Name: Prof. Feldman Dan **Date: 28/04/2021**

CURRICULUM VITAE

1. Personal Details

Permanent Home Address: 2nd Harav Assaf David St., Haifa

Home Telephone Number: 972-4-6575090

Office Telephone Number: 972-4-8240259

Cellular Phone: 972-54-3213908

Electronic Address: danf@cs.haifa.ac.il

2. Higher Education

a. Undergraduate and Graduate Studies

Period of Study	Name of Institution and Department	Degree	Year of Approval of Degree
1998-2002	Tel Aviv University School of Computer Science	BSc	2002
2002-2004	Tel Aviv University School of Computer Science	MSc	2004
2004-2010	Tel Aviv university School of Computer Science	PhD	2010

b. Post-Doctoral Studies

Period of Study	Name of Institution and Department	Name of Host
March 2010-June 2011	California Institute of Technology (Caltech), Center for the Mathematics of Information (CMI), Fellowship member	Head: Prof. Leonard Schulman
July 2012-February 2014	MIT, Computer Science & Artificial Intelligence Laboratory (CSAIL), Distributed Robotics Lab	Director of the MIT Computer Science and Artificial Intelligence Laboratory

Note: * Represents activities and publications since appointment to Senior Lecturer
** Represents activities and publications since granting tenure

c. Specialization

Years	Name of Institution and Department	Rank/Position
2009, April- August	Yahoo! Research, Barcelona	Visiting Student (Internship)
*2017, February	Collège de France, INRIA, Paris,	Visiting Scientist
*2018, January- May	Simons Institute for the Theory of Computing at UC Berkeley (Semester long program on Real-Time Decision Making, Invited by Prof. Richard Karp)	Visiting Scientist

3. Academic Ranks and Tenure in Institutes of Higher Education

Years	Name of Institution and Department	Rank/Position
*Oct 2014-May 2020	The Department of Computer Science at the University of Haifa	Senior Lecturer
** May 2020-present	The Department of Computer Science at the University of Haifa	Senior Lecturer with tenure
***April 2021-present	The Department of Computer Science at the University of Haifa	Associate Professor

4. Offices in Academic Administration

Year	Name of Institution and Department	Role
*2014-Present	The Department of Computer Science at the University of Haifa	Director of the Robotics and Big Data Labs
*2017-Present	The Center for Cyber, Law, and Policy (CCLP) of Israel, in University of Haifa	Scientific Committee Member
*2019	Data Science Center at the University of Haifa	Industrial Liaison

5. Scholarly Positions and Activities outside the University

Years	Memberships in Journals Editions
**2020 - Present	Invited Guest Editor of the Special Issue of the *Sensors* Journal: *Sensor Computing for Big Data Analytics* (1424-8220).

Years	Memberships in Academic Professional Associations
*2017-2018	Member of the COST Action IC1406, EU- European Commission. High-Performance Modelling and Simulation for Big Data Applications (cHiPSet).
*2018	Tnufa Project by the Ramon Foundation: Supervisor for excellent high school student projects
*2019 - Present	Flyviz LTD. Co-Founder.
**2020	Author and designer of the high-school class "math with drones", jointly with the Ofanim Foundation

6. Participation in Scholarly Conferences

a. Active Participation - Abroad

Date	Name of Conference	Place of Conference	Subject of Lecture/Discussion	Role
2007	ACM Symposium on Computational Geometry (SoCG)	South Korea	Bi-criteria Linear-time Approximations for Generalized k-Mean/Median/Center	Presenter
2009	ACM Symposium on Theory of Computing (STOC)	Maryland, USA	Private Coresets	Presenter
2011	ACM symposium on Theory of computing (STOC)	California, USA	A Unified Framework for Approximating and Clustering Data,	Presenter
2011	International Conference on Scale Space and Variational Methods in	Ein Gedi, Israel	From High Definition Image to Low Space Optimization	Presenter

	Computer Vision (SSVM)			
2012	International Conference on Advances in Geographic Information Systems (ACM-GIS)	California, USA	The Single Pixel GIS: Learning Big Data Signals from Tiny Coresets	Presenter
2012	ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)	Beijing, China	An Effective Coreset Compression Algorithm for Large Scale Sensor Networks	Presenter
*2015	ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)	Sydney, Australia	More Constraints, Smaller Coresets: Constrained Matrix Approximation of Sparse Big Data	Presenter
*2016	Neural Information Processing Systems (NIPS)	California, USA	Dimensionality Reduction of Massive Sparse Datasets Using Coresets.	Presenter
*2016	Mathematical and Computational Challenges in Real-Time Decision Making	Simons Institute for the theory of computing, Berkeley University of California	Core-sets for Real-Time Tracking using Caratheodory Theorem, with Applications to Drones	Presenter
*2016	SIAM International Conference on Data Mining (SDM)	Miami, Florida, USA	k-Means for Streaming and Distributed Big Sparse Data	Presenter
*2017	ACM/IEEE International Conference	Pittsburgh, Pennsylvania, USA	Coresets for Differentially Private k-Means Clustering and	Presenter

	on Information Processing in Sensor Networks (IPSN)		Applications to Privacy in Mobile Sensor Networks.	
*2017	International Conference on Machine Learning (ICML)	Sydney, Australia	Coresets for Vector Summarization with Applications to Network Graphs	Presenter
*2018	ACM SIGSAC Conference on Computer and Communications Security (CSS)	Toronto, Canada	Secure Search on Encrypted Data via Multi-Ring Sketch	Presenter

b. <u>Organization of Conferences or Sessions</u>

Date	**Name of Conference**	**Place of Conference**	**Subject of Lecture/Discussion**	**Role**
*2014	ADVCOMP - Inter. Conf. on Advanced Eng. Computing & Applications in Sciences	Italy	Advanced Eng. Computing & Applications in Sciences	Program committee member
*2015	ADVCOMP - Inter. Conf. on Advanced Eng. Computing & Applications in Sciences	France	Advanced Eng. Computing & Applications in Sciences	Program committee member
*2016	AISTATS - International Conference on Artificial Intelligence and Statistics	Spain	Artificial Intelligence and Statistics	Program committee member
*2016	WAFR - International Workshop on the Algorithmic Foundations of Robotics	San-Francisco, USA	Algorithmic Foundations of Robotics	Program committee member
*2016	IoWT - Workshop on the Internet of Wearable Things	California, USA	Internet of Wearable Things	Program committee member
*2016	ADVCOMP - Inter. Conf. on Advanced Eng. Computing & Applications in Sciences	Italy	Advanced Eng. Computing & Applications in Sciences	Program committee member

*2017	International Conference on Knowledge Discovery and Information Retrieval (KDIR)	Funchal, Portugal	Knowledge Discovery and Information Retrieval	Program committee member
*2018	International Conference on Machine Learning (ICML)	Stockholm, Sweden	Machine Learning	Program committee member
*2018	International Conference on Information Processing in Sensor Networks (IPSN)	Stony Brook University, New-York	Sensor Networks	Program committee member
*2018	International Conference on Knowledge Discovery and Information Retrieval	Seville, Spain	Knowledge Discovery and Information Retrieval	Program committee member
*2018	International Conference on Machine Learning (ICML)	Stockholm, Sweden	Machine Learning	Program committee member
*2019	International Joint Conferences on Artificial Intelligence (IJCA)	Macao, China	Artificial Intelligence	Program committee member
*2019	Association for the Advancement of Artificial Intelligence (AAAI)	New York, New York USA	Artificial Intelligence	Program committee member
*2020	International Conference on Machine Learning (ICML)	Virtual Conference (due to COVID-19)	Machine Learning	Program committee member
**2021	International Conference on Machine Learning (ICML)		Machine Learning	Reviewer

Note: Conference with no organization mentioned are independent conferences.

7. Invited Lectures

a. Academy -Abroad

Year	Name of Forum	Place of Lecture	Subject of Lecture	Role
2009	IITK Workshop on Algorithms for Processing Massive Data Sets	Department of Computer Science and Engineering, IIT , Kanpur, India	Coresets for K-Means Clustering and Generalizations	Invited Speaker
2012	Workshop on Algorithms for Data Streams	Technical University of Dortmund, Germany	From Big (GPS) Data to a Searchable text Diary	Invited Speaker
2012	Workshop on Algorithms for Modern Massive Data Sets	Stanford University	Google your life: Learning Sensors Data	Invited Speaker
*2014	Workshop on Kernelization	University of Warsaw, Poland	Turning Big Data into Tiny Data	Invited Speaker
*2014	Machine learning Summer school	Carnegie Mellon University (CMU)	Core-sets Tutorials	Invited Speaker
*2014	Workshop on Sublinear Algorithms	Centro Universitario Residenziale di Bertinoro, Bertinoro, Italy	Sublinear systems using Core-Sets	Invited Speaker
*2017	National Institute of Informatics (NII) Shonan Meeting	Tokyo, Japan	Processing Big Data Streams	Invited Speaker
*2018	National Institute of Informatics (NII) Shonan Meeting	Tokyo, Japan	Analyzing Large Collections of Time Series	Invited Speaker
*2018	Closed Workshop on Computational Geometry	Freie Universitat, Berlin, Netherlands	Computational Geometry for Machine Learning	Invited Speaker
**2021 (To appear)	National University of Singapore	Singapore	Workshop on Foundations of Data Science	Invited Speaker
**2021 (To appear)	ICML (The International Conference on Machine Learning)	Virtual Conference	Subset Selection in Machine Learning: From Theory to Applications	Invited Speaker

**2021 (To appear)	International Conference on Robotics Artificial Intelligence & Automation	Osaka, Japan	Extending the Frontiers of Technologies and Advancements in Robotic and Artificial Intelligence	**Invited Speaker**

b. <u>Academy –Israel</u>

Year	Name of Forum	Place of Lecture	Subject of Lecture	Role
*2017	AmosFest	Tel-Aviv University	Navigation Algorithms for Autonomous Drones in the Supermarkets of Rami Levi	Invited Speaker
*2017	6th French-Israeli Workshop on Foundations of Computer Science	Tel-Aviv University	Secure search on the cloud via Coresets	Invited Speaker
*2019	Civil law in the digital space	Supreme Court venue, Kibutz Maale Hahamisha	Robotics and IoT (Internet of Things)	Invited Speaker
*2019	Technion Computer Engineering (TCE) conference: Autonomous Systems	Technion, Haifa	Core-sets for Nano Drones	Invited Speaker
*2019	Theory Day	The Open University of Raanana, Israel	Theory of Drones	Invited Speaker

c. <u>Industry</u>

Year	Name of Forum	Place of Lecture	Subject of Lecture	Role
2008	Oracles Business Intelligence Day	Tel-Aviv	Data Reduction for Campaign Management	Invited Speaker
*2016	Uber, Data Science Research Department	Uber's Headquarters, San-Francisco	Streaming Algorithms for Big (GPS) Data	Invited Speaker
*2016	Samsung Data Science Research	Samsung, Korea	Coresets for Deep Learning	Invited Speaker

*2016	CVVC - Computer Vision Validation Conference	Intel, Israel	Core-sets for Computer Vision Validation	Invited Speaker
*2017	4th China IoT (Internet of Things) Conference	Beijing, China	Computational Geometry meets Internet of Things	Invited Speaker
*2018	SAMSUNG, Vice Presidents Forum	Korea	On-Device Deep Learning using Coresets	Invited Speaker
*2018	Google X	Mountain View, Silicon Valley, California	Real-Time Robotics using Coresets	Invited Speaker
*2018	Uber, Data Science Research Department	Uber's Headquarters, San-Francisco	Coresets for Deep Learning	Invited Speaker
*2018	Airbus Research	Munich, Germany	Anomaly detection in IoT (Internet of Things)	Invited Speaker
*2019	Samsung Research	Seoul, Korea	Deep Learning On-Device	Invited Speaker
**2020	Elbit LTD	Haifa, Israel	Coresets for Deep Learning	Invited Speaker
**2020	TracxPoint LTD	Haifa, Israel	Coresets for Object Recognition	Invited Speaker

8. Colloquium Talks

Year	**Name of Forum**	**Place of Lecture**	**Presentation**
2011	Computer Science	Ben-Gurion University.	Low-cost and Faster Tracking Systems Using Core-sets for Pose-Estimation
2013	Computer Science	Carnegie Mellon University (CMU)	Learning Patterns in Big Data from Small Core Sets
2013	Computer Science	University of Boston, USA	Learning Patterns in Big Data from Small Core Sets
*2017	Department of Statistics	Department of Statistics University of Haifa	Coresets and Sketches for High Dimensional Subspace Approximation Problems
*2017	Computer Science	Technion, Israel Institute of Technology	From Theorems to Autonomous Toy-Drones

*2017	Computational Geometry and Topology in the Sciences	Collège de France, INRIA, Paris, France	Epsilon-nets meets coresets
*2017	Machine Learning Group	Tel-Aviv University	Real-time Machine Learning using Core-Sets: Autonomous Toy-Drones for Rami Levy
*2019	Computer Science Department	Bar-Ilan University	Machine Learning, Computer Vision, and Drones at the Supermarket
*2019	Robotics Seminar	Technion, Haifa	Modern Computational Geometry for Robotics
*2019	Computer Science Department	University of Tel-Aviv	k-Means Clustering of Lines for Big Data (student's talk)
**2020	Seminar	Center of Mathematical Sciences and Applications, Harvard University	Fast and Accurate Least-Mean-Squares Solvers

9. Scholarships, Awards and Prizes

- 2012-2014 Foxconn Scholarship Award for Postdoc at MIT
- 2018, Samsung's Best Open Innovation Research Proposal.
- 2019, **Honorable Mention Outstanding Paper Award**, Published paper [F31]**,** 34rd Conference on Neural Information Processing Systems (NeurIPS'19). Out of 6743 submissions.

10. Teaching

a. **Courses Taught in Recent Years**

Years	**Name of Course**	**Type of Course**	**Level**	**Number of Students**
*2014-2017	Big Data: Theory and Applications	Seminar	BSc	2014: 30 2015: 34 2016: 22 2017: 26
*2015-2016	Big Data and the "Internet of Things" for ETGAR	Workshop/Lab	BSc/MSc	2015: 32 2016: 31

*2015-2016	Big Data and the "Internet of Things"	Workshop/Lab	BSc/MSc	2015: 54 2016: 62
*2014 - present	Object Oriented Programming	Introduction Course (Mandatory)	BSc	~160
*2014 - present	Advanced Robotics and Big Data	Workshop/Lab	BSc/MSc	2014: 14 2015: 49 2016: 36 2017:40 2018:20 2019:36
*2017 - present	Streaming Algorithms for Learning Big-data on the cloud	Graduate class	BSc/MSc	2017: 60 2018:50 2019:18
*2017 - present	Data analysis in Real-Time systems	Graduate class	BSc/MSc	2017: 60 2018:75
*2018 - present	Musicological Robotics Lab	Undergraduate class	BSc	2018: 15 2019: 11

11. Miscellaneous

My group aims to publish only to top conferences. In case of rejection I encourage my students to improve the results and resend. This may explain why publication time is relatively slow. Our research problems usually arose in the lab and real-world. We then formalize the theoretical problem statement, hopefully solve it, implement the solution and compare to the state-of-the-art in both academy and industry. This is why most of our papers have “loop closure” between theorems and systems. This also helps to attract sponsors and explain the motivation behind our papers.

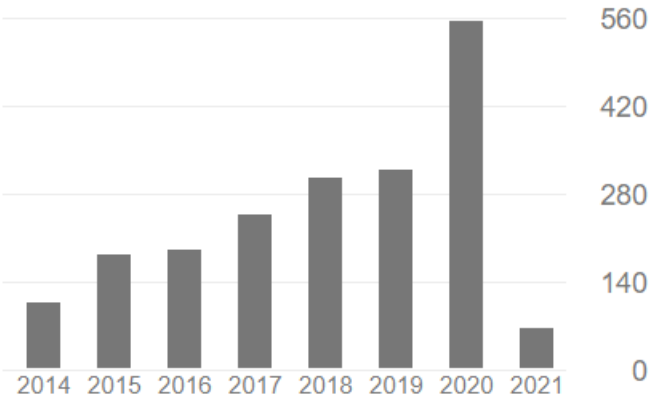
- **Photo with part of the RBD Lab group, taken about a year ago.**

- **Photo with part of the RBD Lab group, taken at Jan. 2020.**

PUBLICATIONS

Notes: The order of the authors is by contribution, unless stated otherwise
(#) denotes that the author was a research student or a postdoc at the time of publication, under my supervision.
IF and ranking were taken from JCR, year of publication.

	All	Since 2016
Citations	2372	1684
h-index	23	19
i10-index	38	34



A. Ph.D. Dissertation

Title: Coresets and Their Applications
Date of submission: 2010
Number of pages: 84
Language: English
Name of supervisors: Prof. Amos Fiat and Prof. Micha Sharir
University: Tel-Aviv University

B. Scientific Books (Refereed)

Authored Books

None

C. Monographs

None

D. Articles in Refereed Journals

Published

1. **Feldman** D., Feigin M. (#) & Sochen N., Learning big (image) data via coresets for dictionaries. (Based on Conf. paper. F.7). *Journal of mathematical Imaging and Vision (JMIV)* 2013, *46*(3), 276-291. IF(2013) = 2.330 R=24/121 Q1.

*2. **Feldman, D.**, Sung, C. (#), Sugaya, A. (#) & Rus, D., iDiary: From GPS signals to a text-searchable diary. (Based on Conf. papers F.12 & F13). *ACM Transactions on Sensor Networks (TOSN)* 2017, *11*(4), pp.60:1-60:41. IF(2017): 1.448. R= 55/144 Q2.

*3. Lucic M. (#), Faulkner M. (#), Krause A. & **Feldman** D., Training Gaussian mixture models at scale via coresets. (Based on Conf. paper F.8) *The Journal of Machine Learning Research (JMLR)* 2018, 18(1), 5885-5909. IF(2018)= 4.091 R= 12/133 Q1. (*Author sorted according to a group leader)*

*4. Epstein D. (#) & **Feldman D.**, Quadcopter Tracks Quadcopter via Real-Time Shape Fitting. *IEEE Robotics and Automation Letters (RA-L)* 2018, *3*(1), pp. 544-550. IF (2018) =4.25, 4/26 (Q1) (Robotics). (*According to a group leader)*

*5. Akavia A., **Feldman D.** & Shaul, H. (#), Secure Data Retrieval on the Cloud: Homomorphic Encryption meets Coresets, *Transactions on Cryptographic Hardware and Embedded Systems (TCHES)* 2019, Issue 2, 80-106. **New Journal (since 2018)** based on the well-known *CHES* conference. *(In alphabetical order)*

*6. Danial J. (#), **Feldman D.** & Hutterer, A. (#), Position Estimation of Multiple Robots: Provable, Practical Approximation Algorithm. *IEEE Robotics and Automation Letters (RA-L)* 2019. *Vol. 4, pp. 1985-1992.* IF (2018)=4.25, 4/26 (Q1) (Robotics). *(In alphabetical order)*

Also accepted and presented in *IEEE/RSJ International Conference on Intelligent Robots and Systems (ICRA)* 2019.)

***7. Feldman, D.,** Core-sets: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.* pp. 1—18, 2020. IF (2018)=2.541, 61/134 (Q2) (Computer Science Artificial Intelligence), 26/105 (Q1) (Computer Science, Theory and Methods).

*8. **Feldman, D**. & Haber, E., Measuring Privacy in the Always-On Era. *The Berkeley Technology Law Journal* (BTLJ) 2019. (53 pages). W&L Rank (**Law Journals:** Washington and Lee University): No.1 out of 76 in IP, and No.1 out of 58 in Science, Technology and computing.

*9. Netzer E. (#), Frid A. (#) & **Feldman D.,** Real-time EEG classification via coresets for BCI applications. *Engineering Applications of Artificial Intelligence (2020),* Elsevier, 8 pages, IF (2018)=3.526, 19/62(Q2) (Automation & Control Systems). 36/134 (Q2) (Computer Science, Artificial Intelligence).

*10. Rozenberg L. (#), Lotan S. (#) & **Feldman D.**, Finding Patterns in Signals Using Lossy Text Compression, *Algorithms* (18 pages).

NOTE: this journal is not ranked in CORE. According to SCIMAGO ranking, it was, in 2019, Q3 in Computational Mathematics, Numerical Analysis, and Theoretical Computer Science. Impact Factor: 1.51.

*11. **Feldman D.,** Schmidt M. (#), & Sohler C., Turning Big data into tiny data: Constant-size coresets for k-means, PCA and projective clustering. *SIAM Journal on Computing (SICOMP) 2020.* (Based on conference paper no.F.14) (88 pages). IF (2018)=1.563, 47/105 (Q2) (Computer Science, Theory & Methods), 67/254 (Q2) (Mathematics, Applied).

12. **Feldman D., Shaul H. (#) & Rus, D., Secure k-ish Nearest Neighbors, in *Proceedings on Privacy Enhancing Technologies (PoPETs) 2020.* (29 pages)

This is a quote from the journal homepage: “Submitted papers undergo a journal-style reviewing process and accepted papers are published in the journal Proceedings on Privacy Enhancing Technologies (PoPETs). PoPETs, a scholarly, open access journal for timely research papers on privacy, has been established as a way to improve reviewing and publication quality while retaining the highly successful PETS community event. PoPETs is published by Sciendo, part of De Gruyter, which has over 260 years of publishing history. PoPETs does not have article processing charges (APCs) or article submission charges." (from: https://content.sciendo.com/view/journals/popets/popets-overview.xml)

13. Barger A. (#) & **Feldman D.. Deterministic Coresets for *k*-Means of Big Sparse Data. *Algorithms 13(4):92 (2020).* (Based on conference paper no. F.22). (16 pages).

NOTE: this journal is not ranked in CORE. According to SCIMAGO ranking, it was, in 2019, Q3 in Computational Mathematics, Numerical Analysis, and Theoretical Computer Science. Impact Factor: 1.51.

14. Nasser S. (#), Jubran I. (#) & **Feldman D.. Autonomous Toy Drone via Coresets for Pose Estimation. *Sensors 20(11): 3042 (2020).* (20 pages).

NOTE: Not ranked in CORE. Q2 in Electrical and Electronic Engineering, Q1 in Instrumentation. IF 3.28, 15/64 (Q1) among all titles in the 'Instruments & Instrumentation', category and 77/266 (Q2) in the 'Engineering, Electrical & Electronic' category.

Accepted for Publication

*15. Jubran I. (#) & **Feldman D..** Aligning Points to Lines: Provable Approximations. *IEEE Transactions on Knowledge and Data Engineering (TKDE).* (24 pages). IF (2018)=3.857, 29/134 (Q1) (Computer Science, Artificial Intelligence), 28/155 (Q1) (Computer Science, Information Systems), 58/266 (Q1) (Engineering, Electrical & Electronic).

** 16. Epstein D. (#) & **Feldman D**. Sphere Fitting with Applications to Machine Tracking. *Algorithms* (18 pages).

NOTE: this journal is not ranked in CORE. According to SCIMAGO ranking, it was, in 2019, Q3 in Computational Mathematics, Numerical Analysis, and Theoretical Computer Science. Impact Factor: 1.51.

****17.** Statman A. (#), Rozenberg L. (#) & **Feldman D**. k-Means+++: Outliers-Resistant Clustering. *Algorithms* (21 pages).

NOTE: this journal is not ranked in CORE. According to SCIMAGO ranking, it was, in 2019, Q3 in Computational Mathematics, Numerical Analysis, and Theoretical Computer Science. Impact Factor: 1.51.

18. Mussay B. (#), **Feldman D., Zhou S. (#), Braverman V. & Osadchy M., IEEE Transactions on Neural Networks, Rank A^. (Based on conference paper no F.34) (16 pages).

E. Articles or Chapters in Scientific Books (Refereed)

*1. **Feldman, D.,** Sampling techniques for handling Big Data via Coresets. In Frederic Ros and Serge Guillaume (eds.) *Sampling Techniques for Supervised or Unsupervised Tasks*. 25 pages. Springer publication

F. Articles in Conference Proceedings

Ranking is from CORE: Computing Research & Education (http://www.core.edu.au/). It states the percentage of acceptations: A^ – 4%, A – 14%, B – 26%, C – 49%.

Published

1. **Feldman D.,** Fiat A. & Sharir, M., Coresets for weighted facilities and their applications. *Annual IEEE Symposium on Foundations of Computer Science (FOCS)* 2006, pp. 315-324, rank A^. (Extended version submitted as KB.1)*(In alphabetical order).*
2. **Feldman D.,** Fiat A., Sharir M. & Segev D. (#), Bi-criteria linear-time approximations for generalized k-mean/median/center. *Annual ACM symposium on Computational geometry (SoCG)* 2007, pp. 19-26, rank A. *(In alphabetical order).*

3. **Feldman D.,** Monemizadeh M. (#) & Sohler C., A PTAS for k-means clustering based on weak coresets. *Annual ACM symposium on Computational geometry (SoCG)* 2007, pp. 11-18, rank A..*(In alphabetical order).*
4. **Feldman D.,** Fiat A., Kaplan H. & Nissim K., Private coresets. *Annual ACM Symposium on Theory of Computing (STOC) 2009*, pp. 361-370, rank A^.*(In alphabetical order)*
5. **Feldman D.,** Monemizadeh M. (#), Sohler C. & Woodruff D. P., Coresets and sketches for high dimensional subspace approximation problems. *ACM-SIAM symposium on Discrete Algorithms (SODA) 2010*, pp. 630-649, rank A^. *(In alphabetical order)*
6. **Feldman D.** & Langberg M., A unified framework for approximating and clustering data. *Annual ACM symposium on Theory of Computing* (STOC) *2011*, pp. 569-578, rank A^.*(In alphabetical order)*
7. Feigin M. (#), **Feldman D.** & Sochen, N., From high definition image to low space optimization. *International Conference on Scale Space and Variational Methods in Computer Vision (SSVM)* 2011, pp. 459-470. (No rank found). (Extended version published as D.1). .*(In alphabetical order)*
8. **Feldman, D**., Faulkner M. (#), & Krause, A., Scalable training of mixture models via coresets. *Advances in neural information processing systems (NIPS) 2011*., pp. 2142-2150, rank A^. Poster+oral presentation track, (Extended version published as D.3) *(In alphabetical order)*
9. **Feldman D. &** Schulman, L. J., Data reduction for weighted and outlier-resistant clustering. *Annual ACM-SIAM symposium on Discrete Algorithms (SODA)* 2012, pp. 1343-1354, rank A^. .*(In alphabetical order)*
10. Gil S. (#), **Feldman D**., & Rus D., Communication coverage for independently moving robots. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 2012, pp. 4865-4872, rank A.
11. Sung C. (#), **Feldman, D.** & Rus, D., Trajectory clustering for motion prediction. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 2012, pp. 1547-1552, rank A.
12. **Feldman D.,** Sugaya A. (#) & Rus, D., An effective coreset compression algorithm for large scale sensor networks. *Information Processing in*

Sensor Networks (IPSN) 2012, pp. 257-268, rank A^. (Extended version published as D.2)

13. **Feldman D.,** Sung, C. (#) & Rus, D., The single pixel GPS: learning big data signals from tiny coresets. *International Conference on Advances in Geographic Information Systems (GIS) 2012*, pp. 23-32, rank B. (Extended version published as D.2).
14. **Feldman D.,** Schmidt M. (#) & Sohler C., Turning big data into tiny data: Constant-size coresets for k-means, PCA and projective clustering. *Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) 2013*, pp. 1434-1453, rank A^. (Extended version submitted as Kc.1). .*(In alphabetical order)*
15. **Feldman D.,** Gil S. (#), Knepper R. (#), Julian B. (#) & Rus D., *K*-Robots clustering of moving sensors using coresets. *IEEE International Conference on Robotics and Automation (ICRA) 2013*, pp. 881-888, rank B.

*16. **Feldman D.,** Sugaya A. (#), Sung C. R. (#), & Rus D., iDiary: from GPS signals to a text-searchable diary. *The 11th ACM Conference on Embedded Networked Sensor Systems (SenSys) 2013,* Pages 6:1-6:12. Rank A^.

*17. Rosman G. (#), Volkov, M. (#), **Feldman, D.,** Fisher III, J. W. & Rus, D., Coresets for k-segmentation of streaming data. *Advances in Neural Information Processing Systems (NIPS)* 2014, pp. 559-567, rank A^. (Extended version submitted as Kb.2).

*18. Paul R (#)., **Feldman D**., Rus D. & Newman, P. Visual precis generation using coresets. *IEEE International Conference on Robotics and Automation (ICRA) 2014*, pp. 1304-1311, rank B.

*19. Munteanu A. (#), Sohler C. & **Feldman D**., Smallest enclosing ball for probabilistic data. *Annual ACM symposium on Computational Geometry (SoCG) 2014*, pp. 214-223, rank A.

*20. **Feldman D.** & Tassa T., More constraints, smaller coresets: Constrained matrix approximation of sparse big data. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* 2015, pp. 249-258, rank A^. (Extended version submitted as Kb. 3).

*21. Nasser S. (#), Barry A. (#), Doniec, M. (#), Peled G. (#), Rosman G. (#), Rus D., & **Feldman D.,** Fleye on the car: big data meets the internet of things. *International Conference on Information Processing in Sensor Networks (IPSN)* 2015, pp. 382-383, rank A^. (*According to a group leader)*

*22. Volkov M. (#), Rosman G. (#), **Feldman D**., Fisher, J. W. & Rus D., Coresets for visual summarization with applications to loop closure. *IEEE International Conference on Robotics and Automation (ICRA)* 2015, pp. 3638-3645, rank B**.**

*23. Barger A. (#) & **Feldman D**., *k*-Means for Streaming and Distributed Big Sparse Data. *SIAM International Conference on Data Mining* (SDM) 2016, pp. 342-350, rank A. (*According to a group leader)*

*****24**. Feldman D.,** Volkov M. (#) & Rus D., Dimensionality reduction of massive sparse datasets using coresets. *Advances in Neural Information Processing Systems (NIPS)* 2016, pp. 2766-2774. rank A^.

*****25**. Feldman D.,** Xiang C. (#), Zhu R. (#) & Rus, D., Coresets for differentially private k-means clustering and applications to privacy in mobile sensor networks. *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN) 2017*, pp. 3-16, rank A^.

*****26**. Feldman D.**, Ozer S. (#) & Rus D., Coresets for vector summarization with applications to network graphs. *International Conference on Machine Learning (ICML) 2017*, pp. 1117-1125, rank A^.

*27. Akavia, A., **Feldman, D.** & Shaul H. (#), Secure Search on Encrypted Data via Multi-Ring Sketch. *ACM SIGSAC Conference on Computer and Communications Security (CSS)* 2018, pp. 985-1001, rank A^.*(In alphabetical order)*

*28. Baykal C. (#), Liebenwein L. (#), Gilitschenski I. (#), **Feldman D**., & Rus, D., Data-Dependent Coresets for Compressing Neural Networks with Applications to Generalization Bounds. *International Conference on Learning Representations (ICLR) 2019.* (35 pages).

*29. Lang H. (#), Baykal C. (#), Abu-Samra N. (#), Tannous T. (#), **Feldman D.,** & Rus D., Deterministic Coresets for Stochastic Matrices with Applications to Scalable Sparse PageRank. *Theory and Applications of Models of Computation (TAMC) 2019: 410-423*, rank B.

*30. Braverman, V., **Feldman, D**. & Lang, H. (#), Streaming Coresets for M-Estimators, *International Conference on Randomization and Computation* (RANDOM) 2019, pp. 62:1—62:16.

*31. Malouf A. (#), Jubran, I. (#), & **Feldman, D.,** Fast and Accurate Least-Mean-Squares Solvers. Conference on Neural Information Processing Systems (NeurIPS, formerly NIPS), 2019: 8305-8316. (12 pages). Oral

(36/6743=0.05% acceptance rate)+Outstanding Paper Award, rank A^. (Extended version submitted as Kc. h10).

*32. Marom Y. (#) & **Feldman D.** *k*-Means Clustering of Lines for Big Data. Conference on Neural Information Processing Systems (NeurIPS, formerly NIPS), 2019: 12797-12806, (10 pages), rank A^. (Extended version submitted as Kc. 9).

33. Liebenwein L., Baykal C., Lang H., **Feldman D., & Rus, D., Provable Filter Pruning for Efficient Neural Networks. *8th International Conference on Learning Representations (ICLR) 2020.* 24 pages. (accepted paper presented in the poster session)

34. Mussay B. (#), Osadchy M., Braverman V., Zhou S. (#) & **Feldman D., Data-Independent Neural Pruning via Coresets. *8th International Conference on Learning Representations (ICLR) 2020.* 14 pages. (Extended version submitted as Kb.11).

*35. Tukan M. (#), Baykal C. (#), **Feldman D.** & Rus D.. On Coresets for Supported Vector Machines, *Theory and Applications of Models of Computation (TAMC) 2020*. 27 pages. rank B.

36. Malouf A. (#), Statman A. (#) & **Feldman D. Tight Sensitivity Bounds for Smaller Coresets. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* 2019, pp. 249-258, rank A^. (Extended version submitted as Kb. 5).

37. Malouf A. (#), Jubran, I. (#), Tukan M. (#) & **Feldman, D.. Sets Clustering. *International Conference on Machine Learning (ICML) 2020.* (25 pages). Rank A^.

**38. M. Tukan (#), A. Malouf (#), & Dan Feldman: Coresets for Near-Convex Functions. Conference on Neural Information Processing Systems (NeurIPS, formerly NIPS), 2020, rank A^.

G. Entries in Encyclopedias

None

H. Other Scientific Publications

Patents

*1. Rus, D. & **Feldman D**., "Data coreset compression." U.S. Patent No. 9,286,312. 2016.

*2. **Feldman D.**, Rus D., Sugaya A. (#) & Sung C. (#), "Text characterization of trajectories." U.S. Patent No. 10176244. 2019

*3. **Feldman D.,** Levin O. & Marom Y. (#), "Security breaches detection by utilizing clustering of weighted outliers", application 506/1.2., 2019.

I. Publications on my work (in public media)

*1. Jonathan Vizcarra, Science, and Medical News, Speeding Up GPS Algorithms Through Data Compression, Line Simplification, and Signal Clustering, November 2012.

*2. Larry Hardesty, Communication of the ACM, Speeding algorithms by shrinking data, November 2012.

*3. Kevin Murnane, Forbes News, Turning Big Data Into Manageable Data, December 2016.

*4. Ralph Anzarouth, Computer Vision News, Guest Interview, Pages 20–23 September 2016.

*5. Robin.ly (https://www.robinly.info/), [NeurIPS 2019 Honorable Mention Outstanding Paper] Jubran & Malouf on Least-Mean-Squares Solvers
Also in: https://www.youtube.com/watch?v=lmxyv55t_fg

*6. Fangyu Cai, Synced Review, AI & Machine Intelligence, "NeurIPS 2019 Opens; Outstanding Paper Awards and 2021 Conference Announced", also in: https://medium.com/syncedreview/neurips-2019-opens-outstanding-paper-awards-and-2021-conference-announced-59f140a3ed6e

*7. Matthew Mayo, KDNuggets, "NeurIPS 2019 Outstanding Paper Awards", https://www.kdnuggets.com/2019/12/neurips-2019-paper-awards.html

J. Other Works Connected with my Scholarly Field

None

K. Submitted Publications

Articles

K_a. Revised and Resubmitted

*1. **Feldman D.,** Fiat A., & Sharir M., Coresets for weighted facilities and their applications. *Journal of the ACM*. (50 pages). (based on conference paper F1), IF (2018) =2.17, 21/52 (Q2) (Computer Science, Hardware & Architecturescie), 78/155 (Q3) (Computer Science, Information Systems) 38/107 (Q2) (Computer Science, Software Engineering), 34/105 (Q2) (Computer Science, Theory & Methods).

*2. **Feldman D.,** Rosman G. (#), Volkov M. (#), Fisher III, J. W., & Rus, D., Coresets for *k*-segmentation of streaming data. *Journal of Machine Learning Research (JMLR).* (Based on conference paper F.16). (60 pages)
IF (2018) =4.091, 16/62 (Q2) (Automation & Control Systems), 27/134 (Q1) (Computer Science, Artificial Intelligence).

*3. **Feldman D.** & Tassa, T., Non-negative Matrix Factorization of sparse big data. *Journal of Machine Learning Research (JMLR).* (Based on conference paper F.20). (65 pages).
IF (2018) =4.091, 16/62 (Q2) (Automation & Control Systems), 27/134 (Q1) (Computer Science, Artificial Intelligence).

K_b. Submitted to Journals

*4. Kfir, Z. (#), Wu X., (#) & **Feldman D.**, Coresets for Scalable Learning of any Mixture of Gaussians and Radial Basis Networks. *Journal of Machine Learning Research (JMLR).* (87 pages).
IF (2018) =4.091, 16/62 (Q2) (Automation & Control Systems), 27/134 (Q1) (Computer Science, Artificial Intelligence).

*5. Braverman V., **Feldman D.**, H. Lang (#) & Statman A. (#). New Framework for Smaller Coreset. *Journal of Machine Learning Research (JMLR).* (35 pages).
IF (2018) =4.091, 16/62 (Q2) (Automation & Control Systems), 27/134 (Q1) (Computer Science, Artificial Intelligence).

*6. Tolochinsky E. (#) & **Feldman D.,** Coresets For Monotonic Functions with Applications to Deep Learning. *Journal of Machine Learning Research (JMLR).* (40 pages).
IF (2018) =4.091, 16/62 (Q2) (Automation & Control Systems), 27/134 (Q1) (Computer Science, Artificial Intelligence).

*7. Jubran I. (#), Malouf A. (#) & **Feldman D.**, Accurate Coresets, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI),* (25 pages).
IF (2018) =3.857, 29/134 (Q1) (Computer Science, Artificial Intelligence), 28/155 (Q1) (Computer Science, Information Systems).

*8. Jubran I. (#), Cohn D. (#) & **Feldman D.,** Provable Approximations for Constrained L_p Regression, *Linear Algebra and its Applications. (39 pages).*
IF (2018) =0.977, 148/254 (Q3) (Mathematics Applied), 97/314 (Q2) (Mathematics).

*9. Marom Y. (#) & **Feldman D.,** k-Means Clustering of Lines for Big Data. *The Journal of Machine Learning Research (JMLR),* 2020, (Based on conference paper no. F.32) (48 pages).
IF (2018) =4.091, 16/62 (Q2) (Automation & Control Systems), 27/134 (Q1) (Computer Science, Artificial Intelligence).

*10. Malouf A. (#), Jubran I. (#), & **Feldman D..** Fast and Accurate Least-Mean-Squares Solvers. *The Journal of Machine Learning Research (JMLR)*, 2020, (Based on conference paper no. F.31) (32 pages).
IF (2018) =4.091, 16/62 (Q2) (Automation & Control Systems), 27/134 (Q1) (Computer Science, Artificial Intelligence).

K_C. Submitted to Conferences

Ranking is from CORE: Computing Research & Education (http://www.core.edu.au/ . It states the percentage of acceptations: A^ – 4%, A – 14%, B – 26%, C – 49%.

*1. Malouf A. (#), Jubran I. (#), Tukan M. (#), & Feldman, D., Faster PAC Learning and Smaller Coresets via Smoothed Analysis. *Conference on Neural Information Processing Systems (NeurIPS, formerly NIPS), 2020.* (pages). Rank A^.

*2. Lang H. (#), A. Statman (#), V. Braverman, D., Feldman D. & Rus D. Sparse Coresets for SVD on Infinite Streams. *Conference on Neural Information Processing Systems (NeurIPS, formerly NIPS), 2020.* (25 pages). Rank A^.

*3. Baykal C. (#), Liebenwein L. (#), Gilitschenski I. (#), **Feldman D. &** Rus D.. SiPPing Neural Networks: Sensitivity-informed Provable Pruning of Neural Networks. *Conference on Neural Information Processing Systems (NIPS), 2020.* (18 pages). Rank A^.

**4. Tukan M. (#), Malouf A. (#), & Feldman, D., Coresets for Near-Convex Functions. *Conference on Neural Information Processing Systems (NeurIPS, formerly NIPS), 2020.* (pages). Rank A^.

5. Jubran I. (#), Cohn D. (#) & **Feldman D., Provable Approximations for ConstrainedL_p Regression, *Conference on Neural Information Processing Systems (NIPS), 2020.* (18 pages). Rank A^.

6. Braverman V. (#), **Feldman D., Lang H. (#), Statman A. (#), & Zhou S. (#), Efficient Coresets Constructions via Sensitivity Sampling, *Conference on Neural Information Processing Systems (NIPS), 2020.* (25 pages). Rank A^.

7. Jubran I.(#), Maalouf A. (#), Kimmel R., & **Feldman D., Provably Approximated ICP, *International Conference on Computer Vision(ICCV), 2021*. Rank A^.

8. Liebenwein L., Maalouf A. (#), **Feldman D., & Rus D., Train Less, Prune More: Decomposition-based Structured Pruning via SVD, *International Conference on Machine Learning (ICML), 2021*. Rank A^.

9. Maalouf A. (#), Eini G. (#), Mussay B. (#), **Feldman D. & Osadchy M. AutoCL: Automatic Coreset Learning, *Knowledge Discovery and Data Mining (KDD), 2021*. Rank A^.

L. Summary of my Activities and Future Plans

The Robotics & Big Data Laboratory aims to be the most practical theory group in the world. It was established in 2014 and now it is has two branches: in the main campus and downtown. The group of 30-40 students is focused in designing and implementing novel data reduction algorithms (e.g. core-sets) for learning “Big data” sets in real-time systems. The data is usually collected from sensors on robots or drones, reduced on the

mini (“Internet of Things”) boards, and then being sent to the machines on the cloud that run existing algorithms on the reduced data.

My research combines theory and experimental results with real-world applications and impacts in many different fields: from machine/deep learning, geometry, privacy, crypto, vision, robotics & even law. On the same time, it focuses on my main expertise which is provable data summarization algorithms (branded as core-sets), usually based on modern techniques in computational geometry.

I plan to expand my research in the following directions:

- Book: after being asked so many times by both students and colleagues regarding introduction book on coresets, I started to write one and hope to finish it soon. It will be based on my lecture notes and presentations from class. May be split into journal surveys.
- Unified theory for coreset for data science: instead of having more papers about coresets for specific problems, I intend to focus on a unified framework based on problem classifications and characters, as in complexity or convex optimization.
- Deep Learning and AI: my recent coresets for deep learning seems to be the tip of the iceberg, especially when it comes to provable compressions. I intend to continue research this field and have serious impact on modern AI by allowing classification on smartphones, and training in minutes instead of weeks.
- Computer Vision: techniques in computer vision today are usually based on heuristics with no provable approximations. I expect that approximation algorithms and coresets may change this. My recent papers in tracking and localization show promising results.
- Nano-Drones: The RBD lab is one of the few labs in the world that focuses on safe low-cost nano drones (<250 gram) which are very hard to control due to lack of sensors and weak computation power. Here my goal is to have the first safe and legal swarm that will be spread in all the streets of Haifa on a daily basis.
- Autonomous cars: My main goal is to show that using better real-time algorithms we can use less sensors and devices. Toy drones are especially relevant since they require very fast real-time computations. Via the Magneton with LiveU (today, DriveU) LTD I learned that this is the main challenge in autonomous cars.

- Library Code: Coresets algorithms are extremely practical, while the theory behind them and the papers are less accessible for non-experts. I intend to make them more accessible to a wider audience by developing source code with my group that will allow researchers, engineers, and students to evaluate and understand them via source code and experiments.
- Running open code that will demonstrate them for real-world problems may change this. My group is working also in this direction, with the help of undergraduates.