# Synthetic turbulent velocity field using statistics

Samy Braik

September 11, 2025

# Contents

# 1   Introduction

Turbulence is the term used to describe the unpredictable and complex behavior of a fluid's motion. It is characterized by vortex. It is opposed to a laminar flow, where the particles' movement is pre. Formally, the behavior of a fluid is quantified by the Reynolds number (Re), the higher this number is the more turbulence the fluid is.

In this work, we study an ideal type of turbulence flow called Homogeneous and Isotropic Turbulence (HIT).
There are few properties that characterized this type of flow. It is homogeneous which means that the statistical properties (mean, variance, correlation functions) do not depend on position in space, and it is isotropic which means that these statistics are invariant under rotations or reflections of the coordinate.
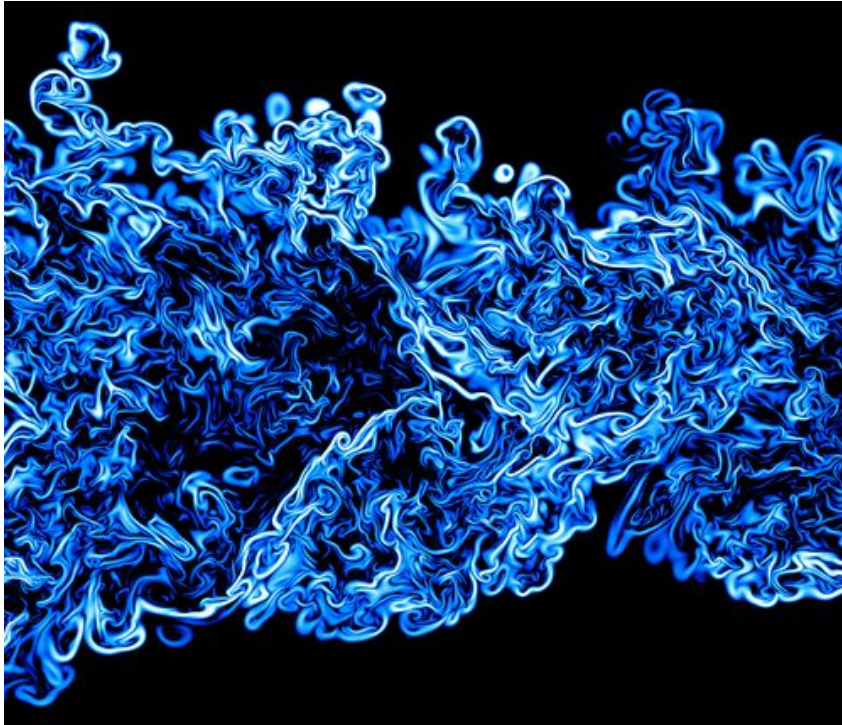


Figure 1: Slice through the field of the scalar dissipation reveals the small-scale structure of turbulence from CNRS UMR 6614 CORIA and JSC

When the goal is to use statistical methods to produce turbulent velocity field two paths are worth considering. The first one is to simply use a statistical method that generate the velocity field directly. This is done in [7],[6] or [4].
Here, generative models seem the most suitable. Two major difficulties could be encountered, the huge cost and the poor generalization of the method. Indeed, in order to train such a model, there is a need for highly resolved turbulence usually produced with Direct Numerical Simulation (DNS) which are costly numerical methods. Apart from that, there is the natural cost of the method itself that can't be sidelined. The poor generalization of the method stem from the fact that usually, the models are trained on fields with a specific Reynolds number which leads to poor robustness if the turbulence parameters

are changed.

The other path would be to start from an already existent method. For example in the case of Reynolds-averaged Navier–Stokes equations (RANS) technique (see [<**empty citation**>] for details), we can use statistical methods to learn a closure model, like it is done in [1]. This allows a better theoretical guarantee on the produced and also get rid of the robustness problem from the previous case.

# 2 Model

## 2.1 Random Fourier

We set ourselves in the random Fourier model developed in [3] and briefly remind it. Some papers that study turbulence start from a velocity field generated by DNS and use Fourier transform to study the field properties in the spectral space. The motivation behind this model, is to perform an inverse Fourier transform to generate a synthetic turbulent field with a good choice of Fourier coefficient.

This leads to the following expression of the velocity field

$$u^s(x,t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \hat{u}(\kappa,\omega) e^{l\psi(\kappa,\omega)} \sigma(\kappa,\omega) \right] e^{l(\kappa \cdot x + \omega t)} d\kappa d\omega \tag{1}$$

Discretizing the expression in N random Fourier modes and taking into account that the field is real, it could be written

$$u^s(x,t) = 2 \sum_{i=1}^{N} \hat{u}_n \cos(\kappa^n \cdot x + \psi_n + \omega_n t)\sigma^n \tag{2}$$

Furthermore, we place ourselves in the following work in a froze turbulence which means that we ignore the effect of time. Therefore, the expression if simplified and we have

$$u^s(x,t) = 2 \sum_{n=1}^{N} \hat{u}_n \cos(\kappa^n \cdot x + \psi_n)\sigma^n \tag{3}$$

For each $n^{th}$ Fourier mode associated with the wave vector $\kappa^n$, $\hat{u}_n$ is the amplitude, $\psi_n$ the phase, $\sigma^n$ the direction and $\omega_n$ the time-frequency.

Like previously mentioned, a good choice of the Fourier coefficient is needed in order to produce a realistic velocity field. [3] features a discussion on the choice of $\hat{u}, \kappa, \psi_n$ and $\sigma$.
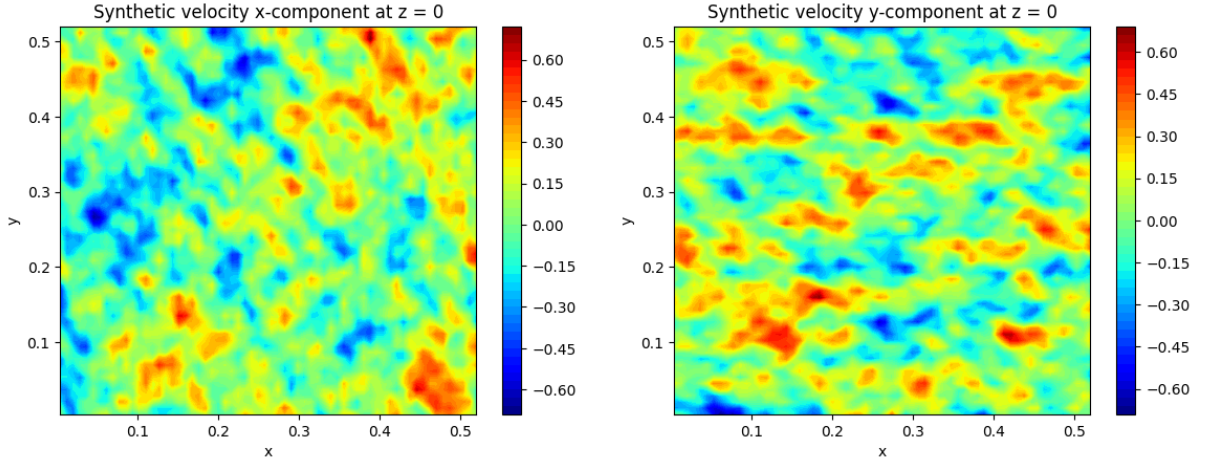
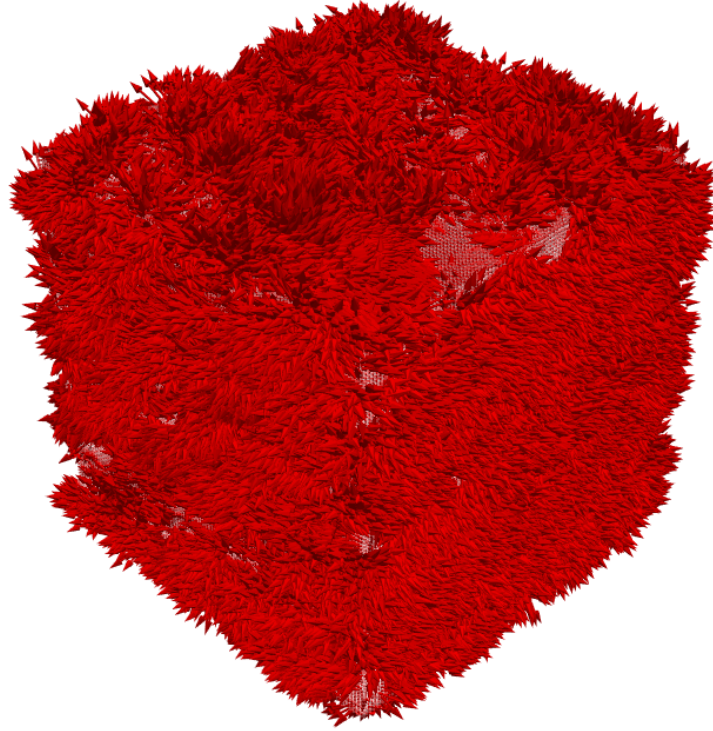Figure 2: Synthetic velocity obtain using the Random Fourier model



Figure 3: Synthetic velocity obtain using the Random Fourier model represented by arrows

## 2.2 Coefficient choice

To ensure the statistical isotropy of the generated field, the wave vector $\kappa$ is chosen randomly on the half-sphere. This leads to the following components choices

$$\kappa_1 = \sin(\theta)\cos(\varphi) \tag{4}$$
$$\kappa_2 = \sin(\theta)\sin(\varphi) \tag{5}$$
$$\kappa_2 = \cos(\theta) \tag{6}$$

with the angles chosen randomly according to the following probability density functions $f_\theta(x) = \frac{sin(x)}{2}$ and $f_\varphi(x) = \frac{1}{2\pi}\mathbb{1}_{[0,2\pi]}$.

Regarding the choice of $\sigma$, it is such that $\kappa \cdot \sigma = 0$ is verified, which is implied by the divergence-free condition $\nabla \cdot u^s = 0$.

Therefore,

$$\sigma_1 = \cos(\varphi)\cos(\theta)\cos(\alpha) - \sin(\varphi)\sin(\alpha) \tag{7}$$
$$\sigma_2 = \sin(\varphi)\cos(\theta)\cos(\alpha) + \cos(\theta)\sin(\alpha) \tag{8}$$
$$\sigma_3 = -\sin(\theta)\cos(\alpha) \tag{9}$$

with $\alpha$ chose according to a uniform distribution on $[0, 2\pi]$ i.e. $f_\alpha(x) = \frac{1}{2\pi}\mathbb{1}_{[0,2\pi]}$.

The phase coefficient $\psi$ is randomly chosen according to $\mathcal{U}[0, 2\pi]$ to ensure spatial homogeneity.

Lastly, the amplitude $\hat{u}_n = \sqrt{E(\kappa_n)\delta\kappa_n}$ where $\delta\kappa_n = \frac{\log(\kappa_N)-\log(\kappa_1)}{N}$ and $\kappa_n = e^{(\log(\kappa_1)+n\delta\kappa_n)}$. The method to compute the prescribed energy spectrum $E(\kappa_n)$ is described in the next subsection.

## 2.3 Energy spectrum

Large eddies that carry energy distribute it through the energy cascade to smaller scales until it's dissipated by viscous effect at Kolmogorov scales. This energy transfer across different scales is characterized by the energy spectrum, noted $E(\kappa)$. It describes how much kinetic energy is attached to a give wave number $\kappa$.
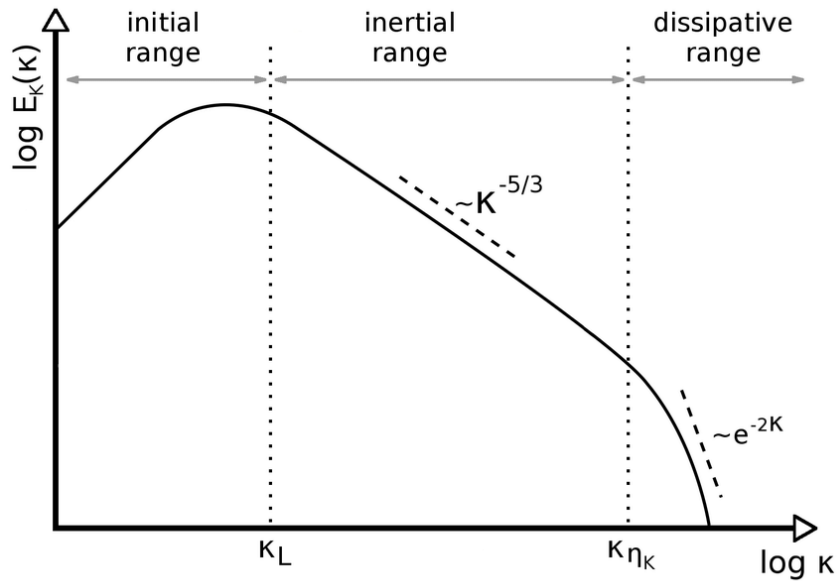


Figure 4: Energy spectrum of a turbulent flow from [5]

In the initial range, this is where the energy is injected into the flow. Could be from an obstacle, a rotating machinery or other phenomenon. The flow is being dominated with big whirlpools.

The inertial range is where the energy is transferred. The big whirlpools are being broke down into smaller whirlpool and so on. Finally, in the dissipative range, the viscous effect dominate and break down the turbulence.

In the context of homogeneous isotropic turbulence, Kolmogorov gave the following formula

$$E(\kappa) = C_k \kappa^{-5/3} \varepsilon^{2/3} \tag{10}$$

where $\varepsilon$ is the turbulent dissipation, $\kappa$ the wave number and $C_k$ the Kolmogorov constant.

Although, this formula captures the behavior in the inertial range and introduces this famous $-\frac{5}{3}$ decreases, it fails at reconstructing the spectrum in the initial and dissipative range i.e. at low and high wave numbers. To obtain a more complete recovery of the energy spectrum, [3] uses, among others, the von-Kármán Pao (VKP) energy spectrum.

It is defined by

$$E_{\text{VKP}}(\kappa) = \frac{2}{3} \alpha_e k L_e \frac{(kL_e)^4}{[(kL_e)^2 + 1]^{17/6}} \exp(-2(\kappa L_\eta)^2) \tag{11}$$
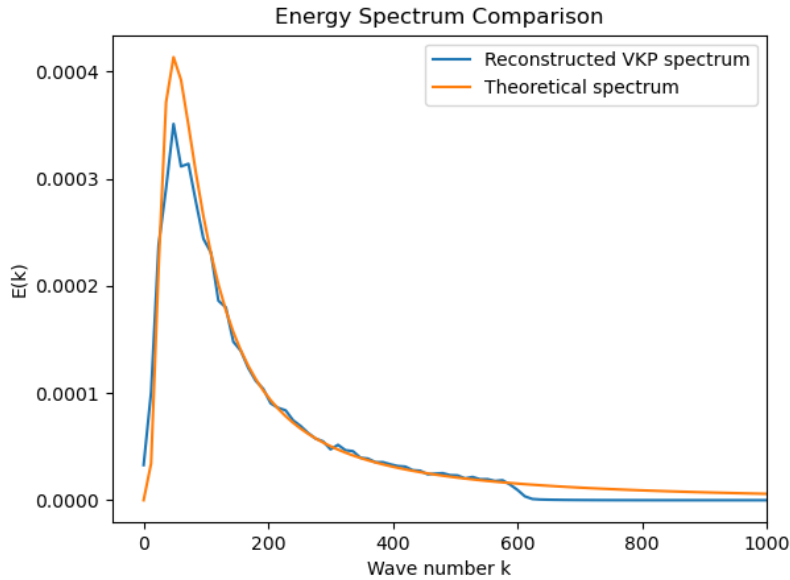


Figure 5: VKP energy spectrum vs theoretical spectrum

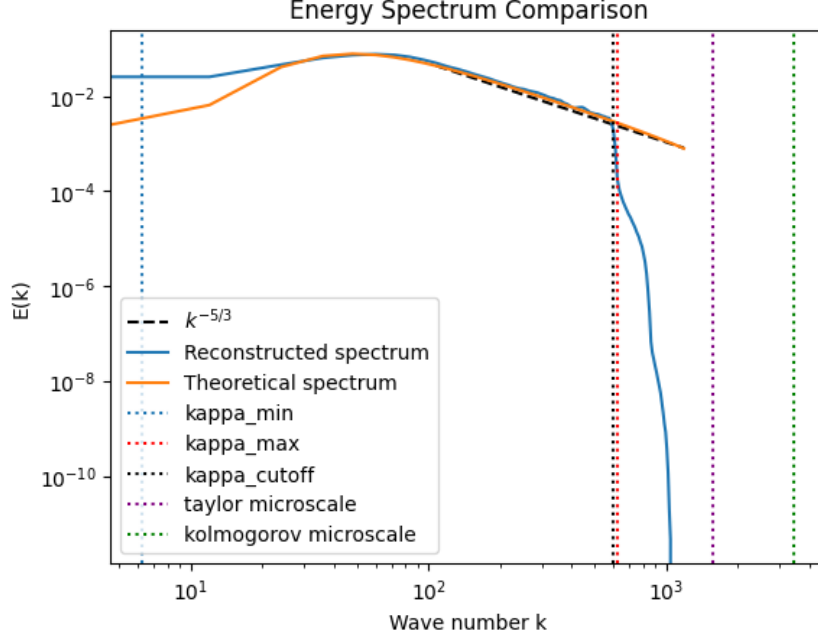In order to reconstruct the theoretical, computational ... is needed. The number of modes...

Figure 6: VKP energy spectrum vs theoretical spectrum with

## 2.4   Limitations and goals

The assessment of the synthetic velocity field's quality could be done by looking at few metrics. First, the reconstructed energy spectrum should match as close as possible the theoretical spectrum. Then, for each component, the average should be 0 and the Root Mean Square (RMS) values should match the one used to build the turbulence in the first place. Lastly, we look at the shape of the velocity increments.

**Definition 1** (RMS). Let $x = (x_1, \ldots x_n)$ a real vector. The RMS is defined by

$$x_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \tag{12}$$

**Definition 2** (Velocity increments). Let $u$ be a velocity field. For $r > 0$ and a point $x$, the velocity increments are defined by

$$\delta_r u = u(x + r) - u(x) \tag{13}$$

The model gives us the following results for each metric

| Direction | Value |
|-----------|-------|
| x | -0.0026918877847492695 |
| y | -0.0001003775978460908 |
| z | -0.0012576839653775096 |

Table 1: Velocity mean

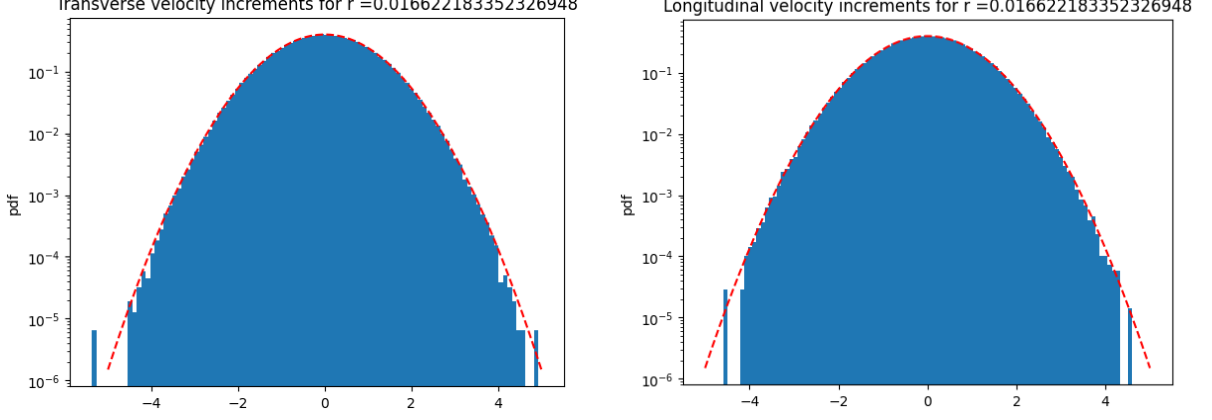| Direction | Value (expected : 0.222) |
|-----------|--------------------------|
| x | 0.19583044946193695 |
| y | 0.17598901689052582 |
| z | 0.19306901097297668 |

Table 2: RMS speed

Figure 7: Velocity increments in blue with standard Gaussian in red dots

Although, the model retrieve a correct energy spectrum (Figure 5) and limited anisotropy, the velocity increments obtained are Gaussian. It is known that velocity increments appear to be non-Gaussian and more particularly heavy-tailed distribution. For instance [<**empty citation**>] shows that velocity increments display heavy taildness when $r$ is small in 13. It is due ...

Furthermore, the following figures show that the more turbulent a flow is, the heavier the tails of its velocity increments are.
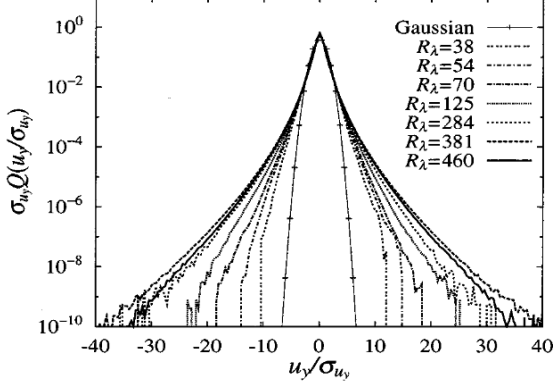


Figure 8: Variation of the transverse velocity derivative PDF with the Reynolds number from [2]
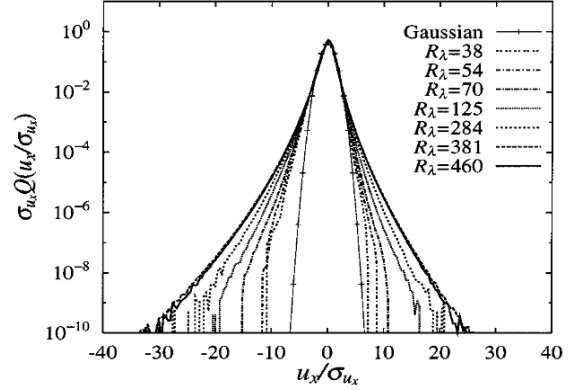


Figure 9: Variation of the longitudinal velocity derivative PDF with the Reynolds number from [2]

With the random Fourier model in mind, the goal is to identify which coefficient could lead to more realistic velocity increments without disrupting the other good properties that are already satisfied.

# 3 Method

To tackle the problem, no obvious path stood out. The one chosen is simply to define a model and a loss and optimize the model's parameters.

Since the goal is to retrieve heavy-tailed velocity increments, the logical criteria to look out is the kurtosis of those increments.

**Definition 3** (Kurtosis)**.** Let $X$ be a real random variable, $\mu$ its mean and $\sigma$ its standard deviation. Its kurtosis is defined by

$$\text{kurt}(X) = \mathbb{E}\left[\left(\frac{X - \mu}{\sigma}\right)\right]^2 \tag{14}$$

**Remark 1.** For a Gaussian distribution, which is the distribution of the velocity increments in the first place, the kurtosis is equal to 3.

**Remark 2.** In practice, we look at the flatness which is simply the kurt $- 3$.

Therefore, in order to push for heavy tailed velocity increments, we came up with a straight forward flatness loss.

**Definition 4** (Flatness loss)**.** Let $F_{\text{target}}$ be the target flatness, $x = (x_1, \ldots, x_n)$ be the input vector.

$$\mathcal{L}_F(x) = \frac{1}{N} \sum_{i=1}^{N} (\text{kurt} - 3 - F_{\text{target}})^2 \tag{15}$$

To compute the energy spectrum as close as the theoretical spectrum, we simply look at the MSE with an added smoothness term. This gave us the following loss

**Definition 5** (Energy spectrum loss)**.** Let $\alpha > 0$, $E_{\text{rec}}$ the reconstructed spectrum and $E$theory the theoretical spectrum. We define a smoothness term by

$$S(E_{\text{rec}}) = \frac{1}{N - 2} \sum_{i=2}^{N-1} (E_{\text{rec},i-1} - 2E_{\text{rec},i} + E_{\text{rec},i+1}) \tag{16}$$

The loss is then defined by

$$\mathcal{L}_{ES} = \frac{1}{N} \sum_{i=1}^{N} (E_{\text{rec},i} - E_{\text{theory},i})^2 + \alpha S(E_{\text{rec}}) \tag{17}$$

Most of the learning and training has been done simply on the parameter by defining them as torch Parameter.
[Mieux décrire le model (nn.Parameter)]

## 3.1 Preliminary work

In order to check the robustness of the random Fourier model in the first place, we conducted a simple experiment. We set the target flatness as 3, meaning we look for Gaussian velocity increments. Since it is the base model velocity increments, we want to know if we will retrieve the parameterization introduce in [3] or will it move away. We look up two situations : starting from Gaussian velocity increments and heavy-tailed velocity increments.
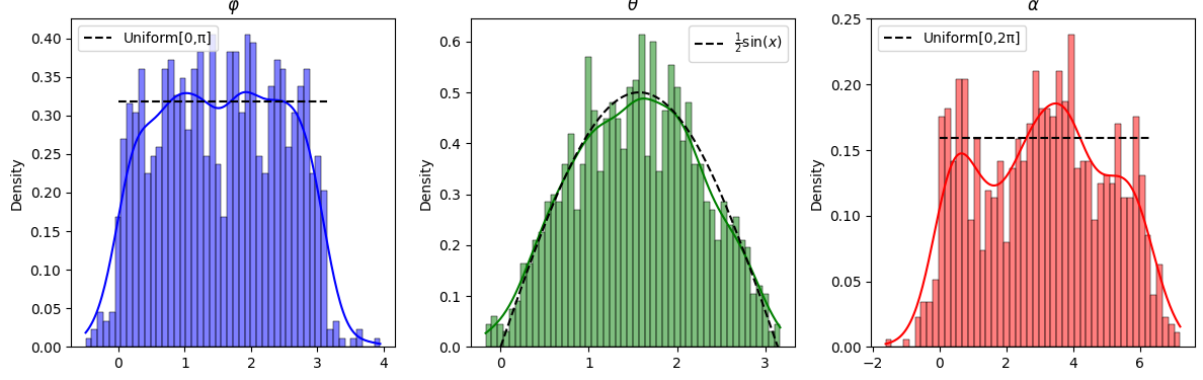
Figure 10: Angles distributions starting from Gaussian velocity increments

Here, the angles distributions stay sensibly the same, with a slight change in $\alpha$. This is expected since we already satisfy the target flatness.
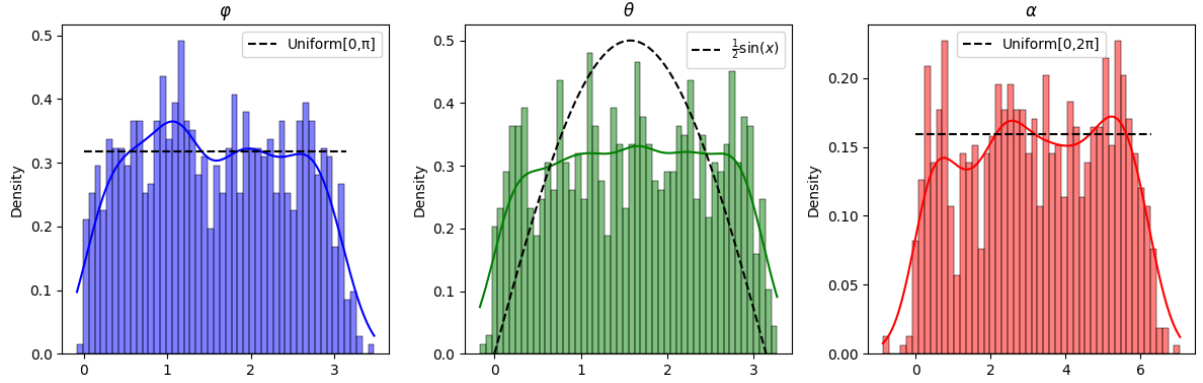


Figure 11: Angles distributions starting from heavy-tailed velocity increments

In this case, $\varphi$ and $\alpha$ stay relatively stable but $\theta$ is drawn to a more spread distribution.

In both cases Gaussian velocity increments are indeed recovered, the RMS speed are in the same range as in 2 and the generated velocity field does average to 0.

## 3.2   Phase coefficient

To improve the situation, our first idea was to work on the phase parameter $\psi$ in 3.

## 3.3   Phase, wave vector and direction

## 3.4   Angles

# 4   Results and Discussion

The work was highly exploratory and featured trials and errors along the way.

# 5   Conclusion

# A    Turbulence parameters

To generate the turbulence few physical and spectral parameters have to be given as input for the model. Throughout the report the parameters that have been used are listed in the following table.

| Parameter | Value | Unit |
|---|---|---|
| Length of the periodic box | $\frac{\pi}{6}$ | m |
| Number of modes | 250 or 500 | – |
| RMS speed | 0.222 | $\mathrm{m\,s^{-1}}$ |
| Integral length scale | 0.024 | m |
| Kinetic energy | $1.5 \times (0.222 \times 2)$ | $\mathrm{m^2\,s^{-2}}$ |
| Viscosity | $1.8 \times 10^{-5}$ | $\mathrm{m^2\,s^{-1}}$ |
| Minimum wave number | $\frac{2\pi}{1.0}$ | $\mathrm{m^{-1}}$ |
| Maximum wave number | $\frac{2\pi}{0.01}$ | $\mathrm{m^{-1}}$ |

# B    Variable selection

# C    Variable selection

To address the question of which angles are the most influential and which one could be negligible, we simply perform a case analysis. In the same training context, we varied individual angles while keeping others froze and untrained. It gave us the following results. Losses obtain for different parameter combination

Table 3: Losses obtain for different parameter combination

(a) 250 modes

(b) 500 modes

| Parameters | Loss | RMS (x, y, z) | Parameters | Loss | RMS (x, y, z) |
|---|---|---|---|---|---|
| $(0,0,\alpha)$ | 123.24456 | (0.24 , 0.08 , 0.12) | $(0,0,\alpha)$ | 0.73864 | (0.14 , 0.20 , 0.06) |
| $(0,\theta,0)$ | 1.50244 | (0.22 , 0.11 , 0.08) | $(0,\theta,0)$ | 0.74741 | (0.20 , 0.11 , 0.10) |
| $(0,\theta,\alpha)$ | 65.67394 | (0.11 , 0.22 , 0.09) | $(0,\theta,\alpha)$ | 0.79737 | (0.14 , 0.21 , 0.09) |
| $(\varphi,0,0)$ | 0.75362 | (0.09 , 0.24 , 0.06) | $(\varphi,0,0)$ | 0.75514 | (0.10 , 0.22 , 0.08) |
| $(\varphi,0,\alpha)$ | 0.74764 | (0.20 , 0.10 , 0.09) | $(\varphi,0,\alpha)$ | 0.74716 | (0.21 , 0.11 , 0.09) |
| $(\varphi,\theta,0)$ | 0.75258 | (0.12 , 0.22 , 0.04) | $(\varphi,\theta,0)$ | 0.74118 | (0.22 , 0.09 , 0.11) |
| $(\varphi,\theta,\alpha)$ | 2.15906 | (0.22 , 0.11 , 0.09) | $(\varphi,\theta,\alpha)$ | 0.75355 | (0.20 , 0.13 , 0.09) |

Overall, the more modes the better it is. A lot of the cases are pretty close together and lead to sensibly the same losses and in practice the desired flatness and a good recovery of the energy spectrum.

RMS values are affected after the learning step, depending on which angles are considered. For the moment, adding a loss on those RMS values tend to lead to less flat tails on the output velocities increments. A trade-off between RMS and flat tails seems to arise. For instance, in the case where all the parameters are learned, like in the previous section, the RMS for the z direction is off (0.09 with an expected value of 0.222).
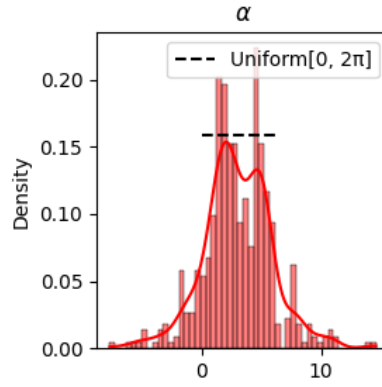
## C.1 Marginal distributions



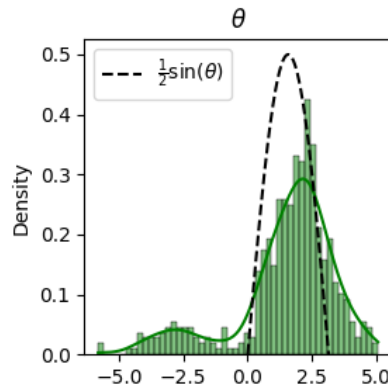Figure 12: Marginal density of the learned parameters $(\alpha)$
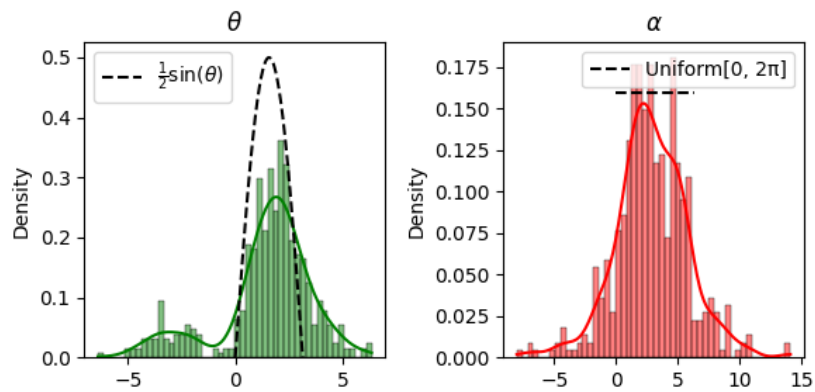


Figure 13: Marginal density of the learned parameters $(\theta)$



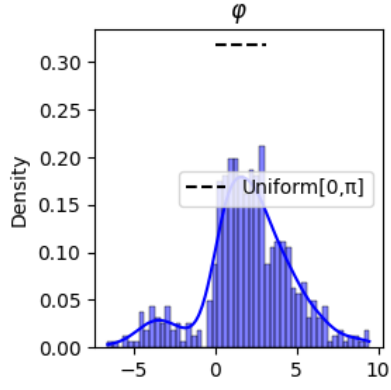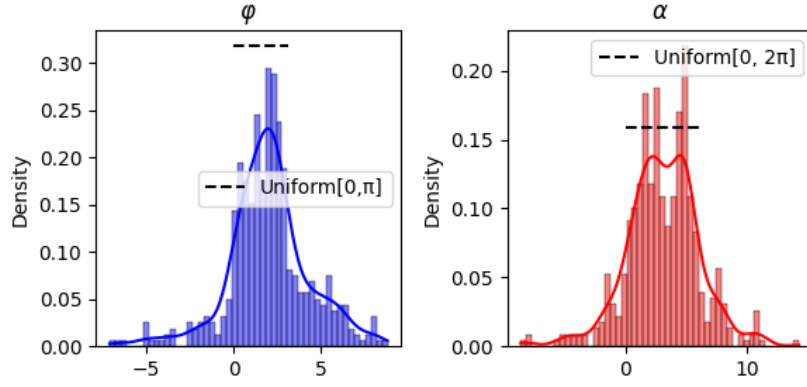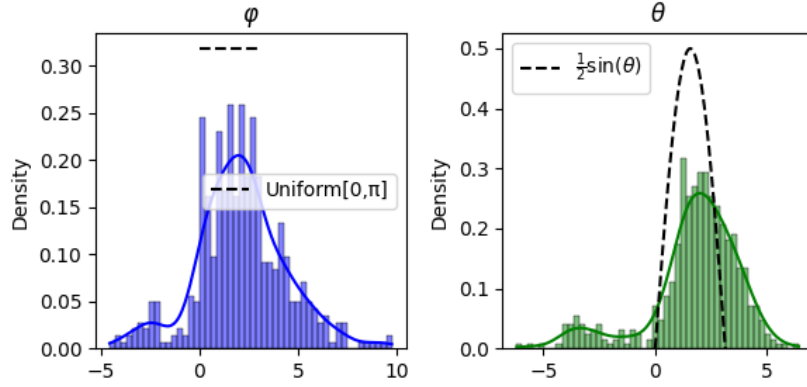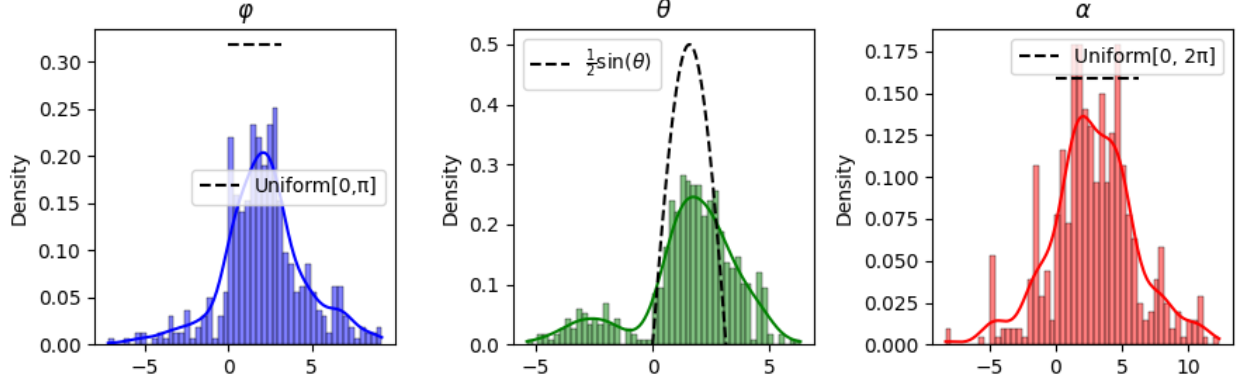Figure 14: Marginal density of the learned parameters $(\theta, \alpha)$

Figure 15: Marginal density of the learned parameters $(\varphi)$



Figure 16: Marginal density of the learned parameters $(\varphi, \alpha)$



Figure 17: Marginal density of the learned parameters $(\varphi, \theta)$

14

Figure 18: Marginal density of the learned parameters $(\varphi, \theta, \alpha)$

We can see that among all the cases, the final marginal distributions seem pretty consistent throughout.

# D   Variable selection with $\psi$

We repeat the previous experiment considering also the phase coefficient $\psi$.

Table 4: Losses obtain for different parameter combination

(a) 250 modes

| Parameters | Loss | RMS (x, y, z) |
|---|---|---|
| $(0, 0, \alpha, \psi)$ | 0.75522 | (0.20 , 0.11 , 0.10) |
| $(\varphi, 0, \alpha, \psi)$ | 0.74388 | (0.09 , 0.22 , 0.04) |
| $(\varphi, 0, 0, \psi)$ | 0.73930 | (0.11 , 0.22 , 0.08) |
| $(0, \theta, 0, \psi)$ | 24.56362 | (0.12 , 0.22 , 0.06) |
| $(0, 0, 0, \psi)$ | 0.74839 | (0.09 , 0.24 , 0.04) |
| $(0, \theta, \alpha, \psi)$ | 0.75160 | (0.21 , 0.12 , 0.10) |
| $(\varphi, \theta, 0, \psi)$ | 0.81957 | (0.09 , 0.21 , 0.08) |
| $(\varphi, \theta, \alpha, \psi)$ | 183.87051 | (0.11 , 0.27 , 0.05) |

(b) 500 modes

| Parameters | Loss | RMS (x, y, z) |
|---|---|---|
| $(0, 0, \alpha, \psi)$ | 0.74805 | (0.11 , 0.25 , 0.07) |
| $(\varphi, 0, \alpha, \psi)$ | 0.74443 | (0.21 , 0.17 , 0.06) |
| $(\varphi, 0, 0, \psi)$ | 0.74356 | (0.09 , 0.27 , 0.06) |
| $(0, \theta, 0, \psi)$ | 0.74534 | (0.20 , 0.11 , 0.10) |
| $(0, 0, 0, \psi)$ | 0.75040 | (0.08 , 0.27 , 0.06) |
| $(0, \theta, \alpha, \psi)$ | 0.74584 | (0.20 , 0.14 , 0.10) |
| $(\varphi, \theta, 0, \psi)$ | 0.73757 | (0.20 , 0.12 , 0.11) |
| $(\varphi, \theta, \alpha, \psi)$ | 3.38473 | (0.23 , 0.11 , 0.09) |
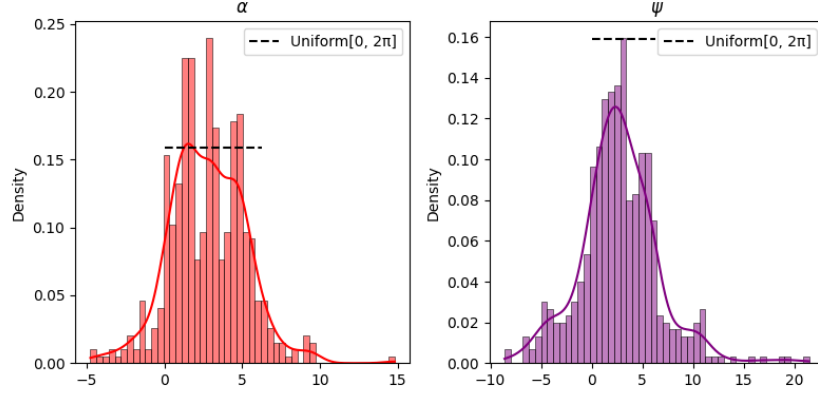
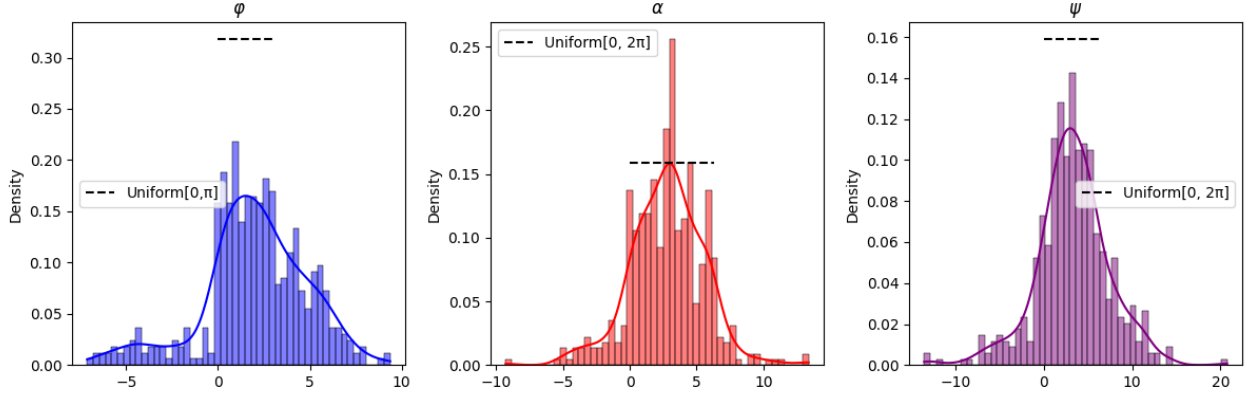Figure 19: Marginal density of the learned parameters $(\alpha, \psi)$



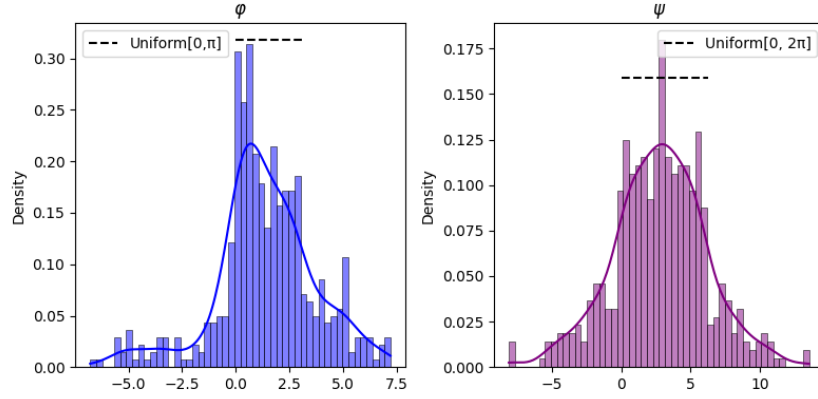Figure 20: Marginal density of the learned parameters $(\varphi, \alpha, \psi)$



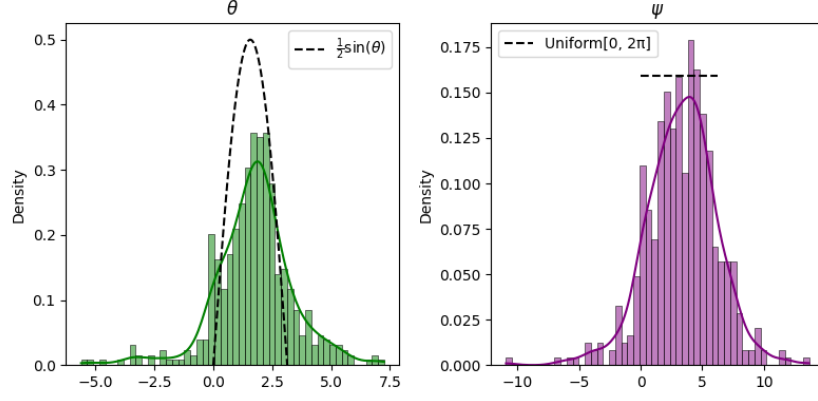Figure 21: Marginal density of the learned parameters $(\varphi, \psi)$

Figure 22: Marginal density of the learned parameters $(\theta, \psi)$



Figure 23: Marginal density of the learned parameters $(\psi)$
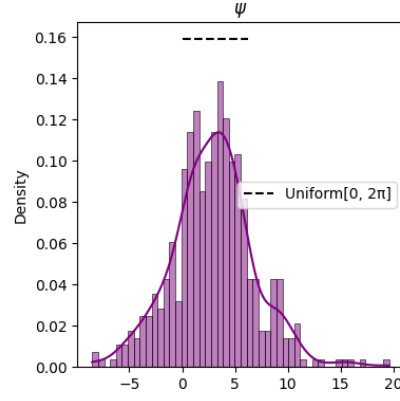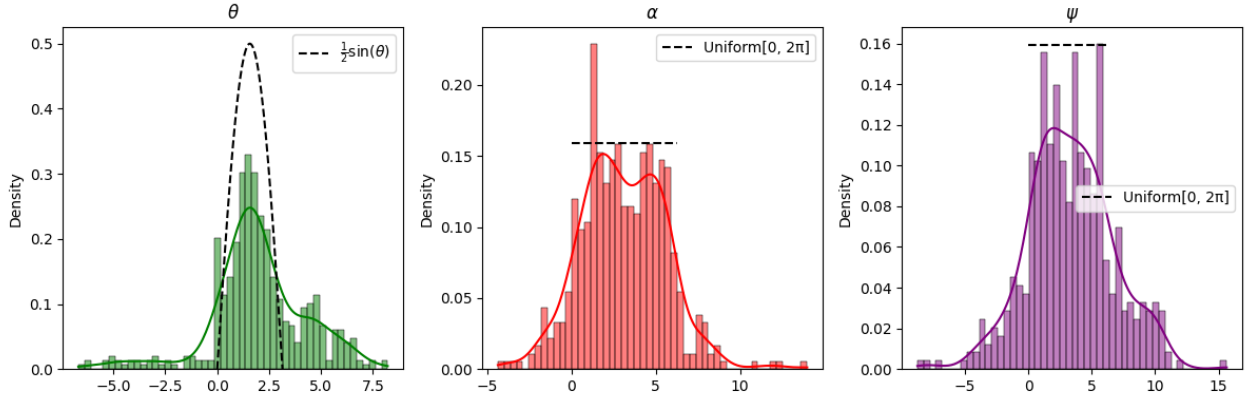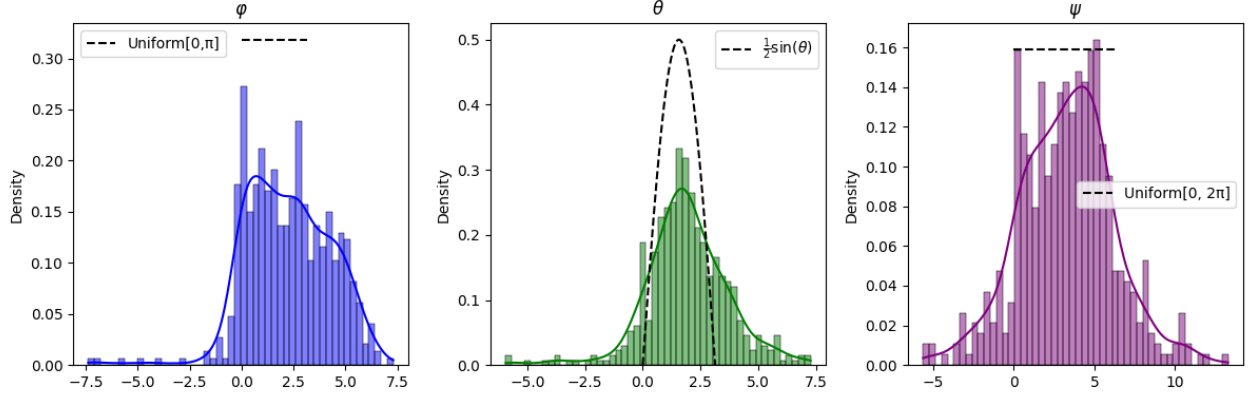


Figure 24: Marginal density of the learned parameters $(\theta, \alpha, \psi)$

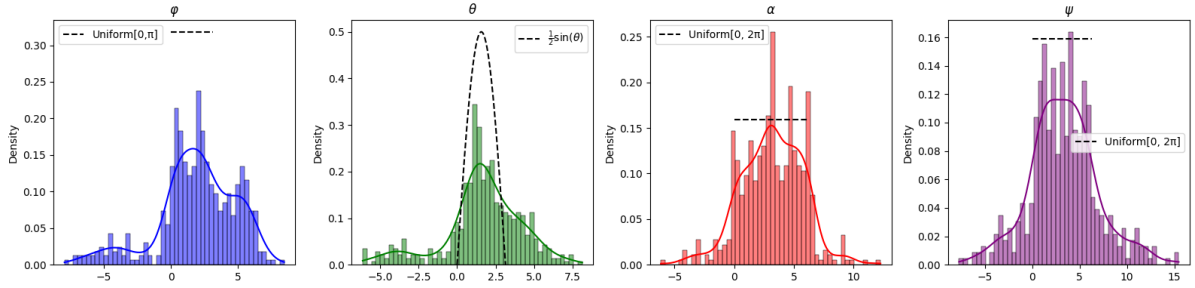Figure 25: Marginal density of the learned parameters $(\varphi, \theta, \psi)$



Figure 26: Marginal density of the learned parameters $(\varphi, \theta, \alpha, \psi)$

Overall, the distribution of the angles appears consistent with the previous case where $\psi$ was not considered. The distribution of $\psi$ itself remains pretty consistent across all cases.
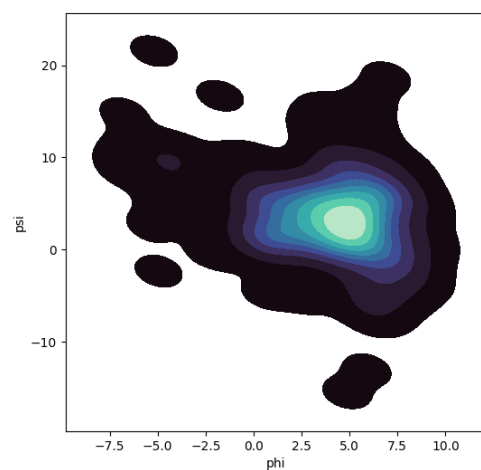
The distribution of $\alpha$ is the most similar to the base distribution $\mathcal{U}([0, 2\pi])$.

However, these figures represent only the marginal densities and do not capture the dependencies between parameters, which could explain the observation made from **??**.
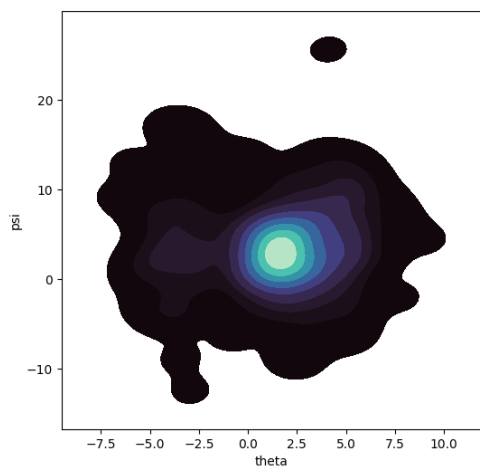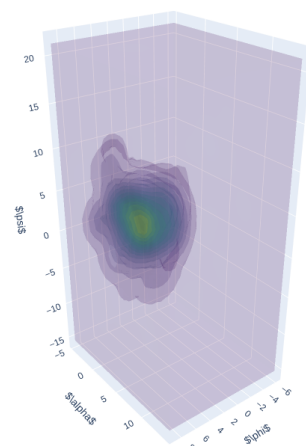
## D.1  Joint distributions
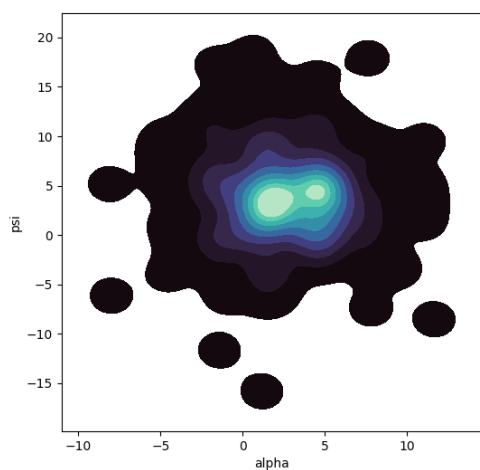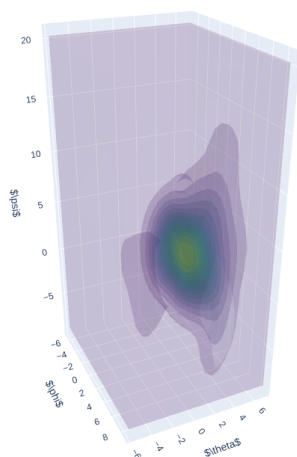
(a) Learned parameters $(\alpha, \psi)$
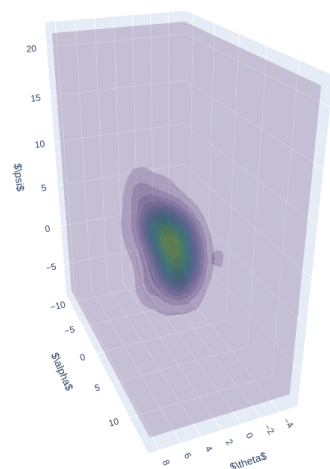


(b) Joint distribution of $(\varphi, \psi)$



(c) Joint distribution of $(\theta, \psi)$



(d) Joint distribution of $(\varphi, \theta, \alpha)$



(e) Joint distribution of $(\theta, \psi)$



(f) Joint distribution of $(\varphi, \theta, \alpha)$

Figure 27: Joint distributions of parameter combination

# References

[1] Deniz Bezgin and Nikolaus Adams. "Normalizing Flows as a Novel PDF Turbulence Model". In: (Jan. 2021). DOI: 10.48550/arXiv.2101.03590.

[2] Toshiyuki Gotoh, Daigen Fukayama, and Tohru Nakano. "Velocity field statistics in homogeneous steady turbulence obtained using a high-resolution direct numerical simulation". In: *Physics of Fluids* 14.3 (Mar. 2002), pp. 1065–1081. ISSN: 1070-6631. DOI: 10.1063/1.1448296. eprint: https://pubs.aip.org/aip/pof/article-pdf/14/3/1065/19111368/1065_1_online.pdf. URL: https://doi.org/10.1063/1.1448296.

[3] Jérémie Janin et al. "A new linear forcing method for isotropic turbulence with controlled integral length scale". In: *Physics of Fluids* 33 (Apr. 2021), p. 045127. DOI: 10.1063/5.0045818.

[4] Meet Hemant Parikh, Xiantao Fan, and Jian-Xun Wang. *Conditional flow matching for generative modeling of near-wall turbulence with quantified uncertainty*. 2025. arXiv: 2504.14485 [physics.flu-dyn]. URL: https://arxiv.org/abs/2504.14485.

[5] Florian Ries. "Numerical Modeling and Prediction of Irreversibilities in Sub- and Supercritical Turbulent Near-Wall Flows". PhD thesis. Nov. 2018.

[6] Haixin Wang et al. *FourierFlow: Frequency-aware Flow Matching for Generative Turbulence Modeling*. 2025. arXiv: 2506.00862 [cs.LG]. URL: https://arxiv.org/abs/2506.00862.

[7] Mustafa Z. Yousif, Linqi Yu, and HeeChang Lim. "Physics-guided deep learning for generating turbulent inflow conditions". In: *Journal of Fluid Mechanics* 936 (2022), A21. DOI: 10.1017/jfm.2022.61.