Appendix

Table A1: Model definitions for the fSIR task.

Model	Layer 1	Layer 2	Layer 3
NPU	NPU$(3,\ h)$	NAU$(h,\ 3)$	–
NPU	NPU$_{\text{real}}(3,\ h)$	NAU$(h,\ 3)$	–
Dense	Dense$(2,\ h, \sigma)$	Dense$(h,\ h,\ \sigma)$	Dense$(h,\ 3)$

Table A2: Testing error on the simple arithmetic task for the different models (i.e. mean of each heatmap in Fig. 6). Each value is obtained by computing median (and median absolute deviation) of the error of 20 models.

Task	NPU	RealNPU	NMU	NALU	iNALU	Dense
$+$	0.2 ± 0.11	**0.08 ± 0.021**	0.2 ± 0.18	2.69 ± 0.22	2.18 ± 0.13	2.103 ± 0.04
$\times$	0.37 ± 0.23	0.066 ± 0.026	**0.005 ± 0.004**	4.55 ± 0.2	3.453 ± 0.065	3.546 ± 0.035
$\div$	0.23 ± 0.13	**0.085 ± 0.038**	11.399 ± 0.035	3.33 ± 0.18	2.54 ± 0.26	14.16 ± 0.23
$\sqrt{\cdot}$	0.031 ± 0.025	**0.004 ± 0.001**	0.16 ± 0.002	0.034 ± 0.006	0.049 ± 0.011	0.084 ± 0.007

Table A3: Model definitions for the simple arithmetic task.

Model	Layer 1	Layer 2	Layer 3
NPU	NAU$(2, 6)$	NPU$(6, 2)$	–
RealNPU	NAU$(2, 6)$	RealNPU$(6, 2)$	–
NMU	NAU$(2, 6)$	NMU$(6, 2)$	–
NALU	NALU$(2, 6)$	NALU$(6, 2)$	–
iNALU	iNALU$(2, 6)$	iNALU$(6, 2)$	–
Dense	Dense$(2, 10, \sigma)$	Dense$(10, 10, \sigma)$	Dense$(10, 2)$

Table A4: Model definitions for the large scale arithmetic task.

Model	Layer 1	Layer 2
NPU	NAU(100, 100)	NPU(100, 1)
NPU	NAU(100, 100)	NPU(100, 1)
NMU	NAU(100, 100)	NMU(100, 1)
NALU	NALU(100, 100)	NALU(100, 1)

Table A5: Dataset parameters for the large scale arithmetic task.

Task	Input size	Subset ratio	Overlap ratio	Training range	Validation range
Add	100	0.5	0.25	Sobol(-1,1)	Sobol(-4,4)
Mult	100	0.5	0.25	Sobol(-1,1)	Sobol(-4,4)
Div	100	0.5	–	Sobol(0,0.5)	Sobol(-0.5,0.5)
Sqrt	100	0.5	–	Sobol(0,2)	Sobol(0,4)

Table A6: Training parameters for the large scale arithmetic task. The β-parameters define the stepwise exponential growth of the L_1 regularization with start, step, growth, and end.

Task	Learning rate	Iterations	β_{start}	β_{end}	β_{step}	β_{growth}
Add	1e-2	1e5	1e-5	1e-4	10 000	10
Mult	5e-3	1e5	1e-5	1e-7	10 000	10
Div	5e-3	1e5	1e-9	1e-7	10 000	10
Sqrt	5e-3	1e5	1e-6	1e-4	10 000	10