Efficient Large Scale Simulation of Stochastic Lattice Models on GPUs

Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming

21st May 2015

NVIDIA GPU CENTER OF EXCELLENCE

HZDR

HELMHOLTZ ZENTRUM DRESDEN ROSSENDORF
NANONET

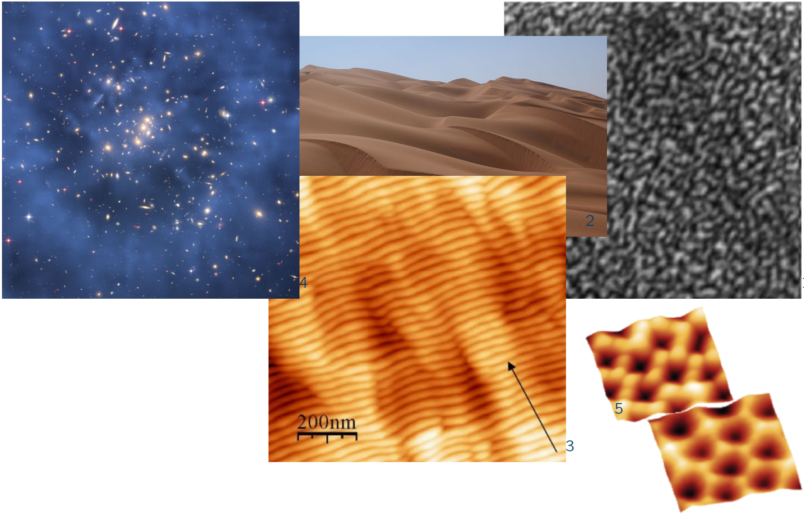
HELMHOLTZ ZENTRUM DRESDEN ROSSENDORF

Acknowledgements

- Henrik Schulz
- Nils Schmeißer
- Michael Bussmann
- Nagy-Egri Máté Ferenc
- Martin Weigel
- my other colleagues

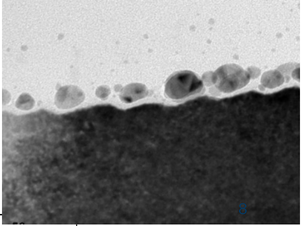
NVIDIA GPU CENTER OF EXCELLENCE

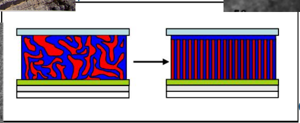
Stochastic Processes in Nature

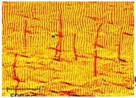


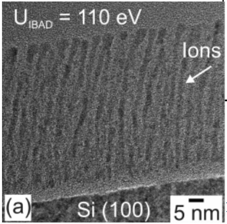
Some Technical Applications


7


8

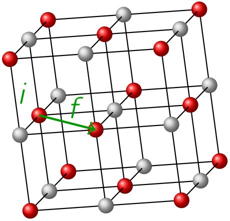

6


9



10

Outline

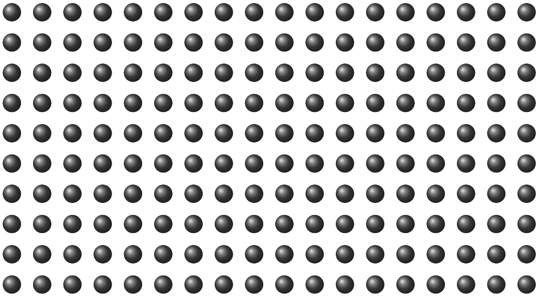


Ising Model

- interacting spins on a lattice

$$H = \sum_{<ij>} J s_i s_j$$

- simple model for magnetic ordering

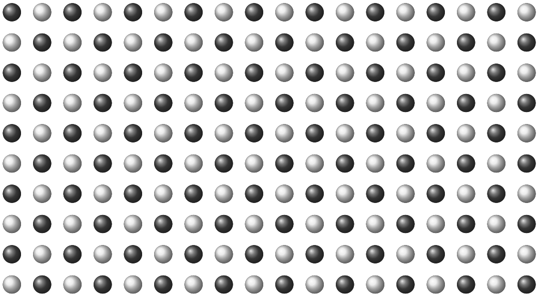


Ising Model

- interacting spins on a lattice

$$H = \sum_{<ij>} J s_i s_j$$

- simple model for magnetic ordering



checkerboard decompostion[1]

[1] Weigel, M. *J. Comp. Phys.* **231**(8) 306 (2012)
Preis, T. et al. *J. Comp. Phys.* **228** 4468 (2009)

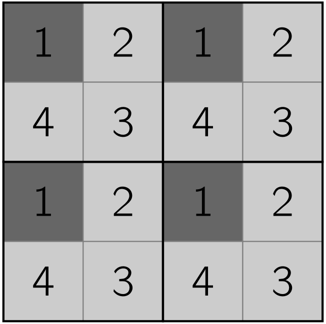
Markov-Chain Stochastic Cellular Automaton

- kinetic model requires sequence of states without memory
- sequence of states ⇒ inherently sequential

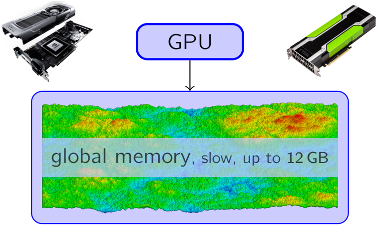
Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Markov-Chain Stochastic Cellular Automaton

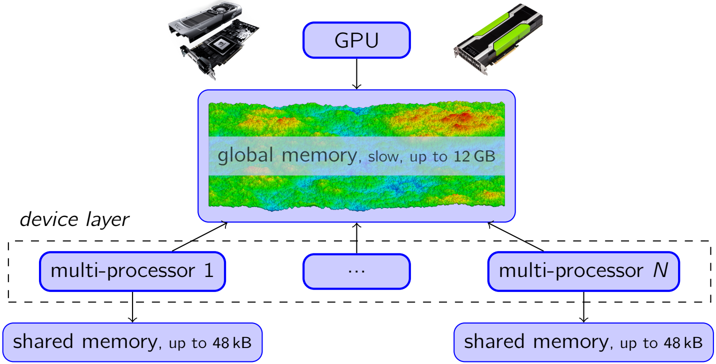
- kinetic model requires sequence of states without memory
- sequence of states $\Rightarrow$ inherently sequential
- apply domain decomposition:
  perform statistically independent updates at the same time
  - treat all sites equally: *mind borders*
  - updates should be uncorrelated: *borders*



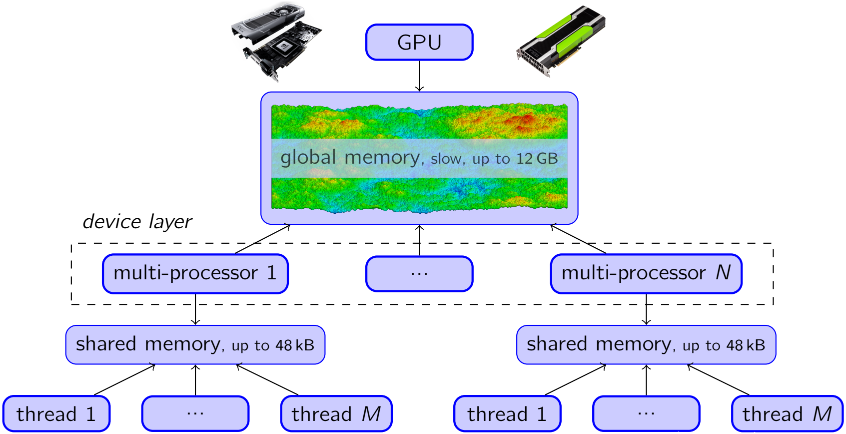
GPU Architecture
—for Stochastic Lattice Models



GPU Architecture
—for Stochastic Lattice Models

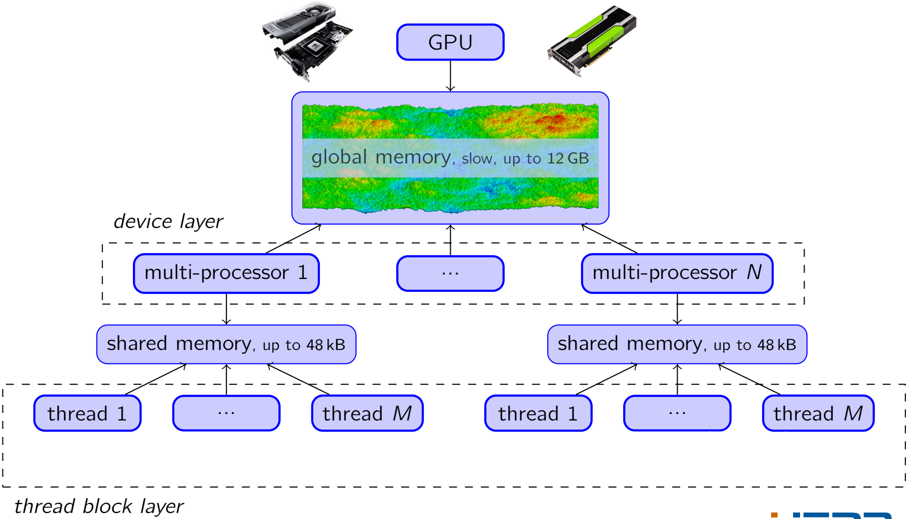


GPU Architecture
—for Stochastic Lattice Models

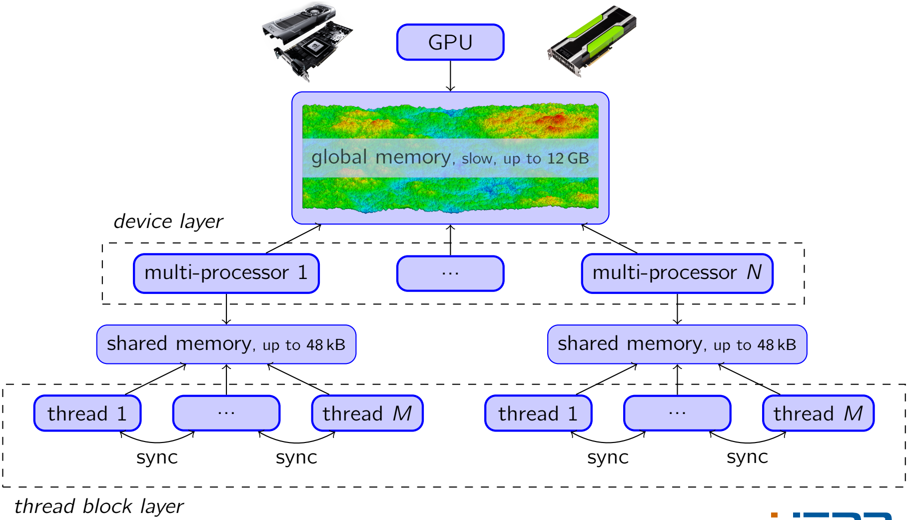


HZDR

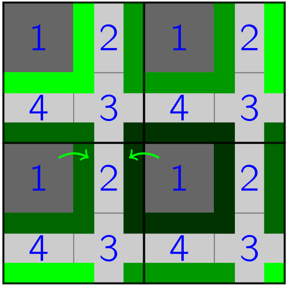
GPU Architecture
—for Stochastic Lattice Models

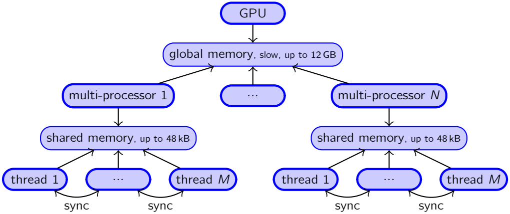


GPU Architecture
—for Stochastic Lattice Models

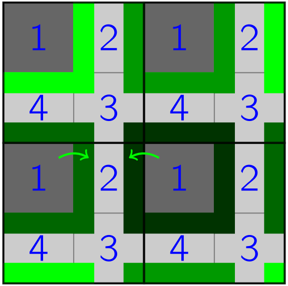


Domain Decomposition



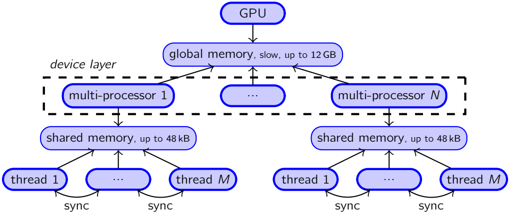


Domain Decomposition



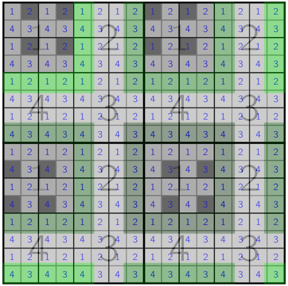
- at *device layer*
  - *example* double tiling decomposition
    random sequence of domain–sets
  - full update sweep

Error: subsystems with fixed boundary



Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

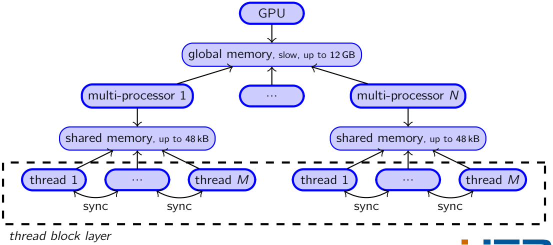
Domain Decomposition



- at *device layer*
  - *example* double tiling decomposition
    random sequence of domain–sets
  - full update sweep

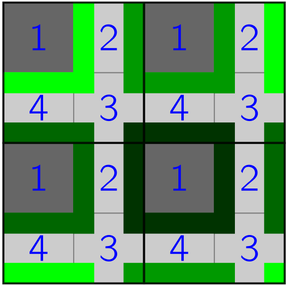
Error: subsystems with fixed boundary

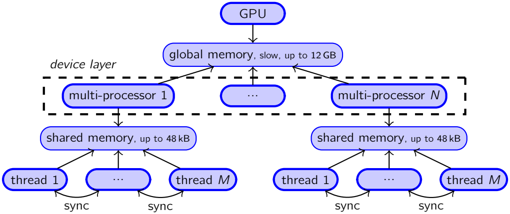
- at *work-group layer*
  - double tiling decomposition
    collective single–hit updates
  - random selection domain–set



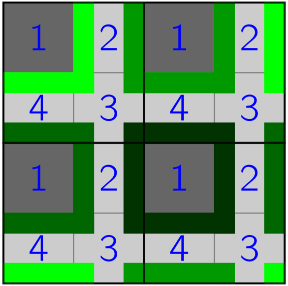
GPU Implementation

- at *device layer*
  - call kernel
  - thread blocks correspons to domains

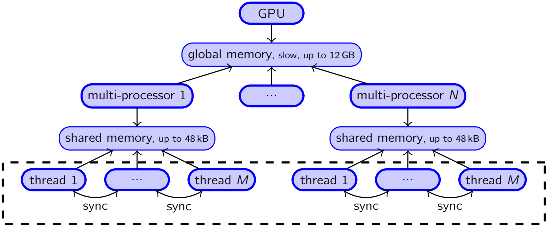




GPU Implementation

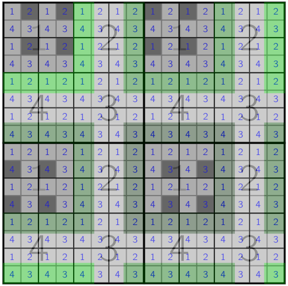


- at *device layer*
  - call kernel
  - thread blocks correspons to domains
- at *work-group layer*
  1. copy domain into shared mem.

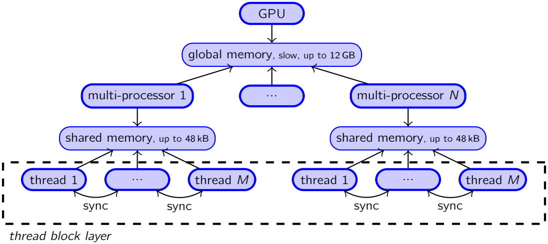


thread block layer

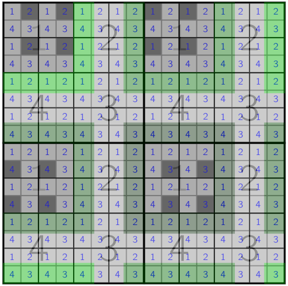
GPU Implementation



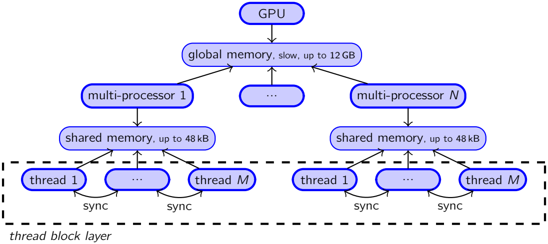
- at *device layer*
  - call kernel
  - thread blocks correspons to domains
- at *work-group layer*
  1. copy domain into shared mem.
  2. generate sequence vector, store to shared



GPU Implementation



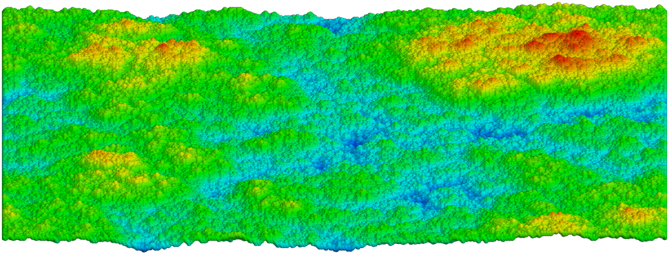
- at *device layer*
  - call kernel
  - thread blocks correspons to domains
- at *work-group layer*
  1. copy domain into shared mem.
  2. generate sequence vector, store to shared
     1. one Metropolis update per thread
     2. `sync`



Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Two Models: KMC and KPZ

KPZ–Equation for Surface Growth



KPZ surface in the steady state

$$\mathrm{d}_t h(\mathbf{x},t) = \underbrace{v}_{\text{mean growth vel.}} + \underbrace{\sigma_2 \nabla^2 h(\mathbf{x},t)}_{\text{surface tension}} + \underbrace{\lambda\left[\nabla h(\mathbf{x},t)\right]^2}_{\text{local growth vel.}} + \underbrace{\eta(\mathbf{x},t)}_{\text{noise}}$$

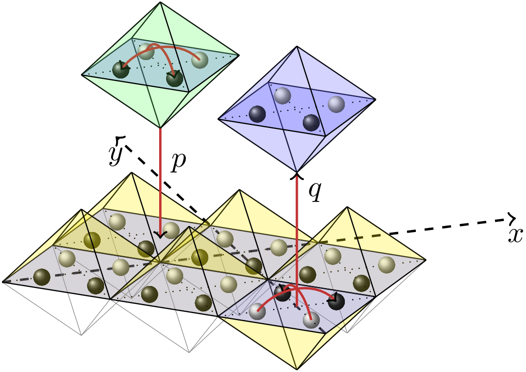
KPZ stochastic differential equation[2]

- growth processes, randomly stirred fluids, directed polymers in random media, dissipative transport, ...

[2] Kardar, M., Parisi, G., Zhang, Y.-C. *Phys. Rev. Lett.* **56** 889 (1986)

Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Model I—Roof-Top-Model for KPZ Growth

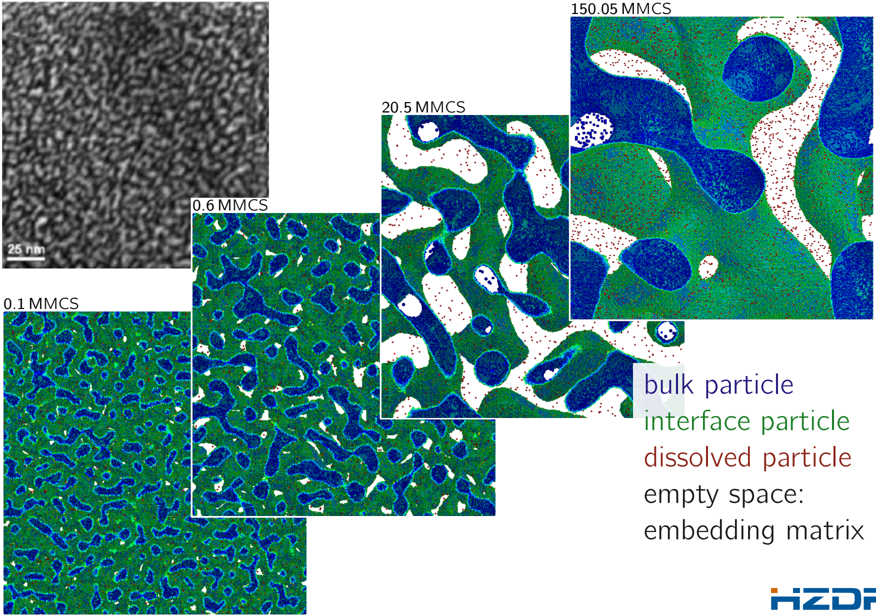


$2+1$D roof-top model—octahedron model[3]

- random depostion
  - $\Rightarrow$ site-selection *only* source of noise for $p = 1$
- \+ lattice gas with directed dimer diffusion

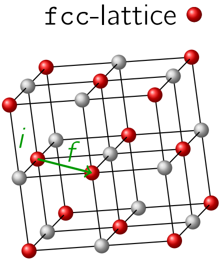
[3] Ódor, G., Liedke, B., Heinig, K.-H. *Phys. Rev. E* **79** 021125 (2009) (Plischke, M., Rácz, Z., Liu, D. *Phys. Rev. B* **35** 3485 (1987))

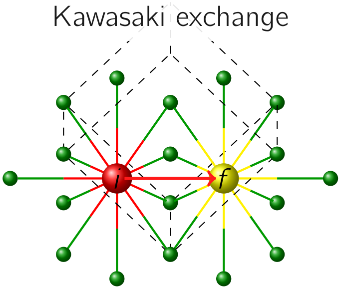
Simulation of Nano Structure Self-Organization



HZDR

Model II—Kinetic Metropolis Lattice Monte-Carlo





- 3D fcc-lattice
- random site selection not the only source of randomness

⇒ more robust against DD errors

Random Number Generation

What about Random Numbers?

KMC need to draw up to three random numbers per update-attempt: initial site, direction, whether to accept jump

KPZ need to draw up to two numbers: site, (whether to accept jump)

[4] Barash, L.Yu., Shchur, L.N. *Comp. Phys. Comm.* **185**(4) 1343 (2014)

What about Random Numbers?

KMC need to draw up to three random numbers per update-attempt: initial site, direction, whether to accept jump

KPZ need to draw up to two numbers: site, (whether to accept jump)

Mersenne Twister (PRNG[4])	KMC updates	KPZ updates
2.9×10^9	1.2×10^9	4×10^9

… per second on NVIDIA C2070

+ random numbers stored in global memory take time to load

[4] Barash, L.Yu., Shchur, L.N. *Comp. Phys. Comm.* **185**(4) 1343 (2014)

Inline Random Number Generators

64–bit LCG

$$x_i = (ax_{i-1} + c) \bmod m$$

+ 2 registers
− correlations, but still good enough
- skip ahead in sequence for parallel use

4

Inline Random Number Generators

64–bit LCG

$$x_i = (ax_{i-1} + c) \bmod m$$

+ 2 registers
– correlations, but still good enough
■ skip ahead in sequence for parallel use

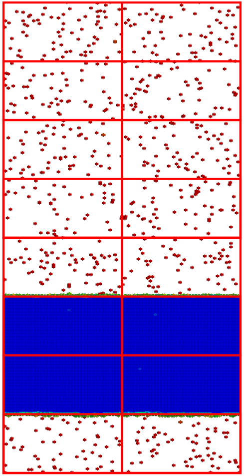
TinyMT[5]

+ good quality random numbers
– 4 registers for internal state + 4 regs for matrices
■ independent sequences for parallel use
■ usually more limited by shared memory

[5] http://www.math.sci.hiroshima-u.ac.jp/~%20m-mat/MT/TINYMT/

Errors due to Domain Decomposition

Broken Detailed Balance at Domain Boundaries

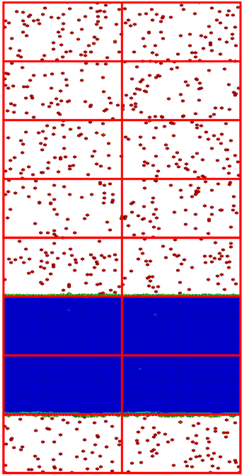


particle solubility in KMC

- device-layer domain boundaries break detailed balance:
  - move cannot be reversed at end of sweep
  - slower sweep at boundary

Member of the Helmholtz Association
Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Broken Detailed Balance at Domain Boundaries



particle solubility in KMC

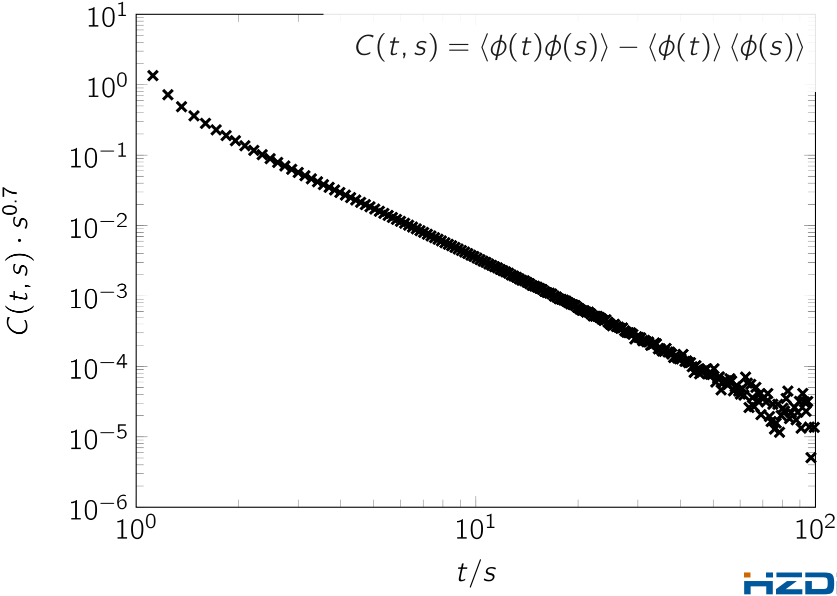
- device-layer domain boundaries break detailed balance:
  - move cannot be reversed at end of sweep
  - slower sweep at boundary

$$c(\varepsilon) = c_0 \cdot \mathrm{e}^{-B\varepsilon}$$

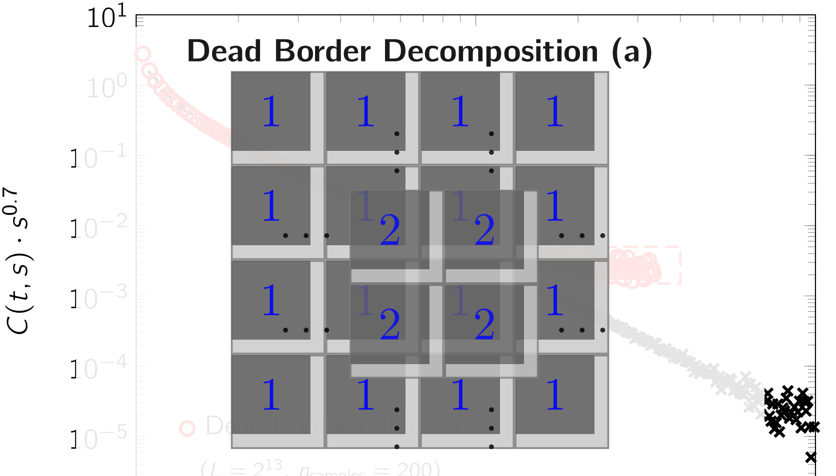
setup	c_0	B
aligned	1.08 ± 0.04	6.04 ± 0.03
random	1.00 ± 0.04	6.00 ± 0.03

$\Rightarrow$ solubility of particles at interface changed

Auto-Correlation of Slopes (Lattice Gas)



Auto-Correlation of Slopes (Lattice Gas)

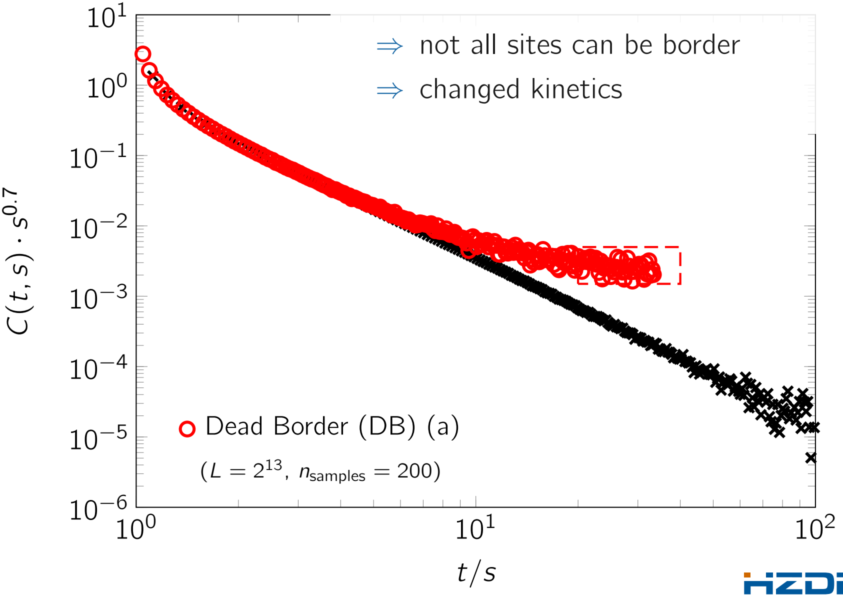


- shift decomposition-origin after each sweep
- ! restricted to 4×4 sites (word boundaries)

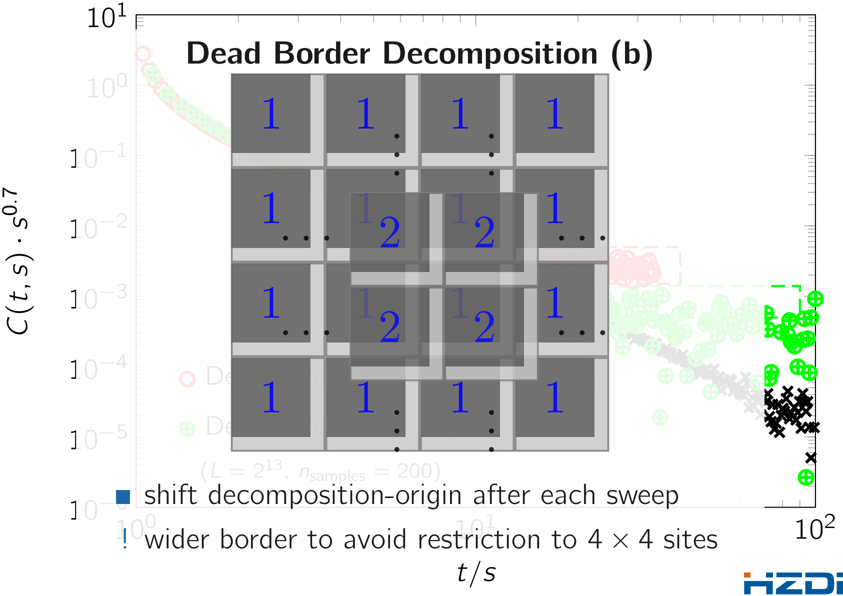
HZDR

Member of the Helmholtz Association
Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http://www.hzdr.de

Auto-Correlation of Slopes (Lattice Gas)



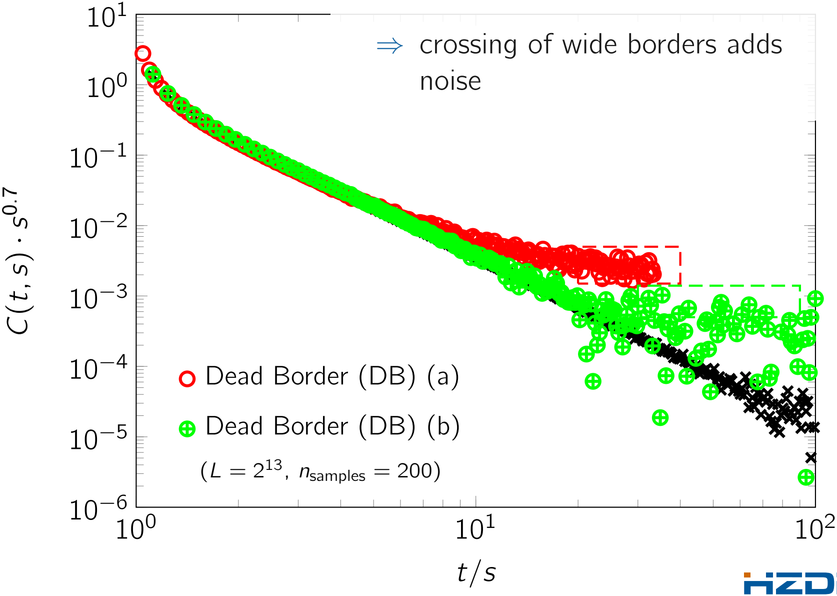
Auto-Correlation of Slopes (Lattice Gas)



- shift decomposition-origin after each sweep
- ! wider border to avoid restriction to 4 × 4 sites

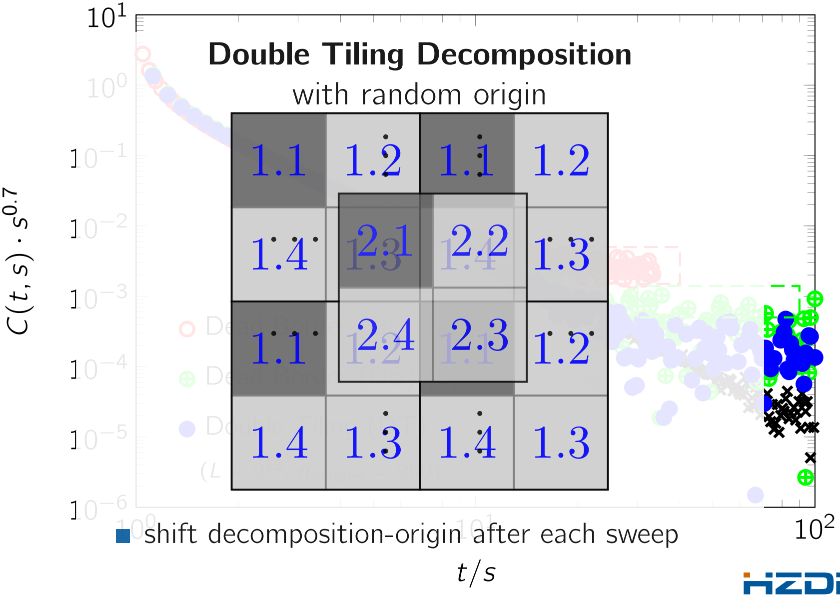
HZDR

Auto-Correlation of Slopes (Lattice Gas)



$\Rightarrow$ crossing of wide borders adds noise

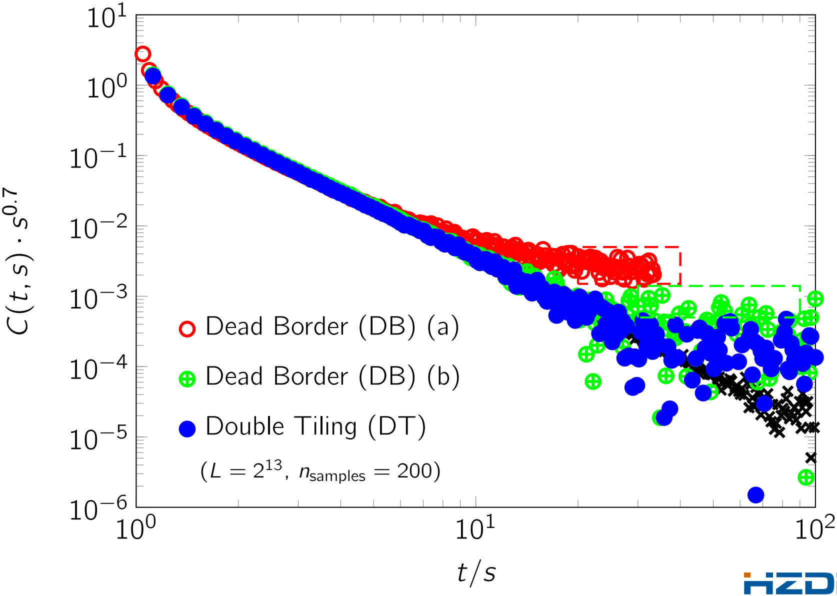
Auto-Correlation of Slopes (Lattice Gas)



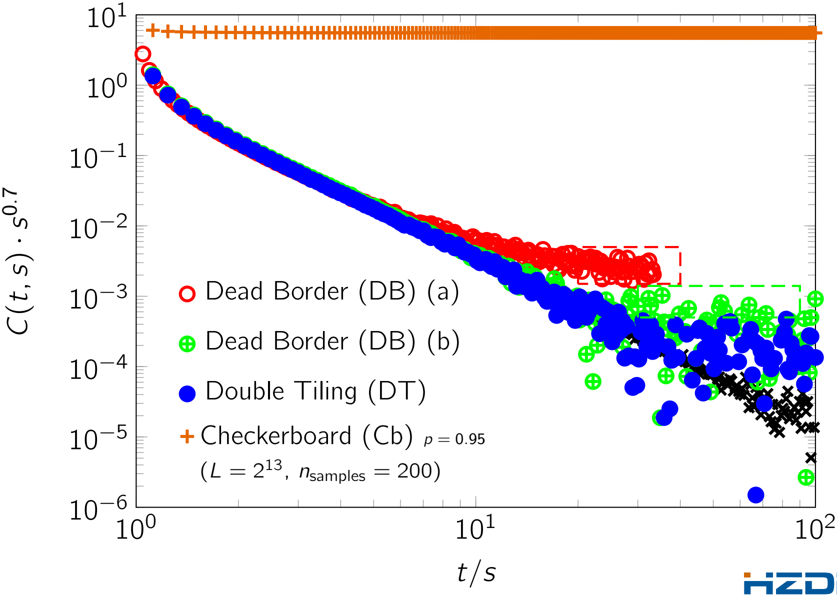
- shift decomposition-origin after each sweep

Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

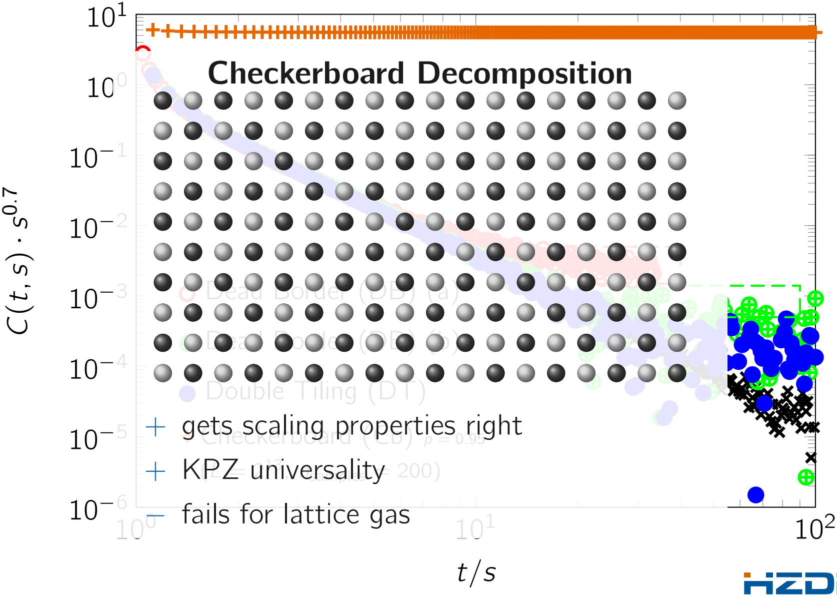
Auto-Correlation of Slopes (Lattice Gas)



Auto-Correlation of Slopes (Lattice Gas)

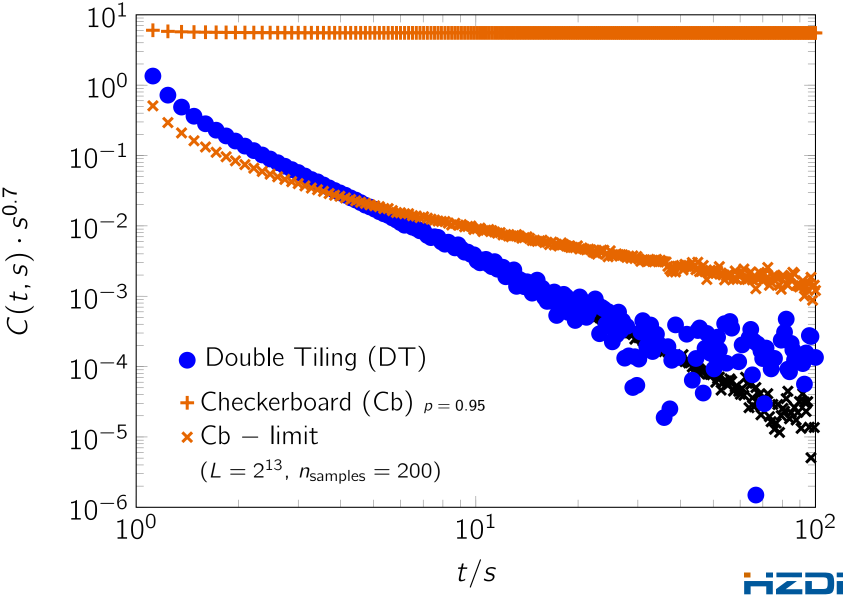


Auto-Correlation of Slopes (Lattice Gas)



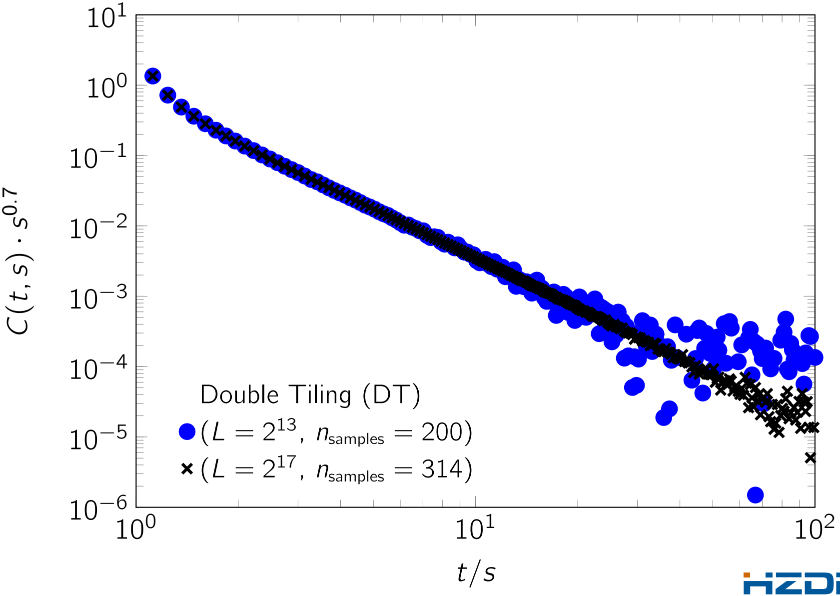
+ gets scaling properties right
+ KPZ universality
− fails for lattice gas

Auto-Correlation of Slopes (Lattice Gas)



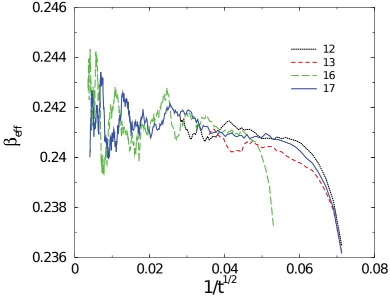
Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Auto-Correlation of Slopes (Lattice Gas)



Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

KPZ–Benchmarks



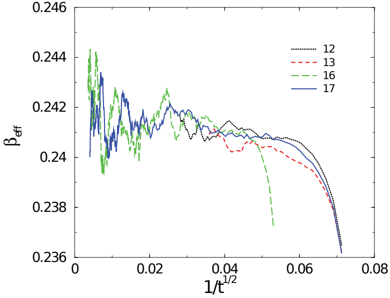
- $2^{17} \times 2^{17} \approx 16 \times 10^9$ lattice sites

$$W(t) \sim t^{\beta}$$

- prediced $\beta = 0.2415(15)$[6]
- excluding field theoretical value $\beta = 0.25$

[6] Kelling, J., Ódor, G. *Phys. Rev. E* **84** 061150 (2011)

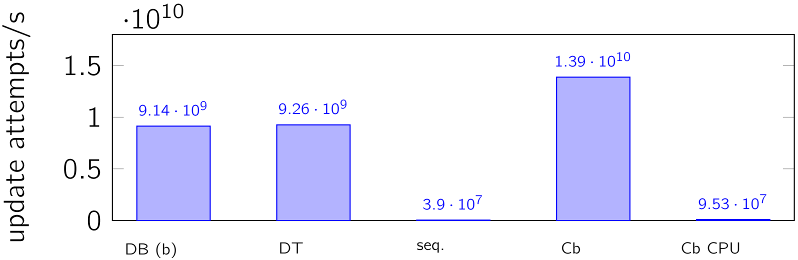
KPZ–Benchmarks



- $2^{17} \times 2^{17} \approx 16 \times 10^9$ lattice sites

$$W(t) \sim t^{\beta}$$

- prediced $\beta = 0.2415(15)$[6]
- excluding field theoretical value $\beta = 0.25$



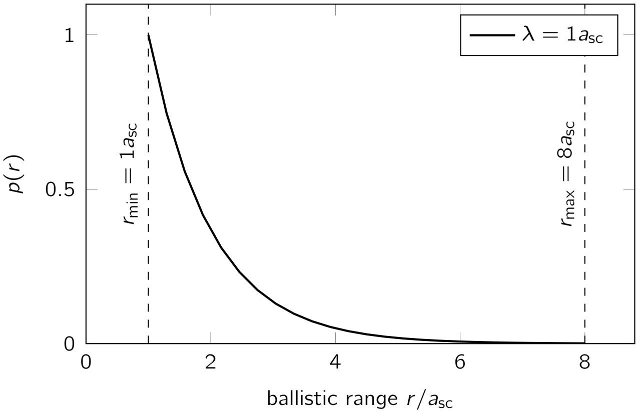
[6] Kelling, J., Ódor, G. *Phys. Rev. E* **84** 061150 (2011)

Long–Range Interaction in KMC

Recoil Mixing: Model

- incident ion hits lattice atoms creating a cascade of displacements
- energy gets divided among displaced atoms ⇒ exponentially distributed range

$$p(r) \sim e^{-r/\lambda}$$



Recoil Mixing: Cellular Automaton

rule ballistic update instead of thermal update with probability

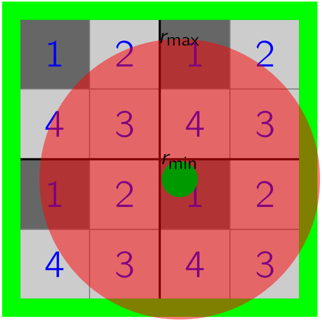
$$\varrho < 1 \times 10^{-2}$$

Recoil Mixing: Cellular Automaton

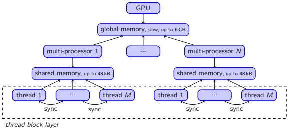
rule ballistic update instead of thermal update with probability

$$\varrho < 1 \times 10^{-2}$$

- maximum range $r_{\text{max}} = 8a_{\text{ac}}$
- rare events

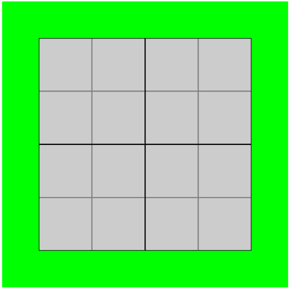


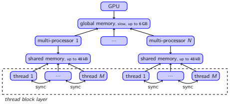
work-group layer, to scale



Recoil Mixing: GPU Implementation

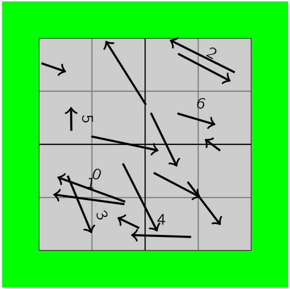
- perform collective ballistic updates
  separate from collective thermal updates
- allocate wider border



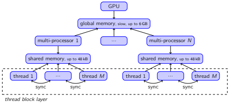


Recoil Mixing: GPU Implementation

- perform collective ballistic updates
  separate from collective thermal updates
- allocate wider border

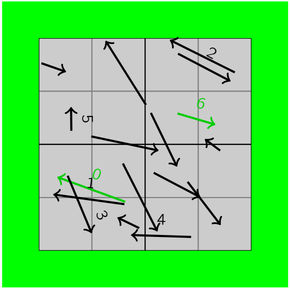


1. generate candidate jumps

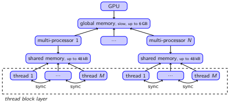


Recoil Mixing: GPU Implementation

- perform collective ballistic updates
  separate from collective thermal updates
- allocate wider border

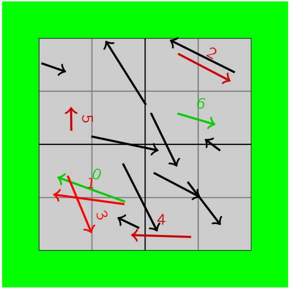


1. generate candidate jumps
2. choose desired number of jumps
   (Poisson distribution)

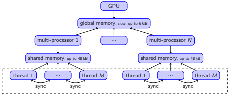


Recoil Mixing: GPU Implementation

- perform collective ballistic updates
  separate from collective thermal updates
- allocate wider border



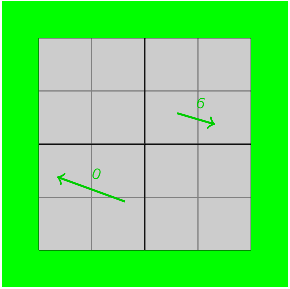
1. generate candidate jumps
2. choose desired number of jumps (Poisson distribution)
3. filtering invalid and colliding jumps (linear reduce, candidates $\gg n$)

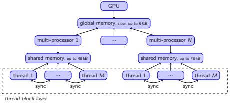


Member of the Helmholtz Association
Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Recoil Mixing: GPU Implementation

- perform collective ballistic updates
  separate from collective thermal updates
- allocate wider border

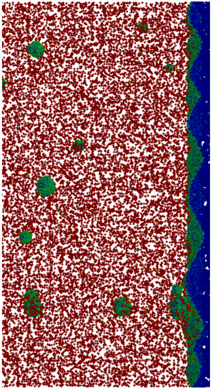




1. generate candidate jumps
2. choose desired number of jumps
   (Poisson distribution)
3. filtering invalid and colliding jumps
   (linear reduce, candidates $\gg n$)
4. carry out chosen jumps

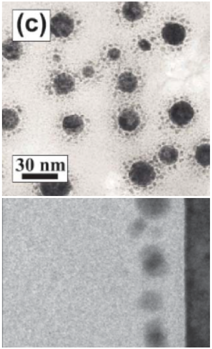
Applications I

Inverse Ostwald Ripening



Simulation of a (100)–interface

$\varrho = 1 \times 10^{-3}$

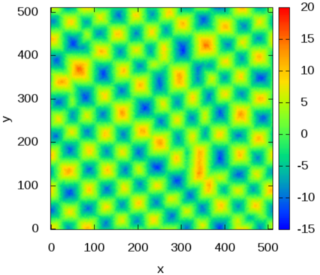


TEM pictures of Au nano clusters in SiO_2 matrix after irradiation with Au ions.

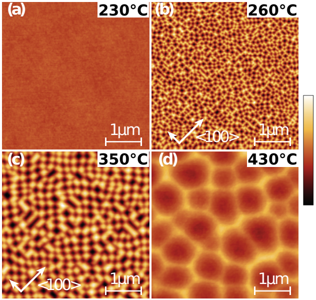
Heinig et al. Appl. Phys. A **77**, *17 (2003)*

Applications II

Instability of buried Interfaces or Surfaces



Simulation of a (100)–interface
$\varrho = 1 \times 10^{-3}$



AFM picture of a Ge (100)–surface after irradiation with Ge ions.
Ou et al. (2013) submitted (arXiv 1303.5133)

Acknowledgements

- Henrik Schulz
- Nils Schmeißer
- Michael Bussmann
- Nagy-Egri Máté Ferenc
- Martin Weigel
- my other colleagues

NVIDIA GPU CENTER OF EXCELLENCE

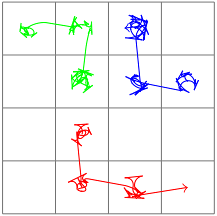
J.Kelling@HZDR.de

Thank You.

Publications

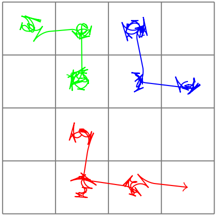
- Kelling, J., Ódor, G.:
  Extremely large-scale simulation of a Kardar-Parisi-Zhang model using graphics cards
  Phys. Rev. E **84** 061150 (2011)
- Kelling, J., Ódor, G., Nagy, M. F., Schulz, H., Heinig, K.-H.:
  Comparison of different parallel implementations of the 2+1-dimensional KPZ model and the 3-dimensional KMC model
  Eur. Phys. J. ST **210** 175 (2012)
- Ódor, G., Kelling, J., Gemming, S.:
  Aging of the (2+1)-dimensional Kardar-Parisi-Zhang model
  Phys. Rev. E **89** 032146 (2014)

Molecular Dynamics



movement of impurity atoms

Molecular Dynamics



movement of impurity atoms

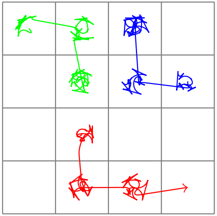
solve Hamilton's equations

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \text{ and } \dot{q}_i = \frac{\partial H}{\partial p_i}$$

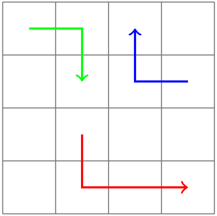
⇒ "nature's random number generator"

Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Molecular Dynamics → Cellular Automaton



movement of impurity atoms



on–lattice movement of impurity atoms

solve Hamilton's equations

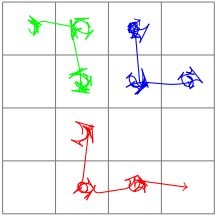
$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \text{ and } \dot{q}_i = \frac{\partial H}{\partial p_i}$$

⇒ "nature's random number generator"

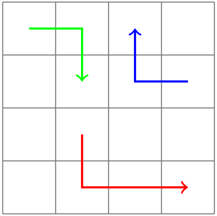
⇒ just use simple RNG

Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Molecular Dynamics → Cellular Automaton



movement of impurity atoms



on–lattice movement of impurity atoms

solve Hamilton's equations

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \text{ and } \dot{q}_i = \frac{\partial H}{\partial p_i}$$

⇒ "nature's random number generator"

⇒ just use simple RNG

- we know how, from *thermodynamics*

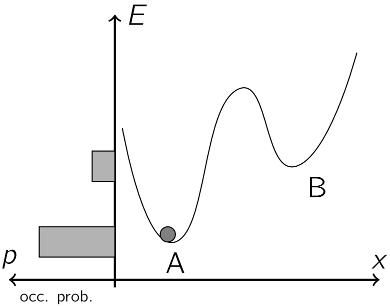
⇒ stochastic cellular automaton based on Metropolis algorithm

Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Metropolis Algorithm

Importance sampling: generate Markov–Chain of most probable states

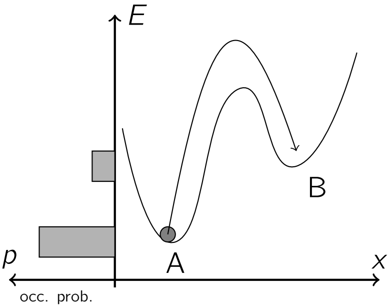
- starting at an initial state ν
- generate ν' through a minimal random modification to ν



Metropolis Algorithm

Importance sampling: generate Markov–Chain of most probable states

- starting at an initial state ν
- generate ν' through a minimal random modification to ν
- calculate the energy E_k needed for a transition $\nu \mapsto \nu'$

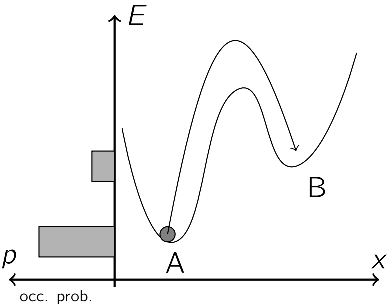


Metropolis Algorithm

Importance sampling: generate Markov–Chain of most probable states

- starting at an initial state ν
- generate ν' through a minimal random modification to ν
- calculate the energy E_k needed for a transition $\nu \mapsto \nu'$
- accept ν' as next state with probability

$$W = \Gamma_0 \cdot \min\left(\mathrm{e}^{-\beta E_k}, 1\right)$$



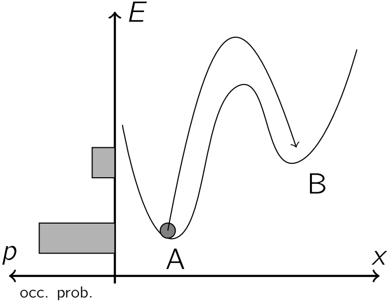
Member of the Helmholtz Association
Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Metropolis Algorithm

Importance sampling: generate Markov–Chain of most probable states

- starting at an initial state ν
- generate ν' through a minimal random modification to ν
- calculate the energy E_k needed for a transition $\nu \mapsto \nu'$
- accept ν' as next state with probability

$$W = \Gamma_0 \cdot \min\left(e^{-\beta E_k}, 1\right)$$



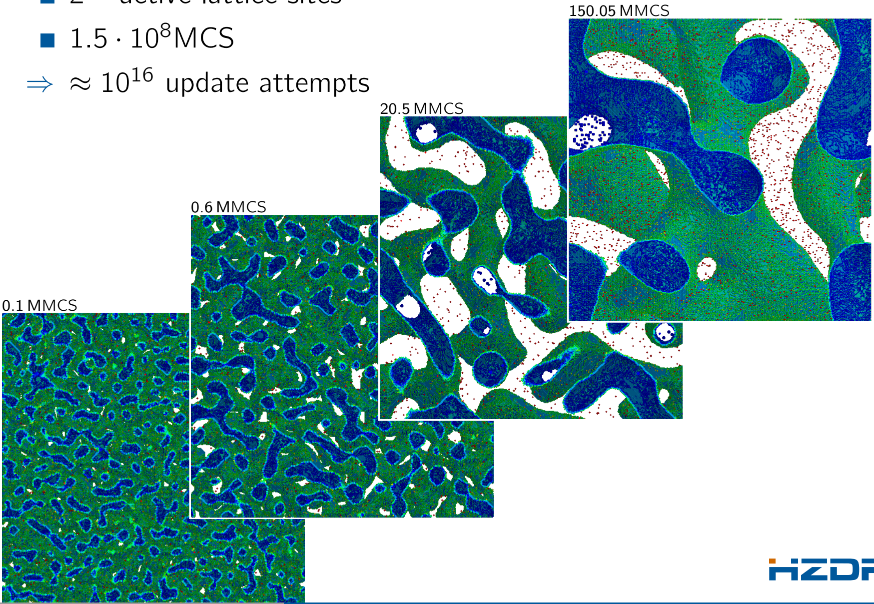
kinetics vs. equilibrium:

- nearest neighbor jumps
- random site-selection...

Large–Scale Simulation of Spinodal Coarsening

- 2^{26} active lattice sites
- $1.5 \cdot 10^{8}$MCS

$\Rightarrow \approx 10^{16}$ update attempts



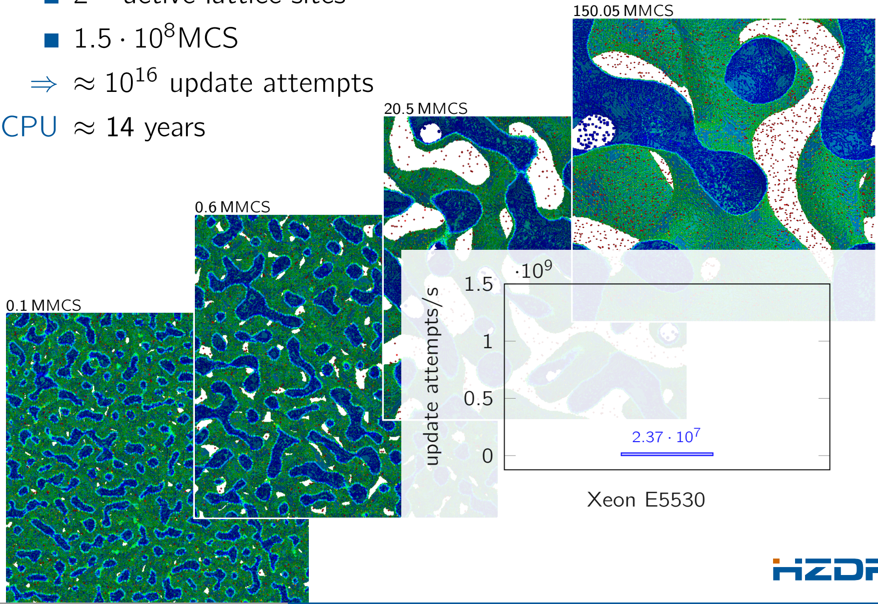
Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

Large–Scale Simulation of Spinodal Coarsening

- 2^{26} active lattice sites
- $1.5 \cdot 10^{8}$MCS

$\Rightarrow$ $\approx 10^{16}$ update attempts

CPU $\approx$ 14 years



Jeffrey Kelling, Géza Ódor, Karl-Heinz Heinig, Sibylle Gemming | FWIO | http//www.hzdr.de

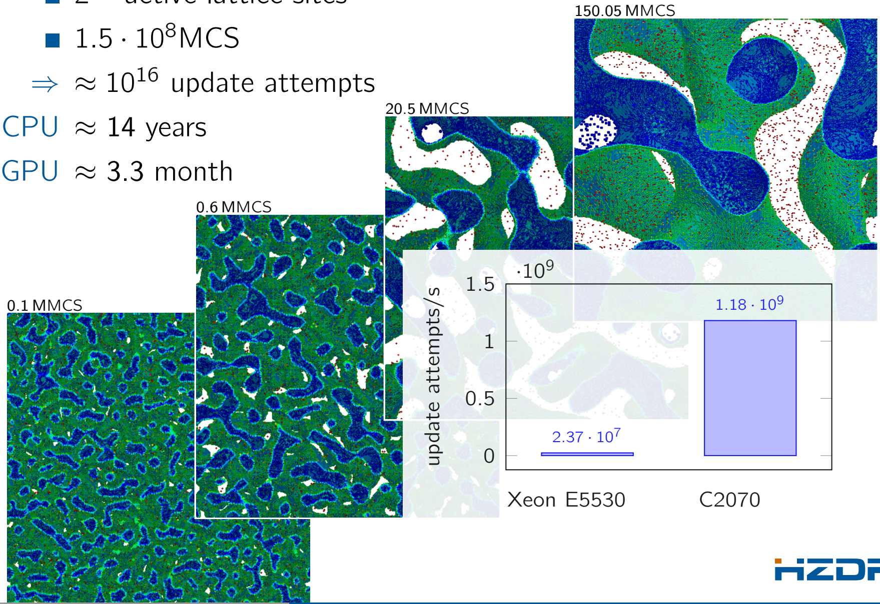
Large–Scale Simulation of Spinodal Coarsening

- 2^{26} active lattice sites
- $1.5 \cdot 10^{8}$MCS

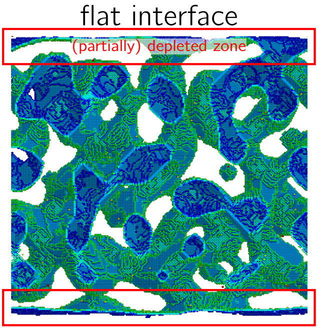
$\Rightarrow \approx 10^{16}$ update attempts

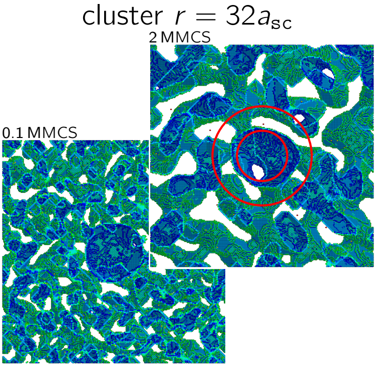
CPU $\approx$ 14 years

GPU $\approx$ 3.3 month



Coarsening in the Presence of Inhomogeneities





$$\varepsilon = 2.0,\ c = 0.325,\ L = 2^8$$

HZDR

Image Sources

1. Müller, T., Heinig, K.-H. et al. *Appl. Phys. Lett.* **85** 2373 (2004) *As referenced in RainbowEnergy project.*
2. http://en.wikipedia.org/wiki/File:Rub_al_Khali_002.JPG
3. https://www.hzdr.de/db/Cms?pOid=24344&pNid=2707
4. http://hubblesite.org/newscenter/archive/releases/2007/17/image/a
5. Ou X., Keller A., Helm M., Fassbender J., Facsko S. *Phys. Rev. Lett.* **111** 016101 (2013)
6. Tseng, Y.-C., Darling, S.B. *Polymers* **2** 470 (2010)
7. Fernando Tomás from Zaragoza, Spain
8. https://www.hzdr.de/db/Cms?pOid=24344&pNid=2707
9. Teshome, B., Facsko, S., Keller, A. *Nanoscale* **6** 1790 (2014)
10. Krause, M., Buljan, M. et al. *Phys. Rev. B* **89** 085418 (2014)