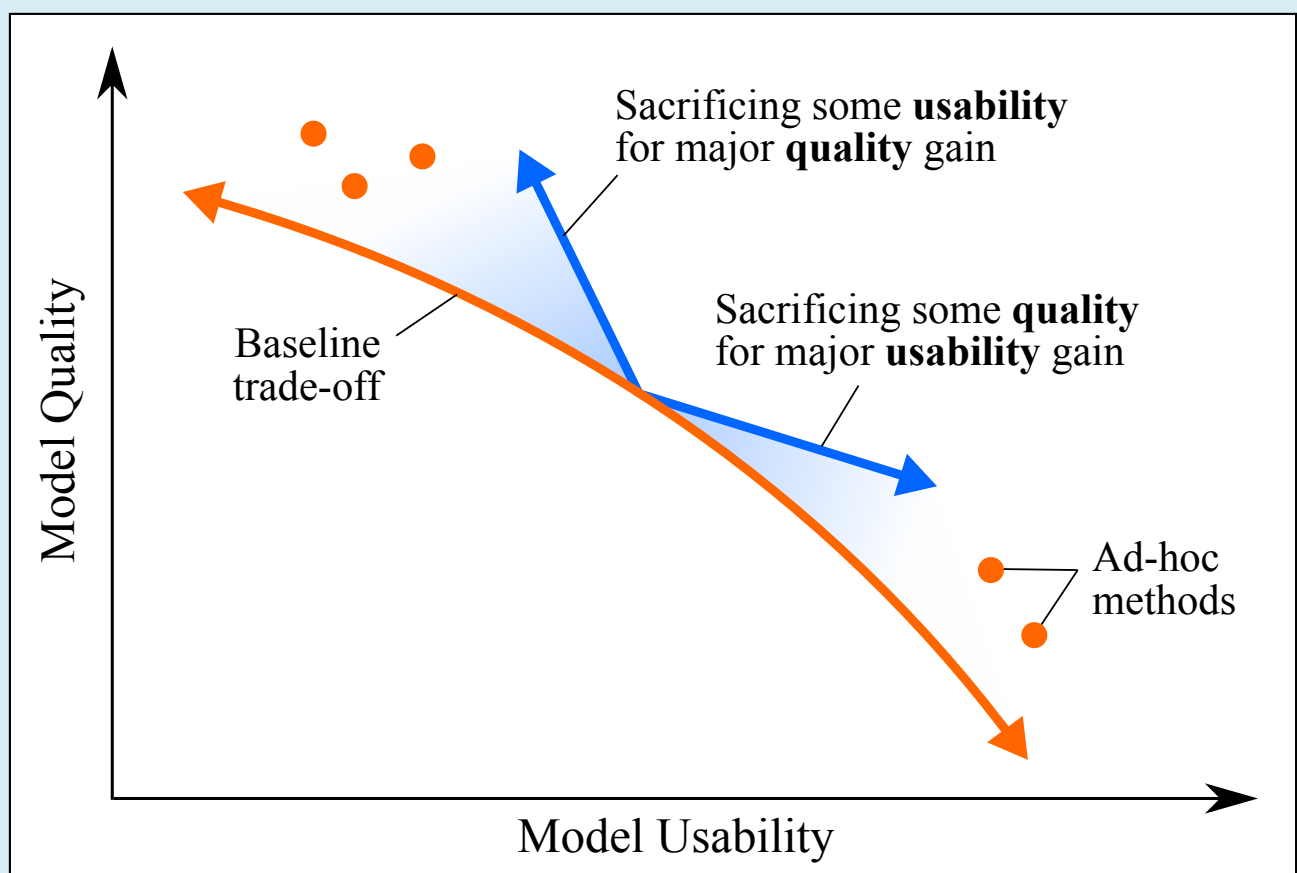# Selective Abstraction and Stochastic Methods for Scalable Power Modelling of Heterogeneous Systems

A. Rafiev, F. Xia, A. Iliasov, R. Gensh, A. M. M. Aalsaud, A. Romanovsky, A. Yakovlev

School of Computing Science and *School of Electrical and Electronic Engineering*
**Newcastle University**

Abstract — With the increase of system complexity in both platforms and applications, power modelling of heterogeneous systems is facing grand challenges from the model scalability issue. To address these challenges, this paper studies two systematic methods: selective abstraction and stochastic techniques. The concept of selective abstraction via black-boxing is realised using hierarchical modelling and cross-layer cuts, respecting the concepts of boxability and error contamination. The stochastic aspect is formally underpinned by Stochastic Activity Networks (SANs). The proposed method is validated with experimental results from Odroid XU3 heterogeneous 8-core platform and is demonstrated to maintain high accuracy while improving scalability.

Motivation and Contributions

- We developed new structuring methods to tackle complexity and scalability in modelling by providing a power-proportionality metric for selective abstraction and methods to retain accuracy by avoiding error contamination.
- We validated these methods using power modelling in SANs and showed their effectiveness in improving the trade-offs between

The Proposed Method
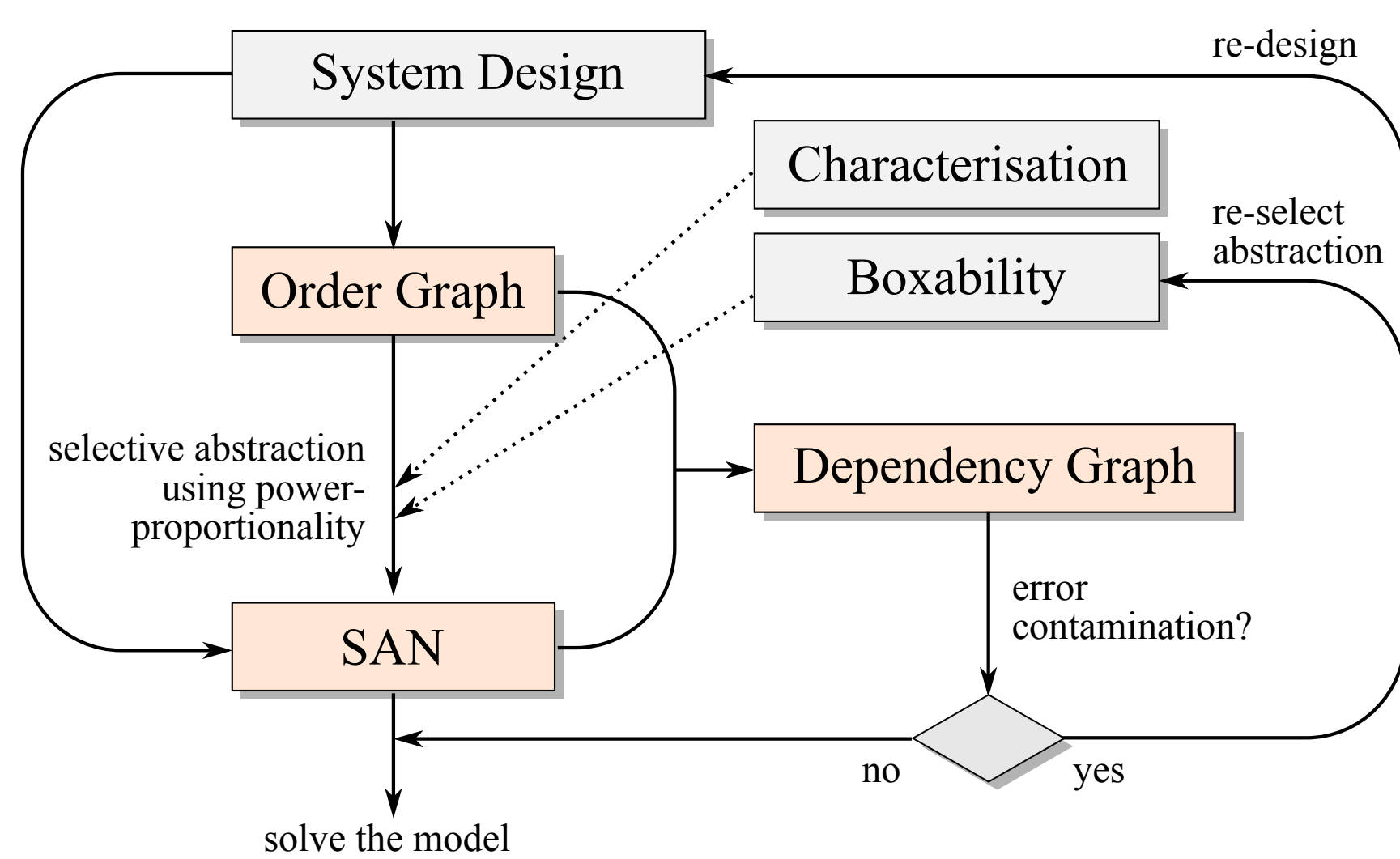
Hierarchical Modelling: Order Graphs (OGs)
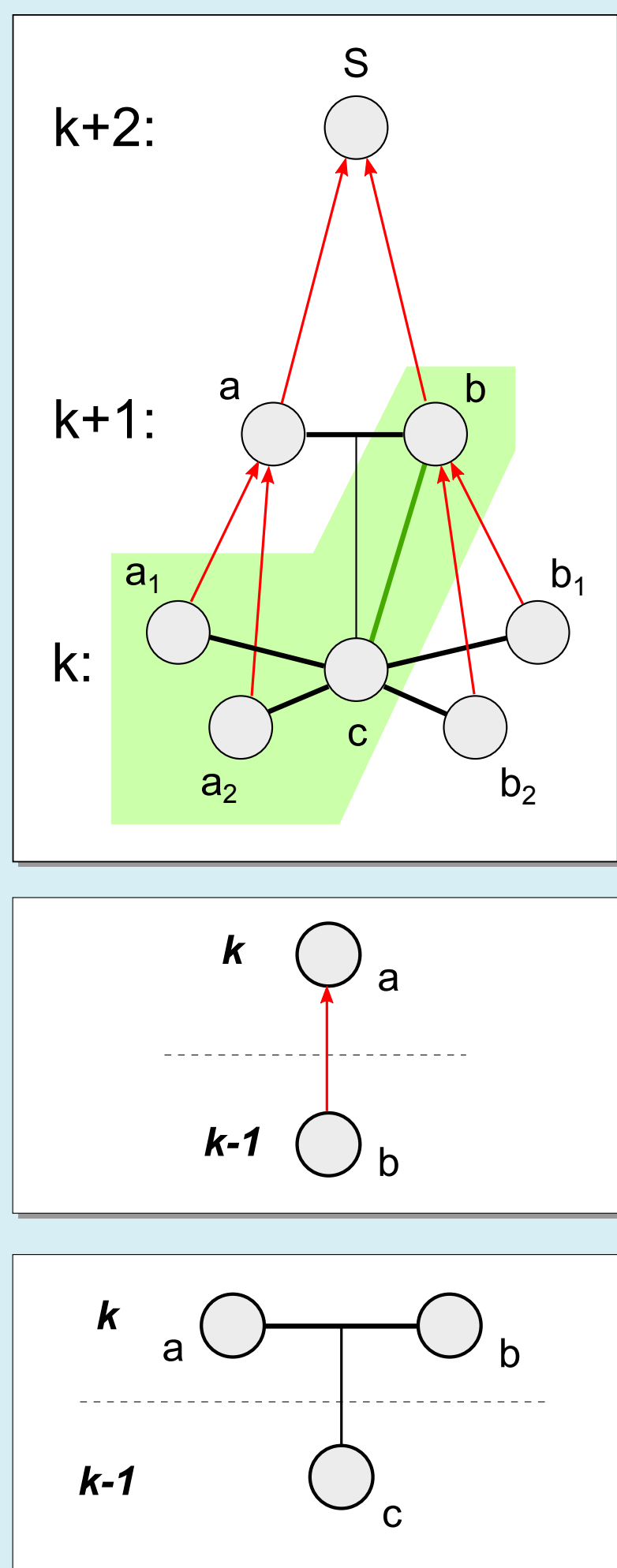
Order graph is:
– an infinite graph-like structure;
– any element of an order graph is an order graph;
– represents an inter-related family of graphs.

***k*-order graph**
– is a specific projection defining a single graph;
– made of nodes and edges of order k;
– all higher orders are (temporarily) disregarded.

Relative graph orders: $(k-1)$ sub-graph, $(k+1)$ super-graph.

◄ Example of an OG with highlighted **cross-layer cut**, which joins elements from different orders (layers of abstraction).

Inclusion / containment arc:
– connects nodes in different orders;
– form a tree.

Dependency arc:
– connects nodes of the same order.

Support arc:
– connects a node and a dependency arc in different orders.

Selective Abstraction Metric

The goal of selective abstraction is to obtain a cut that provides the minimal model while its added error satisfies the given threshold ε: $|\Delta E| < \varepsilon$.

$$|\Delta E| = \left|\Delta e_x \frac{p_x}{p}\right|$$

- Δe_x is the local change of the percentage error, as a result of the black-boxing, in the part x being black-boxed;
- p_x is the power consumed by this part;
- p is the total power consumption.

Case Study: Odroid XU3 Platform

- 28nm 8-core Application Processor Exynos 5422, based on the **ARM big.LITTLE** architecture:
  – high performance **Cortex-A15** quad-core processor block,
  – low power **Cortex-A7** quad-core block,
  – Mali-T628 GPU,
  – 2GB LPDDR3 DRAM.
- Realtime current sensors measuring four separate power domains: **A7**, **A15**, GPU and DRAM.

Model Evaluation Framework
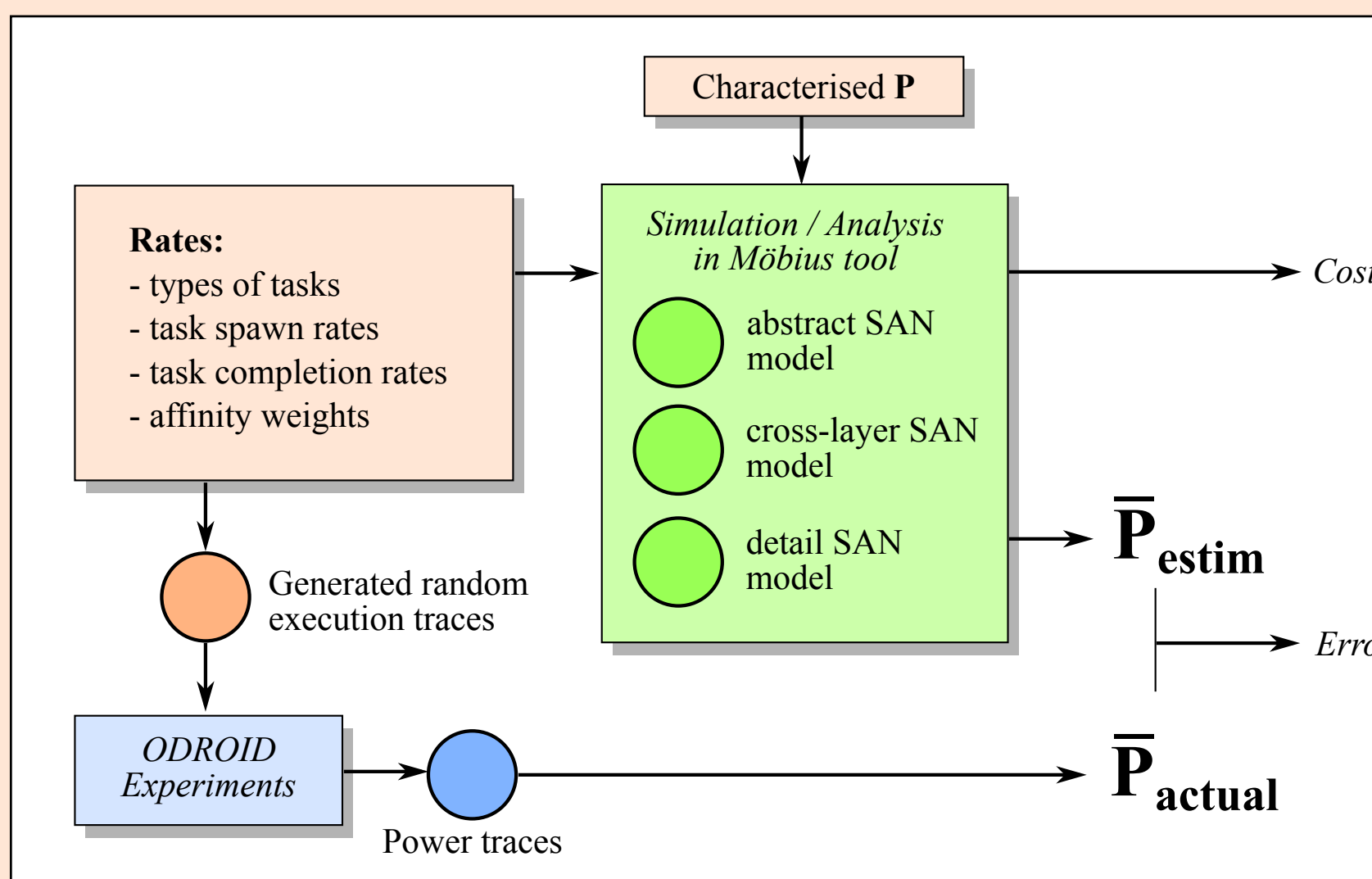
- **Abstract model (1+1)** is the most abstract model with two meta-cores, one representing the A7 domain, the other representing A15 ($n = 2$).
- **Cross-layer model (1+4)** is the model obtained using the proposed method of selective abstraction. Here, three A7 cores are grouped into a single meta-core representing the entire domain ($n = 5$).
- **Detail model (3+4)** is the most detailed model considering each core separately ($n = 7$).

Models

SAN (1+1) model for the naïve scenario:

SAN (1+1) model with fixed affinities:
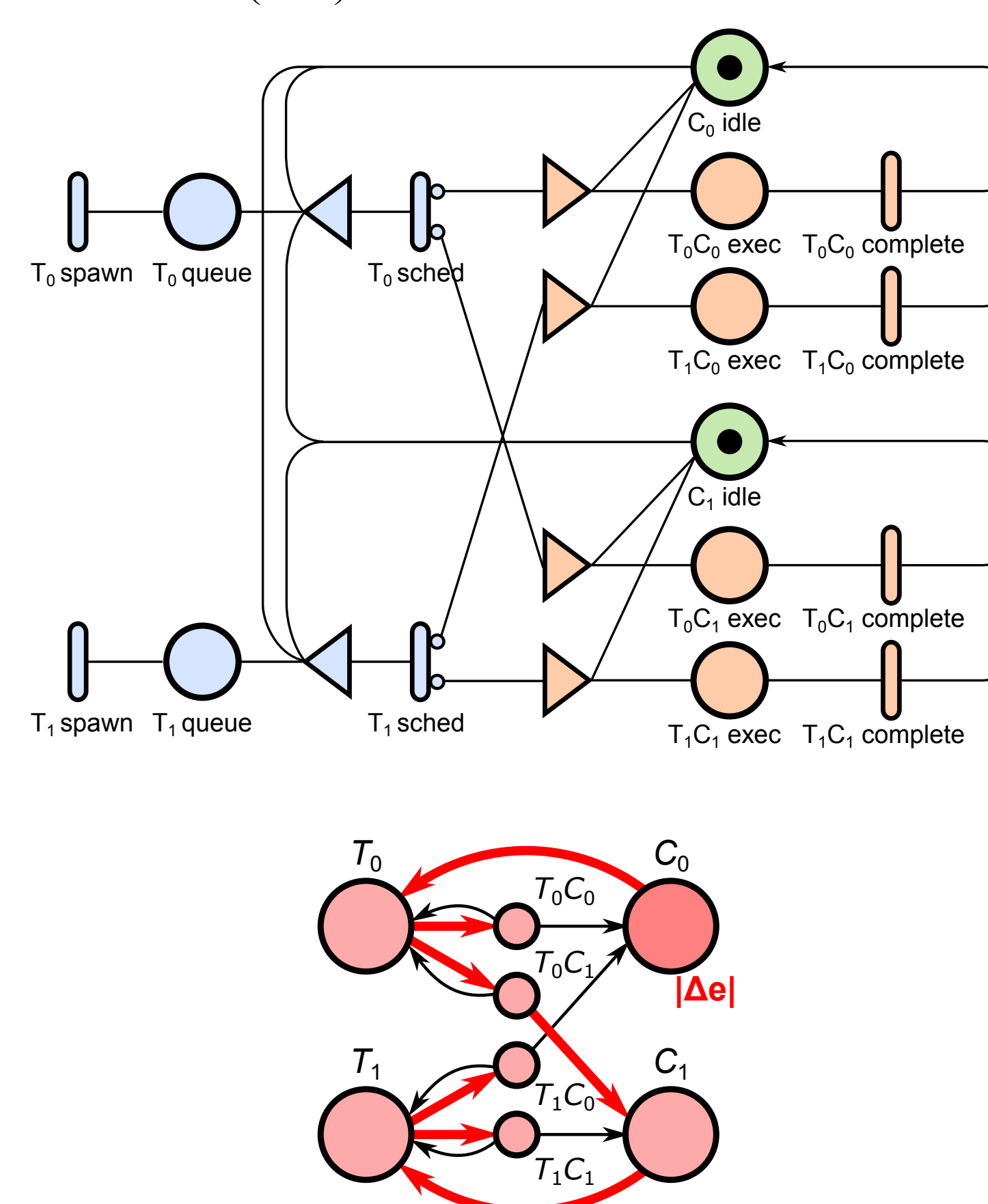
Corresponding dependency graph with highlighted error contamination paths.
Error contamination
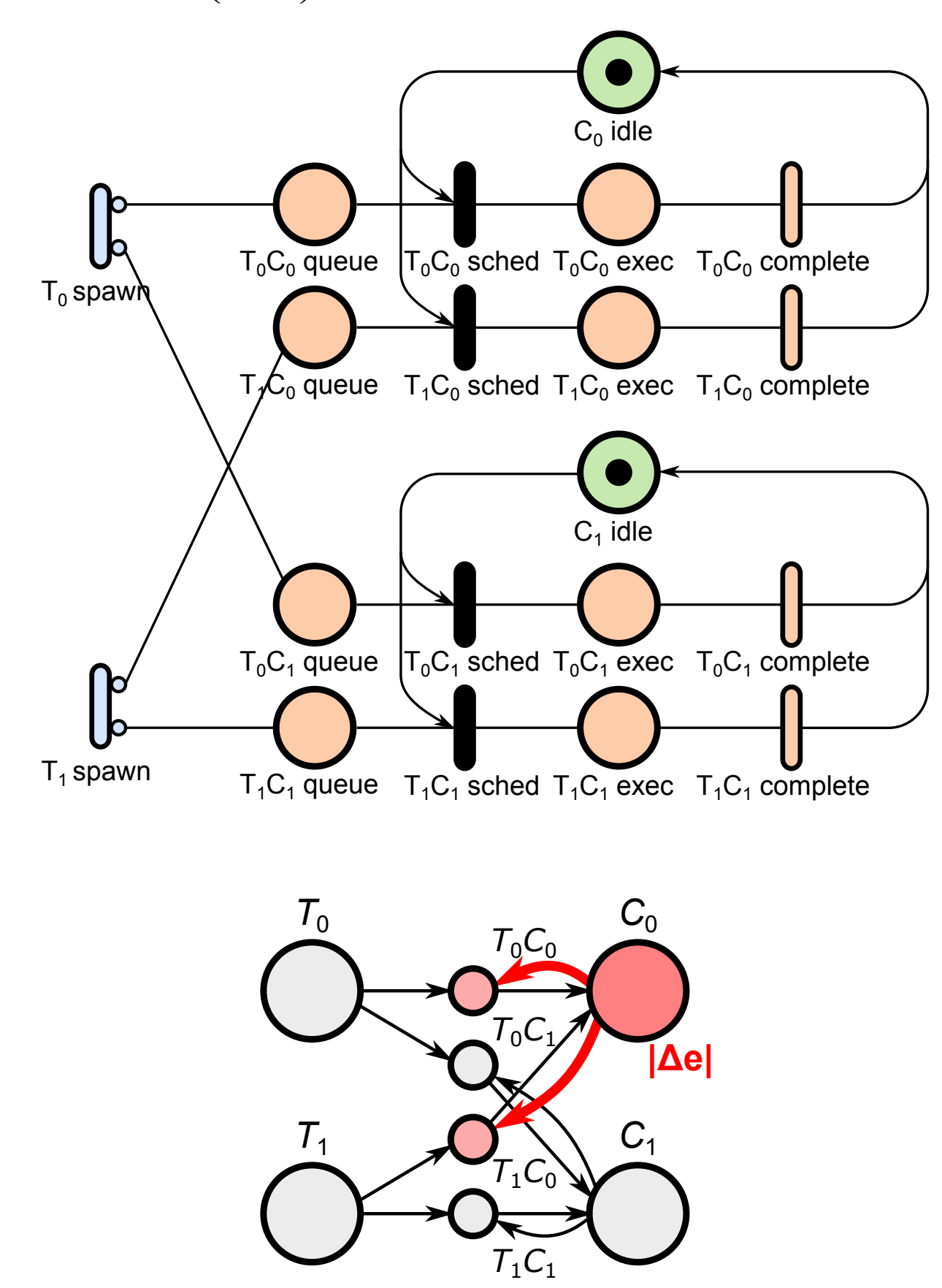
Corresponding dependency graph with highlighted error contamination paths.
No error contamination

Error contamination is a property of system design and not a model artifact. It can be detected through model analysis, and the design can be modified to provide boxability.

Simulation Results

scen	model	pwr, W	pwr var	error	sim, s
EQ	1+1	1.7662	0.0470	6.53%	2.535
	1+4	1.7287	0.0424	4.27%	2.546
	3+4	1.7215	0.0423	3.84%	2.764
	meas.	1.6579	0.0572		
CA	1+1	2.0205	0.0619	7.99%	2.545
	1+4	1.9470	0.0468	4.06%	2.610
	3+4	1.9421	0.0468	3.79%	2.764
	meas.	1.8711	0.0385		
TCA	1+1	2.0038	0.0608	10.41%	2.547
	1+4	1.9274	0.0439	6.21%	2.613
	3+4	1.9245	0.0440	6.05%	2.771
	meas.	1.8148	0.0279		
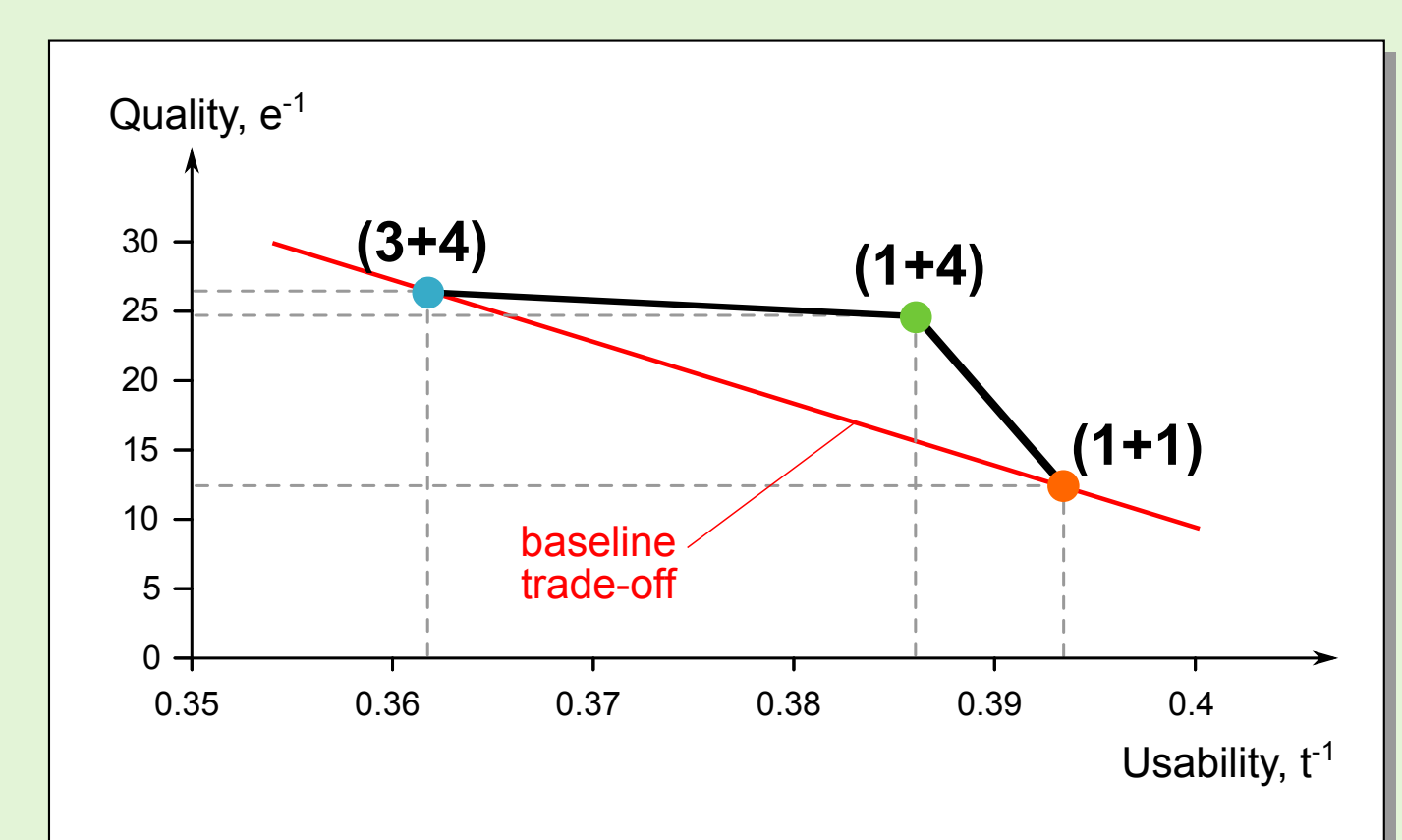
The error added by moving from (3+4) to (1+4) model in comparison to going from (3+4) straight to (1+1) is proportional to the power output of A7 domain in relation to the total power.
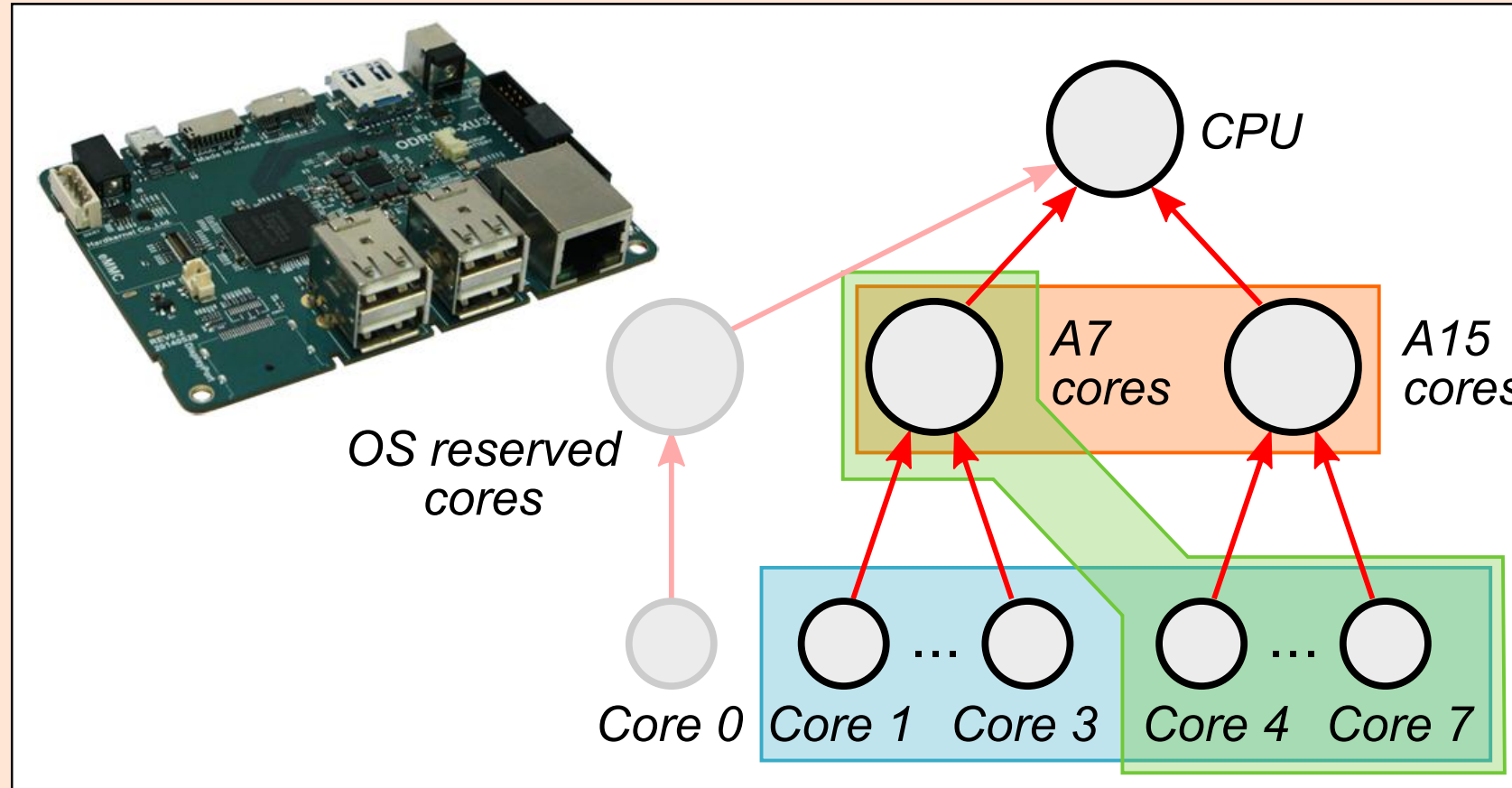
Conclusions

We propose a new method for scalable power modelling of multi-core heterogeneous systems that:

– supports the systematic discovery of good trade-offs between modelling quality and model scalability;
– rationalises model sizes based on power proportional representation and stochastic modelling;
– identifies error contamination and determines boxability.

The method is effective in:

– discovering good trade-offs between modelling quality and model scalability;
– choosing the model size;
– reducing the designer's effort.

Bibliography

[1] A. Rafiev, F. Xia, A. Iliasov, R. Gensh, A. M. M. Aalsaud, A. Romanovsky, A. Yakovlev. *Order Graphs and Cross-layer Parametric Significance-driven Modelling* in Proc. to ACSD, 2015.

[2] M. A. N. Al-hayanni, A. Rafiev, R. Shafik, F. Xia. *Power and Energy Normalized Speedup Models for Heterogeneous Many Core Computing* in Proc. to ACSD, 2016.

[3] W. Sanders, J. Meyer. *Lectures on Formal Methods and Performance Analysis*, volume LNCS2090, chapter Stochastic Activity Networks: Formal Definitions and Concepts, pages 315-343. Springer, 2001.

[4] P. Greenhalgh. *big.LITTLE Processing with ARM Cortex-A15 & Cortex-A7 - Improving Energy Eciency in High-Performance Mobile Platforms*. ARM, 2011. White Paper.

[5] A. Aalsaud, R. Shafik, A. Rafiev, F. Xia, S. Yang, A. Yakovlev. *Power-Aware Performance Adaptation of Concurrent Applications in Heterogeneous Many-Core Systems* in Proc. to ISLPED, 2016.

[6] S. Yang, R. Shafik, G. Merrett, E. Stott, J. Levine, J. Davis, B. Al-Hashimi. *Adaptive Energy Minimization of Embedded Heterogeneous Systems Using Regressionbased Learning* in Proc. to PATMOS, 2015.