Challenges for the (scale-resolved) simulation of transonic turbines

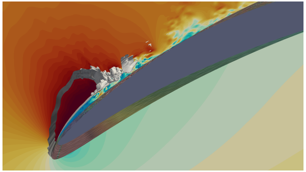
VKI online workshop on Next-Generation High-Speed Low-Pressure Turbines
21 – 23 November 2022

K. Hillewaert[1,2]
contributions by A. Bilocq[1]**, M. Boxho**[1,2,3]**, M. Rasquin**[2]**, T. Toulorge**[2]
Université de Liège[1], Cenaero[2] and Université catholique de Louvain[3]
koen.hillewaert@uliege.be/cenaero.be

Outline

Use of DNS and LES for fundamental studies of turbomachinery

- numerical wind-tunnel: phenomenological understanding
- design of measurement devices
- **reference data sets for calibration and improvement of turbulence models**



DNS and LES complementary to experiments

+ complete control of (boundary) conditions (—— inlet turbulence)
+ all quantities available everywhere
– computational cost: statistical convergence and storage of data

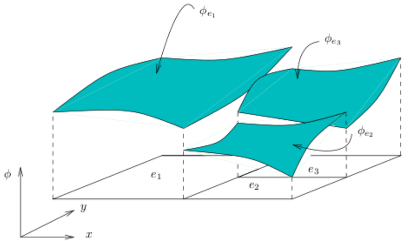
Lucia (Cenaero) - 4 PFlops

Enablers

- Fast-pacing increase in computational power
- high accuracy and highly scalable numerical techniques
- (co-processed) powerful data analysis: machine learning, UQ ...

System of conservative equations

$$\frac{\partial q}{\partial t} + \nabla \cdot \mathbf{g} = 0$$



Elementwise expansion

$$q \approx u = \sum_i \boldsymbol{u}_i \phi_i \ , \ \ \phi_i \in \mathcal{V}$$

Galerkin variational formulation General conservative system $\mathbf{g}$

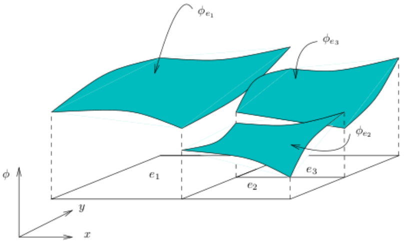
$$\sum_e \int_e v \frac{\partial u}{\partial t} \, dV - \sum_e \int_e \nabla v \cdot \mathbf{g} \, dV + \sum_f \int_f \gamma(u^+, u^-, v^-, v^+, \mathbf{n}) \, dS = 0 \ , \ \ \forall v \in \mathcal{V}$$

Ideal method for DNS and LES on complex geometry

- FEM (ϕ_i, v): accuracy independent mesh quality
- γ impose weak continuity/bc: stability, convergence
- high computational efficiency
- high scalability

System of conservative equations

$$\frac{\partial q}{\partial t} + \nabla\cdot\mathbf{g} = 0$$



Elementwise expansion

$$q \approx u = \sum_i \boldsymbol{u}_i \phi_i \;,\; \phi_i \in \mathcal{V}$$

Galerkin variational formulation Convection subsystem:
FEM-like extension of FVM

$$\sum_e \int_e v \frac{\partial u}{\partial t}\, dV - \sum_e \int_e \nabla v \cdot \mathbf{f}\, dV$$
$$+ \sum_f \int_f (v^+ - v^-)\mathcal{H}(u^+, u^-, \mathbf{n})\, dS = 0 \;,\; \forall v \in \mathcal{V}$$

with $\mathcal{H}$ "FVM" upwind flux Ideal method for DNS and LES on complex geometry

- FEM (ϕ_i, v): accuracy independent mesh quality
- $\mathcal{H}$ FVM upwind flux: stability, convergence and conservation
- high computational efficiency
- high scalability

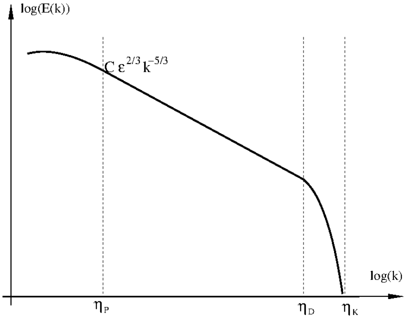
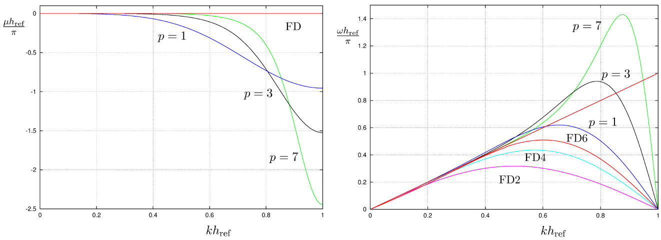


Figure: DNS



- dissipation error → TKE budget
- dispersion error → Kolmogorov cascade
- high order → larger part of resolved scales for same dof
- implicit LES *(Carton et al. 2015)*

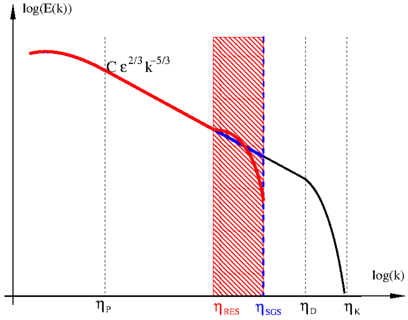
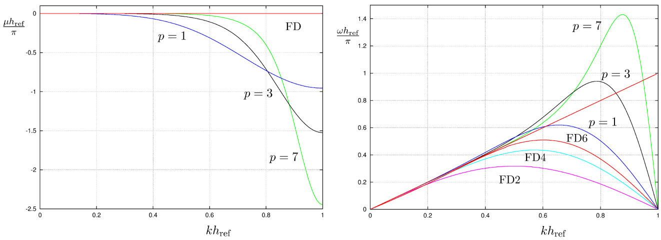
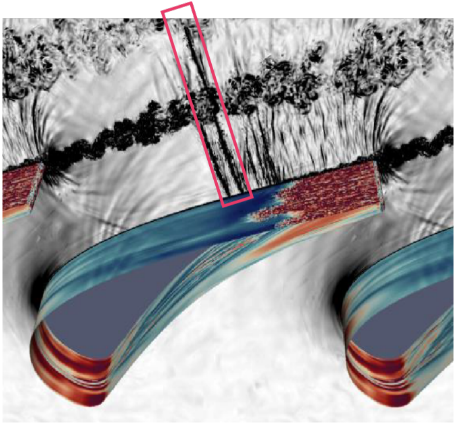


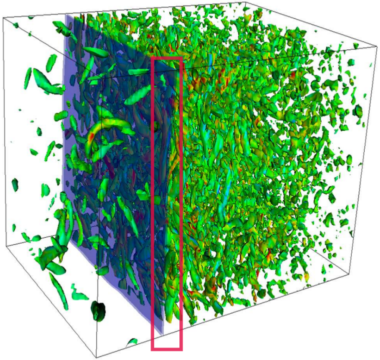
Figure: (I)LES



- dissipation error → TKE budget
- dispersion error → Kolmogorov cascade
- high order → larger part of resolved scales for same dof
- implicit LES *(Carton et al. 2015)*

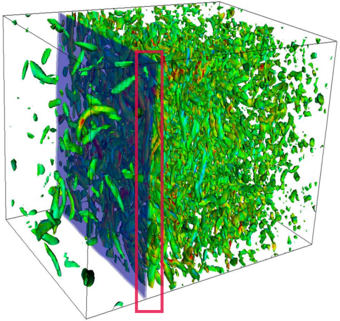


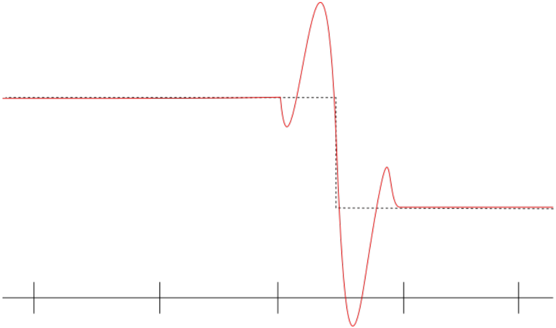
Shock turbulence interaction



HIT through shock (Larson 2006)

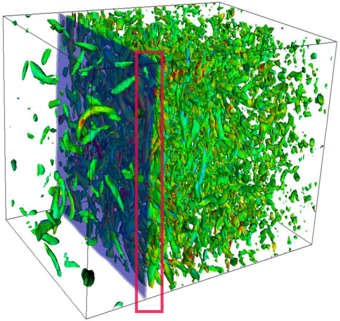
Aliasing of HOT Taylor expansion → Gibbs oscillations

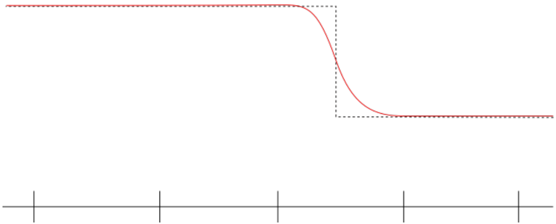


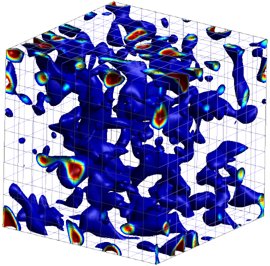


Regularized shock representable by FEM

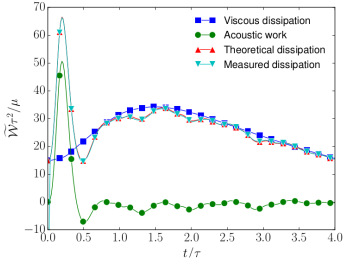
- shock detector
- additional dissipation in troubled cells



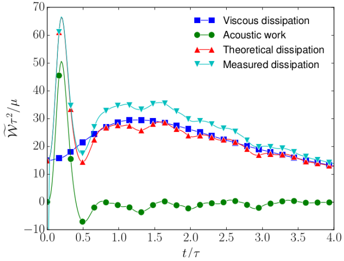




Transonic HIT $M_t = 0.6$



No SCM

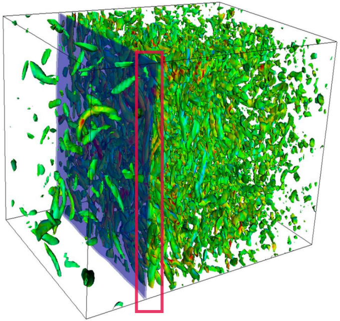


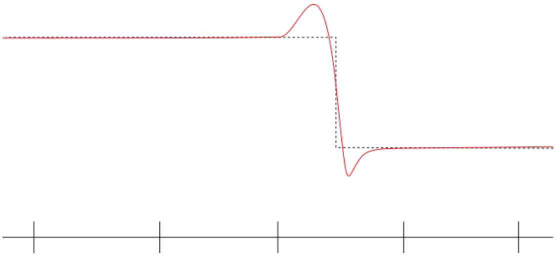
Artificial viscosity

Hillewaert et al., CTR Summer programme 2016

Energy controlled Gibbs oscillations

- no dissipation
- energy conservation/destruction built-in FEM formulation





Elements (Harten, Tadmore, Barth, ...; Carpenter, Gassner, ...)

- *Summation By Parts (SBP):* "discrete" Gauss theorem equivalent removes quadrature inconsistencies

$$\int_e \nabla \cdot \mathbf{g} \, dV \equiv \oint_f \mathbf{g} \cdot \mathbf{n} \, dS$$

- FVM: Entropy $\mathcal{S}$ and entropy variables $w = \mathcal{S}_q$

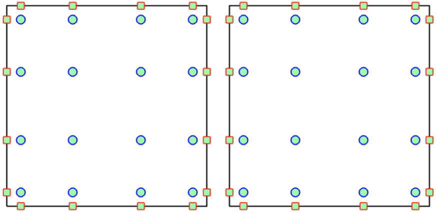
$$w \left(\frac{\partial q}{\partial t} + \nabla \cdot \mathbf{f} \right) = 0 \quad \underbrace{\Rightarrow}_{w\mathbf{f}_q = \boldsymbol{\mathcal{F}}_q} \quad \frac{\partial \mathcal{S}}{\partial t} + \frac{\partial \boldsymbol{\mathcal{F}}}{\partial q} \cdot \nabla \mathbf{q} = 0 \quad \Rightarrow \quad \frac{\partial \mathcal{S}}{\partial t} + \nabla \cdot \boldsymbol{\mathcal{F}} = 0$$

- *Entropy stable schemes (ES)*: Use entropy variables as solution $u = \mathcal{S}_q$ For $v = u$

$$\sum_e \int_e v \frac{\partial q}{\partial t} dV - \sum_e \int_e \nabla v \nabla \cdot \mathbf{f} dV + \sum_f \int_f (v^+ - v^-) \, \mathcal{H}(u^+, u^-, \mathbf{n}) dS = 0$$

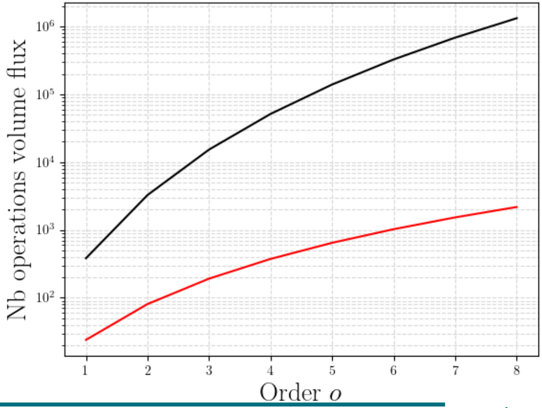
$$\Rightarrow \sum_e \int_e u \frac{\partial q}{\partial t} dV - \sum_f \int_f (u^+ - u^-) \left(\mathcal{H}(u^+, u^-, \mathbf{n}) - \mathbf{f}(\tilde{u}) \cdot \mathbf{n} \right) dS = 0 \quad \Rightarrow \quad \sum_e \int_e \frac{\partial \mathcal{S}}{\partial t} \leq 0$$

entropy stability since $\mathcal{H}(,,)$ is an entropy-consistent flux (e-flux)

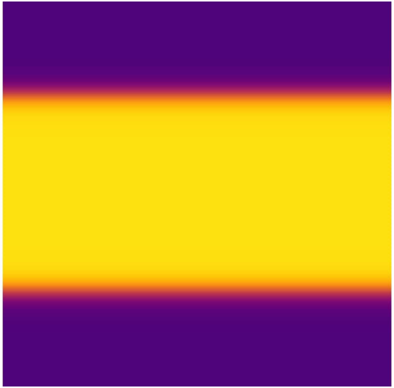


SBP requires that $\mathcal{H}(,,)$ computed for all combinations of solution points

- very high computational const for pure SBP
- reduce cost by using SBP/ES only near shocks
- develop dedicated shock sensor → PhD A. Bilocq



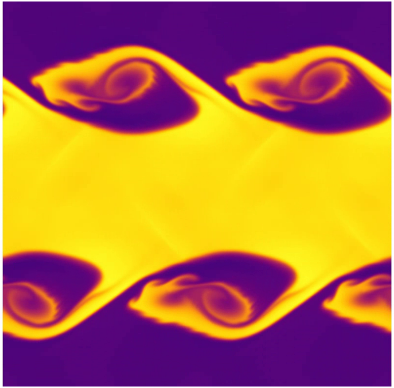
Biperiodic shear layer $Re = \infty$, $M_1 - M_2 = 0.6$, $\rho_1/\rho_2 = 4$



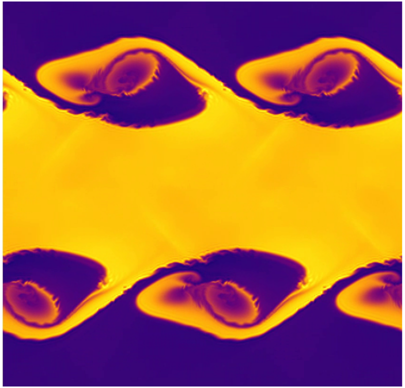
AV

ES

Biperiodic shear layer $Re = \infty$, $M_1 - M_2 = 0.6$, $\rho_1/\rho_2 = 4$

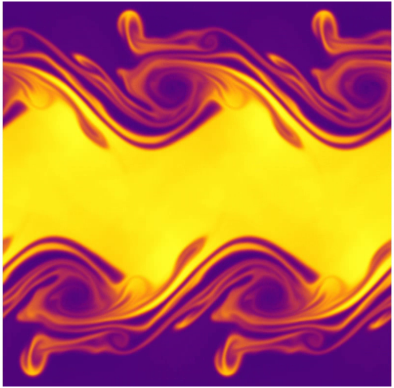


AV

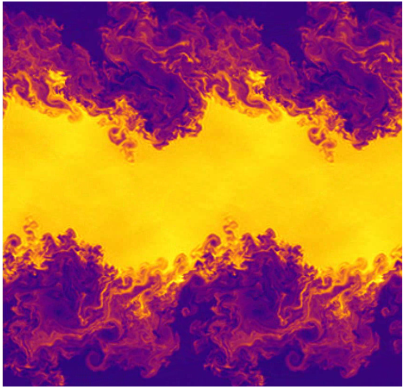


ES

Biperiodic shear layer $Re = \infty$, $M_1 - M_2 = 0.6$, $\rho_1/\rho_2 = 4$



AV



ES

Work of PhD thesis of A. Bilocq

- SBP and ES allows to treat (mild) shocks *without stabilisation*
- maintains order and precision
- has much higher cost per degree of freedom, increases very fast with order
- current ongoing work
  - efficient shock detectors
  - load balancing
  - porting on curved elements

Reynolds $\overline{a}$ and Favre $\widetilde{a}$ averages of a statistically stationary quantity a

$$\overline{a} = \lim_{T\to\infty} \frac{1}{T} \int_0^T a dt \qquad \widetilde{a} = \frac{\overline{\rho a}}{\overline{\rho}}$$

Fluctuations

$$a' = a - \overline{a} \qquad a'' = a - \widetilde{a}$$

Reynolds averaged Navier-Stokes equations

$$\frac{\partial \overline{\rho}}{\partial t} + \nabla \cdot \overline{\rho} \widetilde{\boldsymbol{v}} = 0$$

$$\frac{\partial \overline{\rho} \widetilde{\boldsymbol{v}}}{\partial t} + \nabla \cdot \overline{\rho} \widetilde{\boldsymbol{v}} \widetilde{\boldsymbol{v}} + \nabla \overline{p} = \nabla \cdot (\boldsymbol{\tau} - \boldsymbol{\mathcal{R}})$$

$$\frac{\partial \overline{\rho} \widetilde{E}}{\partial t} + \nabla \cdot \overline{\rho} \widetilde{H} \widetilde{\boldsymbol{v}} = \nabla \cdot \widetilde{\boldsymbol{v}} \cdot (\overline{\boldsymbol{\tau}} - \boldsymbol{\mathcal{R}}) + \nabla \cdot (\overline{\boldsymbol{q}} - \boldsymbol{\mathcal{Q}}) + \nabla \cdot \left(\overline{\boldsymbol{v}'' \cdot \boldsymbol{\tau}} - \frac{1}{2} \overline{\rho \left(\boldsymbol{v}'' \cdot \boldsymbol{v}'' \right) \boldsymbol{v}''} \right)$$

Closing the equations involves

- adapting constitutive equations

$$\widetilde{h} = \mathcal{C}_p \widetilde{T} \qquad \overline{p} = \overline{\rho} R \widetilde{T} \qquad \widetilde{H} = \widetilde{h} + \frac{1}{2} \widetilde{\boldsymbol{v}} \cdot \widetilde{\boldsymbol{v}} + \frac{1}{2} \overline{\boldsymbol{v}'' \cdot \boldsymbol{v}''} = \widetilde{h} + \widetilde{\mathcal{E}_k} + \mathcal{E}_k^t$$

- modeling Reynolds stress $\boldsymbol{\mathcal{R}} = \overline{\rho \boldsymbol{v}'' \boldsymbol{v}''} = \overline{\rho}\ \widetilde{\boldsymbol{v}'' \boldsymbol{v}''}$ and turbulent heat flux $\boldsymbol{\mathcal{Q}} = \overline{\rho \boldsymbol{v}'' h''} = \overline{\rho}\ \widetilde{\boldsymbol{v}'' h''}$
- *approximating* constitutive equations to use available Favre averages

$$\overline{\boldsymbol{\tau}} = \mu \left(\nabla \overline{\boldsymbol{v}} + \nabla \overline{\boldsymbol{v}}^{\mathsf{T}} - 2/3 \nabla \cdot \overline{\boldsymbol{v}}\, \mathcal{I} \right) \approx \mu \left(\nabla \widetilde{\boldsymbol{v}} + \nabla \widetilde{\boldsymbol{v}}^{\mathsf{T}} - 2/3 \nabla \cdot \widetilde{\boldsymbol{v}}\, \mathcal{I} \right) \qquad \overline{\boldsymbol{q}} = \kappa \nabla \overline{T} \approx \kappa \nabla \widetilde{T}$$

Exact Reynolds stress budget

$$\frac{\partial \boldsymbol{\mathcal{R}}}{\partial t} + \nabla\cdot\tilde{\boldsymbol{v}}\boldsymbol{\mathcal{R}} = \underbrace{-\boldsymbol{\mathcal{R}}\cdot\left(\nabla\tilde{\boldsymbol{v}}+\nabla\tilde{\boldsymbol{v}}^{\mathsf{T}}\right)}_{\mathcal{P}} + \underbrace{\overline{\left(\boldsymbol{v}''\nabla p\right)+\left(\boldsymbol{v}''\nabla p\right)^{\mathsf{T}}}}_{\mathcal{R}_p} + \underbrace{\overline{\left(\boldsymbol{v}''\nabla\cdot\boldsymbol{\tau}\right)+\left(\boldsymbol{v}''\nabla\cdot\boldsymbol{\tau}\right)^{\mathsf{T}}}}_{\mathcal{R}_\tau} - \underbrace{\overline{\nabla\cdot\rho\boldsymbol{v}''\left(\boldsymbol{v}''\boldsymbol{v}''\right)}}_{\mathcal{R}_v}$$

with $\boldsymbol{P}$ the production term. Grouping of unclosed terms in $\mathcal{R}_p$, $\mathcal{R}_\tau$, $\mathcal{R}_v$ following Gerolymos and Vallet (2001)

$$\frac{\partial \boldsymbol{\mathcal{R}}}{\partial t} + \nabla\cdot\tilde{\boldsymbol{v}}\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{D}} + \boldsymbol{\Phi} - \boldsymbol{\epsilon} + \frac{2}{3}\overline{p'\nabla\cdot\boldsymbol{v}''\mathcal{I}} + \boldsymbol{\mathcal{K}}$$

with unclosed terms:

- *diffusion:* $\boldsymbol{\mathcal{D}} = -\overline{\nabla\cdot\rho\boldsymbol{v}''\left(\boldsymbol{v}''\boldsymbol{v}''\right)} + \overline{\left(\nabla\cdot\boldsymbol{\tau}'\boldsymbol{v}''\right)+\left(\nabla\cdot\boldsymbol{\tau}'\boldsymbol{v}''\right)^{\mathsf{T}}} - \overline{\left(\nabla p'\boldsymbol{v}''\right)+\left(\nabla p'\boldsymbol{v}''\right)^{\mathsf{T}}}$
- *redistribution/pressure-strain:* $\boldsymbol{\Phi} = \overline{p'\left(\left(\nabla\boldsymbol{v}''\right)+\left(\nabla\boldsymbol{v}''\right)^{\mathsf{T}} - \frac{2}{3}\nabla\cdot\boldsymbol{v}''\mathcal{I}\right)}$
- *dissipation:* $\boldsymbol{\epsilon} = \overline{\left(\boldsymbol{\tau}'\cdot\nabla\boldsymbol{v}''\right)+\left(\boldsymbol{\tau}'\cdot\nabla\boldsymbol{v}''\right)^{\mathsf{T}}}$
- *density fluctuation effects:* $\boldsymbol{\mathcal{K}} = -\overline{\left(\boldsymbol{v}''\left(\nabla\overline{p}-\nabla\cdot\boldsymbol{\tau}\right)\right)+\left(\boldsymbol{v}''\left(\nabla\overline{p}-\nabla\cdot\boldsymbol{\tau}\right)\right)^{\mathsf{T}}} \approx 0$

Exact Reynolds stress budget

$$\frac{\partial \boldsymbol{\mathcal{R}}}{\partial t} + \nabla\cdot\tilde{\boldsymbol{v}}\boldsymbol{\mathcal{R}} = \underbrace{-\boldsymbol{\mathcal{R}}\cdot\left(\nabla\tilde{\boldsymbol{v}}+\nabla\tilde{\boldsymbol{v}}^{\mathsf{T}}\right)}_{\mathcal{P}} + \underbrace{\overline{\left(\boldsymbol{v}''\nabla p\right)+\left(\boldsymbol{v}''\nabla p\right)^{\mathsf{T}}}}_{\mathcal{R}_p} + \underbrace{\overline{\left(\boldsymbol{v}''\nabla\cdot\boldsymbol{\tau}\right)+\left(\boldsymbol{v}''\nabla\cdot\boldsymbol{\tau}\right)^{\mathsf{T}}}}_{\mathcal{R}_\tau} - \underbrace{\overline{\nabla\cdot\rho\boldsymbol{v}''\left(\boldsymbol{v}''\boldsymbol{v}''\right)}}_{\mathcal{R}_v}$$

with $\boldsymbol{P}$ the production term. Grouping of unclosed terms in $\mathcal{R}_p$, $\mathcal{R}_\tau$, $\mathcal{R}_v$ following Knight (1997)

$$\frac{\partial \boldsymbol{\mathcal{R}}}{\partial t} + \nabla\cdot\tilde{\boldsymbol{v}}\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{P}} + \boldsymbol{\mathcal{D}} + \boldsymbol{\Phi} - \boldsymbol{\epsilon}$$

with

- *diffusion:* $\boldsymbol{\mathcal{D}} = -\overline{\nabla\cdot\rho\boldsymbol{v}''\left(\boldsymbol{v}''\boldsymbol{v}''\right)} + \overline{\left(\nabla\cdot\boldsymbol{\tau}\boldsymbol{v}''\right)+\left(\nabla\cdot\boldsymbol{\tau}\boldsymbol{v}''\right)^{\mathsf{T}}} - \overline{\left(\nabla p\boldsymbol{v}''\right)+\left(\nabla p\boldsymbol{v}''\right)^{\mathsf{T}}}$
- pressure-strain: $\boldsymbol{\Phi} = \overline{\left(p\nabla\boldsymbol{v}''\right)+\left(p\nabla\boldsymbol{v}''\right)^{\mathsf{T}}}$
- *dissipation:* $\boldsymbol{\epsilon} = \overline{\left(\boldsymbol{\tau}\cdot\nabla\boldsymbol{v}''\right)+\left(\boldsymbol{\tau}\cdot\nabla\boldsymbol{v}''\right)^{\mathsf{T}}}$

DNS can provide all non-closed terms in RANS turbulence models, but

- *Eddy Viscosity*, *Reynolds stress models*, ... and more to come
- grouping of unclosed terms to ease modeling
  - best combinations may not be independent of flow configuration
  - correlations considered with $\boldsymbol{\tau}$ and p in full, or split in $\boldsymbol{\tau} = \overline{\boldsymbol{\tau}} + \boldsymbol{\tau}'$ and $p = \overline{p} + p'$
- implicit assumptions may not always hold in general case
  - neglected terms due to near-incompressible $\boldsymbol{v}$''
  - use of $\widetilde{\boldsymbol{v}}$ and $\widetilde{T}$ for $\overline{\boldsymbol{\tau}}$ and $\overline{\boldsymbol{q}}$
- formulations can not be converted in to one another . . .
- . . . but share many similar terms

Generic reference data sets for turbulence modeling
Standardized DNS data sets (ERCOFTAC, HiFiTurb)

Turbulent terms can be decomposed in simple averages, *e.g.* Reynolds stress tensor

$$\mathcal{R} = \overline{\rho \mathbf{v}'' \mathbf{v}''} = \overline{\rho \mathbf{v}\mathbf{v}} - \overline{\rho}\widetilde{\mathbf{v}}\widetilde{\mathbf{v}} = \overline{\rho \mathbf{v}\mathbf{v}} - \overline{\rho \mathbf{v}}\ \overline{\rho \mathbf{v}}/\overline{\rho}$$

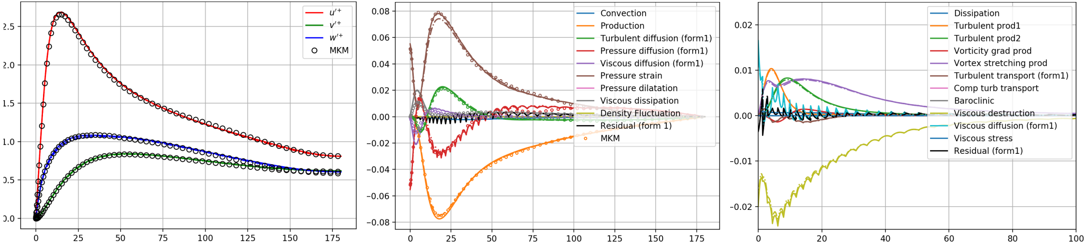
All known RANS models can be reconstructed using 180 basic averages

- level 1* - averaged Navier-Stokes equations
- level 2* - Reynolds stress equations and turbulent heat flux vector
- level 3 - TKE (solenoidal) dissipation
- level 3' - *Reynolds stress dissipation equations: TBD*

Data sets* and best practices developed in HiFiTurb (H2020), part of ERCOFTAC KB Wiki [1]

[1] https://kbwiki-images.s3.amazonaws.com/8/80/List_of_desirable_and_minimum_quantities_to_be_entered_into_the_KB_Wiki.pdf

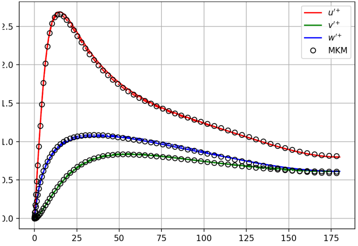
Verification on channel $Re_\tau = 180$



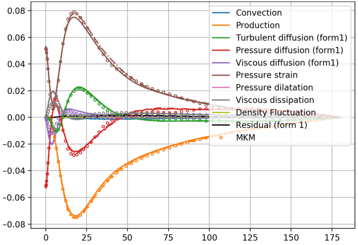
(a) $\sqrt{\mathcal{R}_{uu}}, \sqrt{\mathcal{R}_{vv}}, \sqrt{\mathcal{R}_{ww}}$ (b) budget $\mathcal{R}_{uv}$ (c) budget dissipation ϵ

resolution 75x75x50 - p=3

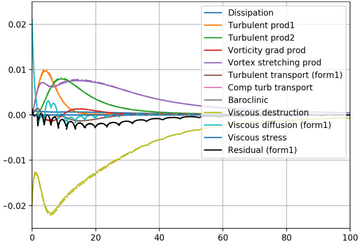
Verification on channel $Re_\tau = 180$



(a) $\sqrt{\mathcal{R}_{uu}}, \sqrt{\mathcal{R}_{vv}}, \sqrt{\mathcal{R}_{ww}}$



(b) budget $\mathcal{R}_{uv}$

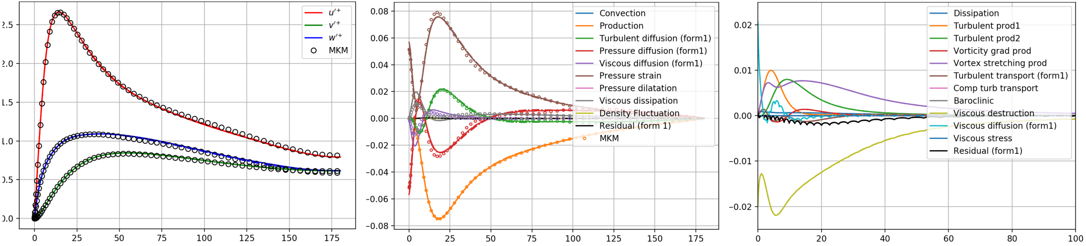


(c) budget dissipation ϵ

resolution 75x75x50 - p=4

DNS verification: closed budget

Verification on channel $Re_\tau = 180$



(a) $\sqrt{\mathcal{R}_{uu}}, \sqrt{\mathcal{R}_{vv}}, \sqrt{\mathcal{R}_{ww}}$ (b) budget $\mathcal{R}_{uv}$ (c) budget dissipation ϵ

resolution 90x72x47 - p=4

ERCOFTAC
European Research Community On Flow, Turbulence And Combustion

- ERCOFTAC KB Wiki DNS
  - standardized data set for level 1 and 2 RANS equations part of
  - a number of additional statistics (*e.g.* rms of density)
  - about 92 “basic” averages to be computed / stored / accumulated
- currently investigating generalised dissipation equations
- mesh and statistical convergence becomes harder with level $\sim$ correlation order
- third order derivatives in dissipation equation $\rightarrow$ at least 5^{th} order accurate
- convergence verification $\sim$ budget closed ?

Direct Numerical Simulations of Turbine Blade Cascades for the Improvement of Turbulence Models through Database Generation, M. Rasquin et al., submitted to ETC2023

LES and DNS are increasingly used **in complement to experiments** for the fundamental study of flows and the development of turbulence models.

LES and DNS are increasingly used **in complement to experiments** for the fundamental study of flows and the development of turbulence models.

DNS and LES results are typically **not accompanied** by an error estimate, *i.c.* confidence intervals on the statistical convergence of the averages.

The lack of convergence of statistical fields is **a source of error** in the Reynolds stress tensor.

LES and DNS are increasingly used **in complement to experiments** for the fundamental study of flows and the development of turbulence models.

DNS and LES results are typically **not accompanied** by an error estimate, *i.c.* confidence intervals on the statistical convergence of the averages.

The lack of convergence of statistical fields is **a source of error** in the Reynolds stress tensor.

We can not apply statistical average error estimates commonly used in experimental studies; need for **alternative** methods, having:

LES and DNS are increasingly used **in complement to experiments** for the fundamental study of flows and the development of turbulence models.

DNS and LES results are typically **not accompanied** by an error estimate, *i.c.* confidence intervals on the statistical convergence of the averages.

The lack of convergence of statistical fields is **a source of error** in the Reynolds stress tensor.

We can not apply statistical average error estimates commonly used in experimental studies; need for **alternative** methods, having:

- **Low memory footprint**
- **Fast convergence**
- **Recurrence formula**

Consider we want statistical averages of a turbulent flow which is

statistically **stationary**	**ergodic**	with realizations: $x_i = \rho_i, \rho_i \boldsymbol{v}_i, p_i, \ldots$ in time series with average μ and variance σ

Consider we want statistical averages of a turbulent flow which is

statistically **stationary**

ergodic

with realizations: $x_i = \rho_i, \rho_i \boldsymbol{v}_i, p_i, \ldots$ in time series with average μ and variance σ

Quantification of the maximum deviation between:

the *actual average* $\mu = \mathbb{E}[x]$

the *sample average* $\overline{x}_n = \frac{1}{n} \sum_i^n x_i$

Consider we want statistical averages of a turbulent flow which is

statistically **stationary**

ergodic

with realizations: $x_i = \rho_i, \rho_i \mathbf{v}_i, p_i, \ldots$ in time series with average μ and variance σ

Quantification of the maximum deviation between:

the *actual average* $\mu = \mathbb{E}\left[x\right]$

the *sample average* $\overline{x}_n = \frac{1}{n}\sum_i^n x_i$

Central theorem: $\overline{x}_n$ is normally distributed $\mathcal{N}(\mu, \sigma_{\overline{x}_n})$

Confidence interval: $\mathcal{P}\left(\mu \in [\overline{x}_n - N_\alpha \sigma_{\overline{x}_n}, \overline{x} + N_\alpha \sigma_{\overline{x}_n}]\right) = \alpha$

Question: How can we estimate $\sigma_{\overline{x}_n}$ in a practical way ?

Standard approach for *non-correlated samples*

$$\sigma_{\overline{x}_n} \approx \frac{s_n}{\sqrt{n-1}} \quad \text{where,} \quad s_n^2 = \frac{1}{n}\sum_i (x_i - \overline{x})^2$$

Difficulties specific to DNS and LES

- consecutive samples highly **correlated** → classical estimate not valid, correct for *auto-correlation function (ACF)*
- relatively short duration in physical time → fast convergence required, use as many data as possible
- the whole flow field and all statistical quantities → can not store whole time signal for estimation

Standard approach for *non-correlated samples*

$$\sigma_{\overline{x}_n} \approx \frac{s_n}{\sqrt{n-1}} \quad \text{where,} \quad s_n^2 = \frac{1}{n}\sum_i (x_i - \overline{x})^2$$

Difficulties specific to DNS and LES

- consecutive samples highly **correlated** → classical estimate not valid, correct for *auto-correlation function (ACF)*
- relatively short duration in physical time → fast convergence required, use as many data as possible
- the whole flow field and all statistical quantities → can not store whole time signal for estimation

Approach

- Development of three **estimators** based on **short recurrences** with m data points
- m will depend on the **correlation time scale** $\mathcal{T}$.
- undersampling if m too large to limit the storage.

The sample mean is an **unbiased estimator** of μ,

$$\overline{x}_n = \frac{1}{n}\sum_{t=1}^{n} x_t$$

The sample mean is an **unbiased estimator** of μ,

$$\overline{x}_n = \frac{1}{n}\sum_{t=1}^{n} x_t$$

Its **variance** is computed as,

$$\mathbb{V}\text{ar}\left(\overline{x}_n\right) = \mathbb{E}\left[\left(\overline{x}_n - \mu\right)^2\right] = \frac{\sigma^2}{n}\left[1 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\rho_k\right]$$

Confidence intervals on statistical data

Variance of the sample mean $\mathbb{V}\mathrm{ar}\left(\overline{x}_n\right)$

The sample mean is an **unbiased estimator** of μ,

$$\overline{x}_n = \frac{1}{n}\sum_{t=1}^{n} x_t$$

Its **variance** is computed as,

$$\mathbb{V}\mathrm{ar}\left(\overline{x}_n\right) = \mathbb{E}\left[\left(\overline{x}_n - \mu\right)^2\right] = \frac{\sigma^2}{n}\left[1 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\rho_k\right]$$

The **auto-correlation function** (ACF) at τ_k is given by,

$$\rho_k = \sigma^{-2}\mathbb{E}\left[(x(t) - \mu)(x(t + \tau_k) - \mu)\right]$$

The sample mean is an **unbiased estimator** of μ,

$$\overline{x}_n = \frac{1}{n}\sum_{t=1}^{n} x_t$$

Its **variance** is computed as,

$$\mathbb{V}\text{ar}(\overline{x}_n) = \mathbb{E}\left[(\overline{x}_n - \mu)^2\right] = \frac{\sigma^2}{n}\left[1 + 2\sum_{k=1}^{n-1}\left(1 - \frac{k}{n}\right)\rho_k\right]$$

The **auto-correlation function** (ACF) at τ_k is given by,

$$\rho_k = \sigma^{-2}\mathbb{E}\left[(x(t) - \mu)(x(t + \tau_k) - \mu)\right]$$

How can we construct $\sigma_{\overline{x}_n}$ knowing that σ and ρ_k are **unknown** and hence need to be **approximated** ?

Variance of the sample mean $\mathbb{V}\mathrm{ar}\left(\overline{x}_n\right)$

σ^2

the variance of the process can be approximated with the **sample variance**:

$$s_n^2 = \frac{1}{n-1}\sum_{t=1}^{n}\left(x_t - \overline{x}_n\right)^2$$

ρ_k

the ACF can be approximated in various ways using selected statistical quantities:

$$\hat{\gamma}_k,\ \hat{\delta}_k,\ \hat{\varphi}_k$$

σ^2

the variance of the process can be approximated with the **sample variance**:

$$s_n^2 = \frac{1}{n-1} \sum_{t=1}^{n} \left(x_t - \overline{x}_n\right)^2$$

ρ_k

the ACF can be approximated in various ways using selected statistical quantities:

$$\hat{\gamma}_k, \hat{\delta}_k, \hat{\varphi}_k$$

For instance, combining $\hat{\delta}_k$ with the sample variance one obtains an **asymptotically unbiased estimator**:

$$\mathbb{V}\mathrm{ar}\left(\overline{x}_n\right) \approx \alpha \mathbb{E}\left[\beta s_n^2 - \sum_{k=1}^{m} \hat{\delta}_k\right]$$, with α and β functions of n and m.

Truncation to the m first terms to reduce the memory storage.

All sample-based statistical data can be computed recursively, *e.g.* **sample mean** and the **sample variance**:

$$\overline{x}_n = \frac{n-1}{n}\overline{x}_{n-1} + \frac{x_n}{n}$$

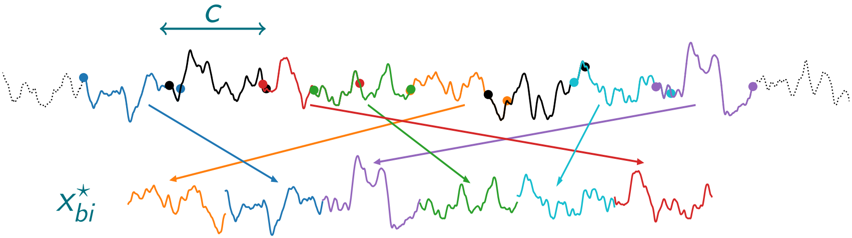
$$s_n^2 = \frac{(n-2)s_{n-1}^2 + (n-1)\overline{x}_{n-1}^2 + x_n^2 - n\overline{x}_n^2}{n-1}$$

Advantages:

- minimal **memory** storage
- low **computational** effort

$\Rightarrow$ Storage ACF estimators $\sim$ m

- **Resampling** algorithm used when the probability distribution of the data is unknown;
- Infer **any statistics** (e.g., mean, variance, ...), from a single time series;
- **Initially** proposed by [Efron, 1979];
- Extended by [Kunsch, 1989] to preserve **correlations** in time series.



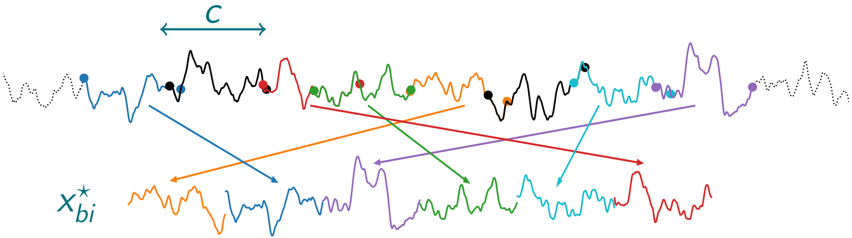
1) **Evaluate** the optimal block size c

2) **Construct** of $N - c + 1$ random overlapping blocks of size c

3) Rrandomly **concatenate** N/c blocks to get a new data series $x^*_{b,i}$

4) **Compute** the statistics of interest

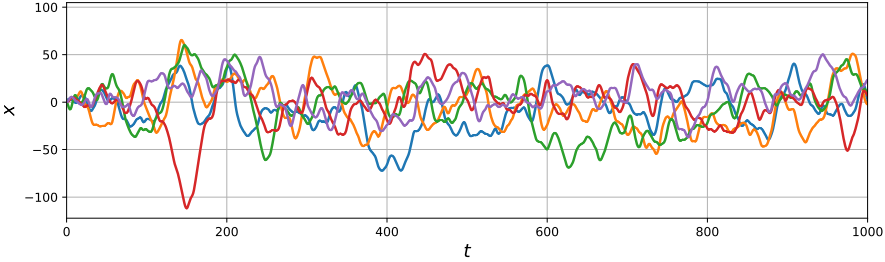
6) **Repeat** the process B times

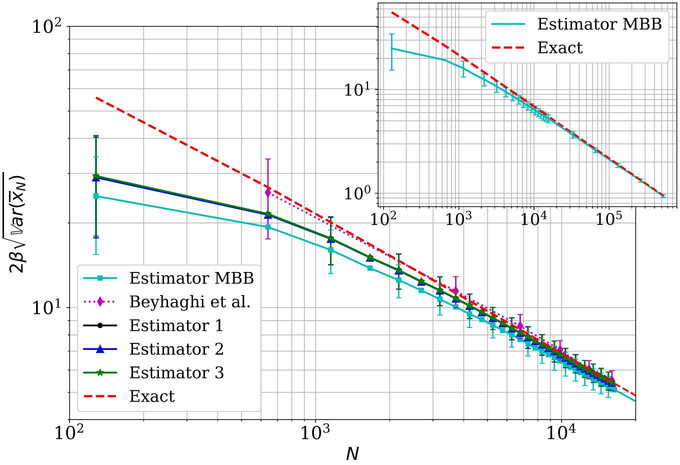


Auto-regressive (AR) process

$$x_i = \sum_{k=1}^{n} \alpha_k x_{i-k} + \epsilon_i, \text{ where } \epsilon_i = \mathcal{N}(0, \sigma_\epsilon^2) \text{ and } \alpha_k \in \mathbb{R}.$$

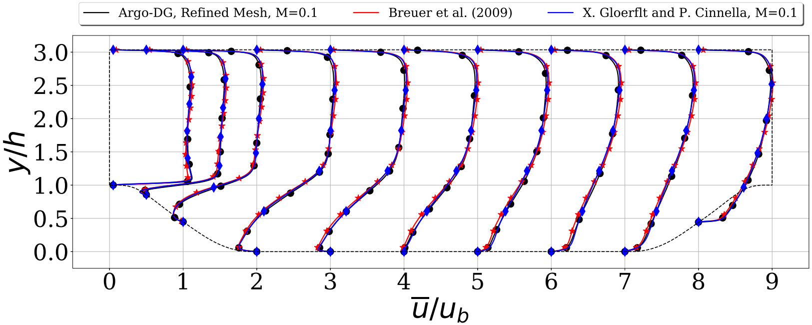
with $n = 6$ and $x_{-5} = \cdots = x_0 = 0$. We have generated 100 realizations of the process.



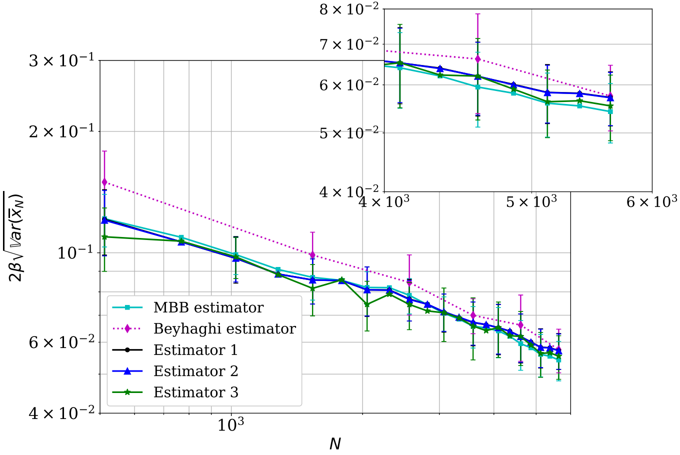


- the three estimators are **unbiased** estimators of the mean.
- they **coincide** well with the estimator of [Beyhaghi *et al.*, 2018] at moderate and large N
- MMB is not biased, it just converges more **slowly**.

- **Bi-periodic** flow evolving between two walls featuring a streamwise **constriction**[2]
- **Controlled** pressure gradient to match the bulk Reynolds number ($Re_b = 10{,}595$) combined with a **low bulk Mach** number $M_b = 0.1$



2 https://www.kbwiki.ercoftac.org/w/index.php/Abstr:2D_Periodic_Hill_Flow



Near the separation:

- Location characterized by a thin boundary layer;
- Rapid and random displacement of the separation over the hill;
- The thee estimators are framed by the MBB (below) and the [Beyhaghi *et al.*, 2018] (above) ones.

Part of PhD Margaux Boxho

Current results

- confidence intervals required for DNS, and therefore estimate of $\mathbb{V}\mathrm{ar}\left(\overline{x}_n\right)$
- new estimators are required
  - high correlation in time series $\rightarrow$ classical estimate based upon s_n^2 not usable
  - large data set $\leftrightarrow$ current refined estimators (MBB, Beyaghi, ...) based upon full time history not practical
- three new estimators are proposed
  - correction with m correlation terms
  - estimated in situ using dedicated statistical quantities
  - computed using recursive formula
- Validation on
  - auto-regressive (AR) process,
  - stochastic solution of the Kuramoto–Sivashinsky (KS) equation,
  - velocity signals extracted on the two-dimensional periodic hill.
- number of terms m is automatically and locally refined $\sim$ estimated **correlation time scale** $\mathcal{T}$.
- normally all timesteps are used but for long correlation times (*e.g.* reirculation bubble) m can be reduced reliably by **undersampling**

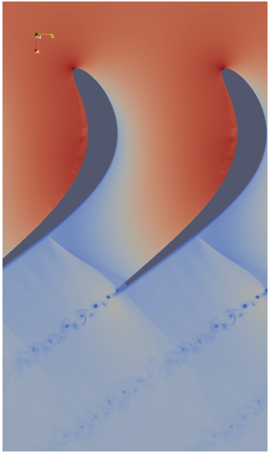
Ongoing work and perspectives

Current development challenges

- integration of error estimators in co-processing
- Reynolds stress dissipation equation breakdown and implementation
- turbulence injection strategy

Applications in current projects

- development of wall models using machine learning on DNS / LES reference data (M. Boxho)
- shock capturing for DNS and LES and study of transonic turbulence (A. Bilocq)
- *study of flow in spleen cascade without and with rotating bars (G. Lopes)*
- *LES of active turbulence grids (F. Bertelli)*
- *receptivity of boundary layer to passing wake (G. Pastorino)?*



Clean inlet flow

- conditions $Re_{2s} = (70k, 120k) \times M_{2s} = (0.7, 0.9, 0.95)$
- RANS reference data: flow field, Reynolds stress, budgets
- blade $\overline{M_s}$, $\overline{C_f}$ distributions
- blade $\overline{p'^2}$, $\overline{\rho'^2}$ (and γ ?)
- wake distributions of p°, α, $\mathcal{E}_k^t$
- wake distributions of $\overline{\rho'^2}$ and length/time scales

LES active grid (collab. F. Bertelli)