Heston Benchmark v2.0 – Specification

# Motivation and Goals

Developing accelerators for financial computations is currently a hot research topic. Many publications are out presenting dedicated SIMD, GPU, FPGA, ASIP, or ASIC implementations aimed to increase either the speed or the energy efficiency of one or a set of tasks. However, due to the large variety of underlying problems, models, algorithms, implementations, and evaluation metrics, it is nearly impossible to fairly tell which implementation is “best” for a given scenario.

Evaluating and comparing different implementations is a challenge, especially if the target platforms differ. The point is that not all algorithms are suitable for all architectures, e.g. Monte Carlo methods might not converge as fast as quadrature or finite difference methods, but can easily be parallelized and therefore fit to highly parallel architectures like FPGAs or clusters.

At the same time, cloud computing is moving into the finance domain and more and more institutes are going to offload compute intensive applications to cloud services. It is hard to evaluate how “good” those service perform compared to optimized GPU or FPGA solutions that can be kept in house.

We therefore see a strong demand for a portable benchmark set that allows to evaluate the following performance criteria of a pricing system:

* the number of products priced per Joule
* the number of products priced per second
* the achieved numerical accuracy

More goals are:

* The benchmark should provide an executable implementation example that generates the correct reference results (we propose a Python implementation here). This reference implementation is freely available.
* The model parameters have to reflect meaningful and realistic scenarios on markets, but also cover corner cases as observed in the days just before a crash or another peculiar event.
* We need meaningful people, institutes, and companies involved that ensure the necessary credibility of this approach and support the use of the benchmark in the finance and implementation domains.
* The benchmark framework should be modular and easily extendable to more products and models.

# Challenges and Questions

## Scalability

The performance of product pricers can differ in a wide range, e.g. from a non-optimized single-threaded CPU program up to dedicated high-speed dataflow architectures or huge clusters. The benchmark needs to address the complete range, e.g. by providing several complex test batteries with different computational complexity.

## Metrics and Scoring

The easiest way for comparing the benchmark results of different systems is to introduce an artificial scoring system that gives a single number of points to characterize an implementation. However, this comes at the cost of transparency over the different results dimensions (speed vs. energy efficiency vs. accuracy). Another option is to provide a set of numbers instead of a single score that gives more information, but makes it again harder to compare different designs. It is mandatory to find the right balance here to ensure complexity handling and meaningful results at the same time.

## Application Area

A big question is which products should be targeted in the benchmark, since there is a large variety observed in the financial area. We propose to strictly distinguish between product and the employed model for the market simulation (like Heston, Black-Scholes, Jump-Diffusion, …), since the same product can be priced with different underlying models.

We propose to start with the showcase “European double barrier option pricing in the Heston model”, since preliminary work on this field has already been done by the University of Kaiserslautern, and Monte Carlo implementations for CPU, GPU, FPGA, and the Maxeler engine already exist.

## Model Input Parameters

It is crucial that the provided model parameters reflect realistic scenarios a) of day-to-day situations in different markets (e.g. liquid stock exchange, FOREX, …), but also critical corner cases that have shown to be hard to handle in the past. The data should therefore be legitimated carefully by one or more of the business partners (see below).

## Soft-Scoring and Non-Functional Data

In addition to the main scoring, the two following aspects shall be included in the benchmark reports, either in textual form or as a second soft-score:

* Design Effort  
  The time it takes to design or re-design a solver is a significant consideration as developers need to get working results quickly. Many applications are also frequently tweaked and modified.
* Portability  
  Related to the previous aspect. How difficult is it to re-deploy an existing solution onto a different device of the same architecture / a new device of a next generation architecture / a completely different architecture?

# Preliminary work

The University of Kaiserslautern has already released a first version of a Heston benchmark for European double barrier option pricing ([www.uni-kl.de/benchmarking](http://www.uni-kl.de/benchmarking)) together with 12 predefined Heston parameter sets taken from literature. However, this version has some weaknesses:

1. The required accuracy of the result is not specified, so the result accuracy can freely be determined and therefore influences runtime and energy consumption and so decreases the comparability.
2. There is no scoring for easy comparison.
3. The benchmark is not challenging enough for high-performance clusters or pricing engines.

# Partners and Contributions

Possible contributors:

* Imperial College
  + implementation competences
  + benchmark construction know-how
  + implementation complexity studies of different algorithms
* University of Kaiserslautern
  + multiple implementations of option pricers based on the Heston model
  + Heston competences
  + interdisciplinary cooperation with finance mathematicians and computational stochastic group
* Maxeler
  + Stephen Weston: high finance competences
* SWIP
  + calibrated Heston parameters from real market data
* Fraunhofer ITWM
  + own Heston implementations
  + meaningful market data
  + contacts to many finance institutes
* J.P. Morgan

# Document Revisions

|  |  |  |
| --- | --- | --- |
| Date | Who | Comment |
|  |  |  |
|  |  |  |
| 2013-08-15 | CdS | added portability and design effort aspects |
| 2013-06-12 | CdS, CB | first draft |