LEARNING STOCHASTIC MODELS OF TURBULENCE

Gautier Tandeau de Marsac*
TAF MCE, IMT Atlantique, France

BACKGROUND

Understanding the statistical structure of fully developed turbulence remains a major challenge in classical physics. The foundational theory, known as K41, was proposed by Kolmogorov [1]. It hypothesized a universal local structure for velocity fluctuations based on a constant mean energy dissipation rate, $\langle \epsilon \rangle$. The fundamental quantity of interest in this framework is the longitudinal velocity increment $\delta_\ell u(x)$ at scale ℓ, defined as the difference in velocity between two points separated by a distance ℓ:

$$\delta_\ell u(x) = u(x+\ell) - u(x) \tag{1}$$

Based on dimensional analysis, Kolmogorov predicted that the variance of these increments scales as a power law with an exponent of 2/3, i.e., $\langle (\delta_\ell u)^2 \rangle \sim \ell^{2/3}$. However, experimental observations quickly revealed that velocity fluctuations at small scales are strongly non-Gaussian, a phenomenon known as intermittency. This implies that the energy dissipation is not uniform but fluctuates significantly in space and time.

In 1962, Kolmogorov-Oboukhov (K62) refined this original theory. This refinement posits that the local energy dissipation rate ϵ_r, averaged over a scale r, is itself a random variable. Specifically, K62 proposes that the logarithm of the dissipation, $\log \epsilon_r$, follows a normal (Gaussian) distribution [2].

This log-normal hypothesis became a cornerstone of modern turbulence modeling, forming the basis of the multifractal formalism [3, 4]. This formalism describes the scaling of velocity increment moments $M_n(\ell) = \langle |\delta_\ell u|^n \rangle \sim \ell^{\zeta_n}$ with a non-linear spectrum ζ_n, capturing intermittency [3]. Chevillard et al. (2012) provided a unified phenomenological theory based on this formalism, modeling velocity increments as a stochastic product of a Gaussian noise and a log-normal multiplier [3]:

$$\delta_\ell u \approx x_1 e^{x_2} \tag{2}$$

This model successfully reproduces the statistics of variance and flatness by defining a quadratic "singularity spectrum" $\mathcal{D}(h)$ [3]. Other approaches, such as Multifractal Random Walks (MRW), have also been proposed to construct stochastic processes that intrinsically generate these log-normal statistics [4].

AIM

This work aims to characterize the laws governing velocity increments in turbulent flows by challenging the classical log-normal hypothesis through a data-driven approach. Specifically, we question whether the components x_1 and x_2 in the stochastic relation $\delta_\ell u = x_1 e^{x_2}$ are necessarily Gaussian. By learning their probability distributions directly from experimental data to fit observed statistics (variance, flatness) across scales, we investigate if a more complex, non-Gaussian structure is required to accurately describe turbulence intermittency.

METHODS

We utilize the Multifractal Random Walk (MRW) framework [4], where velocity increments are modeled as $\delta_\ell u \approx x_1 e^{x_2}$. Here, x_1 represents fractional Gaussian noise (linear correlations), while x_2 (log-dissipation/intermittency) is derived from the convolution of an uncorrelated source noise $n(x)$ over the integral scale L. While standard theory assumes $n(x)$ is Gaussian, our approach aims to learn its true probability density function (PDF) from data.

The study uses a Modane wind tunnel dataset ($f_s = 25$ kHz, $\langle v \rangle = 20.5$ m/s, $R_\lambda = 2500$) containing 256 samples of velocity increment moments across 100 scales ranging from 4×10^{-5} s to 0.24 s. Starting from a reference distribution for $n(x)$, we simulate turbulent velocity samples and optimize the model by minimizing the Mean Squared Error (MSE) between computed and experimental moments.

The optimization proceeds in three stages:

1. **Baseline Log-Normal Model:** We assume $x_1, x_2 \sim \mathcal{N}$ and optimize physical parameters (c_1, c_2) to fit only the variance (S_2).
2. **Higher-Order Constraints:** We incorporate flatness into the loss function to test the Gaussian assumption against intermittency metrics.

* Corresponding author. E-mail: gautier.tandeau-de-marsac@imt-atlantique.net

3. **Learned Distribution:** We relax the Gaussian assumption for x_2, employing the Gumbel-Softmax relaxation technique [5] to differentiably learn the optimal shape of the x_2 distribution, ultimately identifying and validating a trimodal model.

RESULTS

For conciseness, we only present the results of the third stage. We relaxed the standard Gaussian noise assumption by directly learning the probability density function (PDF) of the intermittency noise $n(x)$ using a non-parametric approach. The distribution was discretized into 300 bins, and its shape was optimized alongside the physical parameters (c_1, c_2) to simultaneously minimize the error on both the variance (S_2) and the flatness $(\mathcal{F})$ of velocity increments.

Learned Noise Distribution

The optimization converged to a non-trivial noise distribution, significantly deviating from the Gaussian prior. As shown in Figure 1, the learned PDF (red curve) exhibits a trimodal uniform shape. This structural deviation from Gaussianity appears to be a key feature required to correctly capture the intermittency statistics at small scales.

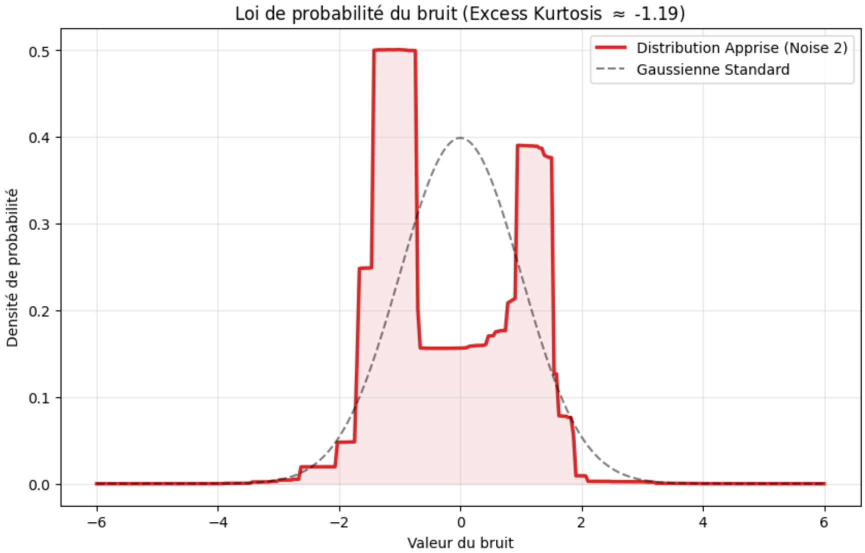

Figure 1. The learned probability density function of the intermittency noise $n(x)$ (red) compared to a standard Gaussian (dashed grey). The optimized distribution is clearly non-Gaussian, characterized by a negative excess kurtosis (-1.19).

Fit Quality and Convergence

This learned distribution allows the MRW model to achieve a remarkable agreement with experimental data for both second- and fourth-order statistics. Figure 2 demonstrates that the optimized model perfectly reproduces the scaling of the variance (S_2) across the inertial range, governed by a Hurst exponent $c_1 \approx 0.36$. Furthermore, and most importantly, it accurately captures the scale-dependence of the flatness, a feat that constrained Gaussian models struggle to achieve. The optimization process was stable, with the intermittency coefficient converging to a value of $c_2 \approx 0.02$, consistent with the accepted phenomenology of fully developed turbulence.

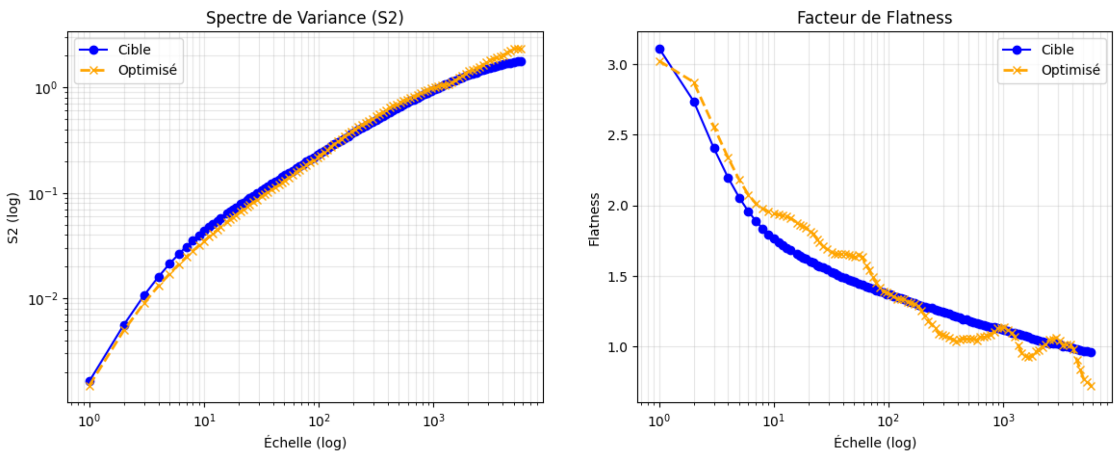

Figure 2. Comparison of experimental (blue) and simulated (orange) statistical moments. Left: The variance spectrum S_2 shows perfect scaling agreement. Right: The flatness scaling is accurately reproduced, validating the capability of the non-Gaussian noise model to capture intermittency.

References

[1] Kolmogorov A.N.: The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Doklady Akademii Nauk SSSR* **30**:299–303, 1941.

[2] Kolmogorov A.N.: A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.* **13**:82–85, 1962.

[3] Chevillard L., Castaing B., Arneodo A., Lévêque E., Pinton J.F., Roux S.G.: A phenomenological theory of Eulerian and Lagrangian velocity fluctuations in turbulent flows. *C. R. Physique* **13**:899–928, 2012.

[4] Bacry E., Delour J., Muzy J.F.: A multifractal random walk. *arXiv:cond-mat/0005405*, 2000.

[5] Jang E., Gu S., Poole B.: Categorical Reparameterization with Gumbel-Softmax. *Proc. ICLR*, 2017.