

$\beta_{m^{k+1}}^{k+1}$   $\beta_{m^{k-k+1}}^{k-k+1}$

**MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP**

**Microelectronics Design**

**TECHNISCHE UNIVERSITÄT KAISERSLAUTERN**

本日の動き  
日経平均株価 9661.79 -494.11

12カ月の動き  
日経平均株価 9661.79 -494.11

12カ月の動き  
TOPIX 931.50 -46.11

新発10年国債 -0.025

**Hardware Accelerators for Financial Mathematics**

Norbert Wehn  
University Kaiserslautern

**(CM)<sup>2</sup>**  
Center for Mathematical + Computational Modelling

$\beta_{m^{k+1}}^{k+1}$   $\beta_{m^{k-k+1}}^{k-k+1}$

**MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP**

**Microelectronics Design**

**TECHNISCHE UNIVERSITÄT KAISERSLAUTERN**

**The Computational Problem in Financial Mathematics**

Modern (electronic) financial markets are as vivid as never before

- Calculation of market prices for financial products like options/derivatives
- Solid mathematical background based on stochastic differential equations
- Asset prices change within several microseconds today
- Accurate and fast calculations are key

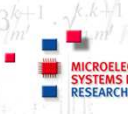
Deutsche Bank announcement March 2011


“Whereas most low latency trading solutions are built in software, Deutsche Bank’s solution is based upon a field programmable gate array (FPGA)...”

“...this is a bit of a revolution, since it’s breaking a barrier...we’re getting into the low-single digit microseconds. That has never done before”

Automated Trader Journal: “FPGA Acceleration of European Options Pricing”


**But they also have an energy problem!**

$\beta_{m^k+1}^{k-k+1}$ 

**Microelectronic sign**



**MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP**


## The Energy Problem in Financial Mathematics

**Canary Wharf: 250MW**  
**90% growth in Energy**  
**over next 5-7 Years**  
(The Guardian.co.uk)




- Olympic games in London 2012 with estimated 400MW power consumptions
- Olympic games get priority for power infrastructure
- Financial companies get shortages in access to power grid
  - ⇒ Non-essential computing functions such as backup are being moved away from trading center to save power

$\beta_{m^k+1}^{k-k+1}$ 

**Microelectronic sign**


**MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP**

## Financial Mathematics and MPSoC



- High performance computing under power constraints
- Banks start to use hardware accelerators
  - GPUs are already in use
  - FPGAs are currently evolving (MATLAB/HDL generator)
- Stochastic systems are known in MPSoC
  - E.g. communication systems
- Design methods already adapted concepts from financial mathematics
  - Modeling of dependencies: copula concept, upper tail dependency
  - Bounds for distribution functions: Comonotonicity
- Advanced design methodologies for various implementation styles for HW accelerators: FPGA, ASIPs, IP....
- Metrics and benchmarks

$\beta_{m^{k+1}}$   $\gamma_{m^{k+1}}$  **MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP** **Pricing for (European) Call Options with Barriers**

Financial contract between buyer and seller  
 Buyer has option to buy e.g. a share from the seller at a certain time for a certain price (strike )

**Arbitrage in finance: possibility of a risk-free profit at zero cost**  
**Risk neutral price calculation for derivatives: no arbitrage possibility**

Source: publickitt / www.flickr.com

$\beta_{m^{k+1}}$   $\gamma_{m^{k+1}}$  **MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP** **Black-Scholes Model (Nobel Prize 1997)**

Most option price calculation are based on Black Scholes Model


- Stochastic Differential Equations based on Brownian Motion
- Asset/share price not differentiable  $\Rightarrow$  integral formulation

$$dS_t = \mu S_t dt + V_t dW_t$$

$S_t$ : price of the asset/share  
 $\mu$ : expected rate of the return of the asset  
 $V_t$ : volatility  
 $W_t$ : Brownian motion

...in many cases lack of analytical solutions  
 $\Rightarrow$  Monte Carlo simulations

- precision  $\sim$  #paths<sup>2</sup>, precision  $\epsilon = \frac{1}{N} \Rightarrow O(N^3)$
- x millions of path simulations necessary

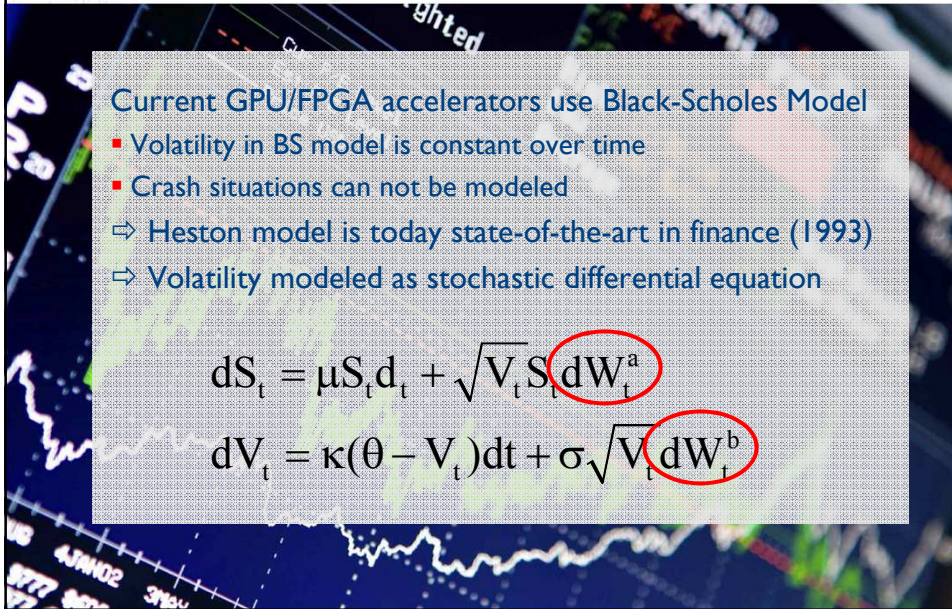

**Heston's Model of Stochastic Volatility**


Current GPU/FPGA accelerators use Black-Scholes Model

- Volatility in BS model is constant over time
- Crash situations can not be modeled

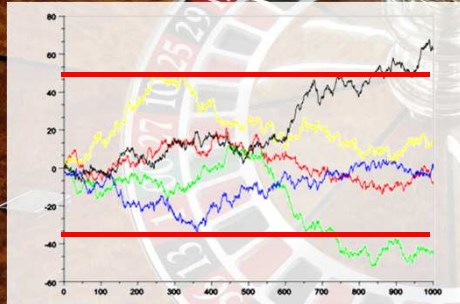
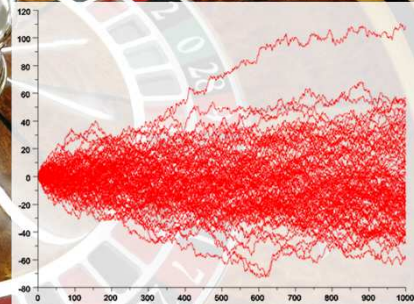
⇒ Heston model is today state-of-the-art in finance (1993)  
 ⇒ Volatility modeled as stochastic differential equation

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_t^a$$

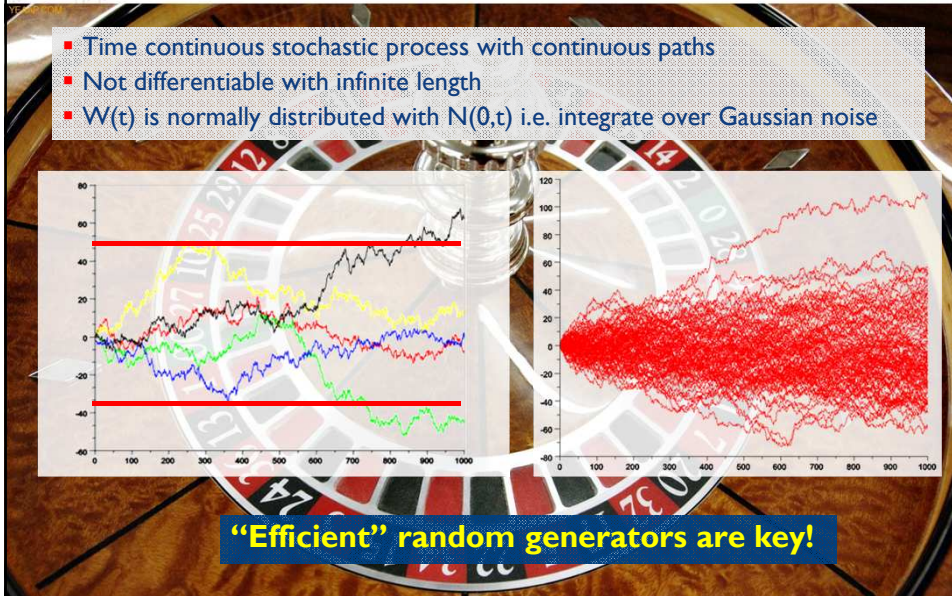
$$dV_t = \kappa(\theta - V_t) dt + \sigma \sqrt{V_t} dW_t^b$$



**Brownian Motion (Particle Theory)**

- Time continuous stochastic process with continuous paths
- Not differentiable with infinite length
- $W(t)$  is normally distributed with  $N(0,t)$  i.e. integrate over Gaussian noise

**"Efficient" random generators are key!**



**New Non-Uniform Random Number Generator**

Mersenne Twister, Box-Muller: high complexity

	Virtex-5 FX70T-3	Virtex-4 LX100-12		Cheung et al. 2007	
	converter + ICDF lookup	converter	ICDF lookup	address decoding	ICDF lookup
Slices	44	55	180	346	125
Flip-flops	62	62	180	346	125
LUTs	40	55	180	346	125
DSP	1 (DSP48E)	0	2 (DSP48)	0	2 (DSP48)
BRAM	1 (36kb)	0	2 (18kb)	0	2 (18kb)
$F_{max}$ / MHz	381	> 400	< 375	< 375	< 375

Parameters optimized for Xilinx Virtex-5 DSP48E

**What about quality?**

180 instead of 346 slices  
 ⇒ 48% less area @ higher precision  
 @ proven quality

**Importance of Quality Tests**

Uniformly distributed random numbers

Standard tests

Non-uniform random numbers

No standard tests!

**Various  $\chi^2$  (goodness-of-fit) tests**

- Bin generated values into groups
- Compare with theoretical values

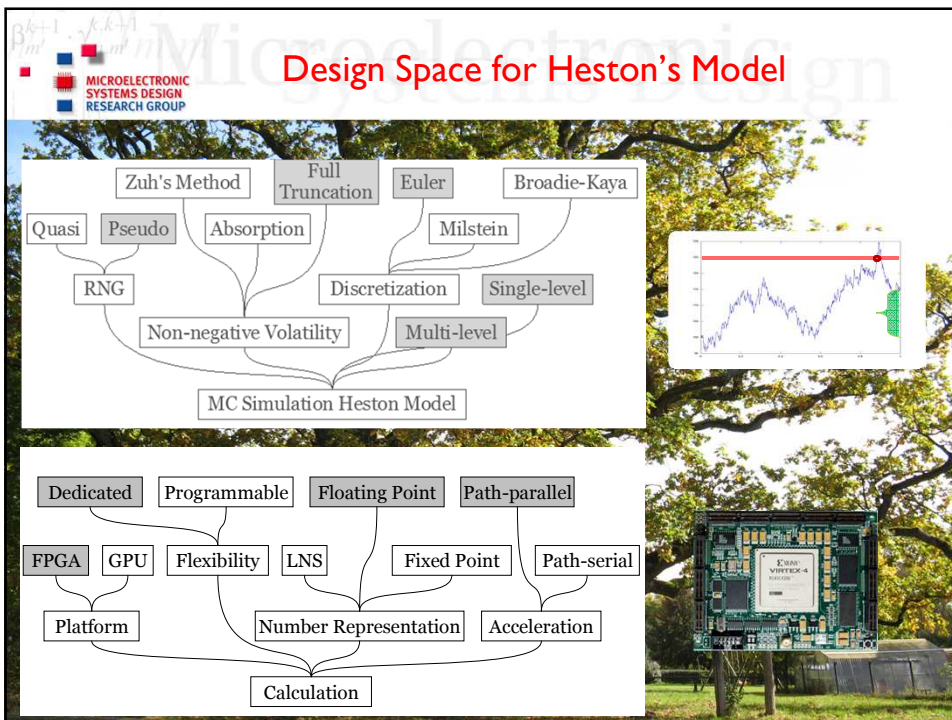
**Kolmogorov-Smirnov test**


- Compares generated values with theoretical cumulative distribution function

**Normal Random Number Generation**

Implementation	Architecture	Power	Throughput [Msamples / s]	Energy per sample
Fast-Mersenne Twister optimized for SIMD Box-Muller	Intel Core 2 Duo PC 2.0 GHz, 3 GB RAM one core only	~ 100 W	<b>600</b>	<b>166.67 pJ</b>
Nvidia Mersenne Twister + Box-Muller CUDA	Nvidia GeForce 9800 GT	~ 105 W	<b>1510</b>	<b>69.54 pJ</b>
Nvidia Mersenne Twister + Box-Muller OpenCL			<b>1463</b>	<b>71.77 pJ</b>
Inversion method for general non-uniform distributions, only one instance	Xilinx FPGA Virtex- 5FX70T3 @380 MHz	~ 1.3 W	<b>379</b>	<b>3.43 pJ</b>
Inversion method for general non-uniform distributions, 100 instances		~ 1.8 W	<b>37900</b>	<b>0.05 pJ</b>

Source: ~H. Crimson / www.flickr.com




Benchmarks

**Fair comparison**


- Benchmarks and metrics
- Not known in financial hardware accelerator community

Set up a benchmark set for Heston's model in cooperation with financial mathematician

- Comprehensive pricing tasks with real world scenarios
- In total 12 different call options: European call, single/double barrier...

Download on [www.uni-kl.de/benchmarking](http://www.uni-kl.de/benchmarking)  
(includes Octave model)

Source: Mike Johnson - TheBusyBrain.com


First Results

**Multi-level Monte Carlo** accelerator of Heston model (**full pricing**)

Precision  $\varepsilon = \frac{1}{N} \Rightarrow MC O(N^3)$ , Multilevel MC up to  $O(N^2)$

	<b>Nvidia GPU Tesla 2050 (Opt. Open-CL)</b>	<b>Intel Xeon CPU 1 core@3.07 GHz (Matlab)</b>
Real-time / s	0.332	22.41
Energy per Simulation / J	96.28	3742.5
Number Format	Floating Point Double Precision	Floating Point Double Precision

Multi-level simulation up to level 4 with antithetic variance reduction, continuity correction, precision of 0.01 and a total of 4,731,944 paths

Best of our knowledge first Multi-level MC implementation of Heston Model on CPU and GPU

$\beta_{m^k+1}^{k-k+1}$   $\gamma_{m^k+1}^{k-k+1}$  **MicroelectronicSign**  
**MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP**

## FPGA Acceleration: Heston Model

Multi-level Monte Carlo Heston model (**only path acceleration**)

	Xilinx Virtex-5* FPGA@100MHz	Nvidia GPU Tesla 2050	Intel Xeon CPU 1 core@3.07 GHz
Real-time / s	<b>15.37</b>	<b>22.269</b>	1600
Energy per Path / mJ	<b>0.023</b>	<b>1.076</b>	44.55
Number Format	Fixed Point 32 bits	Floating Point Double Precision	Floating Point Double Precision

Fixed to a level 5 simulation (1024 fine and 256 coarse steps in parallel) with antithetic variance reduction on 6,000,640 paths  
 \*preliminary results, FPGA capacity allows 5 instances

**Best of our knowledge first Heston model implementation on FPGA**

$\beta_{m^k+1}^{k-k+1}$   $\gamma_{m^k+1}^{k-k+1}$  **MicroelectronicSign**  
**MICROELECTRONIC SYSTEMS DESIGN RESEARCH GROUP**

## Fixed Point versus Float

Multi-level Monte Carlo Heston model (**only path simulation**)

	Fixed Point* 32 bits		Floating Point* Single Precision	
	Number	Percentage	Number	Percentage
LUTs	6645	14%	10291	22%
Flip-Flops	5783	12%	12075	26%
LUT-FF pairs	8729	19%	13397	29%
DSP48E	25	19%	4	3%
BRAM36	7	4%	1	1%
MHz	96		104	

\* preliminary results Virtex 5 – XC5VFX70T (on ML-507 evaluation board)





**Summary**

**Financial Mathematics exciting application for FPGA accelerators  
Can profit from MPSoC community  
Requires interdisciplinary teams**

**Many thanks to  
Christian de Schryver, Matthias Jung, Ivan Shcherbakov  
Henning Marxen, Anton Kostiuk, Elke Korn, Ralf Korn**

<http://ems.eit.uni-kl.de>

