Alessandro Rudi

INRIA, École Normale Supérieure
45 rue d'Ulm, Paris, France

alessandro.rudi@ens.fr
www.di.ens.fr/~rudi

Activity

Permanent researcher, INRIA & École Normale Supérieure, Paris, 2019 - now.
Starting Research position, INRIA & École Normale Supérieure, Paris, 2017-2019.
Post-doc, IIT & MIT, 2014-2017.
Visiting Ph. D., Massachusetts Institute of Technology, 2012-2013.
Ph. D., University of Genova & Istituto Italiano di Tecnologia, Italy, 2011-2014.

Grants & Awards

ERC Starting Grant
1.5 million € for the period 2021-2026

Paris AI Research Institute - Chair
200k € for the period 2024-2027
186k € for the period 2020-2023

Recognitions

3 Long Plenary Talks at NeurIPS (2020 [41], 2017 [62], 2015 [66])
NeurIPS (∼20k participants) is the most important venue for AI & Machine Learning (ML).
Papers accepted with oral presentation (0.5% of submitted papers)

5 Short Plenary Talks at NeurIPS (2023 [9], 2021 [24], [25], 2020 [40], 2018 [59])
Papers accepted with spotlight presentations (1.5% of the submitted papers)

Consulting

EliseAI, New York, USA.
ML models to predict tenants' behavior to reduce delinquency rates – 2023-now

EPHOS Quantum Technologies, Milan, Italy.
ML techniques to improve the mathematical models of optical quantum chips – 2022-now

SNCF (French National train company) Paris, France.
ML models to improve traffic prediction for offer optimization – 2022-2023

Light-On extreme-scale AI, Paris, France.
Derive faster ML algorithms based on Light-On optical computing device – 2019-2022

Kamet Ventures (incubator from AXA Insurance) Paris, France.
Integration of ML techniques in price optimization. – 2019-2020

Teaching

MVA master class on Kernel Methods @ ENS, Paris – 2023-now
B.Sc. class on statistical machine learning @ ENS, Paris – 2018-now
M.Sc. class on Optimization & ML @ IASD, Paris-Dauphine – 2018-2022

Software

I designed and developed Falkon: a multi-GPU library for large scale ML
Fastest library to date for non-parametric large scale ML with complete guarantees [63]
awarded with NeurIPS oral presentation in 2020 [41].

My team

I lead a group of more than 10 people including interns, PhD and Post-docs,
affiliated at INRIA and École Normale Supérieure, rue d'Ulm, Paris.

Scientific and Technical Leadership

1. I am one of the leading international experts in *large-scale machine learning* and *convex/non-convex optimization*. The excellence of my contributions can be evaluated by the fact that (1) I was awarded the

prestigious ERC grant on the topic (2) by the number of publications that have been selected as oral presentations at the NeurIPS conference (the top venue for the AI & Machine Learning scientific and industrial community) and by the reference international journals on the topic. (3) and by the reception of the community, in terms of the number of citations & talk invitations by senior colleagues in well-established and prestigious venues.

2. The ambition of my work is to solve real-world problems with algorithms that are at the same time effective in practice and guaranteed by the theory. In particular, one of the main goals of my research so far has been to develop both theoretical tools and practical algorithms that are (a) scalable in terms of number of examples and dimensionality of each example (b) provably robust to noise and with guarantees on the quality of the predictions & with explicit control of the resulting uncertainty (c) adaptive to different computational architectures, especially distributed settings, GPUs and low-precision hardware. An example is the award-winning Falkon algorithm that I designed and then developed with my team. On the one hand, we provide full theoretical guarantees on the prediction quality (statistical and computational optimality). On the other hand, we developed a library that allowed to scale up learning of a factor 1000x (with Falkon now it is possible to learn with guarantees from billions of data in minutes). The crucial aspect there is not only the careful implementation part, coded in Python + C for multi-GPU, but also the refined mathematical analysis that allowed to dramatically improve the computational complexity of the state-of-the-art methods on the topic. This led to an algorithm that is still the best, to date, in terms of precision and computational cost, for non-parametric regression with confidence intervals.

3. My research, conducted over the past two years within the framework of my ERC, is now extending this kind of approach to other significant problems in applied mathematics, like *non-convex optimization* and *simulation of stochastic differential equations* and opens the path to adaptive scaling algorithms in large-dimensional instances across various classes of applied mathematics problems to address contemporary applications in engineering, finance, operations research, economics.

Experience in Industry

In my consulting activity, which I started five years ago, I focus my advanced problem-solving skills on critical real-world challenges where I can bring quantitative and measurable improvements. I approach problem-solving with an open mind, ready to learn the matter at hand, but also to bring my deep knowledge and skills in terms of mathematical modeling and design/integration/development. The solution must be not only effective in practice but also, when possible, guaranteed (i.e. quantify the error/uncertainty of the solution; characterize the possible modes of failure) so it can be used flawlessly in the client's business.

I complement my technical expertise with key methodological and organizational skills, fundamental for achieving a complex project. In particular, strategic thinking; the ability to analyze, formulate a plan, prioritize & make milestones that provide incremental and breakthrough improvements, and achieve the main goal on time; coordinating and guiding a team of technical experts from other fields with different skills and perspectives.

Scientific and Technical Expertise.

- *AI & Machine Learning*, *Convex/Non-convex Optimization*, *Mathematical Modeling*.
  In these fields I am an internationally recognized expert on the subject; I substantially contributed to the advancement of the field; I have a deep scientific & theoretical comprehension and extensive practical experience, especially in real-world scenarios.
- *Large Language Models; Transformers; Deep Neural Networks; Reinforcement Learning; Online Learning; High-performance Computing; GPU-Computing*.
  In these fields I am an expert: I have a deep scientific & theoretical comprehension and extensive practical experience; especially in real-world scenarios.

Major Collaborations

- Francis Bach (*INRIA & École Normale Supérieure, Paris*). since 2017 stochastic methods for ML. 15 papers.
- Lorenzo Rosasco (*U. Genova & MIT, USA*). since 2012. Large scale ML. 12 papers.
- Simone Severini (*Director of Quantum Computing at Amazon*), since 2017, quantum comp. & ML, 2 papers.
- Quentin Berthet, (*Google DeepMind*), since 2023, randomized linear algebra & machine learning.

- Volkan Cevher (*EPFL, Lausanne, Swisse*), since 2019, on non-parametric methods. 1 paper.
- Ernesto De Vito (*U. Genova*), Inverse problems & Machine Learning
- Jean Ponce (*NYU, USA & INRIA, Paris*), since 2019, ML applied to computer vision, 1 paper.
- Massimiliano Pontil (*IIT & University College of London, UK*), since 2017, multitask learning, 3 papers.
- Florent Krazkala (*EPFL, Lausanne, Swisse*), since 2018, large scale ML.
- Carlo Ciliberto (*UCL, UK*), since 2015. unified framework for structured prediction. 8 papers.
- Florence D'Alche-Buc (*Telecom ParisTech, Paris*), since 2019, non-parametric regression methods.
- Emanuele Borgonovo (*Bocconi University, Milan*), since 2017, ML applications to econometrics, 1 paper.
- Jonathan Weed (*Courant Institute & NYU, USA*), since 2018, optimal transport & ML, 2 papers.
- Jason Altschuler (*U. Penn. USA*), since 2018, optimal transport & ML, 2 papers.

Recent International Journals (last 4 years).

R. Bonalli, A. Rudi. "Non-Parametric Learning of Stochastic Differential Equations with Fast Rates of Convergence". *Foundations of Computational Mathematics (submitted) 2024*

A. Rudi, U. Marteau-Ferey, F. Bach. "Finding Global Minima via Kernel Approximations".
Mathematical Programming (to appear) 2024

P. Aubin, A. Rudi. "Approximation of optimization problems with constraints through kernel Sum-Of-Squares".
SIAM Journal on Optimization (to appear) 2024

U. Marteau-Ferey, F. Bach, A. Rudi. "Second order conditions to decompose smooth functions as sums of squares". *SIAM Journal on Optimization, Vol. 34, Iss. 1, 2024*

F. Bach, A. Rudi. "Exponential convergence of sum-of-squares hierarchies for trigonometric polynomials".
SIAM Journal on Optimization, Vol. 33, Iss. 3, 2023

L. Brogat-Motte, A. Rudi, C. Brouard, J. Rousu, F. d'Alché-Buc. "Vector-valued least-squares regression under output regularity assumptions". *Journal of Machine Learning Reasearch, Vol 23, Iss. 1, 2022*

X. Lu, A. Rudi, E. Borgonovo, L. Rosasco. "Faster Kriging: Facing High-Dimensional Simulators".
Operations Research, Vol. 68, Iss. 1, 2020

J. Lin, A. Rudi, L. Rosasco, V. Cevher. "Optimal Rates for Spectral-regularized Algorithms with Least-Squares Regression over Hilbert Spaces".
Applied and Computational Harmonic Analysis, Vol. 48, Iss. 3, 2020

C. Ciliberto, L. Rosasco, A. Rudi."A General Framework for Consistent Structured Prediction with Implicit Loss Embeddings". *Journal of Machine Learning Research, Vol. 21, Iss. 1, 2020*

Selected invited presentations (last 2 years).

(excluding contributed conference presentations)

- StatML Summer School in Causality and Statistical Learning, Oxford, July 2023
- 7th London Symposium on Information Theory, University College of London, May 2023
- PhD school on PSD Models, Scuola Normale Superiore, Italy, Jan 2023
- Optimization and Statistical Learning Workshop, Les Houches, France, Jan 2023
- Bocconi University, Milano, Italy, Sept 2022
- Czech-French Workshop in AI, Prague, Sept 2022
- ELLIS Theory Workshop, Arenzano, Italy, June 2022
- Colloquium PRAIRIE, Paris, May 2022
- Séminaire Parisien de Statistique — Institut Henri Poincaré, Paris, April 2022
- Workshop MAS-MODE, SMAI, Paris, March 2022
- Online Seminars of IA and Mathematics by Centro Nazionale di Ricerche, March 2022
- Séminaire Parisien d'Optimisation — Institut Henri Poincaré, Paris, March 2022
- S-DCE Seminar Series - Alan Turing Institute, London, March 2022

- Unione Matematica Italiana, Colloquium Prisma, Feb 2022

Reviewing and area chair activities.

Reviewer for international journals: Journal of Machine Learning research, Machine Learning, Applied and Computational Harmonic Analysis, Annals of Statistics, Inverse Problems, Electronic Journal of Statistics. Area chair in international machine learning conferences: COLT, ICML, NeurIPS.

Publications

I have published more than **70 peer-reviewed papers** (67 without my PhD advisors) in international journals and major peer-reviewed conferences. AI & Machine Learning are led mainly by conferences, where the most important is, by far, NeurIPS and then COLT, ICML, AISTATS and some others. Such conferences have low acceptance rate (usually less than 20%) and publications in their proceedings must be high quality and are considered as being as important as journal publications (e.g. Journal of Machine Learning Research). My publications distribute as follows **NIPS: 24 papers, COLT: 9 papers, ICML: 6 papers, AISTATS: 7 papers**. Overall I have more than **2650 citations** and an **h-index of 24** (both obtained from Google Scholar: scholar.google.com/citations?user=EL-7KFsAAAAJ).

[1] T. Cantelobre, C. Ciliberto, B. Guedj, and A. Rudi, "Closed-form filtering for non-linear systems," *COLT*, submitted 2024.

[2] A. Rudi, U. Marteau-Ferey, and F. R. Bach, "Finding global minima via kernel approximations," *Mathematical Programming*, to appear 2024.

[3] R. Bonalli and A. Rudi, "Non-parametric learning of stochastic differential equations with fast rates of convergence," *Foundations of Computational Mathematics*, to appear 2024.

[4] B. Muzellec, A. Vacher, F. R. Bach, F. Vialard, and A. Rudi, "Near-optimal estimation of smooth transport maps with kernel sums-of-squares," *SIAM Journal on Mathematics of Data Science*, to appear 2024.

[5] P. Aubin-Frankowski and A. Rudi, "Approximation of optimization problems with constraints through kernel sum-of-squares," *SIAM Journal on Optimization*, to appear 2024.

[6] U. Marteau-Ferey, F. R. Bach, and A. Rudi, "Second order conditions to decompose smooth functions as sums of squares," *SIAM Journal on Optimization*, vol. 34, no. 1, pp. 616–641, 2024.

[7] F. R. Bach and A. Rudi, "Exponential convergence of sum-of-squares hierarchies for trigonometric polynomials," *SIAM Journal on Optimization*, vol. 33, no. 3, pp. 2137–2159, 2023.

[8] A. Raj, U. Simsekli, and A. Rudi, "Efficient sampling of stochastic differential equations with positive semi-definite models," *NeurIPS*, vol. 36, 2023.

[9] G. Beugnot, J. Mairal, and A. Rudi, "Gloptinets: Scalable non-convex optimization with certificates," *NeurIPS*, vol. 36, 2023.

[10] L. Brogat-Motte, A. Rudi, C. Brouard, J. Rousu, and F. d'Alché-Buc, "Vector-valued least-squares regression under output regularity assumptions," *Journal of Machine Learning Research*, vol. 23, pp. 344:1–344:50, 2022.

[11] U. Marteau-Ferey, F. R. Bach, and A. Rudi, "Sampling from arbitrary functions via PSD models," in *AISTATS*, vol. 151 of *Proceedings of Machine Learning Research*, pp. 2823–2861, PMLR, 2022.

[12] A. Nowak, A. Rudi, and F. R. Bach, "On the consistency of max-margin losses," in *AISTATS*, vol. 151 of *Proceedings of Machine Learning Research*, pp. 4612–4633, PMLR, 2022.

[13] E. Berthier, J. Carpentier, A. Rudi, and F. R. Bach, "Infinite-dimensional sums-of-squares for optimal control," in *CDC*, pp. 577–582, IEEE, 2022.

[14] G. Beugnot, J. Mairal, and A. Rudi, "On the benefits of large learning rates for kernel methods," in *COLT*, vol. 178 of *Proceedings of Machine Learning Research*, pp. 254–282, PMLR, 2022.

[15] B. E. Woodworth, F. R. Bach, and A. Rudi, "Non-convex optimization with certificates and fast rates through kernel sums of squares," in *COLT*, vol. 178 of *Proceedings of Machine Learning Research*, pp. 4620–4642, PMLR, 2022.

[16] T. Cantelobre, C. Ciliberto, B. Guedj, and A. Rudi, "Measuring dissimilarity with diffeomorphism invariance," in *ICML*, vol. 162 of *Proceedings of Machine Learning Research*, pp. 2572–2596, PMLR, 2022.

[17] A. Chatalic, N. Schreuder, L. Rosasco, and A. Rudi, "Nyström kernel mean embeddings," in *ICML*, vol. 162 of *Proceedings of Machine Learning Research*, pp. 3006–3024, PMLR, 2022.

[18] V. Cabannes, F. R. Bach, V. Perchet, and A. Rudi, "Active labeling: Streaming stochastic gradients," in *NeurIPS*, 2022.

[19] L. Brogat-Motte, A. Rudi, C. Brouard, J. Rousu, and F. d'Alché-Buc, "Vector-valued least-squares regression under output regularity assumptions," *Arxiv*, vol. 2211.08958, 2022.

[20] V. A. Cabannes, F. R. Bach, and A. Rudi, "Fast rates for structured prediction," in *COLT*, vol. 134 of *Proceedings of Machine Learning Research*, pp. 823–865, PMLR, 2021.

[21] A. Vacher, B. Muzellec, A. Rudi, F. R. Bach, and F. Vialard, "A dimension-free computational upper-bound for smooth optimal transport estimation," in *COLT*, vol. 134 of *Proceedings of Machine Learning Research*, pp. 4143–4173, PMLR, 2021.

[22] V. A. Cabannes, F. R. Bach, and A. Rudi, "Disambiguation of weak supervision leading to exponential convergence rates," in *ICML*, vol. 139 of *Proceedings of Machine Learning Research*, pp. 1147–1157, PMLR, 2021.

[23] A. Rudi and C. Ciliberto, "PSD representations for effective probability models," in *NeurIPS*, pp. 19411–19422, 2021.

[24] R. Jézéquel, P. Gaillard, and A. Rudi, "Mixability made efficient: Fast online multiclass logistic regression," in *NeurIPS*, pp. 23692–23702, 2021.

[25] G. Beugnot, J. Mairal, and A. Rudi, "Beyond tikhonov: faster learning with self-concordant losses, via iterative regularization," in *NeurIPS*, pp. 28196–28207, 2021.

[26] V. Cabannes, L. Pillaud-Vivien, F. R. Bach, and A. Rudi, "Overcoming the curse of dimensionality with laplacian regularization in semi-supervised learning," in *NeurIPS*, pp. 30439–30451, 2021.

[27] V. Cabannes, A. Rudi, and F. R. Bach, "Fast rates in structured prediction," *COLT*, 2021.

[28] O. Zadorozhnyi, P. Gaillard, S. Gerchinovitz, and A. Rudi, "Online nonparametric regression with sobolev kernels," *Arxiv*, vol. 2102.03594, 2021.

[29] A. Nowak-Vila, A. Rudi, and F. R. Bach, "Max-margin is dead, long live max-margin!," *ICML*, 2021.

[30] B. Muzellec, F. R. Bach, and A. Rudi, "A note on optimizing distributions using kernel mean embeddings," *Arxiv*, vol. 2106.09994, 2021.

[31] B. Muzellec, F. R. Bach, and A. Rudi, "Learning psd-valued functions using kernel sums-of-squares," *Arxiv*, vol. 2111.11306, 2021.

[32] X. Lu, A. Rudi, E. Borgonovo, and L. Rosasco, "Faster kriging: Facing high-dimensional simulators," *Operations Research*, vol. 68, no. 1, pp. 233–249, 2020.

[33] C. Ciliberto, L. Rosasco, and A. Rudi, "A general framework for consistent structured prediction with implicit loss embeddings," *Journal of Machine Learning Research*, vol. 21, pp. 98:1–98:67, 2020.

[34] A. Rudi, L. Wossnig, C. Ciliberto, A. Rocchetto, M. Pontil, and S. Severini, "Approximating hamiltonian dynamics with the nyström method," *Quantum*, vol. 4, p. 234, 2020.

[35] L. Pillaud-Vivien, F. R. Bach, T. Lelièvre, A. Rudi, and G. Stoltz, "Statistical estimation of the poincaré constant and application to sampling multimodal distributions," in *AISTATS*, vol. 108 of *Proceedings of Machine Learning Research*, pp. 2753–2763, PMLR, 2020.

[36] N. Sterge, B. K. Sriperumbudur, L. Rosasco, and A. Rudi, "Gain with no pain: Efficiency of kernel-pca by nyström sampling," in *AISTATS*, vol. 108 of *Proceedings of Machine Learning Research*, pp. 3642–3652, PMLR, 2020.

[37] R. Jézéquel, P. Gaillard, and A. Rudi, "Efficient improper learning for online logistic regression," in *COLT*, vol. 125 of *Proceedings of Machine Learning Research*, pp. 2085–2108, PMLR, 2020.

[38] V. Cabannes, A. Rudi, and F. R. Bach, "Structured prediction with partial labelling through the infimum loss," in *ICML*, vol. 119 of *Proceedings of Machine Learning Research*, pp. 1230–1239, PMLR, 2020.

[39] A. Nowak, F. R. Bach, and A. Rudi, "Consistent structured prediction with max-min margin markov networks," in *ICML*, vol. 119 of *Proceedings of Machine Learning Research*, pp. 7381–7391, PMLR, 2020.

[40] U. Marteau-Ferey, F. R. Bach, and A. Rudi, "Non-parametric models for non-negative functions," in *NeurIPS*, 2020.

[41] G. Meanti, L. Carratino, L. Rosasco, and A. Rudi, "Kernel methods through the roof: Handling billions of points efficiently," in *NeurIPS*, 2020.

[42] C. Ciliberto, A. Rocchetto, A. Rudi, and L. Wossnig, "Fast quantum learning with statistical guarantees," *Arxiv*, vol. 2001.10477, 2020.

[43] T. Eboli, A. Nowak-Vila, J. Sun, F. R. Bach, J. Ponce, and A. Rudi, "Structured and localized image restoration," *Arxiv*, vol. 2006.09261, 2020.

[44] L. Brogat-Motte, A. Rudi, C. Brouard, J. Rousu, and F. d'Alché-Buc, "Learning output embeddings in structured prediction," *Arxiv*, vol. 2007.14703, 2020.

[45] A. Nowak-Vila, F. R. Bach, and A. Rudi, "Sharp analysis of learning with discrete losses," in *AISTATS*, vol. 89 of *Proceedings of Machine Learning Research*, pp. 1920–1929, PMLR, 2019.

[46] U. Marteau-Ferey, D. Ostrovskii, F. R. Bach, and A. Rudi, "Beyond least-squares: Fast rates for regularized empirical risk minimization through self-concordance," in *COLT*, vol. 99 of *Proceedings of Machine Learning Research*, pp. 2294–2340, PMLR, 2019.

[47] D. M. Ostrovskii and A. Rudi, "Affine invariant covariance estimation for heavy-tailed distributions," in *COLT*, vol. 99 of *Proceedings of Machine Learning Research*, pp. 2531–2550, PMLR, 2019.

[48] J. M. Altschuler, F. R. Bach, A. Rudi, and J. Niles-Weed, "Massively scalable sinkhorn distances via the nyström method," in *NeurIPS*, pp. 4429–4439, 2019.

[49] C. Ciliberto, F. R. Bach, and A. Rudi, "Localized structured prediction," in *NeurIPS*, pp. 7299–7309, 2019.

[50] U. Marteau-Ferey, F. R. Bach, and A. Rudi, "Globally convergent newton methods for ill-conditioned generalized self-concordant losses," in *NeurIPS*, pp. 7634–7644, 2019.

[51] R. Jézéquel, P. Gaillard, and A. Rudi, "Efficient online learning with kernels for adversarial large scale problems," in *NeurIPS*, pp. 9427–9436, 2019.

[52] A. Nowak-Vila, F. R. Bach, and A. Rudi, "A general theory for structured prediction with smooth convex surrogates," *Arxiv*, vol. 1902.01958, 2019.

[53] N. Sterge, B. K. Sriperumbudur, L. Rosasco, and A. Rudi, "Gain with no pain: Efficient kernel-pca by nyström sampling," *AISTATS*, 2020.

[54] L. Pillaud-Vivien, A. Rudi, and F. R. Bach, "Exponential convergence of testing error for stochastic gradient methods," in *COLT*, vol. 75 of *Proceedings of Machine Learning Research*, pp. 250–296, PMLR, 2018.

[55] A. Rudi, C. Ciliberto, G. M. Marconi, and L. Rosasco, "Manifold structured prediction," in *NeurIPS*, pp. 5615–5626, 2018.

[56] A. Rudi, D. Calandriello, L. Carratino, and L. Rosasco, "On fast leverage score sampling and optimal learning," in *NeurIPS*, pp. 5677–5687, 2018.

[57] G. Luise, A. Rudi, M. Pontil, and C. Ciliberto, "Differential properties of sinkhorn approximation for learning with wasserstein distance," in *NeurIPS*, pp. 5864–5874, 2018.

[58] L. Pillaud-Vivien, A. Rudi, and F. R. Bach, "Statistical optimality of stochastic gradient descent on hard learning problems through multiple passes," in *NeurIPS*, pp. 8125–8135, 2018.

[59] L. Carratino, A. Rudi, and L. Rosasco, "Learning with SGD and random features," in *NeurIPS*, pp. 10213–10224, 2018.

[60] J. M. Altschuler, F. R. Bach, A. Rudi, and J. Weed, "Approximating the quadratic transportation metric in near-linear time," *Arxiv*, vol. 1810.10046, 2018.

[61] C. Ciliberto, A. Rudi, L. Rosasco, and M. Pontil, "Consistent multitask learning with nonlinear output relations," in *NIPS*, pp. 1986–1996, 2017.

[62] A. Rudi and L. Rosasco, "Generalization properties of learning with random features," in *NIPS*, pp. 3215–3225, 2017.

[63] A. Rudi, L. Carratino, and L. Rosasco, "FALKON: an optimal large scale kernel method," in *NIPS*, pp. 3888–3898, 2017.

[64] R. Camoriano, T. Angles, A. Rudi, and L. Rosasco, "NYTRO: when subsampling meets early stopping," in *AISTATS*, vol. 51 of *JMLR Workshop and Conference Proceedings*, pp. 1403–1411, JMLR.org, 2016.

[65] C. Ciliberto, L. Rosasco, and A. Rudi, "A consistent regularization approach for structured prediction," in *NIPS*, pp. 4412–4420, 2016.

[66] A. Rudi, R. Camoriano, and L. Rosasco, "Less is more: Nyström computational regularization," in *NIPS*, pp. 1657–1665, 2015.

[67] A. Rudi, F. Odone, and E. D. Vito, "Geometrical and computational aspects of spectral support estimation for novelty detection," *Pattern Recognition Letters*, vol. 36, pp. 107–116, 2014.

[68] A. Rudi, G. D. Cañas, and L. Rosasco, "On the sample complexity of subspace learning," in *NIPS*, pp. 2067–2075, 2013.

[69] A. Rudi, G. Chiusano, and A. Verri, "Adaptive optimization for cross validation," in *ESANN*, 2012.

[70] F. Pirri, M. Pizzoli, and A. Rudi, "A general method for the point of regard estimation in 3d space," in *CVPR*, pp. 921–928, IEEE Computer Society, 2011.

[71] A. Rudi, M. Pizzoli, and F. Pirri, "Linear solvability in the viewing graph," in *ACCV*, vol. 6494 of *Lecture Notes in Computer Science*, pp. 369–381, Springer, 2010.