

Contents

Acknowledgments	v
Abstract	vii
Zusammenfassung	ix
1 Introduction	1
1.1 Energy-Efficient Processing	3
1.1.1 Near-Threshold Operation	3
1.1.2 Parallel Processing	4
1.2 Improving Energy Efficiency	5
1.2.1 SIMD Capabilities	5
1.2.2 Hardware Accelerator	6
1.2.3 ISA Extensions	6
1.3 Required Number Formats	7
1.4 Contributions and Publications	8
1.5 Outline of this Thesis	11
2 Background and Preliminaries	13
2.1 Processing Architectures	13
2.1.1 Energy-Efficient Micro-Controllers	15
2.1.2 Data Parallel Architectures and DSPs	15
2.1.3 Hardware Accelerators	16
2.1.4 Multi-Core Architectures	17
2.1.5 ASIPs	17
2.2 Different Number Formats	18

2.2.1	Fixed-Point and High-Dynamic-Range Arithmetic	18
2.3	Evaluation Platform	19
2.3.1	Requirements on the Processing Elements . . .	21
3	The RISC-V Core Architecture and its ISA Extensions	23
3.1	Pipeline Architecture	24
3.1.1	Advantages of the RISC-V ISA	26
3.2	Multi-Core Integration Challenges	28
3.3	ISA Extensions and Micro-Architectural Optimizations	32
3.3.1	Instruction Fetch Unit	33
3.3.2	Hardware Loops	34
3.3.3	Load/Store Unit	36
3.3.4	EX-Stage: ALU	38
3.3.5	EX-Stage: Multiplication	42
3.4	Toolchain Support	45
3.5	Experimental Results	46
3.5.1	Area, Frequency, and Power	47
3.5.2	Instruction Level Performance	49
3.5.3	Function Kernel Performance	51
3.5.4	Convolution Performance	51
3.5.5	Near-Threshold Operation	56
3.6	Summary	58
4	Exploration of the Logarithmic Number System	61
4.1	Related Work	63
4.2	Preliminaries	66
4.2.1	LNS Number Representation, and Format . . .	66
4.2.2	Arithmetic Operations in the LNS Domain . .	67
4.2.3	Rounding Modes and Precision	67
4.2.4	Cotransformations and Fitting Framework . . .	68
4.3	LNU Architecture	69
4.3.1	Main LNU Blocks	69
4.3.2	Trigonometric Functions	72
4.4	Processor Integration	73
4.4.1	Dispatcher and Interconnect	75
4.4.2	LNS ALU-Extensions	76
4.5	Results	77
4.5.1	LNU Results and Comparison with Related Work	77

4.5.2	Designed Chip Variants	81
4.5.3	Instruction Level Performance	85
4.5.4	Function Kernel Performance	85
4.6	Results of Approximate LNUs	88
4.6.1	LNU Design Space Exploration	89
4.6.2	Impact of Approximation on Performance	91
4.7	Summary	96
5	A Sharing Framework for Complex ISA Extensions	99
5.1	The Sharing Concept	100
5.2	A Shared APU	104
5.2.1	Interconnect	104
5.2.2	Dispatcher	105
5.2.3	Shared Unit and Arbitration	108
5.3	Shared DSP Extensions	109
5.4	FP Support for PULP	111
5.4.1	FP Components	113
5.4.2	Core Extensions for FP Support	116
5.4.3	Scalability of the Shared Unit	117
5.5	Experimental Results	120
5.5.1	Area savings	121
5.5.2	Optimal Configuration for NT Operation	125
5.5.3	Application Kernel Performance	128
5.5.4	Energy Savings	132
5.6	A Case Study: Seizure Detection with the Shared FPU	135
5.6.1	Seizure Detection Pipeline	136
5.6.2	Parallel Seizure Detection Performance and Optimal Number of Subunits	138
5.7	Summary	141
6	Summary and Conclusion	143
6.1	Overview of the Main Results	144
6.2	Outlook	147
A	Chip Gallery	149
A.1	28 nm PULP Chips	150
A.2	65 nm PULP Chips	151
A.3	FPU Chips	152

A.4	LNU Chips	153
A.5	Single Core Chips	154
A.6	Other Chips	155
B	Notation and Acronyms	157
	Operators	157
	Acronyms	158
	Bibliography	163
	Curriculum Vitae	177