

# **2022 IEEE 29th Symposium on Computer Arithmetic (ARITH 2022)**

**Virtual Conference  
12 – 14 September 2022**



**IEEE Catalog Number: CFP22121-POD  
ISBN: 978-1-6654-7828-1**

**Copyright © 2022 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP22121-POD
ISBN (Print-On-Demand):	978-1-6654-7828-1
ISBN (Online):	978-1-6654-7827-4
ISSN:	1063-6889

**Additional Copies of This Publication Are Available From:**

Curran Associates, Inc  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: (845) 758-0400  
Fax: (845) 758-2633  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2022 IEEE 29th Symposium on Computer Arithmetic (ARITH) **ARITH 2022**

## Table of Contents

Foreword .....	vii
Organizing Committee .....	viii
Program Committee .....	ix
Steering Committee .....	x
Sponsors .....	xi

### Applications 1

MiniFloat-NN and ExSdotp: An ISA Extension and a Modular Open Hardware Unit for Low-Precision Training on RISC-V Cores .....	1
<i>Luca Bertaccini (ETH Zürich, Switzerland), Gianna Paulin (ETH Zürich, Switzerland), Tim Fischer (ETH Zürich, Switzerland), Stefan Mach (Axelera AI, Switzerland), and Luca Benini (ETH Zürich and University of Bologna, Switzerland)</i>	
A BF16 FMA is All You Need for DNN Training .....	9
<i>John Osorio (Barcelona Supercomputing Center, Spain), Adrià Armejach (Barcelona Supercomputing Center, Spain), Eric Petit (Intel Corp, USA), Greg Henry (Intel Corp, USA), and Marc Casas (Barcelona Supercomputing Center, Spain)</i>	

### Arithmetic Operators 1

Low-Latency and High-Bandwidth Pipelined Radix-64 Division and Square Root Unit .....	10
<i>Javier D. Bruguera (Arm Cambridge Design Center)</i>	
High-Level Algorithms for Correctly-Rounded Reciprocal Square Roots .....	18
<i>Carlos Borges (Department of Applied Mathematics, Naval Postgraduate School), Claude-Pierre Jeannerod (Inria, ENS de Lyon, France), and Jean-Michel Muller (CNRS, ENS de Lyon, France)</i>	
The CORE-MATH Project .....	26
<i>Alexei Sibidanov (University of Victoria, British Columbia, Canada), Paul Zimmermann (Université de Lorraine, CNRS, Inria, LORIA, France), and Stéphane Gloudu (Université de Lorraine, CNRS, Inria, LORIA, France)</i>	
Enhanced Floating-Point Adder with Full Denormal Support .....	35
<i>Jongwook Sohn (Intel Corporation, USA), David K. Dean (Intel Corporation, USA), Eric Quintana (Intel Corporation, USA), and Wing Shek Wong (Intel Corporation, USA)</i>	

## Datapath Automation

Automatic Datapath Optimization using E-Graphs .....	43
<i>Samuel Coward (Intel Corporation, UK), George A. Constantinides (Imperial College London, UK), and Theo Drane (Intel Corporation, USA)</i>	

## Applications 2

Accelerating Variants of the Conjugate Gradient with the Variable Precision Processor .....	51
<i>Yves Durand (Univ. Grenoble Alpes, CEA, List, France), Eric Guthmuller (Univ. Grenoble Alpes, CEA, List, France), Cesar Fuguet (Univ. Grenoble Alpes, CEA, List, France), Jérôme Fereyre (Univ. Grenoble Alpes, CEA, List, France), Andrea Bocco (Univ. Grenoble Alpes, CEA, List, France), and Riccardo Alidori (Univ. Grenoble Alpes, CEA, List, France)</i>	
The Positive Effects of Stochastic Rounding in Numerical Algorithms .....	58
<i>El-Mehdi El Arar (Université Paris-Saclay, UVSQ, LI-PaRAD), Devan Sohler (Université Paris-Saclay, UVSQ, LI-PaRAD), Pablo de Oliveira Castro (Université Paris-Saclay, UVSQ, LI-PaRAD), and Eric Petit (Intel Corp)</i>	
PERCIVAL: Open-Source Posit RISC-V Core with Quire Capability .....	66
<i>David Mallasén (Universidad Complutense de Madrid, Spain), Raul Murillo (Universidad Complutense de Madrid, Spain), Alberto A. Del Barrio (Universidad Complutense de Madrid, Spain), Guillermo Botella (Universidad Complutense de Madrid, Spain), Luis Piñuel (Universidad Complutense de Madrid, Spain), and Manuel Prieto-Matias (Universidad Complutense de Madrid, Spain)</i>	

## Arithmetic Operators 2

Approximate Recursive Multipliers Using Low Power Building Blocks .....	67
<i>Efstratios Zacharelos (University of Naples Federico II, Italy), Italo Nunziata (University of Naples Federico II, Italy), Gerardo Saggese (University of Naples Federico II, Italy), Antonio G.M. Strollo (University of Naples Federico II, Italy), and Ettore Napoli (University of Naples Federico II, Italy)</i>	
Point-Targeted Sparseness and Ling Transforms on Parallel Prefix Adder Trees .....	68
<i>Teodor-Dumitru Ene (Oklahoma State University, USA) and James E. Stine (Oklahoma State University, USA)</i>	
Towards Quantum Logarithm Number Systems .....	76
<i>Mark Arnold (Lehigh University, USA)</i>	

## Crypto 1

PMNS for efficient arithmetic and small memory cost .....	84
<i>Fangan Yssouf Dosso (EMSE, France), Jean-Marc Robert (Universite de Toulon, France), and Pascal Véron (Universite de Toulon, France)</i>	

An Alternative Approach to Polynomial Modular Number System Internal Reduction .....	85
<i>Nicolas Meloni (Université de Toulon, France)</i>	
A Software Comparison of RNS and PMNS .....	86
<i>Laurent-Stéphane Didier (IMATH, Université de Toulon, France),</i>	
<i>Jean-Marc Robert (IMATH, Université de Toulon, France), Fangan Yssouf</i>	
<i>Dosso (SAS, École des Mines de Saint-Étienne, France), and Nadia El</i>	
<i>Mrabet (SAS, École des Mines de Saint-Étienne, France)</i>	

## Crypto 2

Efficient Word Size Modular Multiplication over Signed Integers .....	94
<i>Daichi Aoki (NEC Corporation, Japan), Kazuhiko Minematsu (NEC</i>	
<i>Corporation, Japan), Toshihiko Okamura (NEC Corporation, Japan), and</i>	
<i>Tsuyoshi Takagi (University of Tokyo, Japan)</i>	
Generating Very Large RNS Bases .....	102
<i>Jean-Claude Bajard (Sorbonne Université, CNRS, INRIA, France),</i>	
<i>Kazuhide Fukushima, (KDDI Research Inc), Thomas Plantard (France), and</i>	
<i>Arnaud Sipasseuth (KDDI Research Inc, Japan)</i>	
Quotient Approximation Modular Reduction .....	103
<i>Aurélien Greuet (IDEMIA, France), Simon Montoya (IDEMIA, France; LIX,</i>	
<i>INRIA, CNRS, École Polytechnique, IPP, France), and Clémence</i>	
<i>Vermeersch (IDEMIA, France)</i>	
Efficient Reduction Algorithms for Special Gaussian Integer Moduli .....	111
<i>Malek Safieh (Siemens AG, Technology) and Fabrizio De Santis (Siemens</i>	
<i>AG, Technology)</i>	

## Error Analysis and Formal Verification

Formal Verification of a Chained Multiply-Add Design: Combining Theorem Proving and Equivalence Checking .....	120
<i>David Russinoff (Arm, Inc., USA), Javier Bruguera (Arm, Inc., USA),</i>	
<i>Cuong Chau (Arm, Inc., USA), Mayank Manjrekar (Arm, Inc., USA),</i>	
<i>Nicholas Pfister (Arm, Inc., USA), and Harsha Valsaraju (Arm, Inc.,</i>	
<i>USA)</i>	
Bounding the Round-Off Error of the Upwind Scheme for Advection .....	127
<i>Louise Ben Salem-Knapp (CEA DAM, France), Sylvie Boldo (INRIA,</i>	
<i>France), and William Weens (CEA DAM, France)</i>	
Formally Verified 32- and 64-bit Integer Division using Double-Precision Floating-Point Arithmetic .....	128
<i>David Monniaux (Univ. Grenoble Alpes, CNRS, Grenoble INP, VERIMAG,</i>	
<i>France) and Alice Pain (Univ. Grenoble Alpes, CNRS, Grenoble INP,</i>	
<i>VERIMAG, France)</i>	
<b>Author Index</b> .....	<b>133</b>