

**CURRICULUM VITAE**

Work Address: Raymond and Beverly Sackler School of  
Mathematical Sciences,  
Tel Aviv University,  
Tel Aviv 69978, Israel  
E-mail: saharon@tauex.tau.ac.il  
Home page: <http://www.tau.ac.il/~saharon>

**EDUCATION**

1992–1994 B.Sc., Mathematics, Tel Aviv University (Summa Cum Laude)  
1995–1999 MSc., Statistics, Tel Aviv University (Summa Cum Laude)  
1999–2003 Ph.D., Statistics, Stanford University, Stanford, CA

**ACADEMIC AND PROFESSIONAL EXPERIENCE**

1993–1997 Mathematical programmer, Ron Soferman Image Processing Ltd., Israel.  
1997–2003 Data Mining researcher/consultant, Amdocs Ltd (USA & Israel).  
1999–2003 Research & teaching assistant, Statistics Dept., Stanford University, CA.  
2003–2007 Research Staff Member, IBM T.J. Watson Research Center, NY.  
2007–2010 Senior Lecturer, Dept. of Statistics and OR, Tel Aviv University.  
2010–2016 Associate Professor, Dept. of Statistics and OR, Tel Aviv University.  
2016– Professor, Dept. of Statistics and OR, Tel Aviv University.  
2023– Chair in Modern Statistics and Data Science, Tel Aviv University.

**ACADEMIC AND PROFESSIONAL AWARDS AND ACCOMPLISHMENTS**

1999 Nimrod Lapid & Arik Peretz prizes for MSc studies.  
1999–2008 Five time winner of major data mining competitions (KDD-Cup, INFORMS challenge)  
2002 Primary author of best application paper award winner at KDD-2002.  
2006–2007 Two IBM Research Outstanding Technical Achievement Awards.  
2006–2018 Action Editor for Journal of Machine Learning Research.  
2011 Sackler chair for advancement of young faculty at TAU.  
2011 Senior author of best research paper award winner at KDD-2011.  
2011–2016 Senior program committee member for the 17th-22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining.

- 2018 Paper “From Fixed-X to Random-X Regression: Bias-Variance Decompositions, Covariance Penalties, and Prediction Error Estimation” with Ryan Tibshirani selected for discussion in Journal of the American Statistical Association.
- 2019– Associate Editor for the Annals of Statistics.
- 2019– Academic Leadership team of TAU AI and Data Science center (TAD).
- 2021 Presentation at the Royal Statistical Society webinar on multiple testing, Dec. 2021.
- 2022– Co-head of TAU Data Science degree program.

## RESEARCH GRANTS

- 2005–2008 NSF grant DMS-0505432 for \$115,597 (co-PI with Ji Zhu of Univ. of Michigan, not eligible for disbursements).
- 2007–2011 EU FP7 International Reintegration Grant for 100,000 Euros.
- 2008–2011 Alon Fellowship for young faculty.
- 2009–2012 ISF grant 1227/09 for \$95,000.
- 2010–2012 Two grants for functional genetics research from IBM Research Israel for \$132,000 (co-PI with Ron Shamir, Eran Halperin, equal parts).
- 2012–2016 ISF grant 1487/12 for \$175,000.
- 2014 Microsoft UNext faculty award for \$15,000.
- 2015 Microsoft Azure Research award for \$20,000.
- 2016–2020 ISF grant 1804/16 for \$155,000.
- 2020–2024 ISF grant 2180/20 for \$310,000 (co-PI with Ruth Heller).
- 2022–2024 Edmond J. Safra Center-Tel Aviv Sourasky Medical Center Clinical Bioinformatics Research Grant Grant for \$40,000 (co-PI with four others).
- 2019– Co-PI of multiple grants under the TAU AI and Data Science center, totalling over \$10M.

## STUDENT SUPERVISION

**Postdoctoral researchers:** Ronny Luss (2009-2011); Amit Moscovich-Eiger (2017-2018)

**Graduated PhD students:** David Golan (2014); Shlomi Lifshits (2015); Amichai Painsky (2016); Omer Weissbrod (2016) ; Assaf Rabinowicz (2021) ; Aya Vituri (2023, expected)

**Current PhD students:** Keren Levinstein-Hallak; Oren Yuval; Giora Simchoni; Rajesh Karmakar

**20 Graduated Master’s students**

**LIST OF PUBLICATIONS**

Google Scholar statistics: 16881 citations, h-index 51, i-10 index 104

**Journal Papers and Discussions**

1. Saharon Rosset, Einat Neumann, Uri Eick, Nurit Vatnik  
*Customer Lifetime Value Models for Decision Support*  
Data Mining and Knowledge Discovery Journal, Vol. 7 (2003), 321–339.
2. Saharon Rosset, Ji Zhu, Trevor Hastie  
*Boosting as a Regularized Path to a Maximum Margin Classifier*  
Journal of Machine Learning Research, 5:Aug. (2004), 941–973.
3. Saharon Rosset, Ji Zhu  
*Discussion of “Least Angle Regression” by Efron et al.*  
Annals of Statistics, Apr. 2004.
4. Saharon Rosset, Ji Zhu  
*A Corrected Proof of the Result of “A Prediction Error Property of the Lasso” by Huang(2003)*  
Australia and New Zealand Journal of Statistics, 46:3 (2004), 505–510.
5. Jerome Friedman, Trevor Hastie, Saharon Rosset, Rob Tibshirani, Ji Zhu  
*Discussion of Three Boosting Papers*  
Annals of Statistics, Feb. 2004.
6. Trevor Hastie, Saharon Rosset, Rob Tibshirani, Ji Zhu  
*The Entire Regularization Path of the Support Vector Machine*  
Journal of Machine Learning Research, 5:Oct. (2004), 1391–1415.
7. Robert Tibshirani, Michael Saunders, Saharon Rosset, Ji Zhu, Keith Knight  
*Sparsity and Smoothness via the Fused Lasso*  
Journal of the Royal Statistical Society, Series B, Vol. 67 (2005), No. 1.
8. Saharon Rosset, Ji Zhu  
*Piecewise Linear Regularized Solution Paths*  
Annals of Statistics, Vol. 35:3. (2007), 1012–1030.
9. Saharon Rosset, Claudia Perlich, Bianca Zadrozny  
*Ranking-Based Evaluation of Regression Models*  
Knowledge and Information Systems, 12:3 (2007), 331-353.

10. Saharon Rosset, Claudia Perlich, Yan Liu  
*Making the Most of Your Data: KDD Cup 2007 "How Many Ratings" Winner's Report*  
SIGKDD Explorations, vol. 9, issue 2, 66-69, 2007.
11. Richard Lawrence, Claudia Perlich, Saharon Rosset, et al. (12 authors)  
*Analytics-driven solutions for customer targeting and sales-force allocation*  
IBM Systems Journal, vol. 46, issue 4, 797-816, 2007.
12. Saharon Rosset  
*Efficient Inference on Known Phylogenetic Trees Using Poisson Regression*  
Bioinformatics 22, Suppl 2, 2007 (Proceedings of ECCB-2006).
13. Doron Behar, Saharon Rosset, et al. (ten authors)  
*The Genographic Project Public Participation Mitochondrial DNA Database*  
PLoS Genetics, 3:6 (2007), e104.
14. Claudia Perlich, Prem Melville, Yan Liu, Grzegorz Swirszcz, Richard Lawrence, Saharon Rosset  
*Breast Cancer Identification: KDD CUP Winner's Report*  
SIGKDD Explorations, vol. 10, issue 2, 39-42, 2008.
15. Doron Behar, Ene Metsaplu, Toomas Kivisild, Saharon Rosset, et al. (15 authors)  
*Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora*  
PLoS ONE 3(4): e2062, 2008.
16. Doron Behar, Richard Villems, – 11 other authors – and Saharon Rosset  
*The dawn of human matrilineal diversity: Insights from complete mitochondrial genomes*  
American Journal of Human Genetics, 82(5) : 1130-1140, 2008.
17. Saharon Rosset, Spencer Wells, David Soria-Hernanz, Chris Tyler-Smith, Ajay Royyuru and Doron Behar  
*Maximum Likelihood Estimation of Site-Specific Mutation Rates in Human Mitochondrial DNA from Partial Phylogenetic Classification*  
Genetics 180, 1511- 1524, 2008.
18. Aurelie Lozano, Naoki Abe, Yan Liu, Saharon Rosset  
*Grouped Graphical Granger Modeling for Gene Expression Regulatory Networks Discovery*  
Bioinformatics, 25(12):i110-i118 (Proceedings of ISMB 2009).
19. Ji Zhu, Hui Zou, Saharon Rosset, Trevor Hastie  
*Multi-class AdaBoost*  
Statistics and its Interface, volume 2, issue 3, 2009.

20. Michael F. Hammer, Doron M. Behar, – 5 other authors –, Saharon Rosset, Karl Skorecki  
*Extended Y chromosome haplotypes resolve multiple and unique lineages of the Jewish priesthood*  
Human Genetics, Volume 126, Number 5 / November, 2009.
21. Saharon Rosset  
*Bi-Level Path Following for Cross Validated Solution of Kernel Quantile Regression*  
Journal of Machine Learning Research, 10(Nov):2473-2505, 2009.
22. Richard Lawrence, Claudia Perlich, Saharon Rosset, et al. (10 authors)  
*Operations Research Improves Sales Force Productivity at IBM*  
INFORMS Interfaces, Vol. 40, No. 1, pp. 33-46. 2010.
23. Saharon Rosset, Claudia Perlich, Grzegorz Swirszcz, Yan Liu, Prem Melville  
*Medical Data Mining: Lessons from Winning Two Competitions*  
Data Mining and Knowledge Discovery Journal (DAMI), Volume 20, Number 3 / May, 2010.
24. Doron M. Behar<sup>a</sup>, Saharon Rosset<sup>a</sup>, Shay Tzur<sup>a</sup>, et al. (11 authors)  
*African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney disease in Hispanic Americans*  
Human Molecular Genetics, Vol. 19, No. 9, pp. 1816 -1827, 2010.  
(<sup>a</sup> equal contribution)
25. Doron M. Behar, Bayazit Yunusbayev, Mait Metspalu, Ene Metspalu, Saharon Rosset, et al. (21 authors)  
*The Genome Wide Structure of the Jewish People*  
Nature, Vol. 466 238– 242. doi:10.1038/nature09103, 2010.
26. Osnat Ravit-Amir, Saharon Rosset  
*Maximum Likelihood Estimation of Locus-Specific Mutation Rates in Y-chromosome Short Tandem Repeats*  
Bioinformatics (proceedings of ECCB10) 26(18): i440–i445, 2010.
27. Shay Tzur<sup>a</sup>, Saharon Rosset<sup>a</sup>, et al. (11 authors)  
*Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene*  
Human Genetics, Vol. 128, Issue 3 345 350, 2010.  
(<sup>a</sup>equal contribution)
28. Sijian Wang, Bin Nan, Saharon Rosset, Ji Zhu  
*Random Lasso*  
Annals of Applied Statistics, 5(1), 468–485, 2011.

29. Ehud Aharoni, Hani Neuvirth, Saharon Rosset  
*The Quality Preserving Database: A Computational Framework for Encouraging Collaboration, Enhancing Power and Controlling False Discovery*  
IEEE Transactions on Computational Biology and Bioinformatics, Vol.8, Issue 5, 1431-1437, 2011.
30. Doron M. Behar, Einat Kedem, Saharon Rosset, et al. (14 authors)  
*Absence of APOL1 risk variants protects against HIV-associated nephropathy in the Ethiopian population*  
American Journal of Nephrology, Vol. 34, 452–459, 2011.
31. Giles Hooker, Saharon Rosset  
*Prediction-Based Regularization Using Data Augmented Regression*  
Statistics and Computing, 22(1), 237–249, 2012.
32. David Golan, Saharon Rosset  
*Accurate Estimation of Heritability in Genome Wide Studies using Random Effects Models*  
Bioinformatics (Proceedings of ISMB-ECCB11), 27(13), i317–i323, 2011.
33. Ronny Luss, Saharon Rosset, Moni Shahar  
*Efficient Regularized Isotonic Regression with Application to Gene-Gene Interaction Search*  
Annals of Applied Statistics, 6(1), 253–283, 2012.  
*Correction:* Annals of Applied Statistics, 9(4), 2266–2267, 2015.
34. Melyssa Gymrek, David Golan, Saharon Rosset, Yaniv Erlich  
*lobSTR: A short tandem repeat profiler for personal genomes*  
Genome Research, 22(6) : 1154–62, June 2012.
35. Doron M. Behar, Mannis van Oven, Saharon Rosset, et al. (9 authors)  
*A Copernican Reassessment of the Human Mitochondrial DNA Tree from its Root*  
The American Journal of Human Genetics, 90(4), 675–684, 6 April 2012.
36. David Golan, Yaniv Erlich, Saharon Rosset  
*Weighted Pooling - Practical and Cost Effective Techniques for Pooled High Throughput Sequencing*  
Bioinformatics (Proceedings ISMB 2012), 28(12), i197- i206, 2012.
37. David Golan, Saharon Rosset  
*Comment on “The Predictive Capacity of Personal Genome Sequencing”*  
Science Translational Medicine 4, 135le4, 2012.

38. Shachar Kaufman, Saharon Rosset, Claudia Perlich, Ori Stitleman  
*Leakage in Data Mining: Formulation, Detection, and Avoidance*  
ACM Transactions on Knowledge Discovery from Data, Vol. 6 No. 4, December 2012.
39. Shay Tzur, Saharon Rosset, Walter Wasser, Doron Behar, Karl Skorecki  
*APOL1 allelic variants are associated with lower age of dialysis initiation and thereby increased dialysis vintage in African and Hispanic Americans with non-diabetic end-stage kidney disease*  
Nephrology, Dialysis and Transplantation, 27(4):1498–505, 2012.
40. Shay Tzur, Walter G Wasser, Saharon Rosset, Karl Skorecki  
*Linkage disequilibrium analysis reveals an albuminuria risk haplotype containing three missense mutations in the cubilin gene with striking differences among European and African ancestry populations*  
BMC Nephrology, 13:142, 2012.
41. Gilad Yahalom, Natalie Kaplan, Aya Vituri, Oren S. Cohen, Rivka Inzelberg, Evgenia Kozlova, Amos D. Korczyn, Saharon Rosset, Eitan Friedman, Sharon Hassin-Baer  
*Dyskinesias in patients with Parkinson's disease: Effect of the leucine-rich repeat kinase 2 (LRRK2) G2019S mutation*  
Parkinsonism and Related Disorders, Vol. 18, Issue 9, November 2012, Pages 1039–1041.
42. Ofer Sarig, Sivan Bercovici, Lilach Zoller, – 20 other authors –, Saharon Rosset, Saleh M. Ibrahim, Eli Sprecher  
*Population-Specific Association between a Polymorphic Variant in ST18, Encoding a Pro-Apoptotic Molecule, and Pemphigus Vulgaris*  
Journal of Investigative Dermatology, Vol. 132, Issue 7, Pages 1798–1805, 2012.
43. Siiri Rootsi, Doron M. Behar, Mari Järve, – 7 other authors –, Saharon Rosset, et al. (20 authors)  
*Phylogenetic applications of whole Y-chromosome sequences and the Near Eastern origin of Ashkenazi Levites*  
Nature Communications 4, 2013.
44. Ifeoma I Ulasi, Shay Tzur, Walter G Wasser, – 13 other authors –, Saharon Rosset, Karl Skorecki  
*High population frequencies of APOL1 risk variants are associated with increased prevalence of non-diabetic chronic kidney disease in the Igbo people from south-eastern Nigeria*  
Nephron Clinical practice, Vol. 123, Issue 1-2, 2013.

45. Doron M. Behar, Mait Metspalu, Yael Baran,, Naama M. Kopelman, – 21 additional authors – , Saharon Rosset, Eran Halperin, Richard Villems, and Noah A. Rosenberg *No Evidence from Genome-Wide Data of a Khazar Origin for the Ashkenazi Jews* Human Biology 85(6):859-900. 2013.
46. Shachar Kaufman, Saharon Rosset  
*Exploiting Population Samples to Enhance Genome Wide Association Studies of Disease*  
Genetics, Vol. 197, 337- 349, May 2014.
47. Ronny Luss, Saharon Rosset  
*Generalized Isotonic Regression*  
Journal of Computational and Graphical Statistics, Vol. 23, No. 1, pp 192–210, 2014.  
*Correction:* Journal of Computational and Graphical Statistics, Vol. 24, No. 4, p 1230, 2015.
48. Amichai Painsky, Saharon Rosset  
*Optimal Set Cover Formulation for Exclusive Row Biclustering of Gene Expression*  
Journal of Computer Science and Technology (JCST), 29(3): 423 -435, May 2014.
49. Ehud Aharoni, Saharon Rosset  
*Generalized Alpha Investing: Definitions, Optimality Results, and Application to Public Databases*  
Journal of the Royal Statistical Society, Series B, Vol. 76, Issue 4, 771- 794, September 2014.
50. Saharon Rosset, Ehud Aharoni, Hani Neuvirth  
*Novel Statistical Tools for Management of Public Databases Facilitate Community-Wide Replicability and Control of False Discovery*  
Genetic Epidemiology, Vol. 38, Issue 5, 477- 481, July 2014.
51. Shachar Kaufman, Saharon Rosset  
*When Does More Regularization Imply Fewer Degrees of Freedom? Sufficient Conditions and Counterexamples*  
Biometrika, 101(4), 771–784, December 2014.
52. David Golan, Saharon Rosset  
*Effective Genetic Risk Prediction Using Mixed Models*  
American Journal of Human Genetics, Vol. 95, Issue 4, 383 -393, 2 October 2014.
53. David Golan, Eric Lander, Saharon Rosset  
*Measuring Missing Heritability: Inferring the Contribution of Common Variants*



- Proceedings of the National Academy of Sciences, 111(49), E5272 -E5281, December 2014.
54. Natalie Kaplan, Aya Vituri, Amos D. Korczyn, – 7 other authors – , Saharon Rosset, Sharon Hassin-Baer  
*Sequence variants in SLC6A3, DRD2, and BDNF genes and time to levodopa-induced dyskinesias in Parkinson s disease*  
Journal of Molecular Neuroscience Vol. 53, No. 2, pp 183–188, 2014.
  55. Amichai Painsky, Saharon Rosset  
*Isotonic Modeling with Non-Differentiable Loss Functions, with Application to Lasso Regularization*  
IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 32, No. 2, Feb. 2016.
  56. Amichai Painsky, Saharon Rosset, Meir Feder  
*Generalized Independent Component Analysis Over Finite Alphabets*  
IEEE Transactions on Information Theory, Vol. 62, No. 2, Feb. 2016.
  57. Regev Schweiger, Shachar Kaufman, Reijo Laaksonen, – 3 other authors – , Saharon Rosset, Eran Halperin  
*Fast and Accurate Construction of Confidence Intervals for Heritability*  
American Journal of Human Genetics, Vol. 98, No. 6, 1181- 1192, 2 June 2016.
  58. Omer Weissbrod, Dan Geiger, Saharon Rosset  
*Multikernel linear mixed models for complex phenotype prediction*  
Genome Research, 26 (7), 969 979, July 2016.
  59. Yaron Granot, Omri Tal, Saharon Rosset, Karl Skorecki  
*On the Apportionment of Population Structure*  
PLoS One, DOI:10.1371/journal.pone.0160413 August 9, 2016.
  60. Sagi Shporer, Benny Chor, Saharon Rosset, David Horn  
*Inversion symmetry of DNA k-mer counts: validity and deviations*  
BMC Genomics, 17:696, 2016.
  61. Shay Tzur, Saharon Rosset  
*Strictly conserved tri-nucleotide motif CAT is associated with TAS DNA protein binding sites in human mitochondrial DNA control region*  
Mitochondrial DNA, Vol. 28, No. 2, 250–253, 2017.
  62. Karl L. Skorecki, Jessica H. Lee, Carl D. Langefeld, Saharon Rosset, et al. (21 authors)

- A null variant in the apolipoprotein L3 gene is associated with non-diabetic nephropathy*  
Nephrology Dialysis Transplantation 33 (2), 323-330, 2017.
63. Ronny Luss, Saharon Rosset  
*Bounded Isotonic Regression*  
Electronic Journal of Statistics 11 (2), 4488-4514, 2017.
64. Omer Weissbrod, Elior Rahmani, Regev Schweiger, Saharon Rosset, Eran Halperin  
*Association testing of bisulfite-sequencing methylation data via a Laplace approximation*  
Bioinformatics 33 (14), i325-i332, 2017.
65. Amichai Painsky, Saharon Rosset, Meir Feder  
*Large Alphabet Source Coding Using Independent Component Analysis*  
IEEE Transactions on Information Theory 63 (10), 6514-6529, 2017.
66. Amichai Painsky, Saharon Rosset  
*Cross-validated variable selection in tree-based methods improves predictive performance*  
IEEE Trans. on Pattern Analysis and Machine Intelligence 39 (11), 2142-2153, 2017.
67. Shlomi Lifshits, Omri Tomer, Ittai Shamir, Daniel Barazany, Galia Tsarfaty, Saharon Rosset, Yaniv Assaf  
*Resolution considerations in imaging of the cortical layers*  
NeuroImage 164, 112-120, 2018
68. Omer Weissbrod, Jonathan Flint, Saharon Rosset  
*Estimating SNP-Based Heritability and Genetic Correlation in Case-Control Studies Directly and with Summary Statistics*  
The American Journal of Human Genetics 103 (1), 89-99, 2018
69. Regev Schweiger, Eyal Fisher, Elior Rahmani, Liat Shenhav, Saharon Rosset, Eran Halperin  
*Using stochastic approximation techniques to efficiently construct confidence intervals for heritability*  
Journal of Computational Biology 25 (7), 794-808, 2018.
70. Keren Levinstein-Hallak, Shay Tzur, Saharon Rosset  
*Big data analysis of human mitochondrial DNA substitution models: a regression approach* BMC genomics 19 (1), 759, 2018.
71. Regev Schweiger, Eyal Fisher, Omer Weissbrod, Elior Rahmani, Martina Mller-Nurasyid, Sonja Kunze, Christian Gieger, Melanie Waldenberger, Saharon Rosset,

- Eran Halperin  
*Detecting heritable phenotypes without a model using fast permutation testing for heritability and set-tests*  
Nature Communications 9 (1), 4919, 2018.
72. Amichai Painsky, Saharon Rosset, Meir Feder  
*Linear Independent Component Analysis Over Finite Fields: Algorithms and Bounds*  
IEEE Transactions on Signal Processing 66 (22), 5875-5886, 2018.
73. Ayala Neudorfer, Saharon Rosset  
*Predicting the NCAA basketball tournament using isotonic least squares pairwise comparison model*  
Journal of Quantitative Analysis in Sports, Volume 14, Issue 4, 2018.
74. Ryan Tibshirani, Saharon Rosset  
*Excess Optimism: How Biased is the Apparent Error of an Estimator Tuned by SURE?*  
Journal of the American Statistical Association 114 (526), 697-712, 2019.
75. Amichai Painsky, Saharon Rosset  
*Lossless compression of random forests*  
Journal of Computer Science and Technology 34 (2), 494-506, 2019.
76. Omer Weissbrod, Shachar Kaufman, David Golan, Saharon Rosset  
*Maximum Likelihood for Gaussian Process Classification and Generalized Linear Mixed Models under Case-Control Sampling*  
Journal of Machine Learning Research 20 (108), 1-30, 2019.
77. Elicor Rahmani, Regev Schweiger, Brooke Rhead, Lindsey A Criswell, Lisa F Barcellos, Eleazar Eskin, Saharon Rosset, Sriram Sankararaman, Eran Halperin  
*Cell-type-specific resolution epigenetics without the need for cell sorting or single-cell biology*  
Nature communications 10 (1), 1-11, 2019.
78. Amichai Painsky, Saharon Rosset, Meir Feder  
*Innovation representation of stochastic processes with application to causal inference*  
IEEE Transactions on Information Theory, 66 (2), 2020.
79. Eyal Fisher, Regev Schweiger, Saharon Rosset  
*Efficient Construction of Test Inversion Confidence Intervals Using Quantile Regression*  
Journal of Computational and Graphical Statistics 29 (1), 140-148, 2020.

- 
80. Saharon Rosset, Ryan Tibshirani  
*From fixed-X to random-X regression: Bias-variance decompositions, covariance penalties, and prediction error estimation*  
Journal of the American Statistical Association 115 (529), 138-151, 2020.  
(With Discussion and Rejoinder)
  81. Aviv Navon, Saharon Rosset  
*Capturing between-tasks covariance and similarities using multivariate linear mixed models*  
Electronic Journal of Statistics 14 (2) 3821 - 3844, 2020. <https://doi.org/10.1214/20-EJS1764>
  82. Assaf Rabinowicz, Saharon Rosset  
*Assessing prediction error at interpolation and extrapolation points*  
Electronic Journal of Statistics 14 (1) 272 - 301, 2020. <https://doi.org/10.1214/19-EJS1666>
  83. Shiran Abadi, Oren Avram, Saharon Rosset, Tal Pupko, Itay Mayrose  
*ModelTeller: Model Selection for Optimal Phylogenetic Reconstruction Using Machine Learning*  
Molecular Biology and Evolution, Volume 37, Issue 11, November 2020, Pages 3338-3352, <https://doi.org/10.1093/molbev/msaa154>
  84. Ruth Heller and Saharon Rosset  
*Optimal control of false discovery criteria in the two-group model*  
Journal of the Royal Statistical Society, Series B, 83: 133-155, 2021.
  85. Assaf Rabinowicz, Saharon Rosset  
*Cross-Validation for Correlated Data*  
Journal of the American Statistical Association, 117 (538), 718-731, 2022.
  86. Trevor Hastie, Andrea Montanari, Saharon Rosset, Ryan Tibshirani  
*Surprises in High-Dimensional Ridgeless Least Squares Interpolation*  
Annals of Statistics, 50(2): 949-986 (April 2022).
  87. Saharon Rosset, Ruth Heller, Amichai Painsky, Ehud Aharoni  
*Optimal and maximin procedures for multiple testing problems*  
Journal of the Royal Statistical Society, Series B, 84 (4), 1105-1128, 2022.
  88. Keren Levinstein-Hallak, Saharon Rosset  
*Modeling SARS-CoV-2 substitution processes: predicting the next variant*  
Communications Biology, (2022) 5:285.

89. Oren Yuval, Saharon Rosset  
*Semi-Supervised Empirical Risk Minimization: When can unlabeled data improve prediction*  
Electronic Journal of Statistics, 16(1): 1434-1460 (2022).
90. Amit Moscovich, Saharon Rosset  
*On the cross-validation bias due to unsupervised preprocessing*  
Journal of the Royal Statistical Society, Series B, 84 (4), 1474-1502, 2022.
91. Assaf Rabinowicz, Saharon Rosset  
*Tree-Based Models for Correlated Data*  
Journal of Machine Learning Research, 23 (258), 1-31, 2022.
92. Ruth Heller, Abba Krieger, Saharon Rosset  
*Optimal multiple testing and design in clinical trials*  
Biometrics, to appear (accepted 7/22).

## Reviews

1. Saharon Rosset<sup>a</sup>, Shay Tzur<sup>a</sup>, Walter Wasser, Doron Behar and Karl Skorecki  
*The population genetics of chronic kidney disease: insights from the MYH9- APOL1 locus*  
Nature Reviews Nephrology, 7, 313-326 (June 2011)  
(<sup>a</sup>equal contribution)

## Refereed Book Chapters

1. Saharon Rosset, Ji Zhu  
*Sparse, Flexible and Efficient Modeling using L1 Regularization*  
In: I. Guyon and S. Gunn (eds.), Feature Extraction, Foundations and Applications.  
Springer, 2006.
2. David Golan, Saharon Rosset  
*Statistical Modeling of coverage in High-Throughput Data*  
Chapter 4 in Deep Sequencing Data Analysis, N. Shomron (ed.), Springer, 2013.
3. Saharon Rosset  
*Practical Sparse Modeling: an Overview and Two Examples from Genetics*  
Chapter 3 in Practical Applications of Sparse Modeling, I. Rish et al. (eds.), MIT Press, 2014.

4. David Golan, Saharon Rosset  
*Mixed Models for Case-Control Genome-Wide Association Studies: Major Challenges and Partial Solutions*  
Handbook of Statistical Methods for Case-Control Studies, Chapman and Hall/CR, 2018.

### Peer-Reviewed Competitive Conference Papers

1. Saharon Rosset  
*Ranking: Methods for Flexible Evaluation and Efficient Comparison of Classification Performance*  
Fourth International Conference on Knowledge Discovery and Data Mining, KDD-98.
2. Saharon Rosset, Uzi Murad, Einat Neumann, Yizhak Idan, Gadi Pinkas  
*Discovery of Fraud Rules for Telecommunications: Challenges and Solutions*  
Fifth Intl. Conf. on Knowledge Discovery and Data Mining, KDD-99.
3. Saharon Rosset, Einat Neumann, Uri Eick, Nurit Vatnik, Shuki Idan  
*Evaluation of Prediction Models for Campaign Planning*  
Seventh Intl. Conf. on Knowledge Discovery and Data Mining, KDD-01.
4. Saharon Rosset, Einat Neumann, Uri Eick, Nurit Vatnik, Shuki Idan  
*Lifetime Value Modeling and Its Use for Customer Retention Planning*  
Eighth Intl. Conf. on Knowledge Discovery and Data Mining, KDD-02.  
**Winner of best application paper award**  
(short version of journal paper 1)
5. Saharon Rosset, Eran Segal  
*Boosting Density Estimation*  
Neural Information Processing Systems 15, 2002.
6. Ji Zhu, Saharon Rosset, Trevor Hastie, Rob Tibshirani  
*1-norm Support Vector Machines*  
Neural Information Processing Systems 16, 2003.
7. Saharon Rosset, Ji Zhu, Trevor Hastie  
*Margin Maximizing Loss Functions*  
Neural Information Processing Systems 16, 2003.
8. Saharon Rosset, Einat Neumann  
*Integrating Customer Value Considerations into Predictive Modeling*  
Third IEEE Conference on Data Mining, ICDM-03.

9. Saharon Rosset  
*Model Selection via the AUC*  
21st International Conference on Machine Learning, ICML-04.
10. Trevor Hastie, Saharon Rosset, Rob Tibshirani, Ji Zhu  
*The Entire Regularization Path of the Support Vector Machine*  
Neural Information Processing Systems 17, 2004.  
(short version of journal paper 6)
11. Saharon Rosset, Ji Zhu, Hui Zou, Trevor Hastie  
*A Method for Inferring Label Sampling Mechanisms in Semi-Supervised Learning*  
Neural Information Processing Systems 17, 2004.
12. Saharon Rosset  
*Tracking Curved Regularized Optimization Solution Paths*  
Neural Information Processing Systems 17, 2004.
13. Saharon Rosset  
*Robust Boosting and Its Relation to Bagging*  
Eleventh Intl. Conf. on Knowledge Discovery and Data Mining, KDD-05.
14. Saharon Rosset, Claudia Perlich, Bianca Zadrozny, Srujana Merugu, Sholom Weiss, Rick Lawrence  
*Customer Wallet Estimation*  
1st NYU workshop on CRM and Data Mining, 2005.
15. Saharon Rosset, Claudia Perlich, Bianca Zadrozny  
*Ranking-Based Evaluation of Regression Models*  
Fifth IEEE Conference on Data Mining, ICDM-05.  
(short version of journal paper 9)
16. Sofus Macskassy, Foster Provost, Saharon Rosset  
*ROC Confidence Bands: An Empirical Evaluation*  
22nd International Conference on Machine Learning, ICML-05.
17. Saharon Rosset, Richard Lawrence  
*Data Enhanced Predictive Modeling for Sales Targeting*  
2006 SIAM Conference on Data Mining.
18. Srujana Merugu, Saharon Rosset, Claudia Perlich  
*A New Multi-View Regression Method with an Application to Customer Wallet Estimation*  
Twelfth Intl. Conf. on Knowledge Discovery and Data Mining, KDD-06.

19. Claudia Perlich, Saharon Rosset  
*Identifying Bundles of Product Options using Mutual Information Clustering*  
2007 SIAM Conference on Data Mining, 2007.
20. Claudia Perlich, Saharon Rosset, Bianca Zadrozny, Richard Lawrence  
*High-Quantile Modeling for Customer Wallet Estimation and Other Applications*  
Thirteenth Intl. Conf. on Knowledge Discovery and Data Mining, KDD-07.
21. Saharon Rosset, Grzegorz Swirszcz, Ji Zhu, Nathan Srebro  
 *$l_1$  Regularization in Infinite Dimensional Feature Spaces*  
Twentieth Conference on Learning Theory (COLT-07), 2007.
22. Saharon Rosset  
*Bi-Level Path Following for Cross Validated Solution of Kernel Quantile Regression*  
25th International Conference on Machine Learning (ICML-08), 2008.  
(short version of journal paper 21)
23. Prem Melville, Saharon Rosset, Richard Lawrence  
*Customer Targeting Models Using Actively-Selected Web Content*  
Fourteenth Intl. Conf. on Knowledge Discovery and Data Mining, KDD-08.
24. Aurelie Lozano, Naoki Abe, Yan Liu, Saharon Rosset  
*Grouped Graphical Granger Modeling Methods for Temporal Causal Modeling*  
Fifteenth Intl. Conf. on Knowledge Discovery and Data Mining, KDD-09.
25. Ronny Luss, Saharon Rosset, Moni Shahar  
*Decomposing Isotonic Regression for Efficiently Solving Large Problems*  
Neural Information Processing Systems 23, 2010.
26. Shachar Kaufman, Saharon Rosset and Claudia Perlich  
*Leakage in Data Mining: Formulation, Detection, and Avoidance*  
Seventeenth Intl. Conf. on Knowledge Discovery and Data Mining, KDD-11.  
**Winner of best paper award**  
(short version of journal paper 38)
27. Amichai Painsky, Saharon Rosset  
*Exclusive Row Biclustering for Gene Expression Using a Combinatorial Auction Approach*  
IEEE Conference on Data Mining (ICDM-2012), 2012.  
(short version of journal paper 48)
28. Amichai Painsky, Saharon Rosset  
*Compressing Random Forests*  
IEEE Conference on Data Mining (ICDM-2016), 2016.



29. Regev Schweiger, Eyal Fisher, Elinor Rahmani, Liat Shenhav, Saharon Rosset, Eran Halperin  
*Using Stochastic Approximation Techniques to Efficiently Construct Confidence Intervals for Heritability*  
Intl. Conf. on Research in Computational Molecular Biology (RECOMB), 2017.  
(short version of journal paper 69)
30. Blake Woodworth, Vitaly Feldman, Saharon Rosset, Nathan Srebro  
*The Everlasting Database: Statistical Validity at a Fair Price*  
Neural Information Processing Systems 31, 2018.
31. Ali Pazokitoroudi, Andy Dahl, Noah Zaitlen, Saharon Rosset and Sriram Sankararaman  
*Biobank-scale estimation of the proportion of trait variance explained by gene-environment interactions*  
Intl. Conf. on Research in Computational Molecular Biology (RECOMB) 2021.
32. Giora Simchoni, Saharon Rosset  
*Using Random Effects to Account for High-Cardinality Categorical Features and Repeated Measures in Deep Neural Networks*  
Neural Information Processing Systems 34, 2021.

Date: March 2023