

# Elena Tuzhilina

**Phone:** +1 (647) 482-91-61

**E-mail:** [elena.tuzhilina@utoronto.ca](mailto:elena.tuzhilina@utoronto.ca)

**Web:** [elenatuzhilina.github.io](https://elenatuzhilina.github.io)

**Mail:** 700 University avenue, 9 floor,  
#9086, Toronto, ON M5G 1X6, Canada

## EMPLOYMENT

### Academic

**June 2022 – current**, *Department of Statistical Sciences, University of Toronto, Toronto, Canada*  
*Assistant Professor in Statistics (tenure-track)*

**October 2023 – May 2024**  
*Parental leave*

### Industrial

**June 2021 – August 2021**, *Microsoft, Redmond, USA*  
*Intern in Data Science*

**October 2015 – September 2017**, *SmartCAT, Moscow, Russia*  
*Researcher in Computer Linguistics*

## EDUCATION

**September 2017 – June 2022**, *Department of Statistics, Stanford University, Stanford, USA*  
*Ph.D. in Statistics* [Advisor: Trevor Hastie]

Research area: Applied Statistics, Biostatistics, Machine learning

**September 2015 – June 2017**, *Department of Data Analysis, Yandex School of Analysis, Moscow, Russia*  
*Two-year program in Data Science* [Advisors: Ilya Muchnik, Boris Polyak, Anatoliy Michalskiy]

Research area: Dimension reduction methods in medical research

**September 2015 – left in 2017**, *Faculty of Mechanics and Mathematics, Moscow State University, Moscow, Russia*  
*Ph.D. in Mathematics* [Advisors: Andrey Raigorodsky, Alexander Bulinski]

Research area: Probability theory, Combinatorics, Random graphs

**September 2010 – June 2015**, *Faculty of Mechanics and Mathematics, Moscow State University, Moscow, Russia*  
*B.Sc and M.Sc. in Mathematics* [Advisor: Alexander Bulinski]

Research area: Probability theory, Statistics and Applications

GPA: 5.0 out of 5.0 (with highest distinction)

## MEMBERSHIPS AND SERVICE

**September 2022 – present**, member of the *University of Toronto Data Sciences Institute*.

**September 2022 - present**, organizer of the *Data Science Applied Research and Education seminar at Toronto*.

## FUNDING

**2023 – 2028**, Discovery Grant, *Natural Sciences and Engineering Research Council of Canada*.

**2023 – 2025**, Accelerator Grant, *The University of Toronto McLaughlin Centre*.

**2022 – 2024**, Catalyst Grant, *The University of Toronto Data Sciences Institute*.

**2020 – 2022**, Stanford Data Science Scholarship, *Stanford University*.

**2016 – 2018**, Grant supporting the SmartCAT project, *Skolkovo Institute of Science & Technology*.

**2014 – 2017**, Grant supporting research on amino acids conformations, *Russian Science Foundation*.

**2011 – 2015**, Outstanding student stipend, *Moscow State University*.

## ACADEMIC AWARDS

- 2023**, Dorothy Shoichet Women Faculty in Science Awards of Excellence, *University of Toronto*.
- 2022**, Student Travel Award, *Joint Statistical Meeting, SFASA*.
- 2022**, Best poster award, *Statistics in the Big Data Era conference, Simons Institute*.
- 2022**, Outstanding Teaching Assistance Award, *Stanford University*.
- 2021**, Stanford Teaching Assistant Award, *Stanford University*.

## PUBLICATIONS

- 2024**, “Canonical Correlation Analysis as Reduced Rank Regression in High Dimensions”, C.Donnat, **E.Tuzhilina**, *submitted to the Journal of the American Statistical Association*.
- 2024**, “Statistical curve models for inferring 3D chromatin architecture”, **E.Tuzhilina**, T.Hastie, M.Segal, *Annals of Applied Statistics*.
- 2023**, “Smooth multi-period forecasting with application to prediction of COVID-19 cases”, **E.Tuzhilina**, T.Hastie, R.Tibshirani, *Journal of Computational and Graphical Statistics*.
- 2022**, “Principal Component Analysis”, M.Greenacre, P.Groenen, T.Hastie, A.D’Enza, A.Markos, **E.Tuzhilina**, *Nature Reviews Methods Primers*, Vol. 2.
- 2021**, “Weighted Low Rank Matrix Approximation and acceleration”, **E.Tuzhilina**, T.Hastie, *available from ArXiv*.
- 2021**, “An Open Repository of Real-Time COVID-19 Indicators”, A. Reinhart, L. Brooks, M. Jahja, A.Rumack, J.Tang, W. Saeed, T.Arnold, A.Basu, J.Bien, A.Cabrera, A.Chin, E.Chua, B.Clark, N.DeFries, J.Forlizzi, S.Gratzl, A.Green, G.Haff, R.Han, A.Hu, S.Hyun, A.Joshi, J.Kim, A.Kuznetsov, W.Motte-Kerr, K.Lee, Y.Lee, Z.Lipton, M.Liu, L.Mackey, K.Mazaitis, D.McDonald, B.Narasimhan, N.Oliveira, P.Patil, A.Perer, C.Politsch, S.Rajanala, D.Rucker, N.Shah, V.Shankar, J.Sharpnack, D.Shemetov, N.Simon, V.Srivastava, S.Tan, R.Tibshirani, **E.Tuzhilina**, A.Nortwick, V.Ventura, L.Wasserman, J.Weiss, K.Williams, R.Rosenfeld, R.Tibshirani, *Proceedings of the National Academy of Sciences*, Vol. 118(5).
- 2021**, “Canonical Correlation Analysis in high dimensions with structured regularization”, **E.Tuzhilina**, L.Tozzi, T.Hastie, *Statistical Modelling SAGE*, Vol. 23(3), pp. 203-227.
- 2021**, “Relating whole-brain functional connectivity to self-reported negative emotion in a large sample of young adults using group regularized canonical correlation analysis”, L.Tozzi, **E.Tuzhilina**, M. Glasser, T.Hastie, L.Williams, *NeuroImage*, Vol. 237.
- 2020**, “Principal curve approaches for inferring 3D chromatin architecture”, E.Tuzhilina, T.Hastie, M.Segal, *Biostatistics*.
- 2017**, “Analyzing the Data Bank of Proteins Space Structures (PDB): A Geometrical Approach”, **E.Vilkul**, A.Ivanov, A.Mishchenko, F.Popelensky, A.Tuzhilin, K.Shaitan, *Springer, Journal of Mathematical Sciences*, 225(4), pp. 555–564.
- 2015**, Addendum to “Critical analysis of amino acids and polypeptides geometry”, A.Ivanov, A.Mishchenko, A.Tuzhilin, *Springer, Continuous and Distributed Systems: Theory and Applications*, Vol. 2, pp. 29–74.
- 2015**, “A geometric approach to the analysis of the data bank of the three-dimensional structures of proteins (PDB)”, **E.Vilkul**, A.Ivanov, A.Mishchenko, F.Popelensky, A.Tuzhilin, K.Shaitan, *Intuit, Pure and Applied Mathematics*, Vol. 20(3), pp. 33-46.
- 2015**, “Conformations of swivel chain as a model of protein folding”, **E.Vilkul**, A.Ivanov, A.Tuzhilin, *The Journal of Nanostructures, Mathematical physics and modelling*, Vol. 13(2), pp. 25-42.
- 2014**, “Geometry of amino acids and polypeptides: the case of X-ray analysis”, **E.Vilkul**, A.Tuzhilin, *The Journal of Nanostructures, Mathematical physics and modelling*, Vol. 11(2), pp. 5-27.

## PATENTS

- 2019**, “Data-driven automated selection of profiles of translation professionals for translation tasks”, A.Ukrainets, V.Gusakov, I.Smolnikov, **E.Tuzhilina**, *patent number US20190065463*.

**2018**, “System and method of intellectual automatic selection of performers of translation”, A.Ukrainets, **E.Tuzhilina**, V.Gusakov, I.Smolnikov, *patent number RU2667030*.

## SOFTWARE

*RCCA (R package)* Implementation of regularized canonical correlation analysis with structured data. Includes three modifications: with standard L2 penalty, with partial L2 penalty, and with group penalty.

*WLRMA (R package)* Performs weighted low-rank matrix approximation. Allows to solve both rank-constraint problem as well as its convex relaxation.

*PoisMS and DBMS (R packages)* Allows to compute 3D chromatin reconstruction using a contact matrix. The approach is based on principal curve technique modeling the chromatin directly by a smooth curve.

## CONFERENCES

### Talks

**June 2024**, “Canonical Correlation Analysis as Reduced Rank Regression in High Dimensions”, *Sixth ICSA-Canada Chapter Symposium*, Niagara Falls, Canada.

**June 2023**, “Statistical curve models for inferring 3D chromatin architecture”, *University of Waterloo*, Waterloo, Canada.

**May 2023**, “Statistical curve models for inferring 3D chromatin architecture”, *Statistical Society of Canada*, Ottawa, Canada.

**February 2023**, “Statistical curve models for inferring 3D chromatin architecture”, *McMaster University*, Hamilton, Canada.

**November 2022**, “Canonical Correlation Analysis in high dimensions with structured regularization”, *University of Mississippi Medical Center*, Virtual.

**August 2022**, “Canonical Correlation Analysis in high dimensions with structured regularization”, *Joint Statistical Meeting*, Washington, USA.

**November 2021**, “Intelligent Rollouts for Office”, *Machine Learning & Data Sciences Conference at Microsoft*, Virtual.

**August 2021**, “Canonical Correlation Analysis in high dimensions with structured regularization”, *Joint Statistical Meeting*, Virtual.

**October 2020**, “Canonical Correlation Analysis in high dimensions with structured regularization”, *Industrial Affiliates Annual Conference at Stanford University*, Virtual.

**August 2020**, “Principal curve approaches for inferring 3D chromatin architecture”, *Summer School of Machine Learning at Skolkovo Institute of Science & Technology*, Virtual.

**August 2020**, “Principal curve approaches for inferring 3D chromatin architecture”, *Joint Statistical Meeting*, Virtual

**November 2019**, “A weighted principal curve approach to inferring 3D chromatin architecture”, *Industrial Affiliates Annual Conference at Stanford University*, Stanford, USA.

**September 2016**, “The geometry of iterations defined on measure metric”, *19th meeting of the Geometrical Seminar*, Zlatibor, Serbia.

**April 2014**, “The continuous case of cancer spread problem”, *Lomonosov Conference at Moscow State University*, Moscow, Russia.

**September 2013**, “Mathematical model of cancer spread”, *Probability, Analysis and Geometry*, Ulm, Germany.

### Poster sessions

**August 2023**, “Canonical Correlation Analysis in high dimensions with structured regularization”, *New Researchers Conference*, Toronto, Canada.

**June 2022**, “Statistical curve models for inferring 3D chromatin architecture”, *Statistics in the Big Data Era at Simons Institute*, Berkeley, USA.

**April 2022**, “Canonical Correlation Analysis in high dimensions with structured regularization”, *Data Science Inaugural Conference at Stanford University*, Stanford, USA.

**June 2021**, “Canonical Correlation Analysis in high dimensions with structured regularization”, *DBDS Scientific Retreat at Stanford University*, Stanford, USA.

**September 2019**, “Chromatin reconstruction via Weighted Principal Curves”, *Machine Learning Summer School at Skolkovo Institute of Science & Technology*, Moscow, Russia.

### Summer schools

**2020**, *Summer School of Machine Learning at Skolkovo Institute of Science & Technology*, Virtual.

**2019**, *Machine learning Summer School at Skolkovo Institute of Science & Technology*, Moscow, Russia.

**2016**, *Advanced Statistics and Data Mining Summer School at Polytechnic University of Madrid*, Madrid, Spain.

## TEACHING

### Instructor

**2023**, Practice in Statistics I (STA 220), Probabilities Everywhere (STA 198), *University of Toronto*.

**2021**, Introduction to R for UGs (STATS 32), *Stanford University*.

**2016 – 2017**, Probability Theory and Statistical Practicum, *Moscow State University*.

### Teaching Assistant

**2017 – 2022**, Data Science (STATS 101), Statistical Methods in Engineering and the Physical Sciences (STATS 110), Biostatistics (STATS 141), Data Mining and Analysis (STATS 202), Introduction to Regression Models and Analysis of Variance (STATS 203), Theory of Statistics II (STATS 300B), Modern Applied Statistics: Data Science (STATS 315B), Applied Statistics I (STATS 305A), Applied Multivariate Analysis (STATS 206), *Stanford University*.

## QUALIFICATIONS

### Computing skills

R, Python.

### Languages

Russian (native), English (fluent), French (intermediate).