

Pearlmutter & Siskind joint work on Automatic Differentiation (AD)

[\[html\]](#) [\[bibliography\]](#)

Code

J.M. Siskind, ‘R6RS-AD,’ Purdue University, 20 May 2010.

[\[github\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Stalingrad,’ Purdue University, 1 April 2016.

[\[github\]](#) [\[BibTeX\]](#)

Grants

J.M. Siskind, PI, ‘Algorithmic Differentiation of Functional Programs,’ NSF award CCF-0438806, May 2005–April 2010, \$420,000.

[\[announcement\]](#)

Journal Papers

B.A. Pearlmutter and J.M. Siskind, ‘Reverse-Mode AD in a Functional Framework: Lambda the Ultimate Backpropagator,’ *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 30(2):1–36, March 2008.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Nesting Forward-Mode AD in a Functional Framework,’ *Higher-Order and Symbolic Computation (HOSC)*, 21(4):361–376 December 2008.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#)

A.G. Baydin, B.A. Pearlmutter, A.A. Radul, and J.M. Siskind, ‘Automatic Differentiation in Machine Learning: a Survey,’ *Journal of Machine Learning Research (JMLR)*, 18(153):1–43, April 2018.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Divide-and-Conquer Checkpointing for Arbitrary Programs with No User Annotation,’ *Optimization Methods and Software (OMS)*, 33(4–6):1288–1330, September 2018.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#)

O. Manzyuk, B.A. Pearlmutter, A.A. Radul, D.R. Rush, and J.M. Siskind, ‘Perturbation Confusion in Forward Automatic Differentiation of Higher-Order Functions,’ *Journal of Functional Programming (JFP)*, 29(e12):1–24, September 2019.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[announcement\]](#) [\[slides\]](#) [\[video\]](#)

Conference Papers

J.M. Siskind and B.A. Pearlmutter, ‘Perturbation Confusion and Referential Transparency: Correct Functional Implementation of Forward-Mode AD,’ *Draft Proceedings of the 17th International Workshop on Implementation and Application of Functional Languages (IFL2005)*, Dublin, Ireland, 19–21 September 2005.

[\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

B.A. Pearlmutter and J.M. Siskind, ‘Lazy Multivariate Higher-Order Forward-Mode AD,’ *Proceedings of the 34th Annual Symposium on Principles of Programming Languages (POPL)*, pp. 155–160, Nice, France, 17–19 January 2007.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘First-Class Nonstandard Interpretations by Opening Closures,’ *Proceedings of the 34th Annual Symposium on Principles of Programming Languages (POPL)*, pp. 71–76, Nice, France, 17–19 January 2007.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

B.A. Pearlmutter and J.M. Siskind, ‘Using Programming Language Theory to Make AD Sound and Efficient,’ *Proceedings of the 5th International Conference on Automatic Differentiation (AD2008)*, pp. 79–90, Bonn, Germany, 11–15 August 2008.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#)

D. Wingate, N. Goodman, A. Stuhlmüller, and J.M. Siskind, ‘Nonstandard Interpretations of Probabilistic Programs for Efficient Inference,’ *Proceedings of the Neural Information Processing Systems Conference (NIPS)*, Granada, Spain, 12–15 December 2011.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#)

A. Radul, B.A. Pearlmutter, and J.M. Siskind, ‘AD in Fortran: Implementation via Preprocessor,’ *Proceedings of the 6th International Conference on Automatic Differentiation (AD2012)*, pp. 273–284, Fort Collins, CO, 23–27 July 2012.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Binomial Checkpointing for Arbitrary Programs with No User Annotation,’ *Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD)*, Oxford, UK, 12–15 September 2016.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Efficient Implementation of a Higher-Order Language with Built-In AD,’ *Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD)*, Oxford, UK, 12–15 September 2016.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

A.G. Baydin, B.A. Pearlmutter, and J.M. Siskind, ‘DiffSharp: An AD Library for .NET Languages,’ *Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD)*, Oxford, UK, 12–15 September 2016.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

A.G. Baydin, B.A. Pearlmutter, and J.M. Siskind, ‘Tricks from Deep Learning,’ *Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD)*, Oxford, UK, 12–15 September 2016.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[slides\]](#)

R. Kelly, B.A. Pearlmutter, and J.M. Siskind, ‘Evolving the Incremental λ Calculus into a Model of Forward AD,’ *Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD)*, Oxford, UK, 12–15 September 2016.

[\[copy of record\]](#) [\[pdf\]](#) [\[BibTeX\]](#) [\[poster\]](#)

Workshop Presentations

J.M. Siskind, ‘AD Nesting, Variable Capture, Programming Language Theory, and AD,’ *2nd European Workshop on Automatic Differentiation*, 18 November 2005.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘AD for Probabilistic Programming,’ *NIPS Workshop on Probabilistic Programming: Universal Languages and Inference; systems; and applications*, 13 December 2008.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘What every machine-learning researcher should know about AD,’ *DALI Workshop on Probabilistic Programming*, Friday 10 April 2015.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#)

B.A. Pearlmutter, ‘Automatic Differentiation: History and Headroom,’ *NIPS Workshop on The Future of Gradient-Based Machine Learning Software*, 10 December 2016.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#) [\[video\]](#)

J.M. Siskind, ‘The tension between convenience and performance in automatic differentiation,’ *NIPS Workshop on The Future of Gradient-Based Machine Learning Software*, 10 December 2016.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#) [\[video\]](#)

J.M. Siskind, ‘Binomial Checkpointing for Arbitrary Programs with No User Annotation,’ *SIAM Conference on Computational Science and Engineering (CSE) minisymposium on AD*, 28 February 2017.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Divide-and-Conquer Checkpointing for Arbitrary Programs with No User Annotation,’ *NIPS Workshop on The Future of Gradient-Based Machine Learning Software & Techniques*, 9 December 2017.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Scheme as a framework for Deep Learning,’ *ICFP Scheme and Functional Programming Workshop*, 27 August 2021.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#) [\[video\]](#)

M. Maleki, B.A. Pearlmutter, and J.M. Siskind, ‘Adding AD to Scheme by differentiation the Interpreter,’ *ICFP Scheme and Functional Programming Workshop*, 27 August 2021.

[\[schedule\]](#) [\[BibTeX\]](#) [\[video\]](#)

Public Seminars

J.M. Siskind, ‘Backpropagation Through Functional Programs: How to do Reverse-Mode AD Correctly in a Functional Framework,’ *New Jersey Programming Languages and Systems (NJPLS)*, 16 September 2005.

[\[schedule\]](#) [\[abstract\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Taking Derivatives of Functional Programs: AD in a Functional Framework,’ *New England Programming Languages and Systems (NEPLS)*, 27 October 2005.

[\[schedule\]](#) [\[abstract\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Algorithmic Differentiation of Functional Programs,’ *NSF Science of Design (SoD) Program PI Meeting*, 28 February–2 March 2007.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus,’ *New England Programming Languages and Systems (NEPLS)*, 21 November 2008.

[\[schedule\]](#) [\[abstract\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs and its use for Probabilistic Programming,’ *New Jersey Programming Languages and Systems (NJPLS)*, 3 April 2009.

[\[schedule\]](#) [\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘AD for CV, AC, ML, AI, and CL,’ *Science Foundation Ireland (SFI) Site Visit*, 23 November 2012.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Probabilistic Programming: From Principled Foundations, Through Efficient Implementation, To Innovative Applications,’ *DARPA Probabilistic Programming for Advancing Machine Learning (PPAML) Program PI Meeting*, 10 July 2014.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘The Next 700 Probabilistic Programming Languages,’ *DARPA Probabilistic Programming for Advancing Machine Learning (PPAML) Program PI Meeting*, 22 January 2015.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Nested Queries in Nondeterministic and Probabilistic Programming Languages,’ *New England Programming Languages and Systems (NEPLS)*, 10 November 2015.

[\[schedule\]](#) [\[abstract\]](#) [\[slides\]](#) [\[BibTeX\]](#)

Seminars at Institutions

J.M. Siskind, ‘Algorithmic Differentiation of Functional Programs,’ Purdue University, 16 June 2004.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Algorithmic Differentiation of Functional Programs,’ University of Maryland, 17 June 2004.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Algorithmic Differentiation of Functional Programs,’ University of Toronto, 25 June 2004.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘Algorithmic Differentiation of Functional Programs,’ Northeastern University, 8 August 2004.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Algorithmic Differentiation of Functional Programs,’ University of Illinois at Urbana-Champaign, 1 September 2004.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Taking Derivatives of Functional Programs: AD in a Functional Framework,’ University of Hertfordshire, 22 November 2005.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation for Probabilistic Programming,’ Indiana University, 20 February 2007.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus,’ University of Chicago, 8 April 2008.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation for Probabilistic Programming,’ Indiana University, 29 January 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs and its use for Probabilistic Programming,’ Jane Street Capital, 12 March 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs and its use for Probabilistic Programming,’ DE Shaw, 13 March 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus,’ MIT, 17 March 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus,’ Yale University, 19 March 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus,’ University of Pennsylvania, 20 March 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs and its use for Probabilistic Programming,’ Purdue University, 30 March 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs and its use for Probabilistic Programming,’ Northeastern University, 1 April 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs and its use for Probabilistic Programming,’ TTI Chicago, 17 April 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘Automatic Differentiation of Functional Programs and its use for Probabilistic Programming,’ University of South Carolina, 11 September 2009.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘What does CPS have to do with deep learning?,’ Northeastern University, 11 September 2017.
[\[slides\]](#) [\[BibT_EX\]](#)

J.M. Siskind, ‘What does CPS have to do with deep learning?,’ Amazon, Boston, 6 June 2018.

[\[slides\]](#) [\[BibTeX\]](#)

J.M. Siskind, ‘A Functional-Programming Framework for Deep Learning,’ Meta, 16 December 2021.

[\[slides\]](#) [\[BibTeX\]](#)

Technical Reports

J.M. Siskind and B.A. Pearlmutter, ‘Using Polyvariant Union-Free Flow Analysis to Compile a Higher-Order Functional-Programming Language with a First-Class Derivative Operator to Efficient Fortran-like Code,’ Technical Report TR-ECE-08-01, School of Electrical and Computer Engineering, Purdue University, January 2008.

[\[pdf\]](#) [\[BibTeX\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Putting the Automatic Back into AD: Part I, What’s Wrong,’ Technical Report TR-ECE-08-02, School of Electrical and Computer Engineering, Purdue University, January 2008.

[\[pdf\]](#) [\[BibTeX\]](#)

B.A. Pearlmutter and J.M. Siskind, ‘Putting the Automatic Back into AD: Part II, Dynamic, Automatic, Nestable, and Fast,’ Technical Report TR-ECE-08-03, School of Electrical and Computer Engineering, Purdue University, January 2008.

[\[pdf\]](#) [\[BibTeX\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Nesting Forward-Mode AD in a Functional Framework,’ Technical Report TR-ECE-08-09, School of Electrical and Computer Engineering, Purdue University, October 2008.

[\[pdf\]](#) [\[BibTeX\]](#)

arXiv

A. Radul, B.A. Pearlmutter, and J.M. Siskind, ‘AD in Fortran, Part 1: Design,’ arXiv:1203.1448, 7 March 2012.

[\[pdf\]](#) [\[BibTeX\]](#)

A. Radul, B.A. Pearlmutter, and J.M. Siskind, ‘AD in Fortran, Part 2: Implementation via Preprocessor,’ arXiv:1203.1450, 7 March 2012.

[\[pdf\]](#) [\[BibTeX\]](#)

O. Manzyuk, B.A. Pearlmutter, A.A. Radul, D.R. Rush, and J.M. Siskind, ‘Confusion of Tagged Perturbations in Forward Automatic Differentiation of Higher-Order Functions,’ arXiv:1211.4892, 20 November 2012.

[\[pdf\]](#) [\[BibTeX\]](#)

A.G. Baydin, B.A. Pearlmutter, A.A. Radul, and J.M. Siskind, ‘Automatic differentiation in machine learning: a survey,’ arXiv:1502.05767, 19 April 2015.

[\[pdf\]](#) [\[BibTeX\]](#)

A.G. Baydin, B.A. Pearlmutter, and J.M. Siskind, ‘DiffSharp: Automatic Differentiation Library,’ arXiv:1511.07727, 26 November 2015.

[\[pdf\]](#) [\[BibTeX\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Binomial Checkpointing for Arbitrary Programs with No User Annotation,’ arXiv:1611.03410, 10 November 2016.

[\[pdf\]](#) [\[BibTeX\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Efficient Implementation of a Higher-Order Language with Built-In AD,’ arXiv:1611.03416, 10 November 2016.

[\[pdf\]](#) [\[BibTeX\]](#)

A.G. Baydin, B.A. Pearlmutter, and J.M. Siskind, ‘DiffSharp: An AD Library for .NET Languages,’ arXiv:1611.03423, 10 November 2016.

[\[pdf\]](#) [\[BibTeX\]](#)

R. Kelly, B.A. Pearlmutter, and J.M. Siskind, ‘Evolving the Incremental λ Calculus into a Model of Forward AD,’ arXiv:1611.03429, 10 November 2016.

[\[pdf\]](#) [\[BibTeX\]](#)

A.G. Baydin, B.A. Pearlmutter, and J.M. Siskind, ‘Tricks from Deep Learning,’ arXiv:1611.03777, 10 November 2016.

[\[pdf\]](#) [\[BibTeX\]](#)

J.M. Siskind and B.A. Pearlmutter, ‘Divide-and-Conquer Checkpointing for Arbitrary Programs with No User Annotation,’ arXiv:1708.06799, 22 August 2017.

[\[pdf\]](#) [\[BibTeX\]](#)