北师香港浸会大学
BEIJING NORMAL-HONG KONG BAPTIST UNIVERSITY

IAS | INSTITUTE FOR ADVANCED STUDY

BNBU CENTRE FOR COMPUTATIONAL CULTURE AND HERITAGE

UNIVERSITY OF GENEVA
FACULTY OF SCIENCE - Computer Science Dept

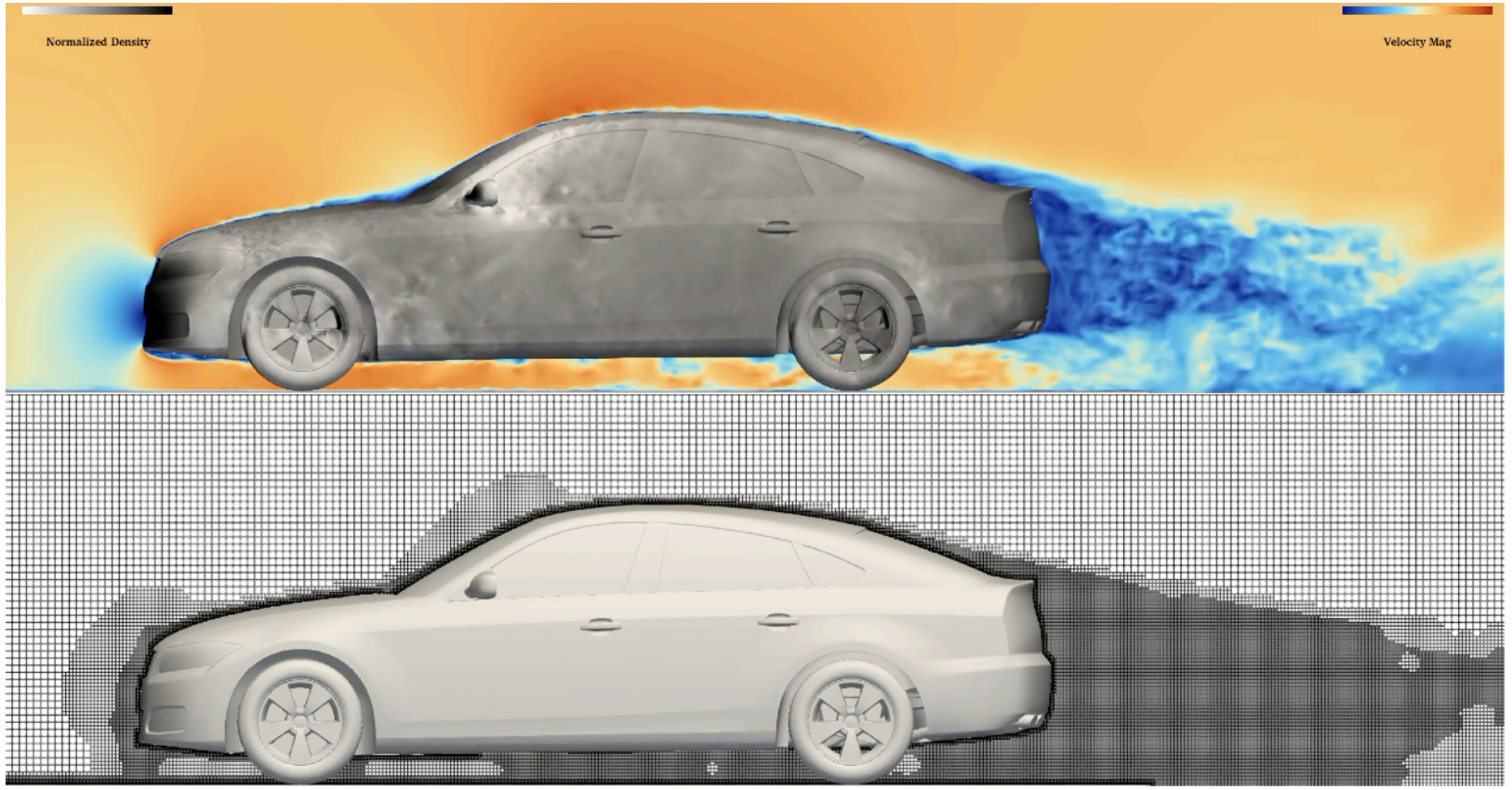
GPU-Accelerated Scale Resolving Simulations for AutoCFD5 on Fixed and Adaptive Octrees



Christophe Coreixas & Jonas Latt

AutoCFD5 Workshop, 8-9 October, Bologna (Italy)

Our target: Overnight realistic simulations (1 GPU)



From CPU clusters to single GPU



Blue Gene/P supercomputer “Intrepid" (2007)
Argonne National Laboratory (**164’000 CPU cores**)

“Brute force”

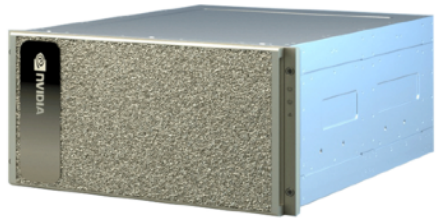
GPU porting



Single rack of NVIDIA GH200 (**72 GPU cards**)
(https://delveloped.nvidia-com/blog)

“GPU-friendly”

numerical method



Single GPU node or DGX-like desktop computer

Low storage approach
Single precision resilience



Single (gaming) GPU

From CPU clusters to single GPU



Blue Gene/P supercomputer “Intrepid" (2007)
Argonne National Laboratory (**164’000 CPU cores**)

100-1000x costs reduction

Anybody can run these simulations
(Democratization / Green Scientific Computing)

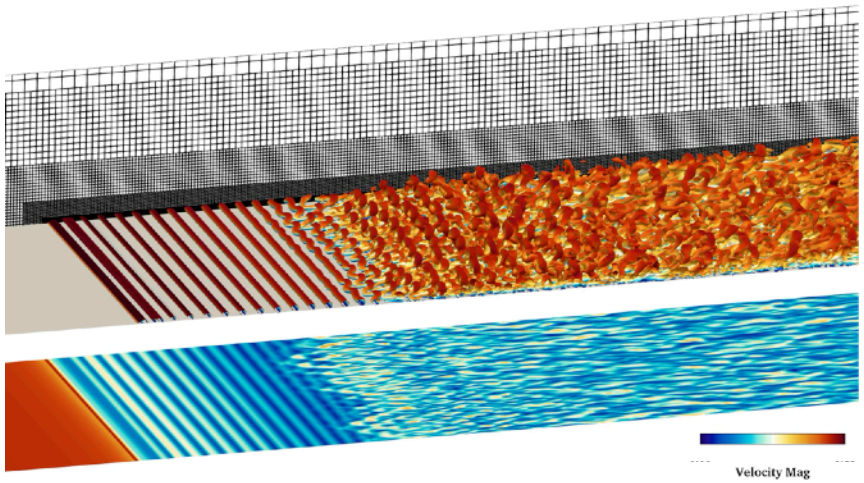


Single (gaming) GPU

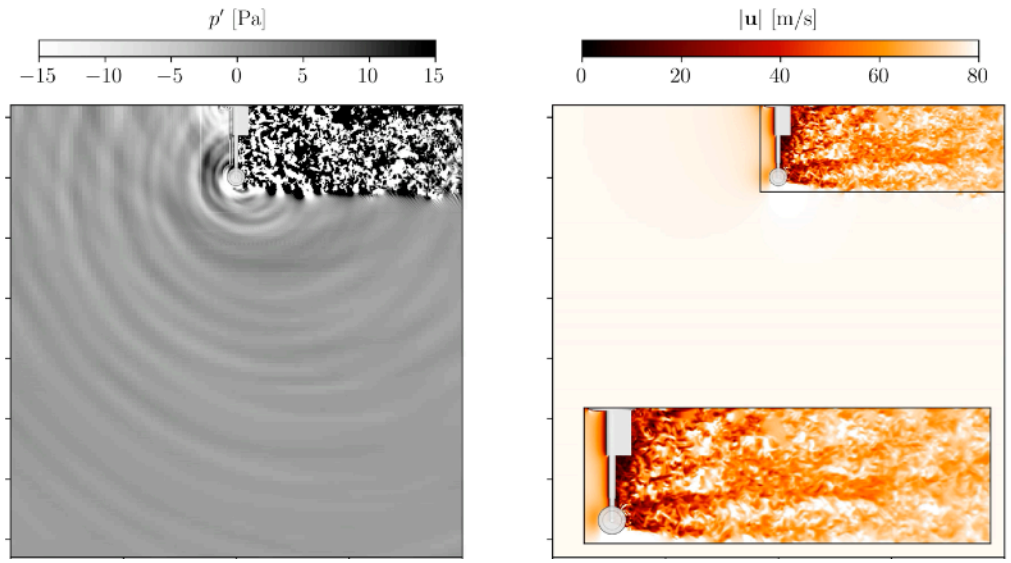
Octoplus

Usually low Mach

- **General Octree solver:** Each leaf can have its own dynamics (e.g., LB, FD and kinetic)



Boundary layer transition



Landing-gear aeroacoustics

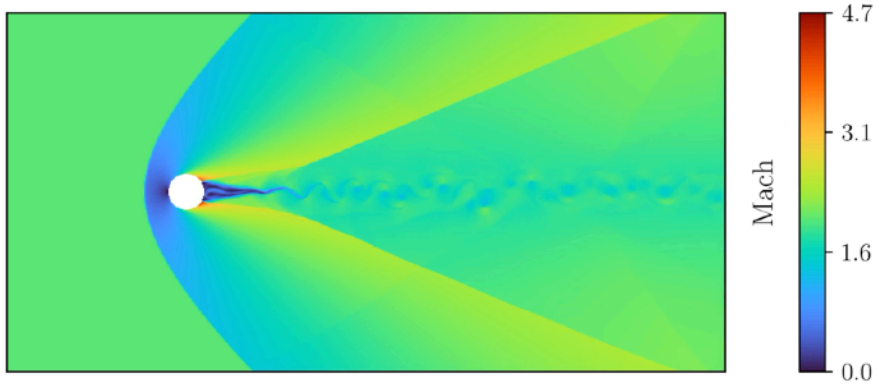
Octoplus

High Mach / Multiphysics (on-going)

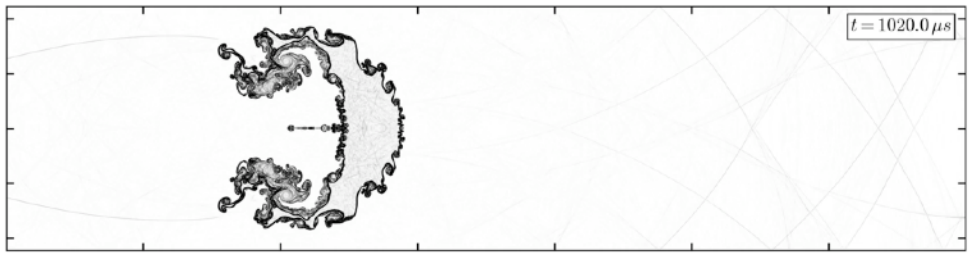
Rarefied (experimental)

- **General Octree solver:** Each leaf can have its own dynamics (e.g., LB, FD and kinetic)

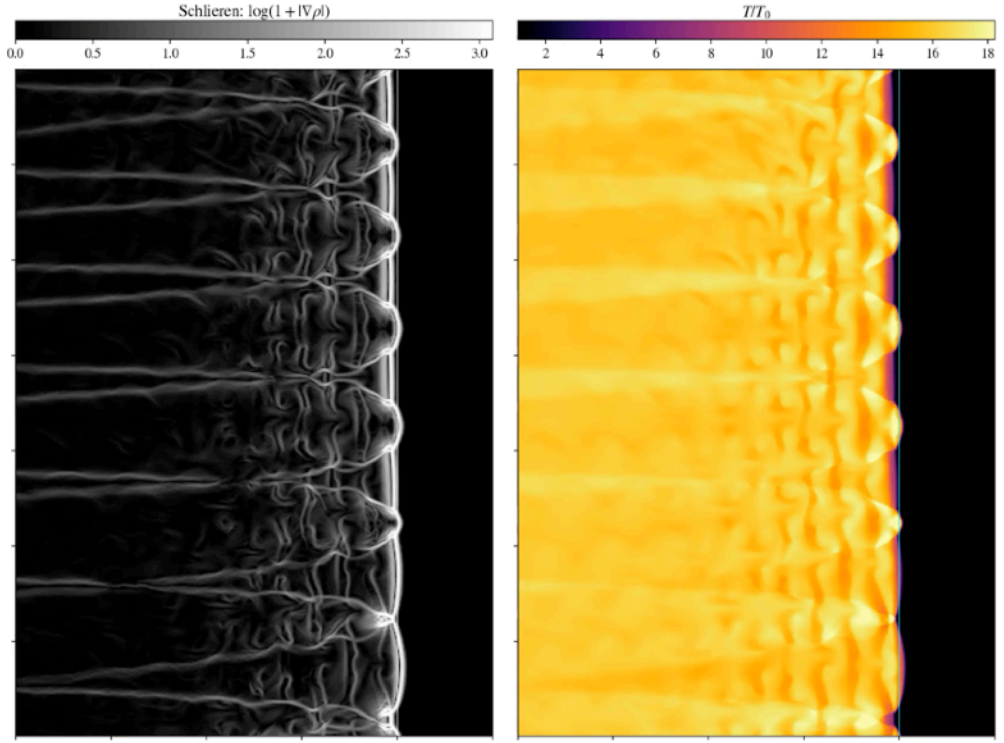
Supersonic flows



Multicomponent flow (shock-bubble)



Multispecies reactive flow (detonation)

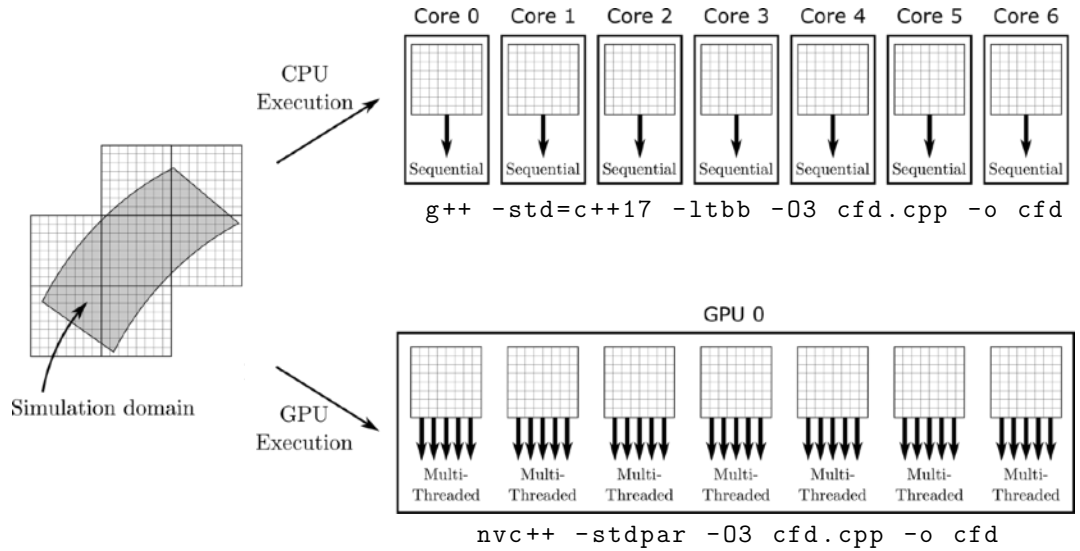


Octoplus

- **General Octree solver:** Each leaf can have its own dynamics (e.g., LB, FD and kinetic)
- **Written in modern C++:** The same code runs on CPUs and GPUs

```
for (int iX=0; iX < nx; ++iX) {
    for (int iY=0; iY < ny; ++iY) {
        for (int iZ=0; iZ < nz; ++iZ) {
            cell(iX, iY, iZ).Solver();
        }
    }
}
```

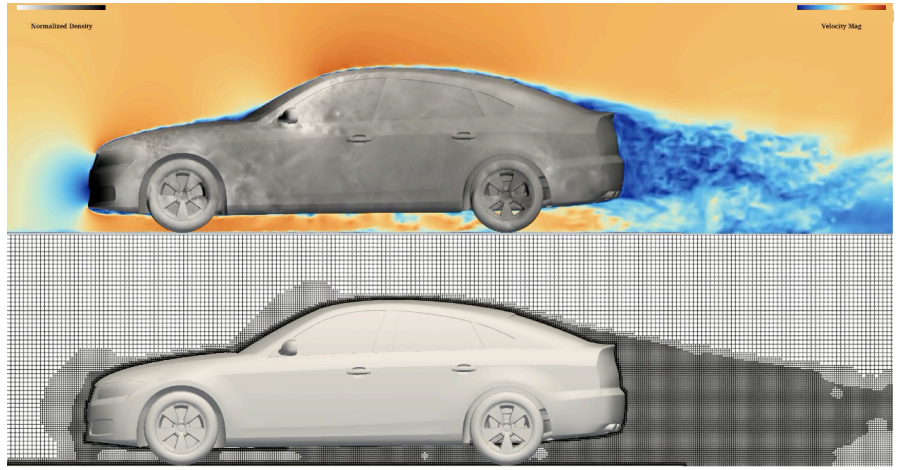
```
for_each(execution::par,
         begin(data), end(data), [](Cell& cell))
{
    cell.Solver();
} );
```

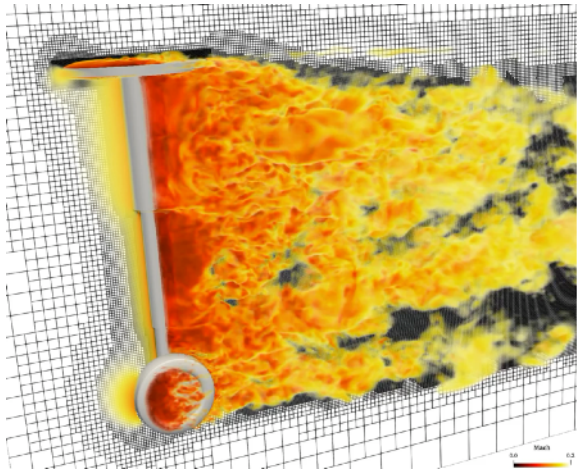


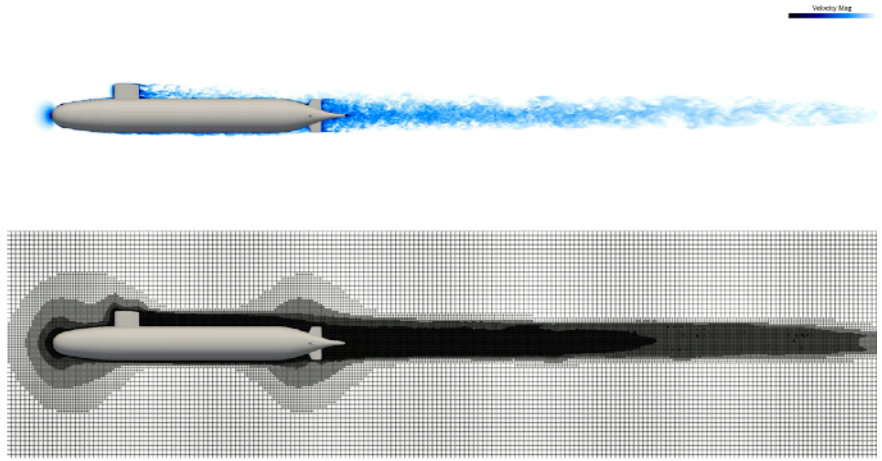
Latt et al., "Cross-platform programming model for many-core lattice Boltzmann simulations", PLoS One, **2021**
Latt & Coreixas, "Multi-GPU acceleration of PALABOS fluid solver using C++ standard parallelism", *Comput. Phys. Commun.*, **2026**
Coreixas & Latt, "Next-generation digital twins: C++ based solvers across CPUs and GPUs", *The PASC26 Conference*, Bern, **2026**

Octoplus

- **General Octree solver:** Each leaf can have its own dynamics (e.g., LB, FD and kinetic)
- **Written in modern C++:** The same code runs on CPUs and GPUs
- **Adaptive mesh refinement:** Transient speedup, wake tracking, etc
- **Additional features:** Cell-centered, local time step, Smago sgs, eq. wall law (since August!)







Coreixas & Latt, "Democratizing Scale-Resolving Aerodynamics with GPU-based CFD on Adaptive Octrees", *60th 3AF (AERO2026)*, 2026
Coreixas & Latt, "Democratizing computational aeroacoustics using GPU-accelerated Octree CFD solvers", *32nd AIAA/CEAS Aeroacoustics Conference*, 2026

Octoplus

- **General Octree solver:** Each leaf can have its own dynamics (e.g., LB, FD and kinetic)
- **Written in modern C++:** The same code runs on CPUs and GPUs
- **Adaptive mesh refinement:** Transient speedup, wake tracking, etc
- **Additional features:** Cell-centered, local time step, Smago sgs, eq. wall law (since August!)

- **Extended modularity:** Decoupling of memory layout, parallelization, bulk physics, BCs, etc
- **Low memory storage + FP32-resilient:** 5-10 M cells per GB of RAM
- **CUDA-comparable performance** on NVIDIA GPUs
- **Targeting cross-vendor portability**

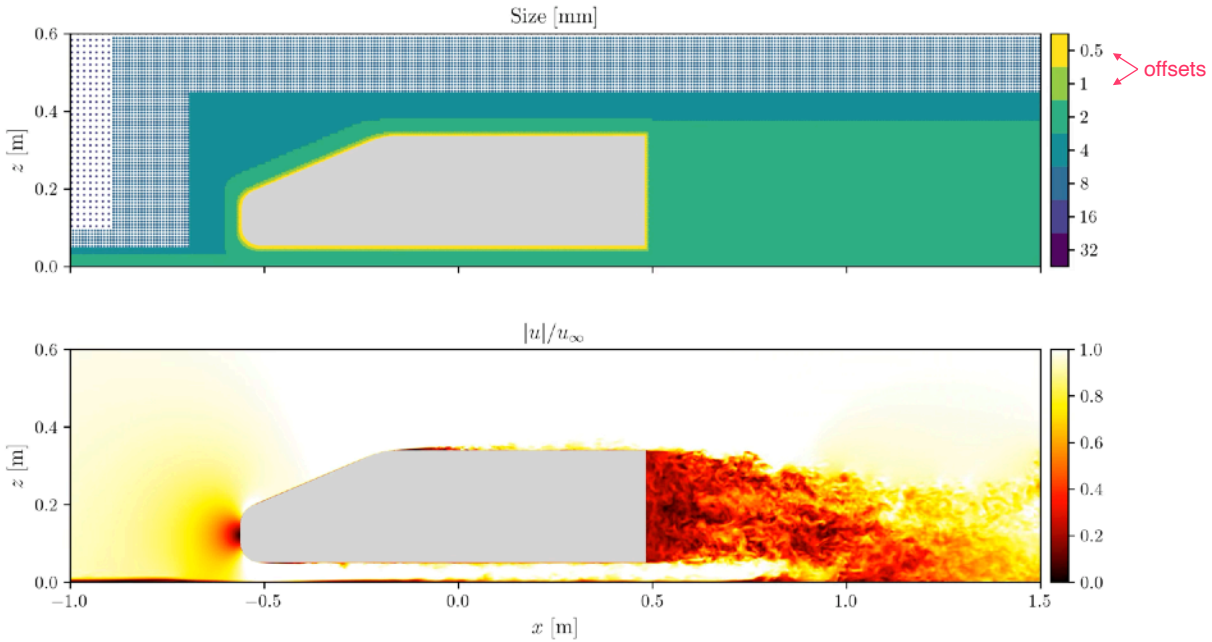
Latt et al., "Cross-platform programming model for many-core lattice Boltzmann simulations", PLoS One, **2021**
Latt & Coreixas, "Multi-GPU acceleration of PALABOS fluid solver using C++ standard parallelism", *Comput. Phys. Commun.*, **2026**
Coreixas & Latt, "Democratizing Scale-Resolving Aerodynamics with GPU-based CFD on Adaptive Octrees", *60th 3AF (AERO2026)*, 2026
Coreixas & Latt, "Democratizing computational aeroacoustics using GPU-accelerated Octree CFD solvers", *32nd AIAA/CEAS Aeroacoustics Conference*, 2026

60 CTUs AutoCFD5 contributions (#99)

Windsor w/o wheels at 2.5° yaw (Case 1)

RTX5090 (1$/h on Runpod)

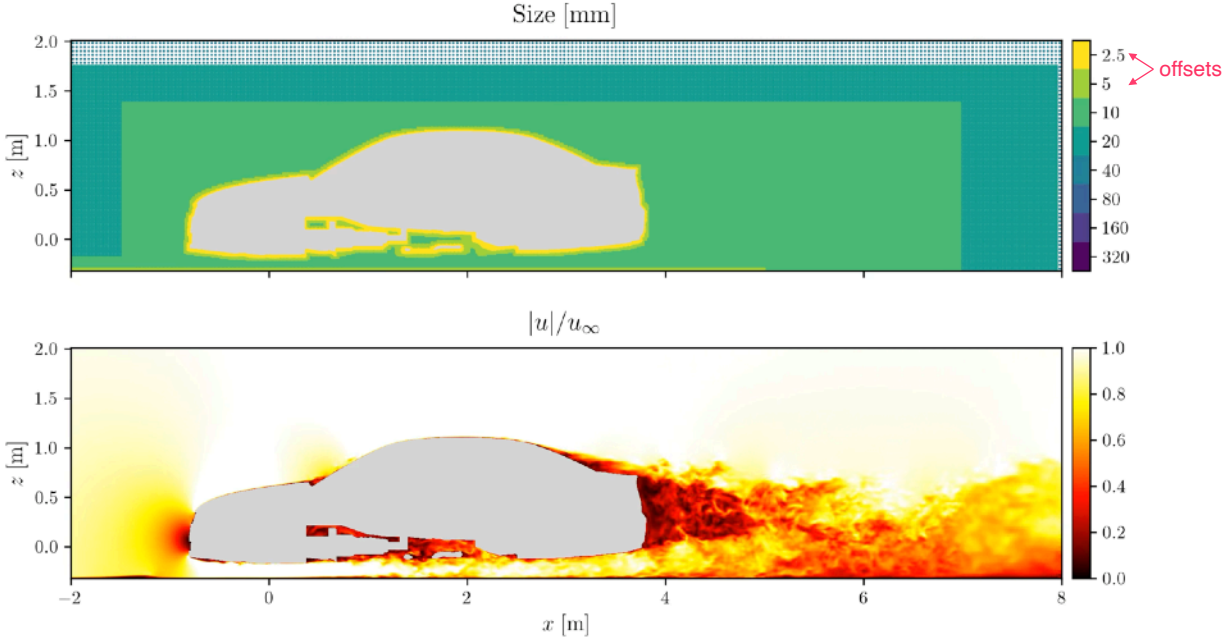
Mesh	Grid (10^6)	Δx_f[mm]	Iterations	Runtime [h]
Coarse (A)	77	2.0	285′000	0.9
Medium (B)	94	1.0	570′000	1.9
Fine (C)	132	0.5	1′140′000	7.8



DrivAer w & w/o FD (Case 2a & 2b)

RTX5090 (1$/h on Runpod)

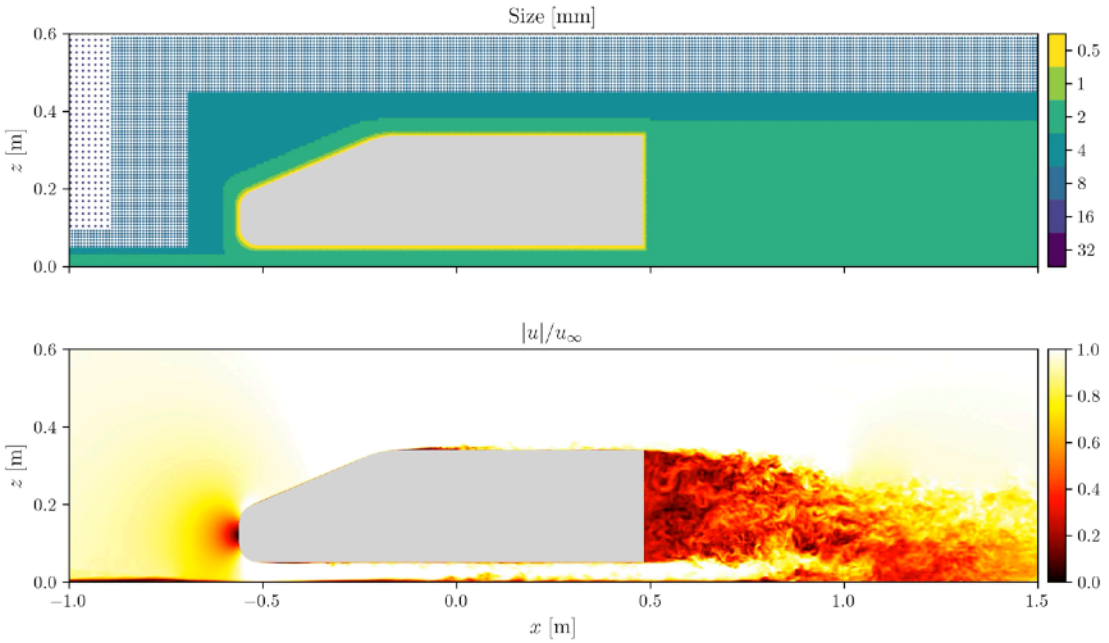
Mesh	Grid (10^6)	Δx_f[mm]	Iterations	Runtime [h]
Coarse	13	10.0	257′000	0.2
Medium	87	5.0	513′000	2.1
Fine	129	2.5	1′026′000	8.1

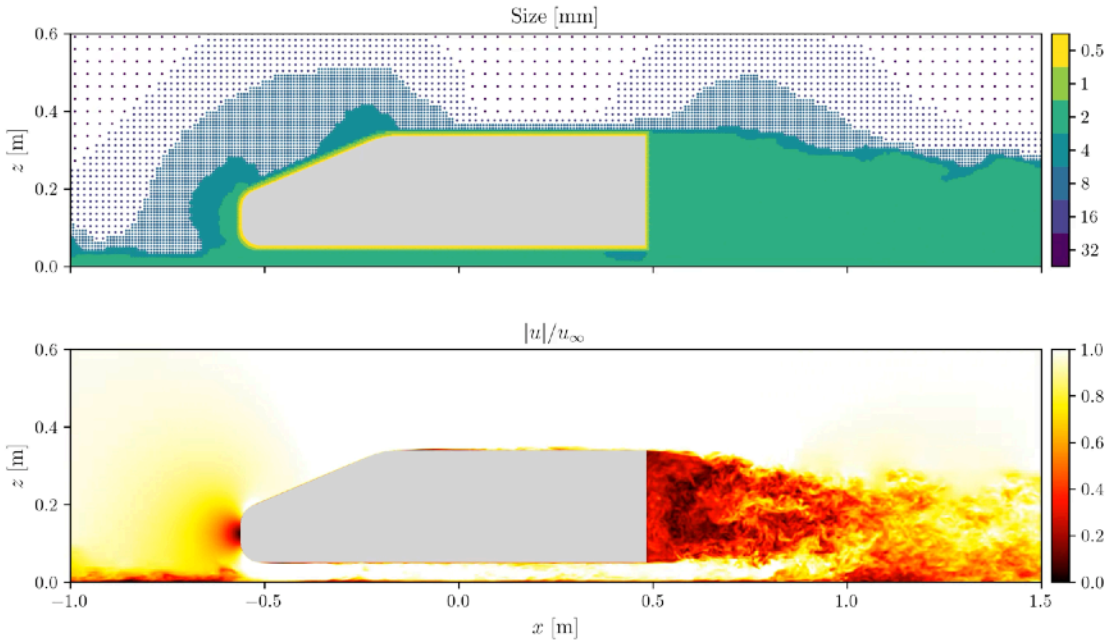


Additional (on-going) studies

- AMR for fully automated design process?
- Impact of precision (FP32 vs FP32/FP64 vs FP64)
- Windsor at different yaw angles (0°, 2.5°, 5°, 10°)
- Design impact on aero of DrivAer (notch vs fast vs estate)

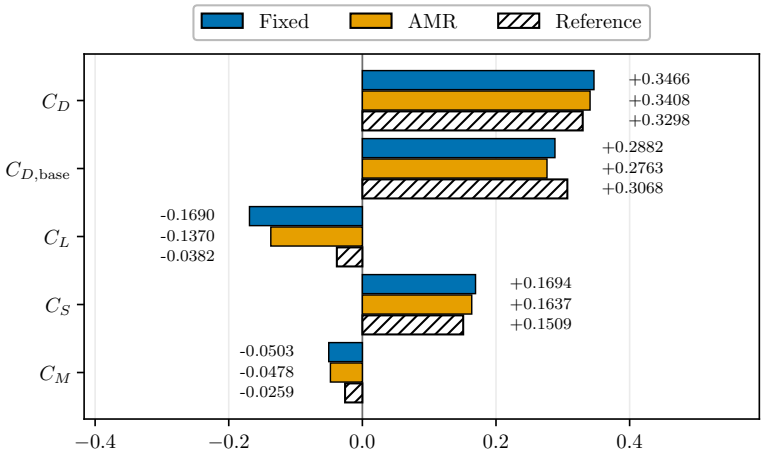
Case 1: Adaptive Mesh Refinement

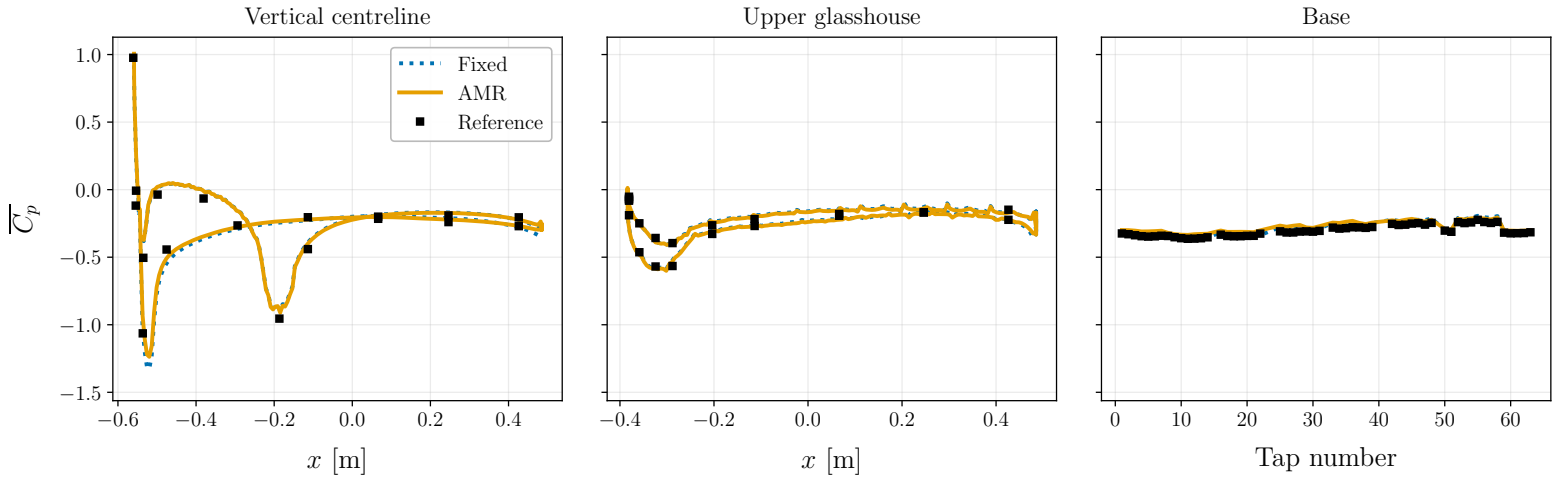


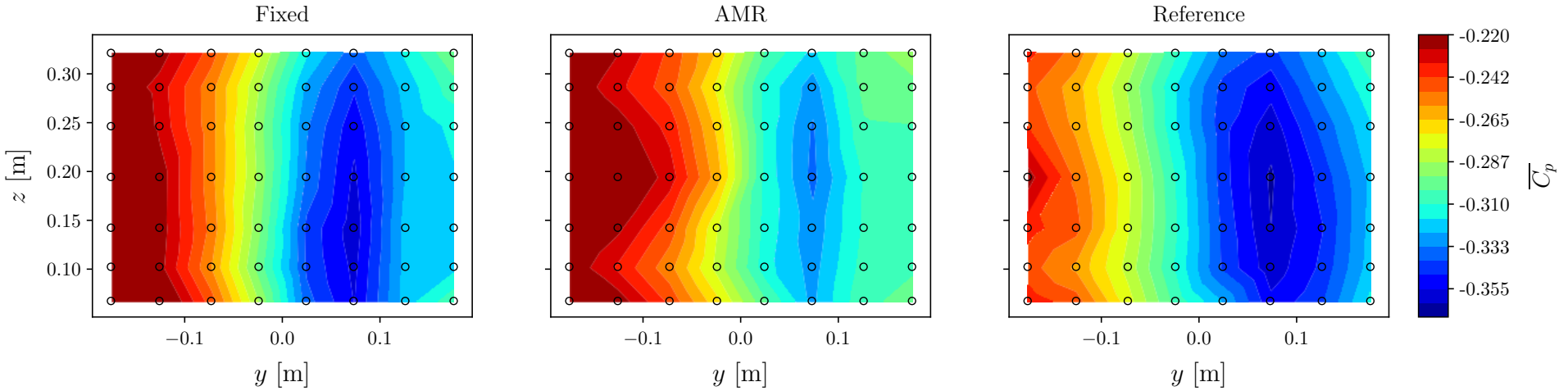


- ☑ AMR captures the wake generation and propagation
- ☑ AMR captures the boundary layer development

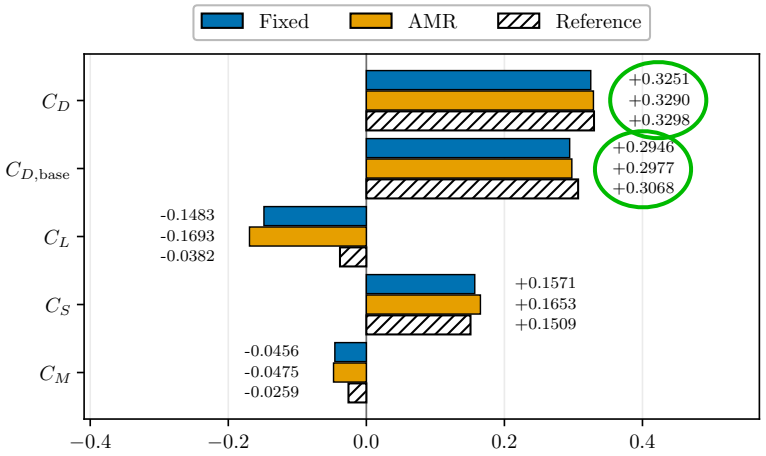
Case 1: Coarse grid 2mm (~1h)

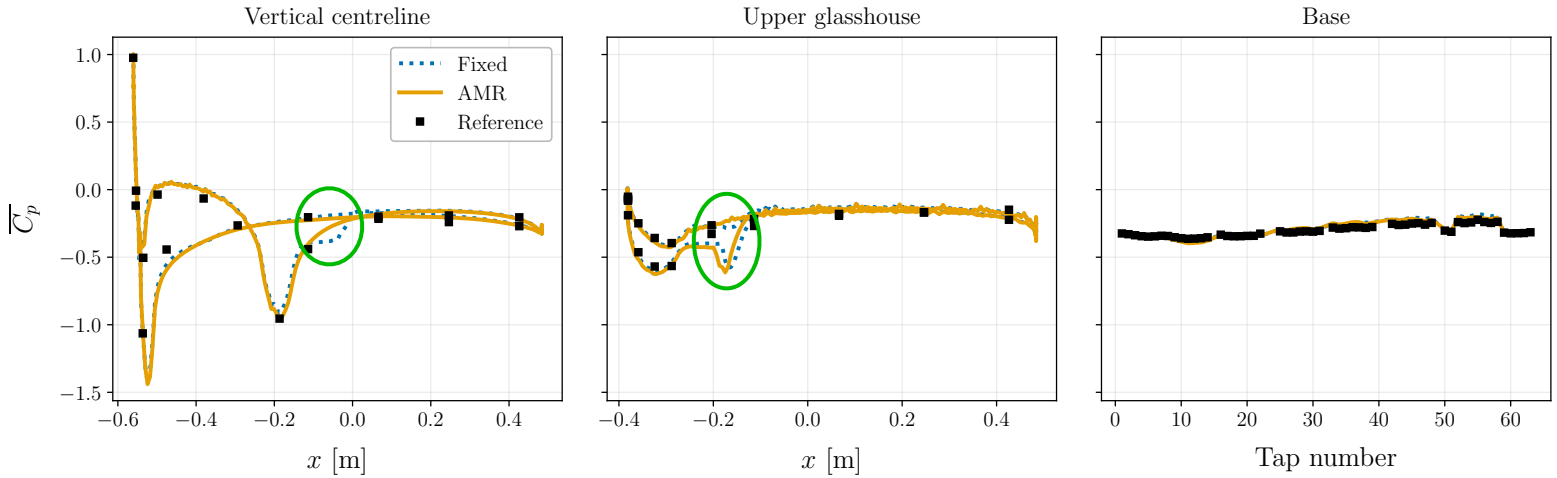


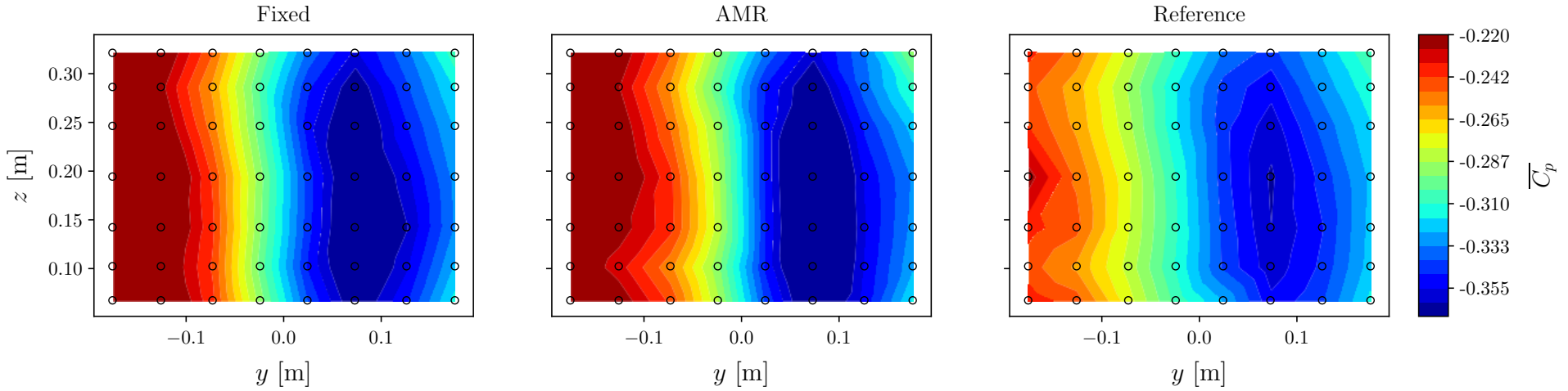




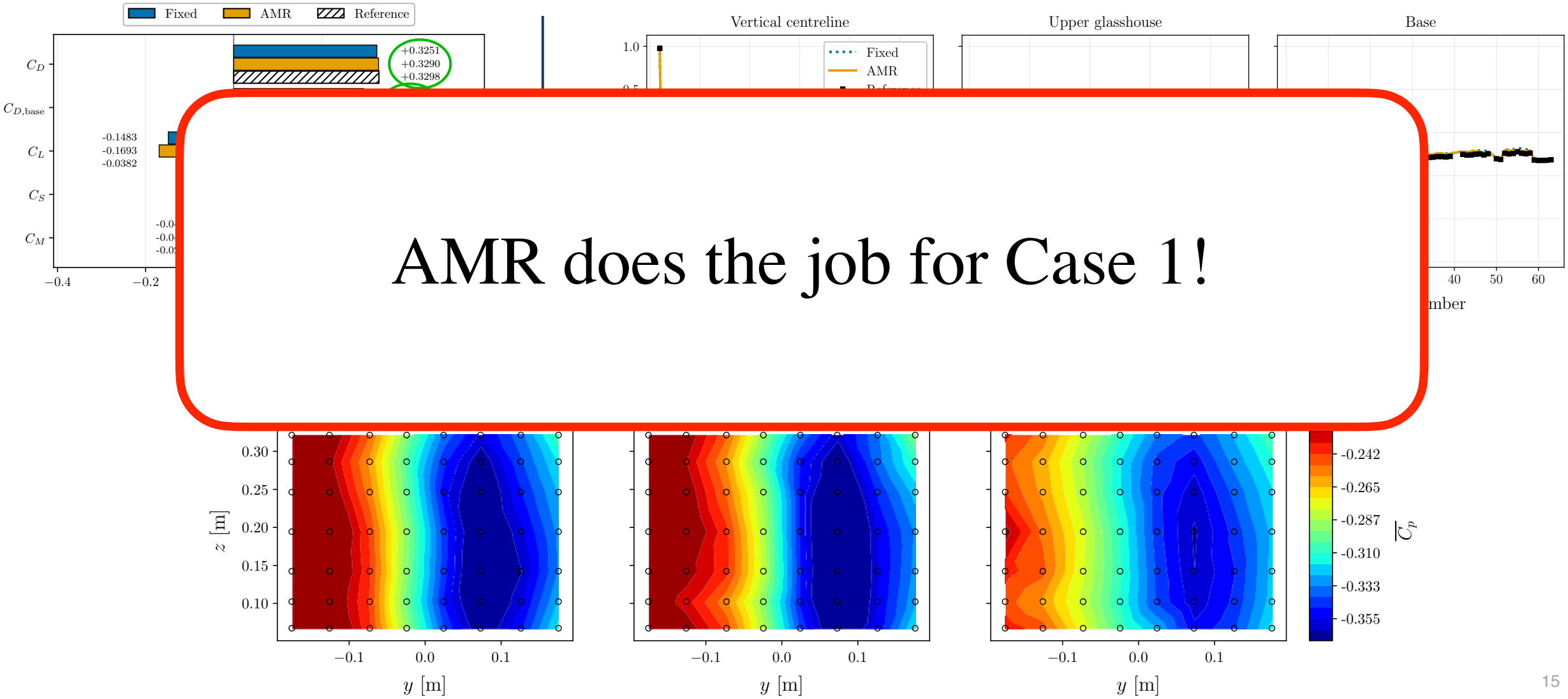
Case 1: Fine grid 0.5mm (~8h)



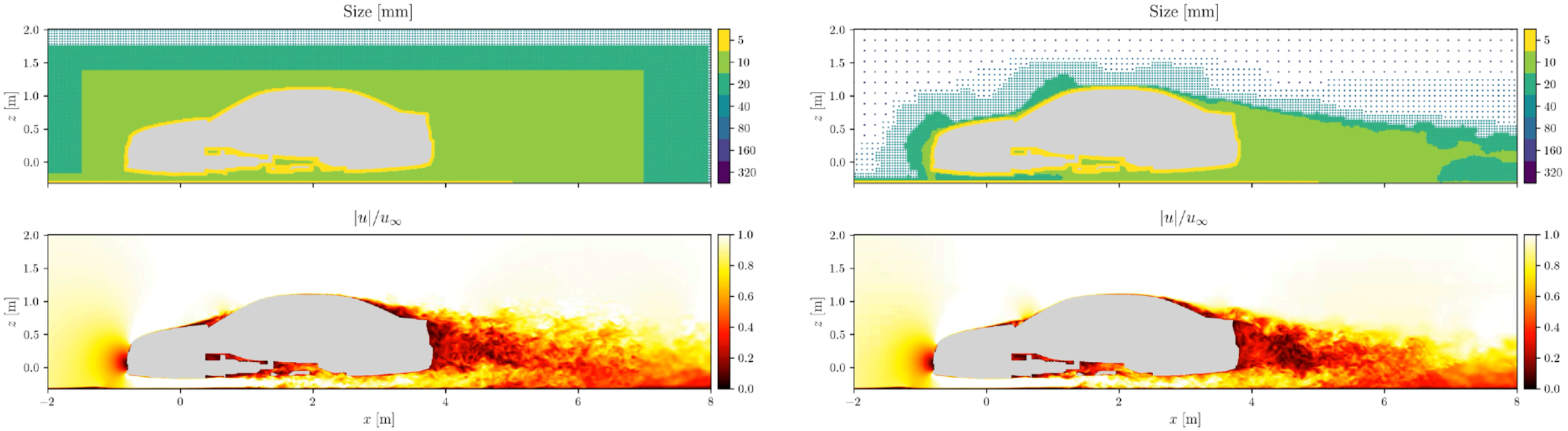




Case 1: Fine grid 0.5mm (~8h)

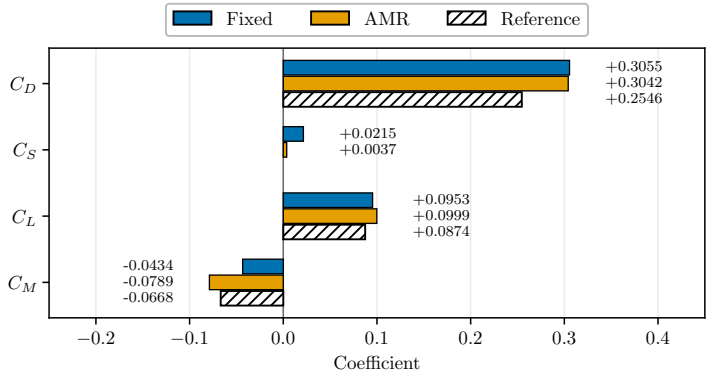


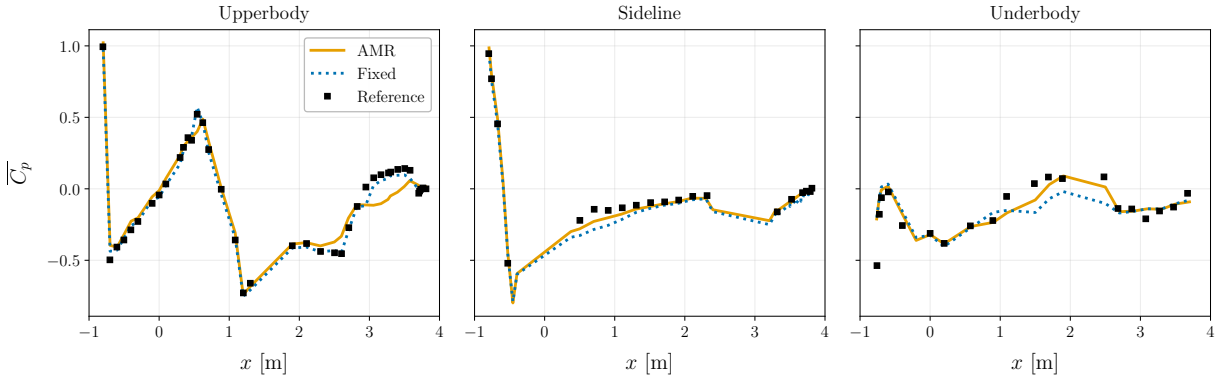
Case 2a: Adaptive Mesh Refinement



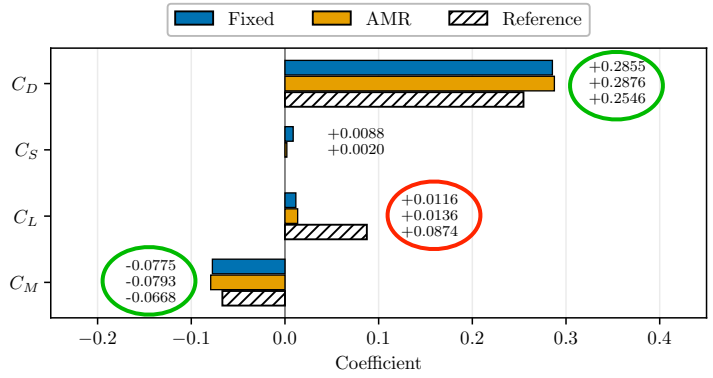
- ☑ AMR captures the wake generation and propagation
- ☑ AMR follows the underbody physics

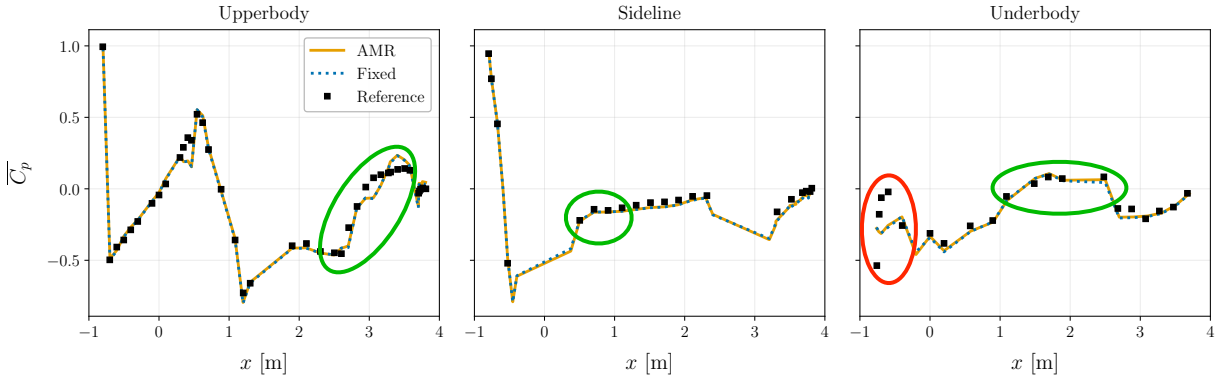
Case 2a: Coarse 10mm (~15min) / Fine 2.5mm (~8h)





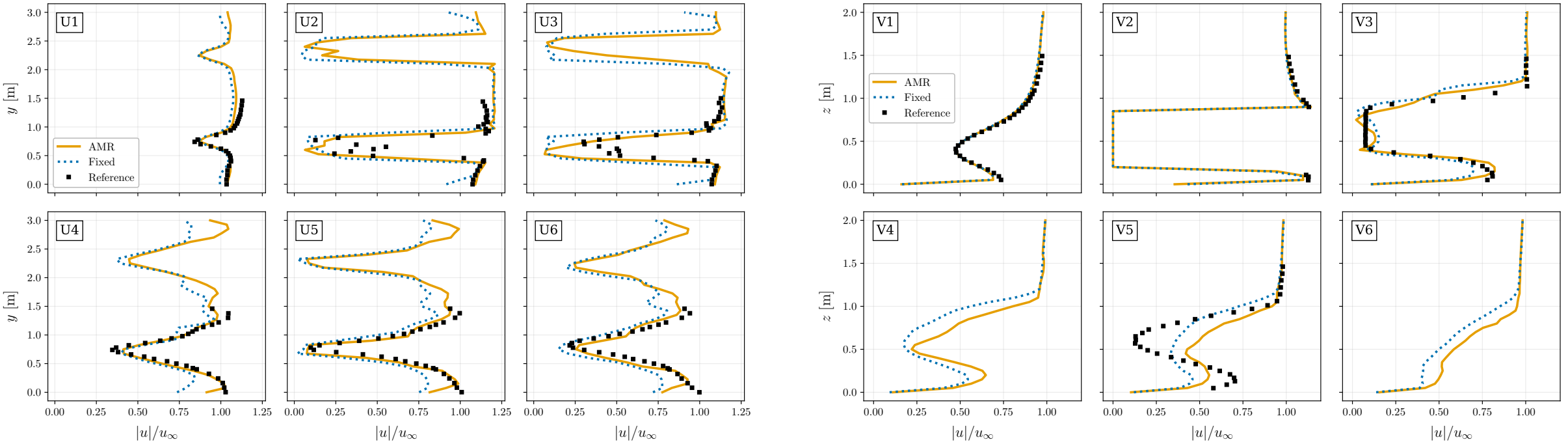
- ☑ Wall model gives impressive results at low resolution





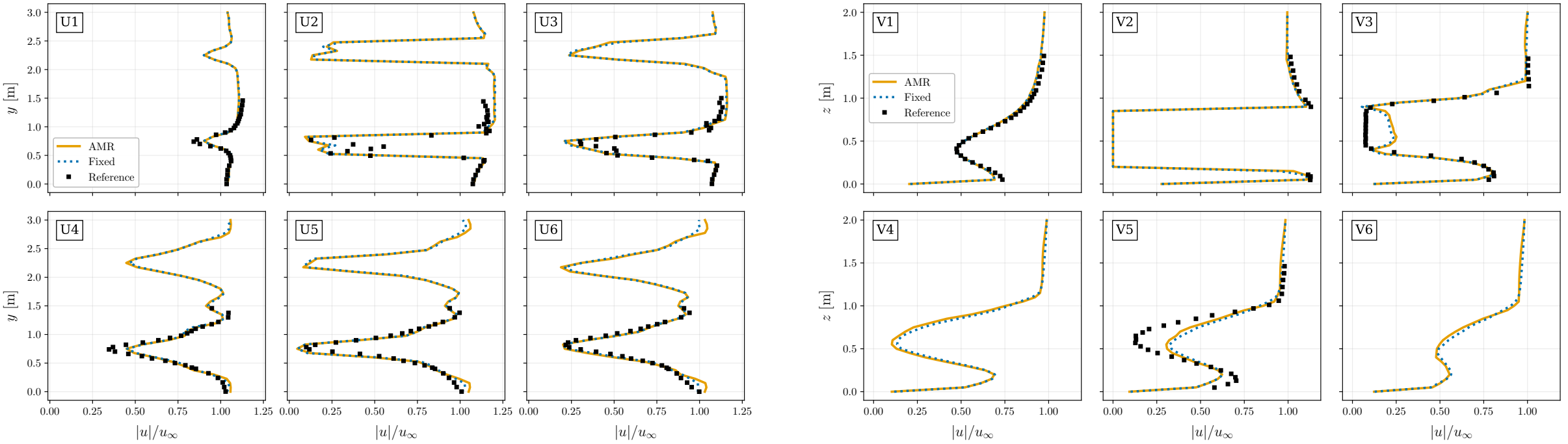
- ☑ Results further improved with increased resolution
- ☑ Gap reduction between fixed and AMR

Case 2a: Coarse 10mm (~15min)



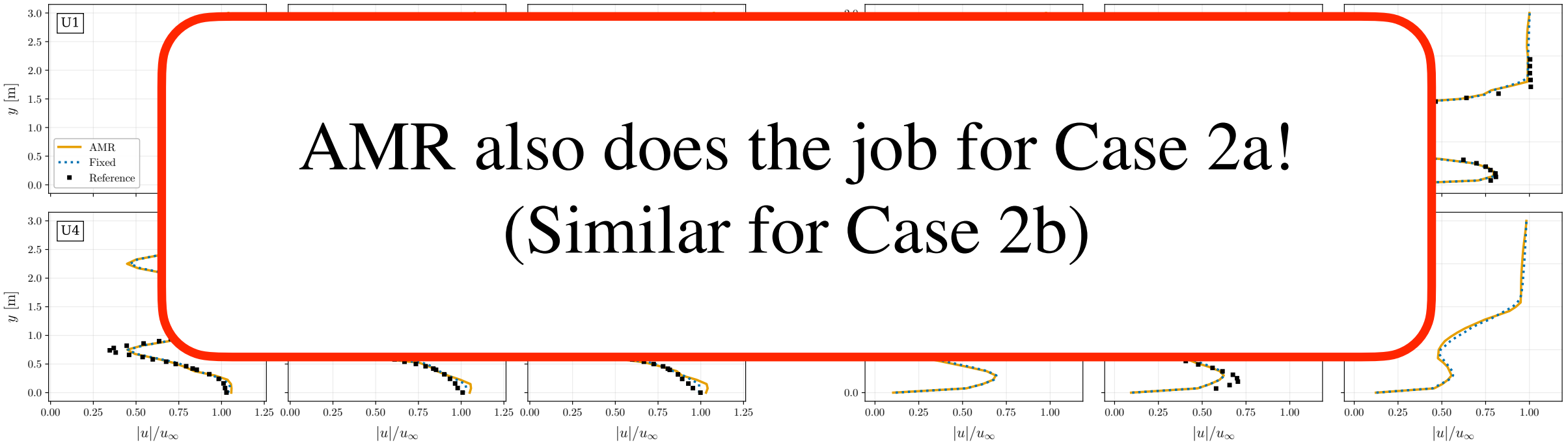
- ☑ Overall good physical behaviour
- ☑ More discrepancies between fixed and AMR

Case 2a: Fine 2.5mm (~8h)



- ☑ Very good results for finest grid
- ☑ Gap reduction between fixed and AMR

Case 2a: Fine 2.5mm (~8h)



AMR also does the job for Case 2a!
(Similar for Case 2b)

- ☑ Very good results for finest grid
- ☑ Gap reduction between fixed and AMR

Take home ideas



CPU Supercomputer (**100k CPU cores**)



Single (gaming) GPU

Doing the same with less ressources

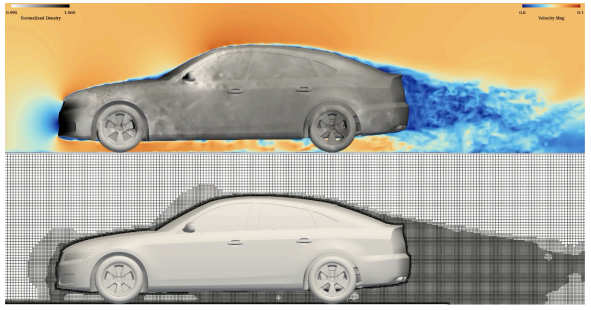
Next-gen CFD solvers will run on a single GPU (node), not a supercomputer!

Octoplus

- Modularity (physics, numerics, data layout, parallelization)
- Octree-grids for real-world geometries
- Multiphysics via multiple solvers (LB, FD, kinetic)
- Getting CUDA comparable performance with C++

Our AutoCFD contributions

- Very good results on Case 1 for < 10 USD
- Good but more work required for Case 2a/b
- AMR for fully automated design studies
- Promising perspectives for low cost physical AI dev



Thank you for your attention!

Interested in Octoplus? Contact us!



Jonas Latt
jonas.latt@unige.ch



Christophe Coreixas
christophecoreixas@bnbu.edu.cn
christophe.coreixas@unige.ch

IAS INSTITUTE FOR ADVANCED STUDY

UNIVERSITY OF GENEVA
FACULTY OF SCIENCE - Computer Science Dept

北师香港浸会大学
BEIJING NORMAL-HONG KONG BAPTIST UNIVERSITY

BNBU CENTRE FOR COMPUTATIONAL CULTURE AND HERITAGE

BNBU Research Fund No UICR0700000094-24
"GPU-accelerated multidisciplinary methods for industrial computational fluid dynamics"