




# Lorenzo De Stefani, Ph.D.

---

|                     |  |  |  |  |  |
|---------------------|--|--|--|--|--|
| CONTACT INFORMATION | Department of Computer Science<br>Brown University<br>115 Waterman Street<br>Providence, RI 02906  | +1 (401) 474-7281<br>lorenzo_destefani@brown.edu<br><a href="https://vivo.brown.edu/display/ld9">https://vivo.brown.edu/display/ld9</a>  |  |  |  |
| CITIZENSHIP         | Italy, US permanent resident   |  |  |  |  |
| CURRENT POSITION    |  <b>Brown University</b> , Providence, USA<br><i>Lecturer of Computer Science</i> | <b>2020–current</b>  |  |  |  |
| EDUCATION           |  <b>Brown University</b> , Providence, USA<br><i>Ph.D. in Computer Science</i>    | <b>2014–2020</b> <ul style="list-style-type: none"><li>• Dissertation: Probabilistic approaches for rigorous and efficient analysis of statistical properties of large datasets</li><li>• Advisor: Eli Upfal</li></ul>  <b>University of Padova</b> , Padova, Italy<br><i>Ph.D. in Information Engineering</i>  | <b>2013–2016</b> <ul style="list-style-type: none"><li>• Dissertation: On space constrained computations</li><li>• Advisor: Gianfranco Bilardi</li></ul> <i>M.Sc. in Computer Engineering, Grade: 110/110, cum laude</i> | <b>2009–2012</b> <ul style="list-style-type: none"><li>• Thesis: On the space complexity of DAG computations</li><li>• Advisor: Gianfranco Bilardi</li></ul> <i>B.Sc. in Computer Engineering, Grade: 110/110, cum laude</i> | <b>2006–2009</b> <ul style="list-style-type: none"><li>• Thesis (translated from Italian): Study on the classification of documents retrieved by information retrieval systems using linear regression analysis</li><li>• Advisor: Giorgio Maria Di Nunzio</li></ul> |
| RESEARCH EXPERIENCE | <b>Graduate Research Assistant</b><br>Upfal Group (BIGDATA), Computer Science, Brown University, Providence, RI  | <b>Sep. 2014–2020</b> <ul style="list-style-type: none"><li>• <i>Statistical learning approaches to frequentist multiple hypothesis testing</i>: Developed methods for multiple hypothesis correction building on results from statistical learning theory improving over classical frequentist methods. Our procedures build on uniform convergence bounds based on Rademacher Complexity and Vapnik-Chervonenkis (VC) dimension. Currently under submission.</li><li>• <i>Statistical learning approaches to adaptive data analysis</i>: Developed efficient and rigorous procedure for controlling the accumulated error in adaptive data analysis building on results from statistical learning theory. Developed novel techniques for Rademacher Complexity estimation based on applications of the Martingale Central Limit theorem and Bernstein's inequality for martingales [2].</li><li>• <i>Rigorous statistical approaches to visual data representation and visualization recommendation</i>: Developed a procedure which allows for adaptive visual exploration while controlling the <i>Marginal False Discovery Rate</i> (mFDR) [7,8,9] building on “<i>alpha investing</i>” testing paradigm. Developed alternative technique with stronger FWER guarantees with extension of the scope of the control to the validation of recommendations of “<i>interesting</i>” visualizations building on use statistical learning analysis tools based on Vapnik-Chervonenkis (VC) dimension [3].</li></ul> |  |  |  |

- *Sub-graph counting in dynamic graph streams*: Developed *Triest*, a suite of one-pass streaming algorithms based on reservoir sampling and the random pairing sampling schemes, which yield high-quality unbiased estimates of the number of triangle motifs in fully dynamic edges streams while using a small local memory. This work was awarded the *Best Student Paper award* for the research track at the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16) [10,13]. Also developed *Tiered Sampling* [5], an extension of the *Triest* paradigm which employs multiple reservoir sample tiers in order to accurately estimate the count of rare and complex patterns such as 4- and 5-cliques.
- *Reconstructing hidden permutations from a sample*: Developed a variant of the *Mallows model* where the distribution is a function of the **Average-Precision (AP)** correlation statistic. We presented a generative model for constructing samples from this distribution, and we developed an efficient algorithm that provably computes an asymptotically unbiased estimate of the center permutation and a faster algorithm that learns with high probability, the hidden central permutation for a wide-range of the parameters of the model. We complement our theoretical analysis with extensive experiments [10].
- *On the I/O complexity integer multiplication and hybrid algorithms* Further developed the G-flow I/O technique [6] to include the formalization of the concept of “*partial Grigoriev’s flow*,” and used it to obtain the first asymptotically tight lower bound on the I/O complexity of the Toom-Cook fast integer multiplication algorithm. In the same work, we presented a matching upper bound [4]. Presented the first asymptotically tight I/O lower bounds for a general class of hybrid non-stationary, non-uniform algorithms for matrix multiplication [1].

#### Graduate Research Assistant

Jan. 2013–Dec. 2015

Advanced Computing Group (ACG), University of Padova, Padova, Italy

- *On the IO complexity of straight line algorithms*: Developed the “*G-flow*” technique which allows to obtain asymptotically tight lower bounds to the Input/Output (I/O) complexity of recursive straight line algorithms. Obtained, using the G-flow technique, the first asymptotically tight I/O complexity lower bound for Strassen’s fast matrix multiplication algorithm which covers computation schedules which allow for multiple evaluations of intermediate results [6].
- *Analysis of the space complexity of DAG computations*: Studied the limits of the *Marking rule technique* by Bilardi et al. Introduced the concept of a “*visit of a Directed Acyclic Graph (DAG)*” and we proved upper bounds for the space requirement of such visits [1, Ph.D. Thesis, University of Padova 2016]. Proposed alternative proof of the general upper bound on space requirement for DAG computations by Hopcroft, Paul, and Valiant [1].
- *Fault resilient algorithms*: Extended *Faulty RAM* model to non-constant safe memory. Developed and analyzed Resilient Mergesort algorithm and Priority Queue data structure in the new model [14].

#### TEACHING EXPERIENCE

*Lecturer of Computer Science*

2020 - Present

Department of Computer Science, Brown University, Providence (RI) USA.

- Data Science (CSCI 1951A) Spring 2021 - Present
- Theory of Computation (CSCI 1010) Fall 2020 - Present
- Design and Analysis of Algorithms (CSCI 1570) Fall 2020 - Present

*Instructor*

Fall 2019

Department of Computer Science, Brown University, Providence (RI), USA

- Theory of Computation (CSCI 1010)

*Co-Instructor*

Fall 2018

Computer Science, Brown University, Providence, RI

- Co-instructed, with Prof. Eli Upfal, “Probability for Computing and Data Analysis” (CSCI 1450)

Graduate Teaching Assistant

Spring 2017-18

Computer Science, Brown University, Providence, RI

- Assisted Prof. Eli Upfal in the instruction of Probability for Computing (CSCI 1540, CSCI 2450)

Graduate Teaching Assistant

July 2016

São Paulo Summer School on Advanced Algorithms, São Paulo, Brazil

- Assisted Prof. Eli Upfal in the instruction of “Sample Complexity and Uniform Convergence”

Graduate Teaching Assistant

A.A. 2013–2014, 2014–2015

Information Engineering, University of Padova, Padova, Italy

- Assisted Prof. Gianfranco Bilardi in the instruction of “Parallel Computing”

AWARDS

- *KDD Best Student Paper Award, Research Track*, San Francisco , CA, 2016.

GRANTS AND FELLOWSHIPS

- *SODA Student Travel Award*, San Diego, CA, 2019.
- *KDD Student Travel Award*, San Francisco , CA, 2016.
- *São Paulo Summer School on Advanced Algorithms Scholarship*, São Paulo, Brazil, 2016.
- *Brown University Graduate Fellowship*, Providence, RI.
- *University of Padova Graduate Fellowship*, admitted with highest score among all applicants, Padova, Italy, 2013.

INVITED TALKS

- **Reconstructing Hidden Permutations Using the Average-Precision (AP) Correlation Statistic.** *Selected for Talk and Poster presentation at the 11th Annual Machine Learning Symposium.* New York Academy of Science, New York (NY), USA. Mar. 3, 2017.
- **Counting Local and Global Triangles in Fully-dynamic Streams with Fixed Memory Size.** *Department of Computer Science, Boston University.* Boston (MA), USA. Sep. 30, 2016.
- **Counting Local and Global Triangles in Fully-dynamic Streams with Fixed Memory Size.** *10th Workshop on Scalable Approaches to High Performance and High Productivity Computing (ScalPerf).* Bertinoro Center for Informatics, Bertinoro, Italy. Sep. 26, 2015.

CONFERENCE PROCEEDINGS

Authors are listed in alphabetic order unless otherwise stated. (\*) equal contributions.

1. G.Bilardi and **L. De Stefani: The DAG Visits technique for Pebbling and I/O Lower Bounds.** *Proceedings of the 42nd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2022).* IIT Madras, India. Dec. 2022.
2. **L. De Stefani: Brief Announcement: The I/O Complexity of Sequential and Parallel Hybrid Integer Multiplication Algorithms.** *Proceedings of the 34th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA 2022).* Philadelphia (PA) USA. Jul. 2022.
3. **L. De Stefani: The I/O complexity of hybrid algorithms for square matrix multiplication.** *Proceedings of the 30th International Symposium on Algorithms and Computation (ISAAC 2019).* Shanghai China. Dec. 2019.
4. **L. De Stefani and E. Upfal: A Rademacher Complexity Based Method for Controlling Power and Confidence Level in Adaptive Statistical Analysis.** *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA).* Washington DC, USA. October. 2019.

5. **L. De Stefani**, L. F. Spiegelberg, T. Kraska and E. Upfal: **VizCertify: A framework for secure data exploration via visual representation.** *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. Washington DC, USA. October. 2019.
  6. G. Bilardi and **L. De Stefani**: **The I/O complexity of Toom-Cook Integer Multiplication.** *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*. San Francisco (CA), USA. Jan. 2019.
  7. **L. De Stefani**, E. Terolli and E. Upfal: **Tiered sampling: An efficient method for approximate counting sparse motifs in massive graph streams.** *Proceedings of 5th IEEE International Conference on Big Data (BigData)*. Boston (MA), USA. Dec. 2017.
  8. G. Bilardi and **L. De Stefani**: **The I/O complexity of Strassen’s Matrix Multiplication with Recomputation.** *Proceedings of the 15th biennial Algorithms and Data Structures Symposium (WADS)*. St. John’s (NL), Canada. Aug. 2017.
  9. Z. Zhao(\*), **L. De Stefani**(\*), E. Zraggen, C. Binnig, E. Upfal and T. Kraska: **Controlling False Discoveries During Interactive Data Exploration.** *Proceedings of the 38th ACM SIGMOD International Conference on Management of Data (SIGMOD/PODS)*. Chicago (IL), USA. May 2017
  10. Z. Zhao, E. Zraggen, **L. De Stefani**, C. Binnig, E. Upfal and T. Kraska: **Safe Visual Data Exploration.** *Proceedings of the 38th ACM SIGMOD International Conference on Management of Data (SIGMOD/PODS)*. Chicago (IL), USA. May 2017.
  11. C. Binning, **L. De Stefani**, T. Kraska, E. Upfal, E. Zraggen and Z. Zhao: **Sustainable Insights, or Why Polygamy is Bad for You.** *Proceedings of the 7th biennial Conference on Innovative Data Systems Research (CIDR)*. Chaminade in Santa Cruz (CA), USA. Jan. 2017.
  12. **L. De Stefani**, A. Epasto, M. Riondato, and E. Upfal: **TRIÈST: Counting Local and Global Triangles in Fully Dynamic Streams with Fixed Memory Size.** *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*. **Best Student Paper Award, Research Track.** San Francisco (CA), USA. Jun. 2016.
  13. **L. De Stefani**, A. Epasto, E. Upfal and F. Vandin: **Reconstructing Hidden Permutations Using the Average-Precision (AP) Correlation Statistic.** *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*. Phoenix (AZ), USA. Feb. 2016.
  14. **L. De Stefani**, G. Di Nunzio, G. Vezzano: **A Visualization Tool of Probabilistic Models for Information Access Components.** *Proceedings of the 13th European Conference on Research and Advanced Technology for Digital Libraries (ECDL)*. Corfu, Greece. Sep. 2009.
- JOURNAL  
ARTICLES
15. **L. De Stefani**, E. Terolli and E. Upfal: **Tiered sampling: An efficient method for counting sparse motifs in massive graph streams.** *ACM Transactions on Knowledge Discovery from Data (TKDD)*. Oct.2021.
  16. **L. De Stefani**, A. Epasto, M. Riondato, and E. Upfal: **TRIÈST: Counting Local and Global Triangles in Fully Dynamic Streams with Fixed Memory Size.** *ACM Transactions on Knowledge Discovery from Data (TKDD)*. Aug. 2017.
  17. **L. De Stefani** and F. Silvestri: **Exploiting non-constant safe memory in resilient algorithms and data structures.** *Theoretical Computer Science (TCS)*. Jun. 2015.
- PREPRINTS
18. **L. De Stefani**: **Communication-Optimal Parallel Standard and Karatsuba Integer Multiplication in the Distributed Memory Model.** <https://arxiv.org/abs/2009.14590>, 2020.

19. **L. De Stefani: The I/O complexity of hybrid algorithms for integer multiplication.**  
<https://arxiv.org/abs/1912.08045>, 2019.

SERVICE TO  
SCIENTIFIC  
COMMUNITY

### **Organizing Committee**

Workshop on Scalable Approaches to High Performance and High Productivity Computing (ScalPerf).  
*Web Co-Chair*. Bertinoro Center for Informatics, Italy. 2012 - present.

### **Program Committee**

- The ACM International Conference on Web Search and Data Mining (WSDM) **2023**
- The IEEE International Conference on Data Science and Advanced Analytics (DSAA) **2018-19.**

#### **Journal Reviewing**

ACM Transactions on Knowledge Discovery from Data (TKDD).

#### **Conference Reviewing**

LATIN'20, AAAI ICWSM'17, ACM SIAM SODA'17, ACM WSDM'17, AAAI'16, IEEE/ACM  
ASONAM'16, ACM SIGKDD'16, CM WebSci'16, IEEE/ACM ASONAM'15, ACM SIGKDD'15,  
ACM ICS'13

REFERENCES

### **Gianfranco Bilardi**

*Professor*

*Department of Information Engineering*

*University of Padova*

📍: Via Gradenigo 6B, Padova, 35131, Italy

☎: +39 049 8277952

✉: [bilardi@dei.unipd.it](mailto:bilardi@dei.unipd.it)

### **Roberto Tamassia**

*Plastech Professor of Computer Science and Department Chair*

*Department of Computer Science*

*Brown University*

📍: 115 Waterman St., Providence, RI 02906

☎: (401) 863-7600

✉: [roberto\\_tamassia@brown.edu](mailto:roberto_tamassia@brown.edu)

### **Eli Upfal**

*Rush C. Hawkins Professor of Computer Science*

*Department of Computer Science*

*Brown University*

📍: 115 Waterman St., Providence, RI 02906

☎: (401) 863-7645

✉: [upfal.rec@gmail.com](mailto:upfal.rec@gmail.com)