

Pearlmutter & Siskind joint work on Automatic Differentiation (AD)

BARAK AVRUM PEARLMUTTER, Maynooth University, Ireland
JEFFREY MARK SISKIND, Purdue University, USA

CCS Concepts: • **Software and its engineering** → **Compilers; Runtime environments**; • **Mathematics of computing** → **Differential calculus**; • **Computing methodologies** → **Neural networks**.

Additional Key Words and Phrases: automatic differentiation

ACM Reference Format:

Barak Avrum Pearlmutter and Jeffrey Mark Siskind. 0. Pearlmutter & Siskind joint work on Automatic Differentiation (AD). *ACM Trans. Program. Lang. Syst.* 0, 0, Article 0 (0), 4 pages. <https://doi.org/0>

REFERENCES

- A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind. 2015b. *Automatic differentiation in machine learning: a survey*. Technical Report 1502.05767. arXiv. <http://arxiv.org/abs/1502.05767>
- A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind. 2018. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research (JMLR)* 18, 153 (April 2018), 1–43. <http://engineering.purdue.edu/~qobi/papers/jmlr2018.pdf>
- A. G. Baydin, B. A. Pearlmutter, and J. M. Siskind. 2015a. *DiffSharp: Automatic Differentiation Library*. Technical Report 1511.07727. arXiv. <http://arxiv.org/abs/1511.07727>
- A. G. Baydin, B. A. Pearlmutter, and J. M. Siskind. 2016a. DiffSharp: An AD Library for .NET Languages. (Sept. 2016). <http://engineering.purdue.edu/~qobi/papers/ad2016c.pdf> Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD).
- A. G. Baydin, B. A. Pearlmutter, and J. M. Siskind. 2016b. *DiffSharp: An AD Library for .NET Languages*. Technical Report 1611.03423. arXiv. <http://arxiv.org/abs/1611.03423>
- A. G. Baydin, B. A. Pearlmutter, and J. M. Siskind. 2016c. Tricks from Deep Learning. (Sept. 2016). <http://engineering.purdue.edu/~qobi/papers/ad2016d.pdf> Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD).
- A. G. Baydin, B. A. Pearlmutter, and J. M. Siskind. 2016d. *Tricks from Deep Learning*. Technical Report 1611.03777. arXiv. <http://arxiv.org/abs/1611.03777>
- R. Kelly, B. A. Pearlmutter, and J. M. Siskind. 2016a. Evolving the Incremental λ Calculus into a Model of Forward AD. (Sept. 2016). <http://engineering.purdue.edu/~qobi/papers/ad2016e.pdf> Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD).
- R. Kelly, B. A. Pearlmutter, and J. M. Siskind. 2016b. *Evolving the Incremental λ Calculus into a Model of Forward AD*. Technical Report 1611.03429. arXiv. <http://arxiv.org/abs/1611.03429>
- M. Maleki, B. A. Pearlmutter, and J. M. Siskind. 2021. Adding AD to Scheme by differentiation the Interpreter. (Aug. 2021). <http://engineering.purdue.edu/~qobi/tweet-talks/icfp2021b-slides.pdf>
- O. Manzyuk, B. A. Pearlmutter, A. A. Radul, D. R. Rush, and J. M. Siskind. 2012. *Confusion of Tagged Perturbations in Forward Automatic Differentiation of Higher-Order Functions*. Technical Report 1211.4892. arXiv. <http://arxiv.org/abs/1211.4892>

Authors' addresses: Barak Avrum Pearlmutter, barak@pearlmutternet, Maynooth University, Computer Science, Maynooth, Ireland; Jeffrey Mark Siskind, qobi@purdue.edu, Purdue University, School of Electrical and Computer Engineering, 465 Northwestern Avenue, West Lafayette, Indiana, 47907-2035, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.

0164-0925/0/0-ART0 \$15.00

<https://doi.org/0>

- O. Manzyuk, B. A. Pearlmutter, A. A. Radul, D. R. Rush, and J. M. Siskind. 2019. Perturbation Confusion in Forward Automatic Differentiation of Higher-Order Functions. *Journal of Functional Programming (JFP)* 29, e12 (Sept. 2019), 1–24. <http://engineering.purdue.edu/~qobi/papers/jfp2019.pdf>
- B. A. Pearlmutter. 2016. Automatic Differentiation: History and Headroom. (Dec. 2016). <http://engineering.purdue.edu/~qobi/tweet-talks/nips2016a-slides.pdf>
- B. A. Pearlmutter and J. M. Siskind. 2007. Lazy Multivariate Higher-Order Forward-Mode AD. In *Proceedings of the Annual Symposium on Principles of Programming Languages (POPL)* (Nice, France). 155–160. <http://engineering.purdue.edu/~qobi/papers/popl2007a.pdf>
- B. A. Pearlmutter and J. M. Siskind. 2008a. *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. <http://docs.lib.purdue.edu/ecetr/369>
- B. A. Pearlmutter and J. M. Siskind. 2008b. Reverse-Mode AD in a Functional Framework: Lambda the Ultimate Backpropagator. *ACM Transactions on Programming Languages and Systems (TOPLAS)* 30, 2 (March 2008), 1–36. <http://engineering.purdue.edu/~qobi/papers/toplas2008.pdf>
- B. A. Pearlmutter and J. M. Siskind. 2008c. Using Programming Language Theory to Make AD Sound and Efficient. In *Proceedings of the International Conference on Automatic Differentiation (AD)* (Bonn, Germany). 79–90. <http://engineering.purdue.edu/~qobi/papers/ad2008.pdf>
- A. Radul, B. A. Pearlmutter, and J. M. Siskind. 2012a. AD in Fortran: Implementation via Preprocessor. In *Proceedings of the International Conference on Automatic Differentiation (AD)* (Fort Collins, CO). 273–284. <http://engineering.purdue.edu/~qobi/papers/ad2012.pdf>
- A. Radul, B. A. Pearlmutter, and J. M. Siskind. 2012b. *AD in Fortran, Part 1: Design*. Technical Report 1203.1448. arXiv. <http://arxiv.org/abs/1203.1448>
- A. Radul, B. A. Pearlmutter, and J. M. Siskind. 2012c. *AD in Fortran, Part 2: Implementation via Preprocessor*. Technical Report 1203.1450. arXiv. <http://arxiv.org/abs/1203.1450>
- J. M. Siskind. 2004a. Algorithmic Differentiation of Functional Programs. (June 2004). <http://engineering.purdue.edu/~qobi/tweet-talks/purdue2004-slides.pdf>
- J. M. Siskind. 2004b. Algorithmic Differentiation of Functional Programs. (June 2004). <http://engineering.purdue.edu/~qobi/tweet-talks/umd2004-slides.pdf>
- J. M. Siskind. 2004c. Algorithmic Differentiation of Functional Programs. (June 2004). <http://engineering.purdue.edu/~qobi/tweet-talks/toronto2004-slides.pdf>
- J. M. Siskind. 2004d. Algorithmic Differentiation of Functional Programs. (Aug. 2004). <http://engineering.purdue.edu/~qobi/tweet-talks/neu2004-slides.pdf>
- J. M. Siskind. 2004e. Algorithmic Differentiation of Functional Programs. (Sept. 2004). <http://engineering.purdue.edu/~qobi/tweet-talks/uiuc2004-slides.pdf>
- J. M. Siskind. 2005a. AD Nesting, Variable Capture, Programming Language Theory, and AD. (Nov. 2005). <http://engineering.purdue.edu/~qobi/tweet-talks/euroad2005-slides.pdf>
- J. M. Siskind. 2005b. Backpropagation Through Functional Programs: How to do Reverse-Mode AD Correctly in a Functional Framework. (Sept. 2005). <http://engineering.purdue.edu/~qobi/tweet-talks/njpls2005-slides.pdf>
- J. M. Siskind. 2005c. Taking Derivatives of Functional Programs: AD in a Functional Framework. (Oct. 2005). <http://engineering.purdue.edu/~qobi/tweet-talks/nepls2005-slides.pdf>
- J. M. Siskind. 2005d. Taking Derivatives of Functional Programs: AD in a Functional Framework. (Nov. 2005). <http://engineering.purdue.edu/~qobi/tweet-talks/hertfordshire2005-slides.pdf>
- J. M. Siskind. 2007a. Algorithmic Differentiation of Functional Programs. (Feb. 2007). <http://engineering.purdue.edu/~qobi/tweet-talks/sod2007-slides.pdf>
- J. M. Siskind. 2007b. Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus or How I Learned to Stop Worrying and Love Map-Closure. (Feb. 2007). <http://engineering.purdue.edu/~qobi/tweet-talks/iu2007-slides.pdf>
- J. M. Siskind. 2008a. AD for Probabilistic Programming. (Dec. 2008). <http://engineering.purdue.edu/~qobi/tweet-talks/nips2008-slides.pdf>
- J. M. Siskind. 2008b. Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus. (Nov. 2008). <http://engineering.purdue.edu/~qobi/tweet-talks/nepls2008-slides.pdf>
- J. M. Siskind. 2008c. Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus. (April 2008). <http://engineering.purdue.edu/~qobi/tweet-talks/uc2008-slides.pdf>
- J. M. Siskind. 2009a. Automatic Differentiation for Probabilistic Programming. (Jan. 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/iu2009-slides.pdf>
- J. M. Siskind. 2009b. Automatic Differentiation of Functional Programs and its use for Probabilistic Programming. (April 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/njpls2009-slides.pdf>

- J. M. Siskind. 2009c. Automatic Differentiation of Functional Programs and its use for Probabilistic Programming. (March 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/janest2009-slides.pdf>
- J. M. Siskind. 2009d. Automatic Differentiation of Functional Programs and its use for Probabilistic Programming. (March 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/deshaw2009-slides.pdf>
- J. M. Siskind. 2009e. Automatic Differentiation of Functional Programs and its use for Probabilistic Programming. (March 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/purdue2009-slides.pdf>
- J. M. Siskind. 2009f. Automatic Differentiation of Functional Programs and its use for Probabilistic Programming. (April 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/neu2009-slides.pdf>
- J. M. Siskind. 2009g. Automatic Differentiation of Functional Programs and its use for Probabilistic Programming. (April 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/tti2009-slides.pdf>
- J. M. Siskind. 2009h. Automatic Differentiation of Functional Programs and its use for Probabilistic Programming. (Sept. 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/sc2009-slides.pdf>
- J. M. Siskind. 2009i. Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus. (March 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/mit2009-slides.pdf>
- J. M. Siskind. 2009j. Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus. (March 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/yale2009-slides.pdf>
- J. M. Siskind. 2009k. Automatic Differentiation of Functional Programs or Lambda the Ultimate Calculus. (March 2009). <http://engineering.purdue.edu/~qobi/tweet-talks/upenn2009-slides.pdf>
- J. M. Siskind. 2010. *R6RS-AD*. Purdue University. <https://github.com/qobi/R6RS-AD>
- J. M. Siskind. 2012. AD for CV, AC, ML, AI, and CL. (Nov. 2012). <http://engineering.purdue.edu/~qobi/tweet-talks/sfi2012-slides.pdf>
- J. M. Siskind. 2014. Probabilistic Programming: From Principled Foundations, Through Efficient Implementation, To Innovative Applications. (July 2014). <http://engineering.purdue.edu/~qobi/tweet-talks/ppaml2014-slides.pdf>
- J. M. Siskind. 2015a. Nested Queries in Nondeterministic and Probabilistic Programming Languages. (Nov. 2015). <http://engineering.purdue.edu/~qobi/tweet-talks/nepls2015-slides.pdf>
- J. M. Siskind. 2015b. The Next 700 Probabilistic Programming Languages. (Jan. 2015). <http://engineering.purdue.edu/~qobi/tweet-talks/ppaml2015-slides.pdf>
- J. M. Siskind. 2015c. What every machine-learning researcher should know about AD. (April 2015). <http://engineering.purdue.edu/~qobi/tweet-talks/dali2015-slides.pdf>
- J. M. Siskind. 2016a. *Stalin V*. Purdue University. <https://github.com/Functional-AutoDiff/STALINGRAD>
- J. M. Siskind. 2016b. The tension between convenience and performance in automatic differentiation. (Dec. 2016).
- J. M. Siskind. 2017a. Binomial Checkpointing for Arbitrary Programs with No User Annotation. (Feb. 2017). <http://engineering.purdue.edu/~qobi/tweet-talks/siam2017-slides.pdf>
- J. M. Siskind. 2017b. Divide-and-Conquer Checkpointing for Arbitrary Programs with No User Annotation. (Dec. 2017). <http://engineering.purdue.edu/~qobi/tweet-talks/nips2017-slides.pdf>
- J. M. Siskind. 2017c. What does CPS have to do with deep learning? (Sept. 2017). <http://engineering.purdue.edu/~qobi/tweet-talks/neu2017-slides.pdf>
- J. M. Siskind. 2018. What does CPS have to do with deep learning? (June 2018). <http://engineering.purdue.edu/~qobi/tweet-talks/amazon2018-slides.pdf>
- J. M. Siskind. 2021a. A Functional-Programming Framework for Deep Learning. (Dec. 2021). <http://engineering.purdue.edu/~qobi/tweet-talks/meta2021-slides.pdf>
- J. M. Siskind. 2021b. Scheme as a framework for Deep Learning. (Aug. 2021). <http://engineering.purdue.edu/~qobi/tweet-talks/icfp2021a-slides.pdf>
- J. M. Siskind and B. A. Pearlmutter. 2005. Perturbation Confusion and Referential Transparency: Correct Functional Implementation of Forward-Mode AD. In *Draft Proceedings of the International Workshop on Implementation and Application of Functional Languages (IFL)* (Dublin, Ireland). <http://engineering.purdue.edu/~qobi/papers/ifl2005.pdf>
- J. M. Siskind and B. A. Pearlmutter. 2007. First-Class Nonstandard Interpretations by Opening Closures. In *Proceedings of the Annual Symposium on Principles of Programming Languages (POPL)* (Nice, France). 71–76. <http://engineering.purdue.edu/~qobi/papers/popl2007b.pdf>
- J. M. Siskind and B. A. Pearlmutter. 2008a. Nesting Forward-Mode AD in a Functional Framework. *Higher-Order and Symbolic Computation (HOSC)* 21, 4 (Dec. 2008), 361–376. <http://engineering.purdue.edu/~qobi/papers/hosc2008.pdf>
- J. M. Siskind and B. A. Pearlmutter. 2008b. *Nesting Forward-Mode AD in a Functional Framework*. Technical Report TR-ECE-08-09. School of Electrical and Computer Engineering, Purdue University. <http://docs.lib.purdue.edu/ecetr/377>
- J. M. Siskind and B. A. Pearlmutter. 2008c. *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. <http://docs.lib.purdue.edu/ecetr/368>
- J. M. Siskind and B. A. Pearlmutter. 2008d. *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. <http://docs.lib.purdue.edu/ecetr/367>
- J. M. Siskind and B. A. Pearlmutter. 2016a. Binomial Checkpointing for Arbitrary Programs with No User Annotation. (Sept. 2016). <http://engineering.purdue.edu/~qobi/papers/ad2016a.pdf> Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD).
- J. M. Siskind and B. A. Pearlmutter. 2016b. *Binomial Checkpointing for Arbitrary Programs with No User Annotation*. Technical Report 1611.03410. arXiv. <http://arxiv.org/abs/1611.03410>
- J. M. Siskind and B. A. Pearlmutter. 2016c. Efficient Implementation of a Higher-Order Language with Built-In AD. (Sept. 2016). <http://engineering.purdue.edu/~qobi/papers/ad2016b.pdf> Extended abstract, presented at the International Conference on Algorithmic Differentiation (AD).
- J. M. Siskind and B. A. Pearlmutter. 2016d. *Efficient Implementation of a Higher-Order Language with Built-In AD*. Technical Report 1611.03416. arXiv. <http://arxiv.org/abs/1611.03416>
- J. M. Siskind and B. A. Pearlmutter. 2017. *Divide-and-Conquer Checkpointing for Arbitrary Programs with No User Annotation*. Technical Report 1708.06799. arXiv. <http://arxiv.org/abs/1708.06799>
- J. M. Siskind and B. A. Pearlmutter. 2018. Divide-and-Conquer Checkpointing for Arbitrary Programs with No User Annotation. *Optimization Methods & Software (OMS)* 33, 03–04 (July 2018), 1288–1330. <http://engineering.purdue.edu/~qobi/papers/oms2018.pdf>
- D. Wingate, N. Goodman, A. Stuhlmüller, and J. M. Siskind. 2011. Nonstandard Interpretations of Probabilistic Programs for Efficient Inference. In *Proceedings of the Neural Information Processing Systems Conference (NIPS)* (Granada, Spain). <http://engineering.purdue.edu/~qobi/papers/nips2011.pdf>