PHYSICAL REVIEW E VOLUME 54, NUMBER 4 OCTOBER 1996

Statistical uncertainty in the analysis of structure functions in turbulence

R. Camussi,* C. Baudet, R. Benzi,† and S. Ciliberto‡
Ecole Normale Supérieure de Lyon, CNRS URA 1325, Physique Recherche, 46 Alleé d'Italie, 69364 Lyon Cedex 06, France
(Received 6 December 1995; revised manuscript received 15 July 1996)

This Rapid Communication is about a procedure for the statistical uncertainty estimation of velocity structure functions measured in turbulent flows. The proposed method is based on the determination of the number of statistically independent samples calculated by the correlation function of the velocity increments and it is applied to experimental data obtained by hot wire measurements in a turbulent jet flow.
[S1063-651X(96)50310-7]

PACS number(s): 47.27.−i, 05.40.+j

The most common approach for the statistical study of locally isotropic turbulent flows, is based on the analysis of the velocity increments field (see, for example, [1]) and, in particular, of the p-order velocity structure function that can be defined as follows:

$$\langle \Delta V(\tau)^p\rangle = \langle [V(\tau+t)-V(t)]^p\rangle, \tag{1}$$

where the symbol $\langle\ \rangle$ represents the average over an appropriate ensemble. In the last few decades, great efforts have been made by experimentalists in order to measure with sufficient accuracy the statistical quantities of Eq. (1), and to obtain reliable data for comparing the results with available theories, mainly with those concerning the scaling properties of the velocity structure functions (see, e.g., [2] for a review). Huge amounts of data have been acquired, mostly by the hot wire anemometry technique, but the correct estimation of the statistical errors remains today an important and still debated task (see, e.g., [3–6]).

It is well known from statistical theories that the statistical accuracy increases for increasing number of acquired samples. The general definition of the error related to the statistical uncertainty, and in particular to the lack of *ergodicity* due to finite length of the time acquisition window, can be given in fact by the following expression, which is valid for stationary statistics:

$$\epsilon_p(\tau) = \left[\frac{\langle \Delta V(\tau)^{2p}\rangle}{\langle \Delta V(\tau)^p\rangle^2} - 1\right]^{1/2} \frac{1}{N_{ind}^{1/2}}. \tag{2}$$

Strictly speaking Eq. (2) is valid if the central limit theorem can be applied; that is, if the correlation function decays much faster than $1/t$. This is not the case for turbulence [7] at small t. Nevertheless, correlation functions go to zero after a certain amount of time. This allows us to use Eq. (2) also for turbulent data if N_{ind} is correctly estimated. The difference between ϵ_p evaluated using independent data records, and ϵ_p, estimated through Eq. (2) with N_{ind} computed following the procedure proposed in this paper, is negligible. Indeed, if Eq. (2) is used, the estimation of ϵ_p, depends not only on the number of samples acquired, but also on the correct evaluation of the number of statistically independent samples, N_{ind}. This aspect has been well focused on in [8] (§ 6.4) and [9]. These authors suggest that a correct estimation of the distance between two successive independent samples should be of the order of twice the integral length. Other authors consider statistically independent samples by choosing the sampling interval of the order of the Taylor microscale (see, e.g., [10]). Other approaches were proposed, e.g., by Anselmet *et al.* [11]. Unfortunately, the proposed procedures do not give univocal methods for the estimation of N_{ind}, and, most important, a correct definition of $N_{ind}(\tau)$, i.e., as a function of the turbulent *eddies* size, is not given. Furthermore, it is well known that intermittency yields non-Gaussian statistics of the velocity increments at small separations (see, e.g., [12]). Thus the definition of N_{ind} should be modified to account for the different statistics observed at different scales. The procedure we propose is a modification of Tennekes and Lumley, [8] approach for the estimation of N_{ind}. Our aim is to evaluate the number of samples that are correlated at each scale τ that is for each characteristic temporal separation of the velocity increments. The distance between two successive samples is determined by the sampling rate adopted for data acquisition whereas two successive statistically independent samples should be separated by a distance larger than their correlation length. Therefore, in order to evaluate the number of correlated samples at each τ, we compute the correlation function of the velocity increments, as follows:

$$C(\delta,\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T [V(t+\tau+\delta)-V(t+\delta)] \times [V(t+\tau)-V(t)]dt, \tag{3}$$

where δ represents the correlation time and τ the velocity scale. At each τ it is possible to calculate the characteristic correlation time scale $T_\Delta(\tau)$ as follows:

$$T_\Delta(\tau) = \frac{\int_0^\infty \delta |C(\delta,\tau)| d\delta}{\int_0^\infty |C(\delta,\tau)| d\delta}, \tag{4}$$

*On leave from Universitá di Roma Tre, Dip. Mecc. ed Autom., Via C. Segre 60, 00146 Roma, Italy.

†Permanent address: Universitá di Roma ''Tor Vergata,'' Dip. di Fisica, Via della Ricerca Scientifica 1, 00133, Italy.

‡Also at Dipartimento di Fisica Universitá di Firenze, Italy.

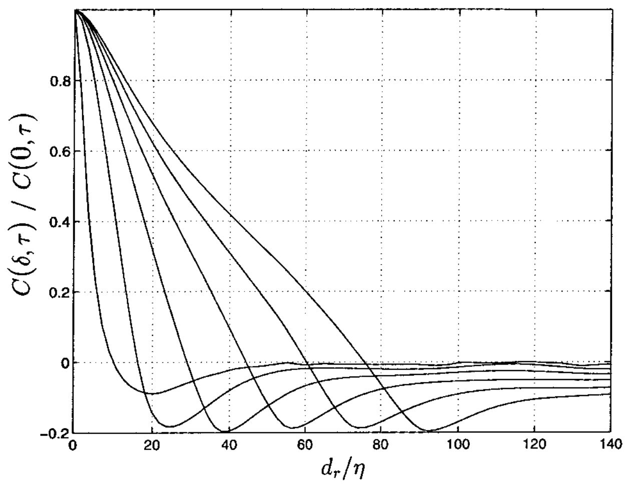

FIG. 1. Correlation function $C(\delta,\tau)$, evaluated at small scales $d_r = \bar{U}\delta$. Each curve is plotted for fixed t_r, which ranges from 5 (first curve on the left) to 250η.

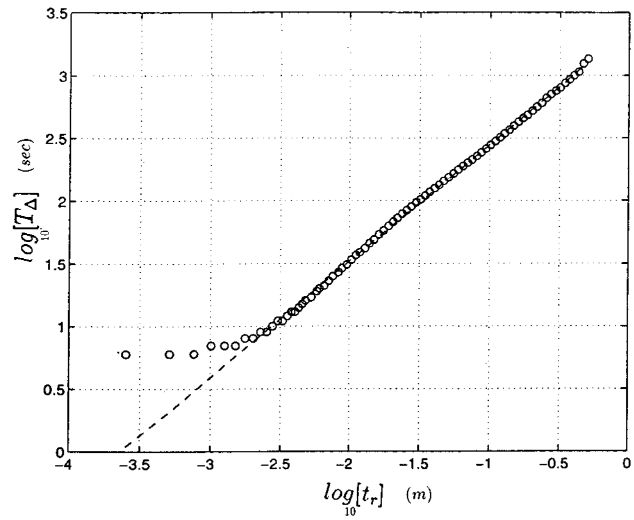

FIG. 2. Log-log plot of T_Δ as a function of $t_r = \bar{U}\tau$.

and the number of statistically independent samples will be also a function of τ, since $N_{ind}(\tau) = [T_{tot}/T_\Delta(\tau)]$. At each scale τ the ratio $T_\Delta/\Delta t$, where Δt is the sampling interval, gives the number of samples which are assumed to be correlated and therefore not independent from the statistical point of view. Once $N_{ind}(\tau)$ is determined, Eq. (2) is used to compute properly the statistical error $\epsilon_p(\tau)$. As indicated in Eq. (2) the value N_{ind} is not dependent on the moment order p. This is obviously related to the possibility of calculating the p-order moments directly from the probability distribution functions (PDFs) of ΔV and not from the ensemble average of ΔV^p. This point has been checked by the estimation of the correlation of order p of the velocity differences, i.e., by using ΔV^p in Eq. (3). We found that the larger the moment order p, the shorter the correlation time. Thus, one will arrive at the conclusion that the number of independent samples increases with the moments order p, which is physically incorrect. Therefore the number N_{ind} calculated from Eq. (3) gives the correct estimation of the statistically independent samples, whereas the dependence from p is accounted for when estimating the error from Eq. (2).

Experimental measurements are conducted in a turbulent flow generated by a jet. Data are acquired by means of a single probe hot wire anemometer (probe TSI 1260 of length $l_w = 500$ μm). The jet diameter D is 12 cm and the probe is positioned in the fully developed region at $x/D \simeq 25$. The mean velocity is $\bar{U} = 8.4$ m/sec and Re_λ (based on the Taylor microscale λ and on the velocity standard deviation) is on the order of 800. 1.6×10^7 samples have been acquired with sampling interval corresponding to about 5 Kolmogorov length that is on the order of the probe size. The anemometer signal is properly filtered to avoid aliasing errors and it is digitized by a 16 bit analog-to-digital converter. The error given by Eq. (2) is then calculated for structure functions of order from 1 to 6 and for different amounts of samples (from 10^6 to 1.6×10^7).

The correlation function $C(\delta,\tau)$ is computed as a function of $d_r = \bar{U}\delta$ and for different separations $t_r = \bar{U}\tau$. In Fig. 1, $C(\delta,\tau)/C(0,\tau)$ calculated at small t_r is presented. Each curve corresponds to different t_r, which ranges from 5 up to 250η. The abscissas corresponding to negative-slope zero-crossings have been checked to be reliable indicators of $\bar{U}T_\Delta(\tau)/\eta$ that are the normalized correlation lengths. For example, at the smallest t_r (that corresponds to a separation equal to the sampling interval), we found a positive correlation function for about the first seven samples that corresponds to a length scale on the order of the Taylor microscale. Therefore, at this scale, the number of statistically independent samples is $N_{tot}/7 \simeq 2.4\times10^6$. Results analogous to those of Fig. 1 are obtained at larger scales (not reported here) and, as in the previous case, the correlation length increases for increasing separations t_r. In Fig. 2 we present the correlation length T_Δ as a function of t_r. Apart from the smallest separations, the increase of T_Δ with τ follows a power law with exponent ~ 1. In order to better analyze the dependence of T_Δ over τ we tried to determine a self-preserving form of $C(\delta,\tau)$. In Fig. 3 we show $C_a(\delta,\tau)/C_a(0,\tau)$ as a function of the normalized separation δ/τ. The correlation function $C_a(\delta,\tau)$ is calculated as shown

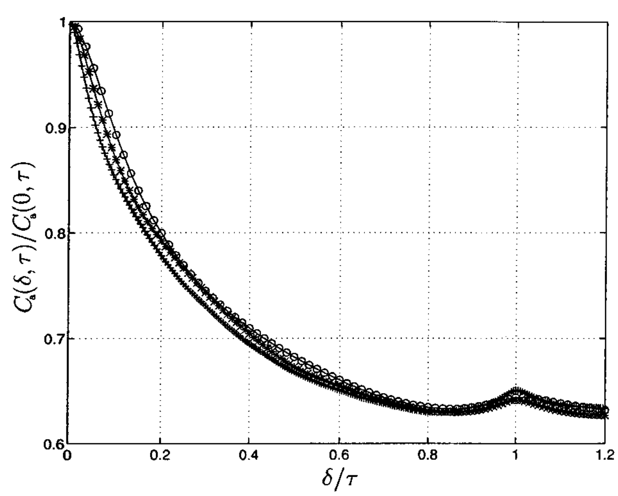

FIG. 3. Self-preservation of $C_a(\delta,\tau)$. o corresponds to $t_r \sim 100\eta$, * to $t_r \sim 200\eta$, and + to $t_r \sim 400\eta$

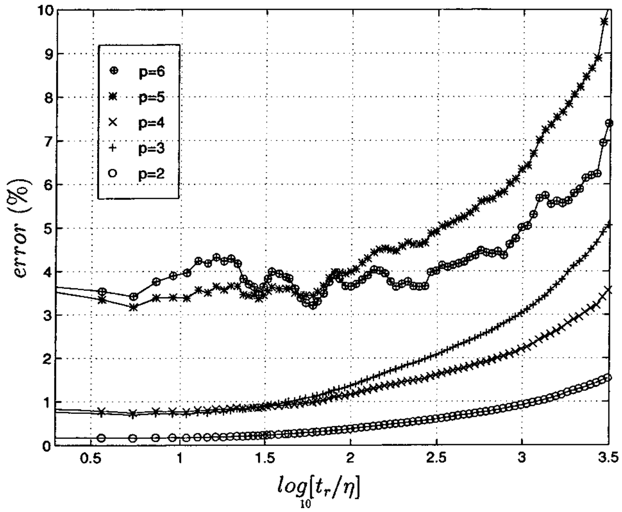

FIG. 4. Statistical error calculated for p from 2 to 6.

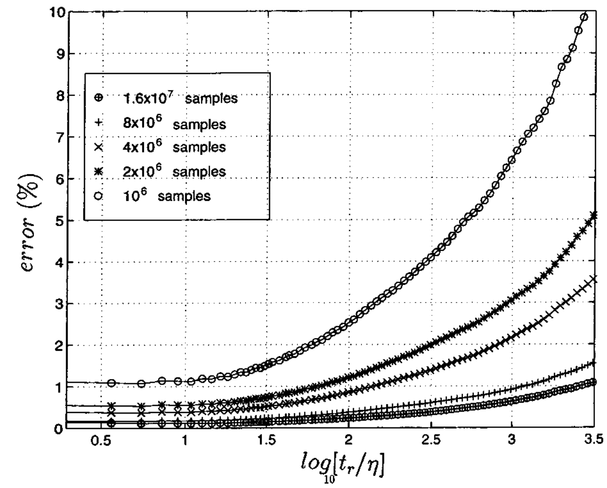

FIG. 5. Statistical error calculated for $p=2$ and increasing data samples.

in Eq. (3) but with the use of the modulus of the velocity differences. The similarity of the curves is achieved with good accuracy over the whole range of δ/τ and for different values of τ (which, also in this case, correspond to different curves). This result suggests that the correlation function can be written in the following form:

$$C_a(\delta,\tau)/C_a(0,\tau)=f(\delta/\tau), \tag{5}$$

where $f(\delta/\tau)$ is a *universal* function whose analytical form is presently under study by the authors [7]. It is important to notice that a necessary condition for Eq. (5) to be used, is that d_r and t_r belong both to the inertial or to the dissipative range of scales.

The error ϵ_p defined in Eq. (2) is shown in Fig. 4 for p ranging from 2 to 6 as a function of the normalized spatial separation t_r/η. It has to be pointed out that the error calculated for $p=6$, which is around 5%, is on the order of previous results (e.g., [6]) only if its averaged value (over the whole range of scales) is considered. As expected, the error increases with p due to the decreasing reliability of the PDFs. In fact, if $P(\Delta V)$ is the PDF of the velocity difference, reliable results are obtained if, for any p, the number of statistically independent samples is large enough to permit the proper calculation of the p-order moments of the velocity difference, which are defined as follows:

$$\langle \Delta V^p\rangle=\int_0^\infty P(\Delta V)\Delta V^p d(\Delta V). \tag{6}$$

The larger p is the shorter the range where $P(\Delta V)\Delta V^p$ can be correctly integrated. In order to obtain better reliability, the statistical error of odd order moments is computed by averaging the absolute values of the velocity differences. We checked that the use of the modulus increases the accuracy of a factor $\geqslant 2$, however, as shown in Fig. 4, even order moments are more precise than the odd ones even when the latter are calculated with the use of the absolute value. For small p $(p<6)$, the error increases for increasing separations t_r. This is related to the increase in magnitude of T_Δ and the consequent decrease of N_{ind}. Nevertheless, for $p=6$, $\epsilon_p(t_r/\eta)$ remains about constant apart from the largest scales. This trend is related to the shape of the PDFs of the velocity differences, that, at the smallest scales, becomes exponential (see, e.g., [12]). Therefore, for high p and small scales, the error evaluated at small t_r becomes higher because, as follows from Eq. (6), the PDFs tails are not correctly resolved. Same conclusions can be achieved from Fig. 5, where ϵ_p is plotted for $p=2$ and increasing amount of data. In this case we observed, for fixed t_r, a decrease of ϵ_p for increasing number of samples. At the largest scales, ϵ_p decreases as $N_{tot}^{-1/2}$ This is the expected trend since Gaussian statistics are the appropriate for large scales. At the smallest scales the decrease of ϵ_2 with N_{tot} is instead faster with a decay exponent close to -1. This is due to the statistics of the small scales that are closer to being exponential than to being Gaussian [13]. These results are evidence that the dependence of ϵ_p over the *eddy* scales τ, significantly affects the uncertainty estimation. Therefore, since such dependence is not trivial, it should always be taken into account for a correct evaluation of the error bars of the velocity structure functions.

As a conclusion, a procedure for the evaluation of error bars related to statistical uncertainties is presented. Attention is focused on the velocity structure functions, and statistical errors are computed by the estimation of the number of statistically independent samples obtained by the correlation function of the velocity increments. The proposed procedure is applied to velocity data acquired in a turbulent jet flow. A self-preserving behavior of the correlation function has been observed, and error bars, for structure functions ranging from the second to the sixth order, have been calculated. The dependence of the statistical error over the length scales has been evidenced and it has been related to the different statistics observed for increasing separations. Finally, we point out that in our data the errors related to the measurement technique (hot wire measures) are negligible with respect to the statistical uncertainty (see, e.g., [14]). Of course, if measurement errors are instead significant, they have to be added to the statistical errors.

We thank G. Stolovitzky for very useful comments which improved the paper a lot. This work has been partially supported by the EEC Contract No. ERBCHRXCT940546.

[1] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics* (MIT, Cambridge, MA, 1975), Vol. 2.
[2] K. R. Sreenivasan, Annu. Rev. Fluid Mech. **23**, 539 (1991).
[3] G. Stolovitzky and K. R. Sreenivasan, Phys. Rev. E **48**, 82 (1993).
[4] R. Benzi, S. Ciliberto, and G. R. Chavarria, Physica D **80**, 385 (1995).
[5] R. Camussi and G. Guj, Exp. Fluids (to be published).
[6] E. Arneodo *et al.*, Europhys. Lett. **34**, 411 (1996).
[7] R. Camussi, S. Ciliberto, R. Benzi, and C. Baudet, Phys. Fluids **8**, 1686 (1996).
[8] H. Tennekes and J. L. Lumley, *A First Course in Turbulence* (MIT, Cambridge, MA, 1972).
[9] H. Tennekes and J. C. Wyngaard, J. Fluid Mech. **55**, 93 (1972).
[10] M. Briscolini and P. Santangelo, J. Fluid Mech. **270**, 94, 199 (1994).
[11] F. Anselmet, Y. Gagne, E. J. Hopfinger and R. A. Antonia, J. Fluid Mech. **140**, 63 (1984).
[12] B. Castaing, Y. Gagne, and E. J. Hopfinger, Physica D **46**, 177 (1990).
[13] N_{ind} is computed for each τ, therefore $N_{ind}(\tau)$ is not proportional to N_{tot}. Thus $\epsilon_p \propto N_{ind}^{-1/2}$ always by construction. However, the dependence of ϵ_p on N_{tot} may be different for different scales.
[14] R. Camussi, G. Guj, D. Barbagallo, and F. Stella, Phys. Fluids **8**, 1181 (1996).