

2020 IEEE 27th Symposium on Computer Arithmetic (ARITH 2020)

**Portland, Oregon, USA
7 – 10 June 2020**



**IEEE Catalog Number: CFP20121-POD
ISBN: 978-1-7281-7121-0**

**Copyright © 2020 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:	CFP20121-POD
ISBN (Print-On-Demand):	978-1-7281-7121-0
ISBN (Online):	978-1-7281-7120-3
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
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2020 IEEE 27th Symposium on Computer Arithmetic (ARITH) **ARITH 2020**

Table of Contents

Foreword	vii
Conference Organization	viii
Program Committee	ix
Steering Committee	x
Sponsors	xi

Regular Papers

Floating-Point Fused Multiply-Add under HUB Format	1
<i>Javier Hormigo (University of Malaga, Spain), Julio Villalba-Moreno (University of Malaga, Spain), and Sonia Gonzalez-Navarro (University of Malaga, Spain)</i>	
A Correctly-Rounded Fixed-Point-Arithmetic Dot-Product Algorithm	9
<i>Sylvie Boldo (Université Paris-Saclay, France), Diane Gallois-Wong (Université Paris-Saclay, France), and Thibault Hilaire (Sorbonne Université, France)</i>	
Heuristics for the Design of Large Multipliers for FPGAs	17
<i>Andreas Boettcher (University of Applied Sciences, Germany), Keanu Kullmann (University of Applied Sciences, Germany), and Martin Kumm (University of Applied Sciences, Germany)</i>	
Efficient, Arbitrarily High Precision Hardware Logarithmic Arithmetic for Linear Algebra	25
<i>Jeff Johnson (Facebook AI Research)</i>	
A Novel Method of Modular Multiplication Based on Karatsuba-Like Multiplication	33
<i>Zhen Gu (Tsinghua University, China) and Shuguo Li (Tsinghua University, China)</i>	
Alternative Split Functions and Dekker's Product	41
<i>Stef Graillat (Sorbonne Université, France), Vincent Lefevre (Univ Lyon, France), and Jean-Michel Muller (Univ Lyon, France)</i>	
Algorithms for Manipulating Quaternions in Floating-Point Arithmetic	48
<i>Mioara Joldes (CNRS, LAAS, France) and Jean-Michel Muller (CNRS, LIP, Université de Lyon, France)</i>	
A Hole in the Ladder: Interleaved Variables in Iterative Conditional Branching	56
<i>Yoann Marquer (Univ Rennes, CNRS, IRISA, France) and Tania Richmond (Univ Rennes, CNRS, IRISA, France)</i>	

Highly Optimized Montgomery Multiplier for SIKE Primes on FPGA .64.....	64
<i>Rami Elkhatib (Florida Atlantic University), Reza Azarderakhsh (Florida Atlantic University), and Mehran Mozaffari Kermani (University of South Florida)</i>	
Fast, Small, and Area-Time Efficient Architectures for Key-Exchange on Curve25519 .72.....	72
<i>Mojtaba Bisheh Niasar (Florida Atlantic University), Reza Azarderakhsh (Florida Atlantic University), Mehran Mozaffari Kermani (University of South Florida), and Rami Elkhatib (Florida Atlantic University)</i>	
An Asymptotically Faster Version of FV Supported on HPR .80.....	80
<i>Jean-Claude Bajard (Sorbonne Université, CNRS, Inria, Institut de Mathématiques de Jussieu – Paris Rive Gauche, France), Julien Eynard (Univ. Grenoble Alpes, CEA, LETI, DSYS, CESTI, F-38000 Grenoble), Paulo Martins (INESC-ID, Instituto Superior Técnico, Universidade de Lisboa), Leonel Sousa (INESC-ID, Instituto Superior Técnico, Universidade de Lisboa), and Vincent Zucca (imec-COSIC KU Leuven)</i>	
Maximum Delay Models for Parallel-Prefix Adders in the Presence of Threshold Voltage Variations .88.....	88
<i>Kleanthis Papachatzopoulos (University of Patras, Patras, Greece) and Vassilis Paliouras (University of Patras, Patras, Greece)</i>	
Variable Precision 16-Bit Floating-Point Vector Unit for Embedded Processors .96.....	96
<i>Alberto Nannarelli (Technical University, Denmark)</i>	
A Framework for Semi-Automatic Precision and Accuracy Analysis for Fast and Rigorous Deep Learning .103.....	103
<i>Christoph Lauter (University of Alaska Anchorage (UAA)) and Anastasia Volkova (Université de Nantes, CNRS, LS2N F-44000 Nantes, France)</i>	
Variable-Radix Coding of the Reals .111.....	111
<i>Peter Lindstrom (Center for Applied Scientific Computing Lawrence Livermore National Laboratory)</i>	
Automatic Design Space Exploration for an Error Tolerant Application .117.....	117
<i>Samuel Coward (Visual Technologies Team Intel Corporation), Theo Drane (Visual Technologies Team Intel Corporation), and Yoav Harel (Visual Technologies Team Intel Corporation)</i>	

Short Papers

Custom-Precision Mathematical Library Explorations for Code Profiling and Optimization .121.....	121
<i>David Defour (University of Perpignan), Pablo de Oliveira Castro (University of Versailles – Li-PaRAD; ECR), Matei Istoan (University of Versailles – Li-PaRAD; ECR), and Eric Petit (Intel Corporation; ÉCR)</i>	
SIMD Multi Format Floating-Point Unit on the IBM z15(TM) .125.....	125
<i>Stefan Payer (Compute Unit Development IBM Deutschland R&D GmbH), Cedric Lichtenau (Compute Unit Development IBM Deutschland R&D GmbH), Michael Klein (Compute Unit Development IBM Deutschland R&D GmbH), Kerstin Schelm (Compute Unit Development IBM Deutschland R&D GmbH), Petra Leber (Compute Unit Development IBM Deutschland R&D GmbH), Nicol Hofmann (Compute Unit Development IBM Deutschland R&D GmbH), and Tina Babinsky (Compute Unit Development IBM Deutschland R&D GmbH)</i>	

Issues with Rounding in the GCC Implementation of the ISO 18037:2008 Standard Fixed-Point Arithmetic .129.....
Mantas Mikaitis (The University of Manchester)

Intel Nervana Neural Network Processor-T (NNP-T) Fused Floating Point Many-Term Dot Product .133.....
Brian Hickmann (Hillsboro, OR, USA), Jiasheng Chen (Folsom, CA, USA), Michael Rotzin (Santa Clara, CA USA), Andrew Yang (Santa Clara, CA USA), Maciej Urbanski (Gdansk, Poland), and Sasikanth Avancha (Bangalore, India)

Author Index 137......