

Références

- ABADI, M. et al. *TensorFlow : Large-Scale Machine Learning on Heterogeneous Systems*. 2015.
- ABELSON, H., G. J. SUSSMAN et J. SUSSMAN. *Structure and interpretation of computer programs*. Justin Kelly, 1996.
- ABERNETHY, J., P. BARTLETT et A. RAKHLIN. “Multitask Learning with Expert Advice”. In : *COLT*. 2007, p. 484-498.
- ABSIL, P., R. MAHONY et B. ANDREWS. “Convergence of the Iterates of Descent Methods for Analytic Cost Functions”. In : *SIAM J. Optim.* 16.2 (2005), p. 531-547.
- ACHLIOPTAS, D. “Database-friendly random projections : Johnson-Lindenstrauss with binary coins”. In : *Journal of computer and System Sciences* 66.4 (2003), p. 671-687.
- ADAMS, A., J. BAEK et A. DAVIS. “Fast High-Dimensional Filtering Using the Permutohedral Lattice”. In : *SIGGRAPH*. 2010.
- ADAMS, A. et al. “Gaussian KD-Trees for Fast High-Dimensional Filtering”. In : *SIGGRAPH*. 2009.
- ADLER, R. J. *An introduction to continuity, extrema, and related topics for general Gaussian processes*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 12. Hayward, CA : Institute of Mathematical Statistics, 1990, p. x+160.
- AFONSO, M. V., J. M. BIOUCAS-DIAS et M. A. T. FIGUEIREDO. “An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems”. In : *IEEE Trans. Image Process.* 20.3 (2011), p. 681-695.
- AHARON, M. et M. ELAD. “Sparse and redundant modeling of image content using an image-signature-dictionary”. In : *SIAM J. Imaging Sci.* 1.3 (2008), p. 228-247.
- AHARON, M., M. ELAD et A. BRUCKSTEIN. “K-SVD : An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation”. In : *IEEE Trans. Signal Process.* 54.11 (2006), p. 4311-4322.
- AHLSWEDE, R. et A. WINTER. “Strong converse for identification via quantum channels”. In : *IEEE Trans. Inf. Theory* 48.3 (2002), p. 569-579.
- AILON, N. et B. CHAZELLE. “The fast Johnson-Lindenstrauss transform and approximate nearest neighbors”. In : *SIAM J. Comput.* 39.1 (2009), p. 302-322.
- AITKEN, A. “On Bernoulli’s numerical solution of algebraic equations”. In : *Proceedings of the Royal Society of Edinburgh* 46 (1926), p. 289-305.
- AKAIKE, H. “A new look at the statistical model identification”. In : *IEEE Trans. Automat. Control* AC-19 (1974), p. 716-723.
- ALAHY, A. et al. “Sparsity driven people localization with a heterogeneous network of cameras”. In : *J. Math. Imaging Vis.* 41.1 (2011), p. 39-58.
- ALEXANDER, S. K. “Multiscale Methods in Image Modelling and Image Processing”. Thèse de doct. University of Waterloo, 2005.
- ALEXANDER, S. K., E. R. VRSCAY et S. TSURUMI. “A Simple, General Model for the Affine Self-similarity of Images”. In : *ICIAR*. 2008, p. 192-203.

- ALFONS, A., C. CROUX et S. GELPER. “Sparse least trimmed squares regression for analyzing high-dimensional large data sets”. In : *Ann. Appl. Stat.* 7.1 (2013), p. 226-248.
- ALQUIER, P. “PAC-Bayesian bounds for randomized empirical risk minimizers”. In : *Math. Methods Statist.* 17.4 (2008), p. 279-304.
- ALQUIER, P. et K. LOUNICI. “PAC-Bayesian Bounds for Sparse Regression Estimation with Exponential Weights”. In : *Electron. J. Stat.* 5 (2010), p. 127-145.
- ALVAREZ, L. et L. MAZORRA. “Signal and image restoration using shock filters and anisotropic diffusion”. In : *SIAM J. Numer. Anal.* 31.2 (1994), p. 590-605.
- ALVAREZ, L. et al. “Axioms and fundamental equations of image processing”. In : *Arch. Rational Mech. Anal.* 123.3 (1993), p. 199-257.
- AMARI, S-I. et al. *Differential geometry in statistical inference*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 10. Hayward, CA : Institute of Mathematical Statistics, 1987.
- AMIT, Y. et D. GEMAN. “Shape quantization and recognition with randomized trees”. In : *Neural Comput.* 9 (7 1997), p. 1545-1588.
- ANDERSON, D. G. “Iterative procedures for nonlinear integral equations”. In : *Journal of the ACM* (1965).
- ANSCOMBE, F. J. “The transformation of Poisson, binomial and negative-binomial data”. In : *Biometrika* 35 (1948), p. 246-254.
- ANTONIADIS, A. “Comments on : ℓ_1 -penalization for mixture regression models”. In : *TEST* 19.2 (2010), p. 257-258.
- ARAVKIN, A. et D. DAVIS. “A SMART Stochastic Algorithm for Nonconvex Optimization with Applications to Robust Machine Learning”. In : *arXiv preprint arXiv :1610.01101* (2016).
- ARGYRIOU, A., T. EVGENIOU et M. PONTIL. “Convex multi-task feature learning”. In : *Machine Learning* 73.3 (2008), p. 243-272.
- “Multi-task feature learning”. In : *NIPS*. 2006, p. 41-48.
- ARIAS-CASTRO, E. et D. L. DONOHO. “Does median filtering truly preserve edges better than linear filtering?” In : *Ann. Statist.* 37.3 (2009), p. 1172-1206.
- ARIAS-CASTRO, E. et K. LOUNICI. “Variable Selection with Exponential Weights and l_0 -Penalization”. In : *arXiv preprint arXiv :1208.2635* (2012).
- ARIAS-CASTRO, E., J. SALMON et R. WILLETT. “Oracle inequalities and minimax rates for non-local means and related adaptive kernel-based methods”. In : *SIAM J. Imaging Sci.* 5.3 (2012), p. 944-992.
- ARLOT, S. et F. BACH. “Data-driven calibration of linear estimators with minimal penalties”. In : *NIPS*. 2009, p. 46-54.
- ARLOT, S. et A. CELISSE. “A survey of cross-validation procedures for model selection”. In : *Statistics surveys* 4 (2010), p. 40-79.
- ARNOLD, T. et R. TIBSHIRANI. “Efficient implementations of the generalized lasso dual path algorithm”. In : *arXiv preprint arXiv :1405.3222* (2014).

- ARTHUR, D. et S. VASSILVITSKII. “k-means++ : The advantages of careful seeding”. In : *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. 2007, p. 1027-1035.
- ARTZNER, P. et al. “Coherent Measures of Risk”. In : *Mathematical Finance* 9.3 (1999), p. 203-228.
- AUDIBERT, J-Y. “Aggregated estimators and empirical complexity for least square regression”. In : *Ann. Inst. H. Poincaré Probab. Statist.* 40.6 (2004), p. 685-736.
- “Fast learning rates in statistical inference through aggregation”. In : *Ann. Statist.* 37.4 (2009), p. 1591-1646.
- “Progressive mixture rules are deviation suboptimal”. In : *NIPS*. 2007, p. 41-48.
- AUDIBERT, J-Y. et A. B. TSYBAKOV. “Fast learning rates for plug-in classifiers”. In : *Ann. Statist.* 35.2 (2007), p. 608-633.
- AVELLA-MEDINA, M. et E. M. RONCHETTI. “Robust and consistent variable selection in high-dimensional generalized linear models”. In : *Biometrika* 105.1 (2018), p. 31-44.
- AWATE, S. P. et R. T. WHITAKER. “Unsupervised, Information-Theoretic, Adaptive Image Filtering for Image Restoration”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 28.3 (2006), p. 364-376.
- AZAÏS, J.-M., Y. DE CASTRO et F. GAMBOA. “Spike detection from inaccurate samplings”. In : *Appl. Comput. Harmon. Anal.* 38.2 (2015), p. 177-195.
- AZZABOU, N. “Variable Bandwidth Image Models for Texture-Preserving Enhancement of Natural Images”. Thèse de doct. École des Ponts, 2008.
- AZZABOU, N., N. PARAGIOS et F. GUICHARD. “Image Denoising Based on Adapted Dictionary Computation”. In : *ICIP*. 2007, p. 109-112.
- BACH, F. “Bolasso : model consistent Lasso estimation through the bootstrap”. In : *ICML*. 2008, p. 33-40.
- “Consistency of the group Lasso and multiple kernel learning”. In : *J. Mach. Learn. Res.* 9 (2008), p. 1179-1225.
- “Consistency of trace norm minimization”. In : *J. Mach. Learn. Res.* 9.Jun (2008), p. 1019-1048.
- BACH, F. et al. “Convex optimization with sparsity-inducing norms”. In : *Foundations and Trends in Machine Learning* 4.1 (2012), p. 1-106.
- BAGCI, H. et al. “Sparsity-regularized Born iterations for electromagnetic inverse scattering”. In : *Antennas and Propagation Society International Symposium, 2008. AP-S 2008. IEEE*. 2008, p. 1-4.
- BAKIN, S. “Adaptive regression and model selection in data mining problems”. Thèse de doct. The Australian National University, 1999.
- BALMAND, S. et A. S. DALALYAN. “On estimation of the diagonal elements of a sparse precision matrix”. In : *Electron. J. Stat.* 10.1 (2016), p. 1551-1579.
- BALZANO, L., R. D. NOWAK et B. RECHT. “Online identification and tracking of subspaces from highly incomplete information”. In : *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*. 2010, p. 704-711.

- BANERJEE, A. et P. BURLINA. “Efficient particle filtering via sparse kernel density estimation”. In : *IEEE Trans. Image Process.* 19.9 (2010), p. 2480-2490.
- BANERJEE, A. et al. “Clustering with Bregman divergences”. In : *J. Mach. Learn. Res.* 6 (2005), p. 1705-1749.
- BARASH, D. “A Fundamental Relationship between Bilateral Filtering, Adaptive Smoothing, and the Nonlinear Diffusion Equation”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 24.6 (2002), p. 844-847.
- BARASH, D. et D. COMANICIU. “A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift”. In : *Image Video Comput.* 22.1 (2004), p. 73-81.
- BARAUD, Y., C. GIRAUD et S. HUET. “Estimator selection in the Gaussian setting”. In : *submitted* (2010).
- “Gaussian model selection with an unknown variance”. In : *Ann. Statist.* 37 (2009), p. 630-672.
- BARBER, R. F. et E. J. CANDÈS. “A knockoff filter for high-dimensional selective inference”. In : *Ann. Statist.* 47.5 (2019), p. 2504-2537.
- BARRON, A. R., L. BIRGÉ et P. MASSART. “Risk bounds for model selection via penalization”. In : *Probab. Theory Related Fields* 113.3 (1999), p. 301-413.
- BARRON, A. R. et al. “Approximation and learning by greedy algorithms”. In : *Ann. Statist.* 36.1 (2008), p. 64-94.
- BARTLETT, P. L. et S. MENDELSON. “Rademacher and Gaussian complexities : risk bounds and structural results”. In : *J. Mach. Learn. Res.* 3.Spec. Issue Comput. Learn. Theory (2002), p. 463-482.
- BAUSCHKE, H. H. et P. L. COMBETTES. *Convex analysis and monotone operator theory in Hilbert spaces*. New York : Springer, 2011, p. xvi+468.
- BECK, A. *First-Order Methods in Optimization*. T. 25. SIAM, 2017.
- BECK, A. et Y. C. ELDAR. “Sparsity constrained nonlinear optimization : Optimality conditions and algorithms”. In : *SIAM J. Optim.* 23.3 (2013), p. 1480-1509.
- BECK, A., E. PAUWELS et S. SABACH. “The cyclic block conditional gradient method for convex optimization problems”. In : *SIAM J. Optim.* 25.4 (2015), p. 2024-2049.
- BECK, A. et M. TEBoulLE. “A fast iterative shrinkage-thresholding algorithm for linear inverse problems”. In : *SIAM J. Imaging Sci.* 2.1 (2009), p. 183-202.
- “Gradient-based algorithms with applications to signal-recovery problems”. In : *Convex optimization in signal processing and communications*. Cambridge Univ. Press, Cambridge, 2010, p. 42-88.
- “Mirror descent and nonlinear projected subgradient methods for convex optimization”. In : *Operations Research Letters* 31.3 (2003), p. 167-175.
- “Smoothing and first order methods : A unified framework”. In : *SIAM J. Optim.* 22.2 (2012), p. 557-580.
- BECK, A. et L. TETRUASHVILI. “On the convergence of block coordinate type methods”. In : *SIAM J. Imaging Sci.* 23.4 (2013), p. 651-694.

- BECKER, S. R., E. J. CANDÈS et M. C. GRANT. “Templates for convex cone problems with applications to sparse signal recovery”. In : *Math. Program. Comput.* 3.3 (2011), p. 165-218.
- BEHNEL, S. et al. “Cython : The Best of Both Worlds”. In : *Computing in Science Engineering* 13.2 (2011), p. 31-39.
- BEKHTI, Y. et al. “A hierarchical Bayesian perspective on majorization-minimization for non-convex sparse regression : application to M/EEG source imaging”. In : *Inverse Problems* 34.8 (2018), p. 085010.
- BELLEÇ, P. C., G. LECUÉ et A. B. TSYBAKOV. “Slope meets Lasso : improved oracle bounds and optimality”. In : *arXiv preprint arXiv :1605.08651* (2016).
- BELLEÇ, P. C., J. SALMON et S. VAITER. “A sharp oracle inequality for Graph-Slope”. In : *Electron. J. Statist.* 11.2 (2017), p. 4851-4870. URL : <https://arxiv.org/pdf/1706.06977.pdf>.
- BELLET, A., R. GUERRAOUÏ et H. HENDRIKX. *Who started this rumor ? Quantifying the natural differential privacy guarantees of gossip protocols*. Rapp. tech. arXiv preprint arXiv :1902.07138, 2019.
- BELLET, A., A. HABRARD et M. SEBBAN. *A Survey on Metric Learning for Feature Vectors and Structured Data*. Rapp. tech. arXiv :1306.6709, 2013.
- *Metric Learning*. Morgan & Claypool, 2015.
- BELLMAN, R. “The theory of dynamic programming”. In : *Bull. Amer. Math. Soc.* 60.6 (nov. 1954), p. 503-515.
- BELLONI, A. et V. CHERNOZHUKOV. “Least squares after model selection in high-dimensional sparse models”. In : *Bernoulli* 19.2 (2013), p. 521-547.
- BELLONI, A., V. CHERNOZHUKOV et L. WANG. “Square-root Lasso : pivotal recovery of sparse signals via conic programming”. In : *Biometrika* 98.4 (2011), p. 791-806.
- BELLONI, A., M. ROSENBAUM et A. B. TSYBAKOV. “An $\{l_1, l_2, l_\infty\}$ -Regularization Approach to High-Dimensional Errors-in-variables Models”. In : *arXiv preprint arXiv :1412.7216* (2014).
- “Linear and Conic Programming Estimators in High-Dimensional Errors-in-variables Models”. In : *arXiv preprint arXiv :1408.0241* (2014).
- BEN-DAVID, S., D. PAL et S. SHALEV-SHWARTZ. “Agnostic Online Learning”. In : *COLT*. 2009.
- BEN-DAVID, S. et R. SCHULLER. “Exploiting task relatedness for multiple task learning”. In : *COLT*. 2003, p. 567-580.
- BEN-ISRAEL, A. et T. N. E. GREVILLE. *Generalized inverses*. Second. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 15. Theory and applications. New York : Springer-Verlag, 2003, p. xvi+420.
- BEN-TAL, A., L. EL GHAOUÏ et A. NEMIROVSKI. *Robust optimization*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2009, p. xxii+542.
- BEN-TAL, A. et M. TEBoulLE. “An old-new concept of convex risk measures : the optimized certainty equivalent”. In : *Math. Finance* 17.3 (2007), p. 449-476.

- BEN-TAL, A. et M. TEBoulLE. “Expected Utility, Penalty Functions, and Duality in Stochastic Nonlinear Programming”. In : *Management Sci.* 32.11 (1986), p. 1445-1466.
- BENDORY, T., S. DEKEL et A. FEUER. “Exact recovery of Dirac ensembles from the projection onto spaces of spherical harmonics”. In : *Constr. Approx.* (2014), p. 1-25.
- “Exact recovery of non-uniform splines from the projection onto spaces of algebraic polynomials”. In : *J. Approx. Theory* 182 (2014), p. 7-17.
- “Robust recovery of stream of pulses using convex optimization”. In : *J. Math. Anal. Appl.* 442.2 (2016), p. 511-536.
- “Super-resolution on the sphere using convex optimization”. In : *IEEE Trans. Signal Process.* 63.9 (2015), p. 2253-2262.
- BENEDETTI, J. K. “On the nonparametric estimation of regression functions”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 39.2 (1977), p. 248-253.
- BENGIO, Y., A. COURVILLE et P. VINCENT. “Representation Learning : A Review and New Perspectives”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 35.8 (2013), p. 1798-1828.
- BENHAMOU, E. et V. MELOT. *Seven proofs of the Pearson Chi-squared independence test and its graphical interpretation*. Rapp. tech. Arxiv preprint arXiv :1808.09171, 2018.
- BENJAMINI, Y. et Y. HOCHBERG. “Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 57.1 (1995), p. 289-300.
- BENOIT, L. et al. “Sparse image representation with epitomes”. In : *CVPR*. 2011, p. 2913-2920.
- BERGEAUD, F. et S. MALLAT. “Matching pursuit of images”. In : *ICIP*. 1995, p. 53.
- BERGH, J. et J. LÖFSTRÖM. *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Berlin : Springer-Verlag, 1976, p. x+207.
- BERMAN, A. et R. J. PLEMMONS. *Nonnegative matrices in the mathematical sciences*. T. 9. Classics in Applied Mathematics. Revised reprint of the 1979 original. Philadelphia, PA : Society for Industrial et Applied Mathematics (SIAM), 1994, p. xx+340.
- BERMAN, S. M. “Limit theorems for the maximum term in stationary sequences”. In : *Ann. Math. Statist.* 35 (1964), p. 502-516.
- BERRADA, L., A. ZISSERMAN et M. P. KUMAR. “Smooth loss functions for deep top-k classification”. In : 2018.
- BERTHET, Q. et al. “Learning with differentiable perturbed optimizers”. In : *arXiv preprint arXiv :2002.08676* (2020).
- BERTRAND, Q. et al. “Handling correlated and repeated measurements with the smoothed multivariate square-root Lasso”. In : *NeurIPS*. 2019.
- BERTRAND, Q. et al. “Implicit differentiation of Lasso-type models for hyperparameter optimization”. In : *ICML*. 2020.
- BERTSEKAS, D. P. *Convex optimization algorithms*. Athena Scientific Belmont, 2015.
- BERTSIMAS, D., D. B. BROWN et C. CARAMANIS. “Theory and applications of robust optimization”. In : *SIAM Rev.* 53.3 (2011), p. 464-501.

- BERTSIMAS, D., A. KING et R. MAZUMDER. “Best subset selection via a modern optimization lens”. In : *Ann. Statist.* 44.2 (2016), p. 813-852.
- BHASKAR, B. N., G. TANG et B. RECHT. *Atomic norm denoising with applications to line spectral estimation*. Rapp. tech. University of Wisconsin-Madison, 2012.
- BHATIA, R. *Matrix analysis*. T. 169. Graduate Texts in Mathematics. New York : Springer-Verlag, 1997.
- BIANCHI, P., W. HACHEM et F. IUTZELER. “A stochastic coordinate descent primal-dual algorithm and applications to large-scale composite optimization”. In : *arXiv preprint arXiv :1407.0898* (2014).
- BIANCHI, P. et J. JAKUBOWICZ. “Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization”. In : *IEEE Trans. Autom. Control* 58.2 (2013), p. 391-405.
- BIAU, G. et K. BLEAKLEY. “Statistical inference on graphs”. In : *Statistics & Decisions* 24 (2006), p. 209-232.
- BIAU, G., L. DEVROYE et G. LUGOSI. “Consistency of random forests and other averaging classifiers”. In : *J. Mach. Learn. Res.* 9 (2008), p. 2015-2033.
- BIAU, G. et al. “COBRA : A Nonlinear Aggregation Strategy”. In : *arXiv preprint arXiv :1303.2236* (2013).
- BICKEL, P. J. et K. A. DOKSUM. *Mathematical statistics*. Basic ideas and selected topics, Holden-Day Series in Probability and Statistics. San Francisco, Calif. : Holden-Day Inc., 1976.
- BICKEL, P. J., Y. RITOV et A. B. TSYBAKOV. “Simultaneous analysis of Lasso and Dantzig selector”. In : *Ann. Statist.* 37.4 (2009), p. 1705-1732.
- BIEN, J., J. TAYLOR et R. TIBSHIRANI. “A lasso for hierarchical interactions”. In : *Ann. Statist.* 41.3 (2013), p. 1111-1141.
- BILCU, R. C. et M. VEHVILAINEN. “Combined non-local averaging and intersection of confiden intervals for image de-noising”. In : *ICIP*. 2008, p. 1736-1739.
- “Fast nonlocal means for image denoising”. In : *Electronic Imaging 2007*. International Society for Optics et Photonics. 2007, 65020R-65020R.
- BINGHAM, E. et H. MANNILA. “Random projection in dimensionality reduction : applications to image and text data”. In : *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM. 2001, p. 245-250.
- BIRGÉ, L. “Model selection via testing : an alternative to (penalized) maximum likelihood estimators”. In : *Ann. Inst. H. Poincaré Probab. Statist.* 42.3 (2006), p. 273-325.
- BIRGÉ, L. et P. MASSART. “Gaussian model selection”. In : *J. Eur. Math. Soc. (JEMS)* 3.3 (2001), p. 203-268.
- “Minimum contrast estimators on sieves : exponential bounds and rates of convergence”. In : *Bernoulli* 4.3 (1998), p. 329-375.
- BISHOP, C. M. *Pattern recognition and machine learning*. Information Science and Statistics. New York : Springer, 2006.

- BLACK, M. J. et G. SAPIRO. “Edges as Outliers : Anisotropic Smoothing Using Local Image Statistics”. In : *SCALE-SPACE '99 : Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision*. 1999, p. 259-270.
- BLONDEL, M., A. FUJINO et N. UEDA. “Convex Factorization Machines”. In : *Machine Learning and Knowledge Discovery in Databases*. Springer, 2015, p. 19-35.
- BLONDEL, M., K. SEKI et K. UEHARA. “Block coordinate descent algorithms for large-scale sparse multiclass classification”. In : *Machine Learning* 93.1 (2013), p. 31-52.
- BLOOMFIELD, P. et W. L. STEIGER. *Least absolute deviations*. T. 6. Progress in Probability and Statistics. Theory, applications, and algorithms. Birkhäuser Boston, Inc., Boston, MA, 1983, p. xiv+349.
- BLU, T. et F. LUISIER. “The SURE-LET approach to image denoising”. In : *IEEE Trans. Image Process.* 16.11 (2007), p. 2778-2786.
- BLUMENSATH, T. et M. E. DAVIES. “Iterative thresholding for sparse approximations”. In : *J. Fourier Anal. Appl.* 14.5 (2008), p. 629-654.
- BOBADILLA, J. et al. “Recommender systems survey”. In : *Knowledge-Based Systems* 46.0 (2013), p. 109-132.
- BOGDAN, M. et al. “SLOPE-adaptive variable selection via convex optimization”. In : *Ann. Appl. Stat.* 9.3 (2015), p. 1103.
- BOISBUNON, A., R. FLAMARY et A. RAKOTOMAMONJY. “Active set strategy for high-dimensional non-convex sparse optimization problems”. In : *ICASSP*. 2014, p. 1517-1521.
- BOLLAPRAGADA, R., D. SCIEUR et A. D’ASPREMONT. “Nonlinear Acceleration of Momentum and Primal-Dual Algorithms”. In : *ArXiv e-print* abs/1810.04539 (2018).
- BOLSTAD, A., B. VAN VEEN et R. NOWAK. “Space-time event sparse penalization for magneto-/electroencephalography”. In : *NeuroImage* 46.4 (2009), p. 1066-1081.
- BONDELL, H. D. et B. J. REICH. “Simultaneous regression shrinkage, variable selection, and supervised clustering of predictors with OSCAR”. In : *Biometrics* 64.1 (2008), p. 115-123.
- BONNEFOY, A. et al. “A dynamic screening principle for the lasso”. In : *EUSIPCO*. 2014, p. 6-10.
- “Dynamic screening : accelerating first-order algorithms for the Lasso and Group-Lasso”. In : *IEEE Trans. Signal Process.* 63.19 (2015), p. 20.
- BONNET, P. et al. “Plant identification : experts vs. machines in the era of deep learning”. In : *Multimedia Tools and Applications for Environmental & Biodiversity Informatics*. Springer, 2018, p. 131-149.
- BOOS, D. D. et R. J. SERFLING. “A Note on Differentials and the CLT and LIL for Statistical Functions, with Application to M -Estimates”. In : *Ann. Statist.* 8.3 (1980), p. 618-624.
- BORKOWSKI, K. J. et al. “Radioactive Scandium in the Youngest Galactic Supernova Remnant G1.9+ 0.3”. In : *The Astrophysical Journal Letters* 724 (2010), p. L161.
- BORWEIN, J. M. et A. S. LEWIS. *Convex analysis and nonlinear optimization. Theory and examples*. 2nd ed. Theory and examples. Springer, New York, 2006.

- BOSER, B. E., I. M. GUYON et V. N. VAPNIK. "A training algorithm for optimal margin classifiers". In : *Proceedings of the fifth annual workshop on Computational learning theory*. ACM. 1992, p. 144-152.
- BOSWELL, D. et T. FOUCHER. *The Art of Readable Code*. O'Reilly Media, 2011.
- BOTELLA, C. et al. "A deep learning approach to species distribution modelling". In : *Multimedia Tools and Applications for Environmental & Biodiversity Informatics*. Springer, 2018, p. 169-199.
- BOTTOU, L. "Online learning and stochastic approximations". In : *On-line learning in neural networks* 17.9 (1998), p. 142.
- BOTTOU, L. et O. BOUSQUET. "The tradeoffs of large scale learning". In : *NIPS*. 2008, p. 161-168.
- BOUCHERON, S., O. BOUSQUET et G. LUGOSI. "Theory of classification : a survey of some recent advances". In : *ESAIM Probab. Stat.* 9 (2005), p. 323-375.
- BOUCHERON, S., G. LUGOSI et P. MASSART. *Concentration Inequalities : A Nonasymptotic Theory of Independence*. Second. Oxford University Press, 2013.
- BOUGLEUX, S., A. ELMOATAZ et M. MELKEMI. "Local and Nonlocal Discrete Regularization on Weighted Graphs for Image and Mesh Processing". In : *Int. J. Comput. Vision* 84.2 (2009), p. 220-236.
- BOULANGER, J. et al. "Patch-based nonlocal functional for denoising fluorescence microscopy image sequences." In : *IEEE Trans. Med. Imag.* 29.2 (2010), p. 442-454.
- BOVIK, A. C. *Handbook of Image and Video Processing (Communications, Networking and Multimedia)*. Orlando, FL, USA : Academic Press, Inc., 2005.
- BOX, G. E. P. "Robustness in the strategy of scientific model building". In : *Robustness in statistics*. Elsevier, 1979, p. 201-236.
- BOYD, S. et L. VANDENBERGHE. *Convex optimization*. Cambridge University Press, 2004, p. xiv+716.
- BOYD, S. et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers". In : *Foundations and Trends in Machine Learning* 3.1 (2011), p. 1-122.
- BOYD, S. et al. "Randomized gossip algorithms". In : *IEEE Trans. Inf. Theory* 52.6 (2006), p. 2508-2530.
- BOYER, C., Y. DE CASTRO et J. SALMON. "Adapting to unknown noise level in sparse deconvolution". In : *Information and Inference* 6.3 (2017), p. 310-348. URL : <http://arxiv.org/pdf/1606.04760v1>.
- BRADLEY, D. M. et J. A. BAGNELL. "Differential sparse coding". In : *NIPS*. T. 22. 2009.
- BREDIES, K. et H. K. PIKKARAINEN. "Inverse problems in spaces of measures". In : *ESAIM : Control, Optimisation and Calculus of Variations* 19.01 (2013), p. 190-218.
- BREGMAN, L. M. "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming". In : *Comput. Math. Math. Phys.* 7.3 (1967), p. 200-217.

- BREHENY, P. et J. HUANG. “Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection”. In : *Ann. Appl. Stat.* 5.1 (2011), p. 232.
- BREIMAN, L. “Bagging Predictors”. In : *Mach. Learn.* 24.2 (1996), p. 123-140.
- “Better subset regression using the nonnegative garrote”. In : *Technometrics* 37 (1995), p. 373-384.
- “Random Forests”. In : *Mach. Learn.* 45.1 (2001), p. 5-32.
- “Stacked regressions”. In : *Mach. Learn.* 24.1 (1996), p. 49-64.
- BREIMAN, L. et al. *Classification and regression trees*. Wadsworth Statistics/Probability Series. Belmont, CA : Wadsworth Advanced Books et Software, 1984, p. x+358.
- BRÉMAUD, P. *Markov chains*. T. 31. Texts in Applied Mathematics. Gibbs fields, Monte Carlo simulation, and queues. New York : Springer-Verlag, 1999, p. xviii+444.
- BREZIS, H. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. New York : Springer, 2011.
- BRIGGS, W. L. et V. E. HENSON. *The DFT An Owners’ Manual for the Discrete Fourier Transform*. An owner’s manual for the discrete Fourier transform. Philadelphia, PA : Society for Industrial et Applied Mathematics (SIAM), 1995.
- BRINKMANN, E.-M. et al. “Bias-reduction in variational regularization”. In : *J. Math. Imaging Vis.* 59.3 (2017), p. 534-566.
- BROWN, L. et al. “The root-unroot algorithm for density estimation as implemented via wavelet block thresholding”. In : *Probab. Theory Related Fields* 146.3-4 (2010), p. 401-433.
- BROX, T. et D. CREMERS. “Iterated Nonlocal Means for Texture Restoration”. In : *SSVM*. T. 4485. Lecture Notes in Computer Science. 2007, p. 13-24.
- BROX, T., O. KLEINSCHMIDT et D. CREMERS. “Efficient Nonlocal Means for Denoising of Textural Patterns”. In : *IEEE Trans. Image Process.* 17.7 (2008), p. 1083-1092.
- BUADES, A. “Image and movie denoising by non local means”. Thèse de doct. Universitat de les Illes Balears, 2006.
- BUADES, A., B. COLL et J.-M. MOREL. “A review of image denoising algorithms, with a new one”. In : *Multiscale Model. Simul.* 4.2 (2005), p. 490-530.
- “Non-local Means Denoising”. In : *Image Processing on Line* (2009).
- “Nonlocal Image and Movie Denoising”. In : *Int. J. Comput. Vision* 76.2 (2008), p. 123-139.
- “The staircasing effect in neighborhood filters and its solution”. In : *IEEE Trans. Image Process.* 15.6 (2006), p. 1499-1505.
- BUADES, A., M. LEBRUN et J.-M. MOREL. “Implementation of the “non-local Bayes” image denoising algorithm”. In : *Image Processing On Line* (2012).
- BUBECK, S. “Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems”. In : *Foundations and Trends in Machine Learning* 5.1 (2012), p. 1-122.
- BUCKHEIT, J. B. et D. L. DONOHO. “WaveLab and Reproducible Research”. In : *Wavelets and Statistics*. Lect. Notes Statist. 103. New York : Springer-Verlag, 1995, p. 55-81.

- BÜHLMANN, P. “High-dimensional statistics, with applications to genome-wide association studies”. In : *EMS Surveys in Mathematical Sciences* 4.1 (2017), p. 45-75.
- “Statistical significance in high-dimensional linear models”. In : *Bernoulli* 19.4 (sept. 2013), p. 1212-1242.
- BÜHLMANN, P. et S. VAN DE GEER. *Statistics for high-dimensional data*. Springer Series in Statistics. Methods, theory and applications. Heidelberg : Springer, 2011.
- BÜHLMANN, P. et B. YU. “Boosting with the L2 loss : regression and classification”. In : *J. Amer. Statist. Assoc.* 98.462 (2003), p. 324-339.
- BUNEA, F., J. LEDERER et Y. SHE. “The Group Square-Root Lasso : Theoretical Properties and Fast Algorithms”. In : *IEEE Trans. Inf. Theory* 60.2 (2014), p. 1313-1325.
- BUNEA, F. et A. NOBEL. “Sequential Procedures for Aggregating Arbitrary Estimators of a Conditional Mean”. In : *IEEE Trans. Inf. Theory* 54.4 (2008), p. 1725-1735.
- BUNEA, F., A. B. TSYBAKOV et M. H. WEGKAMP. “Aggregation for Gaussian regression”. In : *Ann. Statist.* 35.4 (2007), p. 1674-1697.
- “Sparsity oracle inequalities for the Lasso”. In : *Electron. J. Stat.* 1 (2007), 169-194 (electronic).
- BURGER, H. C., C. J. SCHULER et S. HARMELING. “Image denoising with multi-layer perceptrons, part 1 : comparison with existing algorithms and with bounds”. In : *arXiv preprint arXiv :1211.1544* (2012).
- “Image denoising with multi-layer perceptrons, part 2 : training trade-offs and analysis of their mechanisms”. In : *arXiv preprint arXiv :1211.1552* (2012).
- BURGER, M. et S. OSHER. “Convergence rates of convex variational regularization”. In : *Inverse Problems* 20.5 (2004), p. 1411-1421.
- BURGER, M. et al. “Nonlinear inverse scale space methods”. In : *Communications in Mathematical Sciences* 4.1 (2006), p. 179-212.
- BURKE, J. V. et J. J. MORÉ. “On the identification of active constraints”. In : *SIAM J. Numer. Anal.* 25.5 (1988), p. 1197-1211.
- CAI, J.-F., E. J. CANDÈS et Z. SHEN. “A singular value thresholding algorithm for matrix completion”. In : *SIAM J. Optim.* 20.4 (2010), p. 1956-1982.
- CAI, T. T. “Adaptive wavelet estimation : a block thresholding and oracle inequality approach”. In : *Ann. Statist.* 27.3 (1999), p. 898-924.
- CAI, T. T. et W.-X. ZHOU. “A Max-Norm Constrained Minimization Approach to 1-Bit Matrix Completion”. In : *J. Mach. Learn. Res.* 14 (2013), p. 3619-3647.
- “Matrix completion via max-norm constrained optimization”. In : *CoRR* abs/1303.0341 (2013).
- CALAFIORE, G. C., L. EL GHAOUÏ et C. NOVARA. “Sparse identification of polynomial and posynomial models”. In : *IFAC Proceedings Volumes* 47.3 (2014), p. 3238-3243.
- CANDÈS, E. J. et D. L. DONOHO. “Curvelets-a surprisingly effective nonadaptive representation for objects with edges”. In : *Curve and Surface Fitting*. 1999, p. 105-120.
- CANDÈS, E. J. et C. FERNANDEZ-GRANDA. “Super-resolution from noisy data”. In : *J. Fourier Anal. Appl.* 19.6 (2013), p. 1229-1254.

- CANDÈS, E. J. et C. FERNANDEZ-GRANDA. “Towards a mathematical theory of super-resolution”. In : *Comm. Pure Appl. Math.* 67.6 (2014), p. 906-956.
- CANDÈS, E. J. et Y. PLAN. “Matrix completion with noise”. In : *Proceedings of the IEEE* 98.6 (2010), p. 925-936.
- “Near-ideal model selection by ℓ_1 minimization”. In : *Ann. Statist.* 37.5A (2009), p. 2145-2177.
- CANDÈS, E. J. et B. RECHT. “Exact matrix completion via convex optimization”. In : *Found. Comput. Math.* 9.6 (2009), p. 717-772.
- “Simple bounds for recovering low-complexity models”. In : *Math. Program.* (2012), p. 1-13.
- CANDÈS, E. J., J. ROMBERG et T. TAO. “Robust uncertainty principles : exact signal reconstruction from highly incomplete frequency information”. In : *IEEE Trans. Inf. Theory* 52.2 (2006), p. 489-509.
- CANDÈS, E. J. et T. TAO. “Decoding by linear programming”. In : *IEEE Trans. Inf. Theory* 51.12 (2005), p. 4203-4215.
- “Rejoinder : “The Dantzig selector : statistical estimation when p is much larger than n ””. In : *Ann. Statist.* 35.6 (2007), p. 2392-2404.
- “The Dantzig selector : statistical estimation when p is much larger than n ”. In : *Ann. Statist.* 35.6 (2007), p. 2313-2351.
- “The power of convex relaxation : near-optimal matrix completion”. In : *IEEE Trans. Inf. Theory* 56.5 (2010), p. 2053-2080.
- CANDÈS, E. J., M. B. WAKIN et S. P. BOYD. “Enhancing Sparsity by Reweighted l_1 Minimization”. In : *J. Fourier Anal. Applicat.* 14.5-6 (2008), p. 877-905.
- CANDÈS, E. J. et al. “Fast discrete curvelet transforms”. In : *Multiscale Model. Simul.* 5.3 (2006), p. 861-899.
- CANDÈS, E. J. et al. “Robust principal component analysis?” In : *J. ACM* 58.3 (2011), Art. 11, 37.
- CAPPÉ, O. et É. MOULINES. “On-line expectation-maximization algorithm for latent data models”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 71.3 (2009), p. 593-613.
- CARLIER, G. *Programmation Dynamique*. 2007.
- CARRANZA-ROJAS, J. et al. “Going deeper in the automated identification of Herbarium specimens”. In : *BMC evolutionary biology* 17.1 (2017), p. 181.
- CARROLL, R. J. et D. RUPPERT. *Transformation and weighting in regression*. T. 30. CRC Press, 1988.
- CARUANA, R. “Multitask learning”. In : *Learning to learn*. Springer, 1998, p. 95-133.
- CASELLA, G. et R. L. BERGER. *Statistical Inference*. 2^e éd. Duxbury Press, 2001.
- CATONI, O. “Universal” aggregation rules with exact bias bounds. Rapp. tech. Université Paris 6 Pierre et Marie Curie, 1999.
- “Challenging the empirical mean and empirical variance : a deviation study”. In : *Ann. Inst. Henri Poincaré Probab. Stat.* 48.4 (2012), p. 1148-1185.

- CATONI, O. *Pac-Bayesian supervised classification : the thermodynamics of statistical learning*. Beachwood, OH : Institute of Mathematical Statistics, 2007, p. xii+163.
- *Statistical learning theory and stochastic optimization*. T. 1851. Lecture Notes in Mathematics. Berlin : Springer-Verlag, 2004.
- CAUCHY, A. “Méthode générale pour la résolution des systèmes d’équations simultanées”. In : *Comp. Rend. Sci. Paris* 25.1847 (1847), p. 536-538.
- CAVALIER, L. “Nonparametric statistical inverse problems”. In : *Inverse Problems* 24.3 (2008), p. 034004, 19.
- CAVALIER, L. et A. B. TSYBAKOV. “Penalized blockwise Stein’s method, monotone oracles and sharp adaptive estimation”. In : *Math. Methods Statist.* 10.3 (2001), p. 247-282.
- “Sharp adaptation for inverse problems with random noise”. In : *Probab. Theory Related Fields* 123.3 (2002), p. 323-354.
- CAVALIER, L. et al. “Oracle inequalities for inverse problems”. In : *Ann. Statist.* 30.3 (2002), p. 843-874.
- CENSOR, Y. “Computational acceleration of projection algorithms for the linear best approximation problem”. In : *Linear Algebra and its Applications* 416.1 (2006), p. 111-123.
- CESA-BIANCHI, N. et G. LUGOSI. *Prediction, learning, and games*. Cambridge University Press, 2006.
- CESA-BIANCHI, N., Y. MANSOUR et G. STOLTZ. “Improved second-order bounds for prediction with expert advice”. In : *Mach. Learn.* 66 (2-3 2007), p. 321-352.
- CHAFAI, D. “Singular values of random matrices”. In : *Lecture Notes* (2009), p. 8-10.
- CHAMBOLLE, A. et C. DOSSAL. “How to make sure the iterates of FISTA converge”. In : *Preprint* (2014).
- CHAMBOLLE, A. et T. POCK. “A first-order primal-dual algorithm for convex problems with applications to imaging”. In : *J. Math. Imaging Vis.* 40.1 (2011), p. 120-145.
- CHANDRASEKARAN, V. et al. “The Convex Geometry of Linear Inverse Problems”. In : *Foundations of Computational Mathematics* 12.6 (2012), p. 805-849.
- CHANG, C.-C. et C.-J. LIN. “LIBSVM : a library for support vector machines”. In : *ACM transactions on intelligent systems and technology (TIST)* 2.3 (2011), p. 27.
- CHATTERJEE, P. et P. MILANFAR. “A generalization of non-local means via kernel regression”. In : *SPIE*. T. 6814. 2008.
- “Patch-based Near-Optimal Image Denoising”. In : *ICIP*. 2011.
- CHATTERJEE, S. et J. JAFAROV. “Prediction error of cross-validated Lasso”. In : *arXiv preprint arXiv :1502.06291* (2015).
- CHEN, B. et al. “Maximum block improvement and polynomial optimization”. In : *SIAM J. Optim.* 22.1 (2012), p. 87-107.
- CHEN, D., P. MORIN et U. WAGNER. “Absolute approximation of Tukey depth : theory and experiments”. In : *Comput. Geom.* 46.5 (2013), p. 566-573.
- CHEN, L.H.Y. “Poisson approximation for dependent trials”. In : *Ann. Probability* (1975), p. 534-545.

- CHEN, M., C. GAO et Z. REN. “A General Decision Theory for Huber’s ϵ -Contamination Model”. In : *Electron. J. Stat.* 10.2 (2016), p. 3752-3774.
- “Robust Covariance Matrix Estimation under Huber’s Contamination Model”. In : *Ann. Statist.* (2018).
- CHEN, S. et A. BANERJEE. “Alternating estimation for structured high-dimensional multi-response models”. In : *NIPS*. 2017, p. 2838-2848.
- CHEN, S. S. et D. L. DONOHO. “Atomic decomposition by basis pursuit”. In : *SPIE*. 1995.
- CHEN, S. S., D. L. DONOHO et M. A. SAUNDERS. “Atomic decomposition by basis pursuit”. In : *SIAM J. Sci. Comput.* 20.1 (1998), p. 33-61.
- CHEN, Y. et C. CARAMANIS. “Noisy and missing data regression : distribution-oblivious support recovery”. In : *ICML*. 2013, p. 383-391.
- CHEN, Y., C. CARAMANIS et S. MANNOR. “Robust sparse regression under adversarial corruption”. In : *ICML*. 2013, p. 774-782.
- CHENG, Y. “Mean Shift, Mode Seeking, and Clustering”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 17.8 (1995), p. 790-799.
- CHÉTELAT, D., J. LEDERER et J. SALMON. “Optimal two-step prediction in regression”. In : *Electron. J. Stat.* 11.1 (2017), p. 2519-2546.
- CHEVALIER, J.-A., J. SALMON et B. THIRION. “Statistical Inference with Ensemble of Clustered Desparsified Lasso”. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2018, p. 638-646.
- CHICHIGNOUD, M. “Performances statistiques d’estimateurs non-linéaires”. Thèse de doct. Université Aix-Marseille 1, 2010.
- CHICHIGNOUD, M. et J. LEDERER. “A robust, adaptive M-estimator for pointwise estimation in heteroscedastic regression”. In : *Bernoulli* 20.3 (2014), p. 1560-1599.
- CHICHIGNOUD, M., J. LEDERER et M. J. WAINWRIGHT. “Tuning Lasso for sup-norm optimality”. In : *J. Mach. Learn. Res.* 17 (2016).
- CHIQUET, J., P. GUTIERREZ et G. RIGAILL. “Fast tree inference with weighted fusion penalties”. In : *arXiv preprint arXiv :1407.5915* (2014).
- CHRÉTIEN, S. et S. DARSEES. “Sparse recovery with unknown variance : a LASSO-type approach”. In : *IEEE Trans. Inf. Theory* (2011).
- CHUNG, F. R. K. *Spectral Graph Theory*. T. 92. American Mathematical Society, 1997.
- CHZHEN, E. “Plug-in approach for sparse multi-label learning”. In : *(under review)* (2018).
- CHZHEN, E., M. HEBIRI et J. SALMON. “On Lasso refitting strategies”. In : *Bernoulli* 25.4A (nov. 2019), p. 3175-3200.
- CHZHEN, E. et al. “On the benefits of output sparsity for multi-label classification”. In : *ArXiv e-prints* (2018).
- CIARLET, P. G. *Introduction à l’analyse numérique matricielle et à l’optimisation. Cours et exercices corrigés*. Dunod, 2006.
- CLARKE, B. R. *Linear models : the theory and application of analysis of variance*. T. 634. John Wiley & Sons, 2008.
- CLARKE, F. H. *Method of Dynamic and Nonsmooth Optimization*. SIAM, 1989.

- CLARKSON, K. L. “Coresets, sparse greedy approximation, and the Frank-Wolfe algorithm”. In : *ACM Transactions on Algorithms (TALG)* 6.4 (2010), 63 :1-63 :30.
- CLÉMENÇON, S. “On U-processes and clustering performance”. In : *NIPS*. 2011, p. 37-45.
- CLÉMENÇON, S., G. LUGOSI et N. VAYATIS. “Ranking and empirical minimization of U-statistics”. In : *Ann. Statist.* 36.2 (2008), p. 844-874.
- COHEN, A. “All Admissible Linear Estimates of the Mean Vector”. In : *Ann. Math. Statist.* 37.2 (1966), p. 458-463.
- COHEN, A. et al. “Maximal spaces with given rate of convergence for thresholding algorithms”. In : *Appl. Comput. Harmon. Anal.* 11.2 (2001), p. 167-191.
- COIFMAN, R. R. et D. L. DONOHO. *Translation-invariant de-noising*. Rapp. tech. Stanford, 1995.
- COLIN, I. “Adaptation des méthodes d’apprentissage aux U-statistiques”. Thèse de doct. Télécom ParisTech, 2016.
- COLIN, I. et al. “Extending Gossip Algorithms to Distributed Estimation of U-Statistics”. In : *NIPS*. 2015, p. 271-279.
- “Gossip Dual Averaging for Decentralized Optimization of Pairwise Functions”. In : *ICML*. 2016, p. 1388-1396.
- COLLINS, M., S. DASGUPTA et R. E. SCHAPIRE. “A Generalization of Principal Components Analysis to the Exponential Family”. In : *NIPS*. 2002, p. 617-624.
- COMANICIU, D. et P. MEER. “Mean Shift Analysis and Applications”. In : *ICCV*. T. 2. Los Alamitos, CA, USA : IEEE Computer Society, 1999, p. 1197.
- “Mean Shift : A Robust Approach Toward Feature Space Analysis”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 24.5 (2002), p. 603-619.
- COMBETTES, P. L. “Perspective Functions : Properties, Constructions, and Examples”. In : *arXiv preprint arXiv :1610.01552* (2016).
- COMBETTES, P. L. et C. L. MÜLLER. “Perspective Functions : Proximal Calculus and Applications in High-Dimensional Statistics”. In : *J. Math. Anal. Appl.* (2016).
- COMBETTES, P. L. et J.-C. PESQUET. “Proximal splitting methods in signal processing”. In : *Fixed-point algorithms for inverse problems in science and engineering*. T. 49. Springer Optim. Appl. Springer, New York, 2011, p. 185-212.
- COMBETTES, P. L. et V. R. WAJS. “Signal recovery by proximal forward-backward splitting”. In : *Multiscale Modeling & Simulation* 4.4 (2005), p. 1168-1200.
- CONDAT, L. “A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms”. In : *Journal of Optimization Theory and Applications* 158.2 (2013), p. 460-479.
- “A Simple Trick to Speed Up and Improve the Non-Local Means”. In : *submitted,hal-00512801* (2010).
- CORMEN, T. H. et al. *Introduction to algorithms*. MIT press, 2001.
- CORNILLON, P.-A., N. HENGARTNER et E. MATZNER-LØBER. “Recursive Bias Estimation for multivariate regression smoothers”. In : *submitted* (2009).

- CORNILLON, P-A. et E. MATZNER-LØBER. *Régression avec R*. Springer, Collection Pratique R, 2011, p. 242.
- CORNILLON, P-A. et al. *Statistiques avec R*. Didact Statistiques. Presses Universitaires de Rennes, 2010.
- COTTRELL, M. et al. *Exercices de probabilités, Licence - Master - Écoles d'ingénieur*. 3^e. Cassini, 2001.
- COURANT, J. et al. *Informatique pour tous en classes préparatoires aux grandes écoles : Manuel d'algorithmique et programmation structurée avec Python*. Eyrolles, 2013.
- COVER, T. M. et J. A. THOMAS. *Elements of information theory*. Wiley Series in Telecommunications. A Wiley-Interscience Publication. New York : John Wiley & Sons Inc., 1991, p. xxiv+542.
- CRIMINISI, A., P. PÉREZ et K. TOYAMA. “Object Removal by Exemplar-Based Inpainting”. In : *CVPR*. T. 2. 2003, p. 721-728.
- “Region filling and object removal by exemplar-based image inpainting”. In : *IEEE Trans. Image Process.* 13.9 (2004), p. 1200-1212.
- CROW, F. C. “Summed-area tables for texture mapping”. In : *ACM SIGGRAPH Computer Graphics* 18.3 (1984), p. 207-212.
- DABOV, K. et al. “A non-local and shape-adaptive transform-domain collaborative filtering”. In : *LNLA*. Lausanne, 2008, p. 179-186.
- “BM3D Image Denoising with Shape-Adaptive Principal Component Analysis”. In : *Proc. Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS'09)*. 2009.
- “Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering”. In : *IEEE Trans. Image Process.* 16.8 (2007), p. 2080-2095.
- DAI, D., P. RIGOLLET et T. ZHANG. “Deviation optimal learning using greedy Q-aggregation”. In : *Ann. Statist.* 40.3 (2012), p. 1878-1905.
- DAI, D. et T. ZHANG. “Greedy Model Averaging”. In : *NIPS*. 2011, p. 1242-1250.
- DALALYAN, A. S. “SOCP based variance free Dantzig selector with application to robust estimation”. In : *C. R. Math. Acad. Sci. Paris* 350.15-16 (2012), p. 785-788.
- DALALYAN, A. S. et Y. CHEN. “Fused sparsity and robust estimation for linear models with unknown variance”. In : *NIPS*. 2012, p. 1268-1276.
- DALALYAN, A. S., M. HEBIRI et J. LEDERER. “On the Prediction Performance of the Lasso”. In : *Bernoulli* 23.1 (2017), p. 552-581.
- DALALYAN, A. S. et J. SALMON. “Competing against the best nearest neighbor filter in regression”. In : *ALT*. 2011, p. 129-143.
- “Sharp Oracle Inequalities for Aggregation of Affine Estimators”. In : *Ann. Statist.* 40.4 (2012), p. 2327-2355.
- DALALYAN, A. S. et A. B. TSYBAKOV. “Aggregation by Exponential Weighting, Sharp Oracle Inequalities and Sparsity”. In : *COLT*. 2007, p. 97-111.
- “Aggregation by exponential weighting, sharp PAC-Bayesian bounds and sparsity”. In : *Mach. Learn.* 72 (2008), p. 39-61.

- DALALYAN, A. S. et A. B. TSYBAKOV. “Mirror Averaging with Sparsity Priors”. In : *Bernoulli* (2012).
- “Sparse Regression Learning by Aggregation and Langevin Monte-Carlo”. In : *COLT*. 2009.
- “Sparse Regression Learning by Aggregation and Langevin Monte-Carlo”. In : *J. Comput. System Sci.* 78.5 (2012), p. 1423-1443.
- DALALYAN, A. S. et al. “Learning Heteroscedastic Models by Convex Programming under Group Sparsity”. In : *ICML*. 2013, p. 379-387.
- DANIELYAN, A., V. KATKOVNIK et K. EGIAZARIAN. “BM3D frames and variational image deblurring”. In : *IEEE Trans. Image Process.* 21.4 (2012), p. 1715-1728.
- DANIELYAN, A. et al. “Denoising of Multispectral Images via Nonlocal Groupwise spectrum-PCA”. In : *CGIV2010/MCS’10*. 2010, p. 261-266.
- DARBON, J. et al. “Fast Nonlocal Filtering Applied to Electron Cryomicroscopy”. In : *ISBI*. 2008, p. 1331-1334.
- DASGUPTA, S. et A. GUPTA. “An elementary proof of a theorem of Johnson and Lindenstrauss”. In : *Random Structures & Algorithms* 22.1 (2003), p. 60-65.
- DAUBECHIES, I. *Ten lectures on wavelets*. T. 61. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia, PA : Society for Industrial et Applied Mathematics (SIAM), 1992.
- DAUBECHIES, I., M. DEFRISE et C. DE MOL. “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint”. In : *Comm. Pure Appl. Math.* 57.11 (2004), p. 1413-1457.
- DAUBECHIES, I. et al. “Iteratively reweighted least squares minimization for sparse recovery”. In : *Comm. Pure Appl. Math.* 63.1 (2010), p. 1-38.
- DAUWE, A. et al. “A fast non-local image denoising algorithm”. In : *SPIE*. T. 6812. 2008, p. 1033-1038.
- DAVENPORT, M. A. et al. “1-Bit matrix completion”. In : *Information and Inference* 3.3 (2014), p. 189-223.
- DAVIS, G., S. MALLAT et M. AVELLANEDA. “Adaptive greedy approximations”. In : *Constr. Approx.* 13.1 (1997), p. 57-98.
- DAYE, J., J. CHEN et H. LI. “High-Dimensional Heteroscedastic Regression with an Application to eQTL Data Analysis”. In : *Biometrics* 68.1 (2012), p. 316-326.
- DAYE, Z. et X. JENG. “Shrinkage and model selection with correlated variables via weighted fusion”. In : *Comput. Statist. Data Anal.* 53.2 (2009), p. 284-1298.
- DE CASTRO, Y. et F. GAMBOA. “Exact reconstruction using Beurling minimal extrapolation”. In : *J. Math. Anal. Appl.* 395.1 (2012), p. 336-354.
- DE CASTRO, Y. et G. MIJOLE. “Non-uniform spline recovery from small degree polynomial approximation”. In : *J. Math. Anal. Appl.* (2015).
- DELEDALLE, C.-A., L. DENIS et F. TUPIN. “How to compare noisy patches ? Patch similarity beyond Gaussian noise”. In : *Int. J. Comput. Vision* (2012), p. 1-17.

- DELEDALLE, C.-A., L. DENIS et F. TUPIN. “Iterative weighted maximum likelihood denoising with probabilistic patch-based weights”. In : *IEEE Trans. Image Process.* 18.12 (2009), p. 2661-2672.
- “NL-InSAR : Nonlocal Interferogram Estimation”. In : *IEEE Trans. Geosci. Remote Sens.* 49.4 (2011).
- “Poisson NL means : Unsupervised non local means for Poisson noise”. In : *ICIP.* 2010, p. 801-804.
- DELEDALLE, C.-A., V. DUVAL et J. SALMON. “Anisotropic Non-Local Means with Spatially Adaptive Patch Shapes”. In : *SSVM.* 2011.
- “Non-Local Methods with Shape-Adaptive Patches (NLM-SAP)”. In : *J. Math. Imaging Vis.* 43.2 (2012), p. 103-120.
- DELEDALLE, C.-A., N. PAPADAKIS et J. SALMON. “Contrast re-enhancement of Total-Variation regularization jointly with the Douglas-Rachford iterations”. In : *SPARS.* 2015.
- “On debiasing restoration algorithms : applications to total-variation and nonlocal-means”. In : *SSVM.* 2015, p. 129-141.
- DELEDALLE, C.-A., S. PARAMESWARAN et T. Q. NGUYEN. “Image restoration with generalized Gaussian mixture model patch priors”. In : *arXiv preprint arXiv :1802.01458* (2018).
- DELEDALLE, C.-A., J. SALMON et A. S. DALALYAN. “Image denoising with patch based PCA : Local versus global”. In : *BMVC.* 2011.
- DELEDALLE, C.-A., F. TUPIN et L. DENIS. “Patch similarity under non-Gaussian noise”. In : *ICIP.* 2011.
- DELEDALLE, C.-A. et al. “CLEAR : Covariant LEAst-square Re-fitting with applications to image restoration”. In : *SIAM J. Imaging Sci.* 10.1 (2017), p. 243-284.
- DELON, J. et A. DESOLNEUX. *A patch-based approach for removing mixed Gaussian-impulse noise*. Rapp. tech. Telecom Paristech, 2012.
- DELORME, A. et al. “Independent EEG sources are dipolar”. In : *PloS one* 7.2 (2012), e30135.
- DELYON, B. “Régression”. 2015. URL : <https://perso.univ-rennes1.fr/bernard.delyon/regression.pdf>.
- DEMPE, S. et al. “Bilevel programming problems”. In : *Energy Systems. Springer, Berlin* (2015).
- DENG, J. et al. “Imagenet : A large-scale hierarchical image database”. In : *CVPR.* 2009, p. 248-255.
- DENG, Y. et al. “Image-to-Markup Generation with Coarse-to-Fine Attention”. In : *ICML.* T. 70. 2017, p. 980-989.
- DENOYELLE, Q., V. DUVAL et G. PEYRÉ. “Support recovery for sparse deconvolution of positive measures”. In : *J. Fourier Anal. Appl.* (2016).
- DEVORE, R. A. et G. G. LORENTZ. *Constructive approximation*. Berlin : Springer-Verlag, 1993, p. x+449.

- DEVROYE, L., L. GYÖRFI et G. LUGOSI. *A probabilistic theory of pattern recognition*. T. 31. Applications of Mathematics (New York). New York : Springer-Verlag, 1996.
- DEZEURE, R. et al. “High-Dimensional Inference : Confidence Intervals, p -Values and R-Software hdi”. In : *Statist. Sci.* 30.4 (2015), p. 533-558.
- DIMAKIS, A. G., A. D. SARWATE et M. J. WAINWRIGHT. “Geographic Gossip : Efficient Averaging for Sensor Networks”. In : *IEEE Trans. Signal Process.* 56.3 (2008), p. 1205-1216.
- DIMAKIS, A. G. et al. “Gossip algorithms for distributed signal processing”. In : *Proceedings of the IEEE* 98.11 (2010), p. 1847-1864.
- DING, Q. et E. D. KOLACZYK. “A Compressed PCA Subspace Method for Anomaly Detection in High-Dimensional Data”. In : *Arxiv preprint arXiv :1109.4408* (2011).
- DONOHU, D. L. “Breakdown properties of multivariate location estimators”. Thèse de doct. Harvard University, 1982.
- “Compressed sensing”. In : *IEEE Trans. Inf. Theory* 52.4 (2006), p. 1289-1306.
- DONOHU, D. L., A. et A. MONTANARI. “Message-passing algorithms for compressed sensing”. In : *Proceedings of the National Academy of Sciences* 106.45 (2009), p. 18914-18919.
- DONOHU, D. L., M. ELAD et V. TEMLYAKOV. “Stable recovery of sparse overcomplete representations in the presence of noise”. In : *IEEE Trans. Inf. Theory* 52.1 (2006), p. 6-18.
- DONOHU, D. L. et M. GASKO. “Breakdown properties of location estimates based on half-space depth and projected outlyingness”. In : *Ann. Statist.* 20.4 (1992), p. 1803-1827.
- DONOHU, D. L. et P. J. HUBER. “The notion of breakdown point”. In : *A Festschrift for Erich L. Lehmann*. Wadsworth Statist./Probab. Ser. Wadsworth, Belmont, CA, 1983, p. 157-184.
- DONOHU, D. L. et I. M. JOHNSTONE. “Adapting to unknown smoothness via wavelet shrinkage”. In : *J. Amer. Statist. Assoc.* 90.432 (1995), p. 1200-1224.
- “Ideal spatial adaptation by wavelet shrinkage”. In : *Biometrika* 81.3 (1994), p. 425-455.
- DONOHU, D. L., R. C. LIU et B. MACGIBBON. “Minimax risk over hyperrectangles, and implications”. In : *Ann. Statist.* 18.3 (1990), p. 1416-1437.
- DONOHU, D. L. et al. “Maximum entropy and the nearly black object”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 54.1 (1992). With discussion and a reply by the authors, p. 41-81.
- DONOHU, D. L. et al. “Wavelet shrinkage : asymptopia ?” In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 57.2 (1995), p. 301-369.
- DONSKER, M. D. et S. R. S. VARADHAN. “Asymptotic evaluation of certain Markov process expectations for large time. III”. In : *Comm. Pure Appl. Math.* 29.4 (1976), p. 389-461.
- DORÉ, V. et M. CHERIET. “Robust NL-Means Filter With Optimal Pixel-Wise Smoothing Parameter for Statistical Image Denoising”. In : *IEEE Trans. Signal Process.* 57 (2009), p. 1703-1716.

- DOUGLAS, J. et H. H. RACHFORD. “On the numerical solution of heat conduction problems in two and three space variables”. In : *Transactions of the American mathematical Society* 82.2 (1956), p. 421-439.
- DRAPER, N. R. et H. SMITH. *Applied regression analysis*. Third. Wiley Series in Probability and Statistics : Texts and References Section. New York : John Wiley & Sons Inc., 1998, p. xx+706.
- DRINEAS, P. et M. W. MAHONEY. “On the Nyström method for approximating a Gram matrix for improved kernel-based learning”. In : *J. Mach. Learn. Res.* 6 (2005), p. 2153-2175.
- DUBHASHI, D. et A. PANCONESI. *Concentration of measure for the analysis of randomized algorithms*. T. 16. Cambridge University Press Cambridge, 2009.
- DUCHI, J. et al. “Composite objective mirror descent”. In : *COLT*. 2010.
- DUCHI, J. et al. “Efficient projections onto the ℓ_1 -ball for learning in high dimensions”. In : *ICML*. ACM. 2008, p. 272-279.
- DUCHI, J. C., A. AGARWAL et M. J. WAINWRIGHT. “Dual averaging for distributed optimization : convergence analysis and network scaling”. In : *IEEE Trans. Automat. Control* 57.3 (2012), p. 592-606.
- DUDA, R. O., P. E. HART et D. G. STORK. *Pattern classification*. Second. Wiley-Interscience, New York, 2001.
- DUDÍK, M., Z. HARCHAOUI et J. MALICK. “Lifted coordinate descent for learning with trace-norm regularization”. In : *AISTATS*. 2012.
- DUFLO, M. *Algorithmes stochastiques*. T. 23. Mathématiques & Applications (Berlin) [Mathematics & Applications]. Berlin : Springer-Verlag, 1996, p. xiv+319.
- *Random iterative models*. T. 34. Applications of Mathematics (New York). Translated from the 1990 French original by Stephen S. Wilson and revised by the author. Berlin : Springer-Verlag, 1997, p. xviii+385.
- DUMITRESCU, B. A. *Positive trigonometric polynomials and signal processing applications*. Springer, 2007.
- DÜNNER, C. et al. “Primal-Dual Rates and Certificates”. In : *ICML*. T. 48. 2016, p. 783-792.
- DUVAL, V., J.-F. AUJOL et Y. GOUSSEAU. “A bias-variance approach for the nonlocal means”. In : *SIAM J. Imaging Sci.* 4.2 (2011), p. 760-788.
- DUVAL, V. et G. PEYRÉ. “Exact support recovery for sparse spikes deconvolution”. In : *Found. Comput. Math.* (2015), p. 1-41.
- “Sparse spikes deconvolution on thin grids”. In : *ArXiv e-prints* (2015).
- “The non degenerate source condition : support robustness for discrete and continuous sparse deconvolution”. In : *CAMSAP*. 2015, p. 49-52.
- DWORK, C. et A. ROTH. “The algorithmic foundations of differential privacy”. In : *Foundations and Trends in Theoretical Computer Science* 9.3-4 (2014), p. 211-407.
- DYCKERHOFF, R. et P. MOZHAROVSKIY. “Exact computation of the halfspace depth”. In : *Computational Statistics & Data Analysis* 98 (2016), p. 19-30.

- DYKSTRA, R. L. “An algorithm for restricted least squares regression”. In : *J. Amer. Statist. Assoc.* 78.384 (1983), p. 837-842.
- EBRAHIMI, M. “Inverse Problems and Self-similarity in Imaging”. Thèse de doct. University of Waterloo, 2008.
- EFROMOVICH, S. *Nonparametric curve estimation*. Springer Series in Statistics. Methods, theory, and applications. New York : Springer-Verlag, 1999.
- EFROMOVICH, S. Y. “On nonparametric regression for IID observations in a general setting”. In : *Ann. Statist.* 24.3 (1996), p. 1125-1144.
- EFROMOVICH, S. Y. et M. S. PINSKER. “A self-training algorithm for nonparametric filtering”. In : *Avtomat. i Telemekh.* 1.11 (1984), p. 58-65.
- “Sharp-optimal and adaptive estimation for heteroscedastic nonparametric regression”. In : *Statist. Sinica* 6.4 (1996), p. 925-942.
- EFRON, B. “Bootstrap methods : another look at the Jackknife”. In : *Ann. Statist.* 7.1 (1979), p. 1-26.
- EFRON, B., T. J. HASTIE et R. TIBSHIRANI. “Discussion : “The Dantzig selector : statistical estimation when p is much larger than n ” [Ann. Statist. **35** (2007), no. 6, 2313-2351; MR2382644] by E. Candès and T. Tao”. In : *Ann. Statist.* 35.6 (2007), p. 2358-2364.
- EFRON, B. et R. TIBSHIRANI. *An introduction to the bootstrap*. CRC press, 1994.
- EFRON, B. et al. “Least angle regression”. In : *Ann. Statist.* 32.2 (2004). With discussion, and a rejoinder by the authors, p. 407-499.
- EFROS, A. A. et T.K. LEUNG. “Texture Synthesis by Non-parametric Sampling”. In : *ICCV*. 1999, p. 1033-1038.
- EFROYMSON, M. A. “Multiple regression analysis”. In : *Mathematical methods for digital computers*. New York : Wiley, 1960, p. 191-203.
- EGIAZARIAN, K. O., V. KATKOVNIK et J. T. ASTOLA. “Local transform-based image denoising with adaptive window-size selection”. In : *SPIE*. T. 4170. Barcelona, Spain, 2001, p. 13-23.
- EKELAND, I. et R. TEMAM. *Convex analysis and variational problems*. Translated from the French, Studies in Mathematics and its Applications, Vol. 1. North-Holland Publishing Co., Amsterdam-Oxford ; American Elsevier Publishing Co., Inc., New York, 1976, p. ix+402.
- EL GHAOUI, L., V. VIALON et T. RABBANI. “Safe feature elimination in sparse supervised learning”. In : *J. Pacific Optim.* 8.4 (2012), p. 667-698.
- ELAD, M. “On the origin of the bilateral filter and ways to improve it”. In : *IEEE Trans. Image Process.* 11.10 (2002), p. 1141-1151.
- *Sparse and redundant representations. From theory to applications in signal and image processing*. New York, NY : Springer, 2010, p. xx + 376.
- ELAD, M. et M. AHARON. “Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries”. In : *IEEE Trans. Image Process.* 15.12 (2006), p. 3736-3745.
- ELAD, M., P. MILANFAR et R. RUBINSTEIN. “Analysis versus synthesis in signal priors”. In : *Inverse problems* 23.3 (2007), p. 947-968.

- ENGEMANN, D. A. et A. GRAMFORT. “Automated model selection in covariance estimation and spatial whitening of MEG and EEG signals”. In : *NeuroImage* 108 (2015), p. 328-342.
- ENGL, H. W., M. HANKE et A. NEUBAUER. *Regularization of inverse problems*. T. 375. Mathematics and its Applications. Dordrecht : Kluwer Academic publishers Group, 1996.
- ENGLE, R. “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”. In : *Econometrica* 50.4 (1982), p. 987-1007.
- FADILI, J. et al. “Stable Recovery with Analysis Decomposable Priors”. In : *SPARS*. 2013.
- “Stable Recovery with Analysis Decomposable Priors”. In : *SampTA*. 2013.
- FAN, R.-E. et al. “LIBLINEAR : A library for large linear classification”. In : *J. Mach. Learn. Res.* 9 (2008), p. 1871-1874.
- FAN, J. et I. GJIBELS. *Local polynomial modelling and its applications*. T. 66. Monographs on Statistics and Applied Probability. London : Chapman & Hall, 1996, p. xvi+341.
- FAN, J. et R. LI. “Variable selection via nonconcave penalized likelihood and its oracle properties”. In : *J. Amer. Statist. Assoc.* 96.456 (2001), p. 1348-1360.
- FAN, J. et J. LV. “Sure independence screening for ultrahigh dimensional feature space”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 70.5 (2008), p. 849-911.
- FAN, Z. et L. GUAN. “Approximate ℓ_0 -penalized estimation of piecewise-constant signals on graphs”. In : *Ann. Statist.* 46.6B (2018), p. 3217-3245.
- FAZEL, M. “Matrix rank minimization with applications”. Thèse de doct. Stanford University, 2002.
- FERCOQ, O. et P. BIANCHI. “A Coordinate Descent Primal-Dual Algorithm with Large Step Size and Possibly Non-Separable Functions”. In : *SIAM J. Optim.* 29 (1) (2019), p. 100-134.
- FERCOQ, O., A. GRAMFORT et J. SALMON. “Mind the duality gap : safer rules for the lasso”. In : *ICML*. T. 37. 2015, p. 333-342.
- FERCOQ, O. et P. RICHTÁRIK. “Accelerated, parallel and proximal coordinate descent”. In : *ArXiv e-prints* (2013).
- “Accelerated, parallel and proximal coordinate descent”. In : *SIAM J. Optim.* 25.3 (2015), p. 1997-2013.
- FERNANDEZ-GRANDA, C. “Super-Resolution of Point Sources via Convex Programming”. In : *Information and Inference* 5.3 (2016), p. 251-303.
- “Support detection in super-resolution”. In : *SampTA*. 2013.
- FÉVOTTE, C., N. BERTIN et J.-L. DURRIEU. “Nonnegative Matrix Factorization with the Itakura-Saito Divergence. With Application to Music Analysis”. In : *Neural Computation* 21.3 (2009).
- FIGUEIREDO, M. A. T. et J. M. BIOUCAS-DIAS. “Restoration of Poissonian images using alternating direction optimization”. In : *IEEE Trans. Signal Process.* 19.12 (2010), p. 3133-3145.

- FISZ, M. "The limiting distribution of a function of two independent random variables and its statistical application". In : *Colloquium Mathematicum* 3 (1955), p. 138-146.
- FOATA, D. et A. FUCHS. *Calcul des probabilités : cours et exercices corrigés*. Masson, 1996.
- FOI, A. "Anisotropic nonparametric image processing : theory, algorithms and applications". Thèse de doct. Politecnico di Milano, 2005.
- FOI, A., V. KATKOVNIK et K. O. EGIAZARIAN. "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images". In : *IEEE Trans. Image Process.* 16.5 (2007), p. 1395-1411.
- FOI, A. et al. "A novel anisotropic local polynomial estimator based on derirectional multiscale optimizations". In : *Proceedings of the Sixth IMA International Conference on Mathematics in Signal Processing*. Cirencester, United Kingdom, 2004, p. 79-82.
- FOUCART, S. et H. RAUHUT. *An Invitation to Compressive Sensing*. Springer New York, 2013.
- FOYGEL, R. et al. "Learning with the weighted trace-norm under arbitrary sampling distributions." In : *NIPS*. 2011, p. 2133-2141.
- FRANCESCHI, L. et al. "Bilevel Programming for Hyperparameter Optimization and Meta-Learning". In : *ICML*. 2018, p. 1563-1572.
- FREUND, Y. "Boosting a weak learning algorithm by majority". In : *COLT*. Rochester, New York, United States, 1990, p. 202-216.
- FREUND, Y. et R. E. SCHAPIRE. "A decision-theoretic generalization of on-line learning and an application to boosting". In : *Journal of computer and system sciences* 55.1 (1997), p. 119-139.
- FRIEDMAN, J. "Greedy function approximation : a gradient boosting machine". In : *Ann. Statist.* 29.5 (2001), p. 1189-1232.
- FRIEDMAN, J., T. J. HASTIE et Robert R. TIBSHIRANI. "Additive logistic regression : a statistical view of boosting". In : *Ann. Statist.* 28.2 (2000), p. 337-407.
- FRIEDMAN, J., T. J. HASTIE et R. TIBSHIRANI. "Regularization paths for generalized linear models via coordinate descent". In : *J. Stat. Softw.* 33.1 (2010), p. 1-22.
- "Sparse inverse covariance estimation with the graphical lasso". In : *Biostatistics* 9.3 (2008), p. 432-441.
- FRIEDMAN, J. et al. "Pathwise coordinate optimization". In : *Ann. Appl. Stat.* 1.2 (2007), p. 302-332.
- FRYŻLEWICZ, P. et G. P. NASON. "A Haar-Fisz algorithm for Poisson intensity estimation". In : *J. Comput. Graph. Statist.* 13.3 (2004), p. 621-638.
- *Poisson intensity estimation using wavelets and the Fisz transformation*. Rapp. tech. Department of Mathematics, University of Bristol, United Kingdom, 2001.
- FU, W. J. "Penalized Regressions : the Bridge versus the Lasso". In : *J. Comput. Graph. Statist.* 7.3 (1998), p. 397-416.
- FUCHS, J.-J. "On Sparse Representations in Arbitrary Redundant Bases". In : *IEEE Trans. Inf. Theory* 50.6 (2004), p. 1341-1344.

- FUCHS, J.-J. “Recovery of exact sparse representations in the presence of bounded noise”. In : *IEEE Trans. Inf. Theory* 51.10 (2005), p. 3601-3608.
- “Spread representations”. In : *ASILOMAR*. 2011.
- FUKUNAGA, K. et L. HOSTETLER. “The estimation of the gradient of a density function, with applications in pattern recognition”. In : *IEEE Trans. Inf. Theory* 21.1 (1975), p. 32-40.
- GAÏFFAS, S. et O. KLOPP. “High dimensional matrix estimation with unknown variance of the noise.” In : *Stat. Sin.* 27.1 (2017), p. 115-146.
- GAÏFFAS, S. et G. LECUÉ. “Hyper-Sparse Optimal Aggregation”. In : *J. Mach. Learn. Res.* 12 (2011), p. 1813-1833.
- GALLIER, J. et J. QUAINANCE. *Fundamentals of Linear Algebra and Optimization*. 2018. URL : <https://www.seas.upenn.edu/~cis515/linalg.pdf>.
- GALTCHOUK, L. et S. PERGAMENSHCHIKOV. “Sharp non-asymptotic oracle inequalities for non-parametric heteroscedastic regression models”. In : *J. Nonparametr. Stat.* 21.1 (2009), p. 1-18.
- GAO, H.-Y. et A. G. BRUCE. “WaveShrink with firm shrinkage”. In : *Statist. Sinica* (1997), p. 855-874.
- GARCIA, V. et F. NIELSEN. “Simplification and hierarchical representations of mixtures of exponential families”. In : *Signal Processing* 90.12 (2010), p. 3197-3212.
- GASQUET, C. et P. WITOMSKI. *Fourier analysis and applications*. T. 30. Texts in Applied Mathematics. New York : Springer-Verlag, 1999.
- GASSO, G., A. RAKOTOMAMONJY et S. CANU. “Recovering sparse signals with non-convex penalties and DC programming”. In : *IEEE Trans. Signal Process.* 57.12 (2009), p. 4686-4698.
- GAUSS, C. F. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. T. 7. Perthes et Besser, 1809.
- GAUTSCHI, W. “On inverses of Vandermonde and confluent Vandermonde matrices”. In : *Numerische Mathematik* 4 (1 1962), p. 117-123.
- GAZAGNADOU, N., R. GOWER et J. SALMON. “Optimal mini-batch and step sizes for SAGA”. In : *ICML*. 2019.
- GEORGE, E. I. “Combining minimax shrinkage estimators”. In : *J. Amer. Statist. Assoc.* 81.394 (1986), p. 437-445.
- “Minimax multiple shrinkage estimation”. In : *Ann. Statist.* 14.1 (1986), p. 188-205.
- GERCHINOVITZ, S. “Sparsity regret bounds for individual sequences in online linear regression”. In : *Journal of Machine Learning Research - Proceedings Track* 19 (2011), p. 377-396.
- GÉRON, A. *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. O’Reilly Media, 2017.
- *Machine Learning avec Scikit-Learn - Mise en oeuvre et cas concrets*. O’Reilly Media, 2017.

- GIESEN, J., M. JAGGI et S. LAUE. “Approximating parameterized convex optimization problems”. In : *European Symposium on Algorithms*. 2010, p. 524-535.
- GIESEN, J., S. LAUE et P. WIESCHOLLEK. “Robust and efficient kernel hyperparameter paths with guarantees”. In : 2014, p. 1296-1304.
- GIESEN, J. et al. “Approximating concavely parameterized optimization problems”. In : *NIPS*. 2012, p. 2105-2113.
- GILBOA, G. “A total variation spectral framework for scale and texture analysis”. In : *SIAM J. Imaging Sci.* 7.4 (2014), p. 1937-1961.
- GILBOA, G. et S. OSHER. “Nonlocal linear image regularization and supervised segmentation”. In : *Multiscale Model. Simul.* 6.2 (2007), p. 595-630.
- “Nonlocal operators with applications to image processing”. In : *Multiscale Model. Simul.* 7.3 (2008), p. 1005-1028.
- GILCHRIST, W. “Regression Revisited”. In : *International Statistical Review / Revue Internationale de Statistique* 76.3 (2008), p. 401-418.
- *Statistical modelling with quantile functions*. CRC Press, 2000.
- GIMENEZ, J. R. et J. ZOU. “Discovering Conditionally Salient Features with Statistical Guarantees”. In : *ICML*. 2019, p. 2290-2298.
- GIRAUD, C. *Introduction to high-dimensional statistics*. T. 138. CRC Press, 2014.
- “Mixing least-squares estimators when the variance is unknown”. In : *Bernoulli* 14.4 (2008), p. 1089-1107.
- GIRAUD, C., S. HUET et N. VERZELEN. “High-dimensional regression with unknown variance”. In : *Statist. Sci.* 27.4 (2012), p. 500-518.
- GOLDENSHLUGER, A. et O. V. LEPSKI. “Bandwidth selection in kernel density estimation : oracle inequalities and adaptive minimax optimality”. In : *Ann. Statist.* 39 (2011), p. 1608-1632.
- “Structural adaptation via L_p -norm oracle inequalities”. In : *Probab. Theory Related Fields* 143.1-2 (2009), p. 41-71.
- “Universal pointwise selection rule in multivariate function estimation”. In : *Bernoulli* 14.4 (2008), p. 1150-1190.
- GOLDENSHLUGER, A. et A. S. NEMIROVSKI. “On spatially adaptive estimation of nonparametric regression”. In : *Math. Methods Statist.* 6.2 (1997), p. 135-170.
- GOLDSTEIN, T. et S. OSHER. “The split Bregman method for L1-regularized problems”. In : *SIAM J. Imaging Sci.* 2.2 (2009), p. 323-343.
- GOLUB, G. H. et C. F. VAN LOAN. *Matrix computations*. Third. Baltimore, MD : Johns Hopkins University Press, 1996, p. xxx+698.
- *Matrix computations*. Fourth. Johns Hopkins University Press, Baltimore, MD, 2013, p. xiv+756.
- GOLUB, T. R. et al. “Molecular classification of cancer : class discovery and class prediction by gene expression monitoring.” In : *Science* 286.5439 (1999), p. 531-537.
- GOLUBEV, G. K. “Nonparametric estimation of smooth densities of a distribution in L_2 ”. In : *Problemy Peredachi Informatsii* 28.1 (1992), p. 52-62.

- GOLUBEV, Y. “On universal oracle inequalities related to high-dimensional linear models”. In : *Ann. Statist.* 38.5 (2010), p. 2751-2780.
- GONZALEZ, R. C., R. E. WOODS et S. L. EDDINS. *Digital image processing using MATLAB*. Pearson Education India, 2004.
- GOODFELLOW, I., Y. BENGIO et A. COURVILLE. *Deep learning*. MIT press, 2016.
- GOOSSENS, B. et al. “An improved non-local denoising algorithm”. In : *LNLA*. Lausanne, 2008, p. 143-156.
- GORDON, G. J. “Generalized² Linear² Models”. In : *NIPS*. 2003, p. 593-600.
- GORODNITSKY, I. F. et B. D. RAO. “Sparse signal reconstruction from limited data using FOCUSS : A re-weighted minimum norm algorithm”. In : *IEEE Trans. Signal Process.* 45.3 (1997), p. 600-616.
- GOURDON, X. *Les maths en tête : Algèbre*. Ellipses Marketing. Ellipses Marketing, 2009.
- *Les maths en tête : Analyse*. Ellipses Marketing. Ellipses Marketing, 2008.
- GRAMFORT, A., M. KOWALSKI et M. HÄMÄLÄINEN. “Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods”. In : *Phys. Med. Biol.* 57.7 (2012), p. 1937-1961.
- GRAMFORT, A. et al. “MNE software for processing MEG and EEG data”. In : *NeuroImage* 86 (2014), p. 446-460.
- GRAMFORT, A. et al. “Time-frequency mixed-norm estimates : Sparse M/EEG imaging with non-stationary source activations”. In : *NeuroImage* 70 (2013), p. 410-422.
- GRANT, M. et S. BOYD. *CVX : Matlab Software for Disciplined Convex Programming, version 2.1*. 2014.
- “Graph implementations for nonsmooth convex programs”. In : *Recent Advances in Learning and Control*. Sous la dir. de V. BLONDEL, S. BOYD et H. KIMURA. Lecture Notes in Control and Information Sciences. Springer-Verlag Limited, 2008, p. 95-110.
- GRASMAIR, M. “Generalized Bregman distances and convergence rates for non-convex regularization methods”. In : *Inverse Problems* 26.11 (2010), p. 115014.
- “Linear convergence rates for Tikhonov regularization with positively homogeneous functionals”. In : *Inverse Problems* 27 (2011), p. 075014.
- GRASMAIR, M., M. HALTMEIER et O. SCHERZER. “The residual method for regularizing ill-posed problems”. In : *Applied Mathematics and Computation* 218.6 (2011), p. 2693-2710.
- GRASMAIR, M., O. SCHERZER et M. HALTMEIER. “Necessary and sufficient conditions for linear convergence of ℓ_1 -regularization”. In : *Communications on Pure and Applied Mathematics* 64.2 (2011), p. 161-182.
- GREEN, P. J. et B. W. SILVERMAN. *Nonparametric regression and generalized linear models*. T. 58. Monographs on Statistics and Applied Probability. A roughness penalty approach. London : Chapman & Hall, 1994, p. xii+182.
- GREGOR, K. et Y. LECUN. “Learning fast approximations of sparse coding”. In : *ICML*. 2010, p. 399-406.

- GREGORY, R. T. et D. L. KARNEY. *A collection of matrices for testing computational algorithms*. Wiley-Interscience A Division of John Wiley & Sons, Inc., New York-London-Sydney, 1969.
- GRENANDER, U. "Some direct estimates of the mode". In : *Ann. Math. Statist.* 36 (1965), p. 131-138.
- GRIMMETT, G. R. et D. R. STIRZAKER. *Probability and random processes*. Third. New York : Oxford University Press, 2001, p. xii+596.
- GRISHCHENKO, D. et al. "Asynchronous Distributed Learning with Sparse Communications and Identification". In : *arXiv preprint arXiv :1812.03871* (2018).
- GROSS, D. "Recovering low-rank matrices from few coefficients in any basis". In : *IEEE Trans. Inf. Theory* 57.3 (2011), p. 1548-1566.
- GRÜNBAUM, B. *Convex Polytopes*. Second. New York : Springer-Verlag, 2003, p. 471.
- GU, B. et C. X. LING. "A New Generalized Error Path Algorithm for Model Selection". In : *ICML*. 2015, p. 2549-2558.
- GUEDJ, B. et P. ALQUIER. "PAC-Bayesian Estimation and Prediction in Sparse Additive Models". In : *Electron. J. Stat.* 7 (2013), p. 264-291.
- GUILLAUMIN, M., J. VERBEEK et C. SCHMID. "Is that you? Metric learning approaches for face identification". In : *ICCV*. 2009, p. 498-505.
- GULERYUZ, O. G. "Weighted Averaging for Denoising With Overcomplete Dictionaries". In : *IEEE Trans. Image Process.* 16.12 (2007), p. 3020-3034.
- GUNASEKAR, S., P. RAVIKUMAR et J. GHOSH. "Exponential Family Matrix Completion under Structural Constraints". In : *ICML*. 2014.
- GÜNTHER, S., N. N. SCHRAUDOLPH et S. V. N. VISHWANATHAN. "Fast Iterative Kernel Principal Component Analysis". In : *J. Mach. Learn. Res.* 8 (2007), p. 1893-1918.
- GÜNTHER, D., J. SALMON et J. TIERNY. "Mandatory critical points of 2D uncertain scalar fields". In : *Computer Graphics Forum*. T. 33. 3. 2014, p. 31-40.
- GUTTAG, J. V. *Introduction to Computation and Programming Using Python : With Application to Understanding Data*. MIT Press, 2016.
- GYÖRFI, L. et G. LUGOSI. "Strategies for sequential prediction of stationary time series". In : *Modeling uncertainty*. T. 46. Internat. Ser. Oper. Res. Management Sci. Boston, MA : Kluwer Acad. Publ., 2002, p. 225-248.
- GYÖRFI, L. et al. *A distribution-free theory of nonparametric regression*. Springer Series in Statistics. New York : Springer-Verlag, 2002, p. xvi+647.
- HALE, E., W. YIN et Y. ZHANG. "Fixed-Point Continuation for ℓ_1 -Minimization : Methodology and Convergence". In : *SIAM J. Optim.* 19.3 (2008), p. 1107-1130.
- HALKO, N., P. MARTINSSON et J. A. TROPP. "Finding structure with randomness : Probabilistic algorithms for constructing approximate matrix decompositions". In : *SIAM Review* 53 (2011), p. 217.
- HALTMEIER, M. "Stable Signal Reconstruction via ℓ_1 Minimization in Redundant, Non-Tight Frames". In : *IEEE Trans. Signal Process.* 61.2 (2013), p. 420-426.

- HAMERLY, G. et C. ELKAN. “Alternatives to the k-means algorithm that find better clusterings”. In : *ICIKM*. ACM. 2002, p. 600-607.
- HAMILTON, J. D. *Time series analysis*. Princeton, NJ : Princeton University Press, 1994.
- HAMPEL, F. R. “Contributions to the theory of robust estimation”. Thèse de doct. University of California, Berkeley, 1968.
- HAMPEL, F. R. et al. *Robust statistics : The Approach Based on Influence Functions*. Wiley series in probability and statistics. Wiley, 1986.
- HANLEY, J. A. et B. J. MCNEIL. “The meaning and use of the area under a receiver operating characteristic (ROC) curve”. In : *Radiology* 143.1 (1982), p. 29-36.
- HAO, N. et H. H. ZHANG. “Interaction Screening for Ultrahigh-Dimensional Data”. In : *J. Amer. Statist. Assoc.* 109.507 (2014), p. 1285-1301.
- HARCHAOU, Z., A. JUDITSKY et A. NEMIROVSKI. “Conditional gradient algorithms for machine learning”. In : *NIPS workshop*. 2012.
- “Conditional gradient algorithms for norm-regularized smooth convex optimization”. In : *Math. Program.* (2014), p. 1-38.
- HARCHAOU, Z. et C. LÉVY-LEDUC. “Multiple change-point estimation with a total variation penalty”. In : *J. Amer. Statist. Assoc.* 105.492 (2010), p. 1480-1493.
- HÄRDLE, W. *Applied nonparametric regression*. T. 19. Econometric Society Monographs. Cambridge : Cambridge University Press, 1990, p. xvi+333.
- HÄRDLE, W. et al. *Wavelets, approximation, and statistical applications*. T. 129. Lecture Notes in Statistics. New York : Springer-Verlag, 1998, p. xviii+265.
- HARI, R. et A. PUCE. *MEG-EEG Primer*. Oxford University Press, 2017.
- HARMANY, Z., R. MARCIA et R. WILLETT. “This is SPIRAL-TAP : Sparse Poisson Intensity Reconstruction ALgorithms – Theory and Practice”. In : *IEEE Trans. Image Process.* 21.3 (2012), p. 1084-1096.
- HARRISON, X. A. et al. “A brief introduction to mixed effects modelling and multi-model inference in ecology”. In : *PeerJ* 6 (2018), e4794.
- HASTIE, T. J., R. TIBSHIRANI et J. FRIEDMAN. *The Elements of Statistical Learning*. Second. Springer Series in Statistics. New York : Springer, 2009, p. xxii+745.
- HASTIE, T. J., R. TIBSHIRANI et M. WAINWRIGHT. *Statistical Learning with Sparsity : The Lasso and Generalizations*. CRC Press, 2015.
- HASTIE, T. J. et R. J. TIBSHIRANI. *Generalized additive models*. T. 43. CRC press, 1990.
- HAUFE, S. et al. “Combining sparsity and rotational invariance in EEG/MEG source reconstruction”. In : *NeuroImage* 42.2 (2008), p. 726-738.
- HAURY, A.-C. et al. “TIGRESS : Trustful Inference of Gene REgulation using Stability Selection.” In : *BMC systems biology* 6.1 (2012), p. 145.
- HAUSSLER, D., J. KIVINEN et M. K. WARMUTH. “Sequential Prediction of Individual Sequences Under General Loss Functions”. In : *IEEE Trans. Inf. Theory* 44.5 (1998), p. 1906-1925.

- HAUSSLER, D., J. KIVINEN et M. K. WARMUTH. “Tight worst-case loss bounds for predicting with expert advice”. In : *Computational learning theory (EuroCOLT)*. T. 904. Lecture Notes in Comput. Sci. Berlin : Springer, 1995, p. 69-83.
- HEBIRI, M. “Quelques questions de sélection de variables autour de l’estimateur LASSO”. Thèse de doct. Université Paris Diderot, 2009.
- HEBIRI, M. et J. C. LEDERER. “How Correlations Influence Lasso Prediction”. In : *IEEE Trans. Inf. Theory* (2013).
- HEBIRI, M. et S. VAN DE GEER. “The Smooth-Lasso and other $\ell_1 + \ell_2$ -regularized methods”. In : *Electron. J. Stat.* 5 (2011), p. 1184-1226.
- HERBRICH, R., T. GRAEPEL et K. OBERMAYER. “Large margin rank boundaries for ordinal regression”. In : *Advances in Large Margin Classifiers*. Sous la dir. de SMOLA et al. MIT Press, Cambridge, MA, 2000.
- HINTZE, J. L. et R. D. NELSON. “Violin plots : a box plot-density trace synergism”. In : *The American Statistician* 52.2 (1998), p. 181-184.
- HIRAKAWA, K. et T. W. PARKS. “Image denoising using total least squares”. In : *IEEE Trans. Image Process.* 15.9 (2006), p. 2730-2742.
- HIRIART-URRUTY, J.-B. et C. LEMARÉCHAL. *Convex analysis and minimization algorithms. I*. T. 305. Berlin : Springer-Verlag, 1993.
- *Convex analysis and minimization algorithms. II*. T. 306. Berlin : Springer-Verlag, 1993.
- HOCHBERG, Y. et A. C. TAMHANE. *Multiple comparison procedures*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., 1987.
- HOCHREITER, S. et J. SCHMIDHUBER. “Long short-term memory”. In : *Neural computation* 9.8 (1997), p. 1735-1780.
- HOCKING, R. R. “The analysis and selection of variables in linear regression”. In : *Biometrics* 32.1 (1976), p. 1-49.
- HOCKING, T. D. et al. “Clusterpath an algorithm for clustering using convex fusion penalties”. In : *ICML*. 2011, p. 745-752.
- HOEFFDING, W. “Probability inequalities for sums of bounded random variables”. In : *J. Amer. Statist. Assoc.* 58.301 (1963), p. 13-30.
- HOEFLING, H. “A path algorithm for the fused lasso signal approximator”. In : *J. Comput. Graph. Stat.* 19.4 (2010), p. 984-1006.
- HOERL, A. E. et R. W. KENNARD. “Ridge regression : Biased estimation for nonorthogonal problems”. In : *Technometrics* 12.1 (1970), p. 55-67.
- HOLMES, S. et W. HUBER. *Modern statistics for modern biology*. Cambridge University Press, 2018.
- HONG, D. et J.-C. BIRGET. “Deviation bounds for wavelet shrinkage”. In : *IEEE Trans. Inf. Theory* 49.7 (2003), p. 1851-1858.
- HOPFIELD, J. J. “Neural networks and physical systems with emergent collective computational abilities”. In : *Proceedings of the National Academy of Sciences* 79.8 (1982), p. 2554-2558.

- HORN, R. A. et C. R. JOHNSON. *Topics in matrix analysis*. Corrected reprint of the 1991 original. Cambridge : Cambridge University Press, 1994, p. viii+607.
- HORST, R. et N. V. THOAI. “DC Programming : Overview”. In : *Journal of Optimization Theory and Applications* 103.1 (1999), p. 1-43.
- HOYER, P. O. “Independent component analysis in image denoising”. Mém. de mast. Espoo : Helsinki University of Technology, 1999.
- HUBER, P. J. “Robust estimation of a location parameter”. In : *Ann. Math. Statist.* 35 (1964), p. 73-101.
- *Robust Statistics*. John Wiley & Sons Inc., 1981.
- HUBER, P. J. et R. DUTTER. “Numerical solution of robust regression problems”. In : *Compstat 1974 (Proc. Sympos. Computational Statist., Univ. Vienna, Vienna, 1974)*. Physica Verlag, Vienna, 1974, p. 165-172.
- HUBER, P. J. et E. M. RONCHETTI. *Robust statistics*. Second. Wiley series in probability and statistics. Wiley, 2009.
- HUDSON, H. M. “A natural identity for exponential families with applications in multiparameter estimation”. In : *Ann. Statist.* 6.3 (1978), p. 473-484.
- HUHLE, B. et al. “Robust Non-Local Denoising of Colored Depth Data”. In : *IEEE CVPR Workshop on Time of Flight Camera based Computer Vision (ToF-CV)*. Anchorage, Alaska, 2008.
- HUI, J. et al. “Robust video denoising using low rank matrix completion”. In : *CVPR* 0 (2010), p. 1791-1798.
- HÜTTER, J.-C. et P. RIGOLLET. “Optimal rates for total variation denoising”. In : *Conference on Learning Theory*. 2016, p. 1115-1146.
- HYVÄRINEN, A., P. O. HOYER et E. OJA. “Sparse Code Shrinkage : Denoising by Nonlinear Maximum Likelihood Estimation”. In : *NIPS*. 1998, p. 473-479.
- INDYK, P. et R. MOTWANI. “Approximate nearest neighbors : towards removing the curse of dimensionality”. In : *Proceedings of the thirtieth annual ACM symposium on Theory of computing*. ACM. 1998, p. 604-613.
- IOFFE, S. et C. SZEGEDY. “Batch normalization : Accelerating deep network training by reducing internal covariate shift”. In : *arXiv preprint arXiv :1502.03167* (2015).
- İPEK, E. et al. *Efficiently exploring architectural design spaces via predictive modeling*. T. 41. ACM, 2006.
- IUTZELER, F. et al. “Asynchronous distributed optimization using a randomized alternating direction method of multipliers”. In : *CDC*. 2013, p. 3671-3676.
- JACOB, L., G. OBOZINSKI et J.-P. VERT. “Group lasso with overlap and graph lasso”. In : *ICML*. 2009, p. 433-440.
- JAGGI, M. “Revisiting Frank-Wolfe : Projection-free sparse convex optimization”. In : *ICML*. 2013, p. 427-435.
- JAIN, A. K. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ, USA : Prentice Hall, 1989.

- JAIN, H., Y. PRABHU et M. VARMA. “Extreme multi-label loss functions for recommendation, tagging, ranking & other missing label applications”. In : *KDD*. 2016, p. 935-944.
- JAMES, A. T. “The distribution of the latent roots of the covariance matrix”. In : *The Annals of Mathematical Statistics* 31.1 (1960), p. 151-158.
- JAMES, G. et al. *An introduction to statistical learning*. T. 103. Springer Texts in Statistics. With applications in R. Springer, New York, 2013, p. xiv+426.
- JAMES, W. et C. M. STEIN. “Estimation with quadratic loss”. In : *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.* T. 1. Berkeley, Calif. : Univ. California Press, 1961, p. 361-379.
- JAMIESON, K. et A. TALWALKAR. “Non-stochastic best arm identification and hyperparameter optimization”. In : *AISTATS*. 2016, p. 240-248.
- JAVANMARD, A. et A. MONTANARI. “Confidence Intervals and Hypothesis Testing for High-Dimensional Regression”. In : *J. Mach. Learn. Res.* 15 (2014), p. 2869-2909.
- *Model selection for high-dimensional regression under the generalized irrepresentability condition*. Rapp. tech. Stanford, 2013.
- JÉGOU, H., T. FURON et J.-J. FUCHS. “Anti-sparse coding for approximate nearest neighbor search”. In : *ICASSP*. 2012, p. 2029-2032.
- JENATTON, R. et al. “Multi-scale mining of fMRI data with hierarchical structured sparsity”. In : *Pattern Recognition in NeuroImaging (PRNI), 2011 International Workshop on*. IEEE. 2011, p. 69-72.
- JENATTON, R. et al. “Proximal Methods for Hierarchical Sparse Coding”. In : *J. Mach. Learn. Res.* 12 (2011), p. 2297-2334.
- JI, L. et al. “Tensor Completion for Estimating Missing Values in Visual Data”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 35.1 (2013), p. 208-220.
- JI, Z. et al. “A moment-based nonlocal-means algorithm for image denoising”. In : *Information Processing Letters* 109.23-24 (2009), p. 1238-1244.
- JOACHIMS, T. *Making large-Scale SVM Learning Practical*. Rapp. tech. 24. Universität Dortmund, LS VIII-Report, 1998.
- “Optimizing search engines using clickthrough data”. In : *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*. Edmonton, Alberta, Canada : ACM, 2002, p. 133-142.
- “Text categorization with Support Vector Machines : Learning with many relevant features”. In : *Machine Learning : ECML-98*. Sous la dir. de C. NÉDELLEC et C. ROUVEIROL. T. 1398. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1998, p. 137-142.
- JOACHIMS, T., T. FINLEY et C.-N. J. YU. “Cutting-plane training of structural SVMs”. In : *Machine Learning* 77.1 (2009), p. 27-59.
- JOHANSSON, B., M. RABI et M. JOHANSSON. “A randomized incremental subgradient method for distributed optimization in networked systems”. In : *SIAM J. Optim.* 20.3 (2010), p. 1157-1170.

- JOHNSON, D. S. et F. P. PREPARATA. “The densest hemisphere problem”. In : *Theoret. Comput. Sci.* 6.1 (1978), p. 93-107.
- JOHNSON, T. B. et C. GUESTIN. “A Fast, Principled Working Set Algorithm for Exploiting Piecewise Linear Structure in Convex Problems”. In : *arXiv preprint arXiv :1807.08046* (2018).
- “Blitz : A Principled Meta-Algorithm for Scaling Sparse Optimization”. In : *ICML*. T. 37. 2015, p. 1171-1179.
- “StingyCD : Safely Avoiding Wasteful Updates in Coordinate Descent”. In : *ICML*. 2017, p. 1752-1760.
- “Unified methods for exploiting piecewise linear structure in convex optimization”. In : *NIPS*. 2016, p. 4754-4762.
- JOHNSON, W. B. et J. LINDENSTRAUSS. “Extensions of Lipschitz mappings into a Hilbert space”. In : *Contemporary mathematics* 26.189-206 (1984), p. 1.
- JOHNSTONE, I. M. “Oracle inequalities and nonparametric function estimation”. In : *Proceedings of the International Congress of Mathematicians, Vol. III (Berlin, 1998)*. Extra Vol. III. 1998, 267-278 (electronic).
- “Wavelet shrinkage for correlated data and inverse problems : adaptivity results”. In : *Statist. Sinica* 9.1 (1999), p. 51-83.
- JOLLIFFE, I. T. *Principal component analysis*. Second. Springer Series in Statistics. New York : Springer-Verlag, 2002, p. xxx+487.
- JOLY, A. et al. “Crowdsourcing biodiversity monitoring : how sharing your photo stream can sustain our planet”. In : *Proceedings of the 24th ACM international conference on Multimedia*. ACM. 2016, p. 958-967.
- JOLY, A. et al. “Interactive plant identification based on social image data”. In : *Ecological Informatics* 23 (2014), p. 22-34.
- JOLY, E. “Robust estimation of heavy-tailed distributions”. Thèse de doct. Université Paris-Sud, 2015.
- JUDGE, G. G. et T. TAKAYAMA. “Inequality restrictions in regression analysis”. In : *J. Amer. Statist. Assoc.* 61 (1966), p. 166-181.
- JUDITSKY, A. B. et A. S. NEMIROVSKI. “Functional aggregation for nonparametric regression”. In : *Ann. Statist.* 28.3 (2000), p. 681-712.
- “Nonparametric denoising of signals with unknown local structure. I. Oracle inequalities”. In : *Appl. Comput. Harmon. Anal.* 27.2 (2009), p. 157-179.
- “On verifiable sufficient conditions for sparse signal recovery via ℓ_1 minimization”. In : *Math. Program.* 127.1 (2011), p. 57-88.
- JUDITSKY, A. B., P. RIGOLLET et A. B. TSYBAKOV. “Learning by mirror averaging”. In : *Ann. Statist.* 36.5 (2008), p. 2183-2206.
- JUDITSKY, A. B. et al. “Recursive aggregation of estimators by the mirror descent method with averaging”. In : *Problemy Peredachi Informatsii* 41.4 (2005), p. 78-96.
- KAILATH, T. “Lectures on Wiener and Kalman filtering”. In : *Lectures on Wiener and Kalman Filtering*. Springer, 1981, p. 1-143.

- KALAI, A. et al. “Agnostically Learning Halfspaces”. In : *SIAM J. Comput.* 37.6 (2008), p. 1777-1805.
- KARIMIREDDY, P. et al. “Efficient Greedy Coordinate Descent for Composite Problems”. In : *arXiv preprint arXiv :1810.06999* (2018).
- KARNIN, Z., T. KOREN et O. SOMEKH. “Almost Optimal Exploration in Multi-Armed Bandits”. In : *ICML*. 2013, p. 1238-1246.
- KARP, R. et al. “Randomized rumor spreading”. In : *Symposium on Foundations of Computer Science*. IEEE. 2000, p. 565-574.
- KARPATY, A. et L. FEI-FEI. “Deep visual-semantic alignments for generating image descriptions”. In : *CVPR*. 2015, p. 3128-3137.
- KATKOVNIK, V. “A New Method for Varying Adaptive Bandwidth Selection”. In : *IEEE Trans. Image Process.* 47.9 (1999), p. 2567-2571.
- KATKOVNIK, V., K. O. EGIAZARIAN et J. T. ASTOLA. “Adaptive Window Size Image Denoising Based on Intersection of Confidence Intervals (ICI) Rule”. In : *J. Math. Imaging Vis.* 16.3 (2002), p. 223-235.
- “Application of the ICI principle to window size adaptive median filtering”. In : *Signal Processing* 83.2 (2003), p. 251-257.
- “Median filter with varying bandwidth adaptive to unknown smoothness of the signal”. In : *IEEE International Symposium on Circuits and Systems*. 2000, p. 519-522.
- KATKOVNIK, V. et V. G. SPOKOINY. “Spatially Adaptive Estimation via Fitted Local Likelihood Techniques”. In : *IEEE Trans. Image Process.* 56.3 (2008), p. 873-886.
- KATKOVNIK, V. et al. “Directional varying scale approximations for anisotropic signal processing”. In : *EUSIPCO*. 2004, p. 101-104.
- “From Local Kernel to Nonlocal Multiple-Model Image Denoising”. In : *Int. J. Comput. Vision* 86.1 (2010), p. 1-32.
- KAUFMANN, E., N. KORDA et R. MUNOS. “Thompson sampling : An asymptotically optimal finite-time analysis”. In : *Algorithmic Learning Theory*. Springer. 2012, p. 199-213.
- KEARNS, M. J., R. E. SCHAPIRE et L. SELLIE. “Toward Efficient Agnostic Learning”. In : *Mach. Learn.* 17.2-3 (1994), p. 115-141.
- KEMPE, D., A. DOBRA et J. GEHRKE. “Gossip-Based Computation of Aggregate Information”. In : *Symposium on Foundations of Computer Science*. IEEE. 2003, p. 482-491.
- KERVANN, C. et J. BOULANGER. “Optimal Spatial Adaptation for Patch-Based Image Denoising”. In : *IEEE Trans. Image Process.* 15.10 (2006), p. 2866-2878.
- KERVANN, C., J. BOULANGER et P. COUPÉ. “Bayesian Non-local Means Filter, Image Redundancy and Adaptive Dictionaries for Noise Removal”. In : *SSVM*. T. 4485. 2007, p. 520-532.
- KESHAVAN, R. H., A. MONTANARI et S. OH. “Matrix completion from noisy entries”. In : *J. Mach. Learn. Res.* 11 (2010), p. 2057-2078.
- KHAN, J. A., S. VAN AELST et R. H. ZAMAR. “Robust linear model selection based on least angle regression”. In : *J. Amer. Statist. Assoc.* 102.480 (2007), p. 1289-1299.

- KHOTANZAD, A. et Y. H. HONG. “Invariant Image Recognition by Zernike Moments”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 12.5 (1990), p. 489-497.
- KIM, S.-J. et al. “ l_1 trend filtering”. In : *SIAM Rev.* 51.2 (2009), p. 339-360.
- KIM, S.-J. et al. “An interior-point method for large-scale l_1 -regularized least squares”. In : *IEEE J. Sel. Topics Signal Process.* 1.4 (2007), p. 606-617.
- KIM, J. et H. PARK. “Fast Active-set-type Algorithms for L1-regularized Linear Regression.” In : *AISTATS*. 2010, p. 397-404.
- KIM, K. I., M. O. FRANZ et B. SCHÖLKOPF. “Iterative Kernel Principal Component Analysis for Image Modeling”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 27.9 (2005), p. 1351-1366.
- KINDERMANN, S., S. OSHER et P.W. JONES. “Deblurring and denoising of images by non-local functionals”. In : *Multiscale Model. Simul.* 4.4 (2005), 1091-1115 (electronic).
- KIVINEN, J. et M. K. WARMUTH. “Averaging expert predictions”. In : *Computational learning theory (EuroCOLT)*. T. 1572. Lecture Notes in Comput. Sci. Berlin : Springer, 1999, p. 153-167.
- “Exponentiated gradient versus gradient descent for linear predictors”. In : *Inform. and Comput.* 132.1 (1997), p. 1-63.
- KLEMELÄ, J. et E. MAMMEN. “Empirical risk minimization in inverse problems”. In : *Ann. Statist.* 38.1 (2010), p. 482-511.
- KLOPP, O. “Noisy low-rank matrix completion with general sampling distribution”. In : *Bernoulli* 2.1 (2014), p. 282-303.
- “Rank penalized estimators for high-dimensional matrices”. In : *Electron. J. Stat.* 5 (2011), p. 1161-1183.
- KLOPP, O. et al. “Adaptive Multinomial Matrix Completion”. In : *Electron. J. Statist.* 9.2 (2015), p. 2950-2975.
- KLUTCHNIKOFF, N. “On the adaptive estimation of anisotropic functions”. Thèse de doct. University Aix-Marseille, 2005.
- KNAUS, C. et M. ZWICKER. “Dual-domain image denoising”. In : *ICIP*. 2013, p. 440-444.
- KNEIP, Alois. “Ordered linear smoothers”. In : *Ann. Statist.* 22.2 (1994), p. 835-866.
- KOCH, I. “On the asymptotic performance of median smoothers in image analysis and nonparametric regression”. In : *Ann. Statist.* 24.4 (1996), p. 1648-1666.
- KOENKER, R. et G. BASSETT. “Regression quantiles”. In : *Econometrica* 46.1 (1978), p. 33-50.
- KOENKER, R. et K. F. HALLOCK. “Quantile regression”. In : *J. Econ. Perspect.* 15.4 (2001), p. 143-156.
- KOH, K., S.-J. KIM et S. BOYD. “An interior-point method for large-scale l_1 -regularized logistic regression.” In : *J. Mach. Learn. Res.* 8.8 (2007), p. 1519-1555.
- KOLACZYK, E. D. “Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds”. In : *Statist. Sinica* 9.1 (1999), p. 119-135.
- KOLACZYK, E. D. et R. D. NOWAK. “Multiscale likelihood analysis and complexity penalized estimation”. In : *Ann. Statist.* 32.2 (2004), p. 500-527.

- KOLAR, M. et J. SHARPNACK. “Variance Function Estimation in High-dimensions”. In : *ICML*. 2012, p. 1447-1454.
- KOLTCHINSKII, V. *Oracle inequalities in empirical risk minimization and sparse recovery problems*. T. 2033. Lecture Notes in Mathematics. Heidelberg : Springer, 2011.
- “Rademacher penalties and structural risk minimization”. In : *IEEE Trans. Inf. Theory* 47.5 (2001), p. 1902-1914.
- “Sparse recovery in convex hulls via entropy penalization”. In : *Ann. Statist.* 37.3 (2009), p. 1332-1359.
- KOLTCHINSKII, V., K. LOUNICI et A. B. TSYBAKOV. “Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion”. In : *Ann. Statist.* 39.5 (2011), p. 2302-2329.
- KOLTCHINSKII, V. et S. MINSKER. “ L_1 -penalization in functional linear regression with subgaussian design.” In : *J. Éc. Polytech., Math.* 1 (2014), p. 269-330.
- KONEČNÝ, J. et al. “Federated Optimization : Distributed Machine Learning for On-Device Intelligence”. In : *CoRR* abs/1610.02527 (2016).
- KOREN, Y., R. BELL et C. VOLINSKY. “Matrix factorization techniques for recommender systems”. In : *Computer* 42.8 (2009), p. 30-37.
- KOROSTEL'EV, A. P. et A. B. TSYBAKOV. *Minimax theory of image reconstruction*. T. 82. Lecture Notes in Statistics. New York : Springer-Verlag, 1993, p. xii+258.
- KOUACHI, S. “Eigenvalues and eigenvectors of tridiagonal matrices”. In : *Electron. J. Linear Algebra* 15 (2006), p. 115-133.
- KOWALCZYK, W. et N. A. VLASSIS. “Newscast EM”. In : *NIPS*. 2004, p. 713-720.
- KOWALSKI, M. “Sparse regression using mixed norms”. In : *Appl. Comput. Harmon. Anal.* 27.3 (2009), p. 303-324.
- KOWALSKI, M. et al. “Accelerating ISTA with an active set strategy”. In : *OPT 2011 : 4th International Workshop on Optimization for Machine Learning*. 2011, p. 7.
- KRISHNAMURTHY, K., M. RAGINSKY et R. WILLETT. “Multiscale photon-limited spectral image reconstruction”. In : *SIAM J. Imaging Sci.* 3.3 (2010), p. 619-645.
- KULIS, B. “Metric Learning : A Survey”. In : *Foundations and Trends in Machine Learning* 5.4 (2012), p. 287-364.
- KUMAR, A. et al. “A binary classification framework for two-stage multiple kernel learning”. In : *ICML*. 2012, p. 1295-1302.
- KUSHNER, H. J. et G. YIN. *Stochastic approximation and recursive algorithms and applications*. T. 35. Springer Verlag, 2003.
- KYUNG, M. et al. “Penalized regression, standard errors, and Bayesian lassos”. In : *Bayesian Analysis* 5.2 (2010), p. 369-411.
- LAFFERTY, J. et L. WASSERMAN. “Rodeo : sparse, greedy nonparametric regression”. In : *Ann. Statist.* 36.1 (2008), p. 28-63.
- LAFOND, J. “Complétion de matrice de faible rang : aspects statistiques et computationnels”. Thèse de doct. Télécom ParisTech, 2016.

- LAFOND, J. “Low rank matrix completion with exponential family noise”. In : *COLT*. 2015, p. 1224-1243.
- LAFOND, J. et al. “Probabilistic Low-rank Matrix Completion on Finite Alphabet”. In : *NIPS*. 2014, p. 1727-1735.
- LAKSHMINARAYANAN, B., D. M. ROY et Y. W. TEH. “Mondrian Forests : Efficient Online Random Forests”. In : *NIPS*. 2014, p. 3140-3148.
- LAM, S. K., A. PITROU et S. SEIBERT. “Numba : A LLVM-based Python JIT Compiler”. In : *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*. ACM, 2015, p. 1-6.
- LANCKRIET, G. R. G. et al. “Learning the kernel matrix with semidefinite programming”. In : *J. Mach. Learn. Res.* 5 (2004), p. 27-72.
- LANGE, K. *MM optimization algorithms*. Society for Industrial et Applied Mathematics, Philadelphia, PA, 2016, p. ix+223.
- LANGFORD, J. et J. SHAWE-TAYLOR. “PAC-Bayes & Margins”. In : *NIPS*. 2002, p. 423-430.
- LAPIN, M., M. HEIN et B. SCHIELE. “Analysis and optimization of loss functions for multi-class, top-k, and multilabel classification”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 40.7 (2017), p. 1533-1554.
- LAUB, A. J. *Matrix analysis for scientists and engineers*. Society for Industrial Mathematics, 2005.
- LAURENT, B. et P. MASSART. “Adaptive estimation of a quadratic functional by model selection”. In : *Ann. Statist.* 28.5 (2000), p. 1302-1338.
- LE MORVAN, M. et J.-P. VERT. “WHInter : A Working set algorithm for High-dimensional sparse second order Interaction models”. In : *ICML*. 2018, p. 3632-3641.
- LE PENNEC, E. et S. MALLAT. “Sparse geometric image representations with bandelets”. In : *IEEE Trans. Image Process.* 14.4 (2005), p. 423-438.
- LEBRUN, M., A. BUADES et J.-M. MOREL. “A Nonlocal Bayesian Image Denoising Algorithm”. In : *SIAM J. Imaging Sci.* 6.3 (2013), p. 1665-1688.
- LEBRUN, M. et al. “Secrets of image denoising cuisine”. In : *Acta Numerica* 21.1 (2012), p. 475-576.
- LECUÉ, G. “Optimal rates of aggregation in classification under low noise assumption”. In : *Bernoulli* 13.4 (2007), p. 1000-1022.
- LECUÉ, G. et M. LERASLE. “Robust machine learning by median-of-means : theory and practice”. In : *arXiv preprint arXiv :1711.10306* (2017).
- LECUÉ, G. et S. MENDELSON. “Aggregation via empirical risk minimization”. In : *Probab. Theory Related Fields* 145.3-4 (2009), p. 591-613.
- “On the optimality of the aggregate with exponential weights for low temperatures”. In : *submitted* (2010).
- LEDERER, J. “Trust, but verify : benefits and pitfalls of least-squares refitting in high dimensions”. In : *ArXiv e-prints* (2013).
- LEDERER, J. et C. L. MÜLLER. “Don’t Fall for Tuning Parameters : Tuning-Free Variable Selection in High Dimensions With the TREX”. In : *AAAI*. 2015, p. 2729-2735.

- LEDOIT, O. et M. WOLF. “A well-conditioned estimator for large-dimensional covariance matrices”. In : *J. Multivariate Anal.* 88.2 (2004), p. 365-411.
- LEDOUX, M. *Isoperimetry and Gaussian analysis*. Springer, 1996, p. 165-294.
- *The concentration of measure phenomenon*. T. 89. Mathematical Surveys and Monographs. Providence, RI : American Mathematical Society, 2001.
- LEDOUX, M. et M. TALAGRAND. *Probability in Banach spaces*. T. 23. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Isoperimetry and processes. Springer-Verlag, Berlin, 1991, p. xii+480.
- *Probability in Banach spaces*. Classics in Mathematics. Isoperimetry and processes, Reprint of the 1991 edition. Berlin : Springer-Verlag, 2011.
- LEE, J.-S. “Digital image smoothing and the sigma filter”. In : *Computer Vision, Graphics, and Image Processing* 24.2 (1983), p. 255-269.
- LEE, A. J. *U-Statistics : Theory and Practice*. Marcel Dekker, New York, 1990.
- LEE, D. D. et H.S. SEUNG. “Algorithms for non-negative matrix factorization”. In : *NIPS*. T. 13. Citeseer, 2001.
- LEE, H. et al. “Efficient sparse coding algorithms”. In : *NIPS*. 2007, p. 801-808.
- LEE, S., J. ZHU et E. P. XING. “Adaptive Multi-Task Lasso : with Application to eQTL Detection”. In : *NIPS*. 2010, p. 1306-1314.
- LEE, W. et Y. LIU. “Simultaneous multiple response regression and inverse covariance matrix estimation via penalized Gaussian maximum likelihood”. In : *Journal of multivariate analysis* 111 (2012), p. 241-255.
- LEFÈVRE, A., F. BACH et C. FÉVOTTE. “Itakura-Saito nonnegative matrix factorization with group sparsity”. In : *ICASSP*. IEEE. 2011, p. 21-24.
- LEFKIMMIATIS, S., P. MARAGOS et G. PAPANDREOU. “Bayesian inference on multiscale models for Poisson intensity estimation : Applications to photon-limited image denoising”. In : *IEEE Trans. Image Process.* 18.8 (2009), p. 1724-1741.
- LEGENDRE, A.-M. *Nouvelles méthodes pour la détermination des orbites des comètes*. F. Didot, 1805.
- LEHMANN, E. L. et G. CASELLA. *Theory of point estimation*. 2^e éd. Springer Texts in Statistics. New York : Springer-Verlag, 1998, p. xxvi+589.
- LEJEUNE, M. *Statistiques, la théorie et ses applications*. Springer, 2010.
- LEMMENS, B. et R. NUSSBAUM. *Nonlinear Perron-Frobenius theory*. T. 189. Cambridge Tracts in Mathematics. Cambridge : Cambridge University Press, 2012.
- LEPSKI, O. V. “On a problem of adaptive estimation in Gaussian white noise”. In : *Theory Probab. Appl.* 35.3 (1990), p. 454-466.
- “On problems of adaptive estimation in white Gaussian noise”. In : *Topics in nonparametric estimation*. T. 12. Adv. Soviet Math. Providence, RI : Amer. Math. Soc., 1992, p. 87-106.
- LEPSKI, O. V., E. MAMMEN et V. G. SPOKOINY. “Optimal spatial adaptation to inhomogeneous smoothness : an approach based on kernel estimates with variable bandwidth selectors”. In : *Ann. Statist.* 25.3 (1997), p. 929-947.

- LERASLE, M. et R. I. OLIVEIRA. “Robust empirical mean estimators”. In : *arXiv preprint arXiv :1112.3914* (2011).
- LEUNG, G. “Information Theory and Mixing Least Squares Regression”. Thèse de doct. Yale University, 2004.
- LEUNG, G. et A. R. BARRON. “Information theory and mixing least-squares regressions”. In : *IEEE Trans. Inf. Theory* 52.8 (2006), p. 3396-3410.
- LEVIN, A. et B. NADLER. “Natural Image Denoising : Optimality and Inherent Bounds”. In : *CVPR*. 2011.
- LI, K.-C. “From Stein’s unbiased risk estimates to the method of generalized cross validation”. In : *Ann. Statist.* 13.4 (1985), p. 1352-1377.
- LI, L. et al. “Hyperband : A Novel Bandit-Based Approach to Hyperparameter Optimization”. In : *ArXiv e-prints* (2016).
- LI, W., H. DAI et Y. ZHANG. “Location-aided fast distributed consensus in wireless networks”. In : *IEEE Trans. Inf. Theory* 56.12 (2010), p. 6208-6227.
- LI, X. et al. “A First Order Free Lunch for SQRT-Lasso”. In : *ArXiv e-prints* (2016).
- LI, Y. et Y. SINGER. “The Well Tempered Lasso”. In : *ICML* (2018), p. 3030-3038.
- LI, Z., A. USCHMAJEV et S. ZHANG. “On convergence of the maximum block improvement method”. In : *SIAM J. Optim.* 25.1 (2015), p. 210-233.
- LIANG, J., J. FADILI et G. PEYRÉ. “Activity Identification and Local Linear Convergence of Forward-Backward-type Methods”. In : *SIAM J. Optim.* 27.1 (2017), p. 408-437.
- “Local Linear Convergence of Forward-Backward under Partial Smoothness”. In : *NIPS*. 2014, p. 1970-1978.
- LIANG, L. et al. “Real-time texture synthesis by patch-based sampling”. In : *ACM Trans. Graph.* 20.3 (2001), p. 127-150.
- LIEW, C. K. “Inequality constrained least-squares estimation”. In : *J. Amer. Statist. Assoc.* 71.355 (1976), p. 746-751.
- LIN, Y. et H. H. ZHANG. “Component selection and smoothing in multivariate nonparametric regression”. In : *Ann. Statist.* 34.5 (2006), p. 2272-2297.
- LIU, S. et Y. YANG. “Combining models in longitudinal data analysis”. In : *Annals of the Institute of Statistical Mathematics* 64 (2012), p. 233-254.
- LIU, W., Z. YU et M. YANG. “Robust Elastic Net Regression”. In : *ArXiv e-prints* (2015).
- LIU, Y.-L. et al. “A Robust and Fast Non-Local Means Algorithm for Image Denoising”. In : *Journal of Computer Science and Technology* 23.2 (2008), p. 270-279.
- LIU, Z. et al. “Online EM algorithm for mixture with application to internet traffic modeling”. In : *Comput. Statist. Data Anal.* 50.4 (2006), p. 1052-1071.
- LLOYD, S. “Least squares quantization in PCM”. In : *IEEE Trans. Inf. Theory* 28.2 (1982), p. 129-137.
- LOH, P.-L. et M. J. WAINWRIGHT. “Regularized M-estimators with Nonconvexity : Statistical and Algorithmic Theory for Local Optima”. In : *J. Mach. Learn. Res.* 16 (2015), p. 559-616.

- LOPUHAÄ, H. P. et P. J. ROUSSEEUW. “Breakdown Points of Affine Equivariant Estimators of Multivariate Location and Covariance Matrices”. In : *Ann. Statist.* 19.1 (1991), p. 229-248.
- LORIEUL, T. et al. “Toward a large-scale and deep phenological stage annotation of herbarium specimens : Case studies from temperate, tropical, and equatorial floras”. In : *Applications in Plant Sciences* 7.3 (2019), e01233.
- LOTH, M. “Active Set Algorithms for the LASSO”. Thèse de doct. Université des Sciences et Technologie de Lille - Lille I, 2011.
- LOU, Y. et al. “Nonlocal Similarity Image Filtering”. In : *ICIAP '09 : Proceedings of the 15th International Conference on Image Analysis and Processing*. Vietri sul Mare, Italy : Springer-Verlag, 2009, p. 62-71.
- LOUCHET, C. “Variational and Bayesian models for image denoising : from total variation towards non-local means”. Thèse de doct. Université Paris Descartes, 2008.
- LOUCHET, C. et L. MOISAN. “Total variation as a local filter”. In : *SIAM J. Imaging Sci.* 4.2 (2011), p. 651-694.
- “Total Variation denoising using posterior expectation”. In : *EUSIPCO*. 2008.
- LOUNICI, K. “Estimation statistique en grande dimension, parcimonie et inégalités d’oracle”. Thèse de doct. Université Paris Diderot, 2009.
- “Generalized mirror averaging and D -convex aggregation”. In : *Math. Methods Statist.* 16.3 (2007), p. 246-259.
- “Sup-norm convergence rate and sign concentration property of Lasso and Dantzig estimators”. In : *Electron. J. Stat.* 2 (2008), p. 90-102.
- LOUNICI, K. et al. “Oracle inequalities and optimal inference under group sparsity”. In : *Ann. Statist.* 39.4 (2011), p. 2164-2204.
- LUISIER, F. et al. “Fast interscale wavelet denoising of Poisson-corrupted images”. In : *Signal Processing* 90.2 (2010), p. 415-427.
- LUSTIG, M., D. L. DONOHO et J. M. PAULY. “Sparse MRI : The application of compressed sensing for rapid MR imaging”. In : *Magnetic Resonance in Medicine* 58.6 (2007), p. 1182-1195.
- MA, J., G. PLONKA et M. Y. HUSSAINI. “Compressive video sampling with approximate message passing decoding”. In : *IEEE Trans. Circuits Syst. Video Technol.* 22.9 (2012), p. 1354-1364.
- MA, S., D. GOLDFARB et L. CHEN. “Fixed point and Bregman iterative methods for matrix rank minimization”. In : *Math. Program.* (2009), p. 1-33.
- MACKEY, D. J. C. *Information theory, inference and learning algorithms*. New York : Cambridge University Press, 2003, p. xii+628.
- MADSEN, H. et P. THYREGOD. *Introduction to general and generalized linear models*. CRC Press, 2010.
- MAGGIONI, M. et al. “A nonlocal transform-domain filter for volumetric data denoising and reconstruction”. In : *IEEE Trans. Image Process.* 22 (2013).

- MAGGIONI, M. et al. “Video Denoising, Deblocking, and Enhancement Through Separable 4-D Nonlocal Spatiotemporal Transforms”. In : *IEEE Trans. Image Process.* 21.9 (2012), p. 3952-3966.
- MAHMOUDI, M. et G. SAPIRO. “Fast image and video denoising via nonlocal means of similar neighborhoods”. In : *IEEE Trans. Signal Process. Lett.* 12 (2005), p. 839-842.
- MAIRAL, J. “Optimization with first-order surrogate functions”. In : *ArXiv e-prints* (2013).
- “Sparse coding for machine learning, image processing and computer vision”. Thèse de doct. École normale supérieure de Cachan, 2010.
- MAIRAL, J., F. BACH et J. PONCE. “Task-driven dictionary learning”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 34.4 (2012), p. 791-804.
- MAIRAL, J., G. SAPIRO et M. ELAD. “Learning multiscale sparse representations for image and video restoration”. In : *Multiscale Model. Simul.* 7.1 (2008), p. 214-241.
- MAIRAL, J. et B. YU. “Complexity analysis of the Lasso regularization path”. In : *ICML*. 2012, p. 353-360.
- MAIRAL, J. et al. “Convex and network flow optimization for structured sparsity”. In : *J. Mach. Learn. Res.* 12 (2011), p. 2681-2720.
- MAIRAL, J. et al. “Non-local sparse models for image restoration”. In : *ICCV*. 2009, p. 2272-2279.
- MAIRAL, J. et al. “Online learning for matrix factorization and sparse coding”. In : *J. Mach. Learn. Res.* (2010), p. 19-60.
- MÄKITALO, M. et A. FOI. “Optimal inversion of the Anscombe transformation in low-count Poisson image denoising”. In : *IEEE Trans. Image Process.* 20.1 (2011), p. 99-109.
- “Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise”. In : *submitted* (2012).
- “Optimal inversion of the generalized Anscombe transformation for Poisson-Gaussian noise”. In : *IEEE Trans. Image Process.* 22.1 (2013), p. 91-103.
- MALEKI, A., M. NARAYAN et R. G. BARANIUK. “Anisotropic Nonlocal Means”. In : *Appl. Comput. Harmon. Anal.* (2012).
- “Suboptimality of Nonlocal Means for Images with Sharp Edges”. In : *Appl. Comput. Harmon. Anal.* 33.3 (2012), p. 370-387.
- MALGOUYRES, F. *Projecting onto a polytope simplifies data distributions*. Rapp. tech. University Paris 13, 2006.
- “Rank related properties for Basis Pursuit and total variation regularization”. In : *Signal Processing* 87 (11 2007), p. 2695-2707.
- MALLAT, S. “A theory for multiresolution signal decomposition : The wavelet representation”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 11.7 (1989), p. 674-693.
- *A wavelet tour of signal processing*. San Diego, CA : Academic Press Inc., 1998, p. xxiv+577.
- *A wavelet tour of signal processing*. The sparse way, With contributions from Gabriel Peyré. Elsevier/Academic Press, Amsterdam, 2009, p. xxii+805.
- MALLAT, S. et Z. ZHANG. “Matching Pursuit With Time-Frequency Dictionaries”. In : *IEEE Trans. Image Process.* 41 (1993), p. 3397-3415.

- MALLOWS, C. L. “Some Comments on C_p ”. In : *Technometrics* 15.4 (1973), p. 661-675.
- MAMMEN, E. et A. B. TSYBAKOV. “Smooth discrimination analysis”. In : *Ann. Statist.* 27.6 (1999), p. 1808-1829.
- MAMMEN, E. et S. VAN DE GEER. “Locally adaptive regression splines”. In : *Ann. Statist.* 25.1 (1997), p. 387-413.
- MANDOZZI, J. et P. BÜHLMANN. “Hierarchical Testing in the High-Dimensional Setting With Correlated Variables”. In : *J. Amer. Statist. Assoc.* 111.513 (2016), p. 331-343.
- MANN, H. B. et D. R. WHITNEY. “On a test of whether one of two random variables is stochastically larger than the other”. In : *Ann. Math. Stat.* 18.1 (1947), p. 50-60.
- MANNING, C. D., P. RAGHAVAN et H. SCHÜTZE. *Introduction to Information Retrieval*. New York, NY, USA : Cambridge University Press, 2008.
- MARIN, J.-M. et C. P. ROBERT. *Bayesian core : a practical approach to computational Bayesian statistics*. Springer Texts in Statistics. New York : Springer, 2007, p. xiv+255.
- MARKOWITZ, H. “Portfolio selection”. In : *The Journal of Finance* 7.1 (1952), p. 77-91.
- “The optimization of a quadratic function subject to linear constraints”. In : *Naval Res. Logist. Quart.* 3 (1956), p. 111-133.
- MARONNA, R. A., R. D. MARTIN et V. J. YOHAI. *Robust statistics : Theory and methods*. Chichester : John Wiley & Sons, 2006.
- MARQUAND, A. et al. “Bayesian multi-task learning for decoding multi-subject neuroimaging data”. In : *NeuroImage* 92 (2014), p. 298-311.
- MARSHALL, A. W., I. OLKIN et B. C. ARNOLD. *Inequalities : theory of majorization and its applications*. Second. Springer Series in Statistics. Springer, 2011.
- MARTINET, B. “Brève communication. Régularisation d’inéquations variationnelles par approximations successives”. In : *Revue française d’informatique et de recherche opérationnelle. Série rouge* 4.R3 (1970), p. 154-158.
- MASSART, P. “About the constants in Talagrand’s concentration inequalities for empirical processes”. In : *Ann. Probab.* 28.2 (2000), p. 863-884.
- *Concentration inequalities and model selection*. T. 1896. Lecture Notes in Mathematics. Berlin : Springer, 2007.
- MASSART, P. et C. MEYNET. “The Lasso as an l_1 -ball model selection procedure”. In : *Electron. J. Stat.* 5 (2011), p. 669-687.
- MASSIAS, M., A. GRAMFORT et J. SALMON. “Celer : a Fast Solver for the Lasso with Dual Extrapolation”. In : *ICML*. T. 80. 2018, p. 3315-3324.
- “From safe screening rules to working sets for faster Lasso-type solvers”. In : *NIPS-OPT*. 2017.
- MASSIAS, M. et al. “Dual Extrapolation for Sparse Generalized Linear Models”. In : *arXiv preprint arXiv :1907.05830* (2019).
- MASSIAS, M. et al. “Generalized Concomitant Multi-Task Lasso for Sparse Multimodal Regression”. In : *AISTATS*. T. 84. 2018, p. 998-1007.
- MASSIAS, M. et al. “Support recovery and sup-norm convergence rates for sparse pivotal regression”. In : *AISTATS*. 2020.

- MATOUŠEK, J. “On variants of the Johnson–Lindenstrauss Lemma”. In : *Random Structures & Algorithms* 33.2 (2008), p. 142-156.
- MATOUŠEK, J. et B. GÄRTNER. *Understanding and Using Linear Programming*. 1^{re} éd. Universitext. Springer-Verlag Berlin Heidelberg, 2007.
- MAZUMDER, R., J. H. FRIEDMAN et T. J. HASTIE. “*SparseNet* : coordinate descent with nonconvex penalties”. In : *J. Amer. Statist. Assoc.* 106.495 (2011), p. 1125-1138.
- MAZUMDER, R., T. J. HASTIE et R. TIBSHIRANI. “Spectral regularization algorithms for learning large incomplete matrices”. In : *J. Mach. Learn. Res.* 11 (2010), p. 2287-2322.
- MCALLESTER, D. A. “Some PAC-Bayesian theorems”. In : *COLT*. Madison, Wisconsin, United States : ACM, 1998, p. 230-234.
- MCCULLAGH, P. et J. A. NELDER. *Generalized Linear Models*. Second. Chapman and Hall/CRC Monographs on Statistics and Applied Probability Series. Taylor & Francis, 1989.
- MCELREATH, R. *Statistical Rethinking : A Bayesian Course with Examples in R and Stan*. Chapman & Hall/CRC Texts in Statistical Science. Chapman et Hall/CRC, 2015.
- MCGILL, R., J. W. TUKEY et W. A. LARSEN. “Variations of box plots”. In : *The American Statistician* 32.1 (1978), p. 12-16.
- MCINNES, L. et J. HEALY. “UMAP : Uniform Manifold Approximation and Projection for Dimension Reduction”. In : *ArXiv* abs/1802.03426 (2018).
- MCKINNEY, W. *Python for Data Analysis : Data Wrangling with Pandas, NumPy, and IPython*. O’Reilly Media, 2012.
- MCMILLEN, T. “On the eigenvalues of double band matrices”. In : *Linear Algebra and Its Applications* 431.10 (2009), p. 1890-1897.
- MEINSHAUSEN, N. et P. BÜHLMANN. “High-dimensional graphs and variable selection with the lasso”. In : *Ann. Statist.* 34.3 (2006), p. 1436-1462.
- “Stability selection”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 72.4 (2010), p. 417-473.
- MEINSHAUSEN, N., L. MEIER et P. BÜHLMANN. “P-values for high-dimensional regression”. In : *J. Amer. Statist. Assoc.* 104.488 (2009), p. 1671-1681.
- MESBAHI, M. et M. EGERSTEDT. *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.
- MEYER, Y. *Wavelets and operators*. T. 37. Cambridge Studies in Advanced Mathematics. Translated from the 1990 French original by D. H. Salinger. Cambridge : Cambridge University Press, 1992, p. xvi+224.
- MEYN, S. et R. L. TWEEDIE. *Markov chains and stochastic stability*. Second. With a prologue by Peter W. Glynn. Cambridge : Cambridge University Press, 2009, p. xxviii+594.
- MI, F. et D.-Y. YEUNG. “Probabilistic graphical models for boosting cardinal and ordinal peer grading in MOOCs”. In : *Twenty-Ninth AAAI Conference on Artificial Intelligence*. 2015.
- MICCHELLI, C. A., J. M. MORALES et M. PONTIL. “A family of penalty functions for structured sparsity”. In : *NIPS*. 2010, p. 1612-1623.

- MICHELOT, C. “A finite algorithm for finding the projection of a point onto the canonical simplex of \mathbb{R}^n ”. In : *Journal of Optimization Theory and Applications* 50.1 (1986), p. 195-200.
- MIKA, S. et al. “Kernel PCA and De-Noising in Feature Spaces”. In : *NIPS*. 1999, p. 536-542.
- MILANFAR, P. “A tour of modern image filtering : new insights and methods, both practical and theoretical”. In : *IEEE Signal Process. Mag.* 30.1 (2012), p. 106-128.
- “Symmetrizing Smoothing Filters”. In : *SIAM J. Imaging Sci.* 6.1 (2013), p. 263-284.
- MINSKER, S. “Geometric median and robust estimation in Banach spaces”. In : *Bernoulli* 21.4 (2015), p. 2308-2335.
- MITRA, R. et C.-H. ZHANG. “The benefit of group sparsity in group inference with de-biased scaled group Lasso”. In : *Electron. J. Stat.* 10.2 (2016), p. 1829-1873.
- MOREAU, J.-J. “Fonctions convexes duales et points proximaux dans un espace hilbertien”. In : *C. R. Acad. Sci. Paris* 255 (1962), p. 2897-2899.
- “Proximité et dualité dans un espace hilbertien”. In : *Bull. Soc. Math. France* 93 (1965), p. 273-299.
- MOSK-AOYAMA, D. et D. SHAH. “Fast distributed algorithms for computing separable functions”. In : *IEEE Trans. Inf. Theory* 54.7 (2008), p. 2997-3007.
- MOSLER, K. “Depth statistics”. In : *Robustness and complex data structures*. Springer, 2013, p. 17-34.
- MOZHAROVSKIY, P. “Tukey depth : linear programming and applications”. In : *arXiv pre-print arXiv :1603.00069* (2016).
- MUIRHEAD, R. J. *Aspects of multivariate statistical theory*. Wiley Series in Probability and Mathematical Statistics. New York : John Wiley & Sons Inc., 1982.
- MULLER, J.-M. et al. *Handbook of Floating-Point Arithmetic (2nd Ed.)* Springer, 2018.
- MÜLLER, H.-G. et U. STADTMÜLLER. “Variable bandwidth kernel estimators of regression curves”. In : *Ann. Statist.* 15.1 (1987), p. 182-201.
- MÜLLER, C. L. et J. LEDERER. “Don’t fall for tuning parameters : tuning-free variable selection in high dimensions with the TREX”. In : *ArXiv e-prints* (2014).
- MURESAN, D. D. et T. W. PARKS. “Adaptive principal components and image denoising”. In : *ICIP*. 2003, p. 101-104.
- MURPHY, K. P. *Machine learning : a probabilistic perspective*. MIT press, 2012.
- NADARAYA, E. A. “On Estimating Regression”. In : *Theory of Probability and its Applications* 9.1 (1964), p. 141-142.
- NAKAGAWA, K. et al. “Safe Feature Pruning for Sparse High-Order Interaction Models”. In : *ArXiv e-prints* (2015).
- “Safe pattern pruning : An efficient approach for predictive pattern mining”. In : *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 2016, p. 1785-1794.
- NATARAJAN, B. K. “Sparse approximate solutions to linear systems”. In : *SIAM J. Comput.* 24.2 (1995), p. 227-234.

- NDIAYE, E. “Safe optimization algorithms for variable selection and hyperparameter tuning”. Thèse de doct. Université Paris-Saclay, 2018.
- NDIAYE, E. et al. “Efficient Smoothed Concomitant Lasso Estimation for High Dimensional Regression”. In : *Journal of Physics : Conference Series* 904.1 (2017), p. 012006.
- NDIAYE, E. et al. “GAP Safe screening rules for sparse multi-task and multi-class models”. In : *NIPS*. 2015, p. 811-819.
- “GAP Safe Screening Rules for Sparse-Group-Lasso”. In : *NIPS*. 2016.
- “Gap Safe screening rules for sparsity enforcing penalties”. In : *J. Mach. Learn. Res.* 18.128 (2017), p. 1-33.
- NDIAYE, E. et al. “Safe Grid Search with Optimal Complexity”. In : *ICML*. T. 97. 2019, p. 4771-4780.
- NEDIĆ, A., S. LEE et M. RAGINSKY. “Decentralized online optimization with global objectives and local communication”. In : *ACC*. 2015, p. 4497-4503.
- NEDIĆ, A. et A. E. OZDAGLAR. “Distributed subgradient methods for multi-agent optimization”. In : *IEEE Trans. Autom. Control* 54.1 (2009), p. 48-61.
- NEGAHBAN, S. et M. J. WAINWRIGHT. “Restricted strong convexity and weighted matrix completion : optimal bounds with noise”. In : *J. Mach. Learn. Res.* 13 (2012), p. 1665-1697.
- NEGAHBAN, S. et al. “A unified framework for high-dimensional analysis of M -estimators with decomposable regularizers”. In : *ArXiv e-prints* (2010).
- NEMIROVSKI, A. S. *Topics in non-parametric statistics*. T. 1738. Lecture Notes in Math. Berlin : Springer, 2000, p. 85-277.
- NEMIROVSKI, A. S. et D. B. YUDIN. *Problem complexity and method efficiency in optimization*. A Wiley-Interscience Publication. New York : John Wiley & Sons Inc., 1983, p. xv+388.
- NESTEROV, Y. “A method for solving a convex programming problem with rate of convergence $O(1/k^2)$ ”. In : *Soviet Math. Doklady* 269.3 (1983), p. 543-547.
- “Efficiency of coordinate descent methods on huge-scale optimization problems”. In : *SIAM J. Optim.* 22.2 (2012), p. 341-362.
- *Introductory lectures on convex optimization*. T. 87. Applied Optimization. Boston, MA : Kluwer Academic publishers, 2004.
- “Primal-dual subgradient methods for convex problems”. In : *Math. Program.* 120.1 (2009), p. 221-259.
- “Smooth minimization of non-smooth functions”. In : *Math. Program.* 103.1 (2005), p. 127-152.
- NGUYEN, T.-B., J.-A. CHEVALIER et B. THIRION. “ECKO : Ensemble of Clustered Knockoffs for Robust Multivariate Inference on fMRI Data”. In : *International Conference on Information Processing in Medical Imaging*. 2019, p. 454-466.
- NICULESCU-MIZIL, A. et R. CARUANA. “Predicting Good Probabilities with Supervised Learning”. In : *ICML*. 2005, p. 625-632.

- NIELSEN, F. et V. GARCIA. "Statistical exponential families : A digest with flash cards". In : *ArXiv e-prints* (2009).
- NIKOLOVA, M. "Local strong homogeneity of a regularized estimator". In : *SIAM J. Appl. Math.* 61.2 (2000), p. 633-658.
- NIU, F. et al. "Hogwild! : A lock-free approach to parallelizing stochastic gradient descent". In : *NIPS* 24 (2011), p. 693-701.
- NOCEDAL, J. et S. J. WRIGHT. *Numerical optimization*. Second. Springer Series in Operations Research and Financial Engineering. New York : Springer, 2006.
- NOLAN, D. et T. P. SPEED. *Stat labs : mathematical statistics through applications*. Springer Science & Business Media, 2001.
- NUTINI, J. et al. "Coordinate descent converges faster with the Gauss-Southwell rule than random selection". In : *ICML*. 2015, p. 1632-1641.
- OBOZINSKI, G., B. TASKAR et M. I. JORDAN. "Joint covariate selection and joint subspace selection for multiple classification problems". In : *Statistics and Computing* 20.2 (2010), p. 231-252.
- OBOZINSKI, G., M. J. WAINWRIGHT et M. JORDAN. "Support union recovery in high-dimensional multivariate regression". In : *Ann. Statist.* 39.1 (2011), p. 1-47.
- OCHS, P. et al. "On iteratively reweighted algorithms for nonsmooth nonconvex optimization in computer vision". In : *SIAM Journal on Imaging Sciences* 8.1 (2015), p. 331-372.
- OGAWA, K., Y. SUZUKI et I. TAKEUCHI. "Safe screening of non-support vectors in pathwise SVM computation". In : *ICML*. T. 28. 2013, p. 1382-1390.
- ORCHARD, J., M. EBRAHIMI et A. WONG. "Efficient nonlocal-means denoising using the SVD". In : *ICIP*. 2008, p. 1732-1735.
- OSBORNE, M. R., B. PRESNELL et B. A. TURLACH. "A new approach to variable selection in least squares problems". In : *IMA J. Numer. Anal.* 20.3 (2000), p. 389-403.
- OSHER, S. et al. "An iterative regularization method for total variation-based image restoration". In : *Multiscale Model. Simul.* 4.2 (2005), p. 460-489.
- OSHER, S. et al. "Sparse recovery via differential inclusions". In : *Appl. Comput. Harmon. Anal.* (2016).
- OU, W., M. HÄMALÄINEN et P. GOLLAND. "A Distributed Spatio-Temporal EEG/MEG Inverse Solver". In : *NeuroImage* 44.3 (2009), p. 932-946.
- OUVRARD, J.-Y. *Probabilités : Tome 2, Licence - CAPES*. 2^e éd. Enseignement des mathématiques. Cassini, 2007.
- OUVRARD, J.-Y. *Probabilités : Tome 1, Licence - CAPES*. 2^e éd. Enseignement des mathématiques. Cassini, 2008.
- OWEN, A. B. "A robust hybrid of lasso and ridge regression". In : *Contemporary Mathematics* 443 (2007), p. 59-72.
- PAGÈS, J. *Statistique générale pour utilisateurs. Méthodologie*. Presses universitaires de Rennes, 2010.
- PAN, L. "Human Age Estimation by Metric Learning for Regression Problems". In : *EMMCVPR*. Springer. Berlin, Heidelberg, 2009, p. 455-465.

- PANG, B., L. LEE et S. VAITHYANATHAN. "Thumbs Up ? : Sentiment Classification Using Machine Learning Techniques". In : *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*. EMNLP '02. Stroudsburg, PA, USA : Association for Computational Linguistics, 2002, p. 79-86.
- PANG, C. et al. "A Fast NL-Means Method in Image Denoising based on the Similarity of Spatially Sampled Pixels". In : *MMSP*. 2009.
- PARIKH, N. et S. BOYD. "Proximal Algorithms". In : *Foundations and Trends in Machine Learning* 1.3 (2014), p. 127-239.
- PARK, M. Y. "Generalized linear models with regularization". Thèse de doct. Stanford University, 2006.
- PARK, M. Y. et T. HASTIE. "L1-regularization path algorithm for generalized linear models". In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 69.4 (2007), p. 659-677.
- PARK, T. et G. CASELLA. "The Bayesian lasso". In : *J. Amer. Statist. Assoc.* 103.482 (2008), p. 681-686.
- PARTRIDGE, M. et R. CALVO. "Fast dimensionality reduction and simple PCA". In : *Intelligent Data Analysis* 2 (1998).
- PARZEN, E. "On estimation of a probability density function and mode". In : *Ann. Math. Statist.* 33 (1962), p. 1065-1076.
- PASZKE, A. et al. "Automatic Differentiation in PyTorch". In : *NIPS Autodiff Workshop*. 2017.
- PEARSON, K. "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling". In : *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 50.302 (1900), p. 157-175.
- PEDREGOSA, F. "Hyperparameter optimization with approximate gradient". In : *ICML*. T. 48. 2016, p. 737-746.
- PEDREGOSA, F. et al. "Scikit-learn : Machine Learning in Python". In : *J. Mach. Learn. Res.* 12 (2011), p. 2825-2830.
- PELCKMANS, K. et J. SUYKENS. "Gossip algorithms for computing U-statistics". In : *IFAC Workshop on Estimation and Control of Networked Systems*. 2009, p. 48-53.
- PENG, Z., M. YAN et W. YIN. "Parallel and distributed sparse optimization". In : *2013 Asilomar Conference on Signals, Systems and Computers*. IEEE. 2013, p. 659-646.
- PENG, Z. et al. "Coordinate-friendly structures, algorithms and applications". In : *Ann. Math. Sci. Appl.* 1.1 (2016), p. 57-119.
- PEREKRESTENKO, D., V. CEVHER et M. JAGGI. "Faster Coordinate Descent via Adaptive Importance Sampling". In : *AISTATS*. 2017, p. 869-877.
- PERONA, P. et J. MALIK. "Scale Space and Edge Detection Using Anisotropic Diffusion". In : *IEEE Trans. Pattern Anal. Mach. Intell.* 12 (1990), p. 629-639.
- PERRONE, M. P. "Improving regression estimation : Averaging methods for variance reduction with extensions to general convex measure optimization". Thèse de doct. Providence, RI, USA : Brown University, 1993.

- PERRONE, M. P. et L. N. COOPER. “When networks disagree : Ensemble method for neural networks”. In : *Artificial Neural Networks for Speech and Vision*. 1993, p. 126-142.
- PEYPOUQUET, J. *Convex optimization in normed spaces : theory, methods and examples*. Springer, 2015.
- PEYRÉ, G. “Image processing with nonlocal spectral bases.” In : *Multiscale Model. Simul.* 7.2 (2008), p. 703-730.
- “Manifold Models for Signals and Images”. In : *Computer Vision and Image Understanding* 113.2 (2009), p. 249-260.
- PHAM, V. et L. EL GHAOU. “Robust sketching for multiple square-root LASSO problems”. In : *AISTATS*. 2015.
- PIECH, C. et al. “Tuned Models of Peer Assessment in MOOCs”. In : *Proceedings of the 6th International Conference on Educational Data Mining*. 2013, p. 153-160.
- PINHEIRO, J. et D. BATES. *Mixed-effects models in S and S-PLUS*. Springer Science & Business Media, 2006.
- PINSKER, M. S. “Optimal filtration of square-integrable signals in Gaussian noise”. In : *Probl. Peredachi Inf.* 16.2 (1980), p. 52-68.
- PISIER, G. *The volume of convex bodies and Banach space geometry*. T. 94. Cambridge Tracts in Mathematics. Cambridge : Cambridge University Press, 1989.
- PLACKETT, R. L. “Studies in the History of Probability and Statistics. XXIX : The discovery of the method of least squares”. In : *Biometrika* 59.2 (1972), p. 239-251.
- PLAN, Y. et R. VERSHYNIN. “One-bit compressed sensing by linear programming”. In : *Communications on Pure and Applied Mathematics* (2013).
- “Robust 1-bit compressed sensing and sparse logistic regression : a convex programming approach”. In : *IEEE Trans. Inf. Theory* (2012).
- PLATT, J. C. “Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods”. In : *ADVANCES IN LARGE MARGIN CLASSIFIERS*. MIT Press, 1999, p. 61-74.
- POLLARD, D. *A user’s guide to measure theoretic probability*. T. 8. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge : Cambridge University Press, 2002.
- *Convergence of stochastic processes*. Springer Series in Statistics. New York : Springer-Verlag, 1984.
- POLZEHL, J. et V. G. SPOKOINY. “Adaptive weights smoothing with applications to image restoration”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 62.2 (2000), p. 335-354.
- “Image denoising : pointwise adaptive approach”. In : *Ann. Statist.* 31.1 (2003), p. 30-57.
- “Propagation-separation approach for local likelihood estimation”. In : *Probab. Theory Related Fields* 135.3 (2006), p. 335-362.
- POON, C., J. LIANG et C. SCHOENLIEB. “Local Convergence Properties of SAGA/Prox-SVRG and Acceleration”. In : *ICML*. 2018, p. 4124-4132.
- PORIKLI, F. M. “Constant time $O(1)$ bilateral filtering”. In : *CVPR*. 2008, p. 1-8.

- PORIKLI, F. M. “Integral Histogram : A Fast Way To Extract Histograms in Cartesian Spaces”. In : *CVPR*. T. 1. IEEE Computer Society, 2005, p. 829-836.
- PORTILLA, J. et al. “Image denoising using scale mixtures of Gaussians in the wavelet domain”. In : *IEEE Trans. Image Process.* 12.11 (2003), p. 1338-1351.
- POWELL, J. L. “Censored regression quantiles”. In : *J. Econometrics* 32.1 (1986), p. 143-155.
- POWLEY, B. W. “Quantile function methods for decision analysis”. Thèse de doct. Stanford University, 2013.
- PRIESTLEY, M. B. et M. T. CHAO. “Non-parametric function fitting”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 34 (1972), p. 385-392.
- QIN, Z., K. SCHEINBERG et D. GOLDFARB. “Efficient block-coordinate descent algorithms for the group lasso”. In : *Math. Program. Comput.* 5.2 (2013), p. 143-169.
- QUARTERONI, A., R. SACCO et F. SALERI. *Numerical mathematics*. Second. T. 37. Texts in Applied Mathematics. Berlin : Springer-Verlag, 2007.
- QUINLAN, J. R. “Induction of Decision Trees”. In : *Machine Learning* 1 (1986), p. 81-106.
- RAGINSKY, M. et al. “Sequential anomaly detection in the presence of noise and limited feedback”. In : *Submitted* (2010).
- RAHIMI, A. et B. RECHT. “Random features for large-scale kernel machines”. In : *NIPS*. Sous la dir. de J.C. PLATT et al. Curran Associates, Inc., 2008, p. 1177-1184.
- “Weighted Sums of Random Kitchen Sinks : Replacing minimization with randomization in learning”. In : *NIPS*. Sous la dir. de D. KOLLER et al. Curran Associates, Inc., 2009, p. 1313-1320.
- RAI, P., A. KUMAR et H. DAUME. “Simultaneously leveraging output and task structures for multiple-output regression”. In : *NIPS*. 2012, p. 3185-3193.
- RAJ, A. et al. “Screening Rules for Convex Problems”. In : *arXiv preprint arXiv :1609.07478* (2016).
- RAKOTOMAMONJY, A., G. GASSO et J. SALMON. “Screening Rules for Lasso with Non-Convex Sparse Regularizers”. In : *ICML*. T. 97. 2019, p. 5341-5350.
- RAM, S., A. NEDIĆ et V. VEERAVALLI. “Distributed stochastic subgradient projection algorithms for convex optimization”. In : *J. Optimiz. Theory. App.* 147.3 (2010), p. 516-545.
- RAMANI, S., T. BLU et M. UNSER. “Monte-Carlo SURE : a black-box optimization of regularization parameters for general denoising algorithms”. In : *IEEE Trans. Image Process.* 17.9 (2008), p. 1540-1554.
- RAMDAS, A. et R. J. TIBSHIRANI. “Fast and flexible ADMM algorithms for trend filtering”. In : *Journal of Computational and Graphical Statistics* 25.3 (2016), p. 839-858.
- RAMSAY, J. O. et B. W. SILVERMAN. *Functional data analysis*. Second. Springer Series in Statistics. New York : Springer, 2005.
- RANINEN, E. et E. OLLILA. “Scaled and square-root elastic net”. In : *ArXiv e-prints* (2017).
- RAO, C. R. “Estimation of heteroscedastic variances in linear models”. In : *J. Amer. Statist. Assoc.* 65 (1970), p. 161-172.
- RAO, C. R. et Y. WU. “On model selection”. In : *Model selection*. T. 38. IMS Lecture Notes Monogr. Ser. Inst. Math. Statist., 2001, p. 1-64.

- RAPHAN, M. et E. P. SIMONCELLI. “Learning to be Bayesian without supervision”. In : *NIPS*. T. 19. 2007, p. 1145.
- RECHT, B., M. FAZEL et P. A. PARRILO. “Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization”. In : *SIAM review* 52.3 (2010), p. 471-501.
- RECHT, B. et C. RÉ. “Parallel stochastic gradient algorithms for large-scale matrix completion”. In : *Math. Program. Comput.* 5.2 (2013), p. 201-226.
- REID, S., R. TIBSHIRANI et J. FRIEDMAN. “A study of error variance estimation in lasso regression”. In : *Stat. Sin.* 26.1 (2016), p. 35-67.
- RENDLE, S. “Factorization machines with libfm”. In : *ACM Transactions on Intelligent Systems and Technology (TIST)* 3.3 (2012), p. 57.
- RENDLE, Steffen. “Factorization Machines”. In : *ICDM 2010, The 10th IEEE International Conference on Data Mining, Sydney, Australia, 14-17 December 2010*. 2010, p. 995-1000. DOI : 10.1109/ICDM.2010.127.
- RIGOLLET, P. “Kullback-Leibler aggregation and misspecified generalized linear models”. In : *Ann. Statist.* 40.2 (2012), p. 639-665.
- RIGOLLET, P. et A. B. TSYBAKOV. “Exponential Screening and optimal rates of sparse estimation”. In : *Ann. Statist.* 39.2 (2011), p. 731-471.
- “Linear and convex aggregation of density estimators”. In : *Math. Methods Statist.* 16.3 (2007), p. 260-280.
- “Sparse estimation by exponential weighting”. In : *Statist. Sci.* 27.4 (2012), p. 558-575.
- RIPLEY, B. D. *Pattern recognition and neural networks*. Cambridge university press, 1996.
- ROBERT, C. P. *Méthodes de Monte Carlo par chaînes de Markov*. Sticistique Mathématique et Probabilité. [Mathematical Statistics and Probability]. Paris : Éditions Économica, 1996, p. x+340.
- *The Bayesian choice*. Second. Springer Texts in Statistics. New York : Springer-Verlag, 2001, p. xxiv+604.
- ROBERT, C. P. et G. CASELLA. *Monte Carlo statistical methods*. 2^e éd. Springer Texts in Statistics. New York : Springer-Verlag, 2004, p. xxx+645.
- ROCKAFELLAR, R. T. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton, NJ : Princeton University Press, 1997.
- ROCKAFELLAR, R. T., J. O. ROYSET et S. I. MIRANDA. “Superquantile regression with applications to buffered reliability, uncertainty quantification, and conditional value-at-risk”. In : *European J. Oper. Res.* 234.1 (2014), p. 140-154.
- ROCKAFELLAR, R. T. et S. URYASEV. “Conditional value-at-risk for general loss distributions”. In : *J. Bank. Finance* 26.7 (2002), p. 1443-1471.
- “Optimization of conditional value-at-risk”. In : *Journal of Risk* 2 (2000), p. 21-42.
- “The fundamental risk quadrangle in risk management, optimization and statistical estimation”. In : *Surveys in Operations Research and Management Science* 18.1-2 (2013), p. 33-53.

- ROCKAFELLAR, R. T. et R. J.-B. WETS. *Variational analysis*. T. 317. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin : Springer-Verlag, 1998.
- ROMAN, S. *Coding and information theory*. T. 134. Graduate Texts in Mathematics. New York : Springer-Verlag, 1992, p. xviii+486.
- ROMANO, Y. et M. ELAD. “Boosting of Image Denoising Algorithms”. In : *SIAM J. Imaging Sci.* 8.2 (2015), p. 1187-1219.
- ROSENBAUM, M. et A. TSYBAKOV. “Sparse recovery under matrix uncertainty”. In : *The Annals of Statistics* 38.5 (2010), p. 2620-2651.
- ROSENBLATT, M. “Remarks on some nonparametric estimates of a density function”. In : *Ann. Math. Statist.* 27 (1956), p. 832-837.
- ROSSET, S. et J. ZHU. “Piecewise linear regularized solution paths”. In : *Ann. Statist.* 35.3 (2007), p. 1012-1030.
- ROSSET, S., J. ZHU et T. HASTIE. “Boosting as a regularized path to a maximum margin classifier”. In : *J. Mach. Learn. Res.* 5 (2004), p. 941-973.
- ROTH, S. et M. J. BLACK. “Fields of Experts : A Framework for Learning Image Priors”. In : *CVPR*. T. 2. 2005, p. 860-867.
- ROTH, V. et B. FISCHER. “The Group-Lasso for generalized linear models : uniqueness of solutions and efficient algorithms”. In : *ICML*. 2008, p. 848-855.
- ROTHMAN, A. J., E. LEVINA et J. ZHU. “Sparse multivariate regression with covariance estimation”. In : *Journal of Computational and Graphical Statistics* 19.4 (2010), p. 947-962.
- ROUSSEEUW, P. J. “Least median of squares regression”. In : *J. Amer. Statist. Assoc.* 79.388 (1984), p. 871-880.
- ROUSSEEUW, P. J. et A. M. LEROY. *Robust regression and outlier detection*. Wiley Series in Probability and Mathematical Statistics : Applied Probability and Statistics. New York : John Wiley & Sons Inc., 1987, p. xvi+329.
- ROUSSEEUW, P. J. et I. RUTS. “Algorithm AS 307 : Bivariate location depth”. In : *J. R. Stat. Soc. Ser. C. Appl. Stat.* 45.4 (1996), p. 516-526.
- ROUSSEEUW, P. J. et K. VAN DRIESSEN. “Computing LTS Regression for Large Data Sets”. In : *Data Mining and Knowledge Discovery* 12.1 (2006), p. 29-45.
- ROUVIÈRE, F. *Petit guide de calcul différentiel : à l’usage de la licence et de l’agrégation*. Enseignement des mathématiques. Cassini, 2009.
- ROY, N., G. J. GORDON et S. THRUN. “Finding approximate POMDP solutions through belief compression”. In : *J. Artif. Intell. Res.* 23.1 (2005), p. 1-40.
- RUDIN, L. I., S. OSHER et E. FATEMI. “Nonlinear total variation based noise removal algorithms”. In : *Phys. D* 60.1-4 (1992), p. 259-268.
- RUDIN, W. *Real and complex analysis*. Third. New York : McGraw-Hill Book Co., 1987, p. xiv+416.
- RUMELHART, D. E., G. E. HINTON et R. J. WILLIAMS. “Learning representations by back-propagating errors”. In : *Nature* 323.6088 (1986), p. 533.

- RUPPERT, D. et M. P. WAND. “Multivariate locally weighted least squares regression”. In : *Ann. Statist.* 22.3 (1994), p. 1346-1370.
- SADLER, P. M. et E. GOOD. “The impact of self-and peer-grading on student learning”. In : *Educational assessment* 11.1 (2006), p. 1-31.
- SAHA, A. et A. TEWARI. “On the nonasymptotic convergence of cyclic coordinate descent methods”. In : *SIAM J. Optim.* 23.1 (2013), p. 576-601.
- SAJJADI, M. S. M., M. ALAMGIR et U. VON LUXBURG. “Peer Grading in a Course on Algorithms and Data Structures : Machine Learning Algorithms Do Not Improve over Simple Baselines”. In : *Proceedings of the Third (2016) ACM Conference on Learning at Scale*. 2016, p. 369-378.
- SALAKHUTDINOV, R. et G. HINTON. “An efficient learning procedure for deep Boltzmann machines”. In : *Neural computation* 24.8 (2012), p. 1967-2006.
- SALLEE, P. et B. A. OLSHAUSEN. “Image denoising using learned overcomplete representations”. In : *ICIP*. 2003, p. 381-384.
- SALMON, J. “Agrégation d’estimateurs et méthodes à patch pour le débruitage d’images numériques”. Thèse de doct. Université Paris Diderot, 2010.
- “On two parameters for denoising with Non-Local Means”. In : *IEEE Trans. Signal Process. Lett.* 17 (2010), p. 269-272.
- “Perspectives computationnelles et statistiques pour la régression en grande dimension”. Habilitation à diriger des recherches (HDR). École Normale Supérieure Paris-Saclay, 2017.
- SALMON, J. et A. S. DALALYAN. “Optimal aggregation of affine estimators”. In : *COLT*. 2011.
- SALMON, J. et E. LE PENNEC. “Wavelet and Bayesian Models”. Mém. de mast. ENSAE ParisTech et Université Paris 7-Diderot, 2007.
- SALMON, J. et E. LE PENNEC. “An aggregator point of view on NL-Means”. In : *Proceedings of the SPIE Conference on Mathematical Imaging : Wavelet XIII*. T. 7446. San Diego, CA, USA : SPIE, 2009, 74461E.
- “NL-Means and aggregation procedures”. In : *ICIP*. 2009, p. 2977-2980.
- SALMON, J. et Y. STROZECKI. “From Patches to Pixels in Non-Local methods : Weighted-Average Reprojection”. In : *ICIP*. 2010, p. 1929-1932.
- “Patch Reprojections for Non Local Methods”. In : *Signal Processing* 92.2 (2012), p. 447-489.
- SALMON, J., R. WILLETT et E. ARIAS-CASTRO. “A two-stage denoising filter : the preprocessed Yaroslavsky filter”. In : *SSP*. 2012.
- SALMON, J. et al. “Poisson noise reduction with non-local PCA”. In : *ICASSP*. 2012.
- SALMON, J. et al. “Poisson noise reduction with non-local PCA”. In : *J. Math. Imaging Vis.* 48.2 (2014), p. 279-294.
- SANGNIER, M., O. FERCOQ et F. D’ALCHÉ-BUC. “Joint quantile regression in vector-valued RKHSs”. In : *NIPS*. 2016, p. 3693-3701.

- SANKARAN, R., F. BACH et C. BHATTACHARYYA. “Identifying groups of strongly correlated variables through Smoothed Ordered Weighted ℓ_1 -norms”. In : *AISTATS*. 2017.
- SAPIRO, G. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, 2001.
- SCHAIRER, T. et al. “Parallel Non-Local Denoising of Depth Maps”. In : *LNLA*. Lausanne, 2008.
- SCHAPIRE, R. E. “The Strength of Weak Learnability”. In : *Mach. Learn.* 5 (2 1990), p. 197-227.
- SCHEUER, E. M. “Reliability of an m -out-of- n system when component failure induces higher failure rates in survivors”. In : *IEEE Trans. Rel.* 37.1 (1988), p. 73-74.
- SCHMIDT, M., N. LE ROUX et F. BACH. “Minimizing Finite Sums with the Stochastic Average Gradient”. In : *arXiv preprint arXiv :1309.2388* (2013).
- SCHNEIDER, R. *Convex bodies : the Brunn–Minkowski theory*. Second. T. 151. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2013, p. 760.
- SCHÖLKOPF, B. et A. J. SMOLA. *Learning with kernels : Support vector machines, regularization, optimization, and beyond*. MIT press, 2002.
- SCHÖLKOPF, B., A. J. SMOLA et K-R. MÜLLER. “Nonlinear Component Analysis as a Kernel Eigenvalue Problem”. In : *Neural Comput.* 10.5 (1998), p. 1299-1319.
- SCHOTT, J. R. *Matrix analysis for statistics*. Second. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], 2005, p. xvi+456.
- SCHWARZ, G. “Estimating the dimension of a model”. In : *Ann. Statist.* 6.2 (1978), p. 461-464.
- SCIEUR, D. “Acceleration in Optimization”. Thèse de doct. École normale supérieure, 2018.
- SCIEUR, D., A. D’ASPREMONT et F. BACH. “Regularized Nonlinear Acceleration”. In : *NIPS*. 2016, p. 712-720.
- SCOTT, C. et R. D. NOWAK. “Minimax-optimal classification with dyadic decision trees”. In : *IEEE Trans. Inf. Theory* 52.4 (2006), p. 1335-1353.
- SCOTT, D.W. *Multivariate density estimation*. T. 139. Wiley Online Library, 1992.
- SEARLE, S. R., G. CASELLA et C. E. MCCULLOCH. *Variance components*. T. 391. John Wiley & Sons, 2009.
- SEBER, G. A. F. et A. J. LEE. *Linear Regression Analysis, 2nd edition (Wiley Series in Probability and Statistics)*. 2^e éd. Wiley, 2003.
- SEEGER, M. “Covariance Kernels from Bayesian Generative Models”. In : *NIPS*. 2002, p. 905-912.
- “Pac-bayesian generalisation error bounds for Gaussian process classification”. In : *J. Mach. Learn. Res.* 3 (2003), p. 233-269.
- SELESNICK, I. et M. FARSHCHIAN. “Sparse Signal Approximation via Nonseparable Regularization”. In : *IEEE Trans. Signal Process.* 65.10 (2017), p. 2561-2575.
- SENETA, E. *Non-negative matrices and Markov chains*. Springer Series in Statistics. New York : Springer, 2006, p. xvi+287.

- SEO, H.J. et P. MILANFAR. “Training-free, Generic Object Detection using Locally Adaptive Regression Kernels”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* 32.9 (2010), p. 1688-1704.
- SERVAJEAN, M., R. CHAILAN et A. JOLY. “Non-parametric Bayesian annotator combination”. In : *Information Sciences* 436 (2018), p. 131-145.
- SERVAJEAN, M. et al. “Crowdsourcing Thousands of Specialized Labels : a Bayesian active training approach”. In : *IEEE Transactions on Multimedia* 19.6 (2017), p. 1376-1391.
- SERVAJEAN, M. et al. “Increasing Coverage in Distributed Search and Recommendation with Profile Diversity”. In : *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXII*. Springer, 2015, p. 115-144.
- SHAFIEEZADEH-ABADEH, S., P. M. ESFAHANI et D. KUHN. “Distributionally Robust Logistic Regression”. In : *NIPS*. Sous la dir. de C. CORTES et al. Curran Associates, Inc., 2015, p. 1576-1584.
- SHAH, D. “Gossip algorithms”. In : *Foundations and Trends in Networking* 3.1 (2009), p. 1-125.
- SHAH, R. D. et R. J. SAMWORTH. “Variable selection with error control : Another look at Stability Selection”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 75.1 (2013), p. 55-80.
- SHALEV-SHWARTZ, S. “Online learning and online convex optimization”. In : *Foundations and Trends in Machine Learning* 4.2 (2012), p. 107-194.
- SHALEV-SHWARTZ, S. et S. BEN-DAVID. *Understanding machine learning : From theory to algorithms*. Cambridge University Press, 2014.
- SHALEV-SHWARTZ, S. et T. ZHANG. “Accelerated proximal stochastic dual coordinate ascent for regularized loss minimization”. In : *arXiv preprint arXiv :1309.2375* (2013).
- “Accelerated proximal stochastic dual coordinate ascent for regularized loss minimization”. In : *Math. Program.* 155.1 (2016), p. 105-145.
- SHALEV-SHWARTZ, S. et al. “Pegasos : Primal Estimated sub-GrAdient Solver for SVM”. In : *Math. Program.* 127.1 (2011), p. 3-30.
- SHAN, K. et Y. YANG. “Combining regression quantile estimators”. In : *Statist. Sinica* 19.3 (2009), p. 1171-1191.
- SHAO, J. *Mathematical statistics : exercises and solutions*. New York : Springer, 2005, p. xxviii+359.
- SHAO, J. et X. DENG. “Estimation in high-dimensional linear models with deterministic design matrices”. In : *Ann. Statist.* 40.2 (2012), p. 812-831.
- SHAWE-TAYLOR, J. et N. CRISTIANINI. *An introduction to support vector machines : and other kernel-based learning methods*. Cambridge University Press, 2000.
- SHAWE-TAYLOR, J. et R. C. WILLIAMSON. “A PAC analysis of a Bayesian estimator”. In : *COLT*. New York, NY, USA : ACM, 1997, p. 2-9.
- SHE, Y et A. B. OWEN. “Outlier detection using nonconvex penalized regression”. In : *J. Amer. Statist. Assoc.* 106.494 (2011).
- SHE, Y. “Thresholding-based iterative selection procedures for model selection and shrinkage”. In : *Electron. J. Stat.* 3 (2009), p. 384-415.

- SHEVADE, S. K. et S. S. KEERTHI. “A simple and efficient algorithm for gene selection using sparse logistic regression”. In : *Bioinformatics* 19.17 (2003), p. 2246-2253.
- SHI, H.-J. M. et al. “A primer on coordinate descent algorithms”. In : *ArXiv e-prints* (2016).
- SHIBAGAKI, A. et al. “Regularization Path of Cross-Validation Error Lower Bounds”. In : *NIPS*. 2015, p. 1666-1674.
- SHIBAGAKI, A. et al. “Simultaneous safe screening of features and samples in doubly sparse modeling”. In : *ICML*. T. 48. 2016, p. 1577-1586.
- SHUMWAY, R. H. et D. S. STOFFER. *Time series analysis and its applications : with R examples*. Springer, 2017.
- SILVERMAN, B. W. *Density estimation for statistics and data analysis*. Monographs on Statistics and Applied Probability. London : Chapman & Hall, 1986, p. x+175.
- SIMON, N. et R. TIBSHIRANI. “Standardization and the Group Lasso penalty”. In : *Stat. Sin.* 22.3 (2012), p. 983-1001.
- SIMON, N. et al. “A sparse-group lasso”. In : *J. Comput. Graph. Statist.* 22.2 (2013), p. 231-245.
- SINGER, A., Y. SHKOLNISKY et B. NADLER. “Diffusion interpretation of nonlocal neighborhood filters for signal denoising.” In : *SIAM J. Imaging Sci.* 2.1 (2009), p. 118-139.
- SINGER, A. C. et M. FEDER. “Universal linear prediction by model order weighting”. In : *IEEE Trans. Signal Process.* 47 (1999), p. 2685-2699.
- SINGER, Y. et M. K. WARMUTH. “Batch and on-line parameter estimation of Gaussian mixtures based on the joint entropy”. In : *NIPS*. 1999, p. 578-584.
- SINGH, A. P. et G. J. GORDON. “A unified view of matrix factorization models”. In : *Machine Learning and Knowledge Discovery in Databases* (2008), p. 358-373.
- “Relational learning via collective matrix factorization”. In : *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2008, p. 650-658.
- SKIENA, S. S. *The algorithm design manual*. T. 1. Springer Science & Business Media, 1998.
- SLAWSKI, M. et M. HEIN. “Non-negative least squares for high-dimensional linear models : consistency and sparse recovery without regularization”. In : *arXiv preprint arXiv :1205.0953* (2012).
- SMITH, S. M. et J. M. BRADY. “SUSAN-A new approach to low level image processing”. In : *Int. J. Comput. Vision* 23.1 (1997), p. 45-78.
- SMYTH, P. et D. H. WOLPERT. “Linearly Combining Density Estimators via Stacking”. In : *Mach. Learn.* 36.1-2 (1999), p. 59-83.
- SOLO, V. “A sure-fired way to choose smoothing parameters in ill-conditioned inverse problems”. In : *ICIP*. T. 3. 1996, p. 89-92.
- SOUBIES, E., L. BLANC-FÉRAUD et G. AUBERT. “A Unified View of Exact Continuous Penalties for ℓ_2 - ℓ_0 Minimization”. In : *SIAM J. Optim.* 27.3 (2017), p. 2034-2060.
- SOUTHWELL, R. V. “Relaxation methods in engineering science : a treatise on approximate computation”. In : *The Mathematical Gazette* 25.265 (1941), p. 180-182.

- SPIRA, A. et R. KIMMEL. “Enhancing images painted on Manifolds”. In : *Scale Space and PDF Methods in Computer Vision*. 2005, p. 492-502.
- SPOKOINY, V. G. “Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice”. In : *Ann. Statist.* 26.4 (1998), p. 1356-1378.
- SPRECHMANN, P., A. BRONSTEIN et G. SAPIRO. “Real-time online singing voice separation from monaural recordings using robust low-rank modeling”. In : *ISMIR*. 2012.
- SPRECHMANN, P. et al. “C-hilasso : A collaborative hierarchical sparse modeling framework”. In : *IEEE Trans. Signal Process.* 59.9 (2011), p. 4183-4198.
- STÄDLER, N., P. BÜHLMANN et S. VAN DE GEER. “ ℓ_1 -penalization for mixture regression models”. In : *TEST* 19.2 (2010), p. 209-256.
- STARCK, J.-L., E. J. CANDÈS et D. L. DONOHO. “The curvelet transform for image denoising”. In : *IEEE Trans. Image Process.* 11.6 (2002), p. 670-684.
- STEIN, C. “Inadmissibility of the usual estimator for the mean of a multivariate normal distribution”. In : *Proceeding of the 3rd Berkeley Symposium on Mathematical Statistics and Probability*. T. 1. 399. 1956, p. 197-206.
- STEIN, C. M. “Estimation of the mean of a multivariate distribution”. In : *Proc. Prague Symp. Asymptotic Statist.* Prague, 1973.
- “Estimation of the mean of a multivariate normal distribution”. In : *Ann. Statist.* 9.6 (1981), p. 1135-1151.
- STEINHAUS, H. “Sur la division des corps matériels en parties”. In : *Bull. Acad. Polon. Sci. Cl. III*. 4 (1956), 801-804 (1957).
- STICH, S., A. RAJ et M. JAGGI. “Safe Adaptive Importance Sampling”. In : *NIPS*. 2017, p. 4384-4394.
- STRANG, G. *Introduction to linear algebra*. SIAM, 2003.
- *Introduction to linear algebra*. 5th edition. Wellesley, MA : Wellesley-Cambridge Press, 2016, p. x + 574.
- STRONG, D. M. et T. F. CHAN. “Edge-preserving and scale-dependent properties of total variation regularization”. In : *Inverse Problems* 19.6 (2003). Special section on imaging, S165-S187.
- STUCKY, B. et S. VAN DE GEER. “Asymptotic confidence regions for high-dimensional structured sparsity”. In : *IEEE Trans. Signal Process.* 66.8 (2018), p. 2178-2190.
- STURM, J. F. “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones”. In : *Optimization Methods and Software* 11-12 (1999), p. 625-653.
- SUN, T. et C.-H. ZHANG. “Comments on : ℓ_1 -penalization for mixture regression models”. In : *TEST* 19.2 (2010), p. 270-275.
- “Scaled sparse linear regression”. In : *Biometrika* 99.4 (2012), p. 879-898.
- “Sparse Matrix Inversion with Scaled Lasso”. In : *J. Mach. Learn. Res.* 14 (2013), p. 3385-3418.
- SUN, Y. et al. “Are we there yet? Manifold identification of gradient-related proximal methods”. In : *AISTATS*. 2019, p. 1110-1119.

- SZLAM, A. D., M. MAGGIONI et R. R. COIFMAN. “Regularization on graphs with function-adapted diffusion processes”. In : *J. Mach. Learn. Res.* 9 (2008), p. 1711-1739.
- TAKEDA, H., S. FARSIU et P. MILANFAR. “Kernel Regression for Image Processing and Reconstruction”. In : *IEEE Trans. Image Process.* 16.2 (2007), p. 349-366.
- TAKEUCHI, I. et al. “Nonparametric quantile estimation”. In : *Journal of Machine Learning Research* 7.Jul (2006), p. 1231-1264.
- TALEBI, H. et P. MILANFAR. “Global image denoising”. In : *IEEE Trans. Image Process.* 23.2 (2014), p. 755-768.
- TALEBI, H., X. ZHU et P. MILANFAR. “How to SAIF-ly boost denoising performance”. In : *IEEE Trans. Image Process.* 22.4 (2013), p. 1470-1485.
- TALWAR, K., A. G. THAKURTA et L. ZHANG. “Nearly optimal private lasso”. In : *NIPS*. 2015, p. 3025-3033.
- TANG, G., B. N. BHASKAR et B. RECHT. “Near minimax line spectral estimation”. In : *IEEE Trans. Inf. Theory* 61.1 (2015), p. 499-512.
- TANG, G. et al. “Compressed sensing off the grid”. In : *IEEE Trans. Inf. Theory* 59.11 (2013), p. 7465-7490.
- TASDIZEN, T. “Principal components for non-local means image denoising”. In : *ICIP*. 2008, p. 1728-1731.
- “Principal neighborhood dictionaries for nonlocal means image denoising”. In : *IEEE Trans. Image Process.* 18.12 (2009), p. 2649-2660.
- TAYLOR, H. L., S. C. BANKS et J. F. MCCOY. “Deconvolution with the ℓ_1 norm”. In : *Geophysics* 44.1 (1979), p. 39-52.
- TAYLOR, K. M. “The geometry of signal and image patch-sets”. Thèse de doct. University of Colorado at Boulder, 2011.
- TAYLOR, K. M. et F. G. MEYER. “A random walk on image patches”. In : *Arxiv preprint arXiv :1107.0414* (2011).
- TIBSHIRANI, R. “Regression Shrinkage and Selection via the Lasso”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 58.1 (1996), p. 267-288.
- TIBSHIRANI, R., G. WALTHER et T. J. HASTIE. “Estimating the number of clusters in a data set via the gap statistic”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 63.2 (2001), p. 411-423.
- TIBSHIRANI, R. et al. “Sparsity and Smoothness via the fused LASSO”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67.1 (2005), p. 91-108.
- TIBSHIRANI, R. et al. “Strong rules for discarding predictors in lasso-type problems”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 74.2 (2012), p. 245-266.
- TIBSHIRANI, R. J. “Dykstra’s Algorithm, ADMM, and Coordinate Descent : Connections, Insights, and Extensions”. In : *NIPS*. 2017, p. 517-528.
- “The lasso problem and uniqueness”. In : *Electron. J. Stat.* 7 (2013), p. 1456-1490.
- TIBSHIRANI, R. J. et J. TAYLOR. “Degrees of freedom in lasso problems”. In : *Ann. Statist.* 40.2 (2012), p. 1198-1232.
- “The solution path of the generalized lasso”. In : *Ann. Statist.* 39.3 (2011), p. 1335-1371.

- TIKHONOV, A. N. “On the stability of inverse problems”. In : *Dokl. Akad. Nauk SSSR* 39 (1943), p. 176-179.
- TODESCHINI, A., F. CARON et M. CHAVENT. “Probabilistic Low-Rank Matrix Completion with Adaptive Spectral Regularization Algorithms”. In : *NIPS*. 2013, p. 845-853.
- TOMASI, C. et R. MANDUCHI. “Bilateral Filtering for Gray and Color Images”. In : *ICCV*. 1998, p. 839-846.
- TREFETHEN, L. N. et D. III BAU. *Numerical linear algebra*. Philadelphia, PA : Society for Industrial et Applied Mathematics (SIAM), 1997.
- TROPP, J. A. “Greed is good : algorithmic results for sparse approximation”. In : *IEEE Trans. Inf. Theory* 50.10 (2004), p. 2231-2242.
- “Just relax : convex programming methods for identifying sparse signals in noise”. In : *IEEE Trans. Inf. Theory* 52.3 (2006), p. 1030-1051.
- “User-friendly tail bounds for sums of random matrices”. In : *Found. Comput. Math.* 12.4 (2012), p. 389-434.
- TRUONG, Y. K. “Asymptotic properties of kernel estimators based on local medians”. In : *Ann. Statist.* 17.2 (1989), p. 606-617.
- TSCHUMPERLÉ, D. et L. BRUN. “Image Denoising and Registration by PDE’s on the Space of Patches”. In : *LNLA*. Lausanne, 2008, p. 32-40.
- “Non-Local Image Smoothing by Applying Anisotropic Diffusion PDE’s in the Space of Patches”. In : *ICIP*. 2009.
- TSENG, P. “Convergence of a block coordinate descent method for nondifferentiable minimization”. In : *J. Optim. Theory Appl.* 109.3 (2001), p. 475-494.
- TSENG, P. et S. YUN. “Block-coordinate gradient descent method for linearly constrained nonsmooth separable optimization”. In : *J. Optim. Theory Appl.* 140.3 (2009), p. 513.
- TSIANOS, K., S. LAWLOR et M. RABBAT. “Push-Sum distributed dual averaging for convex optimization”. In : *CDC*. 2015.
- TSITSIKLIS, J. N. “Problems in decentralized decision making and computation”. Thèse de doct. Massachusetts Institute of Technology, 1984.
- TSYBAKOV, A. B. *Introduction à l’estimation non-paramétrique*. T. 41. Mathématiques & Applications (Berlin) [Mathematics & Applications]. Berlin : Springer-Verlag, 2004, p. x+175.
- *Introduction to nonparametric estimation*. Springer Series in Statistics. New York : Springer, 2009.
- “Optimal orders of accuracy of the estimation of nonsmooth images”. In : *Problemy Peredachi Informatsii* 25.3 (1989), p. 13-27.
- “Optimal Rates of Aggregation”. In : *COLT*. 2003, p. 303-313.
- *Statistique appliquée*. 2006.
- TUKEY, J. W. *Exploratory data analysis*. Addison-Wesley Publishing Company, 1977.
- “Mathematics and the picturing of data”. In : *Proceedings of the International Congress of Mathematicians, Vancouver, 1975*. T. 2. 1975, p. 523-531.

- TURLACH, B. A., W. N. VENABLES et S. J. WRIGHT. “Simultaneous variable selection”. In : *Technometrics* 47.3 (2005), p. 349-363.
- VAINSENCHER, D., H. LIU et T. ZHANG. “Local smoothness in variance reduced optimization”. In : *NIPS*. 2015, p. 2179-2187.
- VAITER, S. et al. “Local Behavior of Sparse Analysis Regularization : Applications to Risk Estimation”. In : *Appl. Comput. Harmon. Anal.* 35.3 (2013), p. 433-451.
- VAITER, S. et al. “Robust sparse analysis regularization”. In : *IEEE Trans. Inf. Theory* 59.4 (2013), p. 2001-2016.
- VALÉRY, P. *Mauvaises pensées et autres*. Gallimard, 1942.
- VAN AELST, S. et P. J. ROUSSEEUW. “Minimum volume ellipsoid”. In : *Wiley Interdisciplinary Reviews : Computational Statistics* 1.1 (2009), p. 71-82.
- VAN DE GEER, S. *Estimation and testing under sparsity*. T. 2159. Lecture Notes in Mathematics. Lecture notes from the 45th Probability Summer school held in Saint-Flour, 2015, École d’Été de Probabilités de Saint-Flour. Springer, 2016, p. xiii+274.
- VAN DE GEER, S. et P. BÜHLMANN. “On the conditions used to prove oracle results for the Lasso”. In : *Electron. J. Stat.* 3 (2009), p. 1360-1392.
- VAN DE GEER, S. et J. LEDERER. “The Lasso, correlated design, and improved oracle inequalities”. In : *From Probability to Statistics and Back : High-Dimensional Models and Processes—A Festschrift in Honor of Jon A. Wellner*. Institute of Mathematical Statistics, 2013, p. 303-316.
- VAN DE GEER, S. et B. STUCKY. “ χ^2 -confidence sets in high-dimensional regression”. In : *Statistical analysis for high-dimensional data*. Springer, 2016, p. 279-306.
- VAN DE GEER, S. et al. “On asymptotically optimal confidence regions and tests for high-dimensional models”. In : *Ann. Statist.* 42.3 (2014), p. 1166-1202.
- VAN DE GEER, S. A. *Empirical processes in M-estimation*. Cambridge university press, 2000.
- VAN DE VILLE, D. et M. KOCHER. “Non-Local Means with Dimensionality Reduction and SURE-Based Parameter Selection”. In : *IEEE Trans. Image Process.* 99.99 (2011), p. 1-1.
- “SURE-Based Non-Local Means”. In : *IEEE Trans. Signal Process. Lett.* 16 (2009), p. 973-976.
- VAN DE WEIJER, J. et R. VAN DEN BOOMGAARD. “Local Mode Filtering”. In : *CVPR*. T. 2. 2001, p. 428-433.
- VAN DER MAATEN, L. et G. HINTON. “Visualizing data using t-SNE”. In : *J. Mach. Learn. Res.* 9.Nov (2008), p. 2579-2605.
- VAN DER VAART, A. W. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2000.
- VAN DER VAART, A. W. et JON A. WELLNER. *Weak convergence and empirical processes*. Springer Series in Statistics. With applications to statistics. New York : Springer-Verlag, 1996.

- VANDEL, J. et al. “Probing transcription factor combinatorics in different promoter classes and in enhancers”. In : *BMC Genomics* 20.1 (2019).
- VANDERPLAS, J. *Python Data Science Handbook*. O’Reilly Media, 2016.
- VAPNIK, V. N. *Statistical learning theory*. Wiley, 1998.
- VAROQUAUX, G., A. GRAMFORT et B. THIRION. “Small-sample brain mapping : sparse recovery on spatially correlated designs with randomization and clustering”. In : *ICML*. 2012, p. 1375-1382.
- VIOLA, P. et M. JONES. “Rapid Object Detection using a Boosted Cascade of Simple Features”. In : *CVPR*. T. 1. 2001, p. 511-518.
- VISHWANATHAN, S. V. N., A. J. SMOLA et M. N. MURTY. “SimpleSVM”. In : *ICML*. 2003, p. 760-767.
- VON NEUMANN, J. et O. MORGENSTERN. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey, 1944, p. xviii+625.
- VŮ, B. C. “A splitting algorithm for dual monotone inclusions involving cocoercive operators”. In : *Advances in Computational Mathematics* 38.3 (2013), p. 667-681.
- WAGENER, J. et H. DETTE. “Bridge estimators and the adaptive Lasso under heteroscedasticity”. In : *Math. Methods Statist.* 21 (2 2012), p. 109-126.
- WAI, H.-T. et al. “Decentralized Frank–Wolfe algorithm for convex and nonconvex problems”. In : *IEEE Transactions on Automatic Control* 62.11 (2017), p. 5522-5537.
- WAINWRIGHT, M. J. “Sharp Thresholds for High-Dimensional and Noisy Sparsity Recovery Using ℓ_1 -Constrained Quadratic Programming (Lasso)”. In : *IEEE Trans. Inf. Theory* 55.5 (2009), p. 2183-2202.
- WAND, M. P. et M. C. JONES. *Kernel smoothing*. T. 60. Monographs on Statistics and Applied Probability. London : Chapman et Hall Ltd., 1995, p. xii+212.
- WANG, Y.-X. et al. “Trend Filtering on Graphs”. In : *arXiv preprint arXiv :1410.7690* (2014).
- WANG, J., P. WONKA et J. YE. “Lasso Screening Rules via Dual Polytope Projection”. In : *arXiv preprint arXiv :1211.3966* (2012).
- WANG, J. et J. YE. “Multi-Layer Feature Reduction for Tree Structured Group Lasso via Hierarchical Projection”. In : *NIPS*. 2015, p. 1279-1287.
- “Two-Layer Feature Reduction for Sparse-Group Lasso via Decomposition of Convex Sets”. In : *arXiv preprint arXiv :1410.4210* (2014).
- WANG, J. et al. “A Safe Screening Rule for Sparse Logistic Regression”. In : *NIPS*. 2014, p. 1053-1061.
- WANG, J. et al. “Fast Non-Local Algorithm for Image Denoising”. In : *ICIP*. 2006, p. 1429-1432.
- WANG, J. et al. “Lasso screening rules via dual polytope projection”. In : *NIPS*. 2013, p. 1070-1078.
- WANG, W. et M. Á. CARREIRA-PERPIÑÁN. “Projection onto the probability simplex : An efficient algorithm with a simple proof, and an application”. In : *arXiv preprint arXiv :1309.1541* (2013).

- WANG, Z. et al. “Image quality assessment : from error visibility to structural similarity”. In : *IEEE Trans. Signal Process.* 13.4 (2004), p. 600-612.
- WASSERMAN, L. *All of Nonparametric Statistics (Springer Texts in Statistics)*. Springer, 2007.
- WASSERMAN, L. et K. ROEDER. “High-dimensional variable selection”. In : *Ann. Statist.* 37.5A (2009), p. 2178-2201.
- WATSON, G. A. “Characterization of the subdifferential of some matrix norms”. In : *Linear Algebra and its Applications* 170 (1992), p. 33-45.
- WATSON, G. S. “Smooth Regression Analysis”. In : *Sankhya : The Indian Journal of Statistics, Series A* 26.4 (1964), p. 359-372.
- WEI, E. et A. OZDAGLAR. “Distributed alternating direction method of multipliers”. In : *CDC*. 2012, p. 5445-5450.
- “On the $O(1/k)$ Convergence of Asynchronous Distributed Alternating Direction Method of Multipliers”. In : *IEEE GlobalSIP*. 2013.
- WEI, X. et S. MINSKER. “Estimation of the covariance structure of heavy-tailed distributions”. In : *NIPS*. 2017, p. 2859-2868.
- WEICKERT, J. *Anisotropic diffusion in image processing*. European Consortium for Mathematics in Industry. Stuttgart : B. G. Teubner, 1998, p. xii+170.
- WEISSMAN, T. et al. “Universal discrete denoising : known channel”. In : *IEEE Trans. Inf. Theory* 51.1 (2005), p. 5-28.
- WEISZFELD, E. “Sur le point pour lequel la somme des distances de n points donnés est minimum”. In : *Tohoku Mathematical Journal, First Series* 43 (1937), p. 355-386.
- WILK, M. B. et R. GNANADESIKAN. “Probability plotting methods for the analysis for the analysis of data”. In : *Biometrika* 55.1 (1968), p. 1-17.
- WILLETT, R. “Multiscale Analysis of Photon-Limited Astronomical Images”. In : *Statistical Challenges in Modern Astronomy (SCMA) IV*. 2006.
- WILLETT, R. et R. D. NOWAK. “Fast Multiresolution Photon-Limited Image Reconstruction”. In : *Proc. IEEE Int. Sym. Biomedical Imaging — ISBI '04*. 2004.
- “Platelets : A multiscale approach for recovering edges and surfaces in photon-limited medical imaging”. In : *IEEE Trans. Med. Imag.* 22.3 (2003), p. 332-350.
- WILLIAMS, D. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge : Cambridge University Press, 1991.
- WIPF, D. P. et al. “Estimating the Location and Orientation of Complex, Correlated Neural Activity using MEG”. In : *NIPS*. 2008, p. 1777-1784.
- WISHART, J. “The generalised product moment distribution in samples from a normal multivariate population”. In : *Biometrika* (1928), p. 32-52.
- WITHERS, C. S. et S. NADARAJAH. “Expressions for the distribution and percentiles of the sums and products of chi-squares”. In : *Statistics* (2012).
- WOLKE, R. et H. SCHWETLICK. “Iteratively Reweighted Least Squares : Algorithms, Convergence Analysis, and Numerical Comparisons”. In : *SIAM Journal on Scientific and Statistical Computing* 9.5 (1988), p. 907-921.

- WOLPERT, D. H. “Stacked generalization”. In : *Neural Networks* 5.2 (1992), p. 241-259.
- WRIGHT, J. R., C. THORNTON et K. LEYTON-BROWN. “Mechanical TA : Partially automated high-stakes peer grading”. In : *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*. 2015, p. 96-101.
- WRIGHT, S. J. “Accelerated Block-Coordinate Relaxation for regularized optimization”. In : *SIAM J. Optim.* 22.1 (2012), p. 159-186.
- WRIGHT, S. J., R. D. NOWAK et M. A. T. FIGUEIREDO. “Sparse reconstruction by separable approximation”. In : *IEEE Trans. Signal Process.* 57.7 (2009), p. 2479-2493.
- WU, T. T. et K. LANGE. “Coordinate descent algorithms for lasso penalized regression”. In : *Ann. Appl. Stat.* (2008), p. 224-244.
- XIANG, Z. J. et P. J. RAMADGE. “Fast lasso screening tests based on correlations”. In : *ICASSP*. 2012, p. 2137-2140.
- XIANG, Z. J., Y. WANG et P. J. RAMADGE. “Screening tests for lasso problems”. In : *IEEE Trans. Pattern Anal. Mach. Intell.* PP.99 (2016).
- XIANG, Z. J., H. XU et P. J. RAMADGE. “Learning sparse representations of high dimensional data on large scale dictionaries”. In : *NIPS*. 2011, p. 900-908.
- XIAO, L. “Dual averaging methods for regularized stochastic learning and online optimization”. In : *J. Mach. Learn. Res.* 11 (2010), p. 2543-2596.
- XU, H., C. CARAMANIS et S. MANNOR. “Robust regression and Lasso”. In : *IEEE Trans. Inf. Theory* 56.7 (2010), p. 3561-3574.
- XU, H. et al. “Linear Total Variation Approximate Regularized Nuclear Norm Optimization for Matrix Completion”. In : *Abstr. Appl. Anal.* (2014), Art. ID 765782, 8.
- XU, J. et S. OSHER. “Iterative regularization and nonlinear inverse scale space applied to wavelet-based denoising”. In : *IEEE Trans. Image Process.* 16.2 (2007), p. 534-544.
- XU, P. et P. J. RAMADGE. “Three structural results on the lasso problem”. In : *ICASSP*. 2013, p. 3392-3396.
- YANG, X., S. KIM et E. P. XING. “Heterogeneous multitask learning with joint sparsity constraints”. In : *NIPS*. 2009, p. 2151-2159.
- YANG, Y. “Adaptive estimation in pattern recognition by combining different procedures”. In : *Statist. Sinica* 10.4 (2000), p. 1069-1089.
- “Adaptive regression by mixing”. In : *J. Amer. Statist. Assoc.* 96.454 (2001), p. 574-588.
- “Aggregating regression procedures to improve performance”. In : *Bernoulli* 10.1 (2004), p. 25-47.
- “Combining different procedures for adaptive regression”. In : *J. Multivariate Anal.* 74.1 (2000), p. 135-161.
- “Combining forecasting procedures : some theoretical results”. In : *Econometric Theory* 20.1 (2004), p. 176-222.
- “Mixing strategies for density estimation”. In : *Ann. Statist.* 28.1 (2000), p. 75-87.
- “Regression with multiple candidate models : selecting or mixing?” In : *Statist. Sinica* 13.3 (2003), p. 783-809.

- YANG, Y., J. MA et S. OSHER. “Seismic data reconstruction via matrix completion”. In : *Inverse Probl. Imaging* 7.4 (2013), p. 1379-1392.
- YANG, Y. et al. “Deep ADMM-Net for Compressive Sensing MRI”. In : *NIPS*. 2016, p. 10-18.
- YAROSLAVSKY, L. P. *Digital picture processing*. T. 9. Springer Series in Information Sciences. Berlin : Springer-Verlag, 1985, p. xii+276.
- YIN, W. et al. “Bregman iterative algorithms for l_1 -minimization with applications to compressed sensing”. In : *SIAM J. Imaging Sci.* 1.1 (2008), p. 143-168.
- YU, G., G. SAPIRO et S. MALLAT. “Image modeling and enhancement via structured sparse model selection”. In : *ICIP*. 2010.
- “Solving Inverse Problems With Piecewise Linear Estimators : From Gaussian Mixture Models to Structured Sparsity”. In : *IEEE Trans. Image Process.* 21.5 (2012), p. 2481-2499.
- YUAN, D. et al. “Distributed dual averaging method for multi-agent optimization with quantized communication”. In : *Systems & Control Letters* 61.11 (2012), p. 1053-1061.
- YUAN, G, C.-H HO et C.-J LIN. “An improved GLMNET for l_1 -regularized logistic regression”. In : *J. Mach. Learn. Res.* 13 (2012), p. 1999-2030.
- YUAN, M. et Y. LIN. “Model selection and estimation in regression with grouped variables”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 68.1 (2006), p. 49-67.
- YUAN, Zh. et Y. YANG. “Combining linear regression models : when and how?” In : *J. Amer. Statist. Assoc.* 100.472 (2005), p. 1202-1214.
- YUEH, W.C. “Eigenvalues of several tridiagonal matrices”. In : *Applied Mathematics E-Notes* 5 (2005), p. 66-74.
- YUN, S. “On the iteration complexity of cyclic coordinate gradient descent methods”. In : *SIAM J. Optim.* 24.3 (2014), p. 1567-1580.
- ZENG, X. et M. A. T. FIGUEIREDO. “The Ordered Weighted ℓ_1 Norm : Atomic Formulation, Projections, and Algorithms”. In : *ArXiv e-prints* (2014).
- ZENG, Y. et P. BREHENY. “The biglasso Package : A Memory- and Computation-Efficient Solver for Lasso Model Fitting with Big Data in R”. In : *arXiv preprint arXiv :1701.05936* (2017).
- ZEWAIL, A. H. et J. M. THOMAS. *4D electron microscopy : imaging in space and time*. Imperial College Pr, 2009.
- ZHANG, C.-H. “Nearly unbiased variable selection under minimax concave penalty”. In : *Ann. Statist.* 38.2 (2010), p. 894-942.
- ZHANG, C.-H. et S. S. ZHANG. “Confidence intervals for low dimensional parameters in high dimensional linear models”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 76.1 (2014), p. 217-242.
- ZHANG, C.-H. et T. ZHANG. “A general theory of concave regularization for high-dimensional sparse estimation problems”. In : *Statist. Sci.* 27.4 (2012), p. 576-593.
- ZHANG, B., J. FADILI et J.-L. STARCK. “Wavelets, ridgelets, and curvelets for Poisson noise removal”. In : *IEEE Trans. Image Process.* 17.7 (2008), p. 1093-1108.

- ZHANG, J. et al. “Complexity of Finding Stationary Points of Nonsmooth Nonconvex Functions”. In : *CoRR* abs/2002.04130 (2020). eprint : 2002.04130.
- ZHANG, L. et al. “Two-stage image denoising by principal component analysis with local pixel grouping”. In : *Pattern Recogn.* 43.4 (2010), p. 1531-1549.
- ZHANG, T. “Adaptive forward-backward greedy algorithm for learning sparse representations”. In : *IEEE Trans. Inf. Theory* 57.7 (2011), p. 4689-4708.
- “Information-theoretic upper and lower bounds for statistical estimation”. In : *IEEE Trans. Inf. Theory* 52.4 (2006), p. 1307-1321.
- ZHAO, P. et al. “Online AUC Maximization”. In : *ICML*. 2011, p. 233-240.
- ZHOU, S. “Thresholding procedures for high dimensional variable selection and statistical estimation”. In : *NIPS*. 2009, p. 2304-2312.
- ZHU, J. et T. J. HASTIE. “Support vector machines, kernel logistic regression and boosting”. In : *Multiple Classifier Systems* 2364 (2002), p. 16-26.
- ZIMMER, S., S. DIDAS et J. WEICKERT. “A rotationally invariant block matching strategy improving image denoising with non-local means”. In : *LNLA*. 2008.
- ZIMMERT, J. et al. “Safe screening for support vector machines”. In : *NIPS 2015 Workshop on Optimization in Machine Learning (OPT)*. 2015.
- ZONTAK, M. et M. IRANI. “Internal statistics of a single natural image”. In : *CVPR*. 2011, p. 977-984.
- ZORAN, D. et Y. WEISS. “From learning models of natural image patches to whole image restoration”. In : *ICCV*. 2011, p. 479-486.
- ZOU, H. “The adaptive lasso and its oracle properties”. In : *J. Amer. Statist. Assoc.* 101.476 (2006), p. 1418-1429.
- ZOU, H. et T. J. HASTIE. “Regularization and variable selection via the elastic net”. In : *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67.2 (2005), p. 301-320.
- ZOU, H., T. J. HASTIE et R. TIBSHIRANI. “On the “degrees of freedom” of the lasso”. In : *Ann. Statist.* 35.5 (2007), p. 2173-2192.
- “Sparse principal component analysis”. In : *J. Comput. Graph. Statist.* 15.2 (2006), p. 265-286.
- ZOUA, E. P. “Subsampling estimates of the Lasso distribution.” Mém. de mast. ETH, 2011.
- ZUMER, J. M. et al. “A Probabilistic Algorithm Integrating Source Localization and Noise Suppression of MEG and EEG data”. In : *NIPS*. 2007, p. 1625-1632.
- ZUUR, A. et al. *Mixed effects models and extensions in ecology with R*. Springer Science & Business Media, 2009.