA), J. Lee, N. Bertrand, and C. Rozell (Georgia Institute of Technology, USA)

Tracking dynamically evolving signals is a critical part of many applications. Additionally, many dynamic signals have additional structure, such as sparsity in a basis. While a number of methods have been developed to track sparse signals, none of these methods take into account spatial information about the indices of x. For example, consider the application where targets are being tracked in a 2D space. Predicting a target one pixel away from the truth has the same ℓ_p error as predicting a target 20 pixels away, In these cases current methods would over-emphasize certain prediction errors and obtain sub-par performance. In this work we extend the dynamic filtering framework to capture dynamics related to the spatial locations. Specifically, we introduce the earth-mover's distance (EMD) as a dynamics regularizer for use in tracking of sparse signals. We describe here the optimization program and computational considerations, and empirically demonstrate superior performance in tracking targets in two-dimensional visual scenes.

113. Class-adapted Blind Image Deblurring

M. Ljubenovic and M. Figueiredo (Instituto de Telecomunicações, Instituto Superior Técnico, Portugal)

Over the past few decades, significant progress has been made in solving image deblurring problems; however, most of the developed methods are focused on deblurring of natural images. Specific classes of images, like text, face, fingerprints are found in many important applications, such as document analysis or forensics. State-of-the-art blind image deblurring methods are usually based on edge extraction or on typical statistics of natural images. When there is not much texture in a blurred image (e.g., face images), performance of methods based on edge extraction is limited. On the other hand, methods tailored for natural images do not take into consideration the particular structure of images of a specific class (e.g., text or fingerprints). In this work, we proposed a method with a patch-based class-adapted image prior trained from a dataset which contains clean images of a specific class. Results obtained so far show that the method can be used for at least two specific classes of images: text and face. Additionally, the proposed method uses a week prior on the blurring filter (positivity and support) and, because of that, is able to recover a wide variety of blurring filters.

114. Exploiting Joint Array and Spatial Sparsity for Broadband Source Localisation with Fisher Information Matrix Constraints

M. Chen and W. Wang (Uni. of Surrey, UK)

Recently, a Fisher Information Matrix (FIM) constrained joint array and spatial sparsity approach has been presented for narrowband direction of arrival (DoA) estimation. In this paper, we extend the FIM constraints from narrowband to broadband scenario. In the step of sparse array optimisation, a constraint with FIM is considered to reduce the error before scaling the observed signal by the weight coefficients. In the spatial sparsity reconstruction step, the difference between the reconstruted result and the desired beam response is also constrained by adding a statistic expression. The two-step is operated in an iterative process and the current simulation results for narrowband DoA estimation demonstrate the performance of the proposed method.

115. Sparse Support Recovery with Non-smooth Loss Functions

K. Degraux (Uni. Catholique de Louvain, Belgium),
G. Peyré (CNRS, Ecole Normale Supérieure, France),
J. Fadili (ENSICAEN, CNRS, France),
L. Jacques (Uni. Catholique de Louvain, Belgium)

In this work, we study the support recovery guarantees of underdetermined sparse regression using the l1-norm as a regularizer and a non-smooth loss function for data fidelity. More precisely, we focus on the ℓ_1 and ℓ_∞ losses, and contrast them with the usual ℓ_2 smooth loss. We identify an "extended support" for the vector to recover and derive a sharp condition which ensures that it is stable to small additive noise in the observations. We give a detailed numerical analysis of the support stability of compressed sensing recovery with these different losses. This highlights different parameter regimes, ranging from support stability to increasing support instability.

116. Sparse Super-Resolution from Laplace Measurements

Q. Denoyelle (Uni. Paris-Dauphine, France), E. Soubies (EPFL, Switzerland),
V. Duval (INRIA Rocquencourt, France),
G. Peyré (CNRS, Ecole Normale Supérieure, France)

We propose a theoretical analysis of the super-resolution performance of the BLASSO "off-the-grid" recovery method from Laplace transform measurements. This transform is not translation invariant, thus requiring the use of theoretical and algorithmic tools that go beyond traditional deconvolution-based methods. We show that the BLASSO offers a stable and computationally tractable super-resolution of positive spikes. In particular, when the signal-to-noise ratio is of the order of $1/t^{2N-1}$ (where t is the spacing between the N spikes to recover), the BLASSO program outputs the correct number of spikes. This result suggests that the BLASSO should be a method of choice to tackle challenging Laplace inversions, which are at the heart of recently proposed fluorescence imaging methods.

117. $\ell 1/\ell 2$ Regularized Non-Convex Low-Rank Matrix Factorization

P. Giampouras, A. Rontogiannis, and K. Koutroumbas
(National Observatory of Athens, Greece)

Low-rank matrix factorization plays a key role in a plethora of problems commonly met in machine learning applications dealing with big data as it reduces the size of the emerging optimization problems. In this work we introduce a novel low-rank promoting regularization function which gives rise to an algorithm that induces sparsity jointly on the columns of the matrix factors. Apart from the reduced computational complexity requirements it offers, the new algorithm also provides a *basis* of the sought low-rank subspace.

118. An Efficient Direction-Of-Arrival Estimation Method Based on Weighted Sparse Spectrum Fitting

K. Ichige (Yokohama National Uni., Japan)

This paper presents a novel Direction-Of-Arrival (DOA) estimation method for Uniform Linear Array (ULA) based on the weighted Sparse Spectrum Fitting (SpSF) which can be regarded as a weighted ℓ_1-regularization problem.

119. BPConvNet for Compressed Sensing Recovery in Bioimaging
K. H. Jin, M. McCann, M. Unser (EPFL, Switzerland)
Iterative reconstruction methods have become the standard ap- proach to solving inverse problems in imaging including denoising, deconvolution, and interpolation. With the appearance of compressed sensing, our theoretical understanding of these approaches evolved further with remarkable outcomes. These advances have been particularly influential in the field of biomedical imaging, e.g., in magnetic resonance imaging (MRI) and X-ray computed tomography (CT). A more recent trend is deep learning, which has arisen as a promising framework pro- viding state-of-the-art performance for image classification and segmentation, regression-type neural networks. In this paper, we explore the relationship between CNNs and iterative optimization methods for one specific class of inverse problems: those where the normal operator associated with the forward model is a convolution. Based on this connection, we propose a method for solving these inverse problems by combining a fast, approximate solver with a CNN. We demonstrate the approach on low-view CT reconstruction and accelerated MRI using residual learning and multilevel learning.

120. Learning Non-Structured, Overcomplete and Sparsifying Transform
D. Kostadinov, S. Voloshynovskiy, S. Ferdowsi (Uni. of Geneva, Switzerland)
This paper focuses on learning a data adaptive transform for sparse data in a space with dimensions larger (or equal) than the dimensions of the original space. A generalized problem formulation is addressed and iterative alternating algorithm for learning incoherent and well-conditioned overcomplete transform is proposed. The proposed algorithm alternates between transform update step and sparse coding step. Global optimal and ε-close solutions for every step are presented, having low computational complexity. A convergence result for the iterative algorithm is given. Computer simulations for image denoising application are provided, demonstrating promising denoising recovery performance and promoting advantages in computational complexity, memory requirements and training data requirements.

121. Harmonic Mean Iteratively Reweighted Least Squares For Low-Rank Matrix Recovery
C. Kümmerle and J. Sigl (Technische Universität München, Germany)
We propose a new Iteratively Reweighted Least Squares (IRLS) algorithm for the low-rank matrix recovery problem. It introduces enhanced, so-called harmonic mean weight matrices, which lead to an improved minimization of the nonconvex Schatten-p functional compared to previous IRLS approaches to the problem. Our convergence analysis includes a locally superlinear convergence rate of the algorithm. This is remarkable, since to the best of our knowledge, it is the first algorithm for low-rank matrix recovery which achieves a superlinear rate of convergence for low complexity measurements as well as for larger problems. This convergence rate is confirmed very accurately by our numerical experiments. Our simulations also suggest that the algorithm needs fewer measurements for successful recovery than any other tractable algorithm in the literature.

122. Leveraging Union of Subspace Structure to Improve Constrained Clustering
J. Lipor and L. Balzano (Uni. of Michigan, USA)
We present a pairwise-constrained clustering algorithm that leverages the union of subspace structure assumed in subspace clustering. The central step of the

algorithm is in querying points of minimum margin between estimated subspaces; analogous to classifier margin, these lie near the decision boundary. We prove that points lying near the intersection of subspaces are points with low margin. Our procedure can be used after any subspace clustering algorithm that outputs an affinity matrix and is capable of driving the clustering error down more quickly than other state-of-the-art active query algorithms on datasets with subspace structure.

123. Optimization Convergence of Matching Pursuit Algorithms
F. Locatello and M. Tschannen (ETH Zurich, Switzerland),
R. Khanna (Uni. of Texas, Austin, USA), M. Jaggi (EPFL, Switzerland)
The matching pursuit algorithm and its variants are among the most commonly used methods for greedy optimization. In this paper, we present the first explicit convergence rates of matching pursuit methods in an optimization sense, for general sets of atoms. We present sublinear O(1/t) convergence on general smooth objectives, and linear convergence on strongly convex objectives. Our algorithm variants and rates do not need any incoherence or sparsity assumptions. Direct applications of the presented algorithms are structured matrix and tensor factorization problems.

124. Stable Recovery of the Factors From a Deep Matrix Product
F. Malgouyres (Uni. de Toulouse, France), J. Landsberg (Texas A&M Uni., USA)
We study a deep matrix factorization problem. It takes as input the matrix X obtained by multiplying K matrices (called factors) and aims at recovering the factors. When K=1, this is the usual compressed sensing framework; K=2: Examples of applications are dictionary learning, blind deconvolution, self-calibration; $K\geq3$: can be applied to many fast transforms (such as the FFT). In particular, we apply the theorems to deep convolutional network. Using a Lifting, we provide : a necessary and sufficient conditions for the identifiability of the factors (up to a scale indeterminacy); - an analogue of the Null-Space-Property, called the Deep-Null-Space-Property which is necessary and sufficient to guarantee the stable recovery of the factors.

125. Magnetic Resonance Fingerprinting by exploiting Low Rank
G. Mazor and L. Weizman (Technion, Technion),
A. Tal (Weizmann Institute of Science, Israel), Y. Eldar (Technion, Israel)
Magnetic Resonance Fingerprinting (MRF) is a relatively new approach that provides quantitative MRI measures using randomized acquisition. Extraction of physical quantitative tissue parameters is performed off-line, based on acquisition with varying parameters and a dictionary generated according to the Bloch equations. MRF uses hundreds of radio frequency (RF) excitation pulses for acquisition, and therefore high under-sampling ratio in the sampling domain (k-space) is required for reasonable scanning time. This under-sampling causes spatial artifacts that hamper the ability to accurately estimate the tissue's quantitative values. In this work, we introduce a new approach for quantitative MRI using MRF, called magnetic resonance Fingerprinting with LOw Rank (FLOR). We exploit the low rank property of the concatenated temporal imaging contrasts, on top of the fact that the MRF signal is sparsely represented in the generated dictionary domain. %We present an iterative scheme that consists of a gradient step followed by a low rank projection using the singular value decomposition. Experiments on real MRI data, acquired using a spirally-sampled MRF FISP sequence, demonstrate better resolution compared to other compressed-sensing based methods for MRF at 5% sampling ratio.

126. Sparse Estimation in Ordinary Differential Equation Systems

F. Mikkelsen and N. Hansen (Uni. of Copenhagen, Denmark)

Understanding chemical reaction networks in systems biology is an important dicipline, as it provide the means for quantifying downstream effects of chemical or medical interventions. Thus identifying these systems from noisy data is a major challenge with far reaching applications – such as to the inference of phosphoprotein interaction networks. For mass action kinetics the system structure – the network – is encoded via sparsity in a parameter vector, whose dimension increases rapidly with the number of species. We have developed an algorithm for simultaneous system identification and parameter estimation via minimisation of a penalised loss function. The global minimiser is difficult to find, and focus has been on computational aspects, as well as variance-reduction techniques. Several techniques are combined to cope with the computational and statistical aspects. The resulting method is implemented in an R-package, which provides sparse estimates of systems with up to 10^5 reactions.

127. Compressive Hyperspectral Imaging using Coded Fourier Transform Interferometry

A. Moshtaghpour, V. Cambareri, L. Jacques (Uni. Catholique de Louvain, Belgium), P. Antoine and M. Roblin (Lambda-X SA, Belgium)

Fourier Transform Interferometry (FTI) is a Hyperspectral (HS) imaging technique that is specially desirable in high spectral resolution applications, such as spectral microscopy in biology. The current resolution limit of FTI is actually due to the durability of biological elements when exposed to illuminating light. We propose two variants of the FTI imager, i.e., coded illumination-FTI and coded aperture-FTI, that efficiently allocate the illumination distribution with a variable density sampling strategy, so that the exposure time of the biological specimen is minimized while spectral resolution is preserved. We derive a theoretical analysis for both proposed methods. Our results are supported by several experimental simulations.

128. Generalized Approximate Message Passing for Noisy Quantized Compressed Sensing

O. Musa (Technical Uni. of Vienna, Austria), N. Goertz (Vienna Uni. of Technology, Austria)

Compressed sensing (CS) is a novel technique that allows for stable reconstruction with a sampling rate lower than the Nyquist rate if the unknown vector is sparse. In many practical applications CS measurements are first scalar quantized (possibly with high rate) and later corrupted in different ways. Reconstruction by convetional techniques of such highly distorted measurements yields poor results. To overcome this problem, we use the well known generalized approximate message passing (GAMP) algorithm and tailor it for quantized CS measurements corrupted with noise. We provide the necessary expressions for the nonlinear updates of different noise models, namely symmetric discrete memoryless channels (SDMCs) and additive white Gaussian noise (AWGN) channels. Our numerical results show the superiority of the GAMP algorithm compared to conventional reconstruction algorithms in both symmetric discrete memoryless channel (SDMC) and AWGN channels.

129. Efficient Learning of Sparse Image Representations Using Homeostatic Regulation

V. Boutin, F. Ruffier, and L. Perrinet (Aix-Marseille Uni., France)

One core advantage of sparse representations is the efficient coding of complex signals using compact codes. For instance, it allows for the representation of any sample as a combination of few elements drawn from a large dictionary of basis functions. In the context of the efficient processing of natural images, we propose here that sparse coding can be optimised by designing a proper homeostatic rule regulating the competition between the elements of the dictionary. Indeed, a common design for unsupervised learning rules relies on a gradient descent over a cost measuring representation quality with respect to sparseness. The sparseness constraint introduces a competition which can be optimised by ensuring that each item in the dictionary is selected as often as others. We implemented this rule by introducing a gain normalisation similar to what is observed in biological neural networks. We validated this theoretical insight by challenging the matching pursuit sparse coding algorithm with the same learning rule but with or without homeostasis. Simulations show that for a given homeostasis rule, gradient descent performed similarly the learning of a dataset of image patches. While the coding accuracy did not vary much, including homeostasis changed qualitatively the learned features. In particular, homeostasis results in a more homogeneous set of orientation selective filters, which is closer to what is found in the visual cortex of mammals. To further validate these results, we applied this algorithm to the optimisation of a visual system to be embedded in an aerial robot. In summary, this biologically-inspired learning rule demonstrates that principles observed in neural computations can help improve real-life machine learning algorithms.

130. Non-convex Blind Deconvolution Approach for Sparse Radio-Interferometric Imaging

A. Repetti, J. Birdi, and Y. Wiaux (Heriot-Watt Uni., UK)

New generations of imaging devices aim to produce high resolution and high dynamic range images. In this context, the associated high dimensional inverse problems can become extremely challenging from an algorithmic view point. In addition, the imaging procedure can be affected by unknown calibration kernels. This leads to the need of performing joint image reconstruction and calibration, and thus of solving non-convex blind deconvolution problems. In a recently submitted paper, we developed a method to solve the joint imaging and calibration problem in radio interferometry in astronomy. To solve this problem, we leverage a block-coordinate forward-backward algorithm, specifically designed to minimize non-smooth non-convex and high dimensional objective functions. Here we describe the proposed method and show its performance through simulation results.

131. Sequential Learning of Analysis Operators

M. Sandbichler and K. Schnass (Uni. of Innsbruck, Austria)

We present a new way to learn an analysis operator from co-sparse data in an online fashion via a stochastic gradient method.

132. Parameter Learning for Log-supermodular Distributions

T. Shpakova (INRIA, France), F. Bach (INRIA/ENS, France)

We consider log-supermodular models on binary variables, which are probabilistic models with negative log-densities which are submodular. These models provide probabilistic interpretations of common combinatorial optimization tasks such as image segmentation. In this paper, we focus primarily on parameter estimation in the models from known upper-bounds on the intractable log-partition function. We show that the bound based on separable optimization on the base polytope of the submodular function is always inferior to a bound based on “perturb-and-MAP” ideas. Then, to learn parameters, given that our approximation of the log-partition function is an expectation (over our own randomization), we use a stochastic sub-gradient technique to maximize a lower-bound on the log-likelihood. This can also be extended to conditional maximum likelihood. We illustrate our new results in a set of experiments in binary image denoising, where we highlight the flexibility of a probabilistic model to learn with missing data.

133. Sketching With Structured Matrices for Array Imaging

R. S. Srinivasa, M. Davenport, J. Romberg (Georgia Institute of Technology, USA)

We address a novel matrix sketching problem motivated by a coherent imaging application using antenna arrays. We propose a novel trade-off between the number of measurements needed for reconstruction and bandwidth of excitation without assuming any sparsity on the target image. This process can be modeled as a linear sketching problem with a repeated block diagonal sketching matrix. We show that the reduction in dimension depends on the interaction of subspaces of groups of rows of the original matrix. We provide simulation results for general matrices and imaging examples to show the utility of such measurements in coherent imaging

134. High Dimensional Dictionary Learning and Applications

J. Sulam, M. Zibulevsky, M. Elad (Technion, Israel)

In this work, we show how to efficiently handle bigger dimensions and go beyond the small patches in sparsity-based signal and image processing methods. We build our approach based on a new cropped Wavelet decomposition, which enables a multi-scale analysis with virtually no border effects. Employing this as the base dictionary within a double sparsity model, and to cope with the increase of training data, we present an Online Sparse Dictionary Learning (OSDL) algorithm to train this model effectively, enabling it to handle millions of examples. The resulting large trainable atoms - trainlets - not only achieve state of the art performance in dictionary learning when compared to other methods, but it also opens the door to new challenges and problems that remained unattainable until now. In addition to reviewing the capabilities of the OSDL algorithm, we present very recent results on inpainting of large regions of face images, as well as preliminary results on full end-to-end image compression.

135. Image Sharpening using Scene-Adapted Priors

A. Teodoro, J. Bioucas-Dias, and M. Figueiredo

(Instituto de Telecomunicações, Instituto Superior Técnico, Portugal)

With general purpose denoisers approaching performance limits, we turn our attention to methods targeted to specific tasks or image classes. Using the recent plug-and-play (PnP) framework, we plug a denoiser based on Gaussian mixture models (GMM) into the iterations of an alternating direction method of multipliers (ADMM), to tackle a data fusion problem known as image sharpening. Results show that our method performs competitively with other state-of-the-art algorithms.

136. Theoretical Limits of Streaming Inference and Mini-Batch Message-Passing Algorithms

A. Manoel (Neurospin CEA, France), E. Tramel (INRIA, France), T. Lesieur and L. Zdeborová (CNRS & CEA, France), F. Krzakala (ENS, France)

In statistical learning with big data, one is often forced to process only a small chunk of data at a time, resulting in "online" or "streaming" methods. We present an analysis, based on a statistical physics approach, of the information theoretic limits of inference in the streaming setting, precisely characterizing the interpolation between fully online and offline inference. We examine the streaming problem applied to the binary perceptron, sparse linear regression, and Gaussian mixture models, discovering first-order phase transitions in the chosen batch size. We further propose a limit-achieving mini-batch version of approximate message-passing and apply it to real-world data.

137. Class-specific Image Denoising Using Importance Sampling

M. Niknejad, J. Bioucas-Dias, and M. Figueiredo (Instituto de Telecomunicações, Instituto Superior Técnico, Portugal)

We propose a method for image denoising in which the noisy image is known to belong to a certain class (text, face, fingerprints). We fit a set of multivariate Generalized Gaussian (GG) distributions to a database of clean image patches from the certain class. We use these priors to denoise each noisy patch of the image. Due to difficulty of obtaining the MMSE estimation of clean patch using the GG distribution as a prior, we use an importance sampling approach to approximate it. Our experiments show that the proposed method outperforms other general or class-specific methods for image denoising.

138. Network-based Sparse Modeling of Breast Invasive Carcinoma Survival Data

A. Veríssimo and E. Carrasquinha (IDMEC, Portugal), M.-F. Sagot (ERABLE, INRIA, France), A. Oliveira (INESC-ID, Instituto Superior Técnico, Portugal), S. Vinga and M. Lopes (IDMEC, Portugal)

Learning survival models from oncological data has now become a major challenge due to the significant increase of molecular information. The inherent high-dimensionality of these datasets, where the number of features largely exceeds the number of observations, leads to ill-posed inverse problems and, consequently, to models that often lack interpretability. In order to tackle this problem, regularized optimization has emerged as a promising approach, allowing to impose constraints on the structure of the solutions, these include sparsity, for example, using LASSO, or other penalizing functions that use network-based information if the features have a graph-based configuration. We compared how different sparse methods perform when applied to a Breast Invasive Carcinoma dataset and how introducing network knowledge impact model prediction. These include Elastic Net and their coupling with DegreeCox. The results regarding the concordance c-index show an improvement when network information is included, whereas the log-rank tests on the separation between high and low-risk patients exhibit a decrease in performance. It is expected that the obtained models can be further support clinical decision and prognostic assessment of oncological patients.

139. Dynamic One-Bit Matrix Completion

L. Xu and M. Davenport (Georgia Institute of Technology, USA)

In this work we consider a new setting where we aim to recover an underlying and dynamically evolving low-rank matrix from binary observations. First we specify the evolving model of the underlying low-rank matrix and the binary observation model. Then we propose the one-bit LOWEMS (Locally Weighted Matrix Smoothing) estimator by minimizing the weighted negative log-likelihood under both low-rank and spikiness constraints. Furthermore, we provide a constrained alternating gradient descent algorithm to solve the above estimator based on matrix factorization. Finally, both our synthetic simulations and real world experiments show superior performance of the proposed estimator over its static counterpart.

140. Weighted Diffusion Sparse LMP Algorithm in Non-uniform Noise Environment

M. Korki (Swinburne Uni. of Technology, Australia),
H. Zayyani (Qom Uni. of Technology, Iran)

This paper presents an improved version of diffusion least mean p-power (LMP) algorithm for distributed estimation of a sparse parameter vector. We replace the sum of mean square errors with a weighted sum of LMP for global and local cost functions of a network of sensors. The weight coefficients are adaptive and are updated by a simple steepest-descent recursion to minimize the global and local cost functions of the adaptive algorithm. Simulation results show the advantages of the proposed weighted diffusion LMP over the diffusion LMP algorithm specially in the non-uniform noise environments in a sensor network.

141. On Concentration Inequalities for Sparse Vectors

K. Zhang, A. Entezari (Uni. of Florida, USA)

Concentration inequalities play a key role in establishing the suitability of random matrices for compressed sensing as well as dimensionality reduction through random projections. In particular, the Restricted Isometry Property (RIP) for random Gaussian matrices and its relationship to Johnson-Lindenstrauss lemma can be established directly from concentration inequality [3]. While concentration inequality is usually established for general vectors, the distortion (in length) to sparse vectors by random matrices is usually obtained by a combinatorial argument together with union bounds. In this work, we study concentration inequalities specific to sparse vectors when projected by random matrices from compactly supported (e.g., uniform) distributions. From this approach we naturally obtain sharper bounds compared to generic concentration inequalities. These results suggest the superiority of such distributions over the Gaussian distribution for random projection of sparse vectors. Our experiments show this improvement in concentration bound for a special case of sparse binary signals and the results are further corroborated by a higher rate of recovery of general (non-binary) sparse signals from random projections.

142. Hyperspectral Image Denoising and Anomaly Detection Based on Low-Rank and Sparse Representations

L. Zhuang (Instituto de Telecomunicações, Instituto Superior Técnico, Portugal),
L. Gao and B. Zhang (Chinese Academy of Sciences), J. Bioucas-Dias (Instituto de Telecomunicações, Instituto Superior Técnico, Portugal)

The very high spectral resolution of Hyperspectral Images (HSIs) enables the identification of materials with subtle differences and the extraction subpixel information. However, the increasing of spectral resolution often implies an increasing in the noise linked with the image formation process. This degradation mechanism limits the quality of extracted information and its potential applications. Since HSIs represent natural scenes and their spectral channels are highly correlated, they are characterized by a high level of self-similarity and are well approximated by low-rank representations. These characteristic underlies the state-of-the-art in HSI denoising. However, in presence of rare pixels, the denoising performance of those methods is not optimal and, in addition, it may compromise the future detection of the rare pixels. To address these hurdles, we propose a powerful HSI denoiser which implements hard low-rank representation, promotes self-similarity in the representation coefficients, and, by using a form of collaborative sparsity, preserves rare pixels. The denoising and detection effectiveness of the proposed robust HSI denoiser is illustrated using semi-real data.

143. A Sparse Tensor Decomposition with Multi-Dictionary Learning Applied to Diffusion Brain Imaging

C. Caiafa (Indiana Uni., USA), A. Cichocki (LABSP - RIKEN, Japan), F. Pestilli (Indiana Uni., USA)

We use a multidimensional signal representation that integrates diffusion Magnetic Resonance Imaging (dMRI) and tractography (brain connections) using sparse tensor decomposition. The representation encodes brain connections (fibers) into a very-large, but sparse, core tensor and allows to predict dMRI measurements based on a dictionary of diffusion signals. We propose an algorithm to learn the constituent parts of the model from a dataset. The algorithm assumes a tractography model (support of core tensor) and iteratively minimizes the Frobenius norm of the error as a function of the dictionary atoms, the values of nonzero entries in the sparse core tensor and the fiber weights. We use a nonparametric dictionary learning (DL) approach to estimate signal atoms. Moreover, the algorithm is able to learn multiple dictionaries associated to different brain locations (voxels) allowing for mapping distinctive tissue types. We illustrate the algorithm through results obtained on a large in-vivo high-resolution dataset.

144. Exploiting a Hierarchy of Brain Regions for Alzheimer's Disease Classification

H. Barros and M. Silveira (ISR and IST, Portugal)

Sparse methods have gained popularity as an effective way to alleviate the curse of dimensionality in neuroimaging applications such as Alzheimer's disease (AD) and Mild Cognitive Impairment (MCI) classification. By imposing sparsity inducing regularization terms these methods are able to perform feature selection jointly with classification. The simplest of these methods is the Lasso which uses L1 norm regularization to induce sparsity. It has been used for AD classification in [1] and [2]. It is effective but the selected features may be sparsely distributed throughout the whole brain and unstable. To overcome this, one possibility is to consider groups of features. This is the approach of Group Lasso that uses $L_{2,1}$ norm to promote group sparsity. It was used for AD in [3]. It can also be generalized to allow for overlapping groups, including tree structured groups where a hierarchy of relationship between features can be defined. This approach has been proposed in [4] for AD and MCI classification using a pyramid hierarchy. In this extended abstract we propose an alternative tree structure, more consistent with disease related atrophy, where neighboring features are grouped according to anatomically defined regions of the brain and in a hierarchy that joins regions in the left and right hemispheres of the brain to take into account bilateral symmetry which typically occurs in AD. We apply these methods to MRI images from ADNI [5] and evaluate their classification performance and the stability of the obtained feature weights when several runs are performed.

145. Compression-based Acquisition of Structured Signals

S. Beyg (Uni. of Southern California, USA), S. Jalali (Nokia - Bell Labs, USA), A. Maleki (Columbia Uni., USA), U. Mitra (Uni. of Southern California, USA)

We present a novel and general approach that elevates the scope of compressed sensing recovery algorithms far beyond simple structures such as sparsity. The proposed method, referred to as compression-based gradient descent (C-GD), is capable of employing state-of-the-art compression algorithms to solve structured signal recovery problems. This enables C-GD to take advantage of complex structures that are used by the state-of-the-art compression codes, such as JPEG2000 and MPEG4. Our simulation results show that, in many instances, C-GD achieves state-of-the-art signal recovery performance. Furthermore, our theoretical results justify the performance of C-GD observed in our simulation results.

SpaRTaN-MacSeNet Special Poster Session

Monday June 5th 13:00

A. Parameter Learning For Log-supermodular Distributions
Tatiana Shpakova (INRIA, France)

B. Sliced Inverse Regression with Score Functions
Dmitry Babichev (INRIA, France)

C. Structure-exploiting Algorithms for Faster Estimation and Regression
Junqi Tang (Uni. of Edinburgh, UK)

D. Sampling Order Optimization for Contrast Preservation in Accelerated Prospective 3D MRI
Arnold Benjamin (Uni. of Edinburgh, UK)

E. Quantification of Metabolites in MR Spectroscopic Imaging using Machine Learning
Dhritiman Das (TU Muenchen, Germany)

F. A Spatio-temporally Regularised Reconstruction Model for Quantitative Perfusion-Weighted MRI
Cagdas Ulas (TU Muenchen, Germany)

G. Blind Source Separation of functional MRI via Assisted Dictionary Learning
Manuel Morante (Computer Technology Institute & Press, Greece)

H. FMRI Data Analysis via Higher Order Tensor Representations
Christos Chatzichristos (Computer Technology Institute & Press, Greece)

I. Interferometric Phase Image Estimation
Joshin Krishnan (Instituto de Telecomunicações, Portugal)

J. Patch-Based Blind Image Deblurring
Marina Ljubenovic (Instituto de Telecomunicações, Portugal)

K. Sparse Signal Recovery via Correlated Degradation Model
Nasser Eslahi (Tampere Uni. of Technology, Finland)

L. Single Image Super-Resolution Based on Wiener Filter in Similarity Domain
Cristóvão Cruz (Noisless Imaging, Finland)

M. A Greedy Algorithm with Learned Statistics for Sparse Signal Reconstruction
Lucas Rencker (Uni. of Surrey, UK)

N. Masked Non-negative Matrix Factorization for Bird Detection Using Weakly Labeled Data
Iwona Sobieraj (Uni. of Surrey, UK)

O. Skip-filtering Connections for Denoising Autoencoders and Music Source Separation
Stylianos Ioannis Mimilakis (Fraunhofer IDMT, Germany)

P. **Graph Based Learning in Echo State Networks**
Volodymyr Miz (EPFL, Switzerland)

Q. **Structured Sequence Modelling with Graph Convolutional Recurrent Networks**
Youngjoo Seo (EPFL, Switzerland)

R. **3D Point Cloud Denoising**
Zhongwei Xu (Noisless Imaging, Finland)

S. **Audio Source Separation with Deep Neural Networks Using the Dropout Algorithm**
Alfredo Zermini (Uni. of Surrey, UK)

T. **Automatic Music Transcription Using Sparse and Low-Rank Models**
Cian O'Brien (Uni. of Surrey, UK)

U. **Accelerated 3D T2* Mapping with Maximum Likelihood Estimation (MLE) and Parallel Imaging (PI)**
Wajiha Bano (Uni. of Edinburgh, UK)

V. **Transductive Learning with RKHS on Graphs**
Rodrigo Pena (EPFL, Switzerland)

W. **Classification of Unions of Subspaces with Convolutional Dictionaries**
Konstantinos Pitas (EPFL, Switzerland)

X. **Image Restoration by Importance Sampling**
Milad Niknejad (Instituto de Telecomunicações, Portugal)

Y. **Inverse Problems in Hyperspectral Imaging**
Lina Zhuang (Instituto de Telecomunicações, Portugal)

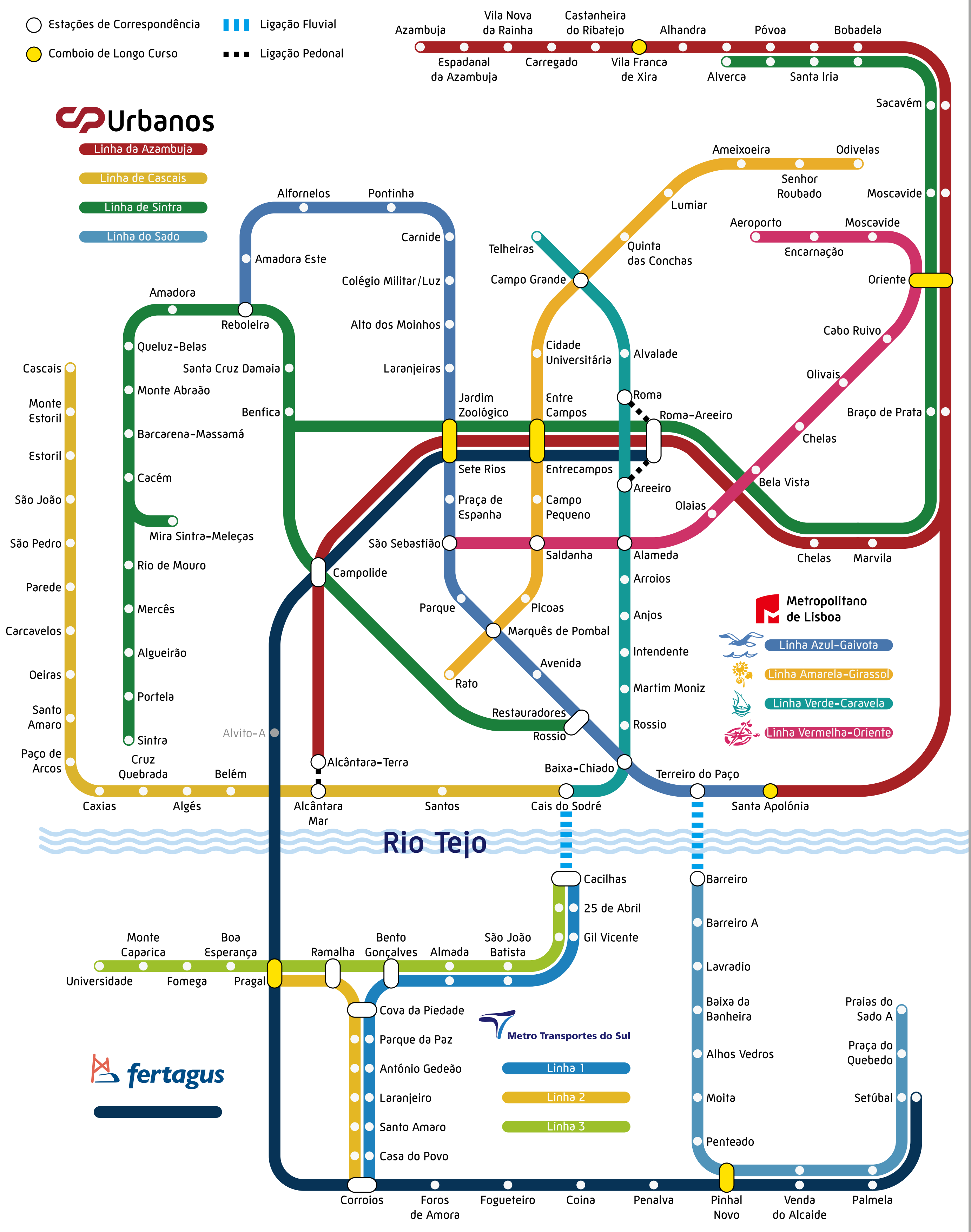
Z. **Acceleration in Convex Optimization**
Damien Scieur (INRIA, France)

SpaRTaN
Sparse Representations and Compressed Sensing Training Network
www.spartan-itn.eu

MacSeNet
Machine Sensing Training Network
www.macsenet.eu

Notes

Metro and Urban Routes for Central Lisbon



Campus Map

Map Data ©2017 Google