SHAPING

TOMORROW

ISC High Performance

JUNE 24 - JULY 2, 2021 | ISC-HPC.COM

Physical Oscillator Model
For Parallel Distributed Computing

AYESHA AFZAL[1,2], GEORG HAGER[1], GERHARD WELLEIN[1,2]

[1] Erlangen National High Performance Computing Center
[2] Department of Computer Science,
University of Erlangen-Nürnberg Germany

36th International Conference on High Performance Computing,
ISC High Performance 2021 Digital, June 29, 2021

Relevance of white-box performance modelling

$e.g., \max(T_{BW}, T_{flops})$ — e.g., $\lambda + \frac{V}{B}$

T_{exec}	$T_{non-exec}$	T_{exec}	$T_{non-exec}$	…
T_{exec}	$T_{non-exec}$	T_{exec}	$T_{non-exec}$	…

Relevance of white-box performance modelling

Afzal et al.

IEEE Cluster 2019

DOI:10.1109/CLUSTER.2019.8890995

ISC HPC 2020

DOI: 10.1007/978-3-030-50743-5_20

ISC HPC 2021

arXiv: 2103.03175 cs.DC]

$e.g., \max(T_{BW}, T_{flops})$ e.g., $\lambda + \frac{V}{B}$

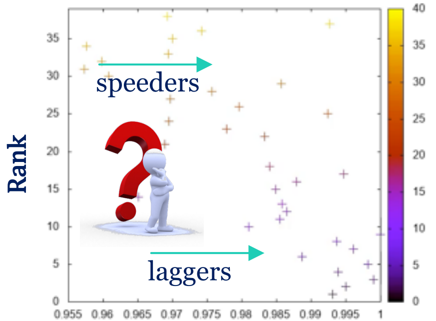
T_{exec}	$T_{non-exec}$	T_{exec}	$T_{non-exec}$	…
T_{exec}	$T_{non-exec}$	T_{exec}	$T_{non-exec}$	…

Accelerator Slowdown

$$T_{dist} = sum(T_{exec}, T_{non-exec}) \pm T_{\delta}$$

Distributed performance model

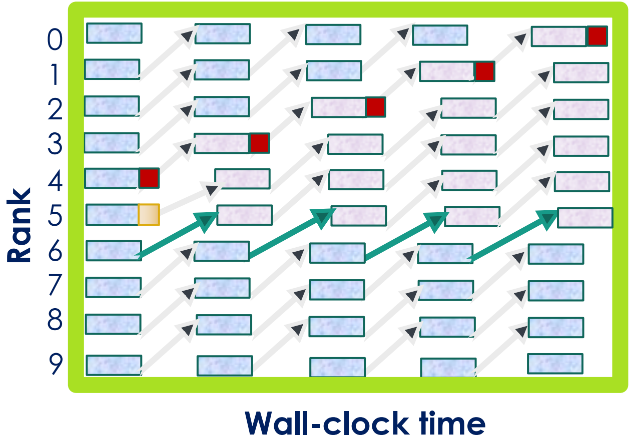
T_{δ}
(Out of) lock-step



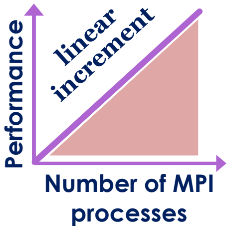
Wall-clock time normalized to slowest process

T_{exec}	$T_{non-exec}$	T_{exec}	$T_{non-exec}$		…
	T_{exec}	$T_{non-exec}$	T_{exec}	$T_{non-exec}$	…

Idle wave: progressing idleness due to processes dependencies



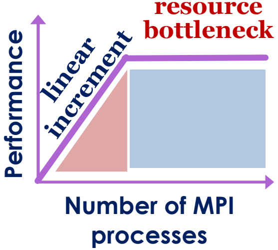
Resource scalability verses bottleneck



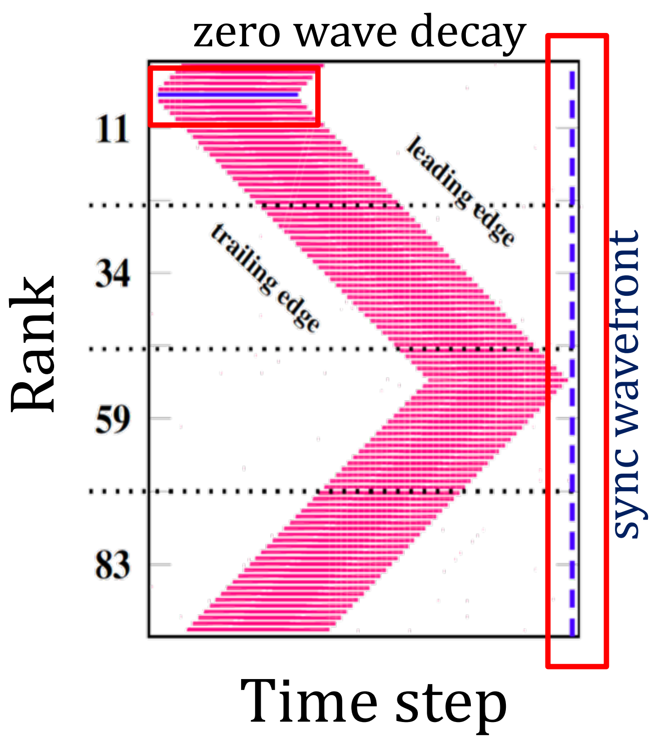
resource-scalable

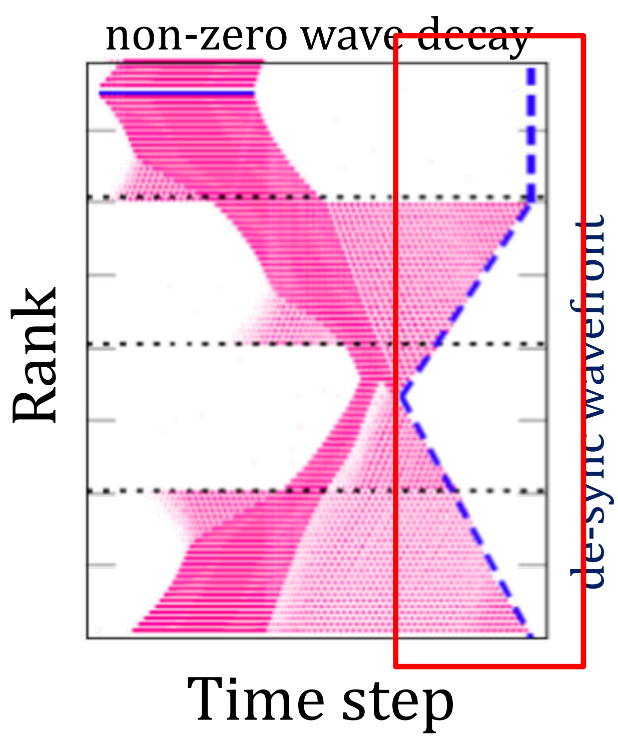
Possible bottlenecks

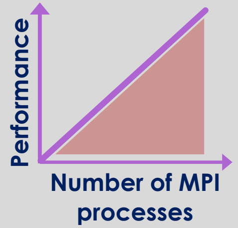
- **memory**
- cache
- on-/inter-chip network
- link b/w host & accelerator



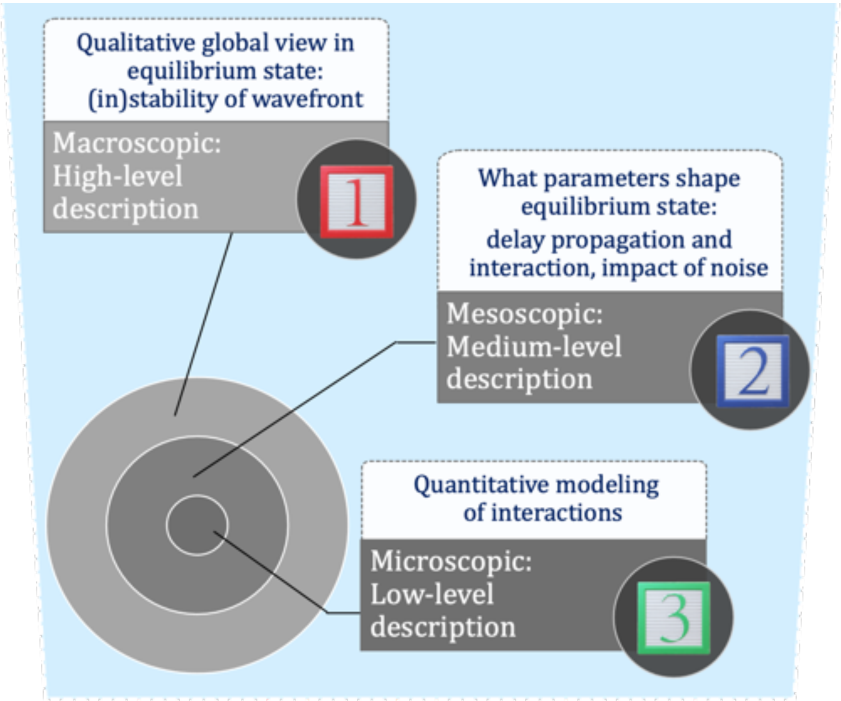
memory-bandwidth bottleneck

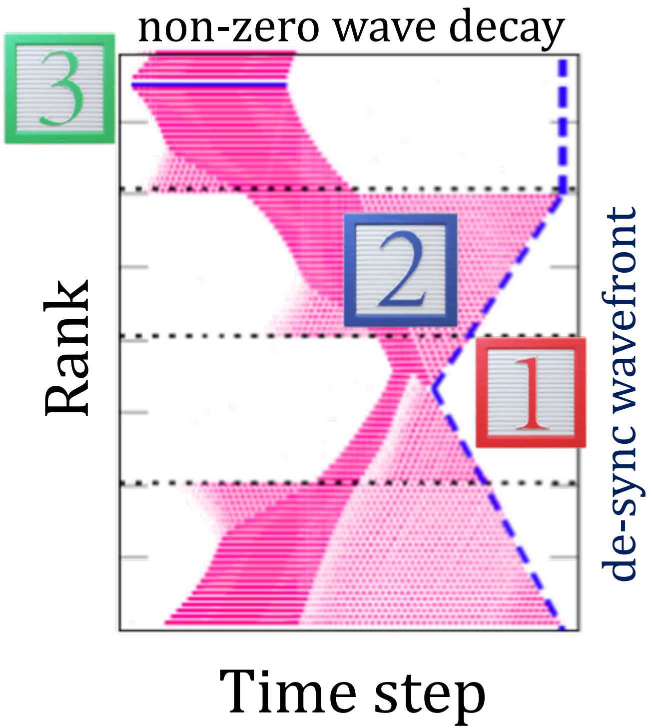

zero wave decay
11
34
59
83
Rank
leading edge
trailing edge
sync wavefront
Time step


non-zero wave decay
Rank
de-sync wavefront
Time step

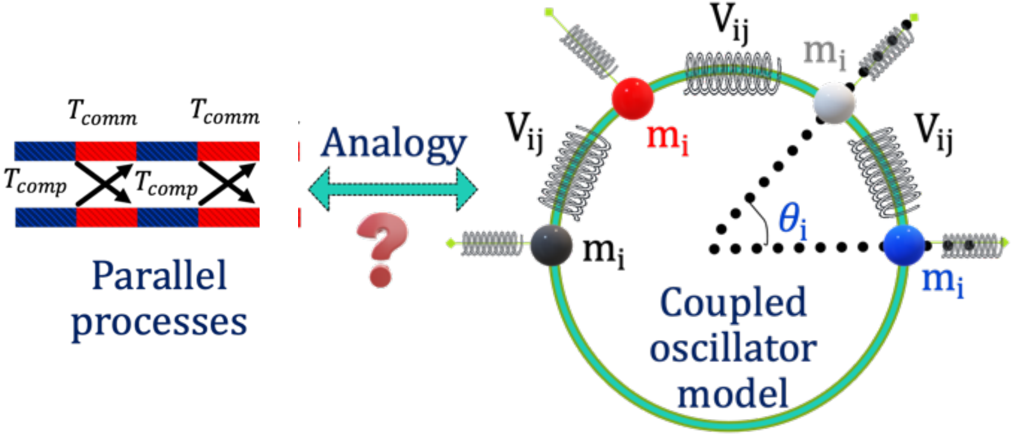

Performance
Number of MPI processes

Model Applicability

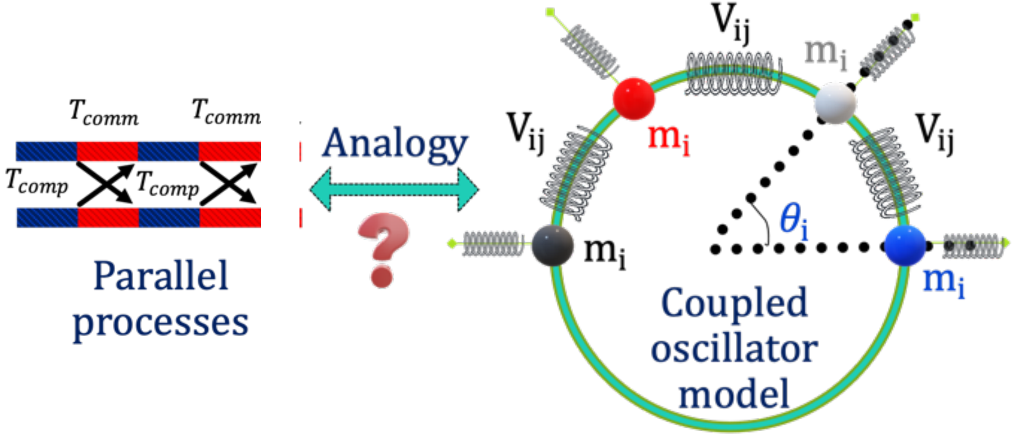




Motivation

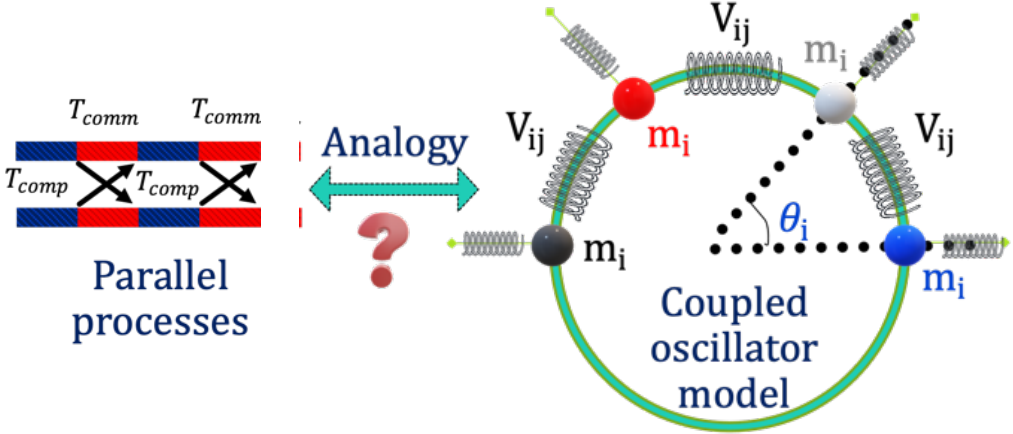


Simple and cheap experiments with solving a system of ODEs than running highly parallel programs

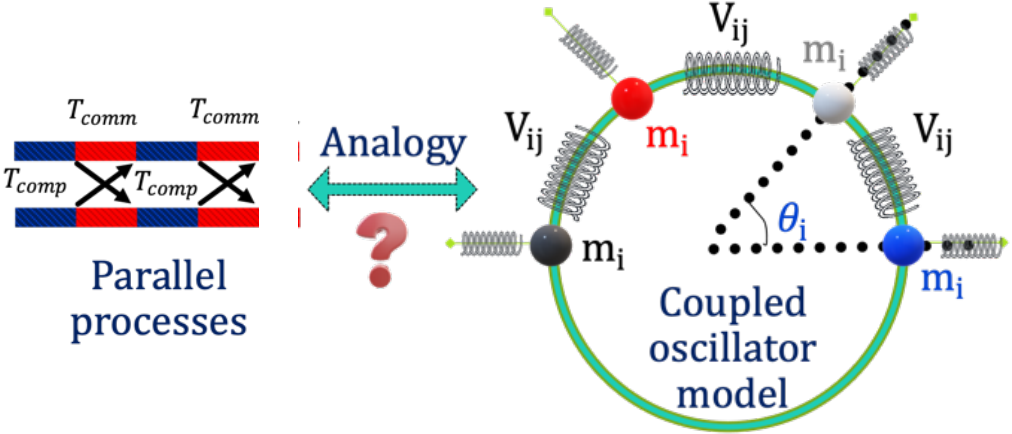


Unknown analogy
in the parallel
computing field

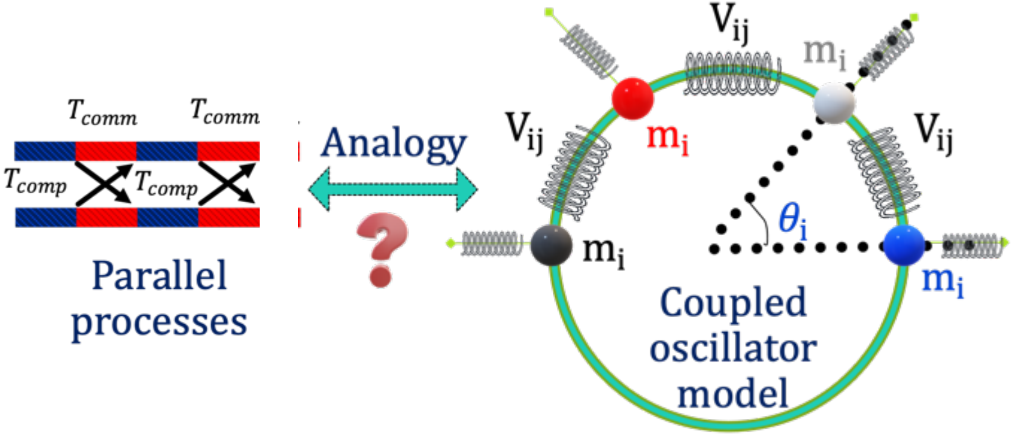
Motivation



Closing the gap: analytical modelling with flexible parameters to mimic the behavior of parallel programs



Equilibrium state: stability of (de/re)sync wavefront



Non-equilibrium state: topology-aware idle wave propagation

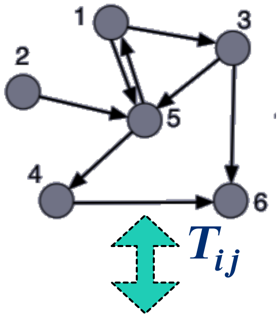
Parallel Programs		Physical Model
number of processes $Comm\_size$	⟷	number of oscillators N
alternate back-and-forth $1/\left(T_{comp} + T_{comm}\right)$	⟷	intrinsic natural frequency ω_i
variability in T_{comp} and T_{comm} of processes	⟷	phase differences (noise) or frequencies spread (load imbalance) of oscillators
delay propagation speed (the distance in ranks travelled in one time step)	⟷	coupling strength K

Physical coupled oscillators model vs. Kuramoto model

Inverse time step	Noise	Interaction term

$$\dot{\theta}_i = \omega_i + \zeta_i(t) + \frac{\sigma.\kappa}{sum(T_{exec}, T_{comm}).Comm\_size} \sum_{j=1}^{N} T_{ij} V_{ij}(\theta(t,\tau))$$

Kuramoto model	Coupled Oscillator Model
$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} sin(\theta_j - \theta_i)$	
global full-connectivity	topology-aware pairwise local interactions
no delay propagation over timestep	delay propagation over timestep
period of an oscillator (one revolution) = 2π	period of a process (one timestep) = $T_{comm} + T_{comm}$
too continuous Ansatz causing built-in noise	discrete Ansatz
equilibrium state and its stability is governed by trigonometric sinusoidal potential	equilibrium state and its stability is governed by scalable and contended processes model



	1	2	3	4	5	6
1	0	0	1	0	1	0
2	0	0	0	0	1	0
3	0	0	0	0	1	1
4	0	0	0	0	0	1
5	1	0	0	1	0	0
6	0	0	0	0	0	0

Physical coupled oscillators model vs. Kuramoto model

Inverse time step	Noise	Interaction term

$$\dot{\theta}_i = \omega_i + \zeta_i(t) + \frac{\sigma.\kappa}{sum(T_{exec}, T_{comm}).Comm\_size} \sum_{j=1}^{N} T_{ij} V_{ij}(\theta(t, \tau))$$

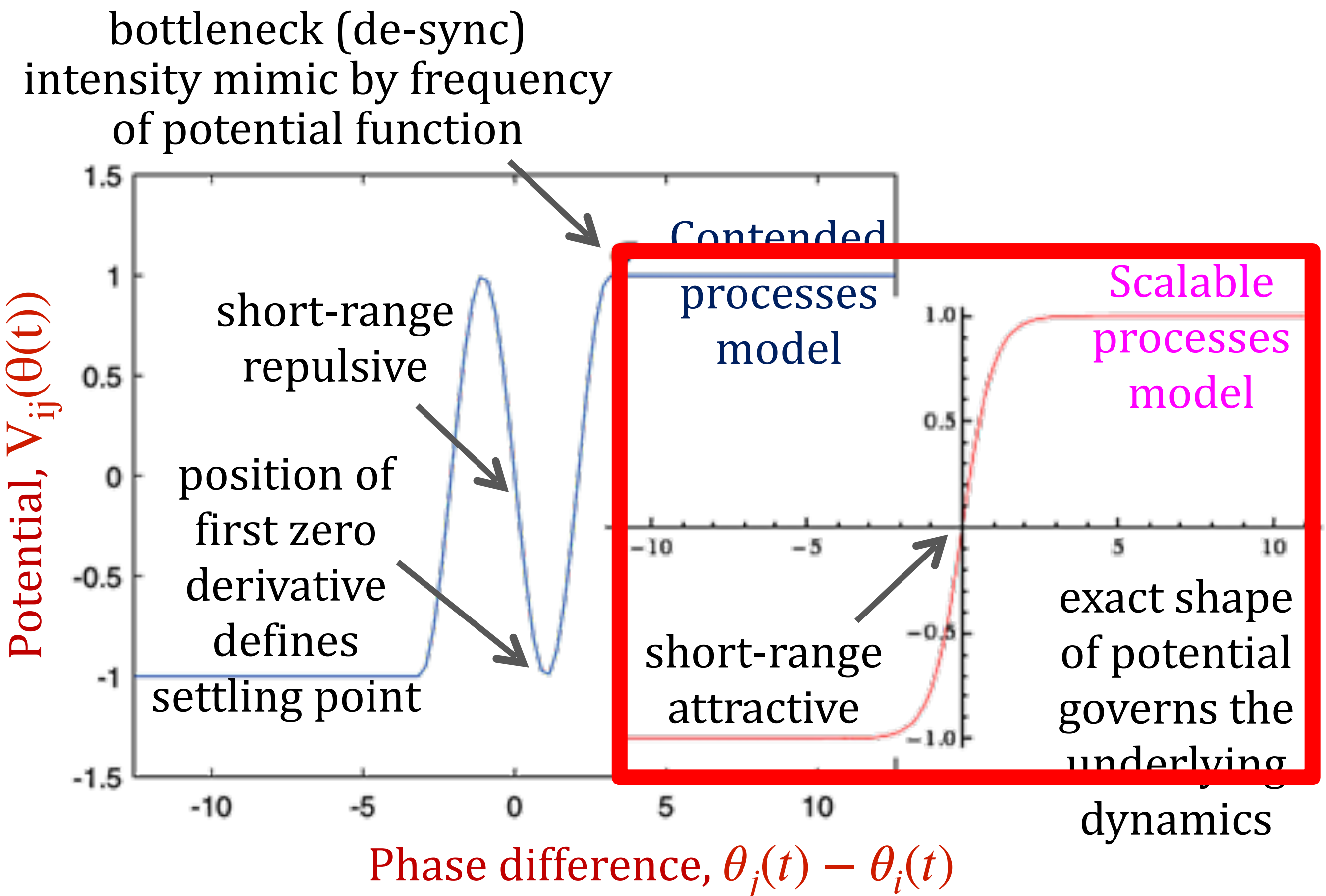
Afzal et al.

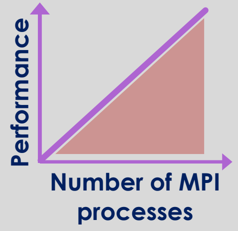
ISC HPC 2021

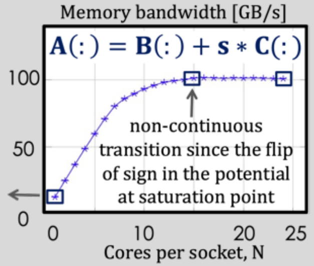
arXiv: 2103.03175 cs.DC]

Kuramoto model		Coupled Oscillator Model
$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} sin(\theta_j - \theta_i)$		
global full-connectivity	⟷	topology-aware pairwise local interactions
no delay propagation over timestep	⟷	delay propagation over timestep
period of an oscillator (one revolution) = 2π	⟷	period of a process (one timestep) = $T_{comm} + T_{comm}$
too continuous Ansatz causing built-in noise	⟷	discrete Ansatz
equilibrium state and its stability is governed by trigonometric sinusoidal potential	⟷	equilibrium state and its stability is governed by scalable and contended processes model

Interaction potential vs. resource scalability visualization

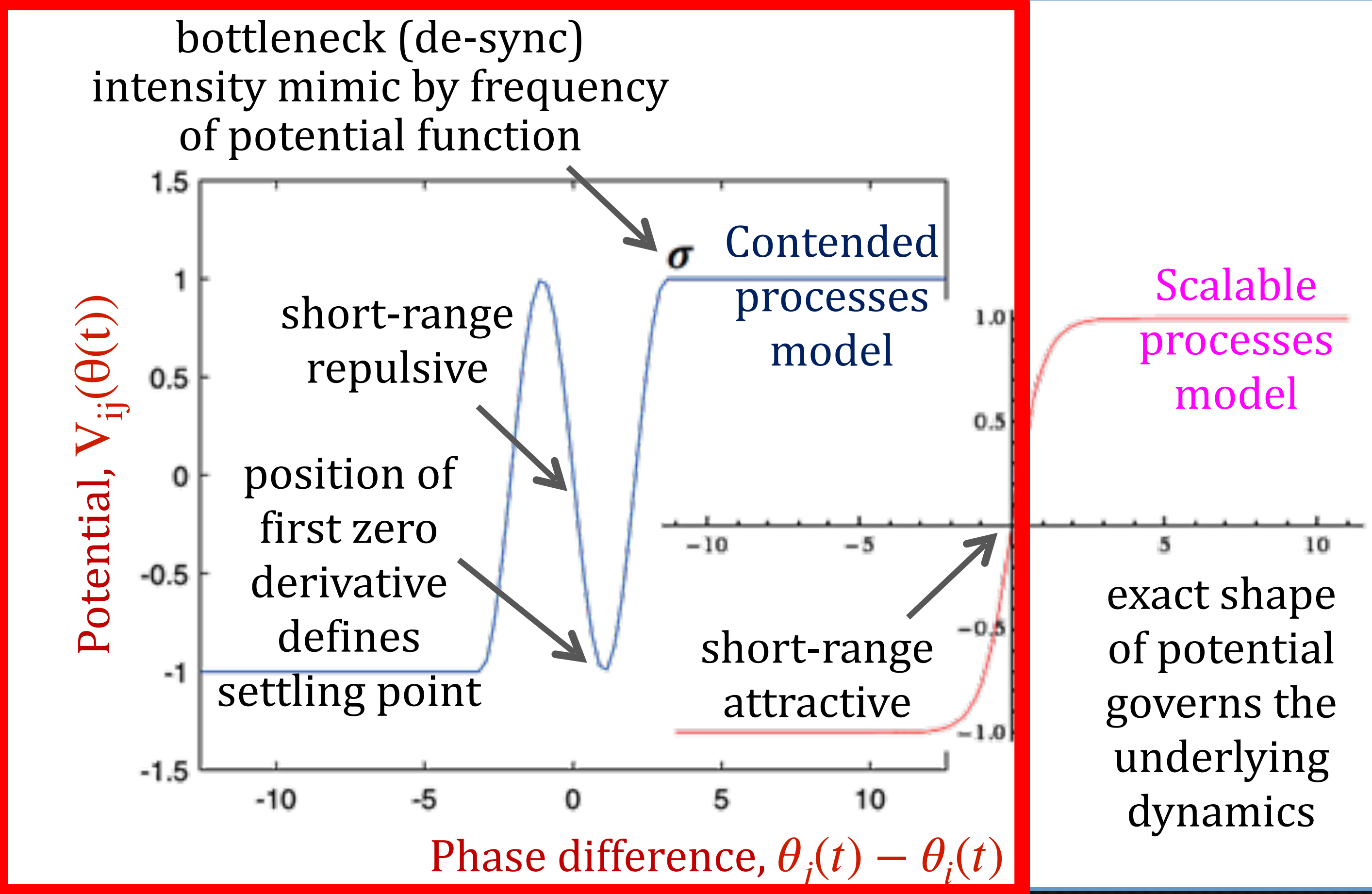


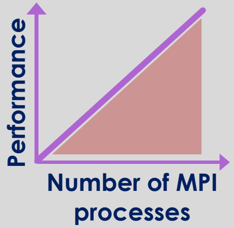


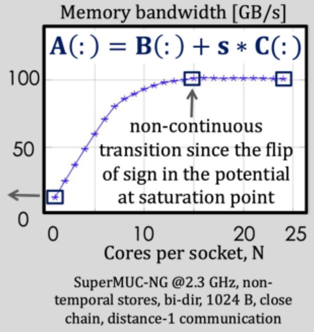


SuperMUC-NG @2.3 GHz, non-temporal stores, bi-dir, 1024 B, close chain, distance-1 communication

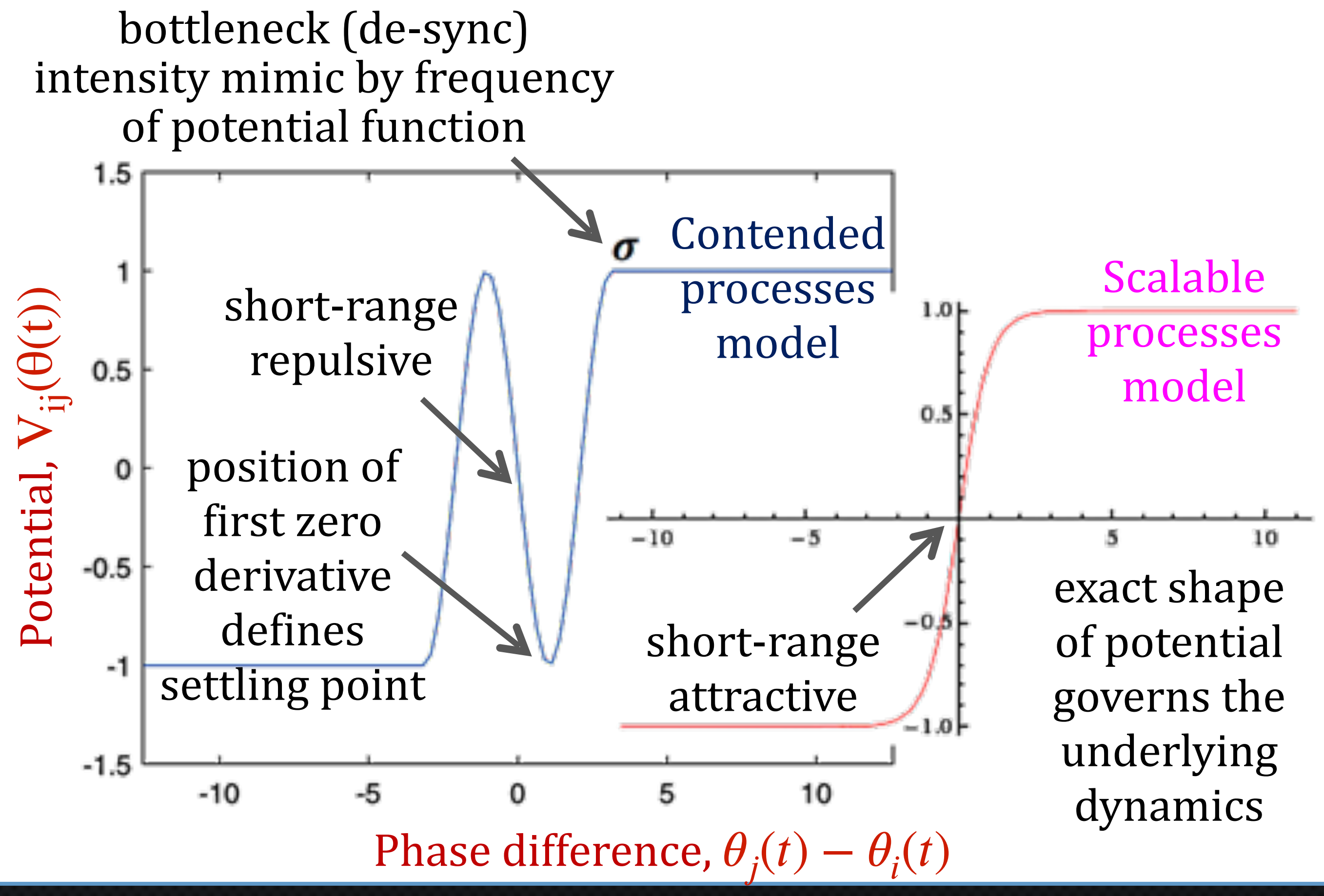
Interaction potential vs. resource scalability visualization

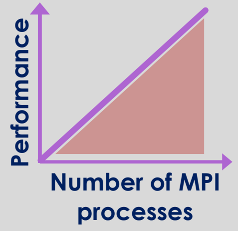


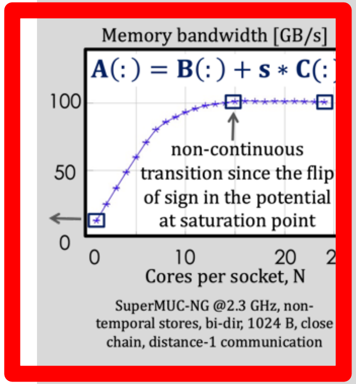




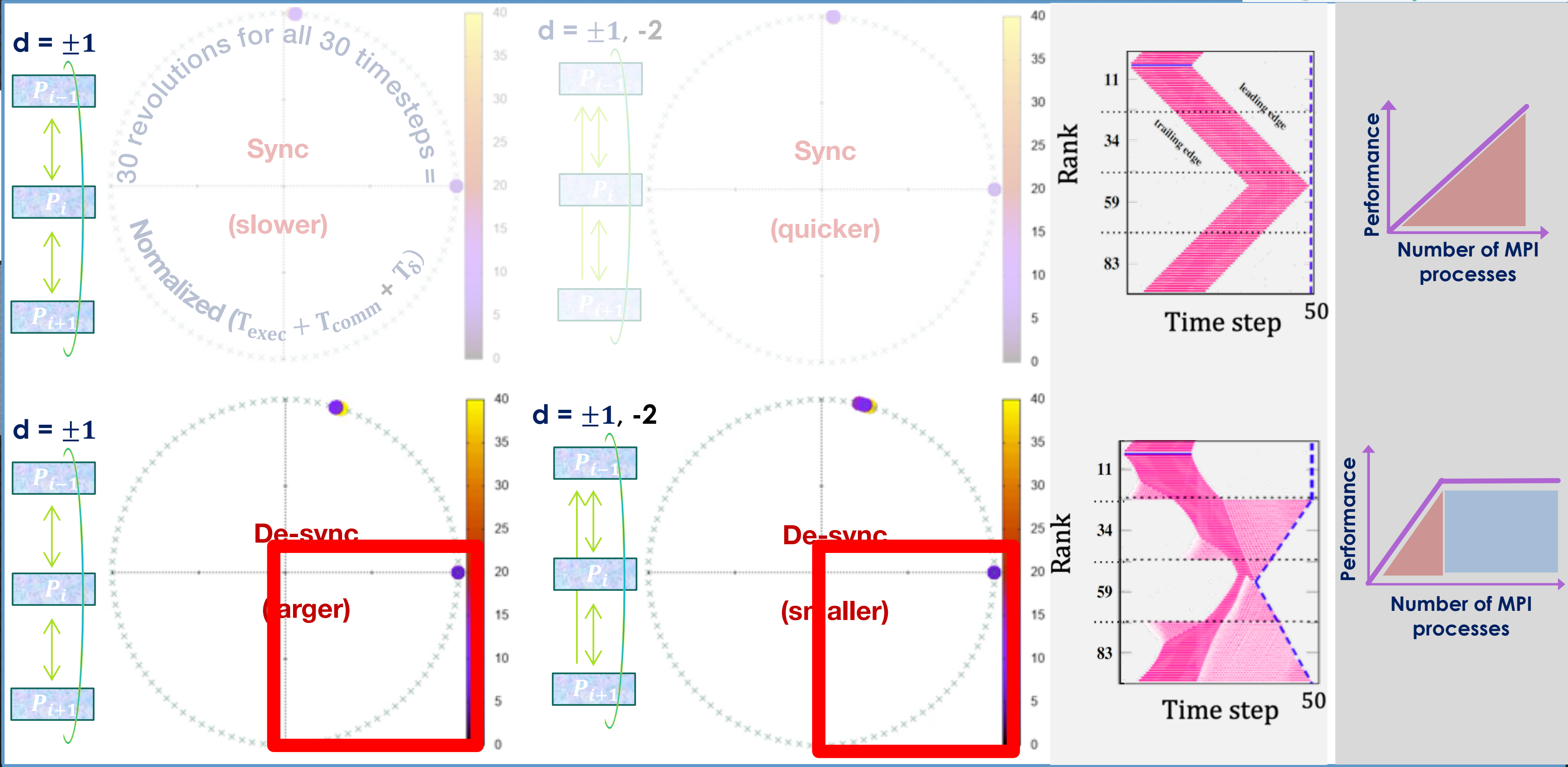
Interaction potential vs. resource scalability visualization





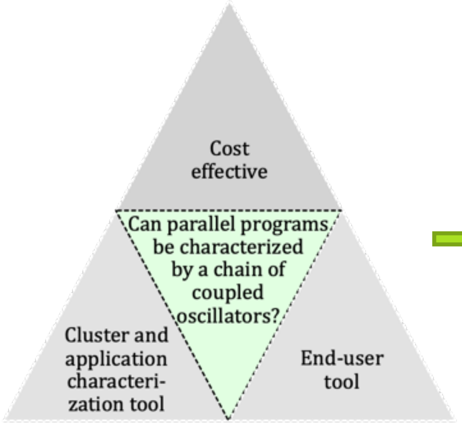


Evaluation: Resource scalability verses bottleneck



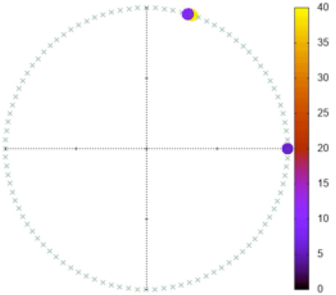
Outlook and future work

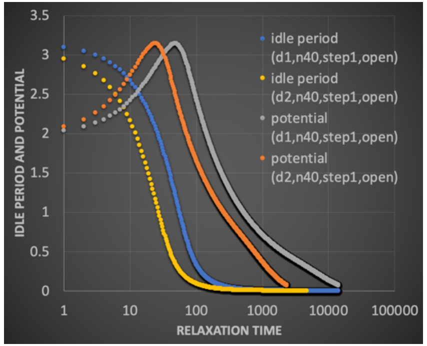
Model Improvement

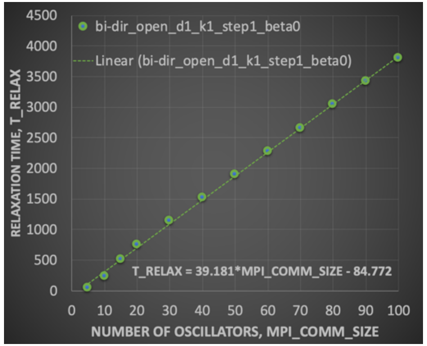


End-user tool

Visualization







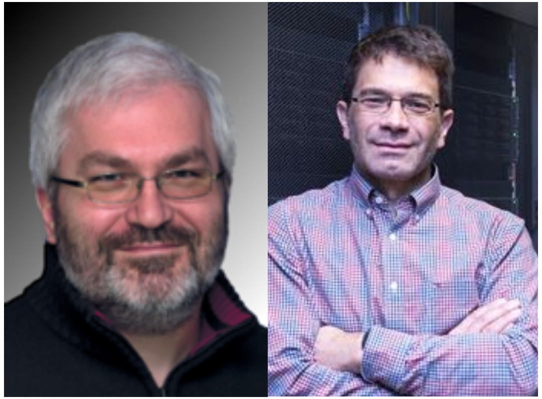
SHAPING TOMORROW

ISC High Performance

Title | Physical Oscillator Model For Parallel Distributed Computing

Contact | Ayesha Afzal
ayesha.afzal@fau.de

Acknowledgement | Co-authors
Georg Hager, Gerhard Wellein



POSTER ARTIFACTS

https://tiny.cc/NDCLPMAE-ISC21

POSTER

https://tiny.cc/NDCLPMAE-ISC21

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

NHR FAU

ISC High Performance — SHAPING TOMORROW

Title: Physical Oscillator Model For Parallel Distributed Computing

Contact: Ayesha Afzal
ayesha.afzal@fau.de

Acknowledgement: Co-authors
Georg Hager, Gerhard Wellein

OMI4papps

KONWIHR

Thank you for Listening!

POSTER ARTIFACTS

https://tiny.cc/NDCLPMAE-ISC21

POSTER

https://tiny.cc/NDCLPMAE-ISC21