

# Jeffrey S. Vitter

*Curriculum Vitæ*

*January 2011*

## 1. Personal

Provost and Executive Vice Chancellor  
Roy A. Roberts Distinguished Professor  
The University of Kansas  
Strong Hall, suite 250  
1450 Jayhawk Blvd.  
Lawrence, KS 66045-7518 USA

Web: [www.provost.ku.edu/jsv/](http://www.provost.ku.edu/jsv/)

Email: [jsv @ ku.edu](mailto:jsv@ku.edu)

Tel.: +1 (785) 864-4904

FAX: +1 (785) 864-4463

Asst.: Ms. Linda Bonebrake, [lbonebrake @ ku.edu](mailto:lbonebrake@ku.edu)

Born November 13, 1955 in New Orleans, Louisiana.

Married, two daughters and one son.



## 2. Biography

Jeff Vitter is the provost and executive vice chancellor and the Roy A. Roberts Distinguished Professor at the University of Kansas. His academic home is the Department of Electrical Engineering and Computer Science, and he is a member of the Information and Telecommunication Technology Center. KU comprises roughly 2,450 faculty members, 5,500 staff, 30,000 students, and an annual budget of \$1 billion, and it serves the main campus in Lawrence (with 10 colleges and schools), the Edwards campus in Overland Park, Kansas, and the Medical Center campuses in Kansas City, Wichita, and Salina, Kansas. As provost, Dr. Vitter is the chief academic and operations officer for the Lawrence and Edwards campuses. He oversees strategic planning and implementation, which focus on excellence in four important areas: energizing the educational environment, elevating doctoral education, driving discovery and innovation, and engaging scholarship for public impact. Key enablers include synergy, people, diversity, external support, efficiency, and infrastructure.

Previously Dr. Vitter was on the faculty in the Department of Computer Science and Engineering at Texas A&M University in College Station, Texas. From 2008–2009, he served at Texas A&M as provost and executive vice president for academics, where he had the responsibility of chief academic officer for a university of roughly 2,700 faculty members, 5,500 staff, 48,000 students, and an annual budget of \$1.2 billion, comprising 10 colleges and schools. In addition, he oversaw the academic mission of Texas A&M University in Doha, Qatar. In collaboration with deans and faculty, Dr. Vitter successfully launched a number of important recruiting efforts and far-reaching faculty initiatives, including those dealing with faculty start-up allocations, multidisciplinary priorities, balanced scorecard reviews and recognition, and diversity. Most significantly, he initiated and led the campus-wide development of the university's Academic Master Plan.

From 2002–2008, Dr. Vitter served as the Frederick L. Hovde Dean of the College of Science and Professor of Computer Science at Purdue University in West Lafayette, Indiana. As dean, he was the chief academic officer and administrator of the College of Science, responsible for overseeing the discovery, learning, engagement, and diversity activities of the college's seven academic departments. In approximate terms, the college comprised 325 faculty, 550 staff, 1,000 graduate students, and 2,800 undergraduate majors, with an annual budget of \$130 million. College courses accounted for about a quarter of the university's 1 million student credit hours. Dr. Vitter led the collaborative development of two college strategic plans, which established a dual focus of excellence in the core departments as well as in multidisciplinary collaborations. The college grew by two major new buildings and 60 faculty members, several hired under the innovative COALESCE faculty program targeting global priorities. He launched a comprehensive study of the undergraduate program, which resulted in an innovative outcomes-based college curriculum approved by the faculty and

implemented in 2007. He proposed and collaboratively piloted the LEAD program for student diversity awareness, subsequently adopted by the university. He initiated the Science Business Partners Program and nurtured many successful funding efforts. Several programs in the college are ranked nationally.

From 1993–2002, Dr. Vitter held a distinguished professorship at Duke University in Durham, North Carolina, where he was the Gilbert, Louis, and Edward Lehrman Professor. He served at Duke as chair of the Department of Computer Science in the College of Arts and Sciences from 1993–2001 and as co-director and a founding member of Duke’s Center for Geometric and Biological Computing from 1997–2002. As chair, he led the department to significant improvements in stature — characterized by a top-20 ranking, stellar faculty hires, dynamic strategic plans, a departmental culture of inclusiveness, comprehensive curriculum redesign, administrative reorganization, substantial increases in both the undergraduate and graduate programs, creation of a successful industry partners program, and rise in sponsored research expenditures to 250%. From 1980–1992, Dr. Vitter progressed through the faculty ranks and served in various leadership roles in the Department of Computer Science at Brown University in Providence, Rhode Island.

His educational degrees include a B.S. with highest honors in mathematics in 1977 from the University of Notre Dame in Notre Dame, Indiana; a Ph.D. in computer science under Don Knuth in 1980 from Stanford University in Stanford, California; and an M.B.A. in 2002 from the Fuqua School of Business at Duke University. His hometown is New Orleans, Louisiana (as everyone who knows him knows!).

Dr. Vitter is a member of the Board of Advisors for the School of Science and Engineering at Tulane University in New Orleans. From 2000–2009, Dr. Vitter served on the Board of Directors of the Computing Research Association (CRA), where he continues to co-chair the Government Affairs Committee. He has served as chair of ACM SIGACT, the Special Interest Group on Algorithms and Computation Theory of the world’s largest computer professional organization, the Association for Computing Machinery. He has served on the executive council of the European Association for Theoretical Computer Science, as well as on various visiting and review committees. Sabbatical sites have included the Mathematical Sciences Research Institute in Berkeley; INRIA in Rocquencourt, France; Ecole Normale Supérieure in Paris; Bell Laboratories in Murray Hill, New Jersey; and INRIA in Sophia Antipolis, France.

Dr. Vitter has been named a Guggenheim Foundation Fellow, a Fellow of the Association for Computing Machinery (ACM), a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), a Fellow of the American Association for the Advancement of Science (AAAS), a National Science Foundation Presidential Young Investigator, a Fulbright Scholar, and an IBM Faculty Development Awardee. He has over 280 book, journal, conference, and patent publications reflecting his research interests described below, and he has given over 180 presentations worldwide. His Google Scholar h-index is 60, and he is an ISI highly cited researcher. His book *Algorithms and Data Structures for External Memory* (Now Publishers, 2008) covers the I/O algorithms field that he helped found. He coauthored the books *Efficient Algorithms for MPEG Video Compression* (Wiley & Sons, 2002) and *Design and Analysis of Coalesced Hashing* (Oxford University Press, 1987). He is coeditor of the collections *External Memory Algorithms* and *Algorithm Engineering*. His editorial board memberships have included *Algorithmica*, *Communications of the ACM*, *IEEE Transactions on Computers*, *Theory of Computing Systems*, and *SIAM Journal on Computing*; in addition, he has edited several special issues and serves on the steering committee of *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. He has consulted widely and is co-holder of patents in the areas of external sorting, parallel I/O, prediction, and approximate data structures. He proposed the concept and participated in the design of what has become the Purdue University Research Expertise database (PURE) and the Indiana Database for University Research Expertise (INDURE), [www.indure.org](http://www.indure.org).

### 3. Research Interests

In his research, Jeff Vitter seeks to exploit the rich interdependence between mathematical computing theory and practice. Dr. Vitter has pioneered the development of important subfields dealing with massive data. He is perhaps best known as a founder of the field of external memory algorithms, which focuses on alleviating the I/O bottleneck between fast internal memory and slow external storage (such as disk). The goal is to design algorithms that exploit locality of reference and parallelism in order to reduce I/O costs, which is important in a variety of data-intensive applications. His recent book, *Algorithms and Data Structures for External Memory*, serves as a reference for the field. He has developed paradigms in several domains for efficient algorithms using external memory and hierarchical memory. His approach for utilizing parallel disks (in which communication with each disk can occur simultaneously) using the notion of read/write duality

has led to state-of-the-art methods for sorting. He has contributed to algorithm engineering via the TPIE system (Transparent Parallel I/O programming Environment) developed by a former student.

A second key area of massive data where Dr. Vitter plays a leadership role is compressed data structures. The goal is to operate directly upon compressed representations of data, yet still achieve fast response time. The wavelet tree data structure he co-developed (not to confuse with wavelets discussed two paragraphs below) is an elegant structure for coding sequences of characters from a multicharacter alphabet; it has become a key component in modern indexing and compression. Until this century, fast data structures for text indexing (such as suffix trees and suffix arrays) required much more space than the data being indexed! Based upon a recursive decomposition of the suffix array, Dr. Vitter and colleagues invented the compressed suffix array, which is substantially smaller — the first fast index provably shown to use only linear space, and then later the first ever whose size per character was proven to be asymptotic (i.e., with constant of proportionality 1) to the higher-order entropy of the text. The index can even reconstruct the original text in a random access manner, and thus the original text can be discarded. The net effect is that the text can be completely replaced by an index structure that has the size of compressed text but can be queried quickly.

In a third aspect of massive data, Dr. Vitter is a leading figure in the data compression community, noted for his analytical bent and influence. He has done fundamental work on data compression for text, images, and video. A provably efficient algorithm for adaptive Huffman coding bears his name. With a former student, Dr. Vitter developed and analyzed fast and practical methods for arithmetic coding. They invented the FELICS algorithm for lossless image compression; it was subsequently implemented in hardware as part of NASA's Mars Reconnaissance Orbiter. It introduced a low-cost prediction framework that influenced algorithms ultimately adopted into the Lossless JPEG standard. In video compression, Dr. Vitter and group proposed the paradigm of minimizing the combined measure of rate plus distortion to significantly improve motion estimation coding; this rate-distortion optimization has been incorporated into the H.264/MPEG-4 AVC standard's reference encoder, used widely in the computing and communications industry.

Fourth, Dr. Vitter and collaborators were the first in the database and systems communities to apply wavelets and compression techniques as key tools for summarizing, approximating, and predicting data. Wavelets have since become heavily used in database optimization, data warehousing, data streams, image processing, and data mining. For his work on wavelets for approximating high-dimensional aggregates, he and his coauthor were the recipients of the 2009 ACM SIGMOD Test of Time award, which recognizes the SIGMOD paper from 10 years earlier that has had the most impact in the following decade in terms of research, products, and methodology. Dr. Vitter has co-developed novel machine learning and prediction mechanisms based upon data compression, using the principle that the more compressible a sequence is, the more predictable it is. His universal prediction algorithms for online prefetching are provably asymptotically optimal (i.e., with constant of proportionality 1). They predict as well as special-purpose methods tuned to the characteristics of the sequence. His learning work includes algorithms for prefetching, caching, data streams, database query optimization, data mining, and power management in mobile computers.

Beginning with his thesis on coalesced hashing, a search method used widely in practice, Dr. Vitter has made many contributions to the analysis of algorithms, using mathematical analysis and asymptotics to derive precise estimates for resource requirements. He has also done much work involving randomized, parallel, and incremental algorithms for a variety of problems in computational geometry, combinatorial optimization, graphics, random sampling, and random variate generation.

## 4. Education

B.S. (with highest honors), mathematics, University of Notre Dame, 1977.

Ph.D., computer science, Stanford University, 1980.

Dissertation topic: *Analysis of Coalesced Hashing* (supervised by Donald E. Knuth).

M.B.A., Duke University, 2002.

## 5. Professional Appointments

1976–1977     Assistant computer performance analyst, Standard Oil Company of California, San Francisco, CA (summers).

1977–1980     Research and teaching assistant, Stanford University.

1979 Teaching Fellow, Stanford University.

1980–1985 Assistant professor, Brown University.

1985–1988 Associate professor (with tenure), Brown University.

1986 Member, Mathematical Sciences Research Institute, Berkeley, CA.

1986–1987 Visiting professor, Ecole Normale Supérieure, Paris, France.

1986–1987 Directeur scientifique, Institut National de Recherche en Informatique et en Automatique (INRIA), Rocquencourt, France.

1987 Lecturer, 2nd Asian School on Computer Science, Bangkok, Thailand.

1988 Visiting professor (short-term), Ecole Normale Supérieure, Paris, France.

1988–1992 Professor, Brown University.

1989 Visiting professor (short-term), Ecole Normale Supérieure, Paris, France.

1989–2000 Associate member, Center of Excellence in Space Data and Information Sciences (CESDIS).

1990–2006 Visiting professor (short-term) and/or adjunct professor, Tulane University.

1993 Professor, Duke University.

1993–2001 Chair of Department of Computer Science, Duke University.

1993–2002 Gilbert, Louis, and Edward Lehrman Professor (distinguished professorship), Duke University.

1993–1999 Adjunct professor, Brown University.

1995–2002 Founding member of Center for Geometric Computing, Brown University.

1997–2002 Founding member and co-director of Center for Geometric and Biological Computing (formerly Center for Geometric Computing), Duke University.

1998 Visiting professor (short-term), University of Aarhus, Århus, Denmark.

1998–1999 Professeur invité, Institut National de Recherche en Informatique et en Automatique (INRIA), Sophia Antipolis, France.

2002–present Adjunct professor, Duke University.

2002–2008 Frederick L. Hovde Dean of the College of Science and Professor, Purdue University.

2008–present Adjunct professor, Purdue University.

2008–2009 Provost and executive vice president for academics, Texas A&M University.

2008–2010 Professor, Texas A&M University.

2010–present Provost and executive vice chancellor, the University of Kansas.

2010 Professor, the University of Kansas.

2011–present Roy A. Roberts Distinguished Professor, the University of Kansas.

## 6. Recent Administrative Assignments

### At Purdue University

2002–2008 Frederick L. Hovde Dean, College of Science.

2002–2008 Convener of College of Science Faculty Council.

2002–2008 Chair of College of Science Promotions Committee.

2002–2008 University Promotions Committee.

2002–2008 Chair of policy board for Center for Education and Research in Information Assurance and Security (CERIAS).

2002 Presidential task force to draft a university statement on integrity.

2002–2003 Chair of College of Science Strategic Planning Committee.

2002–2003 Chair of search committee for associate dean of College of Science.

2002–2005 Executive committee, Environmental Sciences and Engineering Institute.

2002–2008 Executive committee, Computing Research Institute.  
2002–2006 Vice-chair of policy board, Center for Sensing Science and Technology.  
2003–2005 Search committee for vice president for research.  
2003–2004 Search committee for IUPUI dean of science.  
2003–2008 Representative to the University Corporation for Atmospheric Research (UCAR).  
2004 Chair of search committee for associate dean of College of Science.  
2003–2005 College of Science Undergraduate Education Task Force.  
2003–2008 Proposer of concept and participant in design of what has become the Purdue University Research Expertise database (PURE) and the Indiana Database for University Research Expertise (INDURE), www.indure.org.  
2004–2008 Enterprise Resource Planning (ERP) Executive Oversight Committee.  
2004–2008 Executive committee, Center for Wireless Systems and Applications.  
2004–2007 Search committee for director of Purdue Cancer Center.  
2005–2008 Discovery Park Liaison Committee.  
2005–2006 Co-chair of planning committee, Purdue Healthcare CEO Summit: A Healthcare Delivery System for the Next Generation.  
2005–2006 Chair of search committee for dean of the College of Liberal Arts.  
2006 Chair of executive committee, Science Laureates Program.  
2006–2007 Purdue Libraries area committee.  
2006–2008 Chair of policy board, Center for Sensing Science and Technology.  
2007 Co-chair of executive committee, Science Journalism Laureates Program.  
2007–2008 Chair of College of Science Strategic Planning Committee.

### **At Texas A&M University**

2008-2009 Provost and executive vice president for academics.  
2008–2009 President’s Cabinet.  
2008–2009 Convener of Provost’s Administrative Team.  
2008–2009 Chair of Academic Master Planning Committee.  
2009 Board of Directors, Research Valley Partnership.  
2009 Board of Trustees, Texas A&M Research Foundation.  
2009 Co-chair of the Task Force on Enlightened and Shared Governance.

### **At the University of Kansas**

2010–present Provost and executive vice chancellor.  
2010–present Chancellor’s Cabinet.  
2010–present Convener of Academic Council.  
2010–present Chairman of the Board, KU Center for Technology Commercialization.  
2010–present Board of Directors, Kansas Athletics, Inc.  
2010–present Chancellor’s Advisory Committee on Intercollegiate Athletics.  
2010–present Supervisory committee, Institute for Advancing Medical Innovation.  
2011–present Higher Education Task Force, K.C. Area Development Council.

## **7. Honors and Awards**

1973–1977 Notre Dame Scholar, University of Notre Dame.  
1977 General Electric Mathematics Major Award.

- 1977            Graduated with highest honors, University of Notre Dame.
- 1977–present   Phi Beta Kappa.
- 1977–1980    National Science Foundation Graduate Fellow, Stanford University.
- 1983–present   Sigma Xi.
- 1984–1988    IBM Faculty Development Award.
- 1985–1991    National Science Foundation Presidential Young Investigator Award.
- 1986–present   Listed in several national and international Who’s Who publications.
- 1986           Honorary A.M. degree, Brown University.
- 1986–present   Fellow of John Simon Guggenheim Memorial Foundation.
- 1993–present   Fellow of Institute of Electrical and Electronics Engineers (IEEE), for contributions to the theory of sorting and searching and to the design and analysis of computer algorithms.
- 1996–present   Fellow of Association for Computing Machinery (ACM), for contributions to the theory of information storage and retrieval and to the design and mathematical analysis of computer algorithms.
- 1997           Recognition of Service Award, Association for Computing Machinery (ACM).
- 1998           Fulbright Scholar.
- 1999           Medal of the University of Helsinki, Helsinki, Finland.
- 2001           Recognition of Service Award, Association for Computing Machinery (ACM).
- 2002           Graduated as Fuqua Scholar, Fuqua School of Business, Duke University.
- 2009           SIGMOD Test of Time Award, awarded by the Association for Computing Machinery (ACM) Special Interest Group for Management of Data (SIGMOD) to the authors of the most impactful SIGMOD paper from 10 years earlier, “Approximate Computation of Multidimensional Aggregates of Sparse Data Using Wavelets,” by J. S. Vitter and M. Wang.
- 2009–present   Fellow of the American Association for the Advancement of Science (AAAS), for distinguished contributions to the design and analysis of efficient computer algorithms and data structures, particularly those involving massive amounts of data.

## 8. Professional Societies

- 1977–present   Phi Beta Kappa.
- 1979–present   Association for Computing Machinery (ACM) (currently Fellow) and its Special Interest Group on Algorithms and Computation Theory (ACM SIGACT).
- 1980–present   Institute of Electrical and Electronics Engineers (IEEE) (currently Fellow) and IEEE Computer Society.
- 1983–present   Sigma Xi.
- 1984–present   European Association for Theoretical Computer Science (EATCS), with occasional lapses.
- 1986–present   John Simon Guggenheim Memorial Foundation (Fellow).
- 2003–present   American Association for the Advancement of Science (AAAS) (currently Fellow).

## 9. Consultancies

- Alston & Bird, Washington, DC.
- AT&T Labs–Research, Florham Park, NJ.
- Burman, Critton, Luttier & Coleman, West Palm Beach, FL.
- Center for Computing Sciences (formerly Supercomputing Research Center), Bowie, MD.
- Fish & Richardson, Atlanta, GA and San Diego, CA.
- IBM Academic Information Systems, Stamford, CT.
- IBM Palo Alto Scientific Center, Palo Alto, CA.

Institute for Defense Analyses, Alexandria, VA.  
Knowledge Engineering, Inc., Cambridge, MA.  
Lucent Technologies, Bell Laboratories, Murray Hill, NJ.  
Milbank, Tweed, Hadley & McCloy, New York, NY.  
Universities Space Research Association, Columbia, MD.  
Xerox PARC, Palo Alto, CA.

## 10. Service to the Profession

1979–present Referee for *ACM Computing Surveys*, *ACM Conference on Computational Learning Theory (COLT)*, *ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, *ACM International Conference on Mobile Computing and Networking (MobiCom)*, *ACM SIGMETRICS Conference on Measurement & Modeling of Computer Systems (SIGMETRICS)*, *ACM SIGMOD International Conference on Management of Data (SIGMOD)*, *ACM Symposium on Applied Computing (SAC)*, *ACM Symposium on Principles of Database Systems (PODS)*, *ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, *ACM Symposium on Theory of Computing (STOC)*, *ACM Transactions on Database Systems*, *ACM Transactions on Information Systems*, *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, *ACM-IEEE Workshop on I/O in Parallel And Distributed Systems (IOPADS)*, *Acta Informatica*, *Advances in Computing Research*, *Algorithmica*, *CGC Workshop on Computational Geometry*, *Colloquium on Trees in Algebra and Programming (CAAP)*, *Communications of the ACM*, *The Computer Journal*, *Constraint Databases*, *European Conference on Parallel Computing (Euro-Par)*, *Fall Joint Computer Conference (FJCC)*, *Hawaii International Conference on System Sciences (HICSS)*, *IASTED International Conference on Databases and Applications (DBA)*, *IBM Journal of Research and Development*, *IEEE Data Compression Conference (DCC)*, *IEEE International Conference on Image Processing (ICIP)*, *IEEE International Conference on Tools with Artificial Intelligence (TAI)*, *IEEE Symposium on Foundations of Computer Science (FOCS)*, *IEEE Symposium on Parallel and Distributed Processing (SPDP)*, *IEEE Transactions on Circuits and Systems, Part I*, *IEEE Transactions on Circuits and Systems for Video Technology*, *IEEE Transactions on Communications*, *IEEE Transactions on Computers*, *IEEE Transactions on Knowledge and Data Engineering*, *IEEE Transactions on Image Processing*, *IEEE Transactions on Parallel and Distributed Systems*, *Information Processing Letters*, *International Colloquium on Automata, Languages, and Programming (ICALP)*, *International Computing and Combinatorics Conference (COCOON)*, *International Conference on Compression and Complexity of Sequences (SEQUENCES)*, *International Conference on Database Theory (ICDT)*, *International Conference on High Performance Computing*, *International Conference on Parallel Processing (ICPP)*, *International Conference on String Processing and Information Retrieval (SPIRE)*, *International Conference on Very Large Databases (VLDB)*, *International Journal of Imaging Systems and Technology*, *International Journal of Parallel Programming*, *International Parallel Processing Symposium (IPPS)*, *International Symposium on Algorithms and Computation (ISAAC)*, *International Symposium on Spatial and Temporal Databases (SSTD)*, *International Workshop on Data Warehousing and Knowledge Discovery from Sensors and Streams (DKSS)*, *International Workshop on Knowledge Discovery in Data Streams (ICML)*, *International Workshop on Spatio-Temporal Database Management (STDBM)*, *International Workshop on Bioinformatics in Data Mining (BIOKDD)*, *Journal of the ACM*, *Journal of Algorithms*, *Journal of Computer and System Sciences*, *Journal of Parallel and Distributed Computing*, *Mathematical Programming*, *MIT-Brown Conference on Parallelism and VLSI*, *Optical Engineering*, *Parallel Computing Journal*, *Physical Review*, *Real-Time Systems*, *Scandinavian Workshop on Algorithms and Theory (SWAT)*, *Signal Processing*, *Software — Practice and Experience*, *SIAM Journal on Algebraic and Discrete Methods*, *SIAM Journal on Computing*, *Symposium on Theoretical Aspects of Computer Science (STACS)*, *Theoretical Computer Science*, *Theory of Computing Systems (formerly Mathematical Systems Theory:*

- An International Journal on Mathematical Computing Theory*), *VLDB — International Journal on Very Large Databases*, *Workshop on Algorithm Engineering (WAE)*, *Workshop on Algorithms and Data Structures (WADS)*, *Workshop on Massive Data Algorithmics*, *Workshop on Modeling and Specification of I/O (MSIO)*, *Workshop on Randomized Parallel Computing*.
- 1980–present Technical committee for the annual IEEE Symposium on Foundations of Computer Science.
- 1981–present Reviewer of research proposals and/or panel member for various agencies, including the National Science Foundation, Air Force Office of Scientific Research, Army Research Office, NASA, and Natural Sciences and Engineering Research Council of Canada.
- 1987–1991 Executive committee, ACM SIGACT, Special Interest Group on Automata and Computability Theory, Association for Computing Machinery.
- 1991–1997 Vice-chair of executive committee, ACM SIGACT, Special Interest Group on Algorithms and Computation Theory (formerly Automata and Computability Theory), Association for Computing Machinery.
- 1995 Board of Visitors and review committee of Mathematical and Computer Science Division, Army Research Office.
- 1996 External review committee of Department of Computer Science, Georgetown University.
- 1997–1999 Graduate Record Examination Computer Science Committee.
- 1997–2001 Chair of executive committee, ACM SIGACT, Special Interest Group on Algorithms and Computation Theory, Association for Computing Machinery.
- 1997–2001 SIG Governing Board, Association for Computing Machinery.
- 1997–2001 Executive council, EATCS, European Association for Theoretical Computer Science.
- 1999 Coauthor of NSF-sponsored report, “Challenges for Theory of Computing.”
- 1999–2002 Steering committee, Workshop on Algorithm Engineering.
- 2000–2009 Board of Directors, Computing Research Association (CRA).
- 2001–2005 Past chair of executive committee, ACM SIGACT, Special Interest Group on Algorithms and Computation Theory, Association for Computing Machinery.
- 2001–present Co-chair of Government Affairs Committee, Computing Research Association (CRA).
- 2003–2004 External review committee, Institute of Information Sciences, Academia Sinica, Taipei, Taiwan (Republic of China).
- 2003–2008 Proposer of concept and participant in design of what has become the Purdue University Research Expertise database (PURE) and the Indiana Database for University Research Expertise (INDURE), [www.indure.org](http://www.indure.org).
- 2005–2009 Computing Research Funding Task Force, Computing Research Association.
- 2006–present Board of Advisors, School of Science and Engineering, Tulane University. Planning and Operations Committee (2006–2009), Technology Transfer and Business Development Committee (2009–present).
- 2007 External programme review committee, Science Foundation Ireland, Dublin, Ireland.
- 2008–2009 Visiting committee, Institut National de Recherche en Informatique et en Automatique (INRIA), Rocquencourt, France.
- 2009–present Board of Advisors, Center for Massive Data Algorithmics (MADALGO), Danish National Research Foundation, Århus, Denmark.
- 2010 External review committee of Department of Computer Science, Yale University.
- 2010–present Computer Science Task Force, School of Science and Engineering, Tulane University.
- 2010–present Task Force on Graduate Program Disclosure, Council on Academic Affairs, Association of Public and Land-grant Universities (APLU).
- 2010–present Executive committee, Council on Academic Affairs, Association of Public and Land-grant Universities (APLU).

## 11. Editorial Responsibilities

- 1985 Guest co-editor, special issue of *IEEE Transactions on Computers* on the subject of sorting, **C-34**(4), April 1985.
- 1987–1991 Editor, *IEEE Transactions on Computers*.
- 1988 Guest editor, special issues of *Algorithmica* on the subject of parallel and distributed computing, Part I, **3**(1), and Part II, **3**(3), 1988.
- 1988–1995 Editor, *Communications of the ACM*.
- 1989–1997 Editor, *SIAM Journal on Computing*.
- 1991–present Editor, *Theory of Computing Systems* (previously *Mathematical Systems Theory: An International Journal on Mathematical Computing Theory*).
- 1994 Guest editor, special double issue of *Algorithmica* on the subject of Large-Scale Memories, **12**(2–3), 1994.
- 1994–2011 Editor, *Algorithmica*.
- 1996 Co-editor, working group report on strategic directions in storage I/O for large-scale computing, *Computing Surveys*, **28** (4), December 1996.
- 1999 Co-editor, *External Memory Algorithms and Visualization*, DIMACS Series on Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, 1999.
- 1999 Co-editor, *Algorithm Engineering*, Lecture Notes in Computer Science, **1668**, Springer, Berlin, Germany, 1999.
- 2000 Guest co-editor, special issue of *ACM Journal on Experimental Algorithmics* on the subject of algorithm engineering, **5**, 2000.
- 2003 Co-editor, *Proceedings of the ACM Workshop on Paradigms of Computing and Knowledge (PCK-50)*, June 2003.
- 2012–present Steering committee, *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*.

## 12. Conference Responsibilities since 1990

- 1990 Program committee, 2nd Scandinavian Workshop on Algorithm Theory (SWAT '90), Bergen, Norway, July 1990.  
Member of program committee and session chair, 31st annual IEEE Symposium on Foundations of Computer Science (FOCS '90), St. Louis, MO, October 1990.
- 1991 Member of program committee and session chair, 1991 IEEE Data Compression Conference (DCC '91), Snowbird, UT, April 1991.  
Co-chair, 23rd annual ACM Symposium on Theory of Computing (STOC '91), New Orleans, LA, May 1991.
- 1992 Program committee, 1992 IEEE Data Compression Conference (DCC '92), Snowbird, UT, March 1992.  
Program committee, MIT-Brown Conference on Parallelism and VLSI, Providence, R.I., March 1992.
- 1993 Program committee, 1993 IEEE Data Compression Conference (DCC '93), Snowbird, UT, March 1993.  
Session chair, Dagstuhl-Seminar on Average-Case Analysis of Algorithms, Schloss Dagstuhl, Germany, July 1993.  
Program committee, 3rd International Workshop on Algorithms and Data Structures (WADS '93), Montreal, Canada, August 1993.
- 1994 Program committee, 1994 IEEE Data Compression Conference (DCC '94), Snowbird, UT, March 1994.  
Program committee, 7th annual ACM Conference on Computational Learning Theory (COLT '94), New Brunswick, NJ, July 1994.  
Vice-chair, member of program committee, and session chair, 6th IEEE International Conference on Tools with Artificial Intelligence (TAI '94), New Orleans, LA, November 1994.

- 1995 Co-organizer and co-chair of the program committee, Workshop on Modeling and Specification of I/O (MSIO '95), as part of the 7th IEEE Symposium on Parallel and Distributed Processing (SPDP '95), San Antonio, TX, October 1995.
- Program committee, 1st ACM International Conference on Mobile Computing and Networking (MobiCom '95), November 1995.
- Program committee, 6th annual International Symposium on Algorithms and Computation (ISAAC '95), Cairns, Australia, December 1995.
- 1996 Member of program committee and session chair, 7th annual SIAM/ACM Symposium on Discrete Algorithms (SODA '96), Atlanta, GA, January 1996.
- Program committee, Workshop on Randomized Parallel Computing (WRPC '96), as part of the 1996 International Parallel Processing Symposium (IPPS '96), Honolulu, HI, April 1996.
- Session Chair, 1996 ACM Workshop on Parallel Algorithms (WOPA '96), as part of the 2nd ACM Federated Computer Research Conference, Philadelphia, PA, May 1996.
- Member of program committee and session chair, 4th annual ACM-IEEE Workshop on I/O in Parallel And Distributed Systems (IOPADS '96), as part of the 2nd ACM Federated Computer Research Conference, Philadelphia, PA, May 1996.
- Co-chair of working group on storage I/O issues in large-scale computing, ACM Workshop on Strategic Directions in Computing Research, Massachusetts Institute of Technology, Cambridge, MA, June 1996.
- Member of working group on computational geometry, ACM Workshop on Strategic Directions in Computing Research, Massachusetts Institute of Technology, Cambridge, MA, June 1996.
- Member of program committee and session chair, 1st CGC Workshop on Computational Geometry, Baltimore, MD, October 1996.
- 1997 Member of program committee and session chair, 1997 IEEE Data Compression Conference (DCC '97), Snowbird, UT, March 1997.
- Program committee, Workshop on Randomized Parallel Computing (WRPC '97), as part of the 1997 International Parallel Processing Symposium (IPPS '97), Geneva, Switzerland, April 1997.
- Program committee, 16th annual ACM Symposium on Principles of Database Systems (PODS '97), Tucson, AZ, May 1997.
- Program committee, International Conference on Compression and Complexity of Sequences (SEQUENCES '97), Positano, Italy, June 1997.
- Co-chair, member of the program committee, and session chair, 2nd CGC Workshop on Computational Geometry, Durham, NC, October 1997.
- Program committee, 5th annual ACM Workshop on I/O in Parallel And Distributed Systems (IOPADS '97), San Jose, CA, November 1997.
- 1998 Program committee, 1998 IEEE Data Compression Conference (DCC '98), Snowbird, UT, March–April 1998.
- Co-organizer, Workshop on External Memory Algorithms and Visualization, DIMACS, Rutgers University, New Brunswick, NJ, May 1998.
- Session chair, 17th annual ACM Symposium on Principles of Database Systems (PODS '98), Seattle, WA, June 1998.
- Program committee, 3rd CGC Workshop on Computational Geometry, Providence, RI, October 1998.
- 1999 Program committee, 5th annual International Computing and Combinatorics Conference (COCOON '99), Tokyo, Japan, July 1999.
- Session chair and panel member, 26th annual International Colloquium on Automata, Languages, and Programming (ICALP '99), Prague, Czech Republic, July 1999.
- Co-chair of the program committee, session chair, and panel chair, 3rd Workshop on Algorithm Engineering (WAE '99), London, England, July 1999.
- Program committee, 5th annual ACM International Conference on Knowledge Discovery and Data Mining (KDD '99), San Diego, CA, August 1999.

- Member of program committee and session chair, 4th CGC Workshop on Computational Geometry (CGC '99), Baltimore, MD, October 1999.
- 2000 Member of program committee and session chair, 2000 ACM SIGMOD International Conference on Management of Data (SIGMOD '00), Dallas, TX, May 2000.
- Chair, Workshop on reshaping doctoral education, Computing Research Association Conference 2000 (Snowbird '00), Snowbird, UT, July 2000.
- Organizing committee, International Conference on Theoretical Computer Science (IFIP TCS 2000), Sendai, Japan, August 2000.
- Member of program committee and session chair, 26th International Conference on Very Large Databases (VLDB '00), Cairo, Egypt, September 2000.
- 2001 Session chair, 12th annual SIAM/ACM Symposium on Discrete Algorithms (SODA '01), Washington, DC, January 2001.
- Advisory committee, Atlantic Symposium on Computational Biology and Genome Information Systems and Technology, Durham, NC, March 2001.
- Program committee, International Workshop on Bioinformatics in Data Mining (BIOKDD '01), San Francisco, CA, August 2001.
- 2002 Co-organizer, KNUTHfest 2002: Symposium in Honor of Donald E. Knuth's (1,000,000)<sub>2</sub>th Birthday, Stanford University, Stanford, CA, January 2002.
- Co-organizer and session chair, Workshop on Compression Issues in Next-Generation Network Applications, DIMACS, Rutgers University, New Brunswick, NJ, March 2002.
- Co-chair, workshop on research funding, Computing Research Association Conference 2002 (Snowbird '02), Snowbird, UT, July 2002.
- Program committee, 2002 International Conference on Very Large Databases (VLDB '02), Hong Kong, China, August 2002.
- 2003 Co-organizer, ACM Workshop on Paradigms of Computing and Knowledge (PCK-50): Memorial Workshop on the Occasion of Paris C. Kanellakis' 50th Birthday, as part of the 4th ACM Federated Computer Research Conference, San Diego, CA, June 2003.
- Program committee, 8th International Symposium on Spatial and Temporal Databases (SSTD '03), Santorini island, Greece, July 2003.
- Program committee, 8th International Workshop on Algorithms and Data Structures (WADS '03), Ottawa, Canada, July–August 2003.
- Co-organizer, Purdue Cyber-Infrastructure Workshop, West Lafayette, IN, August 2003.
- Program committee, 2003 IEEE International Conference on Image Processing (ICIP '03), Barcelona, Spain, September 2003.
- 2004 Program committee, Workshop on Analytic Algorithmics and Combinatorics (ANALCO '04), as part of the 15th annual SIAM/ACM Symposium on Discrete Algorithms (SODA '04), New Orleans, LA, January 2004.
- Program committee, 19th ACM Symposium on Applied Computing (SAC '04), Nicosia, Cyprus, March 2004.
- Program committee, 2nd International Workshop on Spatio-Temporal Database Management (STDBM '04), as part of the 30th International Conference on Very Large Databases (VLDB '04), Toronto, Canada, August–September 2004.
- Program committee, First International Workshop on Knowledge Discovery from Data Streams (ICML '04), as part of the 15th European Conference on Machine Learning (ECML '04) Pisa, Italy, September 2004.
- Program committee, 11th International Conference on String Processing and Information Retrieval (SPIRE '04), Padua, Italy, October 2004.
- 2005 Panel chair, undergraduate programs—curricular objectives and access issues, annual meeting of the American Association of Universities Arts & Sciences Deans, Irvine, CA, January 2005.
- Program committee, IASTED International Conference on Databases and Applications (DBA '05), Innsbruck, Austria, February 2005.

- Program committee, 20th ACM Symposium on Applied Computing (SAC '05), Santa Fe, NM, March 2005.
- Program committee, 12th International Conference on String Processing and Information Retrieval (SPIRE '05), Buenos Aires, Argentina, October–November 2005.
- 2006 Program committee, IASTED International Conference on Databases and Applications (DBA '06), Innsbruck, Austria, February 2006.
- Co-chair of planning committee, Purdue Healthcare CEO Summit: A Healthcare Delivery System for the Next Generation, West Lafayette, IN, May 2006.
- Program committee, 3rd International Workshop on Spatio-Temporal Database Management (STDBM '06), as part of the 32nd International Conference on Very Large Databases (VLDB '06), Seoul, South Korea, September 2006.
- 2007 Session chair, Information Theory and Applications Workshop (ITA '07), California Institute for Telecommunications and Information Technology, University of California–San Diego, La Jolla, CA, January 2007.
- Program committee, 22nd annual ACM Symposium on Applied Computing (SAC '07), Seoul, South Korea, March 2007.
- 2008 Program committee, 23rd annual ACM Symposium on Applied Computing (SAC '08), Fortaleza, Ceará, Brazil, March 2008.
- 2009 Program committee, Workshop on Massive Data Algorithmics (MASSIVE '09), as part of the 25th Annual ACM Symposium on Computational Geometry (SCG '09), University of Aarhus, Århus, Denmark. June 2009.
- Program committee, First International Workshop on Data Warehousing and Knowledge Discovery from Sensors and Streams (DKSS '09), held in conjunction with the 5th International Conference on Distributed Computing in Sensor Systems (DCOSS '09), Marina Del Rey, CA, June 2009.
- Program committee, 11th International Workshop on Algorithms and Data Structures (WADS '09), Banff, Alberta, Canada, August 2009.
- Program committee, 16th International Conference on String Processing and Information Retrieval (SPIRE '09), Saariselkä, Finland, August 2009.
- 2010 Program committee, Workshop on Massive Data Algorithmics (MASSIVE '10), as part of the 26th Annual ACM Symposium on Computational Geometry (SCG '10), Snowbird, UT, June 2010.
- Chair, Workshop on CRA Guidelines for Enhancing Faculty Recruitment, Computing Research Association Conference 2010 (Snowbird '10), Snowbird, UT, July 2010.
- 2011 Program committee, Workshop on Massive Data Algorithmics (MASSIVE '11), as part of the 27th Annual ACM Symposium on Computational Geometry (SCG '11), Paris, France, June 2011.
- Program committee, 18th International Conference on String Processing and Information Retrieval (SPIRE '11), Pisa, Italy, October 2011.
- 2012 Program committee, 19th International Conference on String Processing and Information Retrieval (SPIRE '12), Cartagena de Indias, Colombia, October 2012.

## 13. Publications

Several of Dr. Vitter's recent publications are available electronically from his online publication library, which is reachable via a link from his home page <http://www.provost.ku.edu/jsv/>.

### Patents

1. E. E. Lindstrom and J. S. Vitter. "External Sorting Using Key Value Distribution and Range Formation," United States Patent No. 4,575,798, IBM Corporation, March 11, 1986.
2. J. S. Vitter, K. M. Curewitz, and P. Krishnan. "Online Background Predictors and Prefetchers," United States Patent No. 5,485,609, Duke University, January 16, 1996.

3. Y. Matias, J. S. Vitter, and N. Young. “Method for Implementing Approximate Data Structures using Operations on Machine Words,” United States Patent No. 5,519,840, AT&T Corporation, May 21, 1996.
4. R. D. Barve, P. B. Gibbons, B. K. Hillyer, Y. Matias, E. Shriver, and J. S. Vitter. “A System and Method for Modeling and Optimizing I/O Throughput of Multiple Disks on a Bus, I” United States patent No. 6,260,108, Lucent Technologies, July 10, 2001.
5. R. D. Barve, P. B. Gibbons, B. K. Hillyer, Y. Matias, E. Shriver, and J. S. Vitter. “A System and Method for Modeling and Optimizing I/O Throughput of Multiple Disks on a Bus, II” United States patent No. 6,301,640, Lucent Technologies, October 9, 2001.
6. L. Arge, O. Procopiuc, S. Ramaswamy, T. Suel, and J. S. Vitter. “System and Method for Performing Scalable Sweep Based Spatial Join,” United States patent pending.

## Books and Edited Special Issues of Journals

7. E. E. Lindstrom, J. S. Vitter, and C. K. Wong, editors. Special issue on sorting in *IEEE Transactions on Computers*, **C-34**(4), April 1985.
8. J. S. Vitter and W.-C. Chen. *Design and Analysis of Coalesced Hashing*, Oxford University Press, New York, NY, 1987.
9. J. S. Vitter, editor. Special issue on parallel and distributed computing, part I, in *Algorithmica*, **3**(1), 1988.
10. J. S. Vitter, editor. Special issue on parallel and distributed computing, part II, in *Algorithmica*, **3**(3), 1988.
11. J. S. Vitter, editor. Special Double Issue on Efficient Algorithms for Large-Scale Memory, *Algorithmica*, **12**(2-3), 1994, 69–71.
12. J. Abello and J. S. Vitter, editors. *External Memory Algorithms*, DIMACS Series on Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, Providence, RI, 1999.
13. J. S. Vitter and C. Zaroliagis, editors. *Algorithm Engineering*, Lecture Notes in Computer Science, **1668**, Springer, Berlin, Germany, 1999.
14. J. S. Vitter and C. Zaroliagis, editors. Special issue on algorithm engineering in *ACM Journal of Experimental Algorithmics*, **5**, 2000.
15. D. T. Hoang and J. S. Vitter. *Efficient Algorithms for MPEG Video Compression*, Wiley Series in Telecommunications and Signal Processing, John Wiley & Sons, New York, NY, 2002.
16. D. Q. Goldin, A. A. Shvartsman, S. A. Smolka, J. S. Vitter, and S. B. Zdonik, editors. *Proceedings of the Paris C. Kanellakis Memorial Workshop on Principles of Computing & Knowledge (PCK50)*, ACM Press, New York, 2003.
17. J. S. Vitter. *Algorithms and Data Structures for External Memory*, Series on Foundations and Trends in Theoretical Computer Science, Now Publishers, Hanover, MA, 2008. Also published as Volume 2, Issue 4 of *Foundations and Trends in Theoretical Computer Science*.

## Book Chapters

18. J. S. Vitter and R. A. Simons. “New Classes for Parallel Complexity: A Study of Unification and Other Complete Problems for  $\mathcal{P}$ ,” in *Computers for Artificial Intelligence Applications* (edited by B. Wah and G. J. Li), IEEE Computer Society Press, Washington, DC, 1986, 135–150.
19. J. E. Savage and J. S. Vitter. “Parallelism in Space-Time Tradeoffs,” in *Advances in Computing Research*, **4**, special volume on parallel and distributed computing (edited by F. P. Preparata), JAI Press, 1987, 117–146.

20. J. S. Vitter and Ph. Flajolet. “Average-Case Analysis of Algorithms and Data Structures,” Chapter 9 in *Handbook of Theoretical Computer Science*, Volume A: *Algorithms and Complexity* (edited by J. van Leeuwen), Elsevier and M.I.T. Press, 1990, 431–524.
21. K. J. Basye, T. L. Dean, and J. S. Vitter. “Coping with Uncertainty in Map Learning,” in *Autonomous Mobile Robots: Perception, Mapping, and Navigation*, (edited by S. S. Iyengar and A. Elfes), IEEE Computer Society Press, 1991.
22. P. G. Howard and J. S. Vitter. “Practical Implementations of Arithmetic Coding,” in *Image and Text Compression* (edited by J. Storer), Kluwer Academic Publishers, Norwell, MA, 1992, 85–112.
23. R. D. Barve, E. F. Grove, and J. S. Vitter. “Simple Randomized Mergesorting on Parallel Disks,” in *Randomization Methods in Algorithm Design* (edited by P. Pardalos, S. Rajasekaran, and J. Rolim), DIMACS Series on Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, Providence, RI, 1998, 1–20.
24. J. S. Vitter. “External Memory Algorithms and Data Structures,” in *External Memory Algorithms* (edited by J. Abello and J. S. Vitter), DIMACS Series on Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, Providence, RI, 1999, 1–38.
25. J. S. Vitter. “Algorithms and Data Structures for External Memory,” Chapter 32 in *The Computer Engineering Handbook* (edited by V. Oklobdzija), CRC Press and IEEE Press, 2002, second edition 2008.
26. J. S. Vitter. “External Memory Algorithms,” Chapter 10 in *Handbook of Massive Data Sets* (edited by J. Abello, P. M. Pardalos, and M. G. C. Resende), Kluwer Academic Publishers, Boston, 2002, 359–418,
27. “Geometric and Spatial Data Structures in External Memory,” Chapter 27 in *Handbook on Data Structures and Applications* (edited by D. Mehta and S. Sahni) CRC Press, 2005.
28. P. G. Howard and J. S. Vitter. “Arithmetic Coding for Data Compression,” in *Encyclopedia of Algorithms* (edited by M.-Y. Kao), Springer Science + Business Media, New York, 2008, 65–68.
29. J. S. Vitter. “External Sorting and Permuting,” in *Encyclopedia of Algorithms* (edited by M.-Y. Kao), Springer Science + Business Media, New York, 2008, 291–297.

## Journal and Conference Articles

This section jointly lists Dr. Vitter’s journal and conference papers. In several cases, the journal paper is an expanded and more detailed version of a conference paper, so for succinctness, their listings are combined into a single entry. Conference papers in computer science are important in their own right; they are generally peer-reviewed, often with very low acceptance ratios.

30. J. S. Vitter. “A Shared Memory Scheme for Coalesced Hashing,” *Information Processing Letters*, **13**(2), November 13, 1981, 77–79.
31. J. S. Vitter. “Deletion Algorithms for Hashing that Preserve Randomness,” *Journal of Algorithms*, **3**(3), September 1982, 261–275. An extended abstract appears in *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computer Science (FOCS ’81)*, Nashville, TN, October 1981, 127–132.
32. J. S. Vitter. “Implementations for Coalesced Hashing,” *Communications of the ACM*, **25**(12), December 1982, 911–926.
33. J. S. Vitter. “Analysis of the Search Performance of Coalesced Hashing,” *Journal of the ACM*, **30**(2), April 1983, 231–258. An extended abstract appears in “Tuning the Coalesced Hashing Method to Obtain Optimum Performance,” *Proceedings of the 21st Annual IEEE Symposium on Foundations of Computer Science (FOCS ’80)*, Syracuse, NY, October 1980, 238–247.
34. J. S. Vitter and W.-C. Chen. “Analysis of Early-Insertion Standard Coalesced Hashing,” *SIAM Journal on Computing*, **12**(4), November 1983, 667–676.

35. J. S. Vitter. "Faster Methods for Random Sampling," *Communications of the ACM*, **27**(7), July 1984, 703–718. An extended abstract appears in "Optimum Random Sampling," *Proceedings of the 17th Annual Conference on Information Sciences and Systems*, Baltimore, MD, March 1983, 158–164. Also in "Optimum Algorithms for Two Random Sampling Problems," *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computer Science (FOCS '83)*, Tucson, AZ, November 1983, 65–75.
36. J. S. Vitter. "US&R: A New Framework for Redoing," *IEEE Software*, **1**(4), October 1984, 39–52. An extended abstract appears in *Proceedings of the ACM Software Engineering Symposium on Practical Software Development Environments*, Pittsburgh, PA, April 1984, 168–176.
37. J. S. Vitter and R. A. Simons. "Parallel Algorithms for Unification and Other Complete Problems for  $\mathcal{P}$ ," *Proceedings of the 1984 ACM Annual Conference: The Fifth Generation Challenge (ACM '84)*, San Francisco, CA, October 1984, 75–84.
38. W.-C. Chen and J. S. Vitter. "Analysis of New Variants of Coalesced Hashing," *ACM Transactions on Database Systems*, **9**(4), December 1984, 616–645. An addendum appears in *ACM Transactions on Database Systems*, **10**(1), March 1985, 127. An extended abstract appears in "New Variants of Coalesced Hashing," *Proceedings of the 21st Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, October 1983, 683–692.
39. E. E. Lindstrom and J. S. Vitter. "The Design and Analysis of BucketSort for Bubble Memory Secondary Storage," *IEEE Transactions on Computers*, **C-34**(3), March 1985, 218–233. Basis for patent in [1]. An extended abstract appears in "The Design and Analysis of BucketSort," *22nd Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, October 1984, 525–534.
40. J. S. Vitter. "Random Sampling with a Reservoir," *ACM Transactions on Mathematical Software*, **11**(1), March 1985, 37–57. An extended abstract appears in "Optimum Algorithms for Two Random Sampling Problems," *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computer Science (FOCS '83)*, Tucson, AZ, November 1983, 65–75.
41. J. S. Vitter and W.-C. Chen. "Optimum Algorithms for a Model of Direct Chaining," *SIAM Journal on Computing*, **14**(2), May 1985, 490–499.
42. J. S. Vitter. "An Efficient I/O Interface for Optical Disks," *ACM Transactions on Database Systems*, **10**(2), June 1985, 129–162. An extended abstract appears in "Computational Complexity of an Optical Disk Interface," *Proceedings of the 11th Annual International Colloquium on Automata, Languages, and Programming (ICALP '84)*, Antwerp, Belgium, July 1984, published in *Lecture Notes in Computer Science*, **172**, Springer, Berlin, Germany, 490–502.
43. C. J. Van Wyk and J. S. Vitter. "The Complexity of Hashing with Lazy Deletion," *Algorithmica* **1**(1), 1986, 17–29.
44. R. Sedgewick and J. S. Vitter. "Shortest Paths in Euclidean Graphs," *Algorithmica*, **1**(1), 1986, 31–48. An extended abstract appears in *Proceedings of the 25th Annual IEEE Symposium on Foundations of Computer Science (FOCS '84)*, West Palm Beach, FL, October 1984, 417–424.
45. J. S. Vitter and R. A. Simons. "New Classes for Parallel Complexity: A Study of Unification and Other Complete Problems for  $\mathcal{P}$ ," *IEEE Transactions on Computers*, **C-35**(5), May 1986, 403–418. Reprinted in [18].
46. W.-C. Chen and J. S. Vitter. "Deletion Algorithms for Coalesced Hashing," *The Computer Journal*, **29**(5), October 1986, 436–450.
47. J. E. Savage and J. S. Vitter. "Parallelism in Space-Time Tradeoffs," *Proceedings of the International Workshop on Parallel Computing and VLSI*, Amalfi, Italy, May 1984, published in *VLSI: Algorithms and Architectures* (edited by P. Bertolazzi and F. Luccio), Elsevier Science Press, 1985, 49–58. A longer version appears in [19].
48. J. T. Stasko and J. S. Vitter. "Pairing Heaps: Experiments and Analysis," *Communications of the ACM*, **30**(3), March 1987, 234–249.
49. J. S. Vitter. "An Efficient Algorithm for Sequential Random Sampling," *ACM Transactions on Mathematical Software*, **13**(1), March 1987, 58–67.

50. J. S. Vitter. "Design and Analysis of Dynamic Huffman Codes," *Journal of the ACM*, **34**(4), October 1987, 825–845. An extended abstract appears in "The Design and Analysis of Dynamic Huffman Coding," *Proceedings of the 26th Annual IEEE Symposium on Foundations of Computer Science (FOCS '85)*, Portland, OR, October 1985, 293–302.
51. C. A. Schevon and J. S. Vitter. "A Parallel Algorithm for Recognizing Unordered Depth-First Search," *Information Processing Letters*, June 24, 1988, 105–110.
52. A. Aggarwal and J. S. Vitter. "The Input/Output Complexity of Sorting and Related Problems," *Communications of the ACM*, **31**(9), September 1988, 1116–1127. An extended abstract appears in "The I/O Complexity of Sorting and Related Problems," *Proceedings of the 14th Annual International Colloquium on Automata, Languages, and Programming (ICALP '87)*, Karlsruhe, West Germany, July 1987, published in *Lecture Notes in Computer Science*, **267**, Springer, Berlin, Germany.
53. J. S. Vitter. "ALGORITHM 673 Dynamic Huffman Coding," *ACM Transactions on Mathematical Software*, **15**(2), June 1989, 158–167. Also appears in *Collected Algorithms of ACM*.
54. R. Tamassia and J. S. Vitter. "Dynamic Graph Algorithms," *Proceedings of the 3rd Australian Workshop on Combinatorial Algorithms*, Bali, Indonesia, June 1990.
55. F. P. Preparata, J. S. Vitter, and M. Yvinec. "Computation of the Axial View of a Set of Isothetic Parallelepipeds," *ACM Transactions on Graphics*, **9**(3), July 1990, 278–300.
56. J. S. Vitter. "Efficient Memory Access in Large-Scale Computation," invited paper in *Proceedings of the 1991 Symposium on Theoretical Aspects of Computer Science (STACS '91)*, Hamburg, West Germany, February 1991, published in *Lecture Notes in Computer Science*, **480**, Springer, Berlin, Germany, 26–41.
57. R. Tamassia, I. G. Tollis, and J. S. Vitter. "Lower Bounds for Planar Orthogonal Drawings of Graphs," *Information Processing Letters*, **39**, 1991, 35–40. An extended abstract appears in "Lower Bounds and Parallel Algorithms for Planar Orthogonal Grid Drawings," *Proceedings of the 3rd IEEE Symposium on Parallel and Distributed Processing (SPDP '91)*, Dallas, TX, December 1991, 386–393.
58. A. L. Buchsbaum, P. C. Kanellakis, and J. S. Vitter. "A Data Structure for Arc Insertion and Regular Path Finding," invited paper in special issue on deductive databases in *Annals of Mathematics and Artificial Intelligence*, **3**(2–4), March 1991. An extended abstract appears in *Proceedings of the 1st Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '90)*, San Francisco, CA, January 1990, 22–31.
59. R. Tamassia and J. S. Vitter. "Parallel Transitive Closure and Point Location in Planar Structures," *SIAM Journal on Computing*, **20**(4), August 1991, 708–725. An extended abstract appears in "Optimal Parallel Algorithms for Transitive Closure and Point Location in Planar Structures," *Proceedings of the 1st Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '89)*, Sante Fe, NM, June 1989, 399–408. Also appears as an invited paper in *Proceedings of the International Workshop on Discrete Algorithms and Complexity*, Fukuoka, Japan, November 1989, 169–178.
60. C. M. Kenyon and J. S. Vitter. "Maximum Queue Size and Hashing with Lazy Deletion," *Algorithmica*, **6**(4), 1991, 597–619. An extended abstract appears in *Proceedings of the 20th Annual Symposium on the Interface of Computing Science and Statistics*, Reston, VA, April 1988, 743–748.
61. M. H. Nodine, D. P. Lopresti and J. S. Vitter. "I/O Overhead and Parallel VLSI Architectures for Lattice Computations," *IEEE Transactions on Computers*, **40**(7), July 1991, 843–852. An extended abstract appears in *Proceedings of the 1990 International Conference on Computing and Information (ICCI '90)*, Niagara Falls, Canada, May 1990, 472–476. Also published in *Advances in Computing and Information*, *Lecture Notes in Computer Science*, **468**, Springer, Berlin, Germany, 1990, 497–506.
62. C. M. Kenyon-Mathieu and J. S. Vitter. "The Maximum Size of Dynamic Data Structures," *SIAM Journal on Computing*, **20**(5), October 1991, 807–823. An extended abstract appears in "General Methods for the Analysis of the Maximum Size of Dynamic Data Structures," *Proceedings of the 16th Annual International Colloquium on Automata, Languages, and Programming (ICALP '89)*, Stresa, Italy, July 1989, published in *Lecture Notes in Computer Science*, **372**, Springer, Berlin, Germany, 473–487.

63. J.-H. Lin and J. S. Vitter. “Complexity Results on Learning by Neural Nets,” *Machine Learning*, **6**, 1991, 211–230. An extended abstract appears in *Proceedings of the 2nd Annual ACM Workshop on Computational Learning Theory (COLT ’89)*, Santa Cruz, CA, July–August 1989, published by Morgan Kaufmann, San Mateo, CA, 118–133.
64. J. S. Vitter and J.-H. Lin. “Learning in Parallel,” *Information and Computation*, **92**(2), February 1992, 179–202. An extended abstract appears in *Proceedings of the 1st Annual ACM Workshop on Computational Learning Theory (COLT ’88)*, Cambridge, MA, August 1988, published by Morgan Kaufmann, San Mateo, CA, 106–124.
65. P. G. Howard and J. S. Vitter. “Practical Implementations of Arithmetic Coding,” invited paper in *Proceedings of the 3rd International Conference on Advances in Communication and Control Systems (COMCON ’91)*, Victoria, Canada, October 1991. A longer version appears in [22].
66. F. P. Preparata, J. S. Vitter, and M. Yvinec. “Output-Sensitive Generation of the Perspective View of Isothetic Parallelepipeds,” *Algorithmica*, **8**, 1992, 257–283. An extended abstract appears in *Proceedings of the 2nd Scandinavian Workshop on Algorithm Theory (SWAT ’90)*, Bergen, Norway, July 1990, published in *Lecture Notes in Computer Science*, **447**, Springer, Berlin, Germany, 71–84.
67. P. G. Howard and J. S. Vitter. “Analysis of Arithmetic Coding for Data Compression,” invited paper in special issue on data compression for image and text in *Journal of Information Processing and Management*, **28**(6), 1992, 749–763. An extended abstract appears in an invited paper in *Proceedings of the 1991 IEEE Data Compression Conference (DCC ’91)*, Snowbird, UT, April 1991, 3–12.
68. P. G. Howard and J. S. Vitter. “New Methods for Lossless Image Compression Using Arithmetic Coding,” invited paper in special issue on data compression for image and text in *Journal of Information Processing and Management*, **28**(6), 1992, 765–779. An extended abstract appears in *Proceedings of the 1991 IEEE Data Compression Conference (DCC ’91)*, Snowbird, UT, April 1991, 257–266.
69. J.-H. Lin and J. S. Vitter. “Approximation Algorithms for Geometric Median Problems,” *Information Processing Letters*, **44**, 1992, 245–249.
70. J.-H. Lin and J. S. Vitter. “Nearly Optimal Vector Quantization via Linear Programming,” *Proceedings of the 1992 IEEE Data Compression Conference (DCC ’92)*, Snowbird, UT, March 1992, 22–31.
71. P. G. Howard and J. S. Vitter. “Error Modeling for Hierarchical Lossless Image Compression,” *Proceedings of the 1992 IEEE Data Compression Conference (DCC ’92)*, Snowbird, UT, March 1992, 269–278.
72. J.-H. Lin and J. S. Vitter. “ $\epsilon$ -Approximations with Minimum Packing Constraint Violation,” *Proceedings of the 24th Annual ACM Symposium on Theory of Computing (STOC ’92)*, Victoria, Canada, May 1992, 771–782.
73. J. S. Vitter and M. H. Nodine. “Large-Scale Sorting in Uniform Memory Hierarchies,” special issue on parallel I/O systems in *Journal of Parallel and Distributed Computing*, **17**, January 1993, 107–114. An extended abstract appears in “Large-Scale Sorting in Parallel Memories,” *Proceedings of the 3rd Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA ’91)*, Hilton Head, SC, July 1991, 29–39.
74. M. H. Nodine and J. S. Vitter. “Paradigms for Optimal Sorting with Multiple Disks,” *Proceedings of the 26th Hawaii International Conference on Systems Sciences (HICSS ’93)*, Maui, HI, January 1993, 50–59.
75. F. P. Preparata and J. S. Vitter. “A Simplified Technique for Hidden-Line Elimination in Terrains,” *International Journal of Computational Geometry & Applications*, **3**(2), 1993, 167–181. An extended abstract appears in *Proceedings of the 1992 Symposium on Theoretical Aspects of Computer Science (STACS ’92)*, Paris, France, February 1992, published in *Lecture Notes in Computer Science*, **577**, Springer, Berlin, Germany, 135–146.
76. P. G. Howard and J. S. Vitter. “Fast and Efficient Lossless Image Compression,” *Proceedings of the 1993 IEEE Data Compression Conference (DCC ’93)*, Snowbird, UT, April 1993, 351–360.

77. M. H. Nodine and J. S. Vitter. "Load Balancing Paradigms for Optimal Use of Parallel Disks and Parallel Memory Hierarchies," invited paper in *Proceedings of the 2nd Annual Dartmouth Institute for Advanced Graduate Studies (DAGS '93)*, Hanover, NH, June 1993.
78. R. F. Cohen, S. Subramanian, R. Tamassia, and J. S. Vitter. "Dynamic Algorithms for Optimization Problems on Bounded Tree-Width Graphs," *Proceedings of the 3rd Conference on Integer Programming and Combinatorial Optimization (IPCO '93)*, Erice, Italy, April–May 1993, 99–112.
79. K. M. Curewitz, P. Krishnan, and J. S. Vitter. "Practical Prefetching via Data Compression," *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data (SIGMOD '93)*, Washington, DC, May 1993, 257–266. Basis for patent in [2].
80. M. T. Goodrich, J.-J. Tsay, D. E. Vengroff, and J. S. Vitter. "External-Memory Computational Geometry," *Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science (FOCS '93)*, Palo Alto, CA, November 1993, 714–723.
81. Y. Matias, J. S. Vitter, and N. Young. "Approximate Data Structures with Applications," *Proceedings of the 5th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '94)*, Alexandria, VA, January 1994, 361–370.
82. P. G. Howard and J. S. Vitter. "Fast Progressive Lossless Image Compression," *Proceedings of the Image and Video Compression Conference, IS&T/SPIE 1994 Symposium on Electronic Imaging Science & Technology*, **2186**, San Jose, CA, February 1994, 98–109.
83. D. T. Hoang, P. M. Long, and J. S. Vitter. "Explicit Bit Minimization for Motion-Compensated Video Coding," *Proceedings of the 1994 IEEE Data Compression Conference (DCC '94)*, Snowbird, UT, March 1994, 175–184.
84. P. G. Howard and J. S. Vitter. "Arithmetic Coding for Data Compression," invited paper in *Proceedings of the IEEE*, **82**(6), June 1994, 857–865.
85. J. S. Vitter and E. A. M. Shriver. "Algorithms for Parallel Memory, I: Two-Level Memories," double special issue on large-scale memories in *Algorithmica*, **12**(2–3), 1994, 110–147. An extended abstract appears in "Optimal Disk I/O with Parallel Block Transfer," *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing (STOC '90)*, Baltimore, MD, May 1990, 159–169.
86. J. S. Vitter and E. A. M. Shriver. "Algorithms for Parallel Memory, II: Hierarchical Multilevel Memories," double special issue on large-scale memories in *Algorithmica*, **12**(2–3), 1994, 148–169. An extended abstract appears in "Optimal Disk I/O with Parallel Block Transfer," *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing (STOC '90)*, Baltimore, MD, May 1990, 159–169.
87. P. G. Howard and J. S. Vitter. "Design and Analysis of Fast Text Compression Based on Quasi-Arithmetic Coding," *Journal of Information Processing and Management*, **30**(6), 1994, 777–790. An extended abstract appears in *Proceedings of the 1993 IEEE Data Compression Conference (DCC '93)*, Snowbird, UT, April 1993.
88. P. B. Miltersen, S. Subramanian, J. S. Vitter, and R. Tamassia. "Complexity Models for Incremental Computation," special issue of *Theoretical Computer Science*, **130**, 1994, 203–236. An extended abstract appears in S. Subramanian, J. S. Vitter, and R. Tamassia, "A Complexity Theoretic Approach to Incremental Computation," *Proceedings of the 1993 Symposium on Theoretical Aspects of Computer Science (STACS '93)*, Würzburg, Germany, February 1993, published in *Lecture Notes in Computer Science*, **665**, Springer, Berlin, Germany, 640–649.
89. J.-H. Lin and J. S. Vitter. "A Theory for Memory-Based Learning," special issue of *Machine Learning*, **17**(2/3), November/December 1994, 143–167. An extended abstract appears in *Proceedings of the 5th Annual ACM Conference on Computational Learning Theory (COLT '92)*, Pittsburgh, PA, July 1992, 103–115.
90. Y.-J. Chiang, M. T. Goodrich, E. F. Grove, R. Tamassia, D. E. Vengroff, and J. S. Vitter. "External-Memory Graph Algorithms," *Proceedings of the 6th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '95)*, San Francisco, CA, January 1995, 139–149.

91. S. Subramanian, R. Tamassia, and J. S. Vitter. “An Efficient Parallel Algorithm for Shortest Paths in Planar Layered Digraphs,” *Algorithmica*, **14**, 1995, 322–339. An extended abstract appears in “A Divide and Conquer Approach to Shortest Paths in Planar Layered Digraphs,” *Proceedings of the 4th IEEE Symposium on Parallel and Distributed Processing (SPDP '92)*, Dallas, TX, December 1992, 176–183.
92. M. H. Nodine and J. S. Vitter. “Greed Sort: An Optimal Sorting Algorithm on Parallel Disks,” *Journal of the ACM*, **42**(4), July 1995, 919–933. An extended abstract appears in “Large-Scale Sorting in Parallel Memories,” *Proceedings of the 3rd Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '91)*, Hilton Head, SC, July 1991, 29–39.
93. E. F. Grove, M.-Y. Kao, P. Krishnan, and J. S. Vitter. “Online Perfect Matching and Mobile Computing,” *Proceedings of the 4th Biannual Workshop on Algorithms and Data Structures (WADS '95)*, Kingston, Ontario, Canada, August 1995, 194–205.
94. A. Awerbuch, A. Azar, E. F. Grove, P. Krishnan, M.-Y. Kao, and J. S. Vitter. “Load Balancing in the  $L_p$  Norm,” *Proceedings of the 36th Annual IEEE Symposium on Foundations of Computer Science (FOCS '95)*, Milwaukee, WI, October 1995, 383–391.
95. R. Tamassia and J. S. Vitter. “Optimal Cooperative Search in Fractional Cascaded Data Structures,” invited paper in special issue on parallel computing in *Algorithmica*, **15**(2), February 1996, 154–171. An extended abstract appears in *Proceedings of the 2nd Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '90)*, Crete, Greece, July 1990, 307–316.
96. P. G. Howard and J. S. Vitter. “Parallel Lossless Image Compression Using Huffman and Arithmetic Coding,” *Information Processing Letters*, **59**, 1996, 65–73. An extended abstract appears in *Proceedings of the 1992 IEEE Data Compression Conference (DCC '92)*, Snowbird, UT, March 1992, 299–308.
97. D. T. Hoang, P. M. Long, and J. S. Vitter. “Efficient Cost Measures for Motion Compensation at Low Bit Rates,” *Proceedings of the 1996 IEEE Data Compression Conference (DCC '96)*, Snowbird, UT, April 1996.
98. P. Krishnan, J. S. Vitter, and B. Iyer. “Estimating Alphanumeric Selectivity in the Presence of Wildcards,” *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data (SIGMOD '96)*, Montreal, Canada, May 1996, 282–293.
99. D. E. Vengroff and J. S. Vitter. “Efficient 3-D Range Searching in External Memory,” *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC '96)*, Philadelphia, PA, May 1996, 192–201.
100. M. H. Nodine, M. T. Goodrich, and J. S. Vitter. “Blocking for External Graph Searching,” *Algorithmica*, **16**(2), August 1996, 181–214. An extended abstract appears in *Proceedings of the 12th Annual ACM Symposium on Principles of Database Systems (PODS '93)*, Washington, DC, May 1993, 222–232.
101. K. Romanik and J. S. Vitter. “Using Vapnik-Chervonenkis Dimension to Analyze the Testing Complexity of Program Segments,” *Information and Computation*, **128**(2), August 1, 1996, 87–108. An extended abstract appears in “Using Computational Learning Theory to Analyze the Testing Complexity of Program Segments,” *Proceedings of the 17th Annual IEEE International Computer Software and Applications Conference (COMPSAC '93)*, Phoenix, AZ, 1993, 367–373.
102. J. S. Vitter and P. Krishnan. “Optimal Prefetching via Data Compression,” *Journal of the ACM*, **43**(5), September 1996, 771–793. An extended abstract appears in *Proceedings of the 32nd Annual IEEE Symposium on Foundations of Computer Science (FOCS '91)*, San Juan, Puerto Rico, October 1991, 121–130.
103. P. C. Kanellakis, S. Ramaswamy, D. E. Vengroff, and J. S. Vitter. “Indexing for Data Models with Constraints and Classes,” *Journal of Computer and System Sciences*, **52**(3), 1996, 589–612. An extended abstract appears in *Proceedings of the 12th Annual ACM Symposium on Principles of Database Systems (PODS '93)*, Washington, DC, May 1993, 233–243.

104. D. E. Vengroff and J. S. Vitter. "I/O-Efficient Scientific Computation using TPIE," *Proceedings of the Goddard Conference on Mass Storage Systems and Technologies*, College Park, MD, September 1996, published in NASA Conference Publication 3340, Volume II, 553–570. A shorter preliminary version appears in "Supporting I/O-Efficient Scientific Computing in TPIE," *Proceedings of the 7th IEEE Symposium on Parallel and Distributed Processing (SPDP '95)*, San Antonio, TX, October 1995, 74–77.
105. G. Gibson, J. S. Vitter, and J. Wilkes, editors. "Strategic Directions in Storage I/O for Large-Scale Computing," *ACM Computing Surveys*, **28**(4), December 1996, 779–793.
106. J. S. Vitter. "Communication Issues in Large-Scale Geometric Computation," *ACM Computing Surveys*, **28**(4es), Article 20, December 1996.
107. D. E. Vengroff and J. S. Vitter. "I/O-Efficient Algorithms and Environments," *ACM Computing Surveys*, **28**(4es), Article 212, December 1996.
108. M. Wang, J. S. Vitter, and B. Iyer. "Selectivity Estimation in the Presence of Alphanumeric Correlations," *Proceedings of the 13th Annual IEEE International Conference on Data Engineering (ICDE '97)*, Birmingham, England, April 1997, 169–180.
109. P. K. Agarwal, T. M. Murali, and J. S. Vitter. "Practical Techniques for Constructing Binary Space Partitions for Orthogonal Rectangles," *Proceedings of the 13th Annual ACM Symposium on Computational Geometry (SCG '97)*, Nice, France, June 1997, 382–384.
110. R. D. Barve, E. F. Grove, and J. S. Vitter. "Simple Randomized Mergesorting on Parallel Disks," special issue on parallel I/O in *Parallel Computing*, **23**(4), 1997, 601–631. An extended abstract appears in *Proceedings of the 8th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '96)*, Padua, Italy, June 1996, 109–118. An updated extended abstract appears in [23].
111. K. J. Basye, T. L. Dean, and J. S. Vitter. "Coping with Uncertainty in Map Learning," *Machine Learning*, **29**(1), 1997, 65–88. An extended abstract appears in *Proceedings of the 11th Joint Conference on Artificial Intelligence (IJCAI '89)*, Detroit, MI, August 1989, 663–668. Reprinted in [21].
112. L. Arge, P. Ferragina, R. Grossi, and J. S. Vitter. "On Sorting Strings in External Memory," *Proceedings of the 29th Annual ACM Symposium on Theory of Computing (STOC '97)*, El Paso, TX, May 1997, 540–548. Also appears as "Sequence Sorting in Secondary Storage," invited paper in *Proceedings of the 1997 International IEEE Conference on Compression and Complexity of Sequences (SEQUENCES '97)*, Positano, Italy, June 1997, 329–346.
113. D. T. Hoang and J. S. Vitter. "Multiplexing VBR Video Sequences onto a CBR Channel with Lexicographic Optimization," *Proceedings of the 1997 IEEE International Conference on Image Processing (ICIP '97)*, Santa Barbara, CA, October 1997.
114. D. T. Hoang, E. Linzer, and J. S. Vitter. "Lexicographic Bit Allocation for MPEG Video," special issue on high-fidelity media processing in *Journal of Visual Communication and Image Representation*, **8**(4), December 1997, 384–404. An extended abstract appears in D. T. Hoang, J. S. Vitter, and E. Linzer, *Proceedings of the 1997 IEEE International Conference on Image Processing (ICIP '97)*, Santa Barbara, CA, October 1997. An earlier version appears in "A Lexicographic Framework for MPEG Rate Control," *Proceedings of the 1997 IEEE Data Compression Conference (DCC '97)*, Snowbird, UT, March 1997, 101–110.
115. P. K. Agarwal, L. Arge, T. M. Murali, K. R. Varadarajan, and J. S. Vitter. "I/O-Efficient Algorithms for Contour Line Extraction and Planar Graph Blocking," *Proceedings of the 9th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '98)*, San Francisco, CA, January 1998, 117–126.
116. L. Arge, O. Procopiu, S. Ramaswamy, T. Suel, and J. S. Vitter. "Theory and Practice of I/O-Efficient Algorithms for Multidimensional Batched Searching Problems," *Proceedings of the 9th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '98)*, San Francisco, CA, January 1998, 685–694.
117. Y. Matias, J. S. Vitter, and M. Wang. "Wavelet-Based Histograms for Selectivity Estimation," *Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data (SIGMOD '98)*, Seattle, WA, June 1998, 448–459.

118. D. T. Hoang, P. M. Long, and J. S. Vitter. “Rate-Distortion Optimizations for Motion Estimation in Low-Bitrate Video Coding,” *IEEE Transactions on Circuits and Systems for Video Technology*, **8**(4), August 1998, 488–500. An extended abstract appears in *Proceedings of the Digital Video Compression Conference, IS&T/SPIE 1996 Symposium on Electronic Imaging Science & Technology*, **2668**, San Jose, CA, January–February 1996.
119. L. Arge, O. Procopiuc, S. Ramaswamy, T. Suel, and J. S. Vitter. “Scalable Sweep-Based Spatial Join,” *Proceedings of the 24th International Conference on Very Large Databases (VLDB ’98)*, New York, NY, August 1998, 570–581.
120. T. M. Murali, P. K. Agarwal, and J. S. Vitter. “Constructing Binary Space Partitions for Orthogonal Rectangles in Practice,” *Proceedings of the 6th Annual European Symposium on Algorithms (ESA ’98)*, Venice, August 1998, published in *Lecture Notes in Computer Science*, **1461**, Springer, Berlin, Germany, 211–222.
121. J. S. Vitter, M. Wang, and B. Iyer. “Data Cube Approximation and Histograms via Wavelets,” *Proceedings of the 7th International Conference on Information and Knowledge Management (CIKM ’98)*, Washington, DC, November 1998, 96–104.
122. P. Krishnan and J. S. Vitter. “Optimal Prediction for Prefetching in the Worst Case,” *SIAM Journal on Computing*, **27**(6), December 1998, 1617–1636. An extended abstract appears in *Proceedings of the 5th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA ’94)*, Alexandria, VA, January 1994, 392–401.
123. P. Krishnan, P. M. Long, and J. S. Vitter. “Adaptive Disk Spindown via Optimal Rent-to-Buy in Probabilistic Environments,” *Algorithmica*, **23**(1), January 1999, 31–56. An extended abstract appears in “Learning to Make Rent-to-Buy Decisions in Probabilistic Environments with Systems Applications,” *Proceedings of the 12th International Conference on Machine Learning (ML ’95)*, Tahoe City, CA, July 1995, 322–330.
124. P. K. Agarwal, L. Arge, G. Brodal, and J. S. Vitter. “I/O-Efficient Dynamic Point Location in Monotone Subdivisions,” *Proceedings of the 10th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA ’99)*, Baltimore, MD, January 1999, 11–20.
125. R. D. Barve, E. A. M. Shriver, P. B. Gibbons, B. K. Hillyer, Y. Matias, and J. S. Vitter. “Modeling and Optimizing I/O Throughput of Multiple Disks on a Bus,” *Proceedings of the Joint International ACM Conference on Measurement and Modeling of Computer Systems (SIGMETRICS ’99)*, Atlanta, GA, May 1999, 83–92. An abstract appears in *Proceedings of the Joint ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS ’98/PERFORMANCE ’98)*, Madison, WI, June 1998, 264–265.
126. R. D. Barve, E. A. M. Shriver, P. B. Gibbons, B. K. Hillyer, Y. Matias, and J. S. Vitter. “Round-like Behavior in Multiple Disks on a Bus,” *Proceedings of the 7th Annual ACM-IEEE Workshop on I/O in Parallel And Distributed Systems (IOPADS ’99)*, Atlanta, GA, May 1999, 1–9.
127. L. Arge, V. Samoladas, and J. S. Vitter. “Two-Dimensional Indexability and Optimal Range Search Indexing,” *Proceedings of the 18th Annual ACM Symposium on Principles of Database Systems (PODS ’99)*, Philadelphia, PA, May–June 1999, 346–357.
128. J. S. Vitter and M. Wang. “Approximate Computation of Multidimensional Aggregates of Sparse Data Using Wavelets,” *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data (SIGMOD ’99)*, Philadelphia, PA, June 1999, 193–204. Recipient of the 2009 SIGMOD Test of Time award for the most impactful paper from SIGMOD ’99.
129. A. Condon, H. Edelsbrunner, E. A. Emerson, L. Fortnow, S. Haber, R. Karp, D. Leivant, R. Lipton, N. Lynch, I. Parberry, C. Papadimitriou, M. Rabin, A. Rosenberg, J. S. Royer, J. E. Savage, A. L. Selman, C. Smith, E. Tardos, and J. S. Vitter. “Challenges for Theory of Computing,” Report of an NSF-Sponsored Workshop on Research in Theoretical Computer Science, *SIGACT NEWS*, **30**(2), June 1999, 62–76.

130. J. S. Vitter. "Online Data Structures in External Memory," invited paper in *Proceedings of the 6th Biannual Workshop on Algorithms and Data Structures (WADS '99)*, Vancouver, Canada, August 1999, published in *Lecture Notes in Computer Science*, **1668**, Springer, Berlin, Germany. Also appears as an invited paper in *Proceedings of the 26th Annual International Colloquium on Automata, Languages, and Programming (ICALP '99)*, Prague, Czech Republic, July 1999, published in *Lecture Notes in Computer Science*, **1644**, Springer, Berlin, Germany, 119–133.
131. D. T. Hoang, P. M. Long, and J. S. Vitter. "Dictionary Selection using Partial Matching," *Information Sciences*, **119**(1–2), 1999, 57–72. An extended abstract appears in "Multiple-dictionary Compression Using Partial Matching," *Proceedings of the 1995 IEEE Data Compression Conference (DCC '95)*, Snowbird, UT, March 1995, 272–281.
132. E. F. Grove, T. M. Murali, and J. S. Vitter. "The Object Complexity Model for Hidden-Surface Elimination," *International Journal of Computational Geometry & Applications*, **9**, 1999, 207–217. An extended abstract appears in *Proceedings of the 7th Canadian Conference on Computational Geometry (CCCG '95)*, Quebec, Canada, August 1995, 273–278.
133. R. D. Barve and J. S. Vitter, "A Theoretical Framework for Memory-Adaptive Algorithms," *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS '99)*, New York, NY, October 1999, 273–284.
134. P. M. Long, A. I. Natsev, and J. S. Vitter. "Text Compression Via Alphabet Re-Representation," *Neural Networks*, **12** (4–5), 1999, 755–765. An extended abstract appears in *Proceedings of the 1997 IEEE Data Compression Conference (DCC '97)*, Snowbird, UT, March 1997.
135. P. K. Agarwal, E. F. Grove, T. M. Murali, and J. S. Vitter. "Binary Space Partitions for Fat Rectangles," *SIAM Journal on Computing*, **29**(5), 2000, 1422–1448. An extended abstract appears in *Proceedings of the 37th Annual IEEE Symposium on Foundations of Computer Science (FOCS '96)*, Burlington, VT, October 1996, 482–493. Also appears in *Abstracts of the 1st CGC Workshop on Computational Geometry*, Center for Geometric Computing, Johns Hopkins University, Baltimore, MD, October 1996.
136. P. K. Agarwal, L. Guibas, T. M. Murali, and J. S. Vitter. "Cylindrical Static and Kinetic Binary Space Partitions," *Computational Geometry*, **16**(2), 2000, 103–127. An extended abstract appears in *Proceedings of the 13th Annual ACM Symposium on Computational Geometry (SCG '97)*, Nice, France, June 1997, 39–48.
137. R. Tamassia, I. G. Tollis, and J. S. Vitter. "A Parallel Algorithm for Planar Orthogonal Grid Drawings," *Parallel Processing Letters*, **10**(1), March 2000, 141–150. An extended abstract appears in "Lower Bounds and Parallel Algorithms for Planar Orthogonal Grid Drawings," *Proceedings of the 3rd IEEE Symposium on Parallel and Distributed Processing (SPDP '91)*, Dallas, TX, December 1991, 386–393.
138. L. Arge, O. Procopiuc, S. Ramaswamy, T. Suel, J. Vahrenhold, and J. S. Vitter. "A Unified Approach for Indexed and Non-Indexed Spatial Joins," *Proceedings of the 7th International Conference on Extending Database Technology (EDBT '00)*, Konstanz, Germany, March 2000, published in *Lecture Notes in Computer Science*, Springer, **1777**, Berlin, Germany, 413–429.
139. R. D. Barve, E. F. Grove, and J. S. Vitter. "Application-Controlled Paging for a Shared Cache," *SIAM Journal on Computing*, **29**(4), 2000, 1290–1303. An extended abstract appears in *Proceedings of the 36th Annual IEEE Symposium on Foundations of Computer Science (FOCS '95)*, Milwaukee, WI, October 1995, 204–213.
140. R. D. Barve, M. Kallahalla, P. Varman, and J. S. Vitter. "Competitive Analysis of Buffer Management Algorithms for Parallel I/O Systems," *Journal of Algorithms*, **36**(2), August 2000, 152–181. An extended abstract appears in "Competitive Parallel Disk Prefetching and Buffer Management," *Proceedings of the 5th Annual ACM-IEEE Workshop on I/O in Parallel And Distributed Systems (IOPADS '97)*, San Jose, CA, November 1997, 47–56.
141. Y. Matias, J. S. Vitter, and M. Wang. "Dynamic Maintenance of Wavelet-Based Histograms," *Proceedings of the 26th International Conference on Very Large Databases (VLDB '00)*, Cairo, Egypt, September 2000, 101–110.

142. P. K. Agarwal, L. Arge, J. Erickson, P. G. Franciosa, and J. S. Vitter. "Efficient Searching with Linear Constraints," *Journal of Computer and System Sciences*, **61**(2), October 2000, 194–216. An extended abstract appears in *Proceedings of the 17th Annual ACM Symposium on Principles of Database Systems (PODS '98)*, Seattle, WA, June 1998, 169–178.
143. A. Natsev, J. R. Smith, Y. C. Chang, C. S. Li, and J. S. Vitter. "Constrained Querying of Multimedia Databases," *Proceedings of the Storage and Retrieval for Media Databases Conference, IS&T/SPIE 2001 Symposium on Electronic Imaging Science & Technology*, **4315**, San Jose, CA, January 2001.
144. A. Natsev, A. Chadha, B. Soetarman, J. S. Vitter. "CAMEL: Concept Annotated iMagE Libraries," *Proceedings of the Storage and Retrieval for Media Databases Conference, IS&T/SPIE 2001 Symposium on Electronic Imaging Science & Technology*, **4315**, San Jose, CA, January 2001.
145. L. Arge, J. S. Chase, L. Toma, J. S. Vitter, R. Wickremesinghe, P. Halpin, and D. Urban. "Digital Terrain Analysis for Massive Grids," *Proceedings of the 16th Annual Symposium of the U.S. Chapter of International Association of Landscape Ecology (US-IALE)*, Pattern, Process, Scale, & Hierarchy: Interactions in Human-Dominated and Natural Landscapes, Tempe, AZ, April 2001.
146. L. Arge, L. Toma, and J. S. Vitter. "I/O-Efficient Algorithms for Problems on Grid-Based Terrains," *ACM Journal of Experimental Algorithmics*, **6**(1), 2001. An extended abstract appears in *Proceedings of the 2nd Workshop on Algorithm Engineering and Experimentation (ALENEX '00)*, San Francisco, CA, January 2000, and in "External-Memory Algorithms for Drainage Network Computation," Abstracts of the 4th CGC Workshop on Computational Geometry, Center for Geometric Computing, Johns Hopkins University, Baltimore, MD, October 1999.
147. J. S. Vitter. "External Memory Algorithms and Data Structures: Dealing with Massive Data," *ACM Computing Surveys*, **33**(2), June 2001, 209–271. Shorter versions appear in "External Memory Algorithms," invited paper in *Proceedings of the 6th Annual European Symposium on Algorithms (ESA '98)*, Venice, August 1998, published in *Lecture Notes in Computer Science*, **1461**, Springer, Berlin, Germany, 1–25, and in an invited tutorial in *Proceedings of the 17th Annual ACM Symposium on Principles of Database Systems (PODS '98)*, Seattle, WA, June 1998, 119–128.
148. P. K. Agarwal, L. Arge, O. Procopiuc, and J. S. Vitter. "A Framework for Index Bulk Loading and Dynamization," *Proceedings of the 28th Annual International Colloquium on Automata, Languages, and Programming (ICALP '01)*, Crete, Greece, July 2001, published in *Lecture Notes in Computer Science*, **2076**, Springer, Berlin, Germany, 115–127.
149. L. Lim, M. Wang, S. Padmanabhan, J. S. Vitter, and R. Agarwal. "Characterizing Web Document Change," *Proceedings of the 2nd International Conference on Web-Age Information Management*, Xi'an, China, July 2001, published in *Lecture Notes in Computer Science*, **2118**, Springer, Berlin, Germany, 133–144.
150. M. Wang, J. S. Vitter, L. Lim, and S. Padmanabhan. "Wavelet-Based Cost Estimation for Spatial Queries," *Proceedings of the 7th International Symposium on Spatial and Temporal Databases (SSTD '01)*, Redondo Beach, CA, July 2001, 175–196.
151. D. A. Hutchinson, P. Sanders, and J. S. Vitter. "The Power of Duality for Prefetching and Sorting with Parallel Disks," *Proceedings of the 13th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '01)*, Crete, Greece, July 2001, 334–335.
152. A. Natsev, J. R. Smith, Y.-C. Chang, C.-S. Li, and J. S. Vitter. "Supporting Incremental Join Queries on Ranked Inputs," *Proceedings of the 27th International Conference on Very Large Databases (VLDB '01)*, Rome, Italy, September 2001, 281–290.
153. A. Natsev, G. Fuh, W. Chen, C.-H. Chiu, and J. S. Vitter. "Aggregate Predicates and Ranked Search in DBMS," *Proceedings of the 13th Australian Database Conference (ADC '02)*, Melbourne, Australia, January–February 2002.
154. R. Grossi and J. S. Vitter. "Compressed Indexes for Fast Search in Sequences," invited keynote paper in *Proceedings of the 6th Joint Conference on Information Sciences (JCIS '02)*, Durham, NC, March 2002.

155. R. D. Barve and J. S. Vitter. “A Simple and Efficient Parallel Disk Mergesort,” invited paper in special issue on parallel algorithms and architectures in *Theory of Computing Systems*, **35**(2), March/April 2002, 189–215. An extended abstract appears in *Proceedings of the 11th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '99)*, St. Malo, France, June 1999, 232–241.
156. L. Arge, K. H. Hinrichs, J. Vahrenhold, and J. S. Vitter. “Efficient Bulk Operations on Dynamic R-trees,” special issue on experimental algorithmics in *Algorithmica*, **33**(1), May 2002, 104–128. An extended abstract appears in *Proceedings of the 1st Workshop on Algorithm Engineering and Experimentation (ALENEX '99)*, Baltimore, MD, January 1999, 328–348.
157. R. Wickremesinghe, J. S. Chase, and J. S. Vitter. “Distributed Computing with Load-Managed Active Storage,” *Proceedings of the 11th Annual IEEE International Symposium on High Performance Distributed Computing (HPDC '02)*, Edinburgh, Scotland, July 2002, 13–23.
158. L. Lim, M. Wang, S. Padmanabhan, J. S. Vitter, and R. Parr. “XPathLearner: An On-Line Self-Tuning Markov Histogram for XML Path Selectivity Estimation,” *Proceedings of the 28th International Conference on Very Large Databases (VLDB '02)*, Hong Kong, China, August 2002, 442–453.
159. L. Arge, O. Procopiuc, and J. S. Vitter. “Implementing I/O-Efficient Data Structures Using TPIE,” *Proceedings of the 10th Annual European Symposium on Algorithms (ESA '02)*, Rome, Italy, September 2002, published in *Lecture Notes in Computer Science*, Springer, **2461**, Berlin, Germany, 88–100.
160. R. Wickremesinghe, L. Arge, J. S. Chase, and J. S. Vitter. “Efficient Sorting using Registers and Caches,” invited paper in special issue of *ACM Journal of Experimental Algorithmics*, **7**(9), 2002. An earlier and shorter version appears in L. Arge, R. Barve, J. S. Chase, J. S. Vitter, and R. Wickremesinghe. “Efficient Sorting using Registers and Caches,” *Proceedings of the 4rd Workshop on Algorithm Engineering (WAE '00)*, Saarbrücken, Germany, September 2000, published in *Lecture Notes in Computer Science*, **1982** Springer, Berlin, Germany, 51–62.
161. L. Arge and J. S. Vitter. “Optimal External Memory Interval Management,” *SIAM Journal on Computing*, **32**(6), 2003, 1488–1508. An extended abstract appears in “Optimal Dynamic Interval Management in External Memory,” *Proceedings of the 37th Annual IEEE Symposium on Foundations of Computer Science (FOCS '96)*, Burlington, VT, October 1996, 560–569. Also appears in Abstracts of the 1st CGC Workshop on Computational Geometry, Center for Geometric Computing, Johns Hopkins University, Baltimore, MD, October 1996.
162. O. Procopiuc, P. K. Agarwal, L. Arge, and J. S. Vitter. “Bkd-tree: A Dynamic Scalable kd-tree,” *Proceedings of the 8th International Symposium on Spatial and Temporal Databases (SSTD '03)*, Santorini island, Greece, July 2003, 46–65.
163. L. Lim, M. Wang, and J. S. Vitter. “SASH: A Self-Adaptive Histogram Set for Dynamically Changing Workloads,” *Proceedings of the 29th International Conference on Very Large Databases (VLDB '03)*, Berlin, Germany, September 2003, 369–380.
164. Y. Matias, J. S. Vitter and W.-C. Ni. “Dynamic Generation of Discrete Random Variates,” *Theory of Computing Systems*, **36**(4), 2003, 329–358. An extended abstract appears in *Proceedings of the 4th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '93)*, Austin, TX, January 1993, 361–370.
165. L. Arge, J. S. Chase, L. Toma, J. S. Vitter, R. Wickremesinghe, P. Halpin, and D. Urban, “Efficient Flow Computation on Massive Grid Terrain Datasets,” *Geoinformatica*, **7**(4), December 2003, 283–313. An extended abstract appears in “Flow Computation on Massive Grids,” *Proceedings of the 9th ACM International Symposium on Advances in Geographic Information Systems (ACM-GIS '01)* Atlanta, GA, November 2001, 82–87.
166. L. Foschini, R. Grossi, A. Gupta, and J. S. Vitter. “Fast Compression with a Static Model in High-Order Entropy,” *Proceedings of the 2004 IEEE Data Compression Conference (DCC '04)*, Snowbird, UT, March 2004, 62–71.
167. T. M. Ghanem, R. Shah, M. F. Mokbel, W. G. Aref, and J. S. Vitter. “Bulk Operations for Space-Partitioning Trees,” *Proceedings of the 20th Annual IEEE International Conference on Data Engineering (ICDE '04)*, Boston, March–April 2004, 29–41.

168. R. Shah, P. J. Varman, and J. S. Vitter. "Online Algorithms for Prefetching and Caching in Parallel Disks," *Proceedings of the 16th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '04)*, Barcelona, Spain, June 2004, 255–264.
169. I. Ilyas, R. Shah, W. G. Aref, J. S. Vitter, and A. Elmagarmid. "Rank-aware Query Optimization," *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data (SIGMOD '04)*, Paris, France, June 2004, 203–214.
170. S. Muthukrishnan, R. Shah, and J. S. Vitter. "Mining Deviants in Time Series Data Streams," *Proceedings of the 16th International Conference on Scientific and Statistical Database Management (SS-DBM '04)*, Santorini Island, Greece, June 2004, 41–50.
171. R. Cheng, Y. Xia, S. Prabhakar, R. Shah, and J. S. Vitter. "Efficient Indexing Methods for Probabilistic Threshold Queries over Uncertain Data," *Proceedings of the 30th International Conference on Very Large Databases (VLDB '04)*, Toronto, CA, August 2004, 876–887.
172. M. Wang, B. Iyer, and J. S. Vitter. "Scalable Mining for Classification Rules in Relational Databases," *Herman Rubin Festschrift, Lecture Notes Monograph Series*, **45**, Institute of Mathematical Statistics, Hayward, CA, Fall 2004. An extended abstract appears in *Proceedings of the International Database Engineering & Application Symposium (IDEAS '98)*, Cardiff, Wales, July 1998, 58–67. A shorter version appears in *Proceedings of the ACM SIGMOD Data Mining and Knowledge Discovery Workshop (DMKD '98)*, Seattle, WA, June 1998.
173. M. Ouzzani, W. G. Aref, E. Bertino, A. C. Catlin, C. W. Clifton, W.-K. Hon, A. K. Elmagarmid, A. Ghafoor, S. E. Hambrusch, S. Prabhakar, J. S. Vitter, X. Zhang. "The Indiana Center for Database Systems at Purdue University," *SIGMOD Record*, **34**(2), June 2005, 53–58.
174. S. Anastasiadis, P. J. Varman, J. S. Vitter, and K. Yi. "Optimal Lexicographic Shaping of Aggregate Streaming Data," *IEEE Transactions on Computers*, **54**(4), April 2005, 398–408. An extended abstract appears in S. Anastasiadis, P. J. Varman, and J. S. Vitter. "Lexicographically Optimal Smoothing for Broadband Traffic Multiplexing," *Proceedings of the 21st Annual ACM Symposium on Principles of Distributed Computing (PODC '02)*, July 2002, Monterey, CA, 68–77.
175. D. A. Hutchinson, P. Sanders, and J. S. Vitter. "Duality Between Prefetching and Queued Writing with Parallel Disks," *SIAM Journal on Computing*, **34**(6), 1443–1463, June 2005. An extended abstract appears in *Proceedings of the 9th Annual European Symposium on Algorithms (ESA '01)*, Århus, Denmark, August 2001, published in *Lecture Notes in Computer Science*, **2161**, Springer, Berlin, Germany, 62–73.
176. R. Grossi and J. S. Vitter. "Compressed Suffix Arrays and Suffix Trees with Applications to Text Indexing and String Matching," *SIAM Journal on Computing*, **35**(2), 2005, 378–407. An extended abstract appears in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC '00)*, Portland, OR, May 2000, 397–406.
177. R. Shah, P. J. Varman, J. S. Vitter. "On Competitive Online Read-Many Parallel Disks Scheduling," *Proceedings of the 17th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '05)*, Las Vegas, NV, July 2005, 217.
178. L. Lim, M. Wang, and J. S. Vitter. "CXHist: An On-line Classification-Based Histogram for XML String Selectivity Estimation," *Proceedings of the 31st International Conference on Very Large Databases (VLDB '05)*, Trondheim, Norway, August–September 2005, 1187–1198.
179. Y. Matias, E. Segal, and J. S. Vitter. "Efficient Bundle Sorting," *SIAM Journal on Computing*, **36**(2), 2006, 394–410. An extended abstract appears in *Proceedings of the 11th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '00)*, San Francisco, CA, January 2000, 839–848.
180. A. Gupta, W.-K. Hon, R. Shah, and J. S. Vitter. "Compressed Dictionaries: Space Measures, Data Sets, and Experiments," *Proceedings of the 5th International Workshop on Experimental Algorithmics (WEA '06)*, Menorca, Spain, May 2006, 158–169.
181. J. S. Vitter and D. A. Hutchinson. "Distribution Sort with Randomized Cycling," *Journal of the ACM*, **53**(7), July 2006, 656–680. An extended abstract appears in *Proceedings of the 12th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '01)*, Washington, DC, January 2001, 77–86.

182. L. Foschini, R. Grossi, A. Gupta, and J. S. Vitter. “When Indexing Equals Compression: Experiments on Suffix Arrays and Trees,” *ACM Transactions on Algorithms*, **2**(4), 2006, 611–639. An extended abstract appears in R. Grossi, A. Gupta, and J. S. Vitter, “When Indexing Equals Compression: Experiments with Compressing Suffix Arrays and Applications,” *Proceedings of the 15th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '04)*, New Orleans, LA, January 2004, 636–645..
183. R. Cheng, Y. Xia, S. Prabhakar, R. Shah, and J. S. Vitter. “Efficient Join Processing over Uncertain-Valued Attributes,” *Proceedings of the 2006 ACM Conference on Information and Knowledge Management (CIKM '06)*, Arlington, VA, November 2006, 738–747.
184. I. Ilyas, W. G. Aref, A. K. Elmagarmid, H. G. Elmongui, R. Shah, and J. S. Vitter. “Adaptive Rank-aware Query Optimization in Relational Databases,” *ACM Transactions on Database Systems*, **31**(4), December 2006, 1257–1304.
185. L. Arge, D. E. Vengroff, and J. S. Vitter. “External-Memory Algorithms for Processing Line Segments in Geographic Information Systems,” *Algorithmica*, **47**(1), January 2007, 1–25. An extended abstract appears in *Proceedings of the 3rd Annual European Symposium on Algorithms (ESA '95)*, Corfu, Greece, September 1995, published in Lecture Notes in Computer Science, **979**, Springer, Berlin, Germany, 295–310.
186. L. Lim, M. Wang, S. Padmanabhan, J. S. Vitter, and R. Agarwal. “Efficient Update of Indexes for Dynamically Changing Web Documents,” *World Wide Web*, **10**(1), March 2007, 37–69. An extended abstract appears in “Dynamic Maintenance of Web Indexes Using Landmarks,” *Proceedings of the 12th International World Wide Web Conference (WWW '03)*, Budapest, May 2003, 102–111.
187. A. Gupta, W.-K. Hon, R. Shah, and J. S. Vitter. “Compressed Data Structures: Dictionaries and the Data-Aware Measures,” *Theoretical Computer Science*, **387**(3), November 2007, 313–331. An extended abstract appears in *Proceedings of the 2006 IEEE Data Compression Conference (DCC '06)*, Snowbird, UT, March 2006, 213–222.
188. W.-K. Hon, R. Shah, and J. S. Vitter. “Compression, Indexing, and Retrieval for Massive String Data,” invited keynote paper in *Proceedings of the 19th Annual Conference on Combinatorial Pattern Matching (CPM '10)*, New York, NY, June 2010, published in Lecture Notes in Computer Science, **6129**, Springer, Berlin, Germany, 260–274.
189. J. S. Vitter. “Building Synergies,” in *Sustaining and Enhancing the Research Mission of Public Universities* (edited by M. L. Rice), Merrill Series on the Research Mission of Public Universities, **115**, Merrill Advanced Studies Center, the University of Kansas, Lawrence, KS, 26–33.
190. M. O. Külekci, J. S. Vitter, and B. Xu. “Fast Pattern Matching via  $k$ -bit Filtering Based Text Decomposition,” invited paper from *ISCIS '10* in *The Computer Journal*, December 20, 2010. An extended abstract appears in “Boosting Pattern Matching Performance via  $k$ -bit Filtering,” *Proceedings of the 25th International Symposium on Computer and Information Sciences (ISCIS '10)*, London, September 2010, published in Lecture Notes in Electrical Engineering, **1**, **62**, Springer, Berlin, Germany, 27–32.
191. M. O. Külekci, W.-K. Hon, R. Shah, J. S. Vitter, and B. Xu. “ $\Psi$ -RA: a Parallel Sparse Index for Genomic Read Alignment,” invited paper from *BIBM '10* in *BMC Genomics*, to appear. An extended abstract appears in “PSI-RA: A Parallel Sparse Index for Read Alignment on Genomes,” *Proceedings of 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM '10)*, Hong Kong, December 2010, 663–668.
192. M. O. Külekci, J. S. Vitter, and B. Xu. “Efficient Maximal Repeat Finding Using the Burrows-Wheeler Transform and Wavelet Tree,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, to appear. An extended abstract appears in “Time- and Space-efficient Maximal Repeat Finding Using the Burrows-Wheeler Transform and Wavelet Trees,” *Proceedings of 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM '10)*, Hong Kong, December 2010, 622–625.
193. M. H. Nodine and J. S. Vitter. “Optimal Deterministic Sorting on Parallel Disks,” submitted to journal. An extended abstract appears in “Deterministic Distribution Sort in Shared and Distributed Memory Multiprocessors,” *Proceedings of the 5th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '93)*, Velen, Germany, June–July 1993, 120–129.

194. M. H. Nodine and J. S. Vitter. “Optimal Deterministic Sorting on Parallel Processors and Parallel Memory Hierarchies,” submitted to journal. An extended abstract appears in “Deterministic Distribution Sort in Shared and Distributed Memory Multiprocessors,” *Proceedings of the 5th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '93)*, Velen, Germany, June–July 1993, 120–129.
195. R. Grossi, A. Gupta, and J. S. Vitter. “High-Order Entropy-Compressed Text Indexes,” being partitioned with more recent work for submission to journals. An extended abstract appears in *Proceedings of the 14th Annual SIAM/ACM Symposium on Discrete Algorithms (SODA '03)*, Baltimore, MD, January 2003, 841–850.
196. A. Gupta, W.-K. Hon, R. Shah, and J. S. Vitter. “A Framework for Dynamizing Succinct Data Structures,” in preparation. An extended abstract appears in *Proceedings of the 34th Annual International Colloquium on Automata, Languages, and Programming (ICALP '07)*, Wrocław, Poland, July 2007, published in Lecture Notes in Computer Science, **4596**, Springer, Berlin, Germany, 521–532.
197. W.-K. Hon, T.-W. Lam, R. Shah, S.-L. Tam, and J. S. Vitter. “Cache-Oblivious Index for Approximate String Matching,” in preparation. An extended abstract appears in *Proceedings of the 16th Annual Conference on Combinatorial Pattern Matching (CPM '07)*, London, Ontario, Canada, July 2007, published in Lecture Notes in Computer Science, **4580**, Springer, Berlin, Germany, 40–51.
198. R. Grossi, A. Gupta, and J. S. Vitter. “Nearly Tight Bounds on the Encoding Length of the Burrows-Wheeler Transform,” in preparation. An extended abstract appears in *Proceedings of the 5th Workshop on Analytical Algorithmics and Combinatorics (ANALCO '08)*, San Francisco, CA, January 2008, 191–202.
199. M. Y. Eltabakh, W.-K. Hon, R. Shah, W. Aref, and J. S. Vitter. “The SBC-tree: An Index for Run-Length Compressed Sequences,” in preparation. An extended abstract appears in *Proceedings of the 11th International Conference on Extending Database Technology (EDBT '08)*, Nantes, France, March 2008, 523–534.
200. Y.-F. Chien, W.-K. Hon, R. Shah, and J. S. Vitter. “Geometric Burrows-Wheeler Transform: Linking Range Searching and Text Indexing,” in preparation. An extended abstract appears in *Proceedings of the 2008 IEEE Data Compression Conference (DCC '08)*, Snowbird, UT, March 2008.
201. W.-K. Hon, T.-W. Lam, R. Shah, S.-L. Tam, and J. S. Vitter. “Compressed Index for Dictionary Matching,” in preparation. An extended abstract appears in *Proceedings of the 2008 IEEE Data Compression Conference (DCC '08)*, Snowbird, UT, March 2008, 23–32.
202. P. Ferragina, R. Grossi, A. Gupta, R. Shah, and J. S. Vitter. “On Searching Compressed String Collections Cache-Obliviously,” in preparation. An extended abstract appears in *Proceedings of the 27th Annual ACM Symposium on Principles of Database Systems (PODS '08)*, Vancouver, Canada, June 2008, 181–190.
203. W.-K. Hon, R. Shah, P. J. Varman, and J. S. Vitter. “Tight Competitive Ratios for Parallel Disk Prefetching,” in preparation. An extended abstract appears in *Proceedings of the 20th Annual ACM Symposium on Parallel Algorithms and Architectures (SPAA '08)*, Munich, Germany, June 2008, 352–361.
204. W.-K. Hon, R. Shah, S. V. Thankachan, and J. S. Vitter. “On Entropy-Compressed Text Indexing in External Memory,” in preparation. An extended abstract appears in *Proceedings of the 16th International Conference on String Processing and Information Retrieval (SPIRE '09)*, Saariselkä, Finland, August 2009, published in Lecture Notes in Computer Science, **5721**, Springer, Berlin, Germany, 75–89.
205. W.-K. Hon, R. Shah, and J. S. Vitter. “Space-Efficient Framework for Top- $k$  String Retrieval Problems,” in preparation. An extended abstract appears in *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS '09)*, Atlanta, GA, October 2009, 713–722.
206. S.-Y. Chiu, W.-K. Hon, R. Shah, and J. S. Vitter. “I/O-efficient Compressed Text Indexes: From Theory to Practice,” in preparation. An extended abstract appears in *Proceedings of the 2010 IEEE Data Compression Conference (DCC '10)*, Snowbird, UT, March 2010.

207. W.-K. Hon, Tsung-Han Ku, R. Shah, S. V. Thankachan, and J. S. Vitter. “Faster Compressed Dictionary Matching,” in preparation. An extended abstract appears in *Proceedings of the 17th International Conference on String Processing and Information Retrieval (SPIRE '10)*, Los Cabos, Mexico, October 2010, published in Lecture Notes in Computer Science, Springer, Berlin, Germany.
208. S. V. Thankachan, R. Shah, W.-K. Hon, and J. S. Vitter. “String Retrieval for Multipattern Queries,” in preparation. An extended abstract appears in *Proceedings of the 17th International Conference on String Processing and Information Retrieval (SPIRE '10)*, Los Cabos, Mexico, October 2010, published in Lecture Notes in Computer Science, Springer, Berlin, Germany.
209. W.-K. Hon, Tsung-Han Ku, R. Shah, S. V. Thankachan, and J. S. Vitter. “Compressed Dictionary Matching With One Error,” in preparation. An extended abstract appears in *Proceedings of the 2011 IEEE Data Compression Conference (DCC '11)*, Snowbird, UT, March 2011, 113–122.
210. R. Grossi, J. S. Vitter, and B. Xu. “Wavelet Trees: From Theory to Practice,” in preparation. An extended abstract appears in *Proceedings of the International Conference on Data Compression, Communication and Processing (CCP '11)*, Palinuro, Italy, June 2011, 10 pages, published in Lecture Notes in Computer Science, Springer, Berlin, Germany.
211. M. Patil, S. V. Thankachan, R. Shah, W.-K. Hon, Tsung-Han Ku, J. S. Vitter, and S. Chandrasekaran. “Inverted Indexes for Phrases and Strings,” in preparation. An extended abstract appears in *Proceedings of the 34th Annual ACM SIGIR Conference (SIGIR '11)*, Beijing, July 2011.
212. T.-H. Ku, W.-K. Hon, R. Shah, S. V. Thankachan, and J. S. Vitter. “Compressed Text Indexing With Wildcards,” in preparation. An extended abstract appears in *Proceedings of the 18th International Conference on String Processing and Information Retrieval (SPIRE '11)*, Pisa, Italy, October 2011, published in Lecture Notes in Computer Science, Springer, Berlin, Germany.
213. J. S. Vitter. “Structure + Style = Communication,” in preparation.
214. D. Gardy and J. S. Vitter. “A Comparison of the Finite and Infinite Models for Trie Structures,” in preparation.

### Technical Reports (information not available in previous citations)

215. J. S. Vitter. Departmental Reports, Computer Services Department, Standard Oil Company of California, summers 1976–1977.
216. J. S. Vitter. *Analysis of Coalesced Hashing*, Ph.D. dissertation, Stanford University. Also appears as Technical Report STAN-CS-80-817, Department of Computer Science, Stanford University, August 1980.
217. J. S. Vitter. “Search Mechanisms for Optical Disks,” Confidential Memo, IBM Palo Alto Scientific Center, March 1983.
218. J. S. Vitter. “US&R: A New Framework for Redoing,” Technical Report CS-83-18, Department of Computer Science, Brown University, revised February 1984.
219. J. S. Vitter. “A Software I/O Package for Optical Disks,” Confidential Memo, IBM Palo Alto Scientific Center, July 1983.
220. E. Lebras, F. P. Preparata, C. Puech, M. Teillaud, J. S. Vitter, and M. Yvinec. “A New Technique for the Display of Isothetic Three-Dimensional Rectangles” Technical Report LIENS-87-6, Département de Mathématiques et d’Informatique, Ecole Normale Supérieure, May 1987.
221. E. A. Lamagna, R. A. Ravenscroft, and J. S. Vitter. “Sum Amusements: A Case Study from the Analysis of Algorithms” Technical Report CS-90-33, Department of Computer Science, Brown University, November 1990.
222. J.-H. Lin and J. S. Vitter. “ $\epsilon$ -Approximations with Minimum Packing Constraint Violation,” Technical Report CS-92-29, Department of Computer Science, Brown University, June 1992.

- 223. J. S. Vitter and W.-C. Ni. “Dynamic Generation of Discrete Random Variates,” Technical Report CS-92-36, Department of Computer Science, Brown University, August 1992.
- 224. J. S. Vitter. “Average-Case Analysis of Prediction,” Dagstuhl-Seminar on Average-Case Analysis of Algorithms, Report No. 68 (edited by Ph. Flajolet, R. Kemp, and H. Prodinger), Schloss Dagstuhl, Wadern, Germany, July 1993.
- 225. J. S. Vitter. “Algorithms for Processing Line Segments in External Memory, with Applications to Databases and Geographic Information Systems,” Dagstuhl-Seminar on Computational Cartography, Report No. 160 (edited by C. Gold, J. Snoeyink, and F. Wagner), Schloss Dagstuhl, Wadern, Germany, November 1996.
- 226. Y. Matias, J. S. Vitter and W.-C. Ni. “Dynamic Generation of Discrete Random Variates,” Technical Report CS-1997-12, Department of Computer Science, Duke University, May 1997.
- 227. R. D. Barve and J. S. Vitter. “External Memory Algorithms with Dynamically Changing Memory Allocations,” Technical Report CS-1998-09, Department of Computer Science, Duke University, June 1998.
- 228. L. Arge and J. S. Vitter. “External Memory Point Location and Level-Balanced B-trees,” Dagstuhl-Seminar on Computational Geometry, Report No. 233 (edited by M. Goodrich, R. Klein, and R. Seidel), Schloss Dagstuhl, Wadern, Germany, March 1999.
- 229. J. S. Vitter. “Geometric Searching in Massive Data,” Proceedings of the 15th European Workshop on Computational Geometry, Antibes-Juan-les-Pins, France, March 1999.
- 230. A. Condon, H. Edelsbrunner, E. A. Emerson, L. Fortnow, S. Haber, R. Karp, D. Leivant, R. Lipton, N. Lynch, I. Parberry, C. Papadimitriou, M. Rabin, A. Rosenberg, J. S. Royer, J. E. Savage, A. L. Selman, C. Smith, E. Tardos, and J. S. Vitter. “Challenges for Theory of Computing,” Report of an NSF-Sponsored Workshop on Research in Theoretical Computer Science, April 1999.
- 231. J. S. Vitter. “Multiresolution Approximation Techniques for Database Systems,” Dagstuhl-Seminar on Data Structures, Report No. 267 (edited by S. Albers, I. Munro, and P. Widmayer), Schloss Dagstuhl, Wadern, Germany, February-March 2000.
- 232. J. S. Vitter. “Entropy-Compressed Indexes for Sequences,” Dagstuhl-Seminar on Data Structures, Report No. 335 (edited by S. Albers, R. Sedgewick, and P. Widmayer), Schloss Dagstuhl, Wadern, Germany, February-March 2002.
- 233. S. V. Anastasiadis, R. G. Wickremesinghe, J. S. Chase, and J. S. Vitter, “Lerna — An Active Storage Framework for Flexible Data Access and Management,” Technical Report CS-2003-7, Department of Computer Science, Duke University, May 2003.
- 234. W.-K. Hon, R. Shah, and J. S. Vitter, “Ordered Pattern Matching: Towards Full-Text Retrieval,” Technical Report 06-008, Department of Computer Sciences, Purdue University, March 2006.
- 235. A. Gupta, W.-K. Hon, R. Shah, and J. S. Vitter. “Dynamic Rank/Select Dictionaries with Applications to XML Indexing,” Technical Report 06-014, Department of Computer Sciences, Purdue University, July 2006.

## 14. Invited Talks since 1990

- 1990 “Optimal Disk I/O with Parallel Block Transfer,” Brandeis University, Waltham, MA.
- “Geometric Algorithms for High-Performance Computing,” Air Force Office of Scientific Research, Bolling Air Force Base, D.C.
- “Optimal Disk I/O with Parallel Block Transfer,” AT&T Bell Laboratories, Murray Hill, NJ.
- “Optimal Disk I/O with Parallel Block Transfer,” Ecole Normale Supérieure, Paris, France.
- “Optimal Disk I/O with Parallel Block Transfer,” Tulane University, New Orleans, LA.
- 1991 “Theory Curriculum,” Respondent, Computer Science Curriculum Workshop, Dedham, MA.

- “Parallelism in Point Location Problems,” Ecole Normale Supérieure, Paris, France.
- “Efficient Memory Access in Large-Scale Computation,” keynote address at 1991 Symposium on Theoretical Aspects of Computer Science (STACS '91), Hamburg, West Germany.
- “Optimal Algorithms for Parallel Large-Scale Memory,” University of Maryland, College Park, MD.
- “Optimal Algorithms for Parallel Large-Scale Memory,” Center of Excellence in Space Data and Information Sciences, NASA Goddard Space Flight Center, Greenbelt, MD.
- “Dynamic Arithmetic Coding,” IEEE Data Compression Conference (DCC '91), Snowbird, UT.
- “Optimal Algorithms for Parallel Large-Scale Memory,” Washington University at St. Louis, St. Louis, MO.
- “Optimal Prefetching via Data Compression,” University of Washington, Seattle, WA.
- “Practical Implementations of Arithmetic Coding,” 3rd International Conference on Advances in Communication and Control Systems (COMCON '91), Victoria, Canada.
- “Optimal Algorithms for Parallel Large-Scale Memory,” Supercomputing Research Center, Bowie, MD.
- “Optimal Prefetching via Data Compression,” University of Maryland, College Park, MD.
- 1992 “ $\epsilon$ -Approximations with Minimum Packing Constraint Violation,” INRIA, Sophia Antipolis, France.
- “Optimal Prefetching via Data Compression,” University of Notre Dame, Notre Dame, IN.
- “Efficient Memory Access in Large-Scale Computation,” Duke University, Durham, NC.
- “Optimal Prefetching via Data Compression,” University of Victoria, Victoria, Canada.
- “Optimal Prefetching via Data Compression,” University of Southwestern Louisiana, Lafayette, LA.
- “ $\epsilon$ -Approximations with Minimum Packing Constraint Violation,” Georgia Institute of Technology, Atlanta, GA.
- “Mathematical Results Specific to Neural Networks,” panel member, Neural Network Workshop, sponsored by Rutgers, East Brunswick, NJ.
- “A Theory for Memory-Based Learning,” Massachusetts Institute of Technology, Cambridge, MA.
- “Paradigms for Optimal Sorting with Multiple Disks and Memory Hierarchies,” keynote address at 7th Maryland Theoretical Computer Science Day, Johns Hopkins University, Baltimore, MD.
- “The Design of Lossless Image Compression Systems,” Workshop on Data and Image Compression Needs and Uses in the Scientific Community, Center of Excellence in Space Data and Information Sciences, NASA Goddard Space Flight Center, Greenbelt, MD.
- 1993 “Locality, Dynamic, and Prediction Issues in DIS,” panel member, ARO Workshop on Virtual, Distributed Interactive Simulation, Research Triangle Park, NC.
- “Obstacles in the Implementation of Parallel Algorithms,” panel member, Workshop on Parallel I/O and Databases, Dartmouth Institute for Advanced Graduate Studies (DAGS '93), Hanover, NH.
- “Load Balancing Paradigms for Optimal Use of Parallel Disks and Parallel Memory Hierarchies,” Workshop on Parallel I/O and Databases, Dartmouth Institute for Advanced Graduate Studies (DAGS '93), Hanover, NH.
- “Average-Case Analysis of Prediction,” Dagstuhl-Seminar on Average-Case Analysis of Algorithms, Schloss Dagstuhl, Wadern, Germany.
- “Paradigms for Optimal Sorting and Computational Geometry in Large-Scale Parallel Memories,” Max Planck Institute, Saarbrücken, Germany.
- “Models for Parallel Secondary and Hierarchical Storage,” Workshop on Models, Architectures, and Technologies for Parallel Computation, DIMACS, Rutgers University, New Brunswick, NJ.
- “Load Balancing Paradigms for Optimal Use of Parallel Disks and Parallel Memory Hierarchies,” Stanford University, Stanford, CA.
- “Optimal Prediction via Data Compression,” University of Texas at Dallas, Dallas, TX.
- “Load Balancing Paradigms for Optimal Use of Parallel Disks and Parallel Memory Hierarchies,” keynote address at the Workshop on Algorithmic Research in the Midsouthwest (WARM '93), University of North Texas, Denton, TX.
- 1994 “Predictive Techniques for Caching and Locality Management,” Microsoft Corporation, Redmond, WA.

- “Efficient Processing of Large-Scale Data,” Mathematisches Forschungsinstitut Oberwolfach, Germany.
- “Data Compression and Applications,” Air Force Office of Scientific Research, Bolling Air Force Base, D.C.
- “Efficient Processing of Large-Scale Data in External Memory,” Distinguished Lecturer Series, Johns Hopkins University, Baltimore, MD.
- “How to Predict Well,” Tulane University, New Orleans, LA.
- “How to Predict Well,” Supercomputing Research Center, Bowie, MD.
- “Future Communication Issues in Dealing with Large-Scale Data,” Army Research Office Math/CS Investment Strategy Meeting, Lake Buena Vista, FL.
- 1995 “Data Compression Techniques for Networks,” Air Force Office of Scientific Research Initiative Planning Meeting, Raleigh, NC.
- “Future Trends and Issues in I/O,” panel moderator, Workshop on Modeling and Specification of I/O (MSIO '95), as part of the 7th IEEE Symposium on Parallel and Distributed Processing (SPDP '95), San Antonio, TX.
- 1996 “I/O Environments for Geometric Computation,” Army Research Office MURI Advisory Board, Johns Hopkins University, Baltimore, MD.
- “Theory and Practice of I/O-Efficient Computation,” Distinguished Lecturer Series, Northwestern University, Chicago, IL.
- “Predicting Fast and Reliably,” keynote address at the Midwest Theory Day, Washington University at St. Louis, St. Louis, MO.
- “Predicting Fast and Reliably,” University of Venice, Venice, Italy.
- “I/O-Efficient Computation,” University of Rome, La Sapienza, Rome, Italy.
- “Algorithms for Processing Line Segments in External Memory, with Applications to Databases and Geographic Information Systems,” Dagstuhl-Seminar on Computational Cartography, Schloss Dagstuhl, Wadern, Germany.
- 1997 “I/O-Efficient Geometry with TPIE,” Army Research Office MURI Advisory Board, University of Pennsylvania. Philadelphia, PA.
- “On Sorting Strings in External Memory,” INRIA, Sophia Antipolis, France.
- “Sequence Sorting in Secondary Storage,” International Conference on Compression and Complexity of Sequences, Positano, Italy.
- “Efficient Geometric Processing of Massive Data Sets,” ARL/ARO Workshop on Battlespace Visualization, Army Research Laboratory, Adelphi, MD.
- “System Requirements and Support for Battlespace Visualization,” panel member, ARL/ARO Workshop on Battlespace Visualization, Army Research Laboratory, Adelphi, MD.
- “Efficient Geometric Processing of Massive Data Sets,” ARL/CGC Workshop, Johns Hopkins University, Baltimore, MD.
- “Geographic Information Systems, Terrains, and Visualization,” Army Research Office MURI Advisory Board, Duke University, Durham, NC.
- “Dealing with Massive Spatial Data,” ARL Federated Laboratory Workshop on Database and Object Oriented Technologies, Army Research Laboratory, Adelphi, MD.
- 1998 “External-Memory Algorithms,” Workshop on External Memory Algorithms and Visualization, DIMACS, Rutgers University, New Brunswick, NJ.
- “External-Memory Algorithms,” invited tutorial at 17th Annual ACM Symposium on Principles of Database Systems (PODS '98), Seattle, WA.
- “External Memory Algorithms: Dealing with MASSIVE Data,” two talks, BRICS Workshop on the Theory and Practice of Algorithms for Problems Involving Massive Data Sets, University of Aarhus, Århus, Denmark.
- “External Memory Algorithms: Dealing with MASSIVE Data,” two talks, Department of Computer Science, University of Aalborg, Aalborg, Denmark.

- “Recent Developments in External Memory Algorithms,” Department of Computer Science, University of Münster, Münster, Germany.
- “External Memory Algorithms,” keynote address at 6th Annual European Symposium on Algorithms (ESA ’98), Venice, Italy.
- “Designing External Memory Algorithms to Adapt to Changing Internal Memory Allocations,” University of Florence, Florence, Italy.
- “Range Queries on Spatial Data,” INRIA, Sophia Antipolis, France.
- “External Memory Algorithms: Dealing with MASSIVE Data,” Journée Scientifique et Assemblée Générale de l’Association Française d’Informatique Théorique, Cachan, France.
- “External Memory Algorithms: Dealing with MASSIVE Data,” Séminaire J. Morgenstern, INRIA, Sophia Antipolis, France.
- 1999 “External Memory Point Location and Level-Balanced B-trees,” Dagstuhl-Seminar on Computational Geometry, Schloss Dagstuhl, Wadern, Germany.
- “Geometric Searching in Massive Data,” 15th Annual European Workshop on Computational Geometry (CG ’99), Antibes, France.
- “Geometric Searching in Massive Data,” University of London, London, England.
- “Parallel I/O Algorithms,” 1999 ACM/UMIACS Workshop on Parallel Algorithms (WOPA ’99), as part of the 3rd Federated Computing Research Conference (FCRC ’99), Atlanta, GA.
- “Simple Randomized Mresgerot (Mergesort),” 5th International Seminar on the Mathematical Analysis of Algorithms (SMAA ’99), Barcelona, Spain.
- “Online Data Structures in External Memory,” Max Planck Institute, Saarbrücken, Germany.
- “Affichage et Elimination des Parties Cachées,” Ecole Normale Supérieure, Paris, France.
- “Online Geometric Data Structures in External Memory,” University of Genoa, Genoa, Italy.
- “Online Data Structures in External Memory,” University of Tübingen, Tübingen, Germany.
- “Online Data Structures in External Memory,” keynote address at 26th Annual International Colloquium on Automata, Languages, and Programming (ICALP ’99), Prague, Czech Republic.
- “Online Data Structures in External Memory,” 6th Biannual Workshop on Algorithms and Data Structures, (WADS ’99), Vancouver, Canada, August 1999.
- 2000 “Efficient Search through Massive Data,” IBM T. J. Watson Research Center, Hawthorne, NY.
- “Multiresolution Approximation Techniques for Database Systems,” Dagstuhl-Seminar on Data Structures, Schloss Dagstuhl, Wadern, Germany.
- “Online Data Structures in External Memory,” Georgia Institute of Technology, Atlanta, GA.
- 2001 “External Memory Algorithms: Dealing with MASSIVE Data,” Distinguished Lecture Series, Louisiana State University, Baton Rouge, LA.
- “Information Technology Partnerships in Education,” TechVision 2020 Mayor’s Information Technology Summit, New Orleans, LA.
- “Computing on MASSIVE Data,” 16th Clemson Mini-Conference on Discrete Mathematics, Clemson University, Clemson, SC.
- “Computing on MASSIVE Data,” University of Notre Dame, Notre Dame, IN.
- 2002 “The Power of Duality and Randomness in Scheduling and Sorting with Multiple Disks,” KNUTHfest 2002 Symposium, Stanford University, Stanford, CA.
- “Entropy-Compressed Indexes for Sequences,” Dagstuhl-Seminar on Data Structures, Schloss Dagstuhl, Wadern, Germany.
- “Compressed Indexes for Fast Search in Sequences,” keynote address at the 6th Joint Conference on Information Sciences (JCIS ’02), Durham, NC.
- “Prediction via Data Compression,” Workshop on Compression Issues in Next-Generation Network Applications, DIMACS, Rutgers University, New Brunswick, NJ.
- “The Data Explosion,” the University of Kansas, Lawrence, KS.
- “The Data Explosion,” Purdue University, West Lafayette, IN.

- Minicourse on “External Memory Algorithms and Parallel Disk Access,” EEF Summer School on Massive Data Sets, University of Aarhus, Århus, Denmark.
- 2003 “How to Store and Search Massive Data Archives,” XVI Louisiana Distinguished Lecture Series, University of Louisiana–Lafayette, Lafayette, LA.
- “Managing and Strengthening Interdisciplinary Programs,” panel member, 2003 Annual Meeting of the American Association of Universities Arts & Sciences Deans, College Park, MD.
- “Data, Data Everywhere!” Scientech Club, Indianapolis, IN.
- “Entropy-Based Space Models for Algorithms and Data Structures,” NSF Workshop on Information Theory and Computer Science Interface Chicago, IL.
- 2004 “Data, Data Everywhere! Compressed Indexing and Indexed Compression,” Distinguished Lecturer Series, University of Rochester, Rochester, NY.
- “CS Education Après le Crash,” plenary panel member, Computing Research Association Conference 2004 (Snowbird ’04), Snowbird, UT.
- “Best of Both Worlds: Data Compression with Fast Indexing,” IBM Data Management Workshop, IBM Toronto Laboratory, Toronto, Canada.
- 2005 “Undergraduate Programs — Curricular Objectives and Access Issues,” panel organizer, 2005 Annual Meeting of the American Association of Universities Arts & Sciences Deans, Irvine, CA.
- 2006 “Compressed Data Structures: Dictionaries and the Data-Aware Measures,” Information Theory and Applications Workshop, California Institute for Telecommunications and Information Technology, University of California–San Diego, La Jolla, CA.
- “Issues and Challenges with Commercialization of Research,” panel member, 2006 Annual Meeting of the American Association of Universities Arts & Sciences Deans, Minneapolis, MN.
- “A Nearly Tight Analysis of the Burrows-Wheeler Transform,” Workshop on Space-Conscious Algorithms, University of Bologna Residential Center, Bertinoro, Italy.
- “Bringing Engagement to Life in Arts and Sciences Departments and Colleges,” panel member, Outreach Scholarship 2006 Conference, Columbus, OH.
- 2007 “The Impact of Computer Science and Information Technology on our Lives,” Lafayette Rotary Club, West Lafayette, IN.
- “Lower Bounds on Encoding Length with Burrows-Wheeler Compression,” Information Theory and Applications Workshop, California Institute for Telecommunications and Information Technology, University of California–San Diego, La Jolla, CA.
- “The Impact of Information Science and Technology on our Lives,” Town and Gown Club, West Lafayette, IN.
- “From Molecules to Ecosystems: Research Foci in Purdue Life Sciences,” Indiana Health Industry Forum, Barnes & Thornburg LLP and Indiana Economic Development Corporation, Indianapolis, IN.
- “The Entrepreneurial Spirit of Purdue,” Purdue Entrepreneurship Roundtable, Palo Alto, CA.
- “Hestia and Climate Change,” moderator, Roundtable Forum on Climate Change, Barnes & Thornburg LLP, Indianapolis, IN.
- “I/O-Efficient Algorithms and Data Structures,” keynote address for inauguration of Center for Massive Data Algorithmics, Danish National Research Foundation, University of Aarhus, Århus, Denmark.
- 2008 “On Searching Compressed String Collections Cache Obliviously,” Information Theory and Applications Workshop, California Institute for Telecommunications and Information Technology, University of California–San Diego, La Jolla, CA.
- Interview as part of ACM SIGMOD series on Distinguished Profiles in Databases, interviewed by Marianne Winslett, West Lafayette, IN.
- “Multidisciplinary Research in Universities,” U.S.–China Computer Science Leadership Summit, Arlington, VA.
- Interview as part of Purdue University Oral History Program Collection, interviewed by Katherine Markee, West Lafayette, IN.
- “I/O-Efficient Algorithms and Data Structures,” Computing Research Association, Washington, DC.

- 2009 “The Value of Connections,” keynote address at Phi Beta Kappa induction ceremony, Alpha of Texas Chapter, College Station, TX.
- “Searching String Collections for the Most Relevant Documents,” University of Pisa, Pisa, Italy.
- “Applying Wavelets in Database Systems,” ACM SIGMOD Test of Time Award lecture for the most impactful paper from SIGMOD ’99, 2009 ACM SIGMOD International Conference on Management of Data (SIGMOD ’09).
- “Searching String Collections for the Most Relevant Documents,” Texas A&M University, College Station, TX.
- “Searching String Collections for the Most Relevant Documents,” Los Alamos Computer Science Symposium 2009 (LACSS), Santa Fe, NM.
- “Compressed Data Structures and Top- $k$  Document Retrieval Problems,” University of Notre Dame, Notre Dame, IN.
- “Compressed Data Structures and Searching Document Collections for the Most Relevant Documents,” inaugural talk for the University of Texas Computer Science Distinguished Lecture Series, University of Texas, Austin, TX.
- “Compressed Data Structures,” Leaders and Innovators Lecture Series, Texas A&M University, College Station, TX.
- 2010 “Compressed Data Structures and Searching Document Collections for the Most Relevant Documents,” Duke University, Durham, NC.
- “GBWT: Compressed Text Indexing in External Memory,” 5th Stringology Research Workshop, Bar-Ilan University, Ramat Gan, Israel.
- “Compression, Indexing, and Retrieval for Massive String Data,” keynote address at 19th Annual Conference on Combinatorial Pattern Matching (CPM ’10), New York, NY.
- “CRA Guidelines for Enhancing Faculty Recruitment,” workshop chair and panel member, Computing Research Association Conference 2010 (Snowbird ’10), Snowbird, UT.
- “Understanding and Using Graduate Program Rankings in Computer Science,” workshop panel member, Computing Research Association Conference 2010 (Snowbird ’10), Snowbird, UT.
- “Managing Up: Partnering with Your Dean,” workshop panel member, Computing Research Association Conference 2010 (Snowbird ’10), Snowbird, UT.
- “Building Synergies,” 2010 Retreat of the Merrill Advanced Studies Center, the University of Kansas, Lawrence, KS.
- “Ensuring that Public Research Universities Maintain Their Strength,” workshop panel member, Association of Public and Land-grant Universities Council on Academic Affairs Summer Meeting, Portland, OR.
- 2011 “Graduate Program Disclosure Metrics,” workshop panel member, Association of Public and Land-grant Universities Council on Academic Affairs Summer Meeting, Asheville, NC.
- “Behavioral and Social Sciences as Key Components in National Research Initiatives,” panel moderator, 2011 Retreat of the Merrill Advanced Studies Center, Nebraska City, NE.
- “Open Access to Scholarly Communication at KU,” panel member, Berlin 9 Open Access Preconference, Washington, DC.
- “KU Bold Aspirations,” Lafayette Rotary Club, Lawrence, KS.
- 2012 “Compressed Data Structures,” keynote address at the 2012 ACM Conference on Information and Knowledge Management (CIKM ’12), Manoa, Hawaii, October–November 2012.

## 15. Research Funding

Dr. Vitter is the sole principal investigator of the grants below, except for those in which multiple co-principal investigators are listed.

- 1981–1984 National Science Foundation, “The Analysis and Implementation of Algorithms for Information Retrieval,” Brown University, \$33,834.

- 1981–1982 IBM, “The Design and Analysis of Algorithms for Sorting and Searching,” Brown University, \$30,000.
- 1982–1987 National Science Foundation, “An Integrated Experimental Environment for Research in Computer Science” (Co-principal investigators A. van Dam, J. E. Savage, P. Wegner, R. Sedgewick, E. Charniak, T. W. Doemann, S. P. Reiss, J. S. Vitter, and P. C. Kanellakis), CER Program, Brown University, \$2,736,377.
- 1983–1986 DARPA/ONR, “Ideographics” and “A Graphical Approach to Software Development” (Co-principal investigators J. E. Savage, A. van Dam, G. M. Baudet, E. Charniak, B. M. Chazelle, T. W. Doemann, P. C. Kanellakis, S. P. Reiss, R. Sedgewick, J. S. Vitter, and P. Wegner), Brown University, \$1,700,000.
- 1983–1986 IBM, “The Design and Analysis of Algorithms for Sorting and Searching and VLSI,” Brown University, \$100,000.
- 1984–1989 National Science Foundation, “The Design and Analysis of Combinatorial Algorithms,” Brown University, \$179,775.
- 1984–1988 IBM Faculty Development Award, Brown University, \$60,000.
- 1985–1992 National Science Foundation Presidential Young Investigator Award, “The Mathematical Analysis of Algorithms,” Brown University, \$312,000.
- 1986–1987 AT&T, “The Mathematical Analysis of Algorithms,” Brown University, \$17,500.
- 1988–1989 IBM, “Algorithms for Manipulating Data using Systems with Parallel Architecture,” Brown University, \$37,500.
- 1988–1993 National Science Foundation, “Multiparadigm Design Environments” (Co-principal investigators J. E. Savage, E. Charniak, T. L. Dean, T. W. Doemann, P. C. Kanellakis, D. P. Lopresti, L. Morgenstern, S. P. Reiss, A. van Dam, J. S. Vitter, P. Wegner, F. K. Zadeck, and S. B. Zdonik), CISE Institutional Infrastructure Program, Brown University, \$3,481,000.
- 1989–1991 IBM, “Algorithms for Manipulating Data using Systems with Parallel Architecture,” Brown University, \$75,000.
- 1989–1992 NASA Graduate Student Research Program, “Analysis of Coding for Data Compression,” Brown University, \$62,000.
- 1991–1994 National Science Foundation, “Algorithmic Issues in High-Performance Computing” (Co-principal investigators J. S. Vitter and R. Tamassia), Brown University (partly subcontracted to Duke University), \$346,802.
- 1991–1994 Army Research Office, “Algorithmic Issues in High-Performance Computing” (Co-principal investigators J. S. Vitter and R. Tamassia), Brown University (partly subcontracted to Duke University), \$150,000.
- 1991–1994 ARPA/ISTO, “High-Performance Design Environments” (Co-principal investigators E. Charniak, T. W. Doemann, J. F. Hughes, P. C. Kanellakis, P. N. Klein, D. P. Lopresti, S. P. Reiss, J. E. Savage, R. Tamassia, A. van Dam, P. van Hentenryck, J. S. Vitter, P. Wegner, F. K. Zadeck, and S. B. Zdonik), Brown University, \$2,654,835.
- 1992–1994 National Science Foundation, “Management, Analysis and Representation of Large Scientific Databases” (Co-principal investigator with L. Sirovich, F. Bisshopp, R. Everson, and J. S. Vitter), Brown University, \$312,514.
- 1992–1994 Air Force Office of Scientific Research, “Design and Analysis of Lossless and Lossy Data Compression Methods with Applications to Communication and Caching,” Brown University (subcontracted to Duke University), \$112,774.
- 1993–1996 National Science Foundation, “Non-Equilibrium Phenomena in Nematic Liquid Crystals” (Co-principal investigators R. Pelcovits, G. Lorient, and J. S. Vitter), Brown University, \$195,785.
- 1993–1996 National Science Foundation, “SIMD/MIMD Parallel Computing: Computational Theory, Scientific Applications, and Systems Research” (Co-principal investigators J. S. Vitter, C. S. Ellis, C. Gardner, D. W. Loveland, J. H. Reif, and D. J. Rose, H. S. Greenside, M.-Y. Kao, G. Kedem, P. Lanzkron, and R. Wagner), CISE Institutional Infrastructure Program, Brown University, \$1,420,001 (plus institutional support of \$421,209).

- 1993–1996 Army Research Office, “Algorithms for Processing Large-Scale Data,” Duke Univ., \$211,945.
- 1994–1996 Air Force Office of Scientific Research, “Design and Analysis of Lossless and Lossy Data Compression Methods with Applications to Communication and Caching,” Duke University, \$305,311.
- 1995–1998 Army Research Office, AASERT Award, “Breaking the I/O Bottleneck in Large-Scale Computation,” Duke University, \$100,000.
- 1995–1998 National Science Foundation, “Acquisition of a Workstation Cluster Testbed for Next-Generation Collaborative Computing” (Co-principal investigators J. S. Vitter, J. S. Chase, C. S. Ellis, G. Kedem, and J. H. Reif), Academic Research Instrumentation Program, Duke University, \$489,600 (plus institutional support of \$489,599).
- 1995–2001 Army Research Office, Multidisciplinary Research Program of the University Research Initiative, “Applicable and Robust Geometric Computing” (Co-principal investigators J. S. Vitter and P. K. Agarwal), Duke University (subcontracted from the Center for Geometric Computation at Brown University, as part of a \$4,500,000 grant with participating institutions Brown, Duke, and Johns Hopkins Universities), \$1,489,042.
- 1996–1999 National Science Foundation, “Efficient I/O Communication for High-Performance Computing,” Duke University, \$274,999 (includes a \$5,000 Research Experience for Undergraduates supplement in 1997).
- 1996–1999 National Science Foundation, “CURIOUS: Center for Undergraduate education and Research: Integration through performance and visualization,” Duke University, Co-Principal Investigators O. L. Astrachan, S. Rodger, P. K. Agarwal, A. W. Biermann, G. Kedem, A. Lebeck, J. H. Reif, X. Sun, J. S. Vitter, and D. J. Rose, \$405,200.
- 1998 Microsoft Education Development Grant, Duke University, Co-Principal Investigators O. L. Astrachan, S. Rodger, and J. S. Vitter, \$50,000.
- 1998–2001 U.S. Department of Education, “Fellowships in Experimental Computer Science,” Duke University, Co-Principal Investigators C. S. Ellis, L. Arge, O. L. Astrachan, J. S. Chase, G. Kedem, A. Lebeck, M. Littman, and J. S. Vitter, \$500,000.
- 1998–2002 National Science Foundation, “Geographic Information Systems on High-Speed Clusters: A Vertically Integrated Approach,” Duke University, Co-Principal Investigators J. S. Chase, P. K. Agarwal, L. Arge, P. N. Halpin, D. L. Urban, and J. S. Vitter, \$1,185,900.
- 1999 Los Alamos Research Laboratory, “Models and Algorithms for Several Levels of Memory,” Duke University, Co-Principal Investigators P. K. Agarwal, L. Arge, and J. S. Vitter, \$50,000.
- 1999–2003 National Science Foundation, “External Memory Algorithms: Dealing with MASSIVE Data,” Duke University and Purdue University, \$290,000.
- 2000–2003 National Science Foundation, “Algorithms for Active Storage,” Co-Principal Investigators J. S. Vitter and J. S. Chase, Duke University, \$252,000.
- 2000–2003 Microsoft Research, “Interactive Research/Teaching Classroom,” Co-Principal Investigators J. S. Vitter, R. Lucic, L. Arge, O. L. Astrachan, J. S. Chase, C. S. Ellis, D. Ramm, S. Rodger, and A. Vahdat, Duke University, \$1,191,470.
- 2001 IBM, “The IBM/Duke Mass Storage Project, the IBM/Duke e-Education Partnership, and the IBM/Duke Pervasive Computing Initiative,” Co-Principal Investigators J. Harer, K. Trivedi, J. Board, J. S. Vitter, R. Lucic, J. S. Chase, A. Vahdat, S. Rodger, A. Lebeck, and O. L. Astrachan, Duke University, \$384,000.
- 2001–2003 Army Research Office, “External Memory Algorithms: Dealing with MASSIVE Data,” Duke University and Purdue University, \$90,985.
- 2002–2005 IBM, “Dynamic Optimization in Databases and Information Systems,” Purdue University, \$40,000.
- 2003–2005 Army Research Office, “External Memory Algorithms: Dealing with MASSIVE Data,” Purdue University, \$107,529.
- 2004–2007 National Science Foundation, “Entropy-Compressed Data Structures,” Purdue University, \$255,000.

- 2006–2010 National Science Foundation, “Performance Models and Systems Optimization for Disk-Bound Applications,” Co-Principal Investigators M. S. Thottethodi, J. S. Vitter, R. T. Shah, T. N. Vijaykumar, and V. S. Pai, Purdue University, \$889,788.
- 2010–2013 National Science Foundation, “Pattern Matching for Massive Data Sets,” the University of Kansas (subcontracted from Louisiana State University, as part of a \$500,000 collaborative research grant), \$235,773.

## 16. Research Projects

In my research I seek to exploit the rich interdependence between theory and practice. My work in algorithm design and analysis spans several application areas. The challenge in each case is to design algorithms that are both provably efficient and practical to implement. I am also interested in the complementary field of computational complexity, which considers the inherent difficulty of problems and thus allows us to determine when algorithms are optimal.

*Efficient algorithms for external memory.* My primary interest is the design and analysis of algorithms that are optimal in terms of the number of input/output operations (I/Os) between internal (primary) memory and external (secondary) storage. My recent book serves as a reference for much of the work in the field. We have developed models of parallel block transfer that apply to the new high-performance disk and hierarchical memory systems being developed. Our approach using duality provides the state-of-the-art approach to sorting using parallel disks. Practical algorithms have been developed for a variety of information processing applications, such as sorting, FFT, matrix operations, range searching, and a variety of computational geometry and combinatorial problems. Work in geographic information systems is described below. We can often prove that the algorithms are optimal, in terms of both I/O and internal processing. In terms of algorithm engineering, we have worked on programming environments such as TPIE (Transparent Parallel I/O programming Environment), developed by a former student, that allow users to implement efficient external memory algorithms in a high-level manner.

*Compressed text indexes and data structures.* Exhaustive search through large collections of data is prohibitively slow. Efficient indexing of massive documents is critical. Good indexes for textual and sequence data should be efficient to construct, allow fast access and queries on the text, and be compact and space-efficient. The wavelet tree data structure I introduced with colleagues (not to be confused with wavelets discussed below) is a practical and elegant structure for coding sequences of characters from a multicharacter alphabet, and it is a key tool in modern text indexing. Until about 12 years ago, the best known data structures that supported fast pattern matching required several times more space than the data being indexed! Our work was the first to achieve fast indexes that were provably linear in size to the text. We have since developed the first sublinear-space text indexes that provably require only as much space as entropy-compressed text, while still allowing fast and powerful searches. Moreover, the original text is not needed and can be reconstructed from the index. Empirical studies of our data structures show superior performance in practice compared with other methods. We are currently looking at compressed data structures for a variety of other problems.

*Data compression,* which deals with developing efficient methods for compressing source data so as to reduce space consumption. In the online setting, compression has important uses in network communication. Efficient variants of Huffman and arithmetic coding algorithms have been developed and analyzed. In text compression we have developed algorithms that improve the speed of statistical methods such as PPMC as well as algorithms that are comparable with Ziv-Lempel methods in speed but better in compression. We have developed the first provably good compressed search indexes for text and sequence data, and we are currently examining further improvements. Image compression work includes both lossless image compression and lossy vector quantization; often the algorithms we develop are progressive and thus facilitate browsing with no added overhead. Our FELICS method has been implemented in hardware for the Mars Reconnaissance Orbiter. We have also worked on video compression methods, specifically motion compensation algorithms for low bit rate applications like teleconferencing and rate control algorithms for high bit rate scenarios like MPEG. We developed the paradigm of minimizing the combined measure of rate plus distortion; this rate-distortion optimization has since been incorporated into the H.264/MPEG-4 AVC standard’s reference encoder, used widely in the computing and communications industry (e.g., Tandberg Television SD MPEG-2 EN8100, HD MPEG-4 EN8190, and SD & HD MPEG-4 iPlex, Atome H.264 encoder, Grass Valley ViBE SD & HD encoders, libavcodec, Mainconcept H.264 encoder, Microsoft VC-1 encoder, Theora 1.1-alpha1, x264

H.264 encoder, and Xvid MPEG-4 ASP encoder).

*Prediction for caching and prefetching*, which deals with developing prediction methods for optimizing system performance. Caching (or paging) usually refers to the process in which the system decides which page to remove from cache when a new page that the user requests is fetched into cache. Prefetching is the active process in which the system uses the user's idle time to predict which page the user will access next and then moves the likely pages into cache in anticipation of the user. We have applied the predictive aspects of optimal data compression methods to develop provably optimal prefetching methods. The methods predict as well in the limit as special-purpose methods tuned to the characteristics of the sequence of page requests. In one case, for sequences generated by Markov sources, the prediction accuracy converges to the intrinsic fault rate. In another, accuracy is optimal in the worst case, compared with all finite-state predictors.

*Database access and data mining*, which deals with the search for and extraction of useful patterns among data in large depositories, such as in commercial databases, image databases, data warehouses, the World Wide Web, and genomics applications. A theme in our work is the development of methods that are simultaneously efficient in terms of their processing time (often dominated by I/O accesses) and the quality of the results they produce. Much of our I/O work on spatial databases and GIS is described above in the computational geometry section. One key data mining task is learning how to construct rules to classify data items automatically. Prediction, applied above for caching and prefetching, is useful for determining likely future events based upon the past data stored in a database. We have worked on how to cluster data based upon measures of similarity and how to do similarity search, as in image databases. We have introduced wavelets to the database community as a key technique for summarizing, approximating, and predicting data. Wavelets have since become heavily used in database optimization, indexing, data warehousing, image processing, and data mining. We have successfully applied wavelets constructed in an I/O-efficient manner to the problems of selectivity estimation in databases and finding approximate solutions to OLAP (on-line analytic processing) queries. I received the 2009 ACM SIGMOD Test of Time award, for co-authoring the SIGMOD paper from 10 years earlier that has had the most impact in the following decade in terms of research, products, and methodology.

*Machine learning and neural networks*, which deals with the design and quantitative evaluation of algorithms for learning concepts from examples — an important component of artificial intelligence. The goal is to construct algorithms or neural networks that can “learn” a target concept after seeing a reasonable number of typical examples. Emphasis is put on the efficiency of such algorithms, in terms of the degree of correctness of the hypothesis produced and of the number of examples and storage space required. Theoretical studies have led to new algorithms, quantitative models, and bounds on performance, especially in the realm of parallel learning algorithms, neural networks, map learning for robotics, memory-based systems, and program testing methods.

*Computational geometry*, which deals with the design and analysis of algorithms for manipulating geometric quantities, such as lines, planes, rectangles, and convex objects. One problem considered deals with efficient methods for eliminating “hidden” lines and surfaces in two-dimensional graphical displays. We have investigated efficient algorithms in terms of a new object model of complexity that can better model graphics hardware. We have worked on efficient new algorithms for binary space partitions and geometric clustering, with applications to data compression, machine learning, and graphics. We have considered point location and important subcases of graph problems in planar and euclidean graphs. An important area of investigation is to develop I/O-efficient indexing methods for the massive amounts of geometric data that can be stored in geographic information systems (GIS) and other spatial databases. We are developing efficient algorithms for search, clustering, accessing data in triangulated irregular networks, contour-line extraction, terrain processing, and map overlay.

*Parallel processing*. My work includes the design of fast parallel combinatorial algorithms and development of new complexity models for parallel computation. We have developed algorithms for transitive closure, shortest paths, and graph drawing in planar structures, cooperative search, and unification. Results have been obtained that demonstrate interesting tradeoffs between running time and amount of parallelism, with applications to VLSI and the utilization of supercomputers. We have introduced complexity classes for general parallel computation that capture a practical notion of “parallelizability.” Parallel algorithms for unification and other complete problems in the class  $\mathcal{P}$  have been developed.

*Incremental (or dynamic) computation*, which deals with how to update the solution to a problem when the input is changed incrementally. Such techniques are especially important when the problem is of large scale. Complexity models have been developed, and several problems have been classified according to degree of incremental complexity. We have discovered interesting relationships between a problem's potential for

parallelization and its potential for dynamization. Efficient incremental algorithms have been developed for a variety of graph problems in restricted domains, including reachability, connectivity, shortest path, maximum flow, and various problems whose general formulations are  $\mathcal{NP}$ -complete.

*Random sampling and order statistics*, which deals with finding novel and optimum methods for generating various random quantities, such as a random sample of  $n$  records from a file containing  $N$  records, or a random variate whose distribution depends upon a set of  $N$  dynamically changing weights. These methods have important applications in databases, data mining, simulations, statistical surveys, and probabilistic algorithms.

*Analysis of hashing methods*, concentrating on the study of coalesced hashing and an algorithm for processing sweep line information called hashing with lazy deletion. Related questions in queueing theory deal with the distribution of the maximum queue size in stochastic processes and dynamic data structures.

*Design and analysis of data structures and graph algorithms*. This includes the mathematical analysis of fundamental combinatorial structures such as families of trees, the design and analysis of priority queue algorithms that are both theoretically optimal and practical to use, and the development of algorithms for unification and its variants. Work on hashing and Euclidean graph algorithms is described earlier.

## 17. Theses

- 1983 Supervision of Master's thesis of R. A. Simons, *On Bounded Depth Parity Circuits*, Brown University.
- 1985 Supervision of Ph.D. thesis of W.-C. Chen, *The Design and Analysis of Coalesced Hashing*, Brown University.
- 1985 Supervision of Master's thesis of J. T. Stasko, *Pairing Heaps: Experiments and Analysis*, Brown University.
- 1986 Committee member for Ph.D. thesis of P. Celis *Robin Hood Hashing*, University of Waterloo,
- 1988 Committee member and reporter for doctoral thesis of C. M. Kenyon (née Mathieu), *Comparaison de modèles combinatoires et probabilistes : deux exemples en Analyse d'Algorithmes*, Ecole Normale Supérieure and University of Paris.
- 1989 Committee member for Ph.D. thesis of J. T. Stasko, *TANGO: A Framework and System for Algorithm Animation*, Brown University.
- 1989 Committee member and reporter for doctoral thesis of D. Gardy, *Bases de données, allocations aléatoires : quelques analyses de performances*, University of Paris.
- 1990 Supervision of the Master's thesis of E. A. Shriver, *Optimal Disk I/O and Memory Access with Parallel Block Transfer*, Brown University.
- 1990 Committee member for Ph.D. thesis of R. J. Calistri, *Classifying and Detecting Plan-Based Misconceptions for Robot Plan Recognition*, Brown University.
- 1991 Committee member for Ph.D. thesis of R. Ravenscroft, *Generating Function Algorithms for Symbolic Computation*, Brown University.
- 1991 Committee member for Ph.D. thesis of A. K. Agrawal, *Network Design and Network Cut Dualities: Approximation Algorithms and Applications*, Brown University.
- 1992 Supervision of Ph.D. thesis of J.-H. Lin, *Approximation Algorithms and Complexity Results for Machine Learning*, Brown University.
- 1992 Committee member for Ph.D. thesis of A. A. Shvartsman, *Fault-Tolerant and Efficient Parallel Computation*, Brown University.
- 1992 Committee member for Ph.D. thesis of R. F. Cohen, *Combine and Conquer*, Brown University.
- 1993 Supervision of Ph.D. thesis of M. H. Nodine, *Minimizing the Input/Output Bottleneck*, Brown University.
- 1993 Committee member for Ph.D. thesis of K. J. Basye, *A Framework for Map Construction*, Brown University.
- 1993 Supervision of Master's thesis of K. M. Curewitz, *Practical Prefetching via Data Compression*, Brown University.
- 1993 Supervision of Ph.D. thesis of P. G. Howard, *The Design and Analysis of Efficient Lossless Data Compression Systems*, Brown University.

- 1995 Supervision of Ph.D. thesis of S. Subramanian, *Parallel and Dynamic Shortest-Path Algorithms for Sparse Graphs*, Brown University.
- 1995 Supervision of Ph.D. thesis of P. Krishnan, *Online Prediction Algorithms for Databases and Operating Systems*, Brown University.
- 1995 Committee member for Ph.D. thesis of J. Subramanian, *Dynamic Resource Management for a Network under Stochastic Demand*, University of North Carolina at Chapel Hill.
- 1996 Committee member for Ph.D. thesis of A. Purakayasstha, *Characterizing and Optimizing Parallel File Systems*, Duke University.
- 1997 Committee member for Ph.D. thesis of P. Ferragina, *Dynamic Data Structures for String Matching Problems*, University of Pisa,
- 1997 Committee member for Ph.D. thesis of S. Chen, *Algorithmic Applications of Data Compression Techniques*, Duke University.
- 1997 Supervision of Ph.D. thesis of D. E. Vengroff, *The Theory and Practice of I/O-Efficient Computation*, Brown University.
- 1997 Supervision of Ph.D. thesis of D. T. Hoang, *Efficient Algorithms for Text and Video Compression*, Brown University.
- 1997 Reporter for doctoral thesis of F. Bertault, *Génération et tracé de structures décomposables*, Henri Poincaré University.
- 1998 Committee member for Ph.D. thesis of K. R. Varadarajan, *Algorithms for Some Geometric Optimization Problems: Navigation, Simplification, and Matching*, Duke University.
- 1998 Supervision of Ph.D. thesis of R. D. Barve, *Algorithmic Techniques to Overcome the I/O Bottleneck*, Duke University.
- 1999 Supervision of Ph.D. thesis of T. M. Murali, *Efficient Hidden-Surface Removal in Theory and in Practice*, Duke University.
- 1999 Supervision of Ph.D. thesis of M. Wang, *Approximation and Learning Techniques in Database Systems*, Duke University.
- 1999 Supervision of Master's thesis of S. Luoma, *A Robust Algorithm for Constructing a BSP of Polygons in  $\mathbb{R}^3$* , Duke University.
- 2000 Committee Member for Ph.D. thesis of P. Desikan, *Issues in Interactive Rendering: Theory and Practice*, Duke University.
- 2001 Supervision of Ph.D. thesis of A. I. Natsev, *Multidimensional Retrieval by Regions, Concepts, and Constraints*, Duke University.
- 2001 Committee member for Ph.D. thesis of C. Procopiuc, *Efficient Algorithms for Geometric Clustering Problems*, Duke University.
- 2002 Supervision of Ph.D. thesis of O. Procopiuc, *Algorithms for Very Large Spatial Databases*, Duke University.
- 2004 Committee member for Ph.D. thesis of I. Ilyas, *Rank-aware Query Optimization*, Purdue University.
- 2004 Supervision of Ph.D. thesis of L. Lim, *On-line Methods for Database Optimization*, Duke University.
- 2007 Supervision of Ph.D. thesis of A. Gupta, *Succinct Data Structures*, Duke University (revised).
- 2010 External reviewer of Ph.D. thesis of Rossano Venturini, *On Searching and Extracting Strings from Compressed Textual Data*, University of Pisa (expected).

Dr. Vitter's former Ph.D. students are currently employed at Akamai, Avaya Labs, Butler University, Google, Hewlett-Packard, IBM Research, Intrinsicity, Microsoft, National Taiwan University, Navini, richrelevance, University of Hawaii, Veveo.tv, Virginia Tech, and Xilient.

## 18. Postdoctoral Assistants

Paul G. Howard, Edward F. Grove, David Hutchinson, Philip M. Long, Lars Arge, Julien Basch, Stergios Anastasiadis, Rahul Shah, Wing-Kai Hon, Oğuzhan Külekci, Bojian Xu, and Saswata Shannigrahi.