Introduction to the DNS of turbulent flows

— Part II —

June 2012

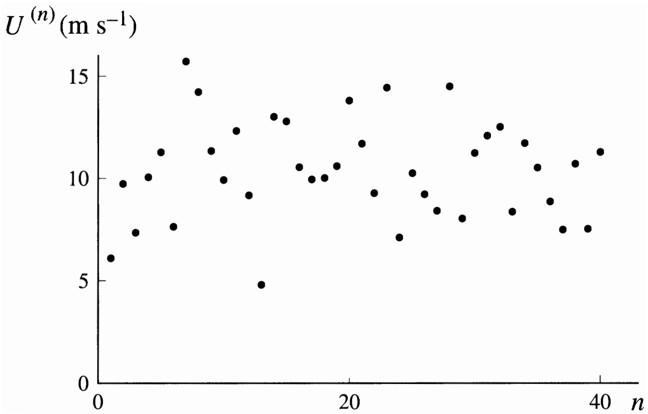
Turbulence, DNS, Post-process of DNS-LES results

- 1-Turbulence - introduction
- 2-DNS at CTTC
- 3-Post-process of DNS-LES results

3 - Post-process of DNS-LES results

- Some theory
- Some practice

Turbulence and deterministic chaos



Imagine that a fluid flow experiment is repeated many times under the *same* conditions.

At a given position and time, one component of the velocity is measuered $u(\mathbf{x}, t)$.

The result is different at each realization of the experiment. But the Navier-Stokes equations are deterministic. How is it possible ?

- In any experiment there are, unavoidably, perturbations in initial conditions, boundary conditions, material properties, etc.
- Due to the nature of the governing equations, turbulent flow fields display a large sensitivity to such perturbations.

Turbulence and deterministic chaos - Butterfly effect

Certain dynamical systems (dynamical = systems whose state evolves with time) may exhibit dynamics that are highly sensitive to initial conditions

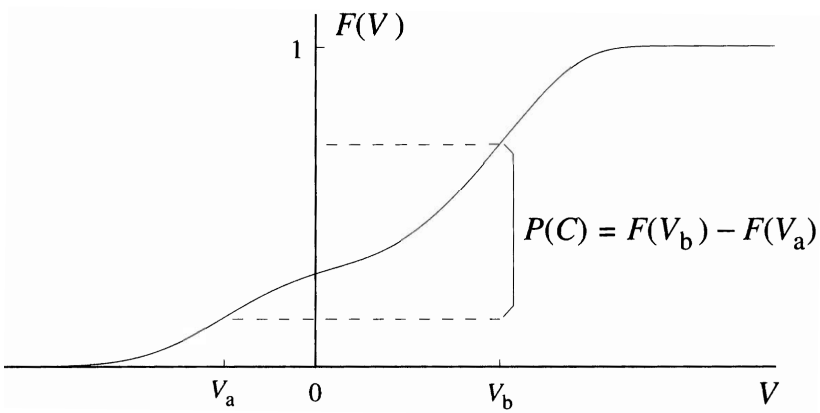
As a result of this sensitivity the behavior of chaotic systems appears to be random.

Butterfly effect: a butterfly's wings might create tiny changes in the atmosphere that may ultimately alter the path of a tornado or delay. The flapping wing represents a small change in the initial condition of the system, which causes a chain of events leading to large-scale alterations of events. *Had the butterfly not flapped its wings, the trajectory of the system might have been vastly different*.

How can we deal with this problem ?

Turbulent velocity as a random variable. CDF

The deterministic but chaotic nature of turbulence precludes correct predictions of magnitudes such as $u(\mathbf{x}, t)$. However, we may treat $u(\mathbf{x}, t)$ as a random variable. Therefore, we need a bit of statistics.



Cummulative distribution function, CDF:

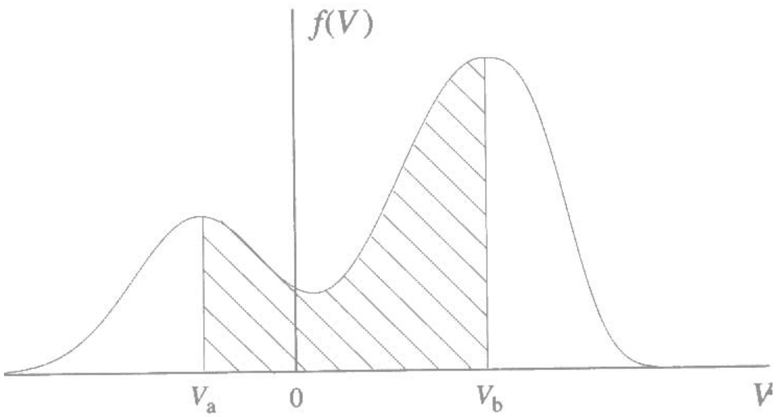
$F(V) = P\left(u < V\right)$

$F(-\infty) = 0,\ F(\infty) = 1$

$P(a \leq u < b) = F(b) - F(a)$

It is non-decreasing, $F(V_b) \geq F(V_a)$ for $V_b < V_a$.

Turbulent velocity as a random variable. PDF



Probability density function, PDF:

$f(V) = \frac{dF(V)}{dV}$

$f(V) \geq 0$

$$\int_{-\infty}^{\infty} f(V)dV = 1$$

(but $f(V)$ can be greater than 1).

$$P(a \leq u < b) = F(b) - F(a) = \int_{a}^{b} f(V)dV$$

Turbulent velocity as a random variable - 2

Mean (or expectation):

$\langle U \rangle = \int_{-\infty}^{\infty} V f(V) dV$

$\langle A + B \rangle = \langle A \rangle + \langle B \rangle$

$\langle cA \rangle = c \langle A \rangle$ (for a constant c)

Fluctuation

$u = U - \langle U \rangle$

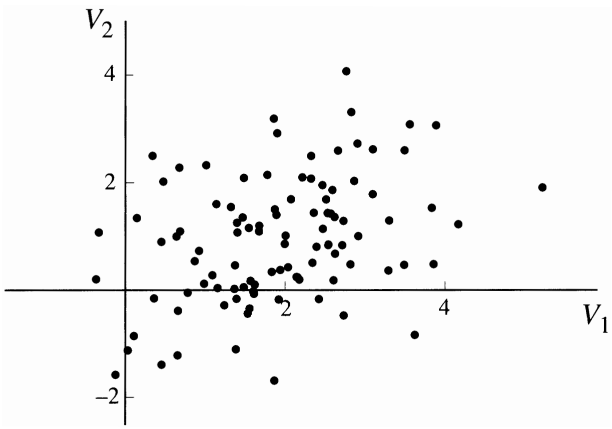
Variance

$var(U) = \langle u^2 \rangle = \int_{-\infty}^{\infty} (V - \langle U \rangle)^2 f(V) dV$

Standard deviation

$\sigma_U = \langle u^2 \rangle^{\frac{1}{2}}$

Joint random variables - 1



Sometimes we are interested in two or more random variables. For instace, the components of the velocity vector $\mathbf{u}(\mathbf{x}, t)$ or the coordinates where a dart reaches the dartboard.

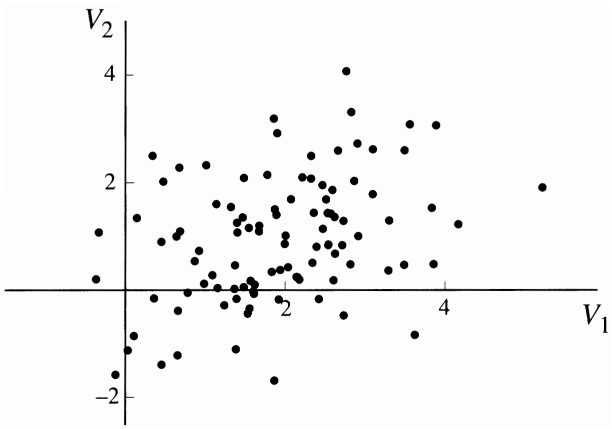
They might be studied separately, as each of them has its own pdf, but perhaps there is a relation between them, so that for instance when one is high the other is low.

The **cummulative distribution function of the joint random variables** is defined by:

$$F_{12}(V_1, V_2) \equiv P(U_1 < V_1, U_2 < V_2)$$

$$F_{12}(-\infty, V_2) = P(U_1 < -\infty, U_2 < V_2) = 0$$

Joint random variables - 2



The **joint PFD** is defined by $f_{12}(V_1, V_2) \equiv \frac{\partial^2}{\partial V_1 \partial V_2} F_{12}(V_1, V_2)$

and

$P(V_{1a} \leq U_1 < V_{1b}, V_{2a} \leq U_2 < V_{2b}) = \int_{V_{1a}}^{V_{1b}} \int_{V_{2a}}^{V_{2b}} f_{12}(V_1, V_2) dV_2 dV_1$

$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{12}(V_1, V_2) dV_1 dV_2 = 1$

Marginal PDF:

$$\int_{-\infty}^{\infty} f_{12}(V_1, V_2) dV_1 = f_2(V_2)$$

Joint random variables - 3

Are they correlated?

The **covariance** is:

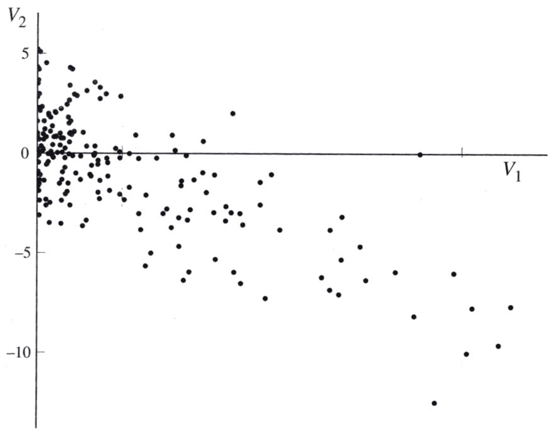
$$cov(U_1, U_2) =< u_1 u_2 >= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (V_1 - \langle U_1 \rangle)(V_2 - \langle U_2 \rangle) f_{12}(V_1, V_2) dV_1 dV_2$$

and the **correlation coefficient** is:

$\rho_{12} \equiv \frac{< u_1 u_2 >}{\sqrt{\langle u_1^2 \rangle \langle u_2^2 \rangle}}$

$-1 \leq \rho_{12} \leq 1$

Joint random variables - 4



$(\rho_{12} = -0{,}81)$

- $\rho_{12} = 0$: there is no correlation
- $\rho_{12} = 1$: perfect correlation
- $\rho_{12} = -1$: perfect negative correlation

Independent random variables are not correlated, but the converse ("non-correlated variables are independent") is not true in general.

Random process - autocorrelation

Consider the initial experiment, the measure of a instantaneous velocity at a given position $u(\mathbf{x}, t)$.

If we consider the same velocity but as a function of time, we have a time-dependant random variable, a **random process**. At each instant it has a pdf.

However, if whe consider $u(\mathbf{x}, t)$ and its value after s seconds, $u(\mathbf{x}, t + s)$ as joint random variables, we can obtain more information using their correlation, here called **autocorrelation**.

$$\rho(s) \equiv \frac{<u(t)u(t+s)>}{\sqrt{<u^2(t)><u^2(t+s)>}}$$

Autocorrelation is a function of s that decreases with s (if the variable is a velocity component).

Statistically stationary flows

A time dependant random variable is **statistically stationary** if *all* its statistics are invariant respect to a time shift.

$$U(\mathbf{x}, t) \equiv U(\mathbf{x}, t + \Delta t)$$

In this case, as an example, $< u^2(t) >=< u^2(t+s) >$ and the autocorrelation becomes:

$$\rho(s) \equiv \frac{<u(t)u(t+s)>}{<u^2(t)>}$$

Statistically stationarity is a very important concept for post-process of DNS data. Turbulent flows are never stationary, but if boundary conditions are constant, they reach **a statistically stationary state** that usually is our region of interest.

Random fields - two point correlation

Usually, we are not interested in a single velocity (or temperature, pressure..) value but in **all the field**. This mathematical object, where at each position of space there is a random variable associated, is called **random field**.

As we did previously with random processes, we can consider the velocities at two points as joint random variables and study how the values at a point are correlated with the values at a certain neighbour.

This is the two-point auto-covariance:

$$R(\mathbf{r}, \mathbf{x}, t) \equiv < \mathbf{u}(\mathbf{x}, t)\mathbf{u}(\mathbf{x}+\mathbf{r}, t) >$$

Random fields - statistically homogeneity

- A random field is **statistically homogeneous** if all its statistics are invariant respect to a shift in position, in any direction.
- A random field is **statistically two-dimensional** if all its statistics are invariant respect to a shift in one particular direction.
- ..

As an example, the flow in a pipe can be **statistically axisymmetric**.

Probability and averaging - “the meaning of it all”

Note that “all” has been constructed based on probability, and then on cummulative distribution functions, pdfs and so on..

However, if velocity in a turbulent field is understood as a random distribution, its pdf is not known even if we can perform a full DNS of it.

So, for instance, we can not evaluate the expectation of the velocity in a given point:

$\langle U \rangle = \int_{-\infty}^{\infty} V f(V) dV$

This is not a big issue, we can *estimate* this expectation. This is what we do in the post-process of DNS-LES results.

Respect to the pdf, it can also be estimated as we will see, but it should be noted that the most important results are the averaged heat transfer coefficients, surface forces, temperatures etc and not their pdfs.

Ensemble average

We considered a experimental measure of $u(\mathbf{x}, t)$ as a random variable, but as we don't know its pdf, we can not evaluate its mean (or standard deviation) as $\langle U \rangle = \int_{-\infty}^{\infty} V f(V) dV$.

All we can do is repeat the experiment, under the same conditions, and measure several times its value, $U^{(1)}$, $U^{(2)}$,..

And then, average the result (**ensemble average**):

$\langle U \rangle_N = \frac{1}{N} \Sigma_{n=1}^{N} U^{(n)}$

It can be seen that

$\langle \langle U \rangle_N \rangle = \langle U \rangle$

Ensemble average of course can not be used in practice.

Time average, space average

The methods actually used to estimate the average of the flow are:

- Time average. If the flow is statistically stationary (SS):
  $< U(t) >_T = \frac{1}{\Delta T} \int_{t_0}^{t_0+\Delta T} U(t')dt'$
  The result is an averaged velocity field, function of $\mathbf{x}$
- Space average. If it is statistically homogeneous (difficult to find in a engineering application):
  $< U(t) >_V = \frac{1}{V} \int_V U(\mathbf{x}, t)dv$
  The result would be a scalar function of time. If it is also SS, after time integration, we get a single scalar.
- If it is homogeneous but only in one direction (eg. x), as in simulations with one periodic direction:
  $< U(t) >_x = \frac{1}{L_x} \int_{x=0}^{x=L} U(x, y, z, t)dx$
  The result is a two-dimensional field (with coordinates y, z), function of time. After time integration, we get a two-dimensional field.

Difficulties of time average

With the definition of time-average:

$$< U(t) >_{t_0}^{t_0+\Delta T} = \frac{1}{\Delta T} \int_{t_0}^{t_0+\Delta T} U(t') dt'$$

We have two problems:

- Select t_0. When is the flow statistically stationary (SS) ?
- Select ΔT. How long must integration period be?

Difficulties of time average-2

It is not obvious when SS state has been reached.

If $\mathbf{u}(t)$ could be saved, integration would be done after the simulation, when t_0 has been identified. But this is totally impossible (too long).

Instead, for each field q of interest, partial integrals are evaluated and stored:

$$\tau < q >_i^{i+1} = \int_{t=i\tau}^{t=(i+1)\tau} q dt$$

Assume now for simplicity that this is done from the beginning of the simulation, $i = 0$.

$\tau << \Delta T$, but long enough so that the number of fields to be stored, $\frac{t_0 + \Delta T}{\tau}$, is reasonable.

From $\tau < q >_i$ we can very easily get the average from $i\tau$ to $(i+1)\tau$, just with a division. Or, more important, combine partial integrals:

$$2\tau < q >_i^{i+2} = \int_{i\tau}^{(i+2)\tau} q dt = \int_{i\tau}^{(i+1)\tau} q dt + \int_{(i+1)\tau}^{(i+2)\tau} q dt = \tau < q >_i^{i+1} + \tau < q >_{i+1}^{i+2}$$

Difficulties of time average-3

When all the $\tau < q >_i^{i+1}$ for $i = 0$ to $i = N - 1$ are available, estimation of $< q >$ can begin (but we let simulation go on during this process, as perhaps SS has not been reached and more data will be needed).

We proceed from the end as we know that the initial data is not SS and we will discard it. The last partial integral considered is:

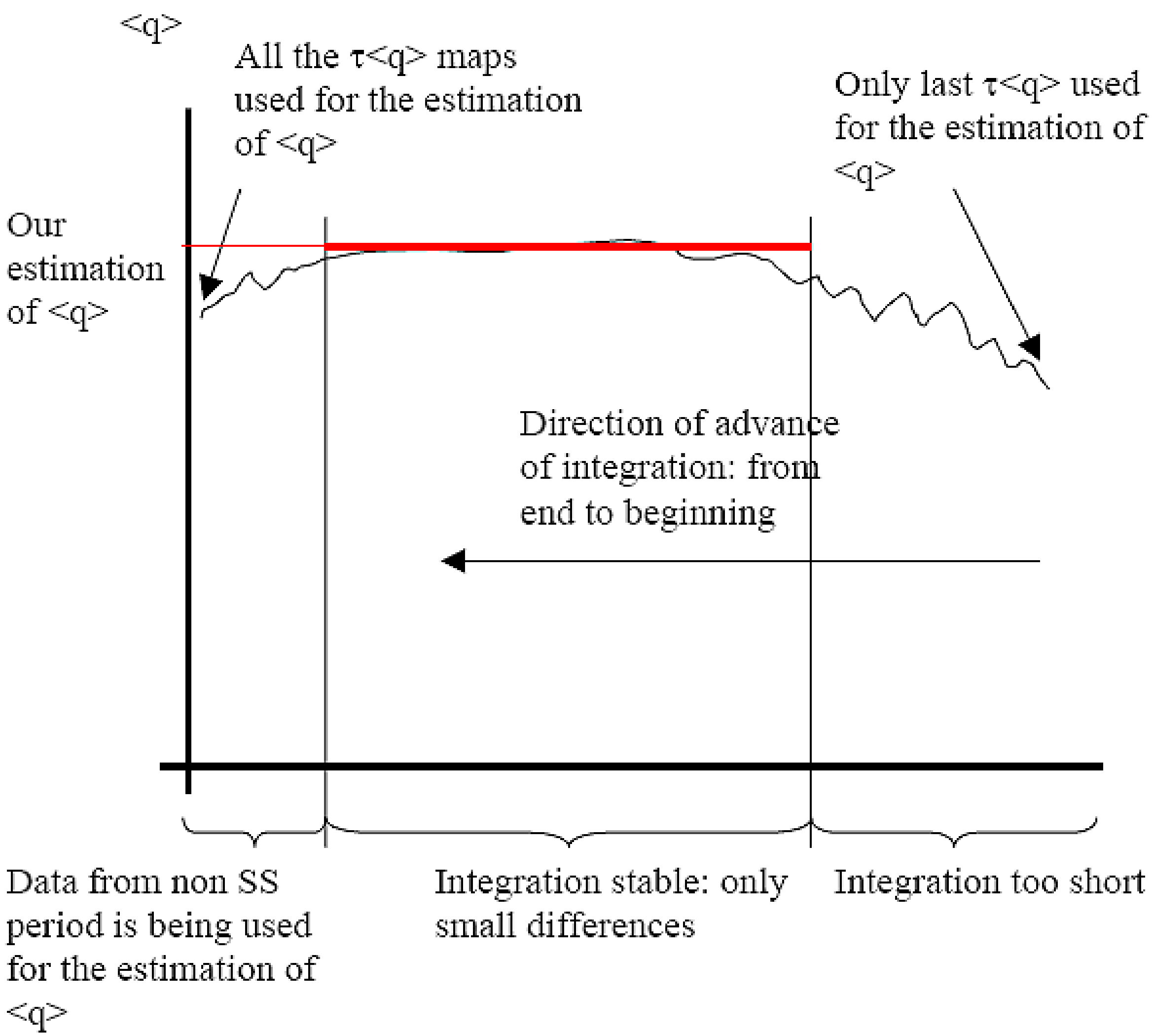
$$\tau < q >_{N-1}^{N} = \int_{\tau(N-1)}^{\tau N} q dt$$

From it, we evaluate the last sub-average. Then, $\tau < q >_{N-1}^{N}$ and $\tau < q >_{N-2}^{N-1}$ are combined to get the second sub-average, and so on.

The sub-averages, beginning from the end, are plotted. If the flow has been SS for a sufficient time and the data from the beginning is kept, three-stages will be identified:

- Total integral period too short. Estimation of average is not good.
- Stable estimation of $< q >$. Result independent of the integration period.
- Period choosen includes initial, non SS data, results change a bit from center.

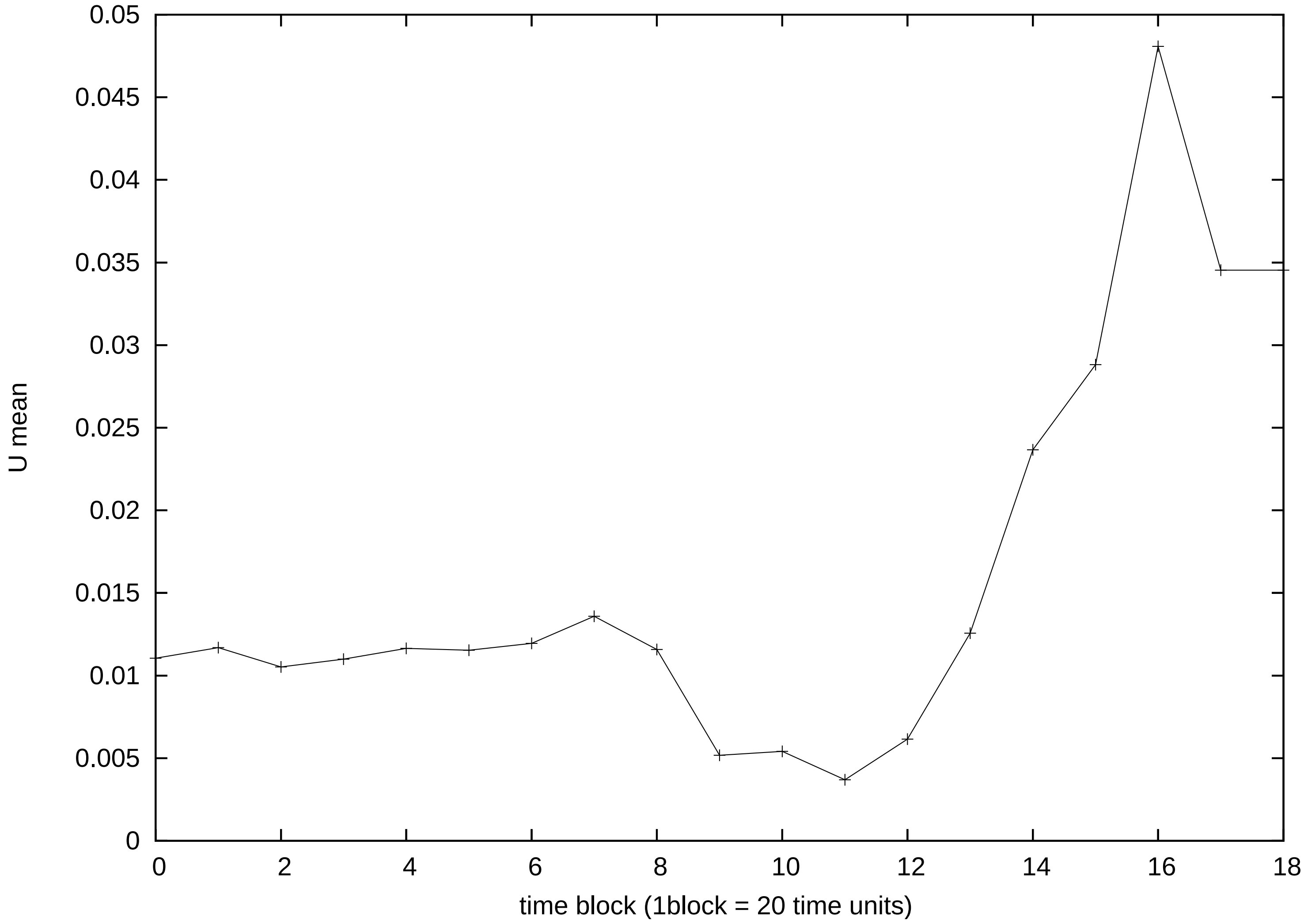
Difficulties of time average-4



If the central part has not been reached, simulation must go on. The left part is not allways visible. As q is a field only a few positions can be examined.

Difficulties of time average-5

A real example:



Evaluation of second order statistics $< a'b' >$

How can $< a'b' >=< (a- < a >)(b- < b >) >$ be estimated if $< a >$ and $< b >$ are not known util the end of the simulation, when Δt and t_0 have been found?

(Remember that $a =< a > +a'$)

$$
\begin{aligned}
< a'b' > &= < (a- < a >)(b- < b >) > \\
&= < ab - a < b > - < a > b+ < a >< b >> \\
&= < ab > - < a >< b > - < a >< b > + < a >< b > \\
&= < ab > - < a >< b >
\end{aligned}
$$

(as $< a < b >>=< a >< b >$ because $< b >$ is a constant and $<< a >< b >>=< a >< b >$)

So, what we do is keep maps of $\tau < a >_i$, $\tau < b >_i$ **and** $\tau < ab >_i$

Typical data to be computed includes the components of the Reynolds Stress tensor, $u_i u_j$.

A large number of fields must be computed and stored.

Example: Reynolds stress tensors in a flow statistically uniform in x-direction

Quiz[1]: Which of the Reynolds Stress Tensor components are null?

$< u_1' >= 0$

$< u_1' u_1' >$?

$< u_2' u_2' >$?

$< u_3' u_3' >$?

$< u_1' u_2' >$?

$< u_1' u_3' >$?

$< u_2' u_3' >$?

And $< u_1' T' >$, $< u_3' T' >$? Hint: Recall that the correlation is $\rho_{ab} = \frac{<ab>}{\dots}$

(Answer is at the end).

[1] Quiz: A riddle or obscure question; an enigma (Webster)

Verification of DNS - non homogeneous directions

How fine must the mesh be to solve the problem?

Even a slight change in the mesh results in a totally different $u(t)$ solution, but perhaps with the same statistics.

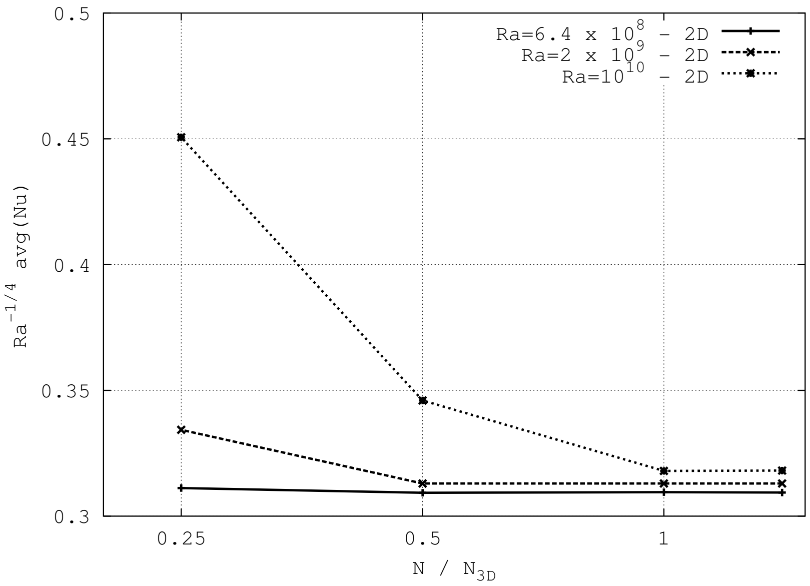
Our hope is to select a mesh fine enough to predict $< u >$ and perhaps the second order statistics.

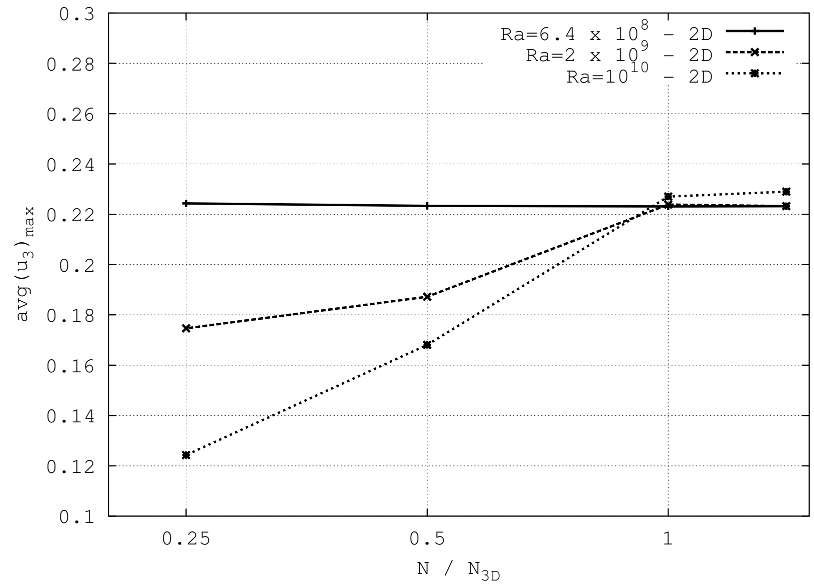
This can be done:

- A priori estimations. In certain cases that are well known, (ie, channel flow) the complexity of the solution can be estimated as a function of Re, and therefore a mesh can be selected.
  Note that Kolmogorov approach can only be used under certain conditions, and that even knowing the size of the smallest structures, the number of nodes to resolve them with appropriated accuracy is to be determined (as depends on the accuracy of the method).
- A posteriori comparisons. A set of meshes is used and DNS is performed in each. The behaviour of the statistic of interest is studied.

Verification of DNS simulations

Grid convergence studies - example of a posteriori analysis



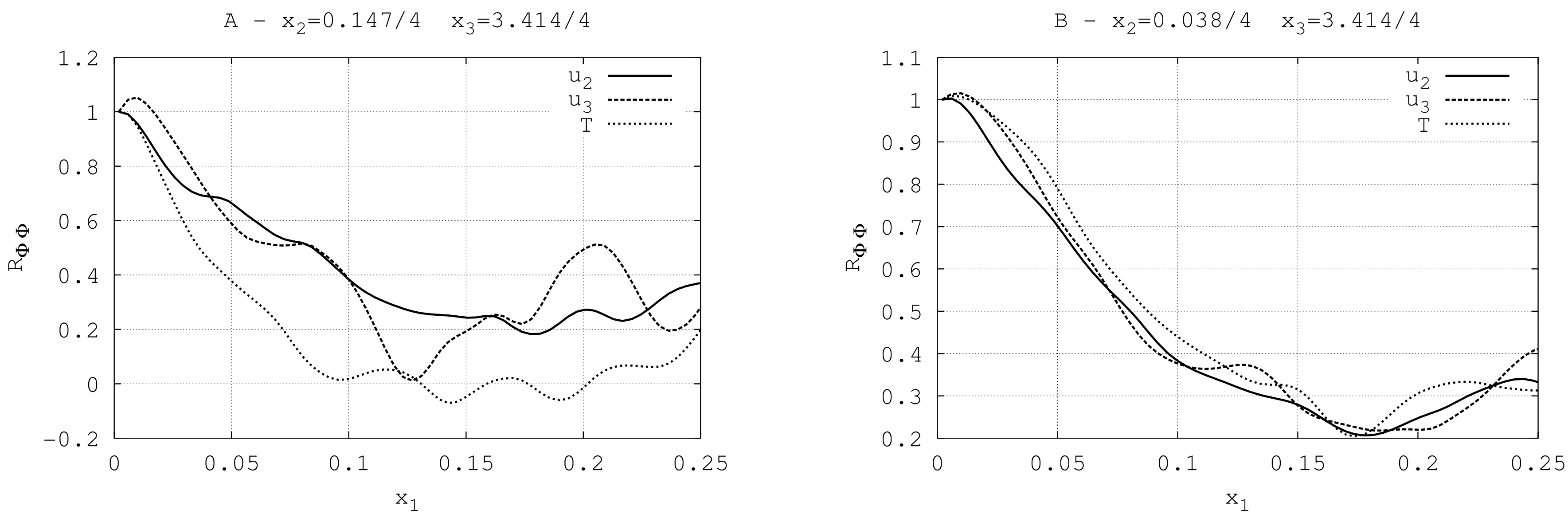


Verification of DNS - Homogeneous directions

Is domain length enough ?

Is the domain long enough to allow the largest structures to develop?

If so, two distant points must be uncorrelated.



Computational domain is considered **large enough** if **uncorrelated turbulent fluctuations at a separation of one-half of the domain size** are obtained. (As the domain is periodic, afer one-half, the correlation must increase).

Verification of DNS - Homogeneous directions

Is the resolution enough at the homogeneous direction ?

The averaged energy spectrum is used to determine if the grid resolution is sufficient in the homogeneous direction (here x_1).

$$E_{\phi\phi}\left(k_1, x_2, x_3\right) = \left\langle \hat{\phi}_{k_1}\left(x_2, x_3\right) \hat{\phi}^*_{k_1}\left(x_2, x_3\right) \right\rangle$$

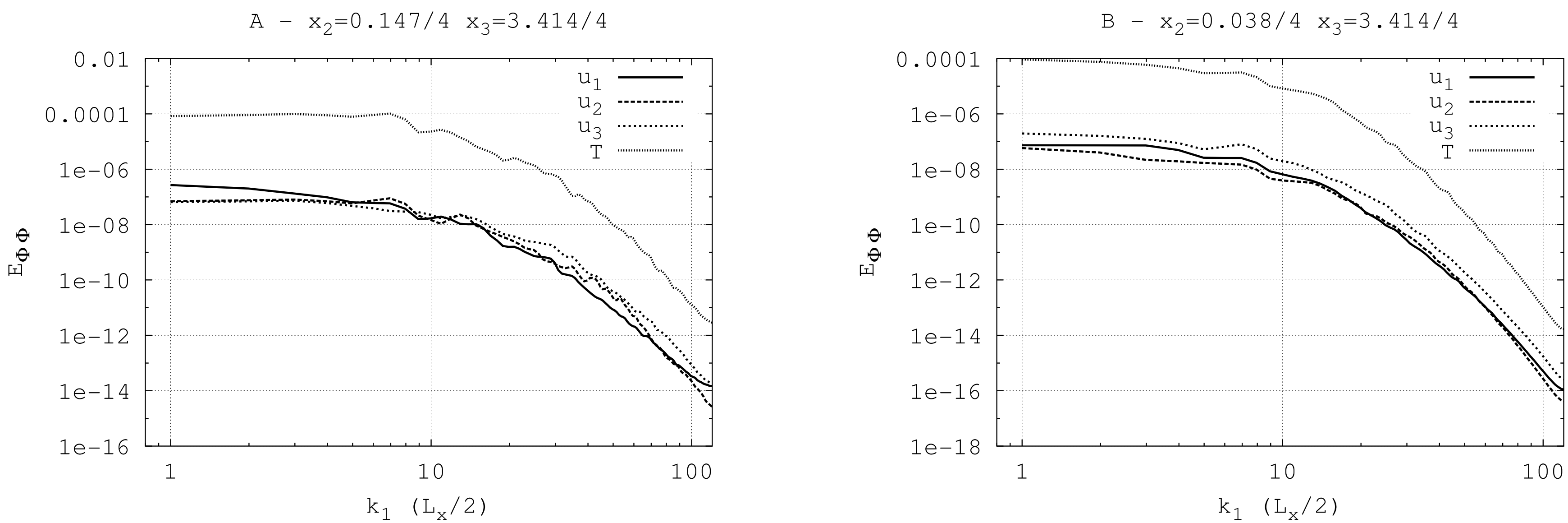
where $(\cdot)^*$ represents the complex conjugate and $\hat{(\cdot)}$ stands for a forward Discrete Fourier Transform (DFT)

$$\hat{\phi}_k = \frac{1}{N} \sum_{j=1}^{N} e^{i\frac{2\pi jk}{N}} \phi_j$$

Verification of DNS simulations

One-dimensional energy spectra

Different (x_2, x_3)-locations are monitored for each case.



Grid resolution would be considered as enough if:

- **No pile-up** at high wavenumbers.
- Energy density between smallest and largest wavenumbers has **dropped several orders of magnitude**.

Solution of the quiz

$< u'_1 u'_1 > \neq 0$ (Of course as it is the average of a non-negative quantity. Unless it was a steady flow, then $u\prime_1 = 0$)

$< u'_1 T' > = 0$ (In some cases, at a given point, a positive fluctuation of u_1 will be associated with a positive fluctuation of T. In other cases, with a negative fluctuation of T. AS THE FLOW IS STATISTICALLY UNIFORM IN X DIRECTION, on average, there is no reason why one should be more frequent than the other.)

$< u'_3 T' > \neq 0$ (This is different, u'_3 IS correlated with T'. If, for instance, near the how wall the vertical velocity increases, it is to be expected that the temperature increases).

$< u'_2 u'_2 > \neq 0$

$< u'_3 u'_3 > \neq 0$

$< u'_1 u'_2 > = 0$

$< u'_1 u'_3 > = 0$

$< u'_2 u'_3 > \neq 0$