Fault Tolerance Interface

Leonardo Bautista Gomez
Argonne National Laboratory

HPC Knowledge Meeting'15


EUROPA
2 JEWELS
Time


Space


Energy


Information

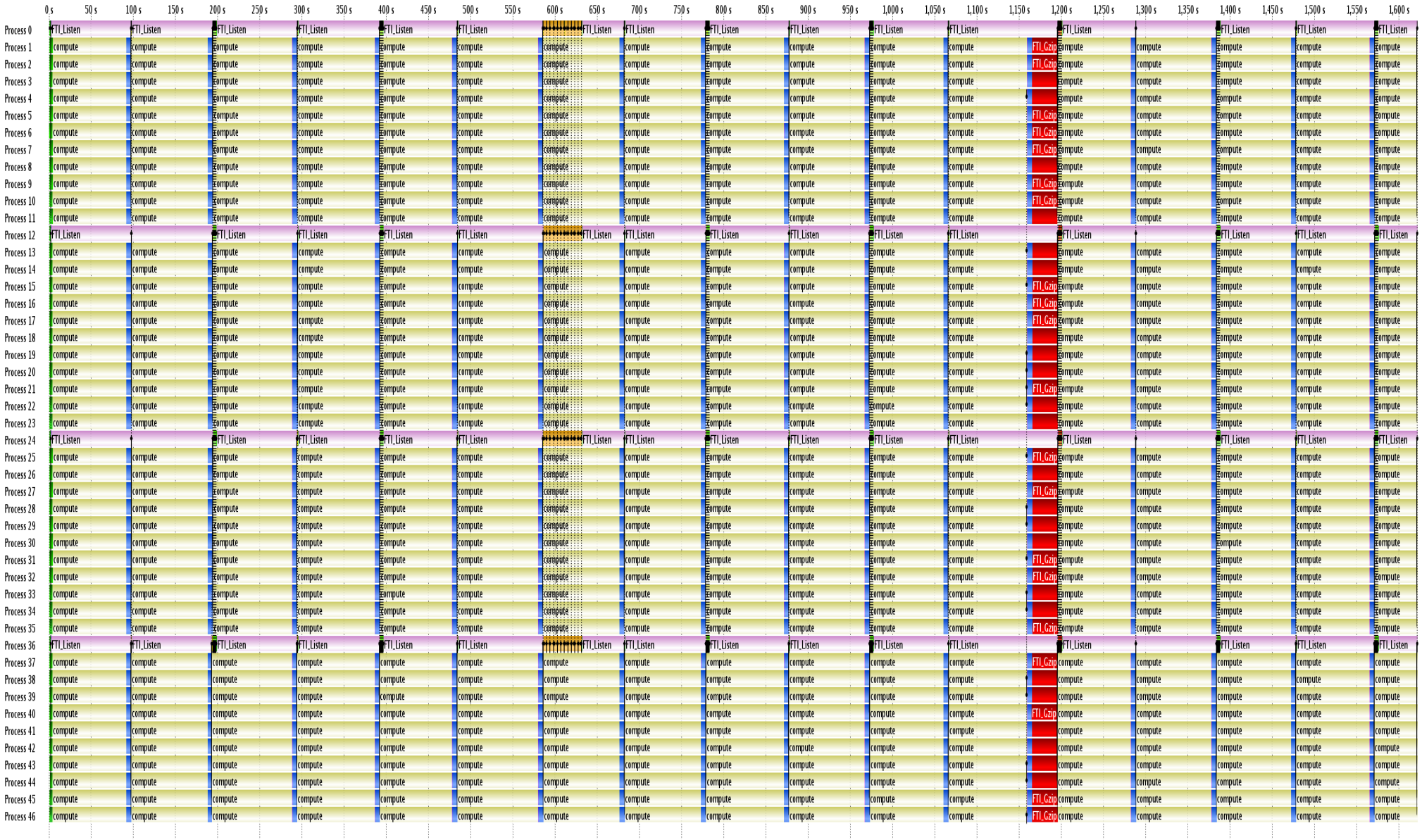
Time

Fault Tolerance Interface

- Application-level
- **Simple API:**
  - FTI_Init
  - FTI_Protect
  - FTI_Snapshot
  - FTI_Finalize

```
int main(int argc, char **argv) {

   MPI_Init(&argc, &argv);
   FTI_Init("conf.fti", MPI_COMM_WORLD);

   double *grid;
   int i, steps=500, size=10000;
   FTI_Protect(0, &i,   1,   FTI_INTG);
   FTI_Protect(1, grid, size,FTI_DFLT);

   for (i=0; i<steps; i++) {
      FTI_Snapshot();
      kernel1(grid);
      kernel2(grid);
   }

   FTI_Finalize();
   MPI_Finalize();
   return 0;
}
```

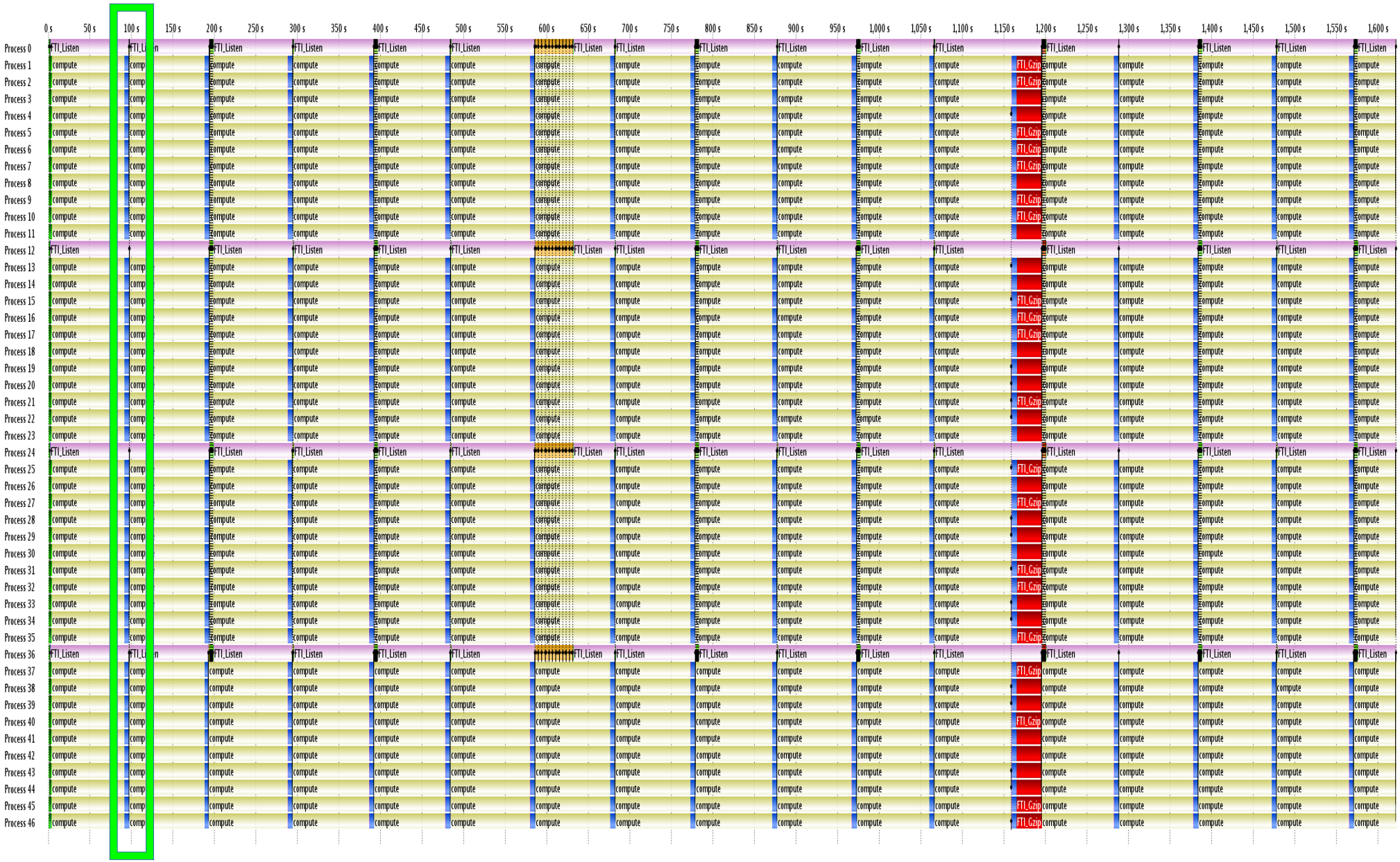
FTI traces on Tsubame2.0



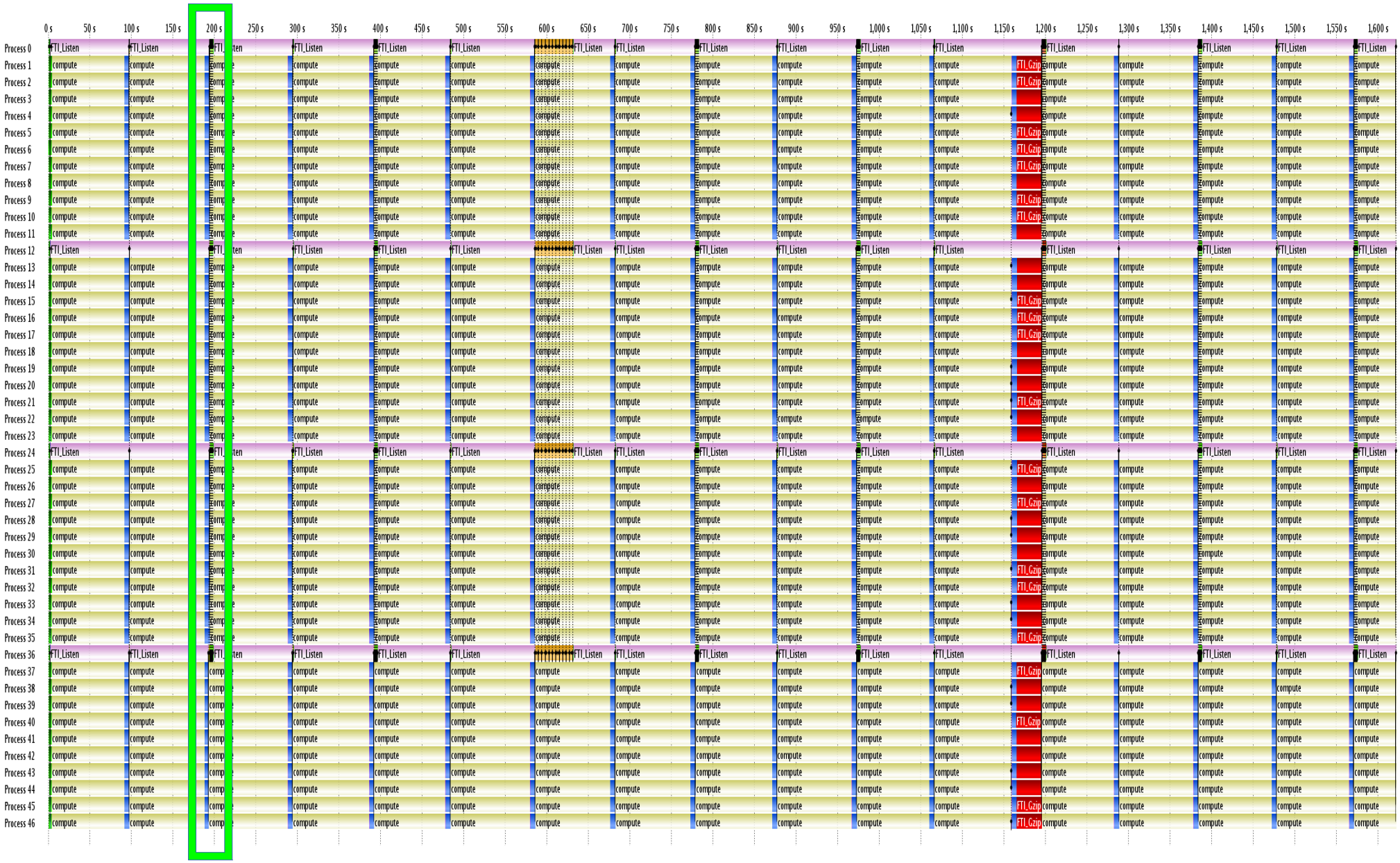
FTI traces on Tsubame2.0

FTI traces on Tsubame2.0

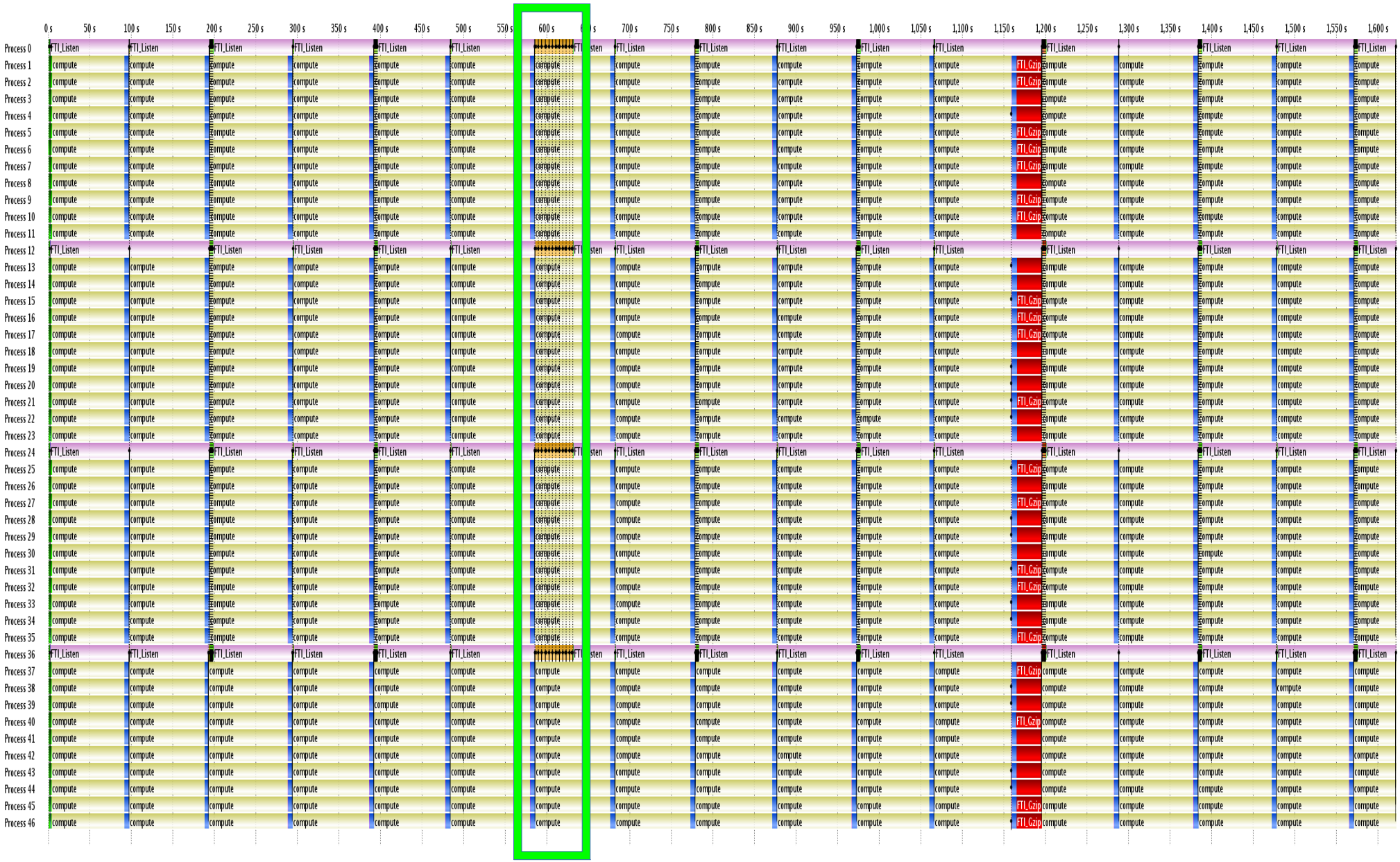
FTI traces on Tsubame2.0



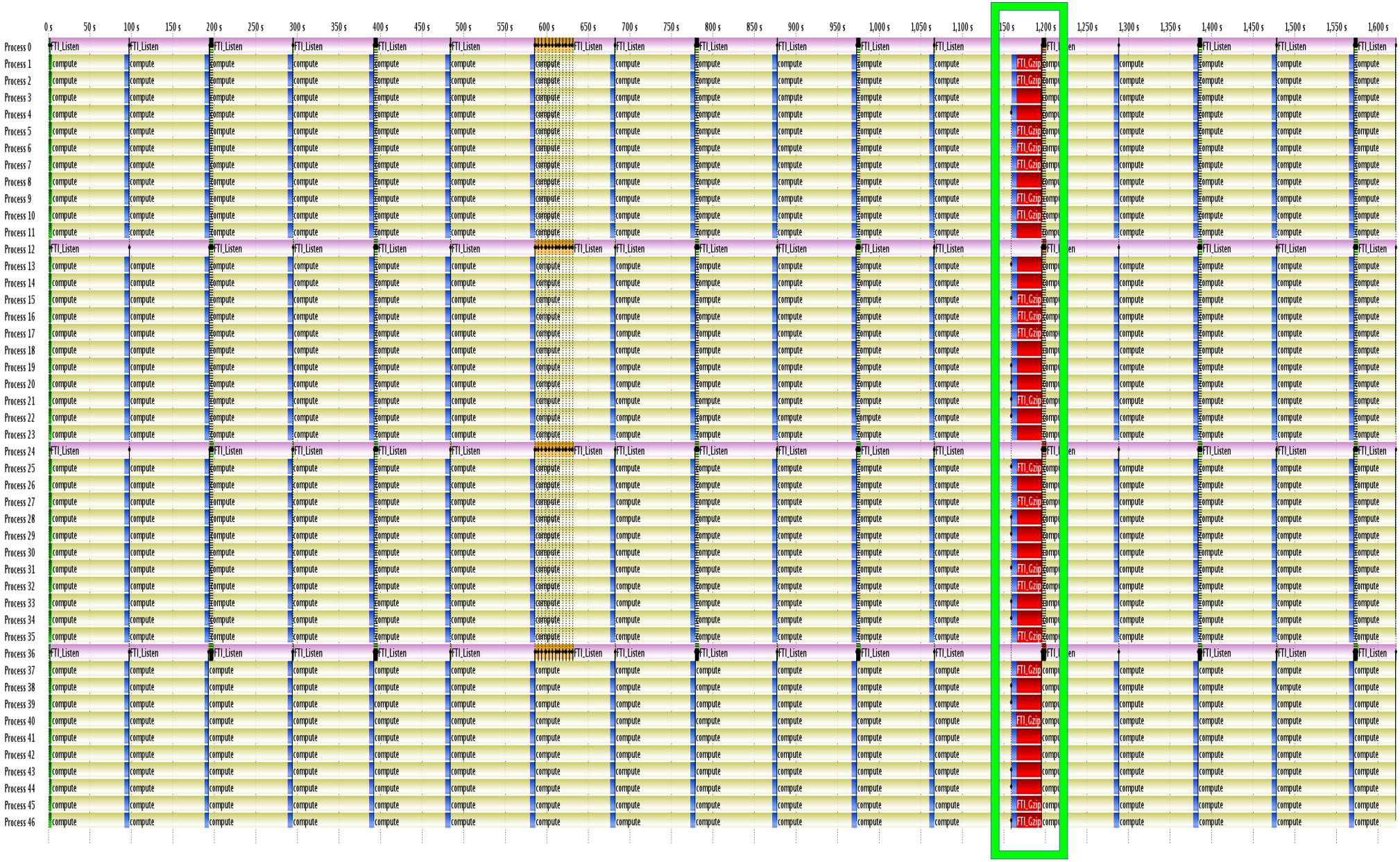
FTI traces on Tsubame2.0



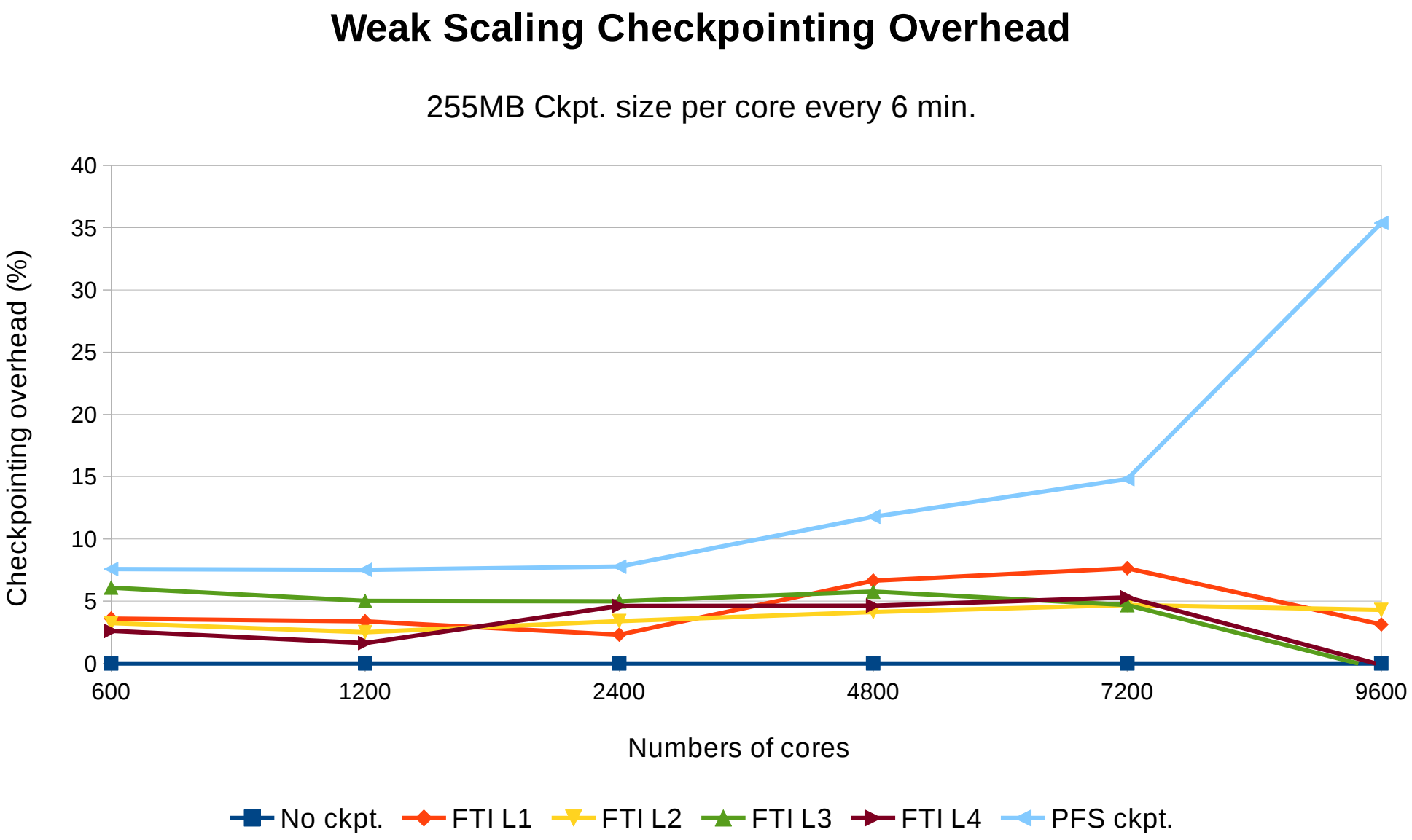
FTI traces on Tsubame2.0



FTI traces on Tsubame2.0



Multilevel Checkpointing




Space

Lossy Floating Point Compression

Lossy Floating Point Compression

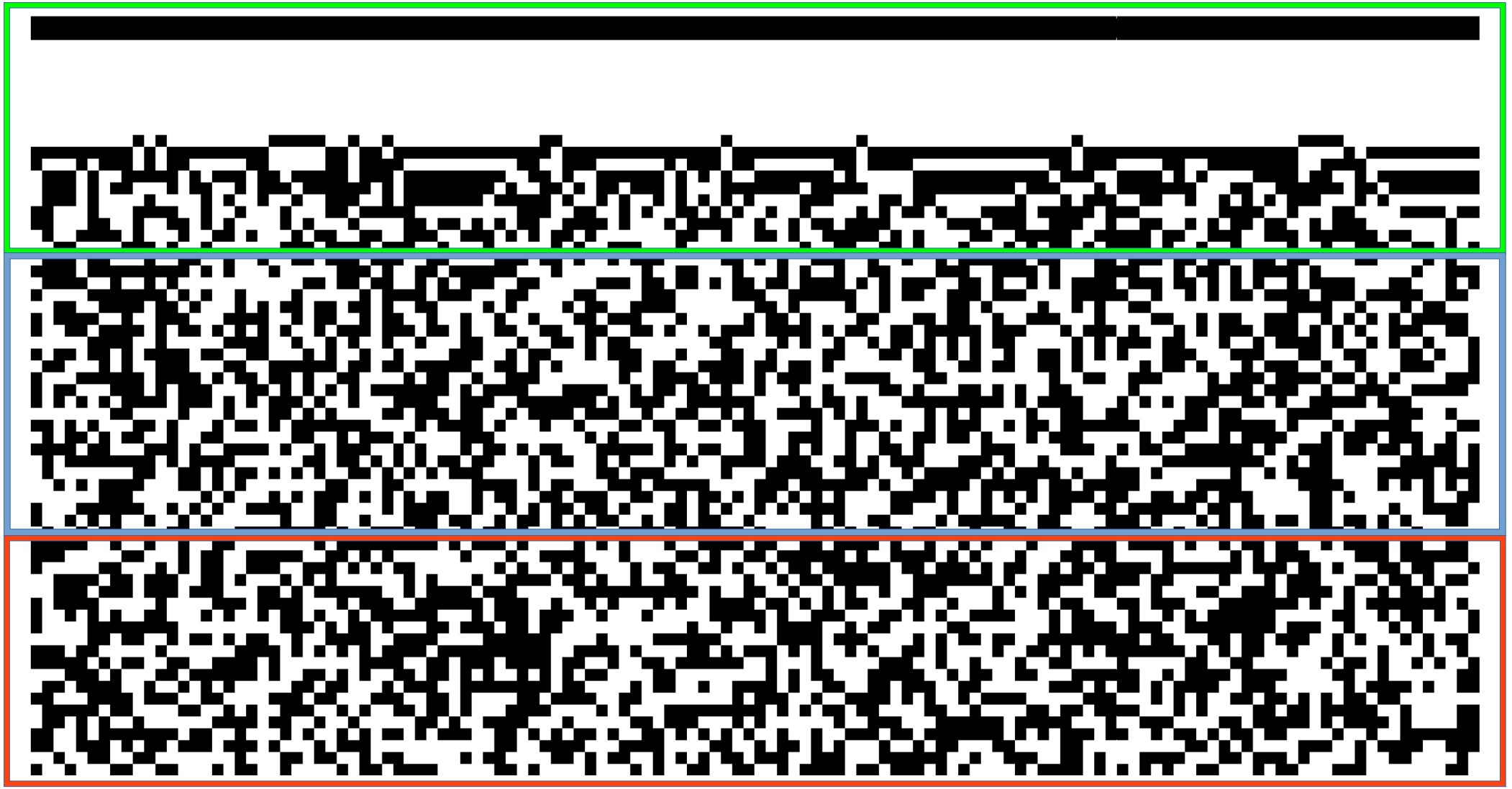
Compress

Lossy Floating Point Compression

Compress Not-to-compress

Lossy Floating Point Compression

Compress Not-to-compress Discard



lz vs fpzip (NRMSE & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		NRMSE(CR)		
fpzip16	5.7e-4(0.15)	4.6e-3(0.10)	1.3e-3(0.04)	1.7e-4(0.12)
fpzip20	3.6e-5(0.27)	2.9e-4(0.22)	8.1e-5(0.09)	1.0e-5(0.24)
fpzip24	2.2e-6(0.39)	1.8e-5(0.34)	5.1e-6(0.19)	6.5e-7(0.36)
fpzip28	1.3e-7(0.52)	1.1e-6(0.47)	3.0e-7(0.32)	3.9e-8(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	5.7e-4(0.28)	4.6e-3(0.16)	1.3e-3(0.09)	1.7e-4(0.26)
lz11	2.2e-6(0.53)	1.8e-5(0.41)	5.1e-6(0.34)	6.5e-7(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (NRMSE & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		NRMSE(CR)		
fpzip16	5.7e-4(0.15)	4.6e-3(0.10)	1.3e-3(0.04)	1.7e-4(0.12)
fpzip20	3.6e-5(0.27)	2.9e-4(0.22)	8.1e-5(0.09)	1.0e-5(0.24)
fpzip24	2.2e-6(0.39)	1.8e-5(0.34)	5.1e-6(0.19)	6.5e-7(0.36)
fpzip28	1.3e-7(0.52)	1.1e-6(0.47)	3.0e-7(0.32)	3.9e-8(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	5.7e-4(0.28)	4.6e-3(0.16)	1.3e-3(0.09)	1.7e-4(0.26)
lz11	2.2e-6(0.53)	1.8e-5(0.41)	5.1e-6(0.34)	6.5e-7(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (NRMSE & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		NRMSE(CR)		
fpzip16	5.7e-4(0.15)	4.6e-3(0.10)	1.3e-3(0.04)	1.7e-4(0.12)
fpzip20	3.6e-5(0.27)	2.9e-4(0.22)	8.1e-5(0.09)	1.0e-5(0.24)
fpzip24	2.2e-6(0.39)	1.8e-5(0.34)	5.1e-6(0.19)	6.5e-7(0.36)
fpzip28	1.3e-7(0.52)	1.1e-6(0.47)	3.0e-7(0.32)	3.9e-8(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	5.7e-4(0.28)	4.6e-3(0.16)	1.3e-3(0.09)	1.7e-4(0.26)
lz11	2.2e-6(0.53)	1.8e-5(0.41)	5.1e-6(0.34)	6.5e-7(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (NRMSE & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		NRMSE(CR)		
fpzip16	5.7e-4(0.15)	4.6e-3(0.10)	1.3e-3(0.04)	1.7e-4(0.12)
fpzip20	3.6e-5(0.27)	2.9e-4(0.22)	8.1e-5(0.09)	1.0e-5(0.24)
fpzip24	2.2e-6(0.39)	1.8e-5(0.34)	5.1e-6(0.19)	6.5e-7(0.36)
fpzip28	1.3e-7(0.52)	1.1e-6(0.47)	3.0e-7(0.32)	3.9e-8(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	5.7e-4(0.28)	4.6e-3(0.16)	1.3e-3(0.09)	1.7e-4(0.26)
lz11	2.2e-6(0.53)	1.8e-5(0.41)	5.1e-6(0.34)	6.5e-7(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (error & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		e_{nmax}(CR)		
fpzip16	3.1e-3(0.15)	9.9e-3(0.10)	6.8e-3(0.04)	5.3e-3(0.12)
fpzip20	2.0e-4(0.27)	6.2e-4(0.22)	4.3e-4(0.09)	3.9e-4(0.24)
fpzip24	1.2e-5(0.39)	3.9e-5(0.34)	2.6e-5(0.19)	2.4e-5(0.36)
fpzip28	7.1e-7(0.52)	2.7e-6(0.47)	1.6e-6(0.32)	1.5e-6(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	3.1e-3(0.28)	9.9e-3(0.16)	6.8e-3(0.09)	5.3e-3(0.26)
lz11	1.2e-5(0.53)	3.9e-5(0.41)	2.6e-5(0.34)	2.4e-5(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (error & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		e_{nmax}(CR)		
fpzip16	3.1e-3(0.15)	9.9e-3(0.10)	6.8e-3(0.04)	5.3e-3(0.12)
fpzip20	2.0e-4(0.27)	6.2e-4(0.22)	4.3e-4(0.09)	3.9e-4(0.24)
fpzip24	1.2e-5(0.39)	3.9e-5(0.34)	2.6e-5(0.19)	2.4e-5(0.36)
fpzip28	7.1e-7(0.52)	2.7e-6(0.47)	1.6e-6(0.32)	1.5e-6(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	3.1e-3(0.28)	9.9e-3(0.16)	6.8e-3(0.09)	5.3e-3(0.26)
lz11	1.2e-5(0.53)	3.9e-5(0.41)	2.6e-5(0.34)	2.4e-5(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (error & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		e_{nmax}(CR)		
fpzip16	3.1e-3(0.15)	9.9e-3(0.10)	6.8e-3(0.04)	5.3e-3(0.12)
fpzip20	2.0e-4(0.27)	6.2e-4(0.22)	4.3e-4(0.09)	3.9e-4(0.24)
fpzip24	1.2e-5(0.39)	3.9e-5(0.34)	2.6e-5(0.19)	2.4e-5(0.36)
fpzip28	7.1e-7(0.52)	2.7e-6(0.47)	1.6e-6(0.32)	1.5e-6(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	3.1e-3(0.28)	9.9e-3(0.16)	6.8e-3(0.09)	5.3e-3(0.26)
lz11	1.2e-5(0.53)	3.9e-5(0.41)	2.6e-5(0.34)	2.4e-5(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (error & CR)

Comp. Method	U	FSDSC	Z3	CCN3
		e_{nmax}(CR)		
fpzip16	3.1e-3(0.15)	9.9e-3(0.10)	6.8e-3(0.04)	5.3e-3(0.12)
fpzip20	2.0e-4(0.27)	6.2e-4(0.22)	4.3e-4(0.09)	3.9e-4(0.24)
fpzip24	1.2e-5(0.39)	3.9e-5(0.34)	2.6e-5(0.19)	2.4e-5(0.36)
fpzip28	7.1e-7(0.52)	2.7e-6(0.47)	1.6e-6(0.32)	1.5e-6(0.49)
fpzip32	0.0e+0(0.64)	0.0e+0(0.59)	0.0e+0(0.44)	0.0e+0(0.61)
lz12	3.1e-3(0.28)	9.9e-3(0.16)	6.8e-3(0.09)	5.3e-3(0.26)
lz11	1.2e-5(0.53)	3.9e-5(0.41)	2.6e-5(0.34)	2.4e-5(0.51)
lz10	0.0e+0(0.78)	0.0e+0(0.66)	0.0e+0(0.59)	0.0e+0(0.76)

lz vs fpzip (CT & CR)

Compression and reconstruction timings and compression ratios (CR) for variables U (3D) and FSDSC (2D)

		U			FSDSC	
Method	comp	reconst.	CR	comp	reconst.	CR
fpzip16	0.100	0.100	0.15	0.003	0.003	0.10
fpzip20	0.110	0.112	0.27	0.003	0.003	0.22
fpzip24	0.119	0.117	0.39	0.004	0.004	0.34
fpzip28	0.127	0.122	0.52	0.004	0.004	0.47
fpzip32	0.143	0.131	0.64	0.005	0.004	0.59
lz10	0.108	0.064	0.78	0.005	0.003	0.66
lz11	0.107	0.065	0.53	0.005	0.003	0.41
lz12	0.097	0.074	0.28	0.004	0.003	0.16

lz vs fpzip (CT & CR)

Compression and reconstruction timings and compression ratios (CR) for variables U (3D) and FSDSC (2D)

		U			FSDSC	
Method	comp	reconst.	CR	comp	reconst.	CR
fpzip16	0.100	0.100	0.15	0.003	0.003	0.10
fpzip20	0.110	0.112	0.27	0.003	0.003	0.22
fpzip24	0.119	0.117	0.39	0.004	0.004	0.34
fpzip28	0.127	0.122	0.52	0.004	0.004	0.47
fpzip32	0.143	0.131	0.64	0.005	0.004	0.59
lz10	0.108	0.064	0.78	0.005	0.003	0.66
lz11	0.107	0.065	0.53	0.005	0.003	0.41
lz12	0.097	0.074	0.28	0.004	0.003	0.16

lz vs fpzip (CT & CR)

Compression and reconstruction timings and compression ratios (CR) for variables U (3D) and FSDSC (2D)

		U			FSDSC	
Method	comp	reconst.	CR	comp	reconst.	CR
fpzip16	0.100	0.100	0.15	0.003	0.003	0.10
fpzip20	0.110	0.112	0.27	0.003	0.003	0.22
fpzip24	0.119	0.117	0.39	0.004	0.004	0.34
fpzip28	0.127	0.122	0.52	0.004	0.004	0.47
fpzip32	0.143	0.131	0.64	0.005	0.004	0.59
lz10	0.108	0.064	0.78	0.005	0.003	0.66
lz11	0.107	0.065	0.53	0.005	0.003	0.41
lz12	0.097	0.074	0.28	0.004	0.003	0.16

lz vs fpzip (CT & CR)

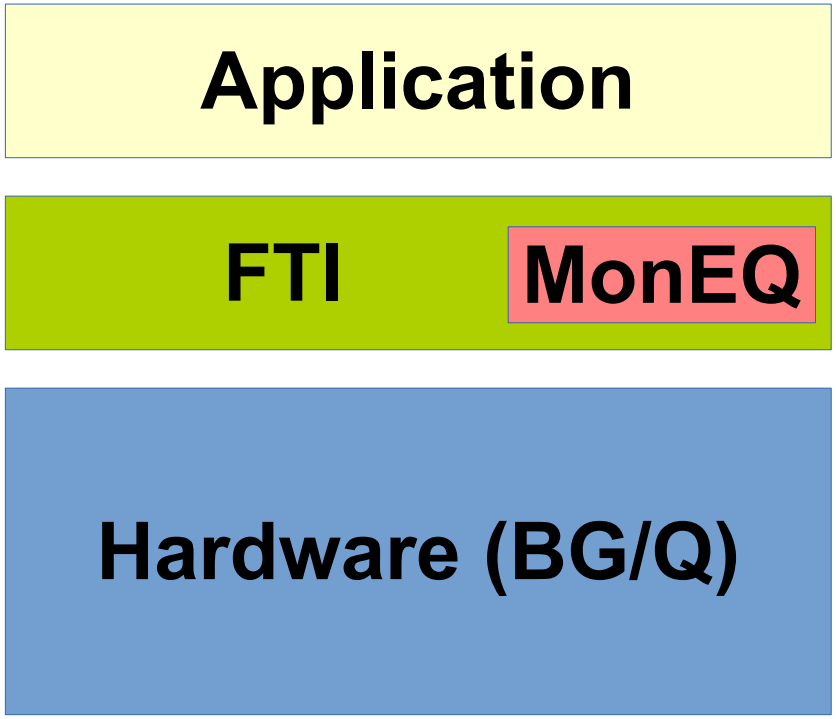
Compression and reconstruction timings and compression ratios (CR) for variables U (3D) and FSDSC (2D)

		U			FSDSC	
Method	comp	reconst.	CR	comp	reconst.	CR
fpzip16	0.100	0.100	0.15	0.003	0.003	0.10
fpzip20	0.110	0.112	0.27	0.003	0.003	0.22
fpzip24	0.119	0.117	0.39	0.004	0.004	0.34
fpzip28	0.127	0.122	0.52	0.004	0.004	0.47
fpzip32	0.143	0.131	0.64	0.005	0.004	0.59
lz10	0.108	0.064	0.78	0.005	0.003	0.66
lz11	0.107	0.065	0.53	0.005	0.003	0.41
lz12	0.097	0.074	0.28	0.004	0.003	0.16

Energy

Power Monitoring on BG/Q

- Time resolution: 560ms.
- Space resolution: Node card level (32 nodes).
- Readings include: Node (DRAM, SRAM, Core, Network, PCIExpress, Optics)



#sec	node_card	chip_core	dram	network	sram	optics	PCIexpress	link_chip_core
6.05	1911.093800	1072.58680	425.240700	53.084900	58.296200	208.650100	51.742000	41.492900
7.17	1612.405100	947.188700	251.289900	53.996800	58.296200	208.997700	51.868900	40.766900
8.29	1605.464300	942.388400	251.582300	52.703600	58.246300	208.390300	51.395900	40.757500
9.41	1605.635600	941.659200	251.841000	53.145800	58.203900	208.561700	51.742000	40.482000
10.53	1610.196400	943.327400	253.116100	53.511300	58.296200	209.913300	51.726100	40.306000
11.65	1615.358800	948.922600	253.116100	53.936800	57.828700	208.650100	52.200300	40.704100
12.77	1603.624900	936.793100	254.385600	53.190000	58.296200	208.002200	52.200300	40.757500
13.89	1602.159500	935.955600	253.116100	53.145800	58.296200	209.479600	51.797500	40.368800

Experimental Platform

Machine:

- Mira (10 PF) IBM Blue Gene/Q
- 49,152 nodes organized in 48 racks
- 16 cores of 1.6 GHz PowerPC A2
- 16 GB of DDR3 memory
- Proprietary 5-D torus network
- Water-cooled.

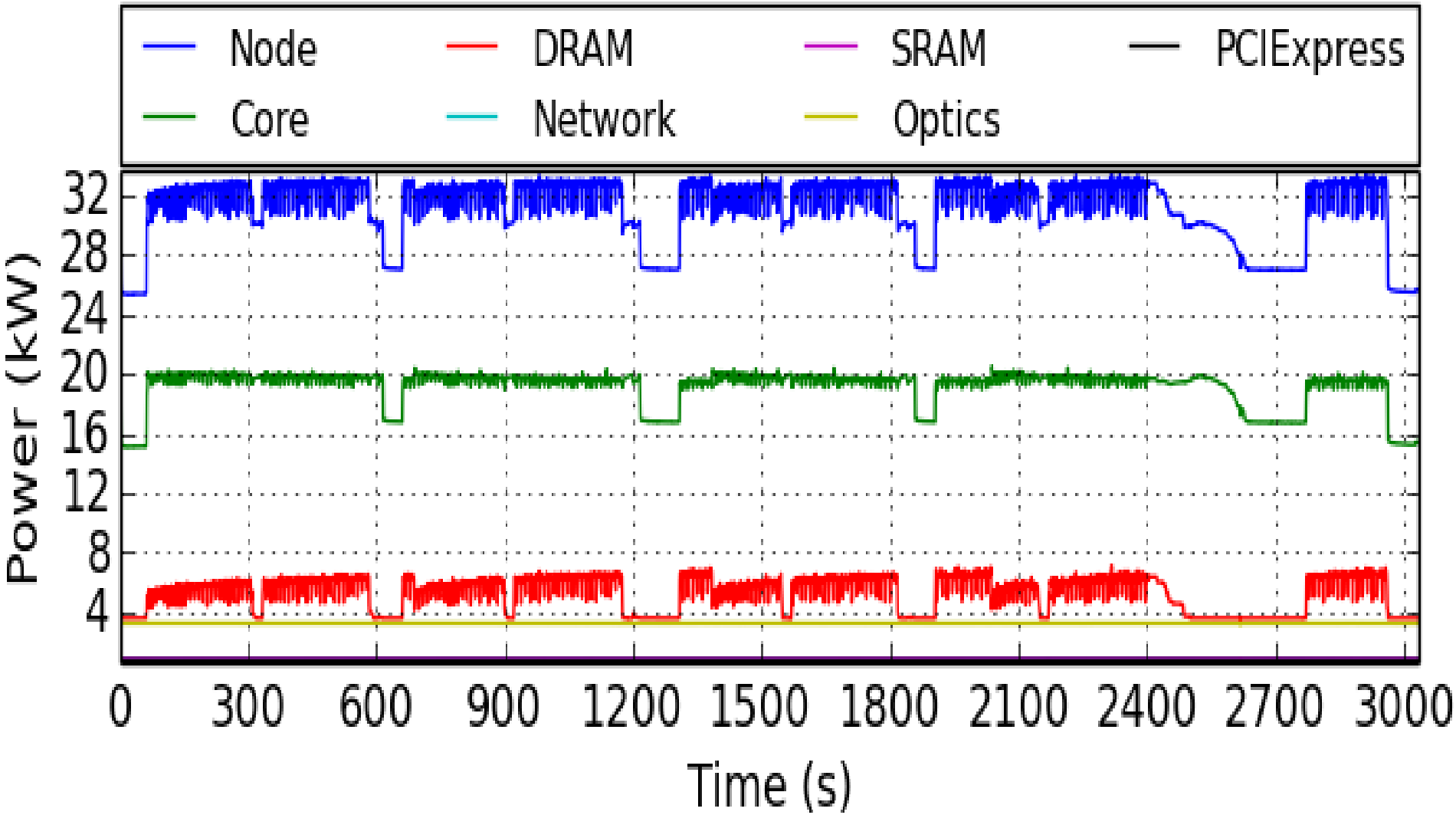
Application

- Lennard-Jones simulation of 1.3 billion atoms
- 512 nodes, 64 MPI processes per node (32,678pr.)

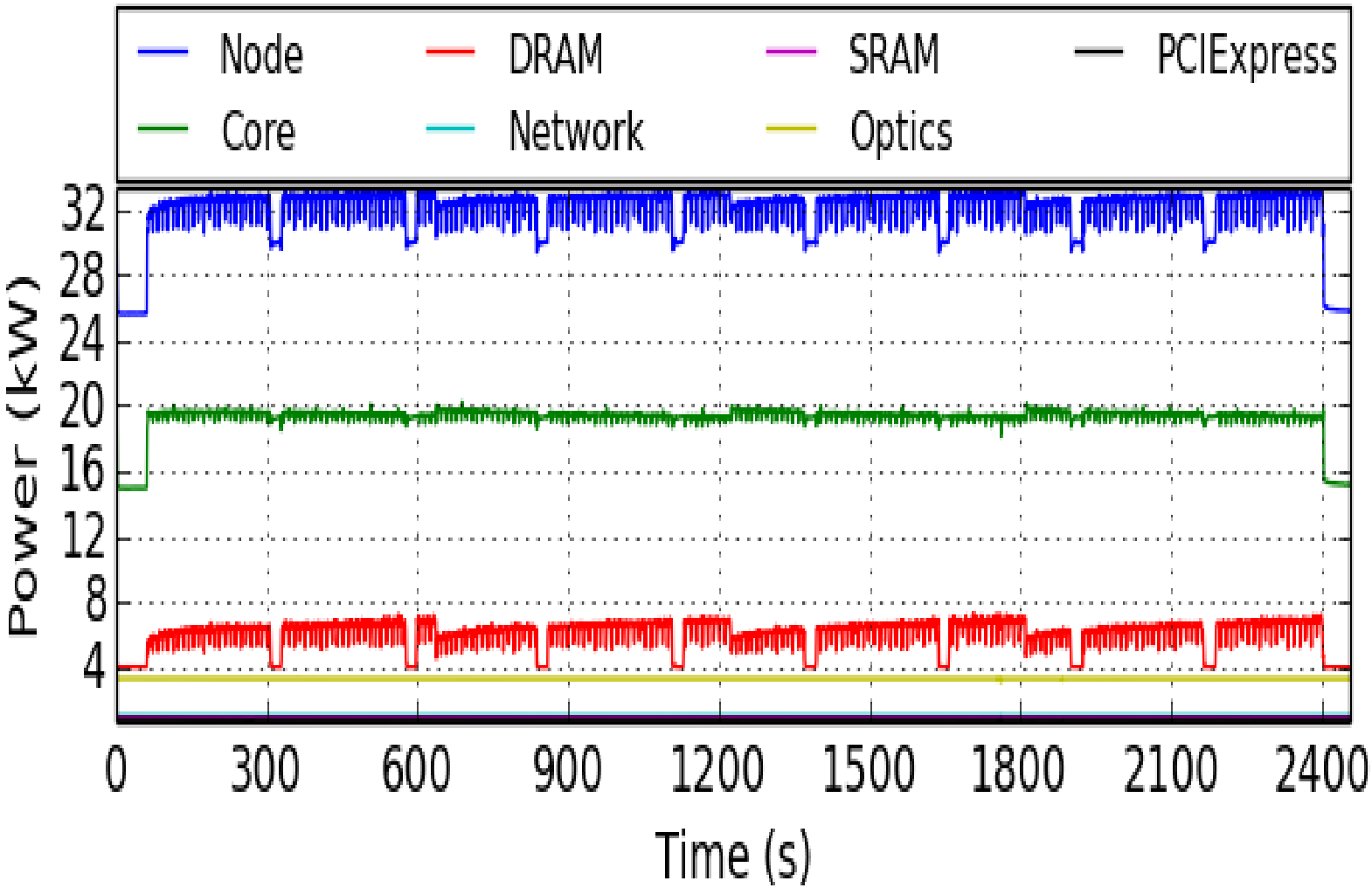
Power Monitoring

- FTI/MonEQ

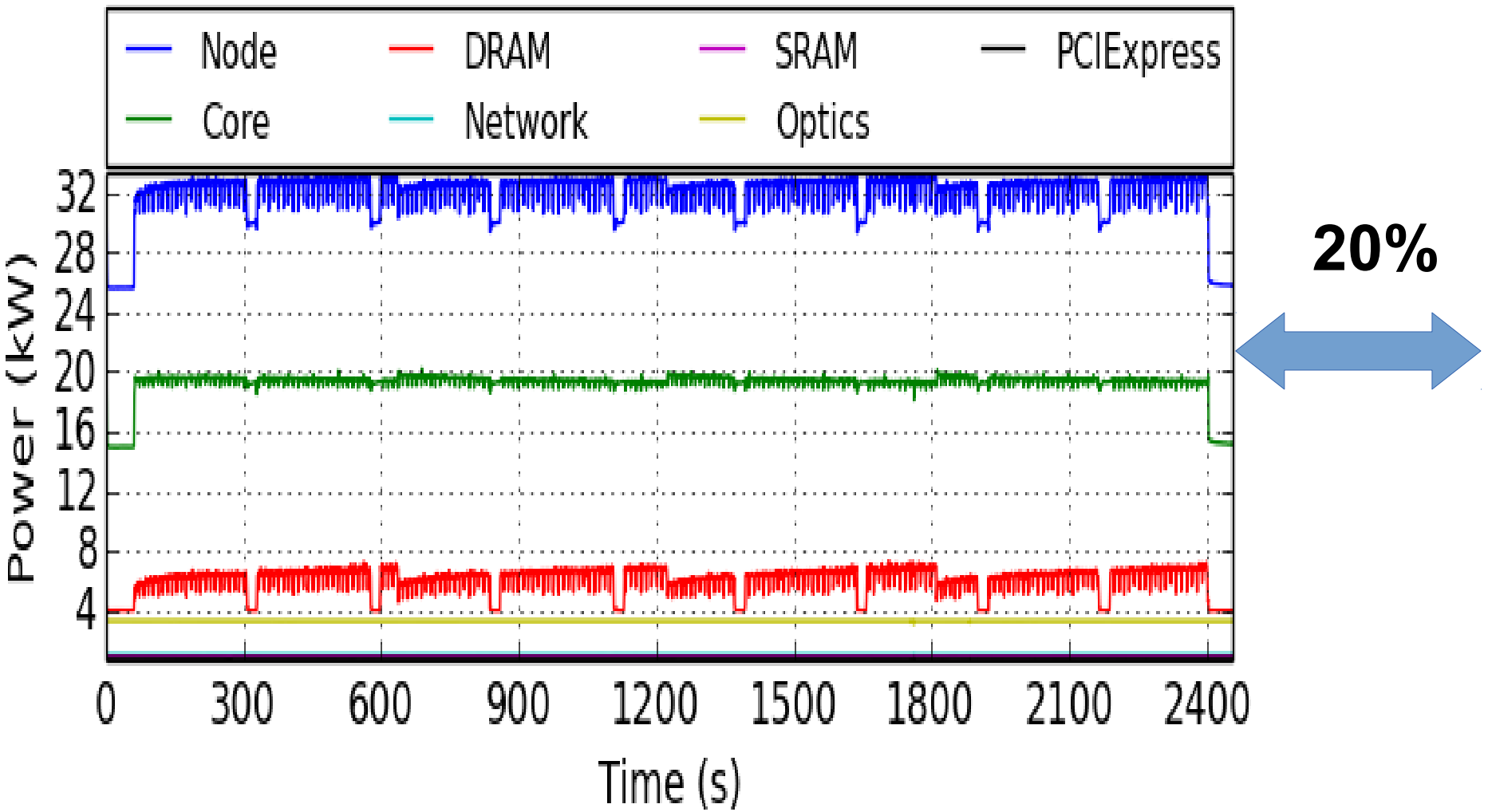
Synchronous Multilevel Checkpointing Power Consumption



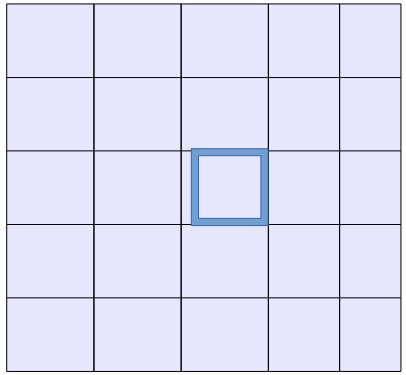
Asynchronous Multilevel Checkpointing Power Consumption

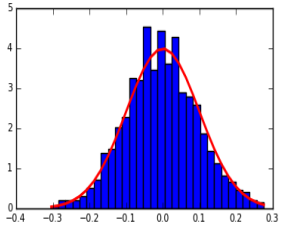


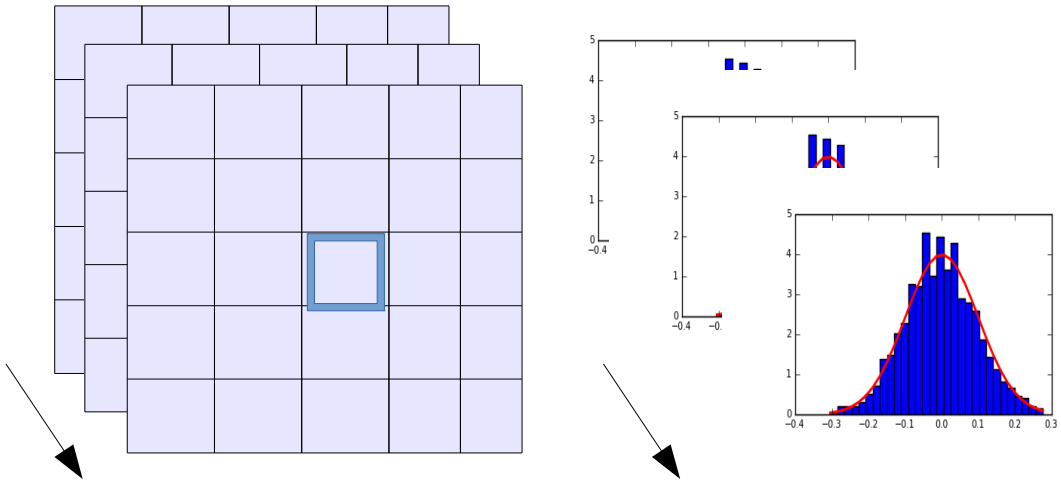
Asynchronous Multilevel Checkpointing Power Consumption



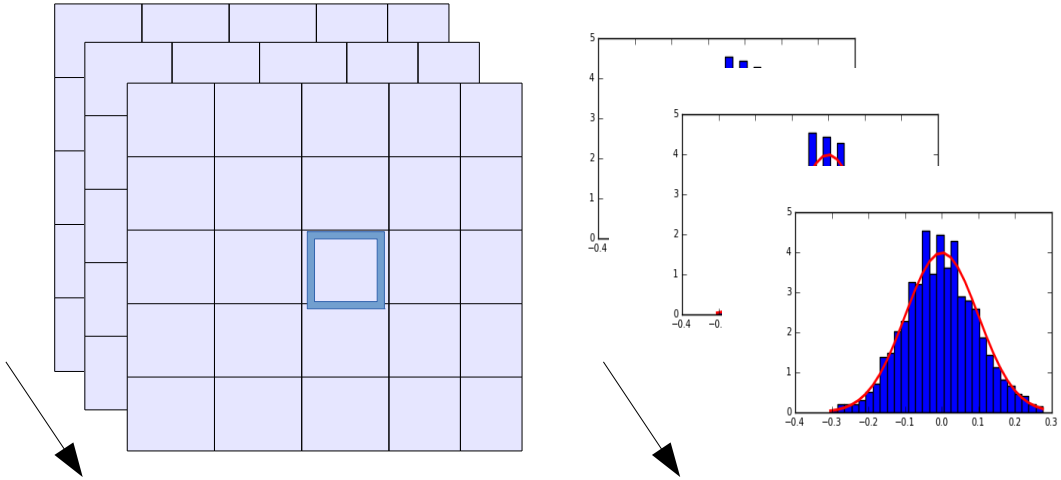
Information



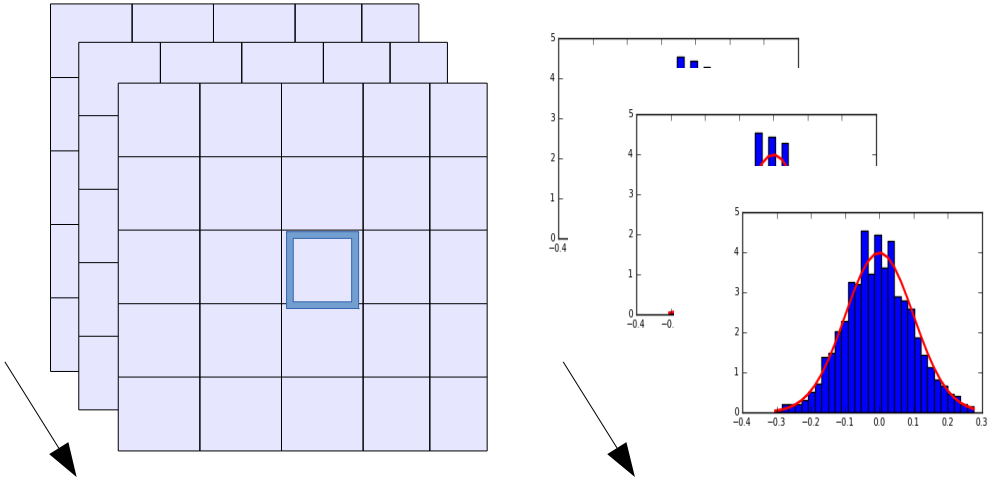

5
4
3
2
1
0
−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3


5
4
3
2
1
0
−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3

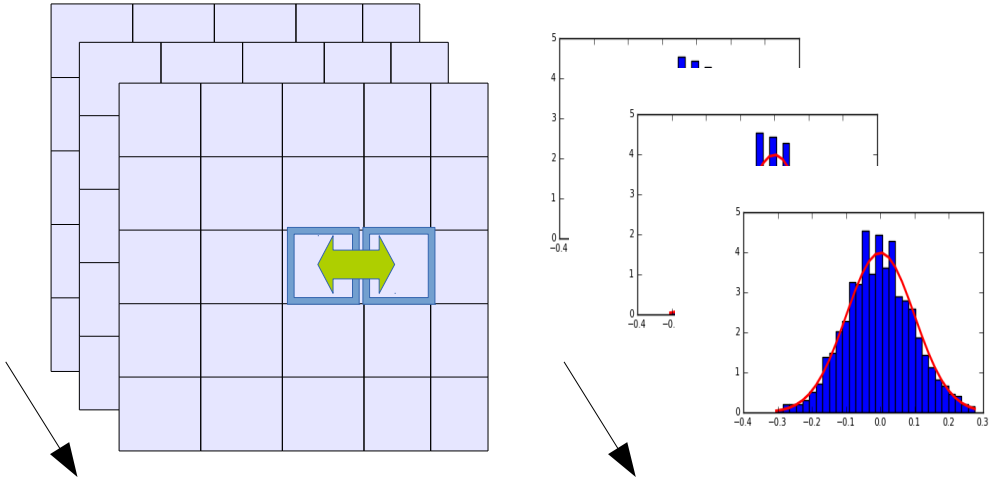
Gamma Detector



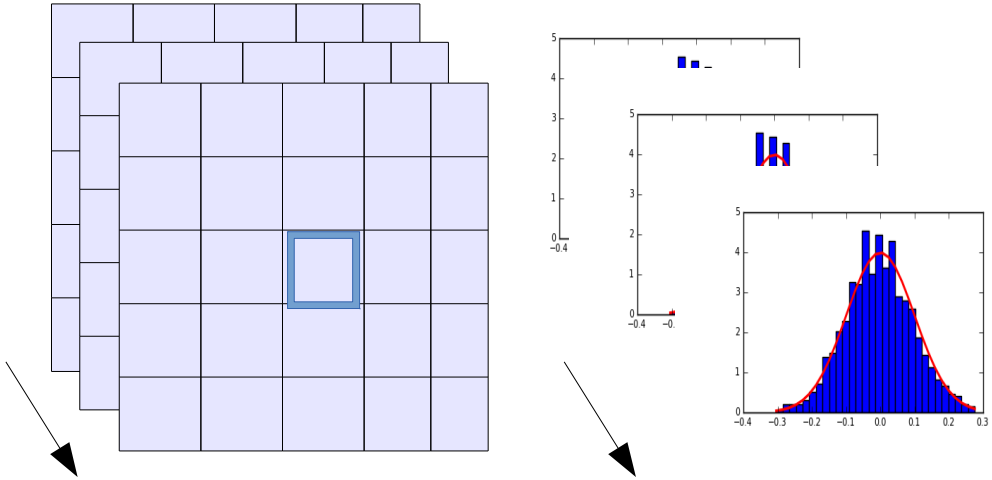
Gamma Detector



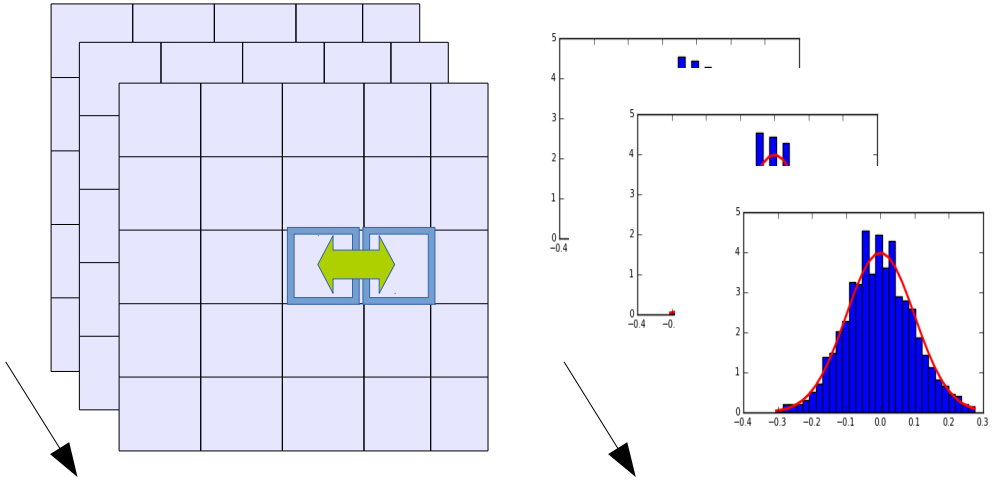
Delta Detector



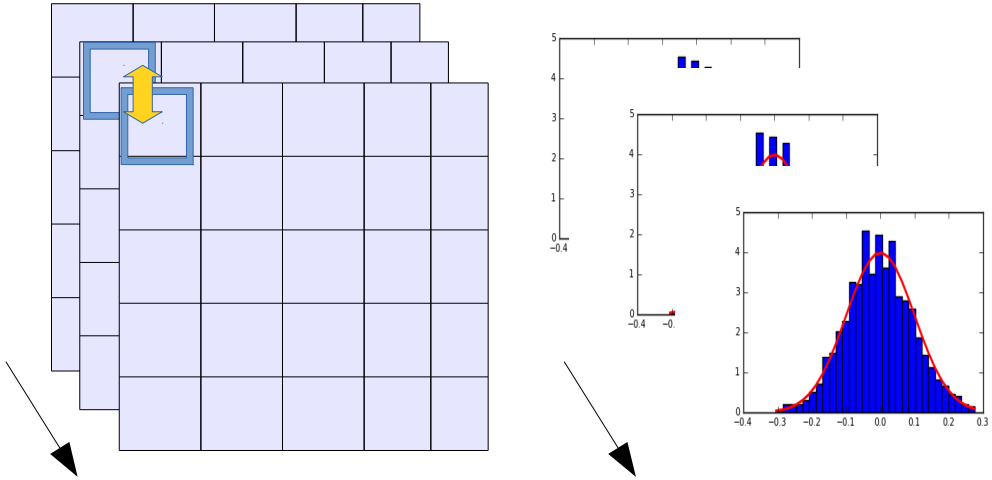
Gamma Detector



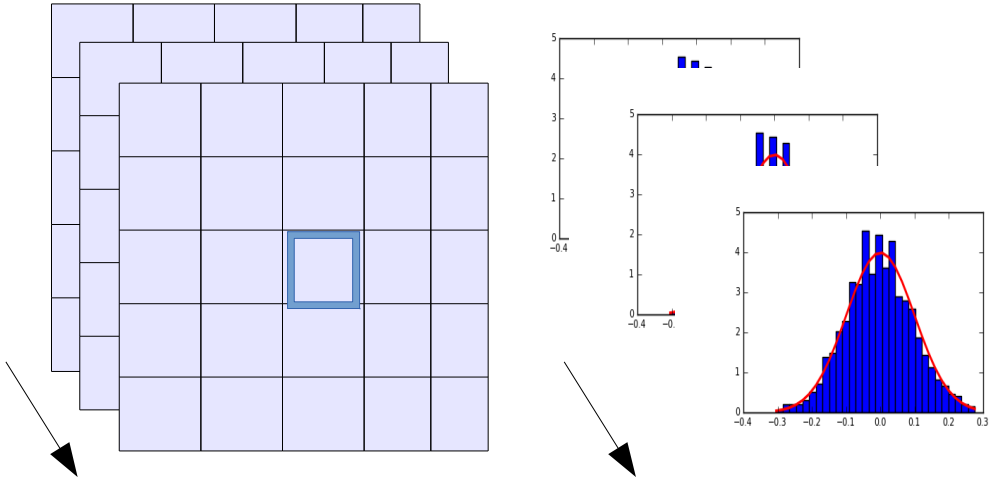
Delta Detector



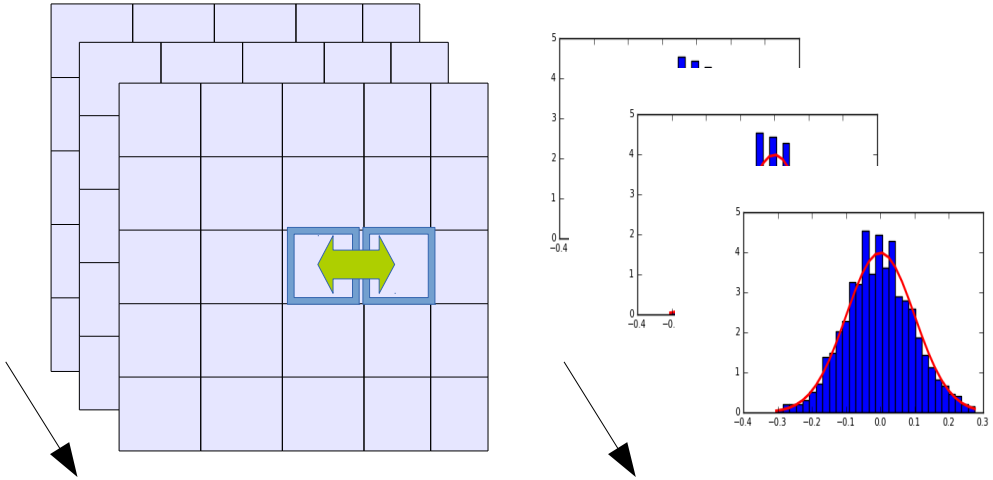
Epsilon Detector



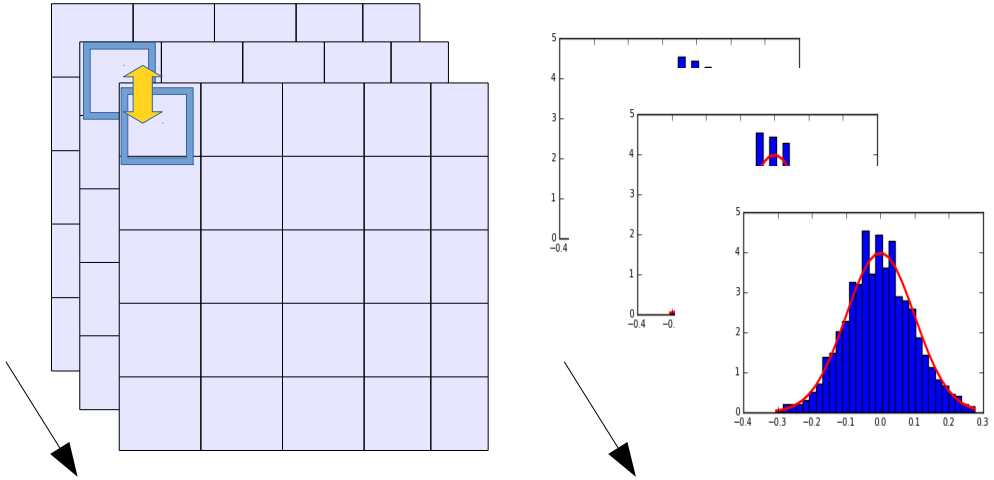
Gamma Detector



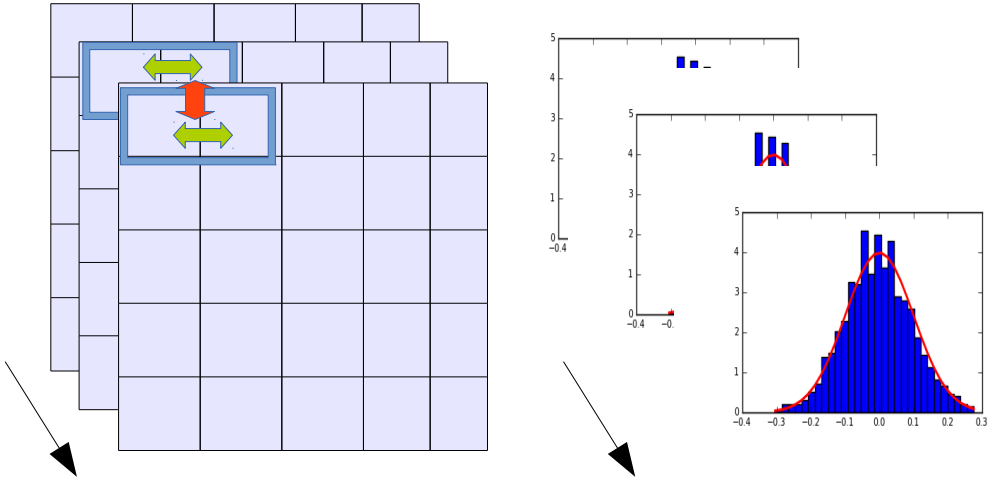
Delta Detector



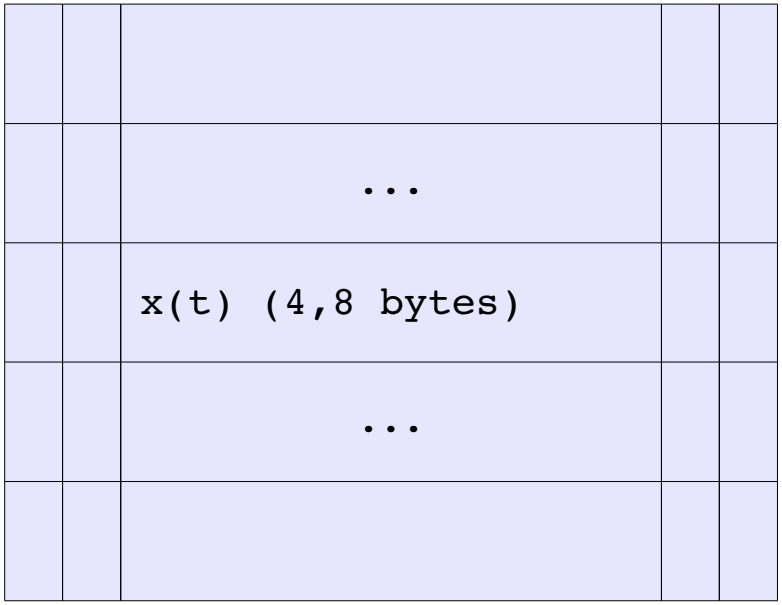
Epsilon Detector



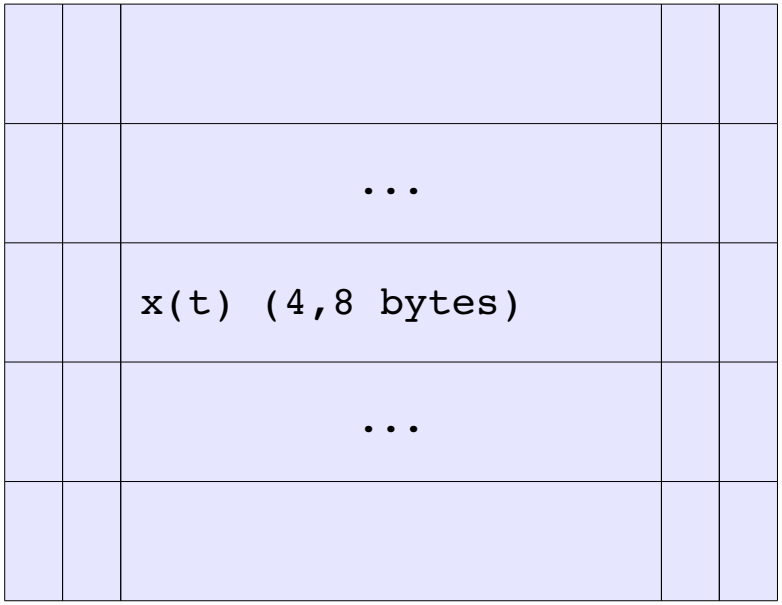
Zeta Detector



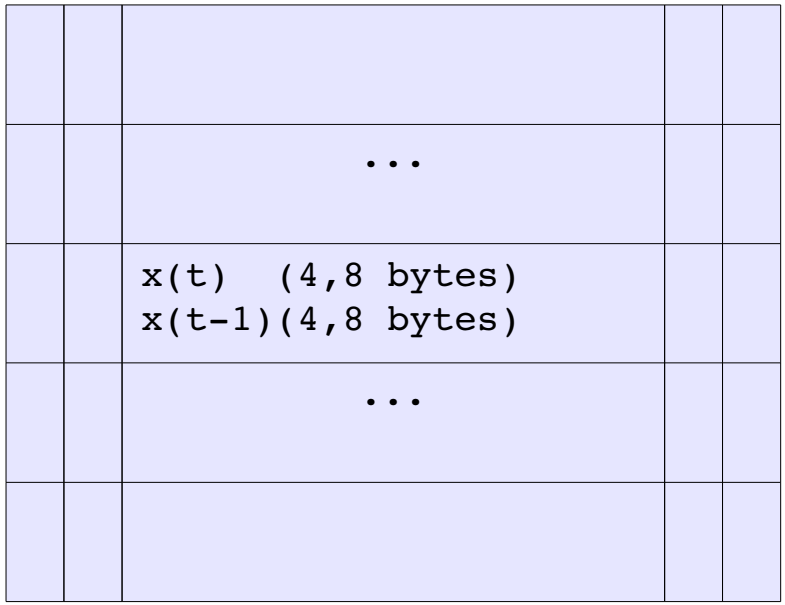
Classic 2D domain



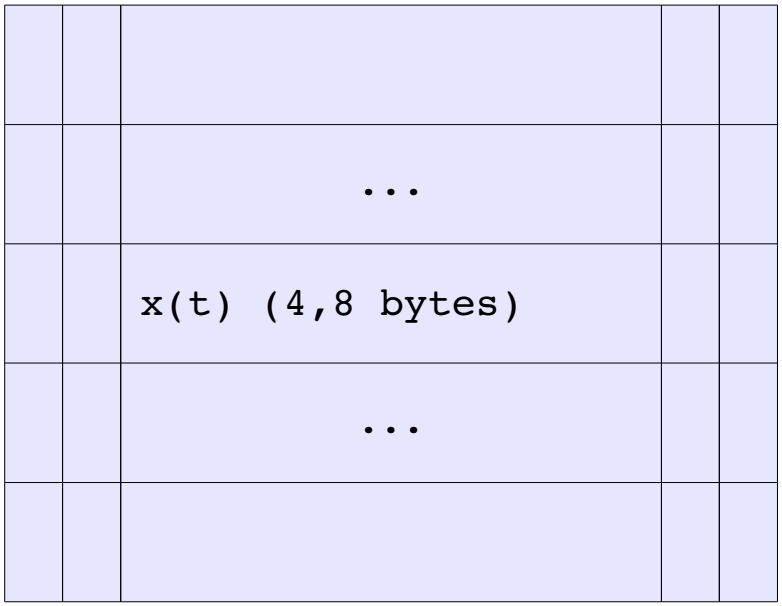
Classic 2D domain



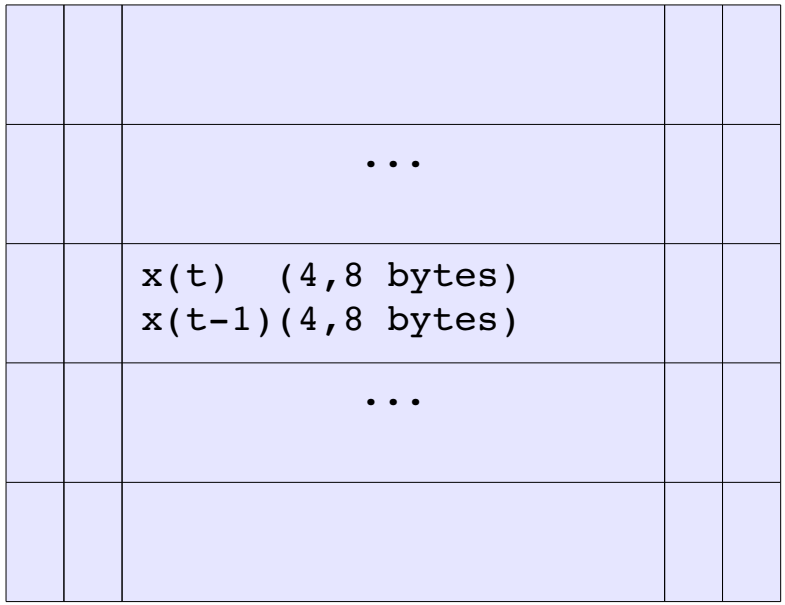
2D domain with 1 step memory



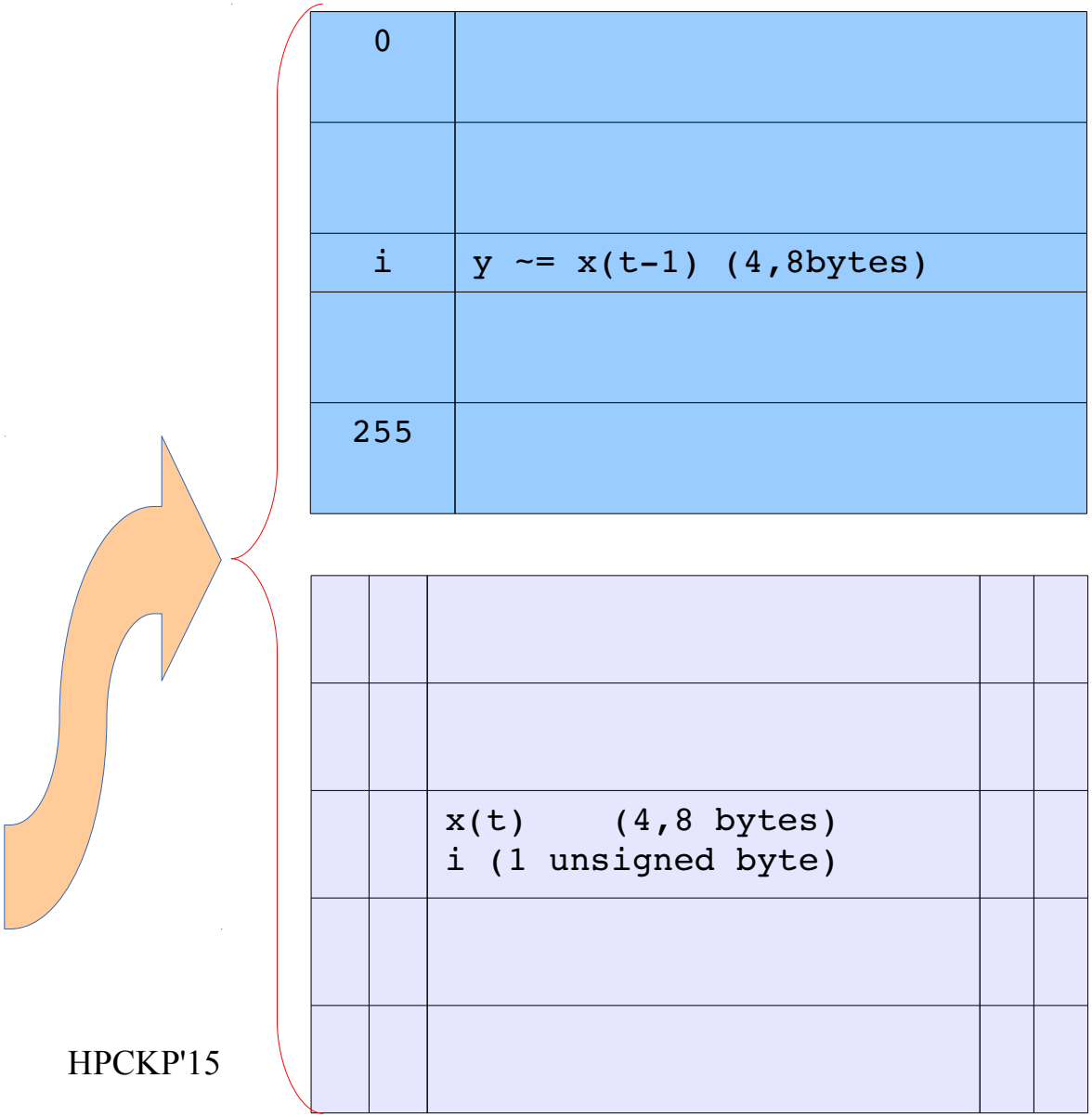
Classic 2D domain



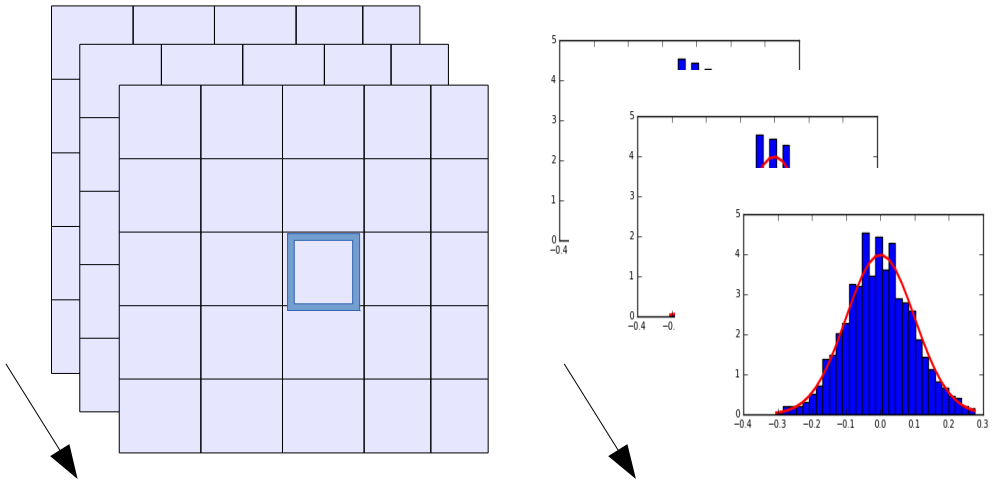
2D domain with 1 step memory



2D domain with approximated 1 step memory

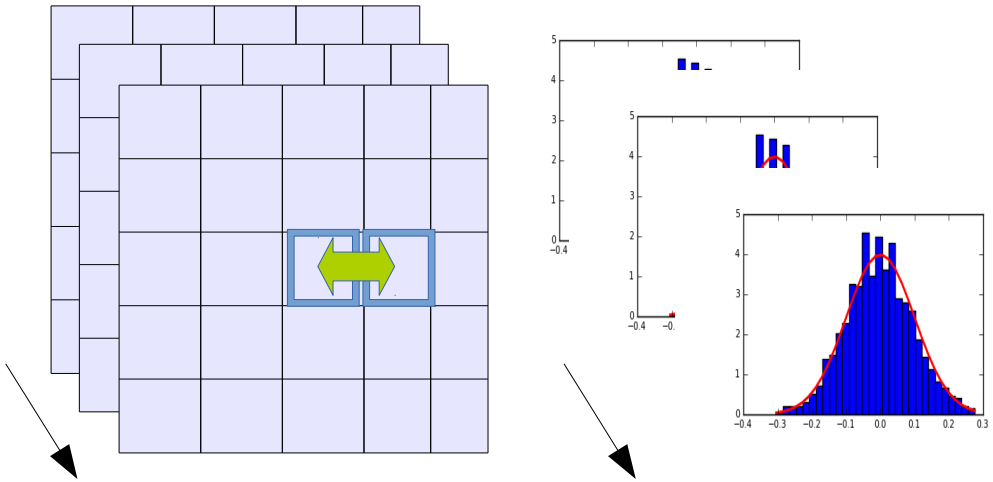


Gamma Detector



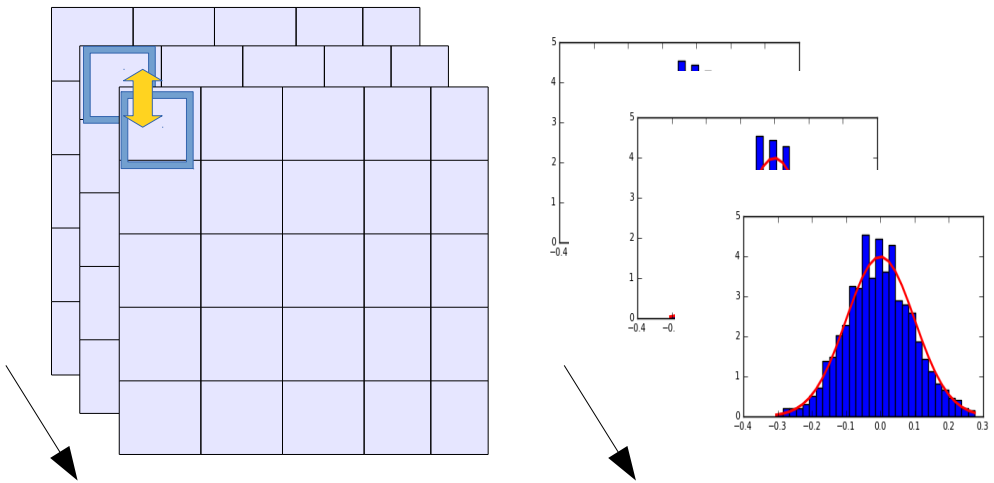
0x + 1KB or 2KB

Delta Detector



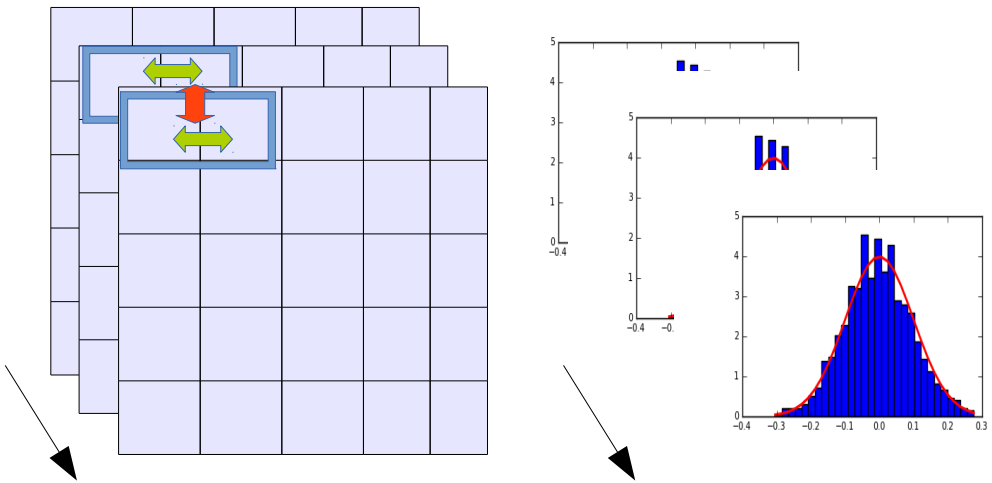
0x + 1KB or 2KB

Epsilon Detector



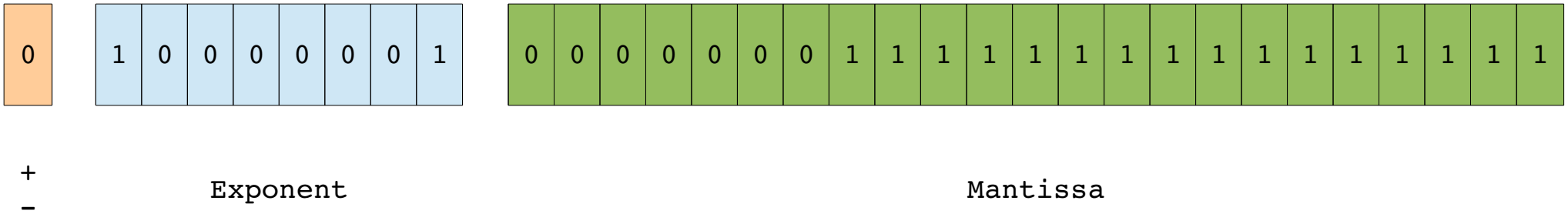
0.25x or 0.12x + 1KB or 2KB

Zeta Detector

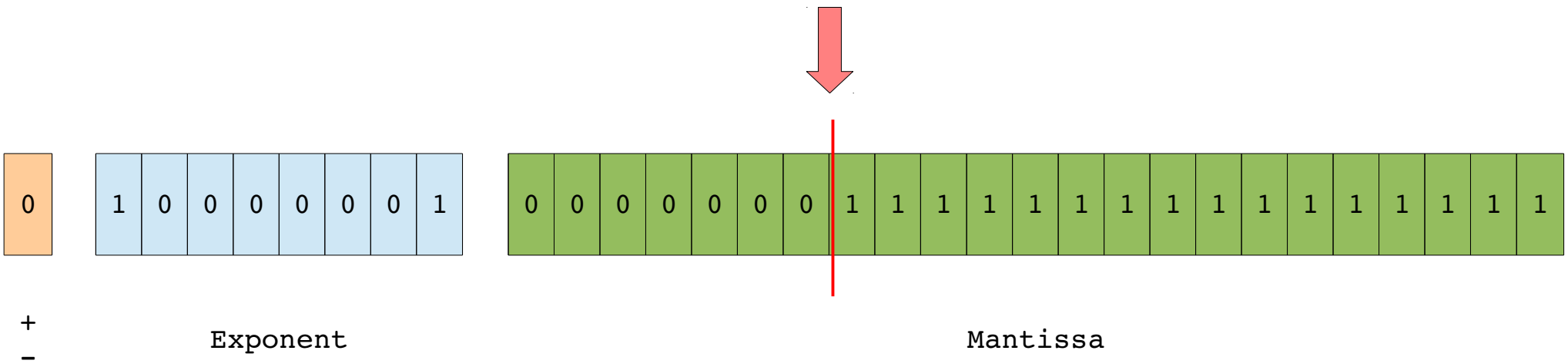


0.25x or 0.12x + 1KB or 2KB

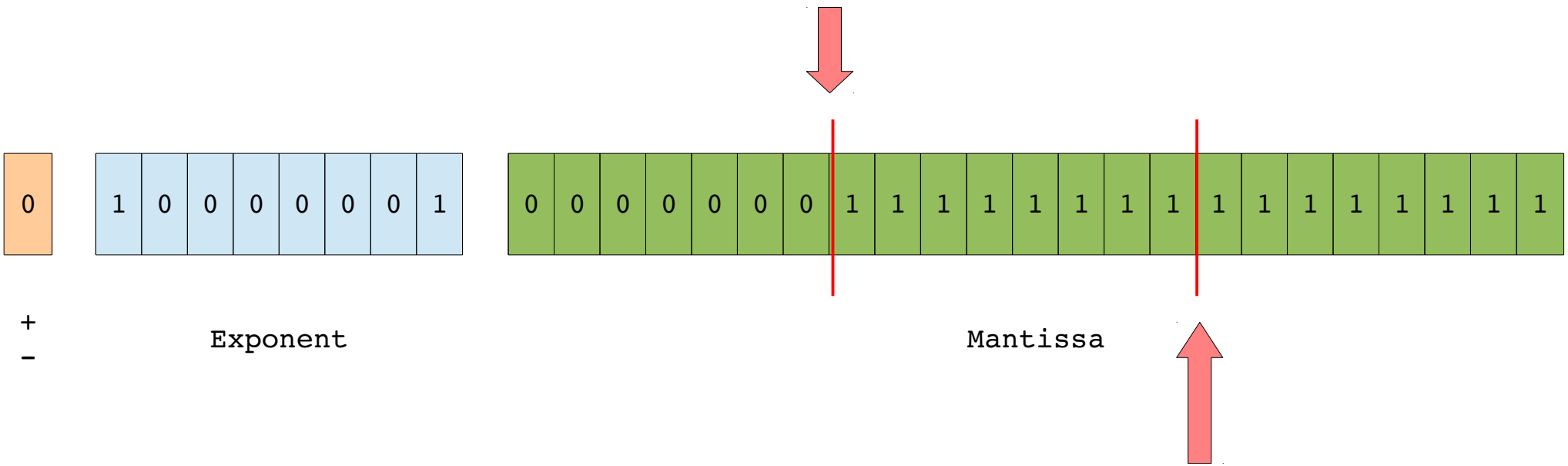
4.000 < 4.0312495 < 4.050



4.000 < 4.0312495 < 4.050 => 16/32 bit-flips detected (50% recall)

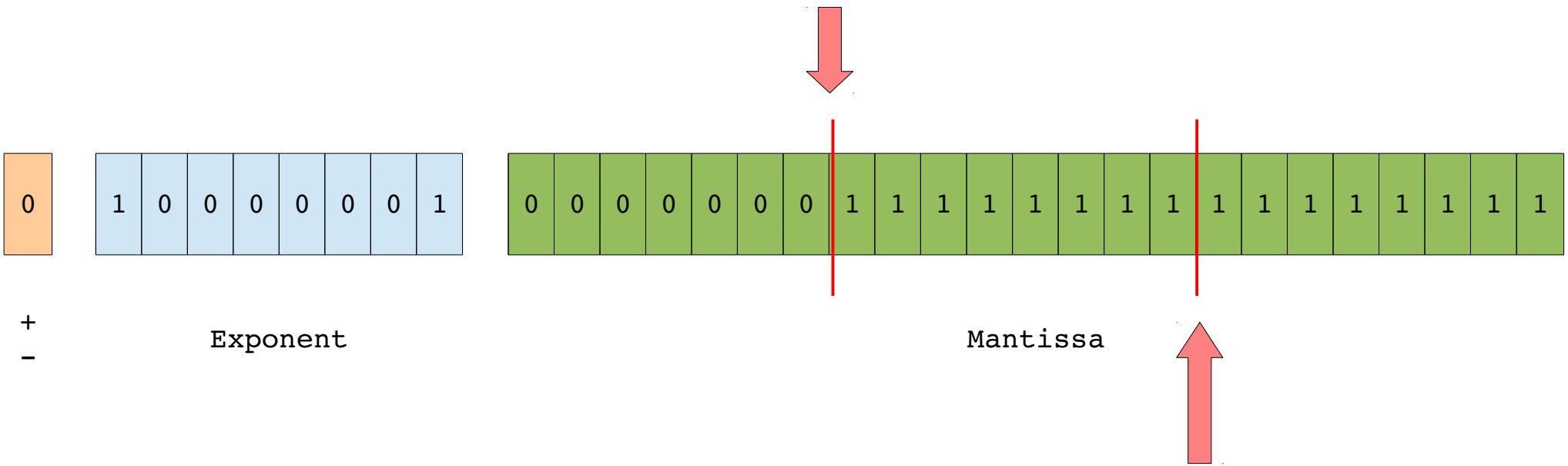


4.000 < 4.0312495 < 4.050 => 16/32 bit-flips detected (50% recall)

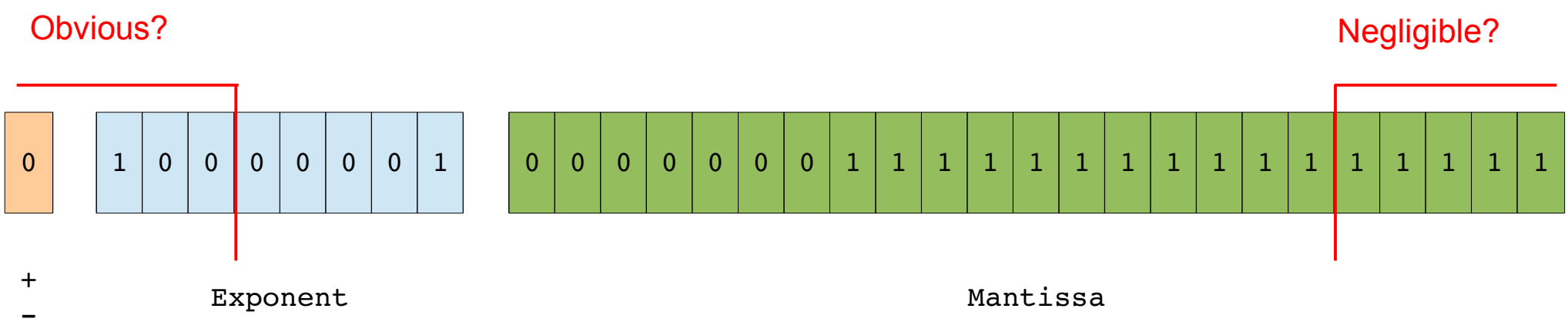


4.03115 < 4.0312495 < 4.03185 => 24/32 bit-flips detected (75% recall)

4.000 < 4.0312495 < 4.050 => 16/32 bit-flips detected (50% recall)

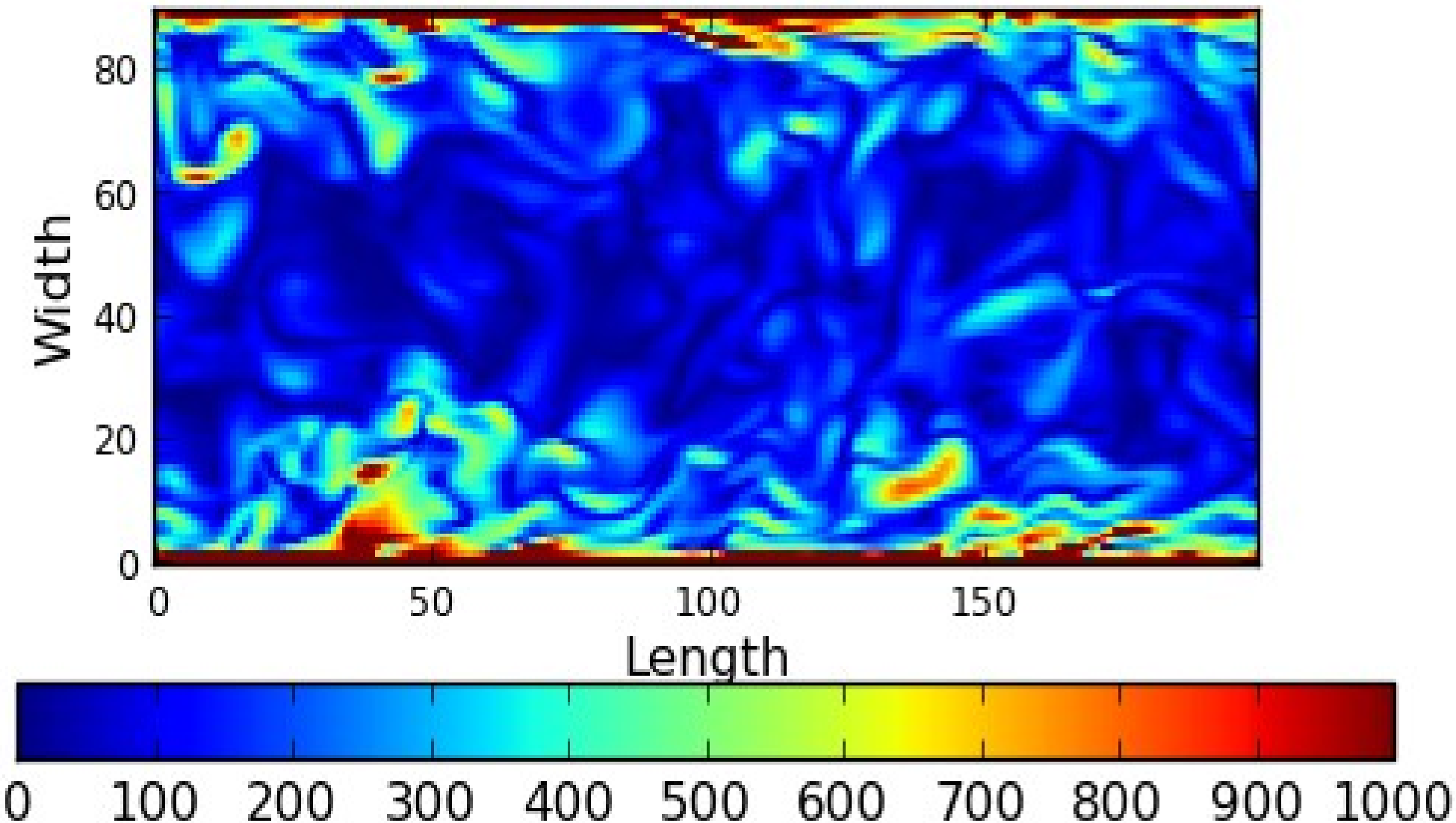


4.03115 < 4.0312495 < 4.03185 => 24/32 bit-flips detected (75% recall)

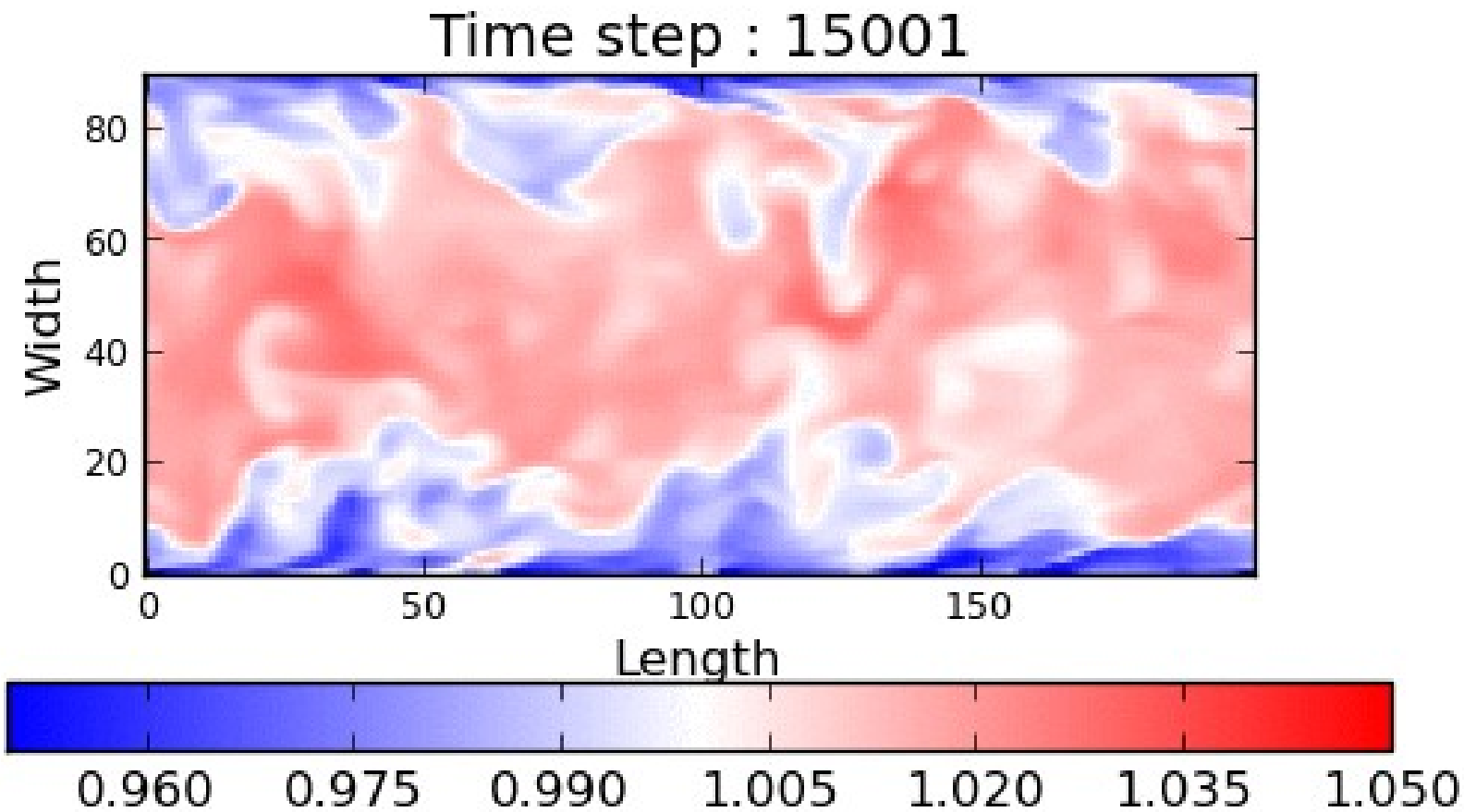


12/24 bit-flips detected (50% coverage) or 20/24 bit-flips detected (83% coverage)

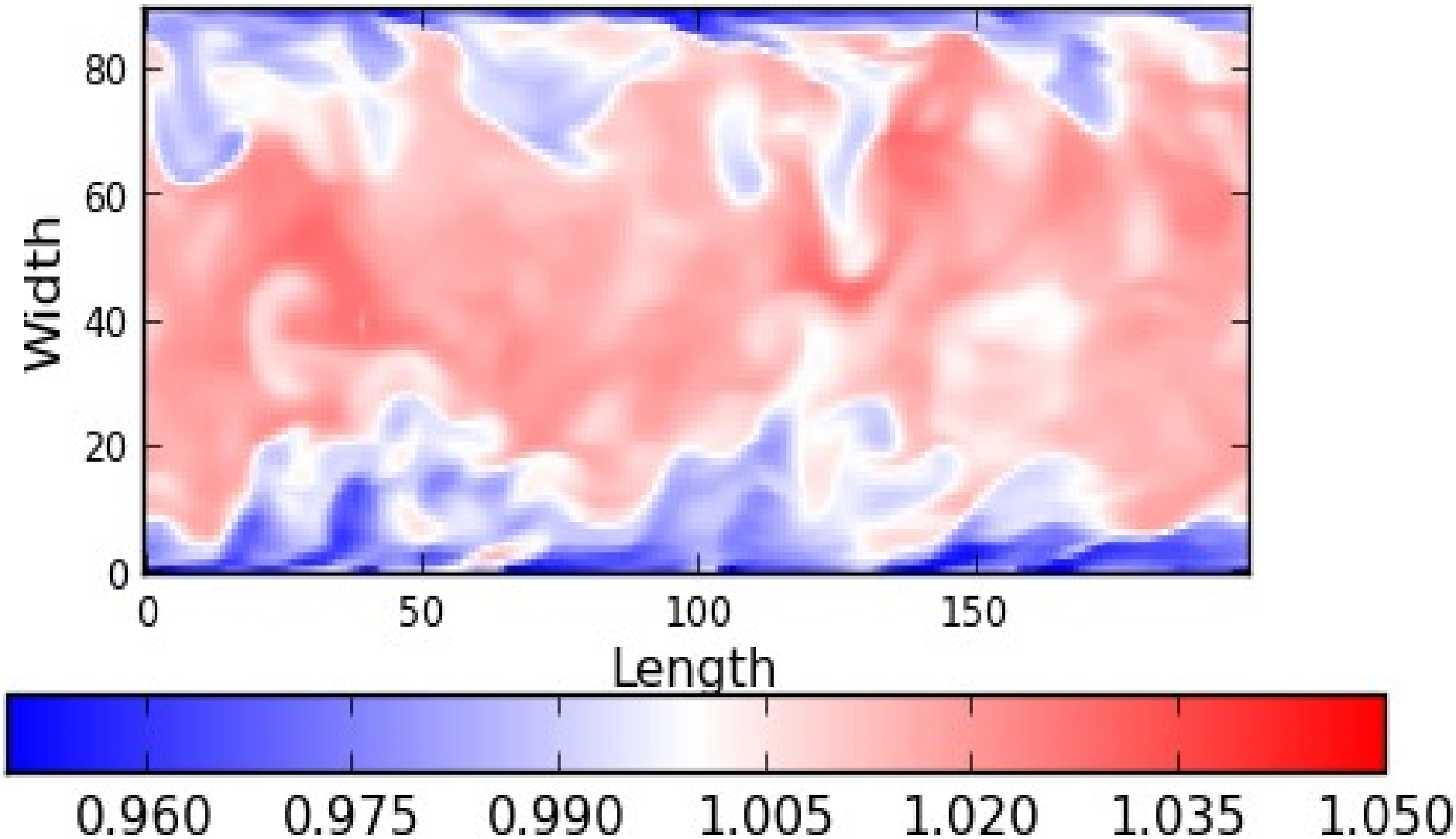
CFD Turbulent Flow (Vorticity)



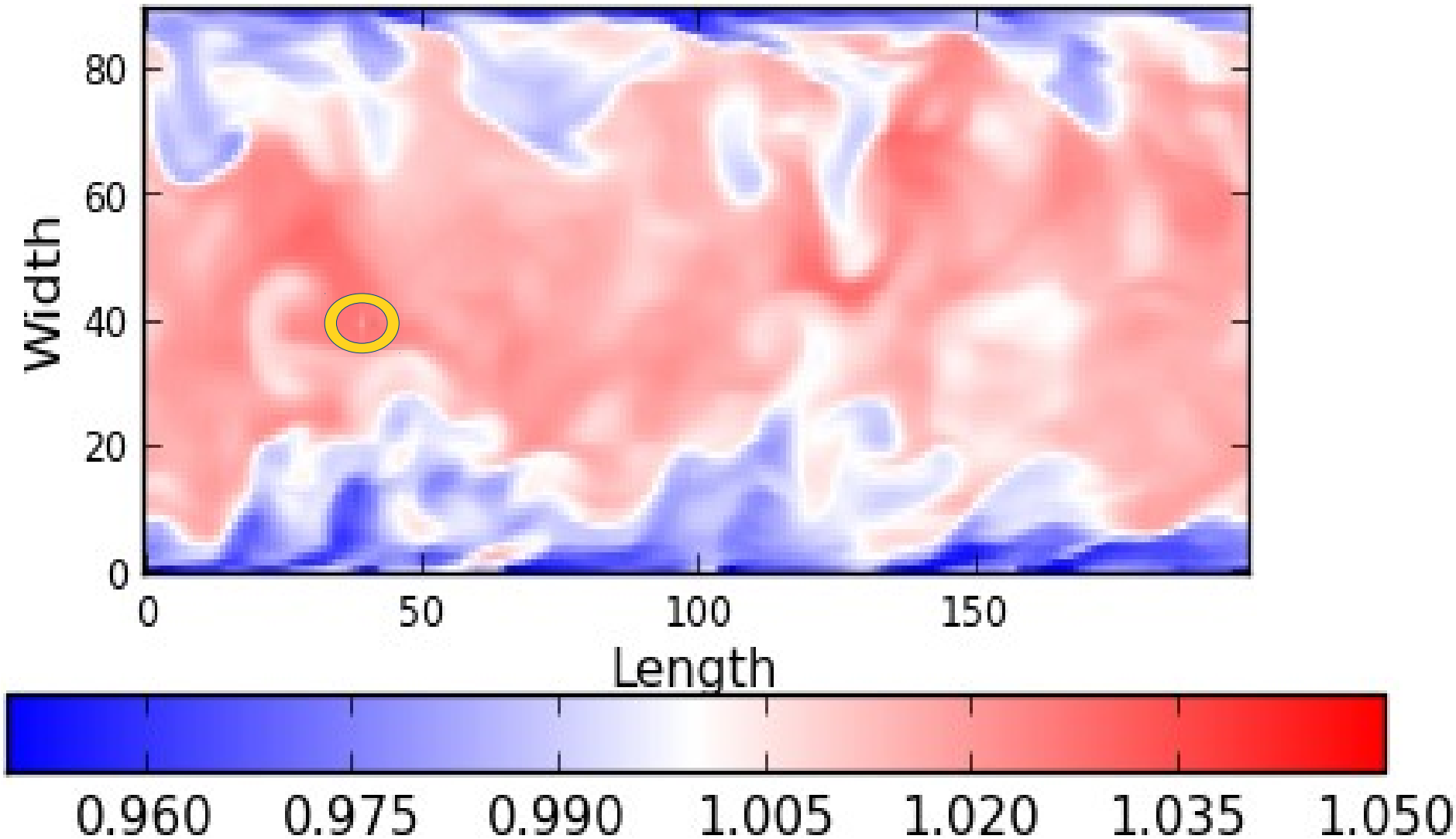
Speed u1


Time step : 15001
Width
80
60
40
20
0
0
50
100
150
Length
0.960
0.975
0.990
1.005
1.020
1.035
1.050

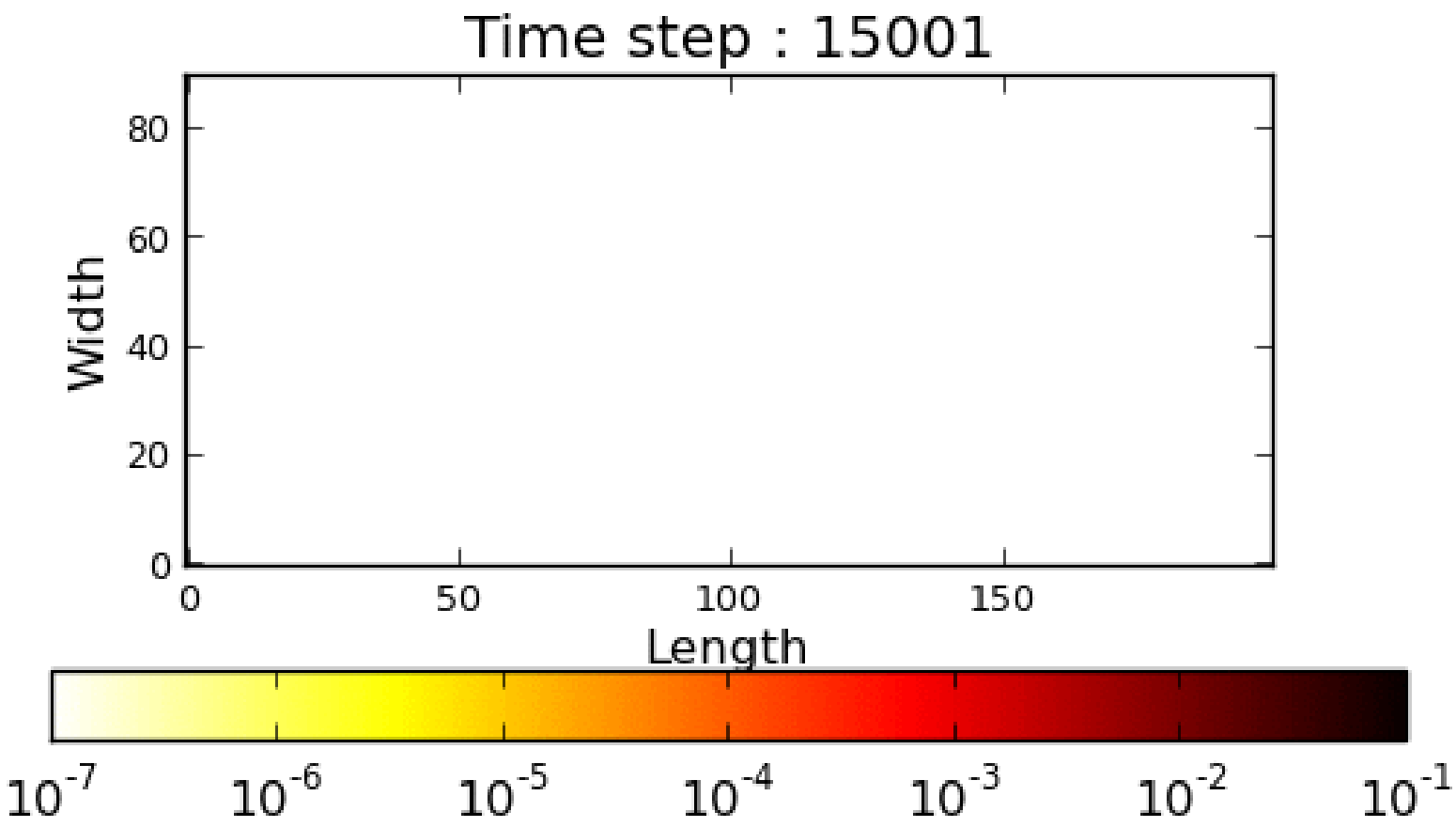
Speed u1


Width
80
60
40
20
0
0
50
100
150
Length
0.960
0.975
0.990
1.005
1.020
1.035
1.050

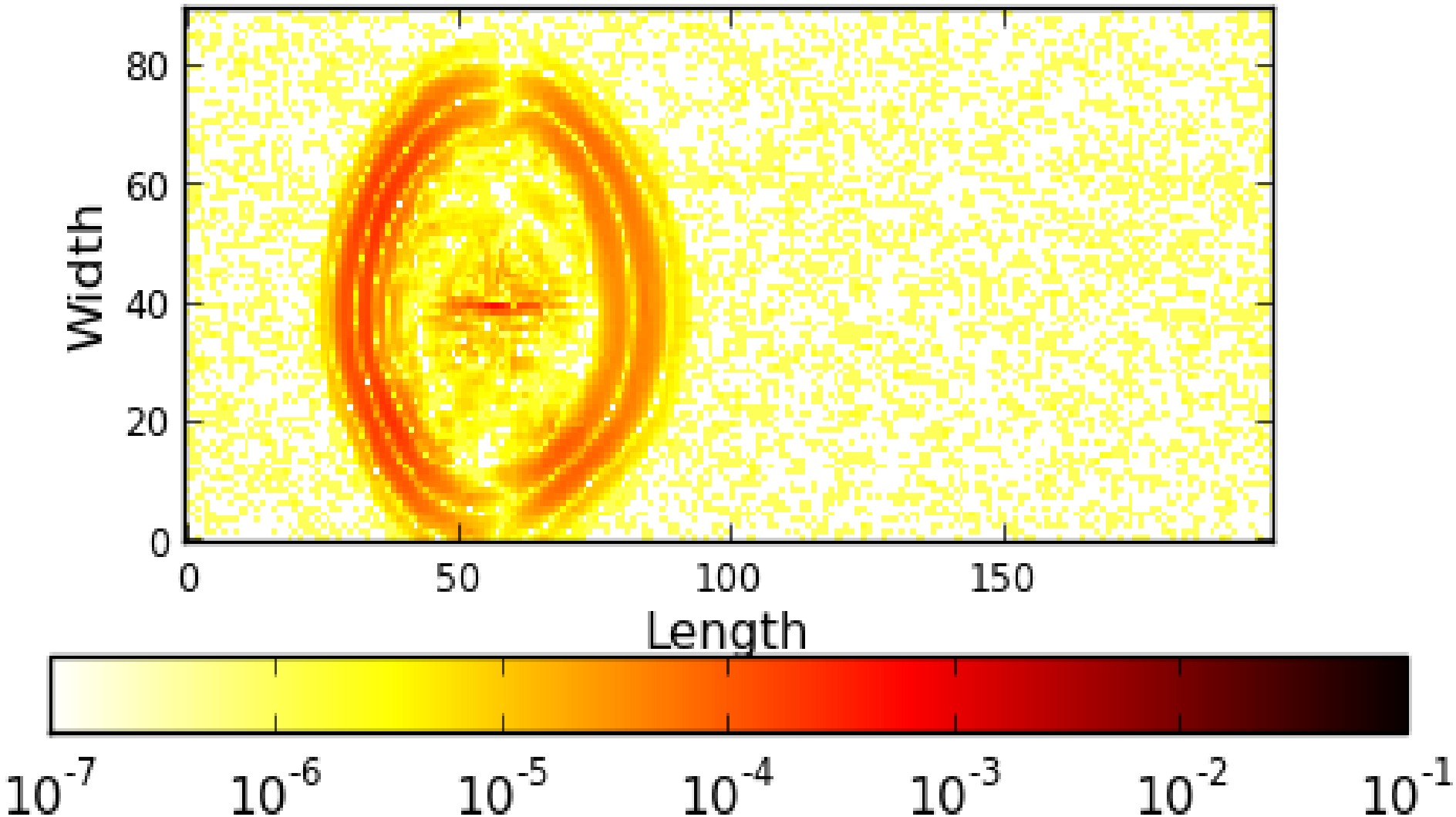
Speed u1


Width
80
60
40
20
0
0
50
100
150
Length
0.960
0.975
0.990
1.005
1.020
1.035
1.050

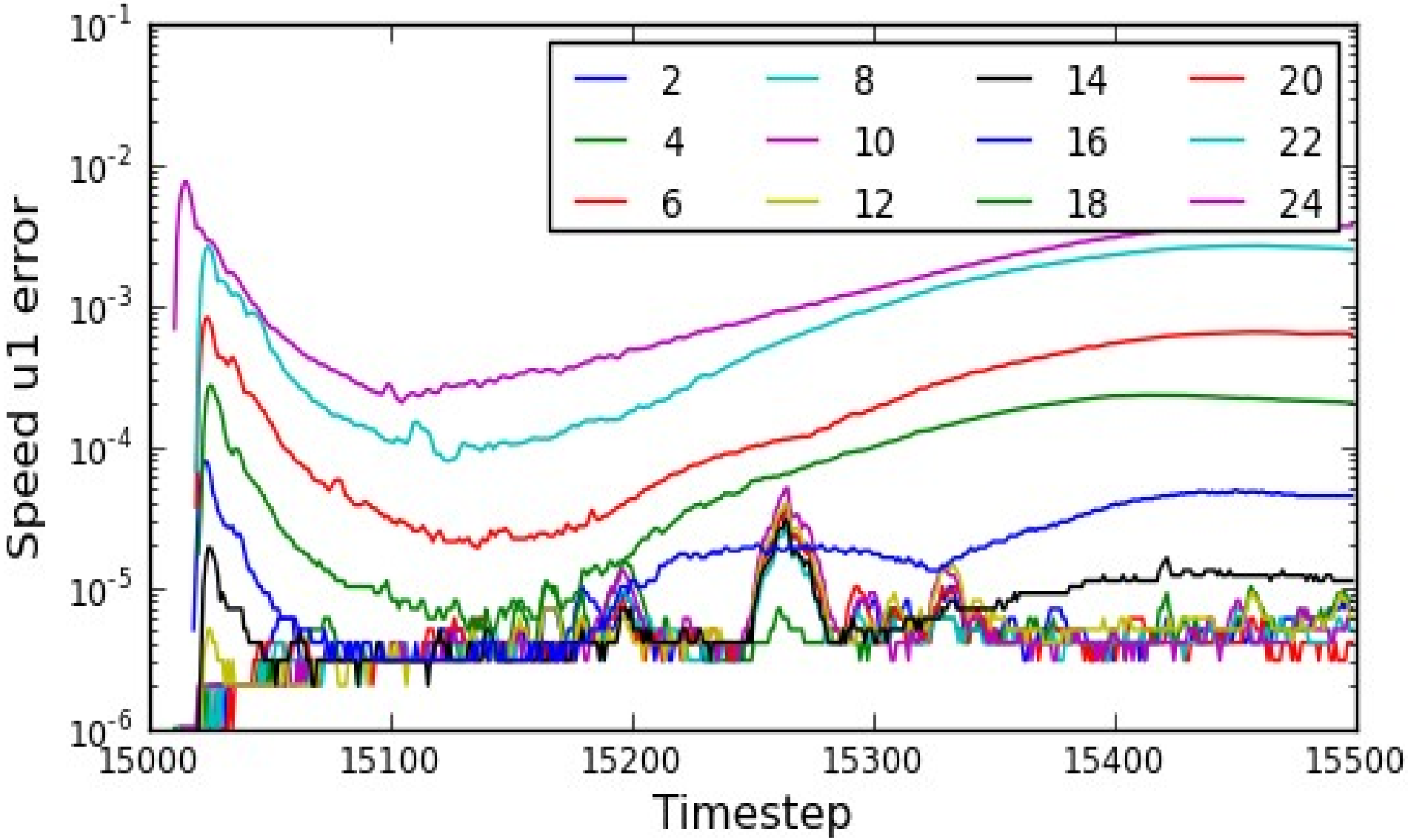
Corruption Propagation



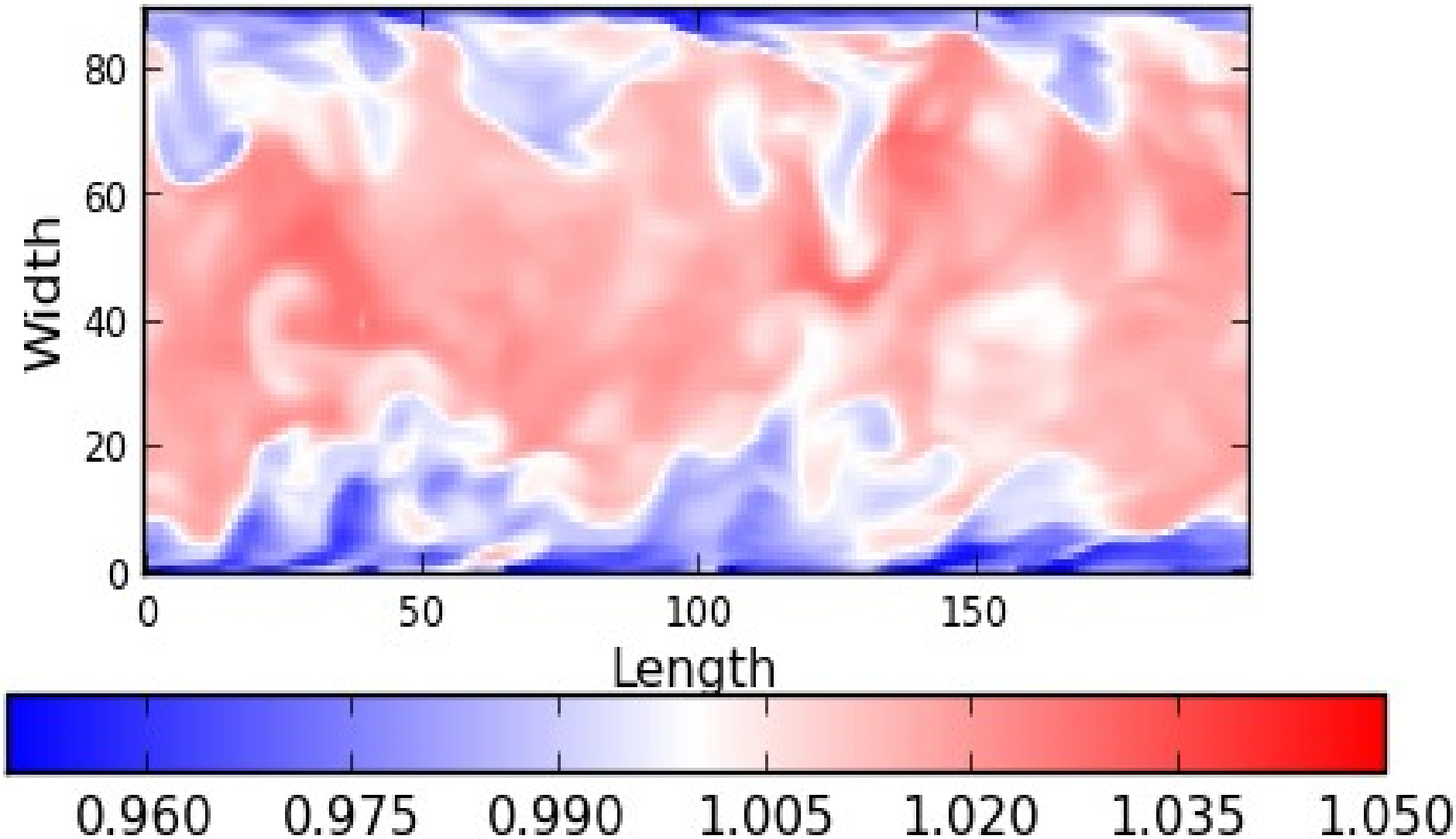
Corruption Propagation



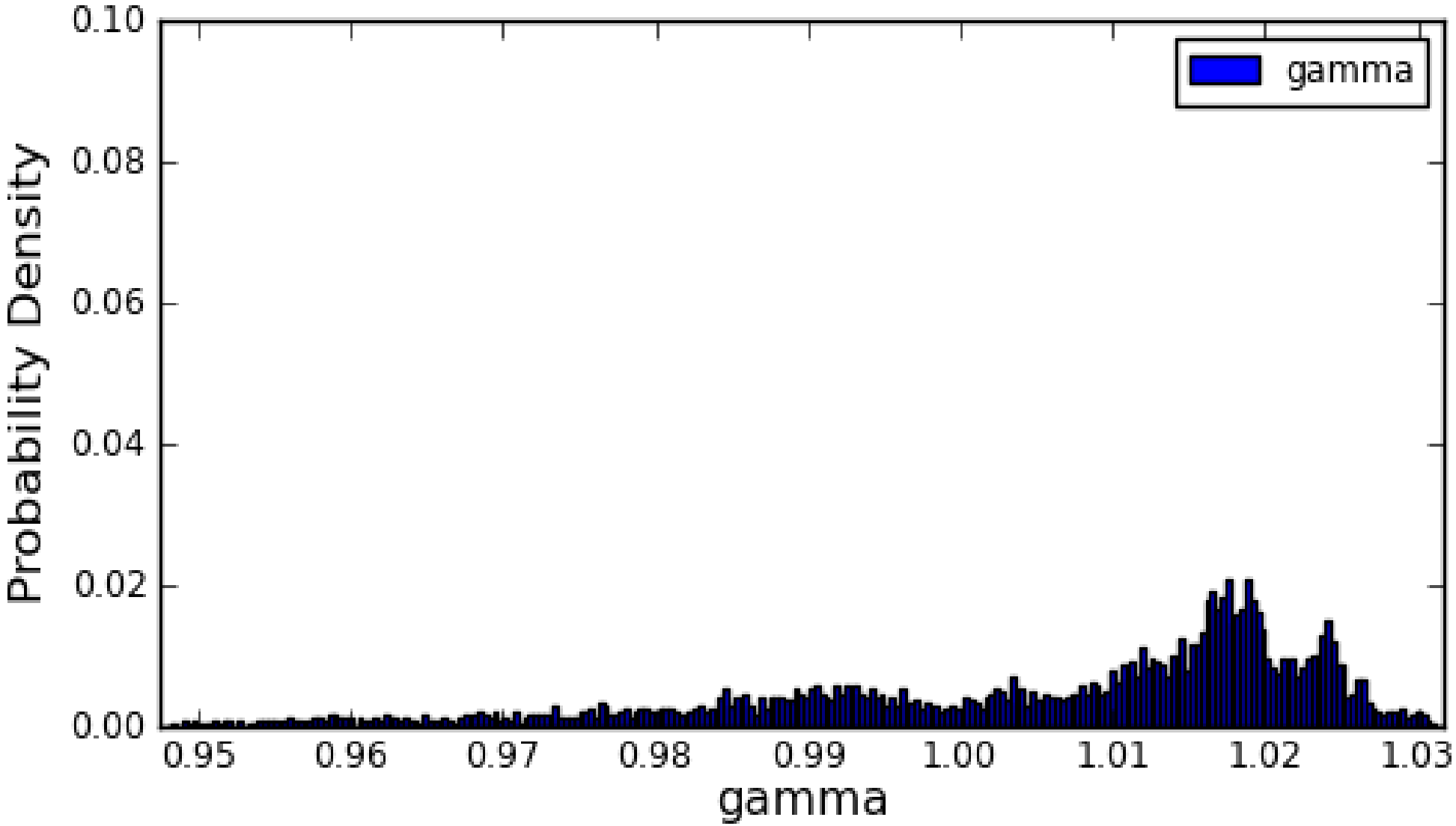
Error Deviation



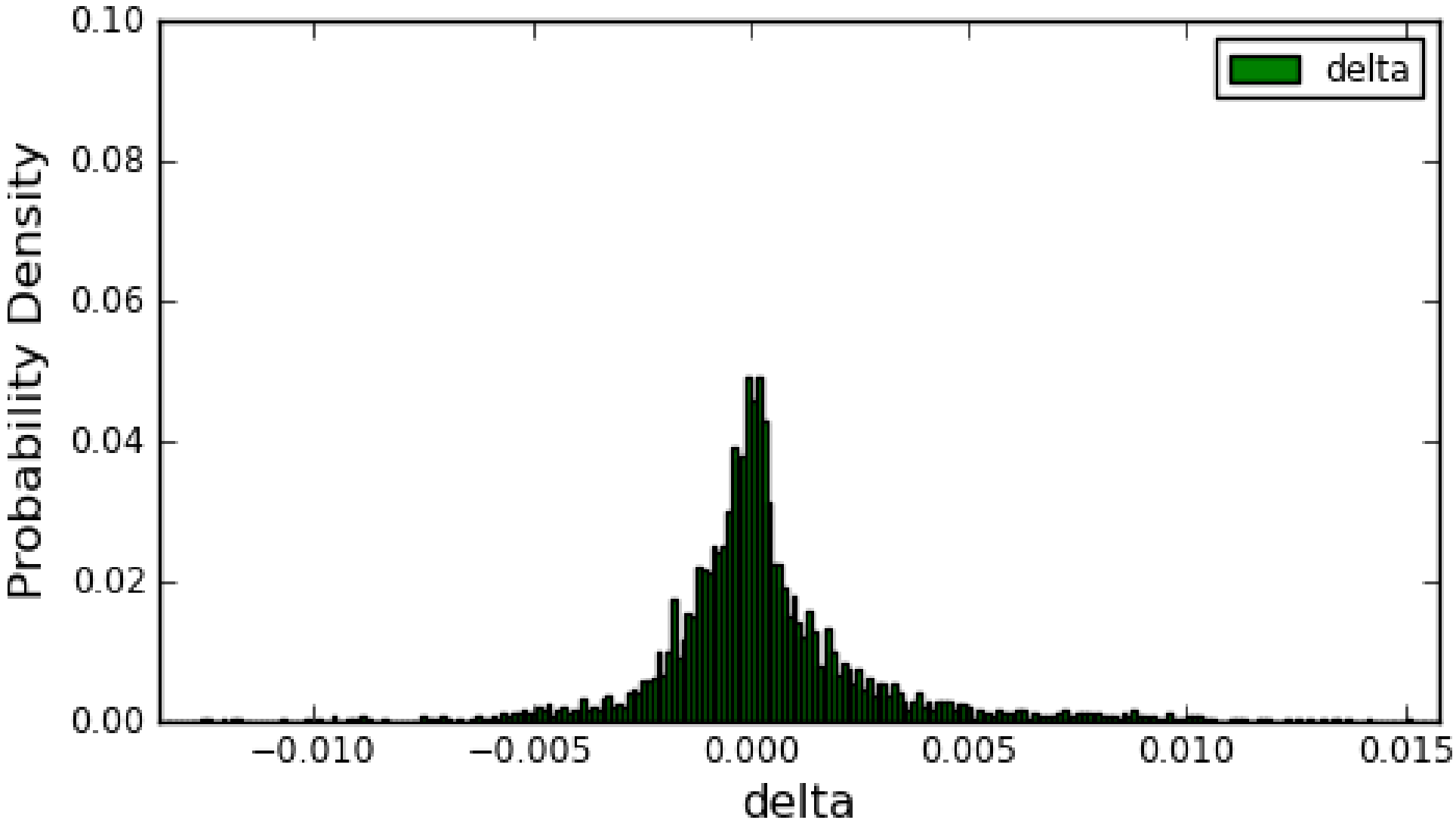
Speed u1


Width
80
60
40
20
0
0
50
100
150
Length
0.960
0.975
0.990
1.005
1.020
1.035
1.050

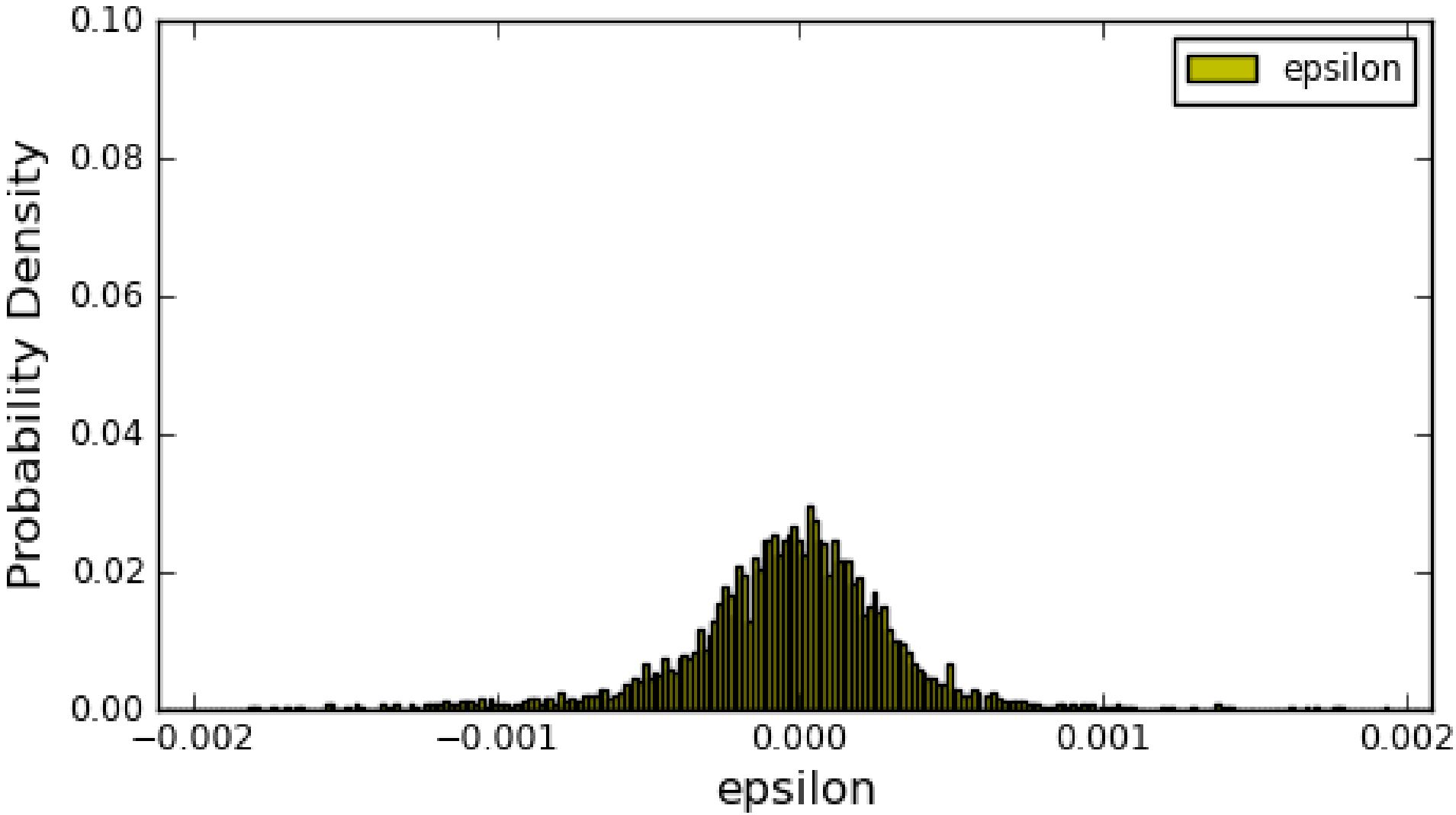
Gamma Detector



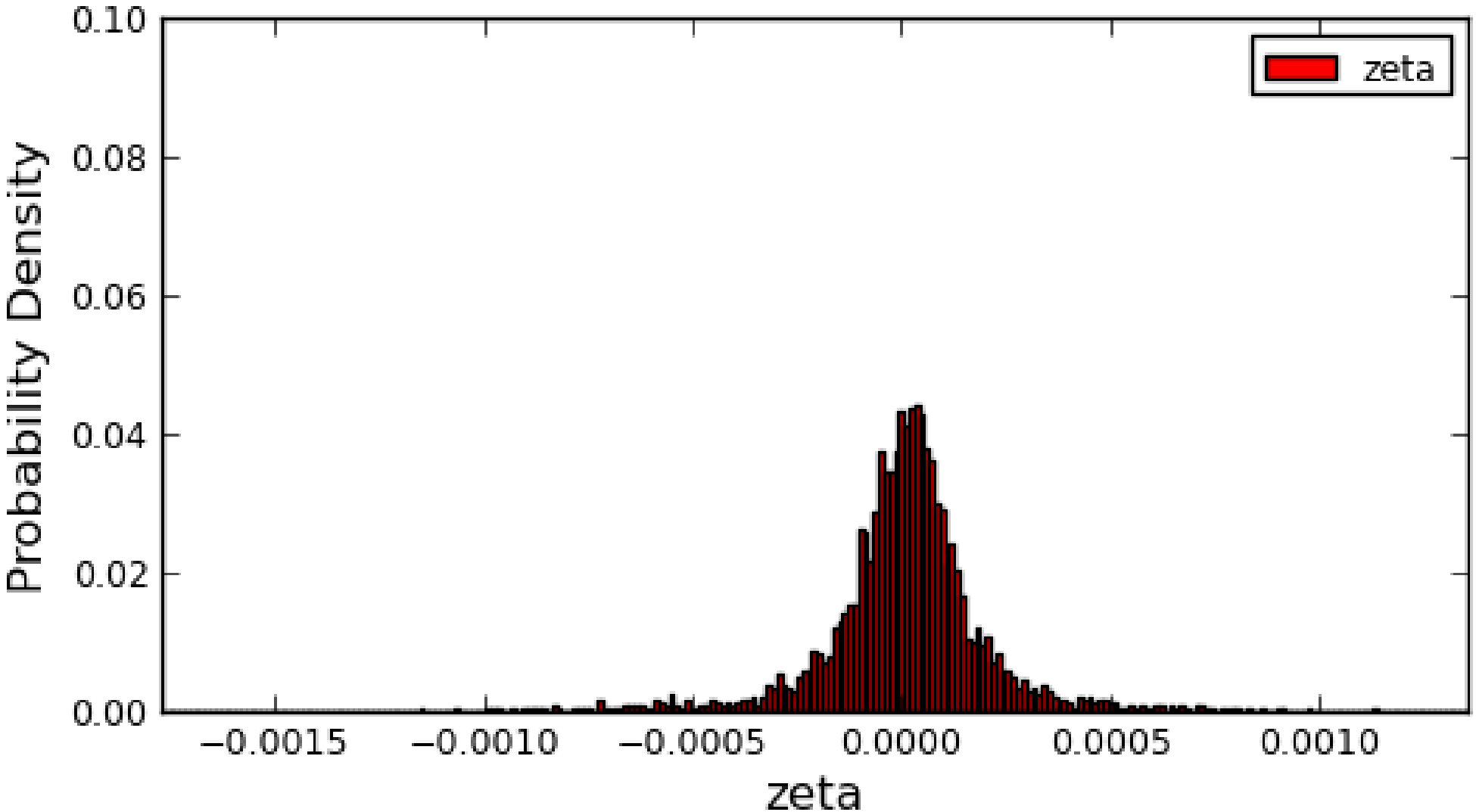
Delta Detector



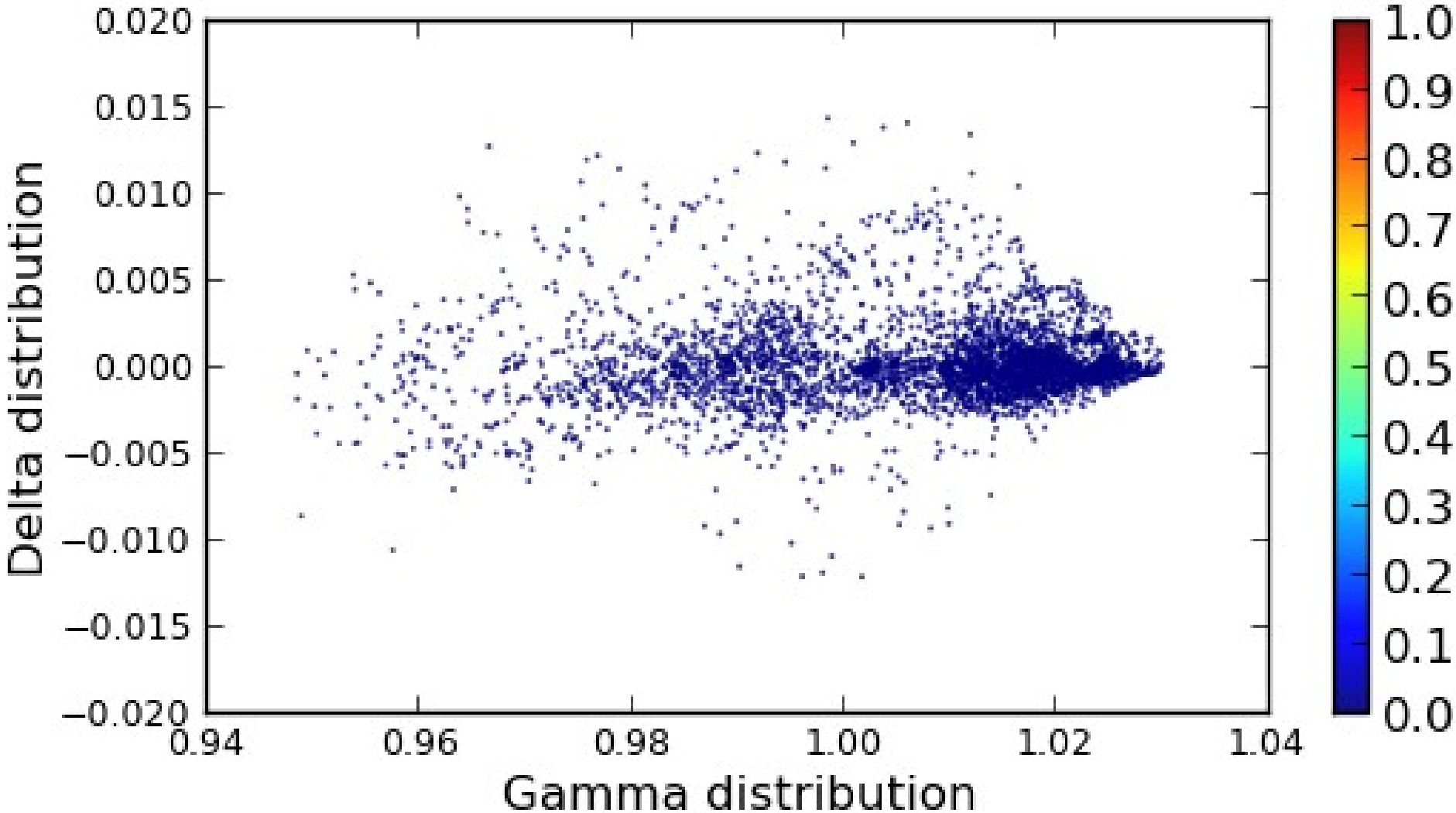
Epsilon Detector



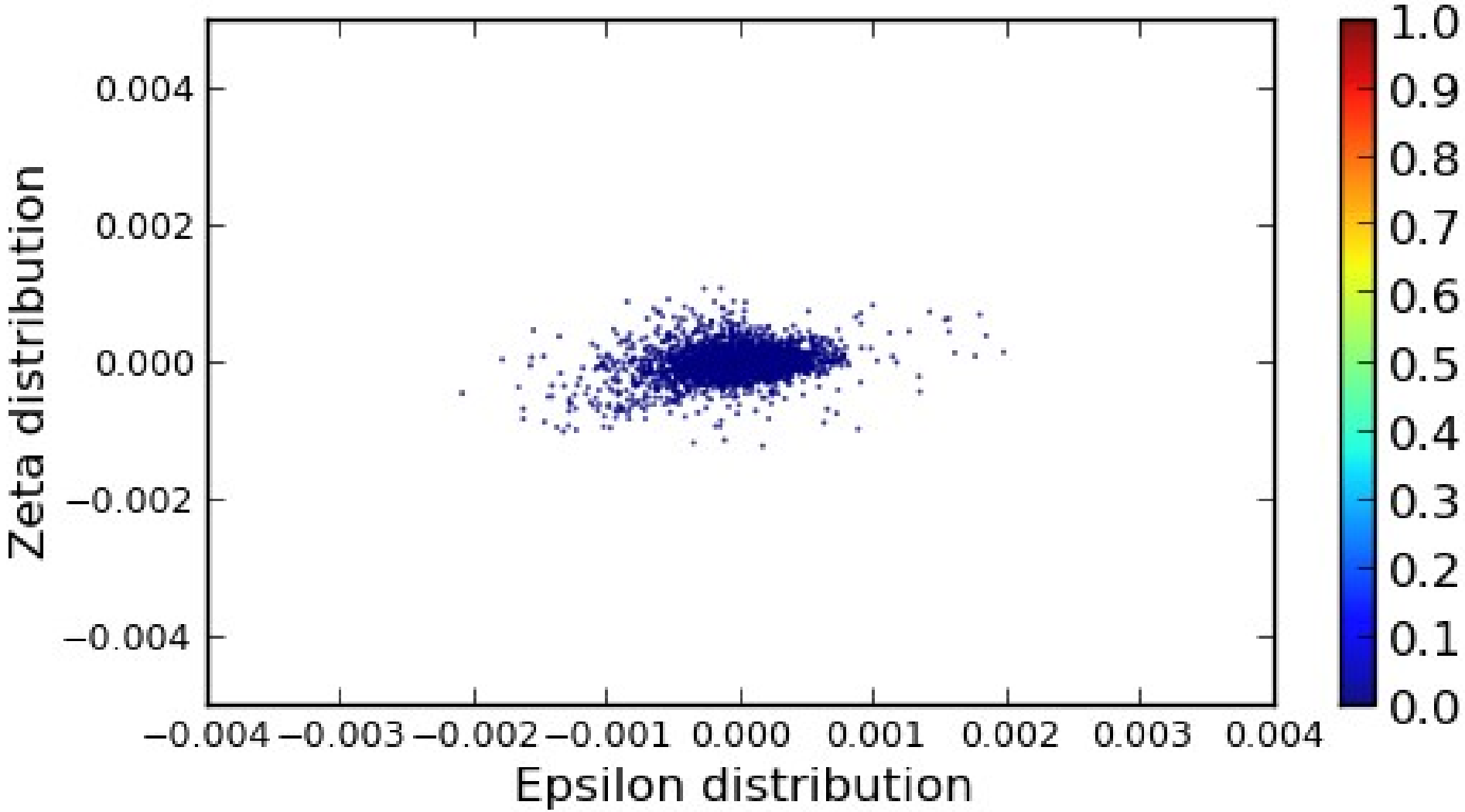
Zeta Detector



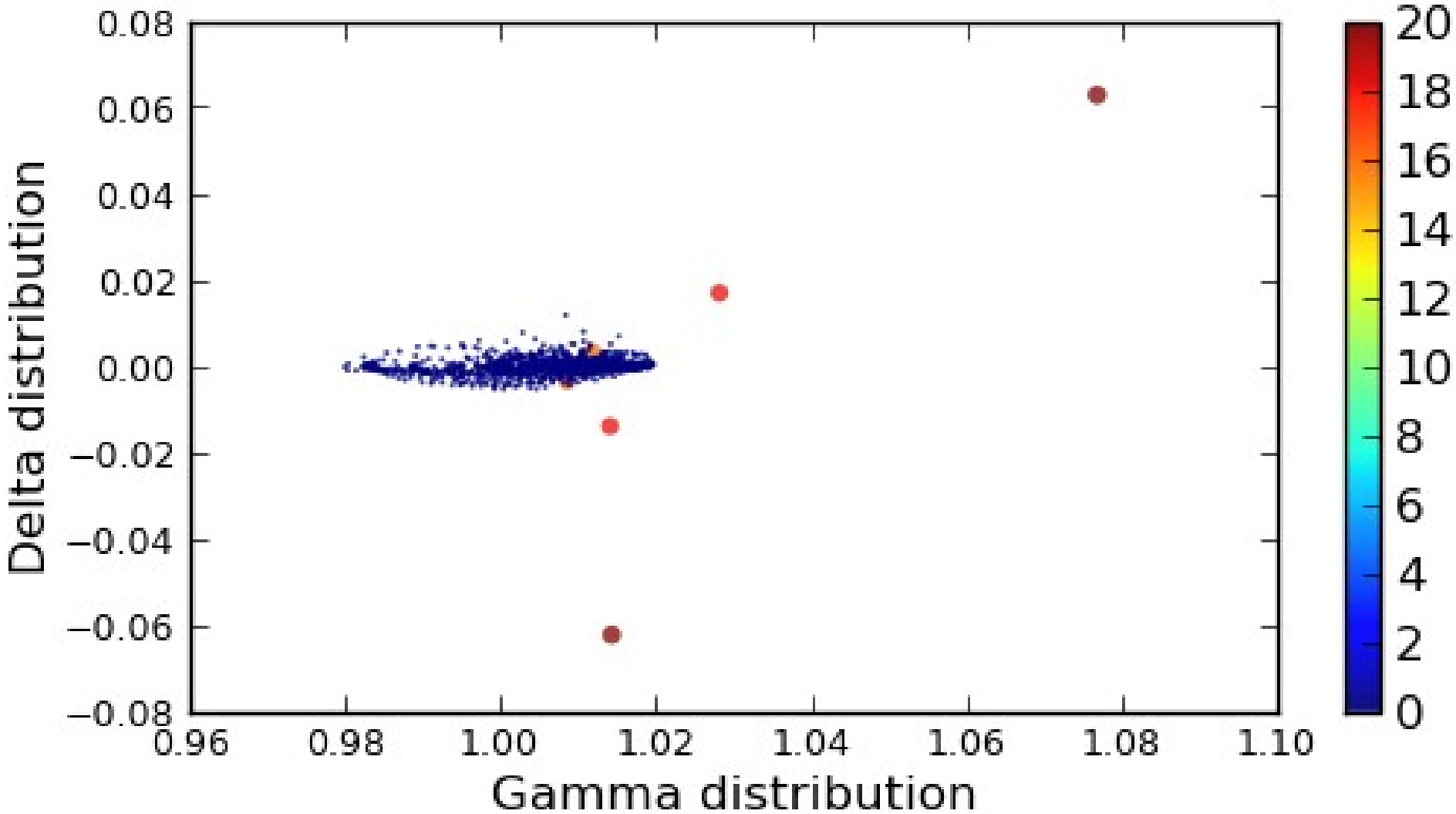
Gamma-Delta Cluster



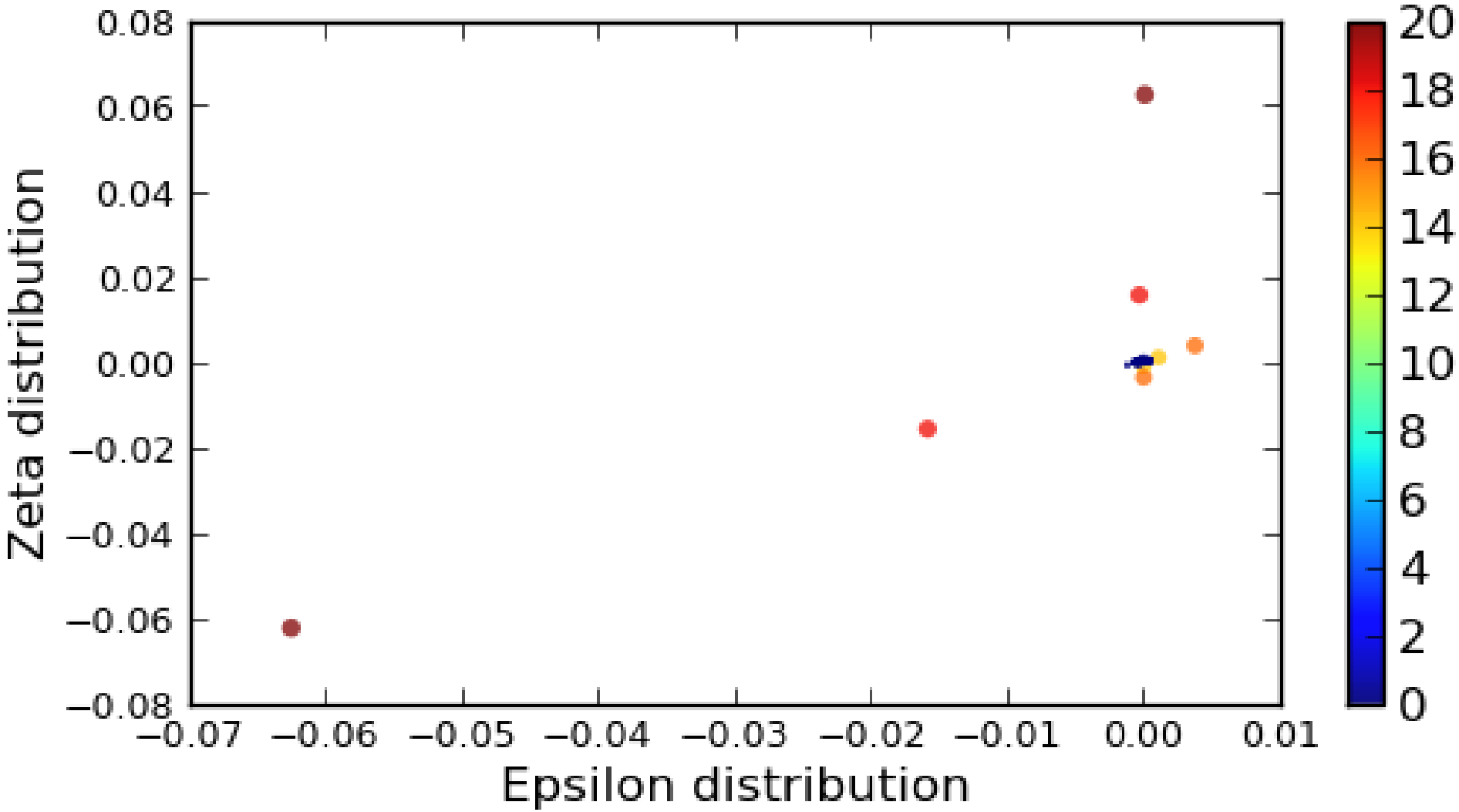
Epsilon-Zeta Cluster



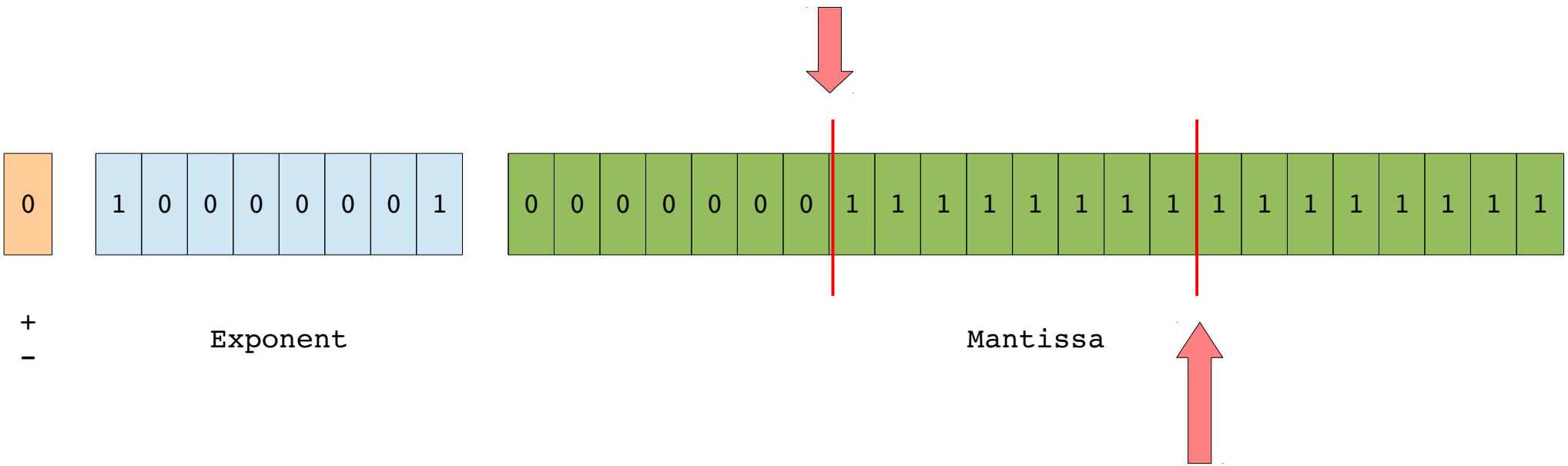
Gamma-Delta Cluster



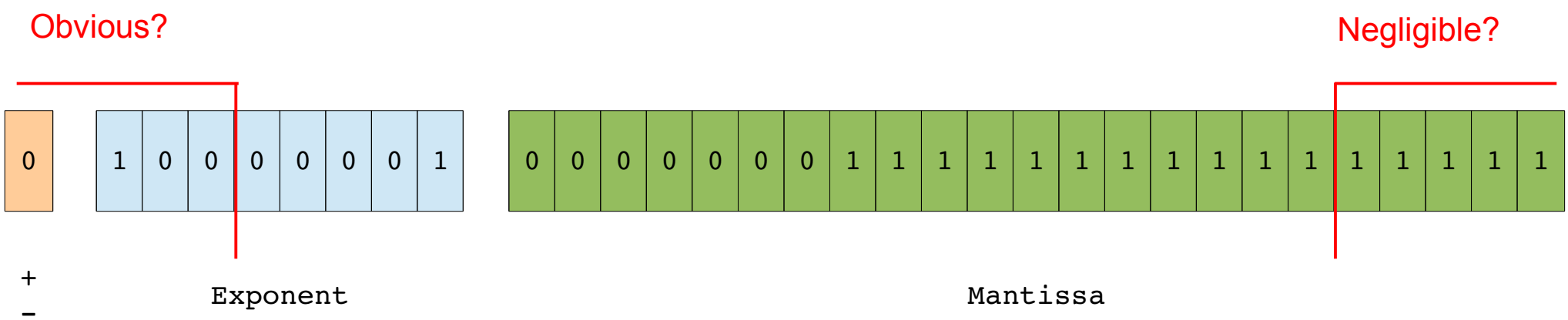
Epsilon-Zeta Cluster



4.000 < 4.0312495 < 4.050 => 16/32 bit-flips detected (50% recall)

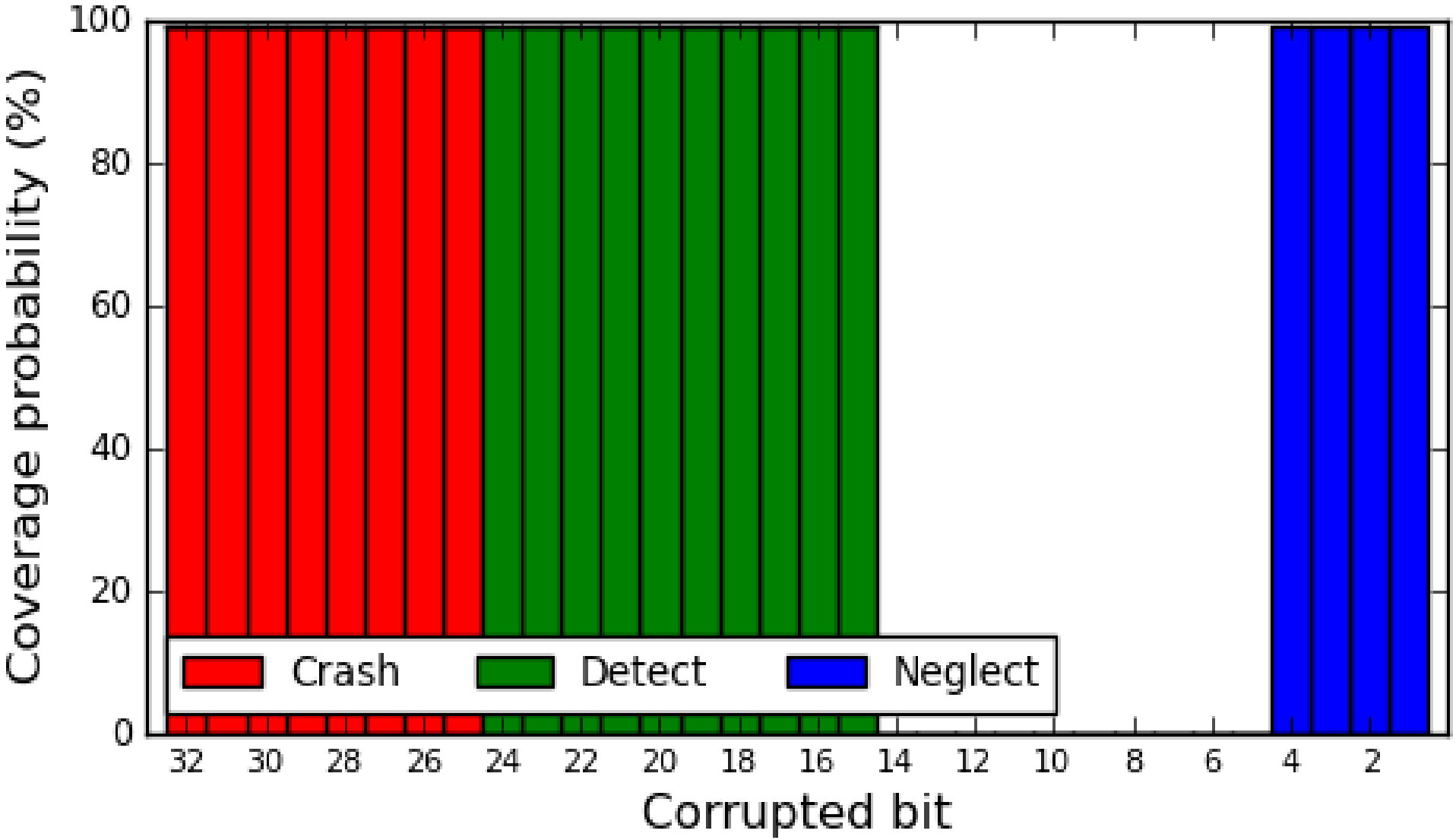


4.03115 < 4.0312495 < 4.03185 => 24/32 bit-flips detected (75% recall)



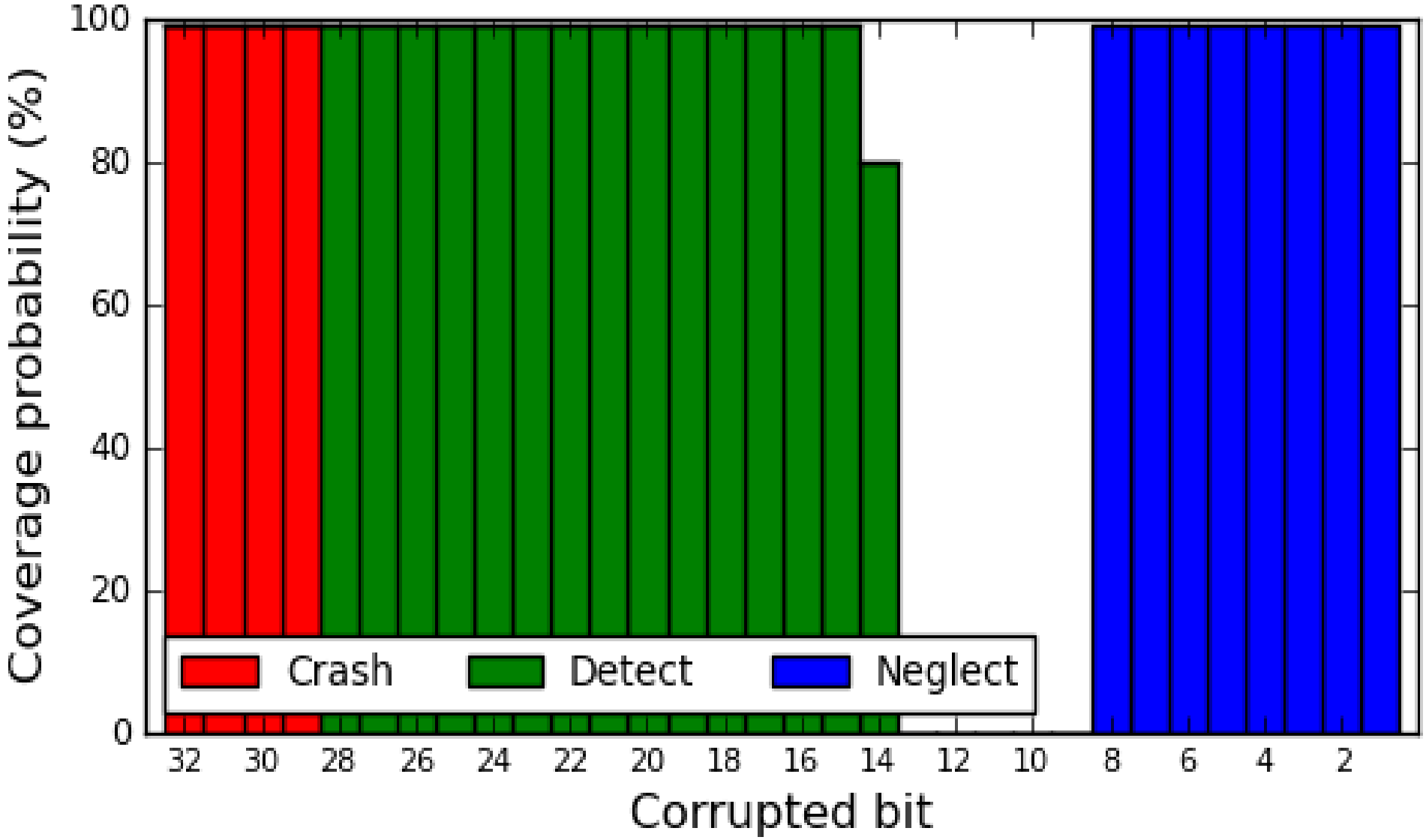
12/24 bit-flips detected (50% coverage) or 20/24 bit-flips detected (83% coverage)

SDC coverage for CFD



Detection recall : 50% => SDC coverage : 68%

SDC coverage for HACC



Detection recall : 70% => SDC coverage : 80%

Thank you!