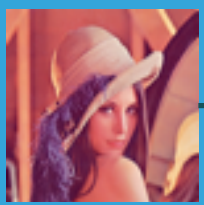
DETECTING SILENT DATA CORRUPTIONS WITH ERROR ESTIMATES

PIERRE-LOUIS GUHUR[1], TOM PETERKA[2], FRANCK CAPPELLO[2]
ENS de Cachan[1], Argonne National Laboratory[2]

SILENT DATA CORRUPTIONS

SDCs = corruptions that remain unnoticed



Only 5 bytes are deleted.

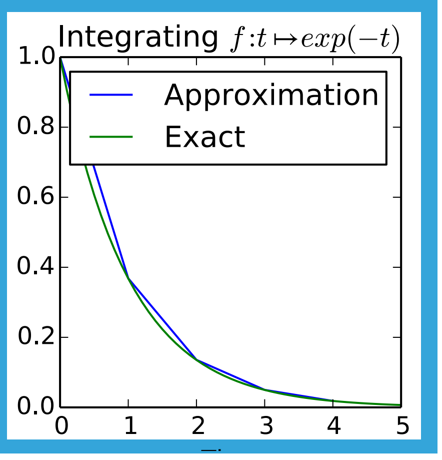
It is still readable though wrong

Origins:

- systematic SDCs: bugs, attacks…
- **nonsystematic**: interferences, aging, ionizing radiation…

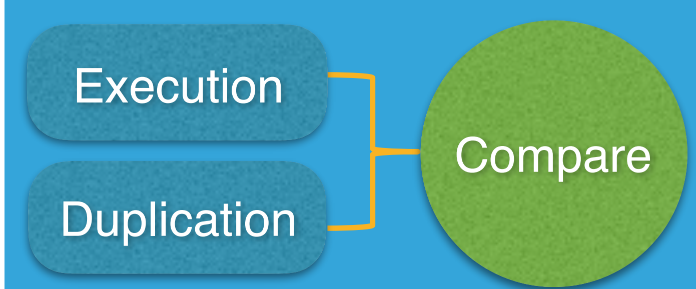
MOTIVATIONS

1. Improving trust of results
2. For next generation of supercomputers, the probability of SDCs is increasing , due to the increases of the number of transistors and of the complexity of power management, software and hardware.
3. **Numerical integration solvers** are particularly sensitive to SDCs: SDCs are propagated, and the solution can diverge.



STATE-OF-THE-ART SDC DETECTORS

Replication



It detects all nonsystematic SDCs, but is often too expensive in memory, in computational and in energy

AID

Di, S. and Cappello, F., "Adaptive-Impact Driven Detection of Silent Data Corruption for HPC Applications", IEEE Transactions on Parallel and Distributed Computing, 2016

BSS14

Benson, A. and Schmit, S. and Schreiber, R., "Silent Error Detection in Numerical Time-Stepping Schemes", IJHPCA (2015)

SOLVERS WITH FIXED STEP

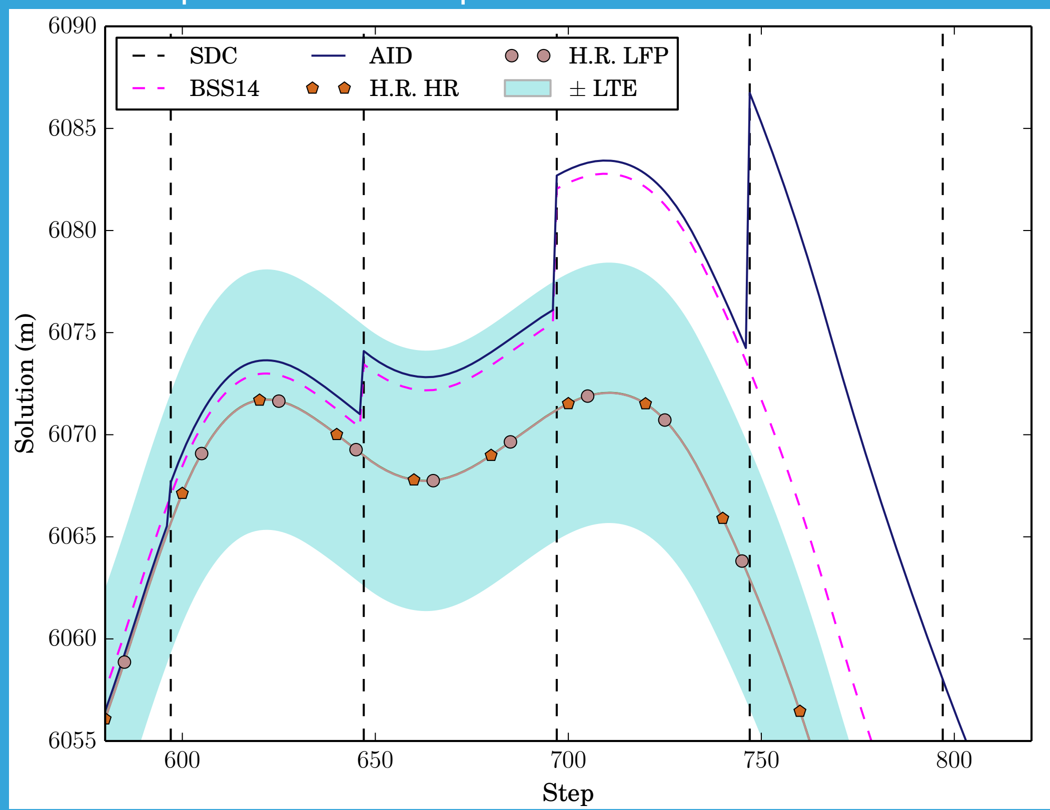
Hot Rod validates a step when 2 estimates of the error agree:

$$D(Est_1 - Est_2) < Threshold$$

Two thresholds are proposed: LFP and HR. LFP does few false positives. HR does few false negatives.

They detect SDCs even smaller than the approximation error of the solver (LTE)., because Est_1 - Est_2 is one order lower than LTE.

Detection performance outperforms the state of the art.



Guhur, Zhang, Peterka, Constantinescu, Cappello, "Lightweight and Accurate Silent Data Corruption Detection in Ordinary Differential Equation Solvers." Euro-Par 2016

SOLVERS WITH VARIABLE STEP

The step size is controlled from an estimate of the error. When the estimate is too high, the step is rejected, which allows to detect some SDCs (called classic controller)

Issue: the estimate may be under-evaluated in the presence of an SDC.

Solution: to double-check the validation of a step with a second estimate.

Method: a backward differentiation formula selected from a trade-off between true positives and false positives.

	Significant FNR (%)	FPR (%)	Computational overhead (%)
Classic controller	13.3	0.0	0.0
Method	1.1	4.2	4.5
Replication	0	100	100

FNR: false negative rate. FPR: false positive rate. The lower, the better