Politecnico di Torino

Master's thesis

Macroscopic Fluctuation Theory applied to turbulence

Author:
Pietro Dragoni

Supervisors:
Lamberto Rondoni
Kirone Mallick
Bérengère Dubrulle

A report submitted in partial fulfillment of the requirements for the degree in M2 in Physics of Complex Systems

at the

Institut de Physique Théorique (IPhT) et Service de Physique de l'État Condensé (SPEC)
Dipartimento di Scienze Matematiche "Giuseppe Luigi Lagrange" and CEA Paris-Saclay

July 13, 2025

Abstract

One of the theoretical challenges in the study of turbulence is to develop an effective, coarse-grained model of its behaviour on a macroscopic scale. Turbulent flows are out-of-equilibrium systems that exhibit stochastic behaviour due to small-scale perturbations affecting macroscopic dynamics within a finite amount of time. This property, known as the spontaneous stochasticity of the incompressible Navier–Stokes equation, challenges traditional coarse-graining approaches because unresolved scales cannot simply be neglected. In this work, these scales are modelled as vanishing, momentum-conserving, multiplicative noise. The resulting stochastic Navier–Stokes equation is analysed using Macroscopic Fluctuation Theory (MFT), a statistical framework for non-equilibrium systems governed by stochastic hydrodynamics. The results and insights of this theory are discussed in the context of Langevin equations. Strategies are presented for evaluating the probability distribution of observables to leading order in the vanishing-noise limit by integrating the instanton equations. The solutions to these equations represent the most likely paths leading to a given rare event. The main contribution of this work is the derivation of the MFT action and instanton equations for the adopted stochastic Navier–Stokes equation. Some insights are presented for the case of isotropic, locally correlated noise, with a particular focus on evaluating the stationary measure and the emergence of Lagrangian phase transitions. These results demonstrate the potential of MFT to provide a parameter-free statistical description of turbulence; however, further work is required to extract practical predictions from the derived model, both analytically and numerically.

Summary

Author:
Pietro Dragoni (Physics of Complex Systems, international track).
Supervisors:
Lamberto Rondoni (DISMA, Politecnico di Torino),
Kirone Mallick (IPhT CEA-Saclay, Université Paris-Saclay),
Bérengère Dubrulle (SPEC CEA-Saclay, Université Paris-Saclay).

Title:
Macroscopic Fluctuation Theory applied to turbulence.

Aim of the project:
Building a statistical theory for turbulence by addressing it with Macroscopic Fluctuation Theory.

Context:
Turbulent flows are ubiquitous in our everyday life but still remain difficult to predict. The reason is the lack of a deterministic or statistical theory of turbulence and the very high computational cost of direct numerical simulations, which even the best super-computers struggle to deal with due to the number of scales involved. Therefore, physicists are trying to build effective models of the macroscopic behaviour of turbulent flows by coarse-graining the small scales details. The aim is to reduce the number of degrees of freedom to simulate. Any coarse-graining introduces a vanishing noise in the coarse-grained hydrodynamics, which is usually disregarded to have a deterministic description of turbulence, following the common belief that determinism rules over classical physics.

However, the phenomenon of spontaneous stochasticity appearing in Navier-Stokes equation at high Reynolds numbers [1, 6, 18, 34, 35, 38] recently questioned the possibility of explaining the macroscopic behaviour of turbulence with a deterministic theory. Phenomenologically, perturbations at small length-scale (even thermal molecular noise) are able to affect the macroscopic scale in finite time, making the effect of the coarse-grained small details non-negligible. From a theoretical perspective, spontaneous stochasticity consists in the existence of a universal finite time predictability barrier beyond which the separation between the two Lagrangian particles approaches a universal power-law growth, where "universal" means independent of their initial distance. This remains valid even if the initial distance is sent to zero: two particles starting with exactly the same initial condition will separate in finite time, making the behaviour of Navier-Stokes equation spontaneously non-deterministic. The mathematical reason of this phenomenon cannot be found in deterministic chaos, not only because it predicts that two particles separates exponentially fast in time, but also because the separation is directly proportional to the initial one; this is in contrast with the observed universality of the time predictability barrier. In fact, the conjectured explanation of spontaneous stochasticity is the loss of uniqueness of solutions of Navier-Stokes equation and the emergence of singular velocity fields in finite time at high Reynolds numbers. In this perspective, the flow becomes stochastic as a Lagrangian particle have to choose randomly which of the various possible solutions to follow. The social relevance of this phenomenon is in meteorology, where spontaneous stochasticity is referred as "the real butterfly effect" and is reported to spoil weather forecasts, which actually need to be more and more accurate in order for us to adapt to our warming world. The detailed discussion of spontaneous stochasticity can be found in section 1.1.

Looking for a statistical model of the macroscopic behaviour of turbulence, this work considers a stochastic version of Navier-Stokes equation for the velocity profile $u(x,t)$, which originates from a coarse-graining of small scales. In adimensional units this reads:

$$\begin{aligned}
&\partial_t u(x,t) + (u(x,t)\cdot\nabla)u(x,t) = -\nabla P(x,t) + \mathrm{Re}^{-1}\Delta u(x,t) + \nabla\cdot\sqrt{\epsilon\sigma(u(x,t))}\eta(x,t)\\
&\nabla\cdot u(x,t) = 0\\
&u(x,0) = u_0(x)\\
&\eta \text{ Gaussian noise of correlation } \langle\eta^{ai}(x,t)\eta^{bj}(y,s)\rangle = C^{ij}_{ab}(x-y,t-s)
\end{aligned}$$

where $\nabla\cdot\sqrt{\epsilon\sigma(u)}\eta$ is a momentum-conserving multiplicative noise representing the coarse-grained small scales in an effective way. This is a vanishing noise ($\epsilon\to 0$) stochastic hydrodynamics. Therefore it

can be studied by the tools of Macroscopic Fluctuation Theory (MFT, [5, 12, 24, 27]), a theory for out of equilibrium systems that has its roots in Large Deviation theory and path integral formulation of stochastic processes. MFT provides with a recipe to compute the stationary probability distribution and the large deviation function of any observable A to the leading order in the vanishing noise limit in systems described by stochastic hydrodynamics. The key ingredient is the existence of a least action principle whose stationary equations give the most likely path leading to a fixed rare state. This optimal path is called the instanton. Due to the fact that the typical path leading a system to an atypical state coincides with the most likely path leading there, the instanton is actually the large deviation path that is observed experimentally and in numerical simulations, up to small fluctuations around it. Therefore, the instanton equations furnish a model to predict how rare events emerge in stochastic hydrodynamics. An introduction to hydrodynamical descriptions, which are the subject of Macroscopic Fluctuation Theory, is given in section 1.2. The key concepts of MFT are discussed in 2 through its application to Langevin equations, which represent a perfect playground to test MFT as everything can be computed in alternative ways.

Original contributions: Besides re-deriving known results of MFT applied to Langevin equations from a Hamiltonian perspective in section 2.2.2, the original part of this work is contained in section 3. It was shown that Macroscopic Fluctuation Theory can formally be applied to the stochastic version of Navier-Stokes equation mentioned above and, for the first time, the following instanton equations were derived:

$$
\begin{aligned}
&\partial_t u^i(x,t) + \mathcal{P}^i[u(x,t)\cdot\nabla\vec{u}(x,t)] - \mathrm{Re}^{-1}\Delta u^i(x,t) = \\
&\qquad -\sum_l \mathcal{P}^{il}\Big\{\sum_{a,b}\frac{\partial}{\partial x^a}\Big[\sqrt{\sigma(u(x,t))}\int_{j,y,s} C^{lj}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^j(y,s)}{\partial y^b}\Big]\Big\} \\
&\partial_t\pi^i(x,t) + \mathcal{P}^i\Big[u(x,t)\cdot\nabla\vec{\pi}(x,t) + \sum_j u^j(x,t)\vec{\nabla}\pi^j(x,t)\Big] + \mathrm{Re}^{-1}\Delta\pi^i(x,t) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,t)}\Big] + \\
&\qquad -\mathcal{P}^i\left[\frac{(\vec{\nabla}_u\sigma(u(x,t)))}{2\sqrt{\sigma(u(x,t))}}\sum_{j,k,a,b}\int_{s,y}\frac{\partial\pi^j(x,t)}{\partial x^a}C^{jk}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^k(y,s)}{\partial y^b}\right] \\
&\quad t=0: \pi^i(x,0) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,0)}\Big] + \mathcal{P}^i\Big[\epsilon\frac{\delta F[u(0)]}{\delta u(x,0)}\Big] \\
&\quad t=t_f: \pi^i(x,t_f) = \lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,t_f)}\Big]
\end{aligned}
$$

where $\mathcal{P} = \mathbb{1} - \nabla\Delta^{-1}\nabla^\top$ is the Leray projector. π is the Hamiltonian conjugated momentum to the velocity field and can be interpreted as the "stochastic force" that the noise has to give to the system to reach the atypical state following the instanton path. These equation furnish an effective model for turbulence, but in this very general formulation it is hard to understand. Therefore, they were specialized to the simpler physical case of a delta correlated isotropic noise:

$$
\langle\eta^{aj}(x,t)\eta^{bk}(y,s)\rangle = \delta(x-y)\delta(t-s)\Big[\delta^{ab}\delta^{jk} + \delta^{ak}\delta^{bj} - \frac{2}{3}\delta^{aj}\delta^{bk}\Big]
$$

The instanton equations result more interpretable and read:

$$
\begin{aligned}
&\partial_t u^i(x,t) + \mathcal{P}^i[u(x,t)\cdot\nabla\vec{u}(x,t)] - \mathrm{Re}^{-1}\Delta u^i(x,t) = -\mathcal{P}^i\Big\{\nabla\cdot\Big[\sigma(u(x,t))S_\pi(x,t)\Big]\Big\} \\
&\partial_t\pi^i(x,t) + \mathcal{P}^i\Big[S_\pi(x,t)u(x,t)\Big] + \frac{\Delta\pi^i(x,t)}{\mathrm{Re}} = -\lambda\mathcal{P}^i\Big(\frac{\delta A[u]}{\delta u(x,t)}\Big) - \mathcal{P}^i\Big\{\frac{1}{4}\big(\vec{\nabla}_u\sigma(u(x,t))\big)\mathrm{tr}\big(S_\pi(x,t)^2\big)\Big\} \\
&\quad t=0: \quad \pi^i(x,0) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,0)}\Big] + \mathcal{P}^i\Big[\epsilon\frac{\delta F[u(0)]}{\delta u(x,0)}\Big] \\
&\quad t=t_f: \quad \pi^i(x,t_f) = \lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,t_f)}\Big]
\end{aligned}
$$

where $S_\pi(x,t) = \nabla\otimes\pi(x,t) + (\nabla\otimes\pi(x,t))^\top$ is the strain tensor for the momentum π. Notice that the first equation is Navier-Stokes equation perturbed with an effective contribution of the noise being

$-\mathcal{P}^i\Big\{\nabla\cdot\Big[\sigma(u(x,t))S_\pi(x,t)\Big]$: this is the shape the noisy term should take in order to make the system follow the instanton path.

These equations are very difficult to study both analytically and numerically because they are as non linear as the Navier-Stokes equation and it is a shooting problem. In principle, integrating them allows to access the large deviation function ϕ_A of any observable A such that $\Pr(A=a)\propto\exp[-\epsilon^{-1}\phi_A(a)]$, to leading order in the vanishing noise. Theoretically, it suffices to apply a Legendre transform to the following cumulant generating function:

$$w_A(\lambda):=\epsilon\ln\langle e^{\frac{\lambda}{\epsilon}A}\rangle=-\epsilon F[u_I(0)]+\lambda A[u_I]-\frac{1}{2}\sum_{a,b}\langle\partial_a\pi_I|\sqrt{\sigma(u_I)}C_{ab}\sqrt{\sigma(u_I)}|\partial_b\pi_I\rangle$$

Numerically, the Legendre transform can be skipped invoking ensemble equivalence as explained in section 2.2.3. With the choice $A=u(t_f)$, the same procedure gives the pseudo-potential ϕ, namely the exponent of the stationary distribution of the process $\mathrm{P}_{\mathrm{ss}}(u)\propto\exp[-\epsilon^{-1}\phi(u)]$.

There was no time to approach the complicated task of integrating these equations seriously in this work, but the groundwork was laid by integrating the instanton equations for a one dimensional Langevin equation in section 2.3, which already shows some nice features. Therefore, this work is far from answering the question whether Macroscopic Fluctuation Theory can produce an effective statistical model of turbulence and, indeed, many paths remain to be explored.

Once predictions of this theory will be available, they can be checked by experiments where rare events are observed [17] in turbulent flows. Would this furnish an approximated but effective model of turbulence, it could be used for weather and climate forecast in place of the empirical and parametric model used nowadays, which could have a positive impact in building resilience against climate change.

Contents

1 Introduction

1.1 Stochasticity in turbulence

The lack of a deterministic or statistical theory of turbulence makes turbulent flows difficult to predict. This section will explain why turbulence cannot be modelled deterministically and why statistical physics approaches are applicable. The justifications will come from a mathematical conjecture (section 1.1.1) and some numerical evidences (section 1.1.2) of a spontaneous stochasticity affecting Navier-Stokes equation (NSE). The social relevance of studying this mathematical property will be discussed in section 1.1.3. Finally, a stochastic version of the NSE that is compatible with these observation is introduced in section 1.1.4.

1.1.1 A mathematical conjecture on the stochasticity of turbulence

Physicists in fluid mechanics are generally assume the existence or uniqueness of solutions to the NSE at all times. However, this has not been proven yet and actually the following scaling argument raises suspicions that this mathematical technicality may have a physical effect. Define the local variation of the velocity field u as $\delta u(x;\ell) := u(x+\ell) - u(x)$ where ℓ is a vector of small norm. [38] finds that in turbulent flows simulations $\delta u \sim t^{1/2}$ for small length-scales $|\ell|$. Considering a particle $x(t)$ advected by this velocity field ($\dot{x}(t) = u(x(t), t)$), the relation $\delta u \sim t^{1/2}$ can be integrated and gives $\delta x \sim t^{3/2}$ for the variation of the position[1]. As a consequence, $\delta u = \frac{d}{dt}\delta x \sim \delta x^{1/3}$. This ordinary differential equation (ODE) does not satisfy the Cauchy-Lipschitz theorem of local existence and uniqueness of solutions and, indeed, $\forall t_J \geq 0$ the function $\delta x(t) = [\frac{2}{3}(t-t_J)]^{3/2}\mathbb{1}\{t > t_J\}$ is a solution[2] for the Cauchy problem of NSE with initial condition $\delta x(0) = 0$. This means that the fluid particle will randomly select the particular trajectory, making its position a stochastic process amenable for statistical physics approaches [16].

Of course, this very simple ODE cannot represent the full complexity of turbulence, but the small calculation above suggests that it is not so unreasonable to assume that the Navier-Stokes equation in the turbulent regime has multiple solution for some initial condition, thus behaving effectively as a stochastic equation. Just like $\delta x(0) = 0$ for the ODE above, the initial velocity profile entailing this loss of determinism in NSE should be sufficiently singular (or non-smooth, meaning that some spatial derivatives should diverge [34]). Mathematically, a velocity profile $u(x)$ is singular if $\exists x_s$ such that the local Hölder exponent[3] h is less than 1 as it implies local divergences of the spatial derivatives of u. Hölder exponents in fully developed turbulence can even be lower than 1/3, as observed in experiments [17], and such irregular profiles might entail multiple solutions for the initial value problem of Navier-Stokes equation.

Proving that the NSE admits non-differentiable or non-unique solutions is one of the unresolved millennial problems. Instead, the 3-dimensional Euler equation, NSE's infinite Reynolds number limit, has been proven numerically to have finite "blow-up time", namely solutions starting from a smooth initial fields develop a singularity (infinite vorticity) in a finite time [9]. As a fully developed turbulent flow has $\mathrm{Re} > 5000$ and even reaches $\mathrm{Re} \sim 10^8$ in the atmosphere, singularities might play a role in turbulence. Indeed, their influence in Navier-Stokes equation is also conjectured to explain dissipation anomaly at low viscosities [19].

The presence of singularities and non-uniqueness of solutions in Navier-Stokes equation is a conjecture that won't be needed in the rest of the report; nevertheless, this is the picture to keep in mind to understand the mathematical reason of the stochasticity emerging in turbulent flows. In the next section stronger evidences of the presence of stochasticity in Navier-Stokes equation at high Reynolds number will be presented.

[1] This law is known in turbulence as the Richardson law.

[2] The indicator function of a clause $\mathbb{1}(\text{clause}) = 1$ if the clause is satisfied and zero otherwise.

[3] The Hölder exponent is a measure of the irregularity of a field and it is defined as the maximum real number $h \leq 1$ such that for l of sufficiently small norm $\exists C \in \mathbb{R} : ||u(x+l) - u(x)|| < C||l||^h$. If $h \in (0, 1]$ the function is called Hölder continuos in x.

1.1.2 Spontaneous stochasticity in Navier-Stokes equation

From a Lagrangian perspective, consider two particles of positions $\dot{x}_i(t) = u(x_i(t), t)$, define their distance vector $\Delta x(t) = x_1(t) - x_2(t)$, assume small initial separation $\Delta x(0)$ and evolve the velocity field integrating the incompressible Navier–Stokes at high Reynolds numbers. It turns out [6] that there exists a finite time predictability barrier t_{PB} beyond which $|\Delta x(t)|^2 \sim 0.52 \cdot E_D t^3$, where E_D is the turbulence kinetic energy dissipation rate. Both this power law and t_{PB} are independent of the initial distance $\Delta x(0)$. As a consequence, even in the vanishing initial distance limit (so starting with exactly the same initial condition), the particles follow different trajectories after t_{PB}, as if the evolution was stochastic. This is the Lagrangian definition of "spontaneous stochasticity".

From an Eulerian perspective, consider the solutions $u(x,t)$ and $\tilde{u}(x,t)$ of the Cauchy problem for the incompressible Navier-Stokes equation corresponding to two slightly different initial conditions $u(x,0) = u_0(x)$ and $\tilde{u}(x,0) = \tilde{u}_0(x)$. By employing numerical simulations of Navier-Stokes equation in the shear layer geometry, it was shown [38] that it exists a time t_{PB}, independent of the initial distance, such that $\forall t > t_{PB}$ $||u(t) - \tilde{u}(t)||^2_{L_2} \sim t$ universally in the initial difference[4]. Again, one could send the initial error to zero and still get different profiles in finite time as Navier-Stokes equation was stochastic.

It is important to underline that this spontaneous stochasticity cannot originate from deterministic chaos. The distance between two chaotic particles at time t follows $|\Delta x(t)| = |\Delta x(0)| \exp[\lambda t]$, so the separation time can be lengthened by decreasing the initial distance and two chaotic particles starting at the same initial condition will follow exactly the same path. The mathematical explanation of spontaneous stochasticity in turbulent flows may be the emergence of locally irregular velocity fields and the consequent loss of uniqueness of solutions, as explained in section 1.1.1. Within this interpretation, the time predictability barrier t_{PB} coincides with the blow-up time of the 3D Euler equation. Notice also that spontaneous stochasticity is a phenomenon appearing in other systems than the Navier-Stokes equation, where it is usually due to the loss of uniqueness of solutions to the Cauchy problem of the differential equation considered [16].

From a physical point of view, spontaneous stochasticity amplifies the small scale perturbations[5] in turbulent flows up to the macro-scale in finite time. It was shown experimentally [17] that this amplification happens in very irregular regions of the turbulent flow, where the local Hölder exponent is less than 1/3. This gives further credit to the connection between spontaneous stochasticity and singularities. It was even proven that thermal noise could propagate to the macro-scale for sufficiently high Reynolds numbers in the SABRA model for turbulence [1], rendering the probability distributions of macroscopic observables very far from delta functions even in the vanishing noise limit. Notice that obtaining experimental or numerical evidence of spontaneous stochasticity in turbulence is very challenging as it requires respectively high resolutions on particle tracking and high Reynolds numbers[6].

1.1.3 The real butterfly effect in weather

This section will discuss the concrete relevance of the study of spontaneous stochasticity in turbulence. It turns out that it is one of the reasons why predicting the weather is difficult.

The concept of the "butterfly effect" is commonly associated with sensitive dependence on initial conditions in chaotic systems. However, [34] argues that Lorenz originally intended the existence of a fundamental finite-time predictability limit that arises in multi-scale turbulent systems, beyond which reducing initial uncertainty does not improve forecasts. This concept is termed the "real butterfly effect". Phenomenologically, this effect is due to the fact that for high Reynolds numbers a perturbation at a small scale l is able to propagate in finite time up to the observation scale[7] $x_0 \gg l$, thus influenc-

[4] The L_p norm of a function $f(x)$ with $x \in D \subset \mathbb{R}^n$ is defined as $\|f\|_{L_p} = \left(\int_D d^n x\, |f(x)|^p\right)^{1/p}$ for $p \geq 1$.

[5] The origin of the small-scale perturbations can be the roughness of surfaces, the boundary layer separation, but even the thermal noise [1] or, in weather and climate perspective, a small storm.

[6] The current maximum Reynolds number achieved in DNS of Navier-Stokes equation is between 5000 and 10^5 depending on the particular geometry considered, while often spontaneous stochasticity needs higher Re to be observed numerically.

[7] The scale l is typically greater than the Kolmogorov length η_K, below which the viscous dissipation is the dominant

ing the macroscopic behaviour and making it effectively unpredictable [34]. In weather forecast, for instance, this amplification of small scale perturbations is believed [34] to be the main reason of many poor quality 6-days forecasts made by the European Centre for Medium-Range Weather Forecasts (ECMWF).

Considering what was discussed in 1.1.2, the real butterfly effect in weather seems the analogue of spontaneous stochasticity in turbulent flows. This makes sense not only because the atmosphere is turbulent, but also as the primitive equations[8] governing the weather are believed [18] to be influenced by singularities present in the inviscid limit, similarly to what was conjectured for NSE in section 1.1.1. If spontaneous stochasticity was the explanation of weather unpredictability, the time predictability barrier would be around 5-10 days, namely the time after which weather forecasts are notoriously unreliable (as also evidenced by the poor-quality 6-days forecasts by ECMWF). This connection to weather makes the study of spontaneous stochasticity not only scientifically interesting but also socially relevant.

1.1.4 A coarse-grained version of the Navier-Stokes equation

Since turbulent flows are multi-scale and deterministically unpredictable, the aim is to coarse-grain the small chaotic scales to get a statistical effective model. Coarse-graining is also a way to treat the presence of singularities in a differential equation.

Suppose that the "resolution" at which the system is observed is x_0 (for weather is $x_0 \sim 10$ Km and for classical turbulence experiments $x_0 \sim 1$ cm) and smaller scales $\ell < x_0$ have to be coarse-grained. The length scales below the Kolmogorov length $\ell < \eta_K$ can be safely coarse-grained[9], as the dissipation prevents fluctuations to affect the macro-scale. The intermediate scales $\eta_K < \ell < x_0$, instead, are sufficiently large to admit irregular fields and cause spontaneous stochasticity, so they cannot be simply disregarded. This section will define and justify a stochastic model for their effect, leading to a stochastic coarse-grained Navier-Stokes equation.

Define the coarse-grained velocity at scale l as $\overline{u}^l(x,t) := (G_l * u)(x,t) := \int_{\mathbb{R}^d} u(x',t)G_l(x-x')\,dx'$ where $\eta_K < l \ll x_0$ and $G_l(x) = l^{-d}g(x/l)$ is a low-pass filter that suppresses variations on scales smaller than l (e.g. any filter that is decaying sufficiently fast in the Fourier space as the wave-vector becomes large). The filter cuts the small scales and smoothens singularities, if present. So use that $u \cdot \nabla u = \nabla \cdot (u \otimes u)$ by incompressibility[10], apply the filter, add and subtract $\overline{u}^l \cdot \nabla \overline{u}^l$, then the coarse grained NSE reads:

$$\partial_t \overline{u}^l + \overline{u}^l \cdot \nabla \overline{u}^l = -\nabla \overline{p}^l + \nu \nabla^2 \overline{u}^l + \nabla \cdot \tau_l \tag{1}$$

where $\overline{p}^l = G_l * p$ is the filtered pressure and $\tau_l = \overline{u}^l \otimes \overline{u}^l - \overline{u \otimes u}^l$ is the sub-grid stress tensor (or turbulent stress tensor) that encapsulates the effects of unresolved small-scales on the resolved scales. This is not negligible because of spontaneous stochasticity and many models of turbulence are built by postulating a physically reasonable dependence of this tensor only on $\overline{u}^l$, so that equation 1 is closed (e.g. in [3] a closure with the turbulent kinetic energy is used). The main assumption in this report is that this sub-grid stress tensor can be interpreted as an effective noise. Referring to [20] for a more rigorous discussion, the reasons are the following ones. First, spontaneous stochasticity and chaos are acting at the coarse-grained scales. A more mathematical justification is that spontaneous stochasticity can be probed by adding a noise to the deterministic equation and checking if in the vanishing noise limit (and $\mathrm{Re} \to \infty$) the system shows non-delta correlated distributions [1, 35, 38]. This is called "stochastic regularization" in the literature [16]. Finally, from a statistical physics perspective, any finite coarse-graining implies a vanishing noise, coherently with the central limit theorem, as it will be discussed in section 1.2.2.

---

term in the Navier-Stokes equation and turbulence cannot develop; the macro-scale x_0 ∼10 Km in weather models, $x_0 \sim$ 1m in turbulence experiments.

[8] Primitive equations are the differential equations that model the weather. They are essentially derived from the Navier-Stokes equation using that the atmosphere has a vertical thickness very smaller than the horizontal extension.

[9] Or, more precisely, the scales below the regularization length $\eta_h \sim \nu^{\frac{1}{1+h}}$ where h is the maximum local Hölder exponent (see section 3.1.2).

[10] Define the outer product $(u \otimes \psi)^{ij} = u^i \psi^j = \partial_a \psi^i$.

This assumption leads to a stochastic coarse-grained Navier-Stokes equation that reads:

$$\begin{cases} \partial_t \overline{u}^l(x,t) + \overline{u}^l(x,t)\cdot\nabla\overline{u}^l(x,t) = -\nabla\overline{p}^l(x,t) + \nu\Delta\overline{u}^l(x,t) + \nabla\cdot\xi_l(x,t,\overline{u}^l) \\ \nabla\cdot\overline{u}^l(x,t) = 0 \quad \text{Incompressibility} \\ \overline{u}^l(x,0) = u_{\text{in}}(x) \quad \text{Initial condition} \end{cases} \tag{2}$$

The hypothesis on the noise are pretty general: the amplitude of the noise should vanish in $l \to 0$, the noise is multiplicative as the sub-grid tensor depends on the velocity field and it is assumed Gaussian (this is a consequence of assuming a generalization of the central limit theorem to hold and it will be clearer in section 1.2.2). The full list of assumptions is discussed in section 3.1.1.

Notice that the noise enters in a momentum-conserving way in the coarse-grained equations 2 (see the divergence which is applied to it). This is physical as the noise is an effective representation of the coarse grained scales that conserve the momentum as they are governed Navier-Stokes equation. However, this is not widely spread in the literature where often the aim is just to probe spontaneous stochasticity and so a vanishing additive noise is sufficient [35]. Thanks to this momentum-conserving noise, equation 2 can be rewritten as a continuity equation, which suggests to address it by the means of Macroscopic Fluctuation Theory. This is a theory for out of equilibrium systems that will be introduced in the next section.

Equation 2 is also known as the Landau-Lifshitz Navier-Stokes equation, derived for the first time in [30]. In that case, the noise is of thermal origin and its statistics are recovered by imposing the Einstein relation between the amplitude of the noise and the temperature in order to ensure that the noise is randomly giving back the energy dissipated by the viscosity. One might also consider the Landau-Lifshitz Navier-Stokes equation to investigate the effect of thermal noise in Navier-Stokes equation. This is done in [1] and the authors conclude that even this tiny noise is sufficient to trigger spontaneous stochasticity[11] as long as the amplitude of the noise is sent to zero not faster than $\text{Re}^{-15/4}$. If even thermal noise is able to make a turbulent flow stochastic at the macroscale, it is clear that the coarse-graining noise, which includes also the thermal scales, cannot be disregarded in the equations.

For completeness, notice that some renormalization group techniques can be also applied to perform this coarse-graining and to furnish a statistical model of turbulence, where the filtering is made at the level of the statistical action that one obtains by path integral formulation [20]. However, the next section 1.2 will discuss why Macroscopic Fluctuation Theory appears as promising as these more sophisticated techniques for turbulence.

1.2 Turbulence is a promising application of Macroscopic Fluctuation Theory

Turbulent flows are inherently out of equilibrium because energy is continuously injected at large spatial scales in order to sustain turbulence which otherwise would fade out due to viscosity. Moreover, models for turbulence need to be stochastic because of spontaneous stochasticity, as discussed in section 1.1. This calls for an out-of-equilibrium statistical mechanics approach to turbulence. A fully established theory of out of equilibrium systems is not available yet, but a widely used idea is to generalize the concept of free energy density f with the one of large deviation function.

Large deviation theory [39] was developed to estimate the probability of a rare event (i.e. a large deviation) in a thermodynamically large system, whose typical behaviour is explained by the central limit theorem (or one of its generalizations). As a simple example, consider a sample $\{X_i\}_{i=1,\dots,N\to\infty}$ of independent random variables X_i identically drawn from q and suppose the aim is to compute the probability of observing $A[\vec{X}] := \frac{1}{N}\sum_{i=1}^{N} a(X_i) = A_0$ very larger than the average value of $a(X)$. It is quite easy to prove that for sufficiently thin tailed distributions q it exists a large deviation function ϕ such that $\Pr(A = A_0) \overset{N\to\infty}{\approx} \exp[-N\phi(A_0)]$ and it turns out that ϕ is the Legendre transform of the cumulant generating function of $a(X)$ (consult [33, 39] for an introduction). As you can see, the shape of the probability of a large deviation is really similar to the Gibbs-Boltzmann distribution, where N corresponds to the inverse temperature and ϕ to the free energy. This result paved the way

[11] In the sense that in the vanishing noise limit some observables such as the largest mode of the velocity has a non-delta distribution.

to the large deviation interpretation of equilibrium statistical mechanics. Large deviation theory can be generalized to Markov processes (including Langevin equation) [13] and fluctuating hydrodynamics [14]. It was also proven [25] that a large deviation function exists for the Landau-Lifshitz Navier-Stokes equation, which is the additive noise version of equation 2. Thus, it is one of the most promising theory to provide a comprehensive description of out of equilibrium systems, such as turbulent flows.

Among the large deviation approaches to out of equilibrium systems, one of the most established is Macroscopic Fluctuation Theory (MFT)[5, 12, 24, 27]. This approach has been successfully applied to some prototypical out-of-equilibrium systems, such as the simple exclusion process and the Kipnis-Marchioro-Presutti model. MFT is a theoretical framework originally used to describe at a macroscopic scale the fluctuations of the density and the current in diffusive transport processes (where the density is conserved), but it is believed to be extensible to a much wider range of systems governed by any stochastic hydrodynamics with vanishing amplitude noise, such as equation 2. The next section will briefly discuss the core ideas of Macroscopic Fluctuation Theory, referring to turbulence as a practical example.

1.2.1 Building a hydrodynamical description from a microscopic model

The first step to introduce Macroscopic Fluctuation Theory is to understand how an hydrodynamics can be derived for a large system like turbulent flows. Suppose that, at the microscopic scale, the system is made of N_p particles with microscopic degrees of freedom $\{x_a, p_a\}_{a=1\dots N_p}$. Let l_{micro} and t_{micro} the typical length scales (e.g. the mean free path and the molecular collision time, for air $l_{\text{micro}} \sim 70\,\text{nm}$ and $t_{\text{micro}} \sim 10^{-10}\,\text{s}$). Assume that there is one or more conserved quantities in the system whose density will be indicated with ρ (e.g. density of particles, velocity, energy density). When the observation resolutions in time t_{macro} and space l_{macro} are very larger than the microscopic ones, the aim is to get an effective macroscopic description of the system through a coarse-graining. The procedure is described below and schematically represented in figure 1.

The first step is to introduce a lattice: divide the system into $N_{\text{cells}} \to \infty$ mesoscopic cells of size l_{meso} such that $l_{\text{macro}} \gg l_{\text{meso}} \gg l_{\text{micro}}$; in turbulence, l_{meso} can be the minimal length scale at which the continuum limit holds, which is of order $1\,\mu\text{m}$ for air. Index each cell with $\boldsymbol{i} \in \mathbb{Z}^d$, where d is the number of space dimensions. Starting from the microstate, define the mesoscopic state of the system at (mesoscopic) time s as the collection of the configurations $C_s^{\boldsymbol{i}}$ of each cell $\boldsymbol{i}$. The configuration of a cell is the "law of large numbers" of some interesting microscopic observables inside the cell: for example, the velocity associated to a cell is the mean velocity of the particles inside it. As the number of particles is limited, the mesoscopic state of each cell is a stochastic process, according to the central limit theorem (in the best case scenario).

Each cell $C_s^{\boldsymbol{i}}$ is assumed to admit an equilibrium state if isolated, with relaxation time t_{meso}. Under a separation of timescales approximation $t_{\text{macro}} \gg t_{\text{meso}} \gg t_{\text{micro}}$, a "local equilibrium assumption"[12] is made: each cell relaxes at equilibrium instantaneously compared to the macroscopic time. Notice that the macro-state is kept far from macroscopic equilibrium thanks to external fields or boundary conditions.

The second step is to obtain the hydrodynamics is to perform a continuum limit. Assume that the mesoscopic dynamics $\{C_s^{\boldsymbol{i}}\}^{\boldsymbol{i}\in\mathbb{Z}^d} := \vec{C}_s$ is Markovian (this can be relaxed, see [5]) and admits a stationary probability measure $\text{P}_{\text{ss}}(\vec{C}_s)$ (which is not the product of the equilibrium distribution that each $C_s^{\boldsymbol{i}}$ would have if isolated because the cells exchange particles between each other). Define the mesoscopic version of the conserved quantity density as $\rho_{\boldsymbol{i},s} = \rho(C_s^{\boldsymbol{i}})$: this is the observable that will completely define the system at the macro-scale, neglecting the other details. Let us proceed with the continuum limit by the rescaling $\boldsymbol{x} := \boldsymbol{i}/N_{\text{cells}}\,, t = s/N_{\text{cells}}^2\,, \rho(\boldsymbol{x}, t) = \rho_{\boldsymbol{i}=N_{\text{cells}}\boldsymbol{x}, s=N_{\text{cells}}^2 t}$ and assume that the system is diffusive. This means that the equation ruling the evolution of the density under

[12] It is important to mention that the local equilibrium assumption has two consequences. First, the generalized Einstein relation $\frac{\partial^2}{\partial\rho^2} f(\rho(x,t)) = 2D(\rho(x,t))/\sigma(\rho(x,t))$ can be established for the free energy density inside each cell $f(\rho)$ [28]. Secondly, if the space domain is finite $x \in \Lambda$ and the system is in contact with a bath of chemical potential $\mu(x,t)$, due to local equilibrium the cells near the boundaries will be constantly at equilibrium with the bath and so $\frac{\partial}{\partial\rho} f(\rho(x,t)) = \mu(x,t)\ \forall x \in \partial\Lambda$. This second consequence does not allow any large deviation at the boundaries.

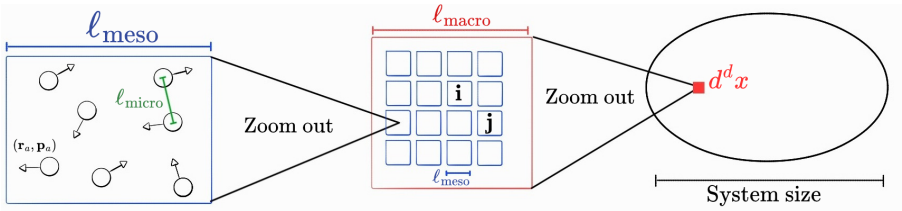

Figure 1: The image displays a schematic representation of the coarse-graining procedure that leads from the microscopic model to its hydrodynamics. The scheme is specialized to the example of a fluid. From left to right: a mesoscopic cell at the microscopic (particle) scale, a macroscopic small volume ($d^d x$) at the mesoscopic scale, and the same at the macroscopic scale.

an external field $E(x,t)$ in the hydrodynamic limit ($N_{\text{cells}} \to \infty$) is the following continuity equation:

$$\partial_t \rho(x,t) = \nabla \cdot J(\rho(x,t),x,t) = \nabla \cdot [-D(\rho(x,t))\nabla\rho(x,t) + \sigma(\rho(x,t))E(x,t)] \tag{3}$$

which is fully described by the diffusivity $D(\rho(x,t))$ and the mobility $\sigma(\rho(x,t))$. In words, the fluctuations of the configuration of the mesoscopic cells self-average at the macro-scale, leading to a deterministic hydrodynamic in the limit $N_{\text{cells}} \to \infty$ or, more precisely, $\epsilon := \left(\frac{l_{\text{meso}}}{l_{\text{macro}}}\right)^d \to 0$. This small parameter represents the "quality" of the coarse-graining as it is proportional to the number of cells in a macroscopic small volume $d^d x$ of size l^d_{macro}.

1.2.2 Any finite coarse-graining comes with a noise

For any finite ϵ the coarse-graining is not perfect and fluctuations around the deterministic hydrodynamics 3 are possible. Assuming that a generalization of the central limit theorem holds $\forall x,t$, these fluctuations are supposed to add a Gaussian noise $\xi(x,t)$ of small amplitude ϵ to the current $J(x,t,\rho(x,t))$. The noise is assumed to be acting on the current and not on the density as the stochastic hydrodynamical description has to respect the conservation laws. In the case of finite ϵ it is possible to introduce the finite coarse-graining density:

$$\rho_\epsilon(x,t) := \epsilon^{-1} \sum_{\boldsymbol{i} \in \text{Vol}(xN_{\text{cells}}\ ,\ l_{\text{macro}})} \rho_{\boldsymbol{i}, s=N^2_{\text{cells}}t} \qquad \text{where Vol}(\boldsymbol{i}\ ,\ l) := \text{Volume around site } \boldsymbol{i} \text{ of size } l$$

and analogously define the i^{th} component of the finite coarse-graining current as the amount of density flowing in the direction of the canonical vector $\boldsymbol{e}_i$ per unit of time and surface. For the central limit theorem assumed above ρ_ϵ will follow a stochastic version of equation 3:

$$\begin{cases} \partial_t \rho_\epsilon(x,t) = \nabla \cdot [J(x,t,\rho_\epsilon(x,t)) + \xi(\rho_\epsilon(x,t),x,t)] \\ \langle \xi^i(\rho_\epsilon(x,t),x,t)\xi^j(\rho_\epsilon(x',t'),x',t')\rangle = \epsilon\sigma(\rho_\epsilon(x,t))C^{ij}\delta(x-x')\delta(t-t') := \chi^{ij}(x-x',t-t',\rho_\epsilon) \end{cases} \tag{4}$$

Notice that the coarse-graining described throughout this section corresponds to the convolution with the filter G_l that was considered in section 1.1.4, which further supports the application of Macroscopic Fluctuation Theory to equation 2. In general, given the hypothesis introduced above, MFT allows to compute the probability distribution (or the cumulants) of any observable $A[\rho]$ (that a priori could depend on the full time evolution of the density) to leading order in ϵ. This corresponds to compute the large deviation function $\phi[\rho]$ of ρ_ϵ at (macroscopic) stationarity, assuming $\mathrm{P}_{\text{ss}}[\rho_\epsilon = \rho] \approx e^{-\epsilon^{-1}\phi[\rho]}$. The procedure is presented for Langevin equation in the section 2 and for the stochastic Navier-Stokes equation 2 in section 3. It can be generalized to any stochastic process with a vanishing noise, which are very common: ϵ is the temperature of the heat bath for Langevin equation [8], in hydrodynamical description it is correspond to the coarse-graining parameter [5], and it is the amplitude of the auxiliary noise introduced in the stochastic regularization approach to systems showing spontaneous stochasticity (section 1.1.4).

2 Methods: discussing Macroscopic Fluctuation Theory by its application to Langevin equations

The aim of this section is to discuss the application of Macroscopic Fluctuation Theory to the more intuitive case of Langevin equations. It will explore the main results of MFT obtained for hydrodynamics [5], some technicalities and give its physical interpretation. Inevitably the results have already been obtained in the literature [5, 8, 24].

Consider the following Langevin equation with delta correlated Gaussian white noise:

$$\begin{cases} \dot{\boldsymbol{r}}(t) \overset{\alpha}{=} \boldsymbol{f}(\boldsymbol{r}(t)) + \sqrt{2\epsilon} g(\boldsymbol{r}(t))\boldsymbol{\eta}(t) & \boldsymbol{r}(0) = \boldsymbol{r}_0 \quad t \in [0, t_f] \\ \langle \eta^i(t)\eta^j(s)\rangle = \delta^{ij}\delta(t-s) \end{cases} \tag{5}$$

where $\boldsymbol{r}, \boldsymbol{f}(\boldsymbol{r}), \boldsymbol{\eta} \in \mathbb{R}^d$ and $g(\boldsymbol{r}(t)) \in \mathbb{R}^{d\times d}$ such that $C(\boldsymbol{r}) := g(\boldsymbol{r})g(\boldsymbol{r})^\top$ is a symmetric positive-defined matrix. The $\overset{\alpha}{=}$ indicates the choice of the discretization scheme $\alpha \in [0,1]$: $\alpha = 0$ corresponds to Itô's discretization, $\alpha = 1/2$ to Stratonovich's one. The functions $\boldsymbol{f}, g$ have to be at least once differentiable and $g(\boldsymbol{r}) > 0\ \forall \boldsymbol{r}$. Let $A[\boldsymbol{r}]$ a generic observable that may depend on the full history of the process and suppose that it admits a large deviation function $\phi_A(a) := -\lim_{\epsilon\to 0} \epsilon \ln \Pr(A = a)$.

Here and in the rest of the report, the following notation is used: given a stochastic process $r(t)$, $\mathrm{P_{ss}}(E)$ is the probability of an event E when the process r is stationary, $\mathbb{P}(\tilde{r})$ is the (path) probability that $r(t) = \tilde{r}(t)\ \forall t \in [0, t_f]$, and $\Pr(E)$ will indicate the generic probability of an event E.

2.1 Some technical aspects treated in one dimension

Consider the easier $d = 1$ case. This section will motivate why it is reasonable to disregard the effect of the choice of the discretization (section 2.1.1), as well as discuss the equivalence between the Lagrangian and Hamiltonian approaches to get to the MFT action (section 2.1.2).

2.1.1 The effect of discretization is sub-leading

This section will show that the discretization effect on the path integral formulation that lies at the basis of Macroscopic Fluctuation Theory is negligible to first order.

As meticulously discussed in [11], the measure of the process defined in equation 5 for $d = 1$ is:

$$\mathbb{P}\Big[r(t),\, t \in [0,t_f] \Big| r(0) = r_0\Big] \overset{\alpha}{=} J[r]\, e^{-F_W[r]} \tag{6a}$$

$$F_W[r] \overset{\alpha}{:=} \int_0^{t_f} dt\, \frac{1}{2}\left[\frac{\dot{r}(t) - f(r(t)) + 2\alpha\epsilon\, g(r(t))g'(r(t))}{\sqrt{2\epsilon}\, g(r(t))}\right]^2 + \alpha f'(r(t)) \tag{6b}$$

$$J[r] \overset{\alpha}{:=} \prod_{t=0}^{t_f/\Delta t} \sqrt{\frac{1}{4\pi\epsilon\Delta t}}\, \frac{1}{|g(\bar{r}_t)|}, \quad \bar{r}_t = r_t + \alpha(r_{t+1} - r_t) \tag{6c}$$

where in the definition of $F_W[r]$ the discretization is implicit while the Jacobian $J[r]$ is represented in a discretized time-grid with spacing Δt. For the sake of brevity the specification of the discretization α will be omitted in the following.

In this report, the aim of using Macroscopic Fluctuation Theory will always be to evaluate the large deviation function of the observable A. It would therefore be natural to start by evaluating $\mathrm{P_{ss}}(A = a)$ and try to use the $\epsilon \to 0$ limit to get $-\phi_A(a) := \epsilon \ln \mathrm{P_{ss}}(A = a) = \epsilon \ln \langle \delta(A - a)\rangle$ in a saddle point approximation. In order to perform it in presence of the constraint $\delta(A - a)$, its Fourier representation $\delta(A(x) - a) = \int_{\mathbb{R}} dk e^{ik(A(x)-a)}/2\pi$ with $k = -i\lambda/\epsilon$ is used, so that:

$$-\phi_A(a) := \epsilon \ln \langle \delta(A-a)\rangle = \epsilon \ln \int_{i\mathbb{R}} \frac{\epsilon d\lambda}{2\pi i} \langle e^{\frac{\lambda}{\epsilon}(A-a)}\rangle = \epsilon \ln \int_{i\mathbb{R}} \frac{\epsilon d\lambda}{2\pi i} e^{-\epsilon^{-1}[\lambda a - w_A(\lambda)]}$$

where $w_A(\lambda) := \epsilon \ln \langle e^{\epsilon^{-1}\lambda A}\rangle$ is the cumulant generating function of the observable A. A saddle point approximation for $\epsilon \to 0$ directly proves that the large deviation function is the Legendre

transform of the cumulant generating function[13] : $\phi_A(a) = \sup_\lambda[\lambda a - w_A(\lambda)] + \mathcal{O}(\epsilon \ln \epsilon)$. This Legendre transform has an interpretation in the statistical mechanics perspective, which is the passage from a microcanonical ensemble to a canonical one (or "tilted" in the language of out of equilibrium systems, [13]). The probability measures are:

$$\mathbb{P}_a^{\text{micro}}[r] = \frac{\mathbb{P}[r]\delta(A[r]-a)}{\int \mathcal{D}r'\mathbb{P}[r']\delta(A[r']-a)}, \qquad \mathbb{P}_\lambda^{\text{can}}[r] = \frac{\mathbb{P}[r]e^{\epsilon^{-1}\lambda A[r]}}{\int \mathcal{D}r'\mathbb{P}[r']e^{\epsilon^{-1}\lambda A[r']}} = \mathbb{P}[r]e^{\epsilon^{-1}\left(\lambda A[r] - w_A(\lambda)\right)}$$

Like in equilibrium statistical physics, the hard constraint on A in the microcanonical ensemble makes the computation of the (microcanonical) partition function too complicated. Therefore, calculations are performed in the canonical one, hoping that the ensemble equivalence is valid at least far from phase transitions. The observable A and the "Lagrange multiplier" $-\lambda/\epsilon$ play the role of the energy and the inverse temperature. In order to compute the large deviation function, it is necessary to find $\lambda^*(a) := \arg\sup_\lambda[\lambda a - w_A(\lambda)]$, which gives the condition $a = \frac{d}{d\lambda} w_A(\lambda)|_{\lambda^*(a)} = \mathbb{E}_{\lambda^*(a)}[A]$ where $\mathbb{E}_\lambda$ is the expectation value with respect to the canonical probability measure $\mathbb{P}_\lambda^{\text{can}}$. The ensemble equivalence is satisfied for the optimal Lagrange multiplier $\lambda^*(a)$ if the fluctuations of A around its average in the canonical ensemble are negligible, namely if $\sqrt{\text{Var}_{\lambda^*}[A]} \ll \mathbb{E}_{\lambda^*}[A] = a$. By definition of $w_A(\lambda)$, this corresponds to $\epsilon \frac{d^2}{d^2\lambda^2} w_A(\lambda)|_{\lambda^*} \ll a^2$, which is true in $\epsilon \to 0$ and far from phase transitions (where the derivatives of the cumulant generating function may diverge). The out of equilibrium ensemble equivalence is discussed in [13] for Markov processes (included Langevin equations) and in [14] for stochastic hydrodynamics, and it will be assumed to hold in the following. Thus, from now on, the large deviation function of an observable will be obtained as the Legendre transform of its cumulant generating function.

Another problem that arises due to this change of ensemble is that this procedure will only return the convex hull of the large deviation function. This is essentially due to the property of the Legendre transform and is coherent with the Gartner–Ellis theorem [39]. It is, though, a minor issue whose solution will be discussed in section 2.2.1 and implemented in section 2.3.

With this in mind, compute the cumulant generating function by path integral formulation:

$$w_A(\lambda) = \epsilon \ln\left[\int \mathcal{D}r\, \mathbb{P}[r|r_0] e^{\frac{\lambda}{\epsilon}A[r]}\right] \overset{(6)}{=} \epsilon \ln\left[\int \mathcal{D}r\, J[r] e^{\frac{\lambda}{\epsilon}A[r] - F_W[r]}\right] := \epsilon \ln\left[\int \mathcal{D}r\, e^{-\frac{1}{\epsilon}L_\lambda[r]}\right] \tag{7}$$

with the "Lagrangian" action $L_\lambda[r]$ defined as:

$$L_\lambda[r] \overset{\alpha}{=} -\lambda A[r] + \frac{1}{4}\int_0^{t_f} dt\left[\frac{\dot r - f + 2\alpha\epsilon g g'}{g}\right]^2 + \epsilon \int_0^{t_f} dt\left[\alpha f' + \frac{1}{2}\ln(4\pi\epsilon dt |g|^2)\right] \tag{8}$$

where a short-hand notation was used inside the integrals: r is always to be intended as evaluated at t and f, g, f', g' are always intended to be evaluated at $r(t)$. Equation 8 clarify immediately that the discretization α plays a subleading role of $\mathcal{O}(\epsilon \ln \epsilon)$. From now on the effect of discretization will be neglected and Itô discretization will be chosen ($\alpha = 0$).

2.1.2 The Hamiltonian approach is equivalent to the Lagrangian one

The approach used above to derive the MFT action is called "Lagrangian approach" because the action 8 shows a Lagrangian structure. Recalling that $\epsilon \to 0$, define the Lagrangian:

$$\mathcal{L}(r(t), \dot r(t)) := \frac{1}{4}\left[\frac{\dot r(t) - f(r(t)) + 2\epsilon\alpha g(r(t)) g'(r(t))}{g(r(t))}\right]^2$$

[13] Recall that the saddle point can be used also for function of complex variables like $\lambda \in i\mathbb{R}$. The idea is to find the saddle point $\lambda_0 \in \mathbb{C}$ of the function at the exponent of the integrand and use Cauchy theorem for analytical functions in order to deform the path $i\mathbb{R}$ into $\lambda_0 + i\mathbb{R}$, assuming that no singularities are encountered in this deformation.

which is such that $L_0[r] = \int_0^{t_f} dt\mathcal{L}(r(t),\dot{r}(t)) + \mathcal{O}(\epsilon\ln\epsilon)$. Using the Legendre transform[14] on the velocity $\dot{r}$ of this Lagrangian, it is straightforward to obtain the following Hamiltonian structure:

$$\text{Action: } S_\lambda[r,\pi] \stackrel{\alpha}{=} -\lambda A[r] + \int_0^{t_f} dt\,[\pi\cdot\dot{r} - \mathcal{H}(r,\pi)] + \epsilon\int_0^{t_f} dt\Big[\alpha f' + \frac{1}{2}\ln(4\pi\epsilon dt|g|^2)\Big] \tag{9}$$

$$\text{Conjugated momentum: } \pi(t) = \frac{\partial\mathcal{L}(r(t),\dot{r}(t))}{\partial\dot{r}(t)} = \frac{\dot{r}(t) - f(r(t)) + 2\epsilon\alpha g(r(t))g'(r(t))}{2[g(r(t))]^2} \tag{10}$$

$$\text{Hamiltonian: } \mathcal{H}\big(r(t),\pi(t)\big) = \pi(t)\big[f(r(t)) - 2\alpha\epsilon g(r(t))g'(r(t)) + g^2(r(t))\pi(t)\big] \tag{11}$$

Notice that the deterministic part of the Langevin equation ($\dot{r} - f(r)$) is clearly appearing inside the dominant part of the Hamiltonian action multiplied by π. For the physical meaning of the Lagrangian and the Hamiltonian formalisms see section 2.2.1.

A more straightforward way to get directly to the Hamiltonian formulation is the "Hamiltonian" approach. It will be discussed briefly here and thoroughly in section 3.2, where this method will save many computations with respect to the Lagrangian approach. The aim here is to prove that these two approaches are equivalent.

Define $\mathbb{E}_\eta[\psi] := \int\mathcal{D}\eta\,\mathbb{P}[\eta]\psi[\eta]$ the expectation value over the noise, let $u[\eta]$ the path for a given realization of the noise according to equation 5 and consider again the cumulant generating function:

$$\begin{aligned} w_A(\lambda) &:= \epsilon\ln\mathbb{E}_\eta\Big[e^{\frac{\lambda}{\epsilon}A[r[\eta]]}\Big] = \epsilon\ln\int\mathcal{D}r\,e^{\frac{\lambda}{\epsilon}A[r]}\mathbb{E}_\eta\Big[\delta(r[\eta]-r)\Big] = \\ &= \epsilon\ln\int\mathcal{D}r\,e^{\frac{\lambda}{\epsilon}A[r]}\mathbb{E}_\eta\Big[\delta\Big(\frac{\dot{r}-f(r)}{\sqrt{2\epsilon}g(r)} - \eta\Big)\cdot\Big|\det\Big\{\mathcal{J}\big[\eta_0 = (\dot{r}-f(r))/(\sqrt{2\epsilon}g(r))\big]\Big\}\Big|^{-1}\Big] \end{aligned} \tag{12}$$

where a generalization of the change of variables for the Dirac delta function $\delta(F(x)) = \delta(x - x_0)/|F'(x_0)|$ (valid if x_0 is the unique simple zero of the function F in the integration domain) was used. More precisely, as $\delta(r[\eta]-r) = \prod_{t\in[0,t_f]}\delta(r[\eta](t) - r(t))$ with $r[\eta](t) = r_0 + \int_0^t ds\Big[f(r(s)) + \sqrt{2\epsilon}g(r(s))\eta(s)\Big]$, take $x = \eta$ and $F[\eta] = r[\eta] - r$; finding η_0 such that $F[\eta_0] = 0$ means to solve $\forall t\; r[\eta_0](t) = r(t)$, which implies $\dot{r}(t) = \frac{d}{dt}r[\eta_0](t) = f(r(t)) + \sqrt{2\epsilon}g(r(t))\eta_0(t)$. So $\eta_0(t) = [\dot{r}(t) - f(r(t))]/(\sqrt{2\epsilon}g(r(t)))$. Then the Jacobian reads $\mathcal{J}[\eta] = \frac{\delta F[\eta]}{\delta\eta}|_{\eta=\eta_0} = \frac{\delta r[\eta]}{\delta\eta}|_{\eta=\eta_0}$. A rigorous way to evaluate this Jacobian taking into account the discretization is by introducing a time-grid $n = 1,...,T/\Delta t$, so that the Langevin equation reads $r_{n+1} - r_n := \Delta r_n = \Delta t f(r_n + \alpha\Delta r_n) + \sqrt{2\epsilon\Delta t}g(r_n + \alpha\Delta r_n)\eta_{n+1}$. It is easy to see that causality implies that the Jacobian is a triangular matrix and so:

$$\det\Big[\mathcal{J}[\eta]_{n,m}\Big] := \det\Big[\frac{\partial r_m}{\partial\eta_n}\Big] = \prod_n\Big[\frac{\partial r_n}{\partial\eta_n}\Big]$$

The determinant of this Jacobian has already been obtained in [11] and it was exactly such that (using that the noise is Gaussian and white):

$$\mathbb{E}_\eta\Big[\delta\Big(\frac{\dot{r}-f(r)}{\sqrt{2\epsilon}g(r)} - \eta\Big)|\det[\mathcal{J}[\eta]]|^{-1}\Big] := \int\mathcal{D}\eta\,e^{-\frac{1}{2}\frac{(\dot{r}-f)^2}{2\epsilon g^2}}\Big[\delta\Big(\frac{\dot{r}-f(r)}{\sqrt{2\epsilon}g(r)} - \eta\Big)\cdot\Big|\det[\mathcal{J}[\eta]]\Big|^{-1}\Big] = J[r]e^{-F_W[r]}$$

as defined in equation 6. This makes clear the equivalence with the Lagrangian and the Hamiltonian approaches introduced here. Any order $\mathcal{O}(\epsilon\ln\epsilon)$ will be neglected from now on, including the Jacobian that has just been computed.

Going on with the Hamiltonian approach, $\delta\Big(\frac{\dot{r}-f(r)}{\sqrt{2\epsilon}g(r)} - \eta\Big) \approx \delta\Big(\dot{r} - f(r) - \sqrt{2\epsilon}g(r)\eta\Big)$ (up to another negligible Jacobian) and using the Fourier representation $\delta(x-y) = \int_{\mathbb{R}}\frac{dk}{2\pi}\exp[ik(x-y)]$ for each t:

$$\begin{aligned} e^{\frac{w_A(\lambda)}{\epsilon}} &\approx \int\mathcal{D}r\,e^{\frac{\lambda}{\epsilon}A[r]}\mathbb{E}_\eta\Big[\delta\Big(\dot{r} - f(r) - \sqrt{2\epsilon}g(r)\eta\Big)\Big] = \int\mathcal{D}r\int\mathcal{D}k\,e^{\frac{\lambda}{\epsilon}A[r]+i\langle k|\dot{r}-f(r)\rangle}\cdot\mathbb{E}_\eta[e^{-i\langle\sqrt{2\epsilon}gk|\eta\rangle}] \stackrel{\pi=i\epsilon k}{=} \\ &= \int\mathcal{D}r\int\mathcal{D}\pi\,e^{\frac{\lambda}{\epsilon}A[r]+\epsilon^{-1}\langle\pi|\dot{r}-f(r)\rangle}\cdot\mathbb{E}_\eta[e^{-\langle\sqrt{2\epsilon^{-1}}g\pi|\eta\rangle}] \stackrel{\text{Gauss}}{=} \int\mathcal{D}r\int\mathcal{D}\pi\,e^{-\epsilon^{-1}S_\lambda[r,\pi]} \end{aligned} \tag{13}$$

where S_λ is the Hamiltonian action introduced in equation 9 and $\langle\psi|\phi\rangle := \int_0^{t_f} dt\,\psi(t)\phi(t)$.

[14] Or the Hubbard-Stratonovich transform.

2.2 Studying a d-dimensional Langevin equation to get the physical meaning

2.2.1 The actions and instanton equations

This section will discuss the application of Macroscopic Fluctuation Theory to the general setting of a d-dimensional Langevin equation in order to give the physical interpretation of its results. With a straightforward generalization of the two approaches described before and neglecting the orders $\mathcal{O}(\epsilon \ln \epsilon)$, the following Lagrangian and Hamiltonian structures are:

$$\text{Lagrangian action: } L_\lambda[\boldsymbol{r}] = -\lambda A[\boldsymbol{r}] + \frac{1}{4}\langle \dot{\boldsymbol{r}} - f(\boldsymbol{r}) | C(\boldsymbol{r})^{-1} | \dot{\boldsymbol{r}} - f(\boldsymbol{r})\rangle \tag{14a}$$

$$\text{Lagrangian: } \mathcal{L}(\boldsymbol{r}(t), \dot{\boldsymbol{r}}(t)) = \frac{1}{4}[\dot{\boldsymbol{r}}(t) - f(\boldsymbol{r}(t))] \cdot C(\boldsymbol{r}(t))^{-1}[\dot{\boldsymbol{r}}(t) - f(\boldsymbol{r}(t))] \tag{14b}$$

$$\text{Hamiltonian action: } S_\lambda[\boldsymbol{r}, \boldsymbol{\pi}] = -\lambda A[\boldsymbol{r}] + \langle \boldsymbol{\pi} | \big(\dot{\boldsymbol{r}} - f(\boldsymbol{r}) - C(\boldsymbol{r})\boldsymbol{\pi}\big)\rangle \tag{14c}$$

$$\text{Conjugated momentum: } \boldsymbol{\pi}(t) = \frac{\partial \mathcal{L}(\boldsymbol{r}(t), \dot{\boldsymbol{r}}(t))}{\partial \dot{\boldsymbol{r}}(t)} = \frac{1}{2} C(\boldsymbol{r}(t))^{-1}\big(\dot{\boldsymbol{r}}(t) - f(\boldsymbol{r}(t))\big) \tag{14d}$$

$$\text{Hamiltonian: } \mathcal{H}(\boldsymbol{r}(t), \boldsymbol{\pi}(t)) = \boldsymbol{\pi}(t) \cdot [f(\boldsymbol{r}(t)) + C(\boldsymbol{r}(t))\boldsymbol{\pi}(t)] \tag{14e}$$

where the scalar product now is generalized to $\langle \boldsymbol{\psi} | \boldsymbol{\phi} \rangle = \sum_{i=1}^{d} \int_0^{t_f} dt\, \psi^i(t)\phi^i(t)$.

The physical meaning of the Lagrangian action is to define a probability measure in the space of trajectories $\{\boldsymbol{r}(t)\}_{t\in[0,t_f]}$ which reads $\mathbb{P}_\lambda^{\mathcal{L}}[\boldsymbol{r}] = \exp[-\epsilon^{-1}L_\lambda[\boldsymbol{r}]] / \int \mathcal{D}\boldsymbol{r}' \exp[-\epsilon^{-1}L_\lambda[\boldsymbol{r}']]$. As already noticed in section 2.1.1, this is the measure of the canonical ensemble, in which the "inverse temperature" $-\epsilon^{-1}\lambda$ is conjugated to the observable A (analogous to the energy); differently from the equal a priori postulate in thermodynamics, the prior (unconstrained) measure here is $\mathbb{P}[\boldsymbol{r}] \propto \exp[-\epsilon^{-1}L_0[\boldsymbol{r}]]$ and it is not uniform on the admissible states.
These interpretations are valid also for the Hamiltonian action, except that it defines a probability measure in the phase space of paths $\{\boldsymbol{r}(t), \boldsymbol{\pi}(t)\}_{t\in[0,t_f]}$. The connection between the two formalism is again a Legendre transform where $\boldsymbol{\pi}$ is the conjugated field to the velocity $\dot{\boldsymbol{r}}$, but it has not an interesting statistical mechanics interpretation as the Hamiltonian and the Lagrangian probability measure have the same partition function which is $\mathbb{E}_\eta\{e^{\frac{\lambda}{\epsilon}A}\}$. The Hamiltonian formalism is more advantageous because the Hamiltonian is a constant of motion of the stationary equations for the action that will be derived soon (the conservation is true in the case $A = A(\boldsymbol{r}(t_f))$ but it could be generalized, see section 2.2.1).

$w_A(\lambda)$ can be obtained by saddle point approximation (see equation 13). The saddle point equations are the stationary equations for the Hamiltonian action and the detailed computation is shown in section 3.2.2 for the much more complicated case of the Navier-Stokes equation. The result is:

$$\frac{d}{dt}\boldsymbol{r}_I(t) = \boldsymbol{f}(\boldsymbol{r}_I(t)) + 2C(\boldsymbol{r}_I(t))\boldsymbol{\pi}_I(t) \tag{15a}$$

$$\frac{d}{dt}\boldsymbol{\pi}_I(t) = -\lambda\frac{\delta A[\boldsymbol{r}_I]}{\delta \boldsymbol{r}_I(t)} - \big[\boldsymbol{\nabla}_r \otimes \boldsymbol{f}(\boldsymbol{r}_I(t))\big]\boldsymbol{\pi}_I(t) - \boldsymbol{\nabla}_r\big[\boldsymbol{\pi}_I(t) \cdot C(\boldsymbol{r}_I(t))\boldsymbol{\pi}_I(t)\big] \tag{15b}$$

$$\boldsymbol{r}_I(0) = \boldsymbol{r}_0 \qquad \boldsymbol{\pi}_I(t_f) = \lambda\frac{\delta A[\boldsymbol{r}_I]}{\delta \boldsymbol{r}_I(t_f)} \tag{15c}$$

where the outer product $\boldsymbol{u}\otimes\boldsymbol{v} := \boldsymbol{u}\boldsymbol{v}^\top$ and $(\boldsymbol{\nabla}_r\big[\boldsymbol{\pi}_I(t)\cdot C(\boldsymbol{r}_I(t))\boldsymbol{\pi}_I(t)\big])^i = \sum_{kj} \pi_I^k(t)\frac{\partial}{\partial r^i}C^{jk}(\boldsymbol{r}_I(t))\pi_I^j(t)$.

Recalling the ensemble equivalence, these equations 15 predict the most likely path leading from $\boldsymbol{r}_0$ to the large deviation $A = a$ (enforced via the Lagrange multiplier λ) in the time interval $[0, t_f]$. Taking the name from a pseudo-particle introduced in quantum field theory [2], this path is called the "instanton". The name underlines that this pseudo-particle is very localized in time, which is a characteristic feature of large deviation paths; for instance, the transition between two stable states in small-noisy systems is very fast compared to the residence time in these states (consult section 2.3.2 and in particular figure 2b).

Notice that equation 15a is the Langevin equation where the noise is replaced by an effective term that is linear in $\boldsymbol{\pi}_I$. Let $\boldsymbol{r}_D(t)$ the trajectory in the zero noise case starting from $\boldsymbol{r}_0$ and assume it is unique. If the final state is chosen to be an atypical value (any different from $\boldsymbol{r}_D(t_f)$), the typical

path leading to the atypical state is the instanton as it is overwhelmingly more probable than any other path. Moreover, the noise makes the particle deviate from the deterministic trajectory and follow the instanton: the term $2C(\boldsymbol{r}_I(t))\boldsymbol{\pi}_I(t)$ in the instanton equation 15a represents the most likely realization of the noise that brings the system to the large deviation. In this perspective, $\boldsymbol{\pi}_I$ is the "stochastic force" that the noise has to give to the system to reach the atypical state following the instanton path. Notice that if the final state chosen is exactly the one reached by the deterministic dynamics, the solution of equations 15 has $\boldsymbol{\pi}_I(t) = 0\ \forall t$ and indeed, intuitively, no stochastic forcing is required.

The minimum action can be rewritten by plugging equation 15a into the equation 14c:

$$S_\lambda[\boldsymbol{r}_I, \boldsymbol{\pi}_I] = -\lambda A[\boldsymbol{r}_I] + \langle \boldsymbol{\pi}_I | C(\boldsymbol{r}_I)\boldsymbol{\pi}_I \rangle \tag{16}$$

In the case $A = A(\boldsymbol{r}(t_f))$ the Hamiltonian is a constant of motion of the instanton equations 15 because they coincide with the Hamilton equations and the Hamiltonian does not explicitly depend on time (equation 14e)[15]. Therefore the action of equation 14c can also be written as:

$$S_\lambda[\boldsymbol{r}_I, \boldsymbol{\pi}_I] = -\lambda A(\boldsymbol{r}_I(t_f)) + \langle \boldsymbol{\pi}_I | \dot{\boldsymbol{r}}_I \rangle - \mathcal{H}(\boldsymbol{r}_I(0), \boldsymbol{\pi}_I(0)) \tag{17}$$

Once the action is evaluated, the cumulant generating function can be obtained to the leading order in ϵ by saddle point approximation on equation 13: $w_A(\lambda) \approx -S_\lambda[\boldsymbol{r}_I, \boldsymbol{\pi}_I]$.

Now it is possible to understand why the effect of a vanishing noise is non negligible, which is one of the main motivations for the use of Macroscopic Fluctuation Theory. Thanks to the small noise, the system can have large deviations from the deterministic dynamics that can be particularly relevant. For instance, if there are many stable fixed points, according to the deterministic dynamics the final state will be exponentially near to the one whose basin of attraction contains initial condition, while the small noise allows for rare jumps between the fixed points leading to completely different behaviour. The effect of a small noise is very important in turbulence where it induces transitions between stable configurations [17], it might induce a smooth solution of the Navier-Stokes equation to develop local singularities, paving the way for spontaneous stochasticity to come to play, or it can allow the transition between non-unique solutions to the deterministic dynamics, if present. In all cases, the vanishing noise has definitely not a negligible effect.

2.2.2 The quasi-potential and the adjoint dynamics

The generalization to a vectorial observable is straightforward: $\boldsymbol{\lambda}$ will be a vector and all the previous results generalize by $\lambda A[\boldsymbol{r}] \mapsto \boldsymbol{\lambda} \cdot \boldsymbol{A}[\boldsymbol{r}]$. This allows to discuss the case of $\boldsymbol{A}[\boldsymbol{r}] = \boldsymbol{r}(t_f)$, thanks to which MFT will furnish the stationary probability distribution[16] of the process $\mathrm{P_{ss}}(\boldsymbol{r}) \approx \exp\{-\epsilon^{-1}\phi(\boldsymbol{r})\}$ by Legendre transform of the cumulant generating function $w(\boldsymbol{\lambda}) \approx -S_\lambda[\boldsymbol{r}_I, \boldsymbol{\pi}_I]$. Due to the similarity of $\mathrm{P_{ss}}$ with the Gibbs-Boltzmann distribution, the large deviation function $\phi(\boldsymbol{r})$ is also called the "quasi-potential".

A first property of the quasi-potential is that, if the deterministic dynamics admits a stable fixed point $\bar{\boldsymbol{r}}$, then it is a minimum of the quasi-potential, because in the limit $\epsilon \to 0$ the probability distribution of the position should converge to a delta over the deterministic solution[17].
Moreover, the quasi-potential is connected to the momentum $\boldsymbol{\pi}_I$. On the one hand, as $\phi = \mathrm{Leg}[w]$, by Legendre duality $\arg\sup_{\boldsymbol{\lambda}}[\boldsymbol{\lambda} \cdot \boldsymbol{r}_f - w(\boldsymbol{\lambda})] := \boldsymbol{\lambda}^*(\boldsymbol{r}_f) = \boldsymbol{\nabla}_{r_f}\phi(\boldsymbol{r}_f)$. On the other hand, the final condition of equation 15c specialized to this observable implies $\boldsymbol{\pi}_I(t_f) = \boldsymbol{\lambda}$. Thus:

$$\boldsymbol{\lambda}^*(\boldsymbol{r}_f) = \boldsymbol{\nabla}_{r_f}\phi(\boldsymbol{r}_f) = \boldsymbol{\pi}_I(t_f) \qquad \text{with } \boldsymbol{\pi}_I \text{ solution of 15 with } \boldsymbol{\lambda} = \boldsymbol{\lambda}^*(\boldsymbol{r}_f) \text{ and } A = \boldsymbol{r}(t_f) \tag{18}$$

[15] In case of an observable depending on the full realization the Hamiltonian $\mathcal{H}$ of equation 14e is not conserved by the instanton equations 15 due to the term $-\lambda\frac{\delta A[\boldsymbol{r}_I]}{\delta \boldsymbol{r}_I(t)}$ in equation 15b. However, if $A[\boldsymbol{r}] = \int_0^{t_f} dt\, a(\boldsymbol{r}(t))$ there is a conserved Hamiltonian which is $\mathcal{H}_\lambda(\boldsymbol{r}, \boldsymbol{\pi}) = \lambda a(\boldsymbol{r}) + \mathcal{H}(\boldsymbol{r}, \boldsymbol{\pi})$. In the most general case, one should redefine also $\boldsymbol{\pi}$ and alter the structure defined in equation 14. This would be a generalization that is not necessary in the following.

[16] The existence of a stationary probability is an assumption [5], it is also the stationary solution of the Fokker-Planck equation and it is not necessary an equilibrium state.

[17] In the case of multiple solutions as conjectured in section 1.1.1, $\exp[-\epsilon^{-1}\phi(\boldsymbol{r})]$ converges a non-trivial probability distribution over them. This is how to probe spontaneous stochasticity through stochastic regularization.

This relationship is useful to introduce the concept of "adjoint dynamics" from a Hamiltonian perspective (in the literature this is derived by time-reversibility arguments, [5]). This is defined such that the most likely path leading the original system from a stationary state to a rare one is the time reversed relaxation path leading the adjoint dynamics from the rare state to the stationary one. It is assumed that the adjoint system is still described by a Langevin equation. Only its deterministic part is needed to find any relaxation path as, by definition, it does not need to be activated by the noise. In order to find the adjoint dynamics, suppose that the system is at stationarity at the beginning, so $\boldsymbol{r}_0 \sim \mathrm{P}_{\mathrm{ss}}$. As detailed in section 3.2.2 for the case of Navier-Stokes equation, drawing the initial condition from a distribution will only change the initial condition of equation 15c to $\boldsymbol{\pi}_I(0) = \boldsymbol{\nabla}_r \phi(\boldsymbol{r}_I(0))$, where initial position $\boldsymbol{r}_I(0)$ is determined by the instanton equations. By definition of stationary measure, the process will remain stationary in $t \in [0, t_f]$ $\forall t_f$ and so, by the property 18, $\boldsymbol{\pi}_I(t_f) = \boldsymbol{\nabla}_r \phi(\boldsymbol{r}_I(t_f)) \quad \forall t_f$, which means $\boldsymbol{\pi}_I(t) = \boldsymbol{\nabla}_r \phi(\boldsymbol{r}_I(t))$ $\forall t$. This can be substituted in equation 15a resulting in:

$$\frac{d}{dt}\boldsymbol{r}_I(t) = \boldsymbol{f}(\boldsymbol{r}_I(t)) + 2C(\boldsymbol{r}_I(t))\boldsymbol{\nabla}_r \phi(\boldsymbol{r}_I(t))$$

Besides allowing to find the instanton path if ϕ was known, this equation furnishes the deterministic part of the adjoint dynamics as, by its definition, it is described by the adjoint position $\boldsymbol{R}(t) := \boldsymbol{r}_I(t_f - t)$, which gives:

$$\frac{d}{dt}\boldsymbol{R}(t) = -\boldsymbol{f}(\boldsymbol{R}(t)) - 2C(\boldsymbol{R}(t))\boldsymbol{\nabla}_r \phi(\boldsymbol{R}(t)) \tag{19}$$

In practice, the adjoint dynamics is useless because ϕ is unknown, but it gives another definition of an equilibrium process as the one such that the adjoint dynamics coincides with the original one [5]. As an example, the potential Langevin equation with $\boldsymbol{f}(\boldsymbol{r}) = -\boldsymbol{\nabla} U(\boldsymbol{r})$ and $C = 1$ respects this definition. Indeed, $\mathrm{P}_{\mathrm{ss}}(\boldsymbol{r}) \propto \exp[-\epsilon^{-1} U(\boldsymbol{r})]$ is an equilibrium state, it is immediate to see that $U = \phi$ and then the adjoint Langevin equation matches with the original one.

Another classical result of Macroscopic Fluctuation Theory is the Hamilton-Jacobi equation for the quasi-potential. This can be derived in many ways, such as by time reversibility arguments [5] or by substituting the large deviation ansatz $\mathrm{P}_{\mathrm{ss}}(\boldsymbol{r}) \propto \exp(-\epsilon^{-1}\phi(\boldsymbol{r}))$ in the Fokker-Planck equation and keeping the leading order [8]. From the Hamiltonian perspective, its derivation requires just to substitute the property $\boldsymbol{\pi}_I(t) = \boldsymbol{\nabla}_r \phi(\boldsymbol{r}_I(t))$ (equation 18, valid for for stationary initial condition), to send $t_f \to \infty$ and use the conservation of the Hamiltonian (valid because the observable only depends on the final position, see section 2.2.1):

$$\boldsymbol{\nabla}\phi(\boldsymbol{r}_I(t)) \cdot [\boldsymbol{f}(\boldsymbol{r}_I(t)) + C(\boldsymbol{r}_I(t))\boldsymbol{\nabla}\phi(\boldsymbol{r}_I(t))] := \mathcal{H}\big(\boldsymbol{r}_I(t), \boldsymbol{\nabla}\phi(\boldsymbol{r}_I(t))\big) = \mathcal{H}\big(\boldsymbol{r}_I(0), \boldsymbol{\nabla}\phi(\boldsymbol{r}_I(0))\big) = 0 \ \forall t > 0$$

where $\mathcal{H}\big(\boldsymbol{r}_I(0), \boldsymbol{\nabla}\phi(\boldsymbol{r}_I(0))\big) = 0$ because $\boldsymbol{\pi}_I(0) \to 0$ as $t_f \to \infty$. The intuitive reason why the initial momentum vanishes is that the initial state is drawn from the stationary distribution (so it is typical) and the rare event $\boldsymbol{r}(t_f) = \boldsymbol{r}_f \in \mathbb{R}^d$ has to be reached in infinite time[18].

2.2.3 The MFT numerical procedure to estimate the large deviation function

Being able to solve the instanton equations analytically is rare as the instanton equations are pretty complicated and non linear, even if in some non-trivial cases such as the simple exclusion process this was done [32]. Therefore it is very important to open a comment on the numerical strategies and the algorithms to integrate them, depending on the quantity of interest.

Case 1. If the aim is to evaluate the cumulant generating function $w_A(\lambda)$ of the observable $A[\boldsymbol{r}]$ and the initial state $\boldsymbol{r}_0$ is known, it suffices to impose $\boldsymbol{r}_I(0) = \boldsymbol{r}_0$, $\boldsymbol{\pi}_I(t_f) = \lambda \frac{\delta A[r_I]}{\delta \boldsymbol{r}_I(t_f)}$ as it was done in equation 15.

Case 2: If the aim is to evaluate large deviation function of the observable, the calculations above suggest to compute the Legendre transform of $w_A(\lambda)$ numerically. However, this can be quite expensive

[18] To be more precise: from equation 15b for any finite t_f $\boldsymbol{\pi}_I(0) = 0$ is not a good initial momentum for the shooting problem if the final state is atypical because with this initial condition the deterministic dynamics is followed. However, as $t_f \to \infty$ at fixed $\boldsymbol{r}_f$, the process has more and more time to reach the final target state and so the initial perturbation to the deterministic dynamics can become less and less strong, approaching 0 in the infinite time limit.

due to the maximization involved, and a short-cut is possible. Intuitively, $S_0[\boldsymbol{r}, \boldsymbol{\pi}]$ defines a probability measure for the paths without any constraint, so

$$\phi_A(a) = -\epsilon \ln \mathrm{P}_{\mathrm{ss}}\{A[\boldsymbol{r}] = a | \boldsymbol{r}(0) = \boldsymbol{r}_0\} = -\epsilon \ln \int_{A[\boldsymbol{r}]=a} \mathcal{D}\boldsymbol{\pi} \int_{\boldsymbol{r}(0)=\boldsymbol{r}_0} \mathcal{D}\boldsymbol{r}\, e^{-\epsilon^{-1} S_0[\boldsymbol{r},\boldsymbol{\pi}]}$$
$$\overset{\text{saddle}}{\approx} \sup \left[S_0[\boldsymbol{r}, \boldsymbol{\pi}] \middle| \{\boldsymbol{r}(t), \boldsymbol{\pi}(t)\}_{t\in[0,t_f]} : \boldsymbol{r}(0) = \boldsymbol{r}_0,\ A[\boldsymbol{r}] = a \right]$$

Therefore, the numerical computation of the Legendre transform can be avoided and the large deviation function of any observable is obtained (up to an additive constant[19]) by solving the instanton equations with final condition $A[\boldsymbol{r}] = a$ and initial one $\boldsymbol{r}(0) = \boldsymbol{r}_0$ (fixed as a reference while varying a).

From a mathematical perspective, this short-cut is a direct consequence of the ensemble equivalence hypothesis. Intuitively, λ is a Lagrange multiplier enforcing $A[\boldsymbol{r}] = a$ (more precisely on average as discussed in section 2.1.1), so it is possible[20] to avoid the maximization over λ by selecting only the instanton paths such that $A[\boldsymbol{r}] = a$. However, this procedure will still return the convex hull of the large deviation function as if the Legendre transform was performed, because it is hidden in the ensemble equivalence assumption[21].

[8] suggests a way to circumvent this problem for the quasi-potential (so $A = \boldsymbol{r}(t_f)$). Suppose that the deterministic dynamics has n stable fixed points $\{\bar{\boldsymbol{r}}_\alpha\}_{\alpha=1\dots n}$ with basin of attraction Λ_α. As mentioned before (section 2.2.1), each stable fixed point will be a local minimum of the pseudo-potential. By the property of the Legendre transform, if the initial condition $\boldsymbol{r}_0$ is in the basin Λ_α, the $\phi(\boldsymbol{r})$ obtained with the instanton procedure will coincide with the quasi-potential only inside Λ_α (up to a constant). The strategy is then to apply this procedure in each basin independently and then match the constants by continuity of the quasi-potential at the borders. In the following, this strategy and the short-cut presented above constitutes the way to compute the pseudo-potential $\phi(r)$ and will be called "MFT procedure" from now on.

In any case, a numerical difficulty of integrating the instanton equations 15 is the fact that the initial momentum (or position) is unknown and should be fixed such that the final condition is satisfied. This is usually called a "shooting problem" and the initial momentum (or position) is found as the optimal one such that the final condition is met, which numerically requires many integrations of the instanton equations for each optimization. Not only does this increase the computational cost, but it also has convergence problems. The solutions to this numerical problem will not be discussed here. A summary of the best techniques can be found in [37].

2.3 A 1-dimensional system to test the MFT procedure in estimating the pseudo-potential

The aim of this section in to show with a practical non-trivial example the effectiveness of the "MFT procedure" described in section 2.2.3 to estimate the leading order of the large deviation function.

2.3.1 Computing the pseudo-potential with a change of variable

Consider again equation 5 with Itô discretization. This paragraph will prove that, in dimension $d = 1$, if $g(r) \neq 0\ \forall r$ and never changes sign (without loss of generality $g(r) > 0 \forall r$), then a candidate

[19] Theoretically the additive constant is $S_0[\boldsymbol{r}_D, 0]$ with $\boldsymbol{r}_D(t)$ solution of the deterministic dynamics with initial condition $\boldsymbol{r}_0$, but it can be fixed by normalization.

[20] Up to corrections of order $\mathcal{O}(\epsilon)$ and far from phase transitions.

[21] A case in which it is evident that this MFT procedure returns the convex hull of the quasi-potential is a Langevin equation with $d = 1$, $C = 1$ and potential U being a double well with minima in $A < B$ and maximum in M. This process has an equilibrium probability distribution $p_{eq}(r) \propto \exp[-\epsilon^{-1}U(r)]$ and so $\phi = U$. Following the MFT procedure to evaluate the large deviation function, fix the initial reference $r_0 = A$ and suppose to apply the instanton procedure to get the quasi-potential in a point $M < r_f < B$ belonging to the other well's basin of attraction. Recall that going "downhill" (with respect to the potential) does not need noise activation if it is done in the amount of time predicted deterministically t_D. Then if the final time is fixed to $t_f > t_D$, the instanton r_I will reach the maximum M in $t_f - t_D$ (waiting near r_0 if required) and then relax with no activation needed to r_f. As a result, $\pi_I(t) = 0 \forall t > t_f - t_D$ and the pseudo-potential predicted with the MFT procedure will be $\phi_{\mathrm{MFT}}(r_f) = \phi_{\mathrm{MFT}}(M)\ \forall r_f \in [M, B]$. As a consequence, ϕ_{MFT} will be the convex hull of U with global minimum in the initial reference r_0. This is further explained in section 2.3.2.

equilibrium state can be found by mapping the multiplicative noise Langevin equation into an additive noise one. Indeed, apply Itô's lemma [23] for change of variables $y = \psi(r)$ in Langevin equation 5 with Itô discretization:

$$\dot{y}(t) \overset{0}{=} \left\{\psi'(r(t))f(r(t)) + \epsilon g^2(r(t))\psi''(r(t)) + \sqrt{2\epsilon}g(r(t))\psi'(r(t))\xi(t)\right\}\Big|_{r(t)=\psi^{-1}(y(t))}$$

Thanks to the assumption $g(r) > 0$, it is possible to define an invertible change of variable ψ such that $\psi'(r)g(r) = 1\,\forall r$, namely $\psi(r) = \int_{r_{\text{ref}}}^{r} dx(g(x))^{-1}$ with r_{ref} arbitrary reference point. As a consequence, the stochastic process y follows an additive noise Langevin equation $\dot{y}(t) = \tilde{f}(y(t)) + \sqrt{2\epsilon}\xi(t)$ with a modified force field:

$$\tilde{f}(y) = \left\{\psi'(r)f(r) + \epsilon g^2(r)\psi''(r)\right\}\Big|_{r=\psi^{-1}(y)} = \frac{f(r(y))}{g(r(y))} - \epsilon g'(r(y)) \tag{20}$$

where $r(y) := \psi^{-1}(y)$. As $d = 1$ it is always possible to define the potentials corresponding to $f, \tilde{f}$ as $f(r) = -V'(r)$ and $\tilde{f}(y) = -\tilde{V}'(y)$ and by equation 20:

$$\tilde{V}(y) - \tilde{V}(y_0) = \int_{y_0}^{y} dY \frac{V'(\psi^{-1}(Y))}{g(\psi^{-1}(Y))} + \epsilon g'(\psi^{-1}(Y)) \overset{Y=\psi(x)}{=} \int_{\psi^{-1}(y_0)}^{\psi^{-1}(y)} \frac{V'(x)}{g^2(x)}dx + \epsilon \ln\left[\left|\frac{g(\psi^{-1}(y))}{g(\psi^{-1}(y_0))}\right|\right]$$

where y_0 is the arbitrary zero of the potential. If $\tilde{V}$ is such that $q_{\text{eq}}(y) \propto \exp[-\epsilon^{-1}\tilde{V}(y)]$ is normalizable, then it is an equilibrium state for the stochastic process y. By change of variable, if $p_{\text{eq}}(r) \propto |\psi'(r)| \exp[-\epsilon^{-1}\tilde{V}(\psi(r))]$ is normalizable then it is an equilibrium[22] state for $r(t)$. Thus the large deviation function of $r(t)$ results:

$$\phi(r) = \tilde{V}(\psi(r)) - \epsilon \ln|\psi'(r)| = \int_{r_{\text{ref}}}^{r} \frac{V'(x)}{g^2(x)}dx + 2\epsilon \ln|g(r)| \tag{21}$$

where $r_{\text{ref}} = \psi^{-1}(y_0)$ is just a reference point for the pseudo-potential and it contributes as a constant that can be reabsorbed in the normalization. Observe that, in the additive noise case[23] $g = 1$, the result $\phi(r) = V(r)$ is correctly recovered. Concerning this large deviation function, notice that the term of order $\mathcal{O}(\epsilon)$ is due to the choice of Itô discretization (it would be zero in Stratonovich's one). This $\mathcal{O}(\epsilon)$ correction is neglected in the MFT procedure to obtain the pseudo-potential, as it will be clear in section 2.3.2. Notice also that the Hamilton-Jacobi equation 2.2.2 in $d = 1$ has a non trivial solution only for $\phi'(r) = V'(r)/g^2(r)$, which gives the leading order of equation 21.

2.3.2 Example: a 1-dimensional double well with multiplicative noise

The aim of this section is to compare the prediction of equation 21 with the one evaluated using the "MFT procedure" described in section 2.2.3 in an example. Consider the asymmetric double-well potential $V(r) = \frac{1}{4}(r^2 - a^2)^2 + br^3$ with maximum in $M = 0$ and minima in $A = -3b - \sqrt{9b^2 + 4a^2} < 0$ and $B = -3b + \sqrt{9b^2 + 4a^2} > 0$. The noise is multiplicative with $g(r) = (1 + g_0 + \tanh(\alpha r))/(1 + g0)$ and $\alpha, g_0 > 0$ and its effect is to change smoothly the noise amplitude across the two minima[24]. A plot of V and g and a realization of the process is shown in figure 2. The MFT procedure to obtain the pseudo-potential (section 2.2.3) requires the computation of many instanton

[23] Any constant $g \neq 0$ can be reabsorbed in ϵ

[23] As the change of variable is invertible, there is no loss of information in the change of variable so, provided q_{eq} and p_{eq} are normalizable, q_{eq} is equilibrium if and only if p_{eq} is equilibrium. This intuitive claim is proven by using the Fokker-Planck equation (FPE): suppose $p(x)$ equilibrium state of the process x, $y = \varphi(x)$ invertible change of variable (without loss of generality, $\varphi'(x) > 0$) with $x = \psi(y) := \varphi^{-1}(y)$ and let $q(y) = p(\psi(y))/\varphi'(\psi(y))$ the induced distribution of y (suppose normalizable). By assumption, the probability current of the Fokker-Planck equation for p vanishes: $(-f(x) + \epsilon\partial_x)p(x) = 0\,\forall x$. By Itô's lemma the current for $q(y)$ reads: $[-f(\psi(y))\varphi'(\psi(y)) - \epsilon\varphi''(\psi(y))]q(y) + \epsilon\partial_y[(\varphi'(\psi(y)))^2 q(y)]$ which is easy to prove to be vanish. Thus $q(y)$ is an equilibrium state for y.

[24] The normalization $1/(1 + g_0)$ is introduced to enforce that the amplitude of the noise is $\sqrt{2\epsilon}$ on average over any interval $[-x, x]$. This choice is to better visualize the effect of the multiplicative noise as compared to the additive noise case $g = 1$.

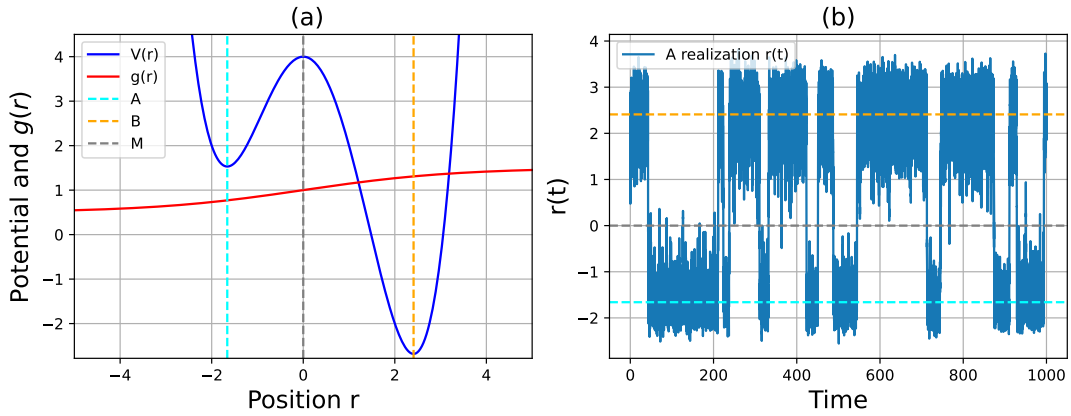

Figure 2: (a) The plot of the potential $V(r)$ and the multiplicative noise function $g(r)$ considered in this particular example. (b) A corresponding realization of the Langevin equation 5 obtained using Itô discretization scheme. Notice that the process jumps many times in A and resides there for long although $V(M)-V(A) < V(M)-V(B)$ because the noise in B is stronger than in A. The parameters are set to $a = 2$, $b = -0.25$, $\epsilon = 1$, $g_0 = 1$, $\alpha = 0.3$. Notice that the noise amplitude ϵ is chosen not so small, as evidenced by the frequent transitions, in order to underline the effect of $\mathcal{O}(\epsilon)$ of equation 21 on the large deviation function.

paths. An example of these is shown in figure 3, with starting point in the minimum A and target in B. The integrations of the instanton equations needed for the shooting method are performed using the standard integrator Explicit Runge-Kutta method of order 5(4) [15], even if it is not a Hamiltonian-conserving (symplectic[25]) integrator, as it is clear from figure 3c.

Notice that the momentum vanishes as soon as the process starts going downhill (from M to B, figure 3b). As a consequence, this portion has null contribution to the Hamiltonian action (recall its expression from equation 16) and $\phi(M) = \phi(B)$ which is false. The inconsistency arises from the Legendre transform involved in the MFT procedure and its explanation is discussed in section 2.2.3. As discussed in that section, for the MFT procedure the initial position is always one of the stable fixed points and the target position is in its basin of attraction. As a consequence, the instantons computed for the pseudo-potential are never going downhill and the instanton of figure 3 is never used.

Figure 4 shows a satisfactory comparison between the pseudo-potential built with the MFT procedure and the one obtained above by the change of variable (equation 21). All the comments are included in the caption. It is remarkable that the effect of the multiplicative noise is able to lower significantly the asymmetry of the potential V. This explains why the realization shown in figure 3b spends approximately the same amount of time in both stable fixed points. The Arrhenius law should be used with the pseudo-potential instead of the potential in out of equilibrium systems [8].

As a concluding remark, in dimension $d > 1$ the change-of-variable procedure described in section 2.3.1 is not applicable because g is a matrix and the deterministic force $f(r)$ might not be potential. Instead, the MFT procedure can be used even when $d > 1$ or for hydrodynamics, and has proven to be successful [8, 32]. The take-home message of this section is that the MFT procedure described in section 2.2.3 is properly working in an example where it was possible to check the result and can promisingly be applied in new cases such as in turbulence.

2.3.3 Choosing the final time t_f

The final time t_f when the instanton is constrained to reach the target r_f is an arbitrary parameter. From a theoretical point of view [5], it should be set to ∞ for the instanton to be as close as possible to the physical path, because the Langevin equation 5 is free to get to the rare position at any time. Indeed, the most likely way for the process to have a large deviation is to remain near the steady state for long, doing typical fluctuations, and then reach r_f very rapidly.

From a numerical perspective, however, not only is $t_f = \infty$ unfeasible, but also instances of

[25] The implementation of symplectic integrators such as Stormer-Verlet or Ruth integrators [21] is not possible because they both need a separable Hamiltonian. This property is not frequent in the Hamiltonian structure of MFT formalism.

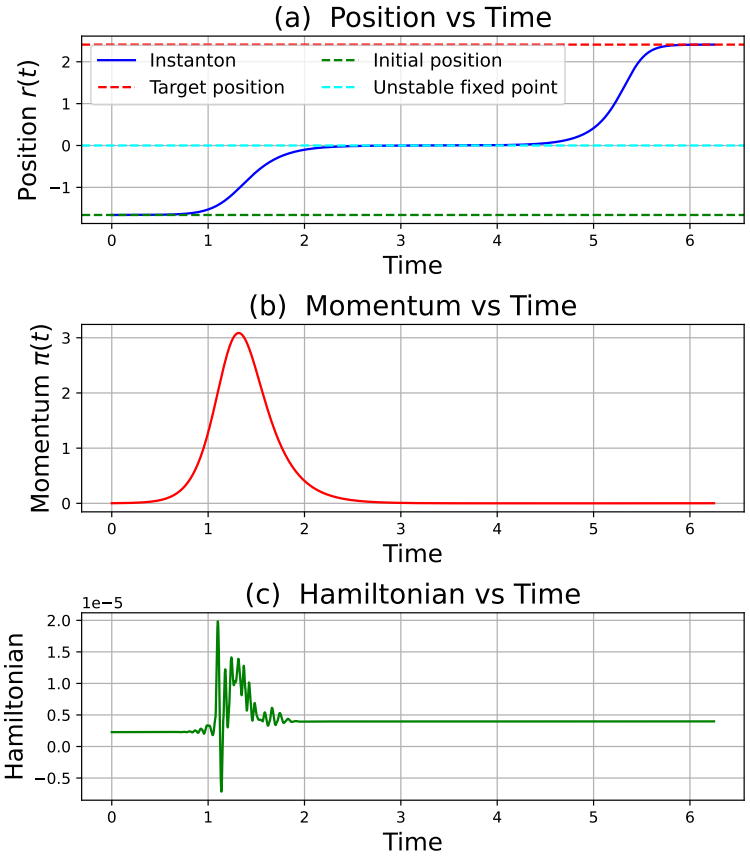

Figure 3: The images show the most likely path leading from the stable fixed point A to the other B (i.e. the instanton from A to B) for the double-well potential $V(r) = \frac{1}{4}(r^2 - a^2)^2 + br^3$. *The computation involved solving the shooting problem defined by the instanton equations 15 with* $r_I(0) = A$, $r_I(t_f) = B$, *as explained in section 2.2.3. The images (a), (b), (c) respectively shows the evolution of the position of the instanton* $r_I(t)$, *the momentum* $\pi_I(t)$ *and the Hamiltonian* $\mathcal{H}(r_I(t), \pi_I(t))$ *(equation 14e for* $\lambda = 0$*). The parameters are set to* $a = 2$, $b = -0.25$, $\epsilon = 1$, $g_0 = 1$, $\alpha = 0.3$ *and the final time* t_f *is optimized via the bisection method explained in section 2.3.3.*
From figure (c) notice that the Hamiltonian is not conserved due to the use of the non-symplectic integrator Runge-Kutta 4(5) [15], but the maximum deviation with respect to the initial value $\sim 10^{-5}$ *is negligible. Notice that* $r_I(t)$ *tends to "wait" near the point A and also "rest" near the maximum M. For any larger* t_f, *the instanton simply spends more time near A and M, while the portions of the trajectory where it climbs or descends the potential remain unchanged. This suggests that the typical large deviation path does not last for all the available time* t_f *and has a typical duration that will be estimated in section 2.3.3. Observe also that the momentum vanishes as soon as the process starts going downhill: from M to B the instanton follows the relaxation path with no activation needed.*

numerical artefacts were observed when t_f was too large. The first artefact is that the optimal $\pi_I(0)$ that solves the shooting problem can result smaller than the numerical precision. In these cases, the shooting optimization does not converge properly and the predicted instanton does not reach the target.
In other instances, in presence of 2 or more stable fixed points, the instanton jumps many times between them before reaching the target, even if the jump was "not needed" to reach it. This phenomenon can be attributed to the fact that if the final time is significantly larger than the average residence times in the fixed points, it is improbable that the process will wait until the final time in the basin of attraction of the fixed point, where it was at the beginning. Both artefacts are deleterious for the evaluation of the pseudo-potential: the former makes the estimate inaccurate because the target is not reached, the latter leads to an overestimation of $\phi(r_f)$ because each jump increases the total action as it requires going up-hill.
A further issue, which though does not affect the pseudo-potential, is that for too large t_f the instanton rests in the a fixed point for long before reaching r_f. This only increases the computational cost.

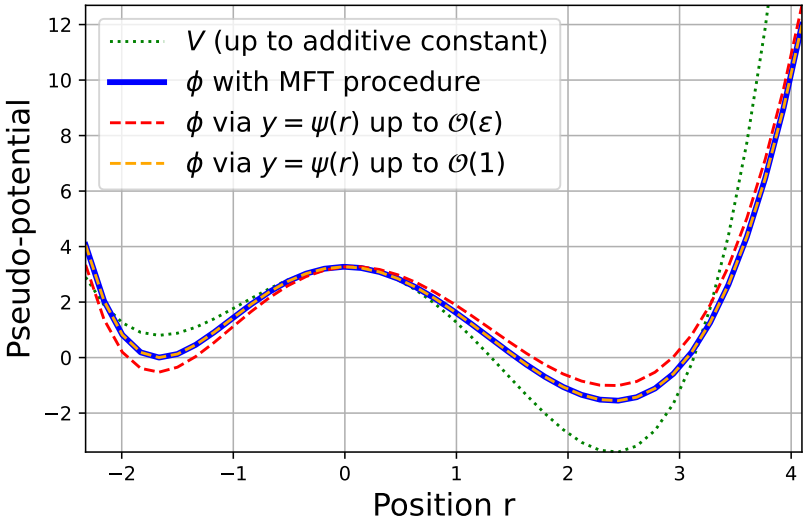

Figure 4: The image reports the asymmetric double-well potential $V(r)$ (green) and three estimates of the large deviation function $\phi(r)$: in blue the result of the MFT procedure (see section 2.2.3), in red the prediction of the change of variable procedure (equation 21) considering the correction of order $\mathcal{O}(\epsilon)$ and in orange the same neglecting this correction. Notice the satisfactory agreement between the MFT procedure and equation 21 up to the small discrepancy due to the fact that the MFT procedure neglects $\mathcal{O}(\epsilon)$. Observe that $\phi(r) \to \infty$ as $|r| \to \infty$, so the distribution $\exp[-\epsilon^{-1}\phi(r)]$ is normalizable. Observe also that the effect of $g(r)$ is to reduce the asymmetry between the two stable fixed points A, B: intuitively the stronger noise in B competes with the potential bias $V(B) - V(A)$. The parameters were set to: $a = 2$, $b = -0.25$, $\epsilon = 1$, $g_0 = 1$, $\alpha = 0.3$. Notice again that the noise amplitude was chosen sufficiently large to evidence the effect of $\mathcal{O}(\epsilon)$ of equation 21 on the pseudo-potential.

A practical solution is to optimize t_f using the "bisection method", an empirical method based on the following observations. It is better to avoid values that are too small so that the initially the instanton is approximately not activated; this is easily spotted as the corresponding $\pi_I(0)$ is "too large". It is also necessary to avoid values of t_f that are too large, which may introduce the artefacts explained above; this is the case when $\pi_I(0)$ is near to the double-precision, or the final condition is not reached, or the r_f is crossed before t_f. Taking into account these considerations, the "bisection method" is the following algorithm 1:

This practical solution is working properly but is computationally expensive and not satisfactory from a theoretical perspective. The following paragraphs will try to find a theoretical estimate for t_f via an "equilibrium assumption". The main idea is that large deviations happen spontaneously during the evolution of a stochastic process: usually the process is performing typical fluctuations around a steady state of the deterministic dynamics when it starts following a large deviation path and very rapidly reaches the rare event (see for instance the transitions between the two stable fixed points in figure 2b). Imagine to integrate the Langevin equation 5 many times and collect a sample of paths going from a steady point r_0 to an atypical position r_f without re-crossing any of them. Define the duration of such a path $t_I(r_0 \to r_f)$. Atypical positions are typically reached following the most likely path leading to them [5, 8], so the distribution of $t_I(r_0 \to r_f)$ would be quite peaked around its average value. Define $\tau_I(r_0 \to r_f) := \langle t_I(r_0 \to r_f) \rangle$ and notice that it is a characteristic time-scales of the process, in the sense that it emerged without constraining on the large deviation to be realised. On the other hand, the instanton predicted by equations 15 represents the most likely path leading from $r_I(0) = r_0$ to $r_I(t_f) = r_f$ in the fixed time interval $[0, t_f]$. As already discussed in section 2.2.1, the instanton is an allowed path, so its probability is $\mathbb{P}[r_I(t) : t \in [0, t_f]] \propto \exp[-\epsilon^{-1} S_0[r_I, \pi_I]]$ where $S_0[r, \pi]$ is the Hamiltonian action[26] introduced in equation 14e (notice that π_I can be obtained from the knowledge of the full path r_I by equation 14d). It is evident that $S_0[r_I, \pi_I]$ depends on t_f both explicitly (t_f is the right extremum of integration over time) and implicitly (as the instanton depends on t_f). However, as the interest is on the properties of the unconstrained Langevin equation 5, the

[26] It was set $\lambda = 0$ because the procedure aims at the pseudo-potential, as explained in section 2.2.3.

Algorithm 1 Bisection method to find the optimal t_f

Initialize t_f^0 (e.g. with the relaxation estimate given in equation 23)
if t_f^0 is too small **then**
  Set $t_f^0 \leftarrow 2 \cdot t_f^0$ and try again
else if t_f^0 is too large **then**
  Set $t_f^{\min} \leftarrow 0$, $t_f^{\max} \leftarrow t_f^0$
  repeat
    $t_f^{\text{cand}} \leftarrow \frac{1}{2}(t_f^{\min} + t_f^{\max})$
    if t_f^{cand} is too small **then**
      $t_f^{\min} \leftarrow t_f^{\text{cand}}$
    else if t_f^{cand} is too large **then**
      $t_f^{\max} \leftarrow t_f^{\text{cand}}$
    else
      Accept t_f^{cand}
    end if
  until t_f^{cand} is suitable
else
  Accept t_f^0
end if

aim is to compute the probability of observing the process going from r_0 to r_f without constraints on the time when r_f is reached. In maths, this probability is:

$$\Pr\{\exists t_f : r(t_f) = r_f | r(0) = r_0\} = \int_0^\infty dt_f \int_{r(0)=r_0}^{r(t_f)=r_f} \mathcal{D}r \int \mathcal{D}\pi \, e^{-\epsilon^{-1} S_0[r,\pi]}$$

which, in saddle point perspective, is dominated by the "most likely instanton" defined by equations 15 with final time $t_f^* \in \arg\min_{t_f} S_0[r_I, \pi_I]$. Since an atypical state is typically reached by the most likely path leading there, this most likely instanton will correspond to the large deviation path observed experimentally. As a consequence, a good estimate for t_f^* is $\tau_I(r_0 \to r_f)$.

Wrapping up, t_f is an arbitrary parameter of the MFT formalism, there exists an optimal t_f and it coincides with $\tau_I(r_0 \to r_f)$. Once the latter is known, the final time t_f of the instanton equations 15 is set according to the final aim. If the purpose is to estimate the pseudo-potential ϕ (the MFT procedure described in section 2.2.3), it is sufficient to choose $t_f \geq \tau_I(r_0 \to r_f)$ (and not too large to avoid the numerical artefacts discussed above); the reason is that the r_0 considered in the MFT procedure are the stable steady states for the deterministic dynamics and so the instanton computed for t_f will "wait" in r_0 for the time in excess $t_f - \tau_I(r_0 \to r_f)$. If instead the aim is to estimate the most likely path leading from r_0 to r_f, then it is important to choose $t_f = \tau_I(r_0 \to r_f)$.

Now the aim is to estimate the average duration of the large deviation path leading from r_0 to r_f without recrossing. The focus will be on the rare transition between the two stable fixed points of the double-well potential $\tau_I(A \to B)$ because it represents the richest case, but the approximation is valid for generic initial and final positions.

As seen in section 2.3.1, if the pseudo-potential of equation 21 is normalizable then $p_{\text{eq}}(r) \propto \exp[-\epsilon^{-1}\phi(r)]$ is an equilibrium state for the process $r(t)$. If equilibrium holds, time reversibility ensures that the instanton path $A \to M$ is the time reversed of the relaxation path $M \to A$ [5, 27]. Notice that this statement is non-trivial for multiplicative noise because $\phi \neq V$, but the following direct calculation proves that it still holds. In general, the instanton path from A to M is the time reversed relaxation path from M to A in the adjoint dynamics (equation 19), so it suffices to prove that the latter has the same deterministic part as the Langevin equation 5. Plugging equation 21 into

equation 19 (with $g(r) > 0$ without loss of generality):

$$\dot{R}(t) = -f(R(t)) - 2C(R(t))\phi'(R(t)) = -f(R(t)) - 2g^2(R(t))\frac{\partial}{\partial R(t)}\left[2\epsilon \ln(g(R(t))) - \int_{-\infty}^{R(t)} dx \frac{f(x)}{g^2(x)}\right] =$$
$$= f(R(t)) - 4\epsilon g(R(t))g'(R(t)) = f(R(t)) + \mathcal{O}(\epsilon)$$

Therefore, up to the negligible order $\mathcal{O}(\epsilon)$, the adjoint dynamics coincides with the original deterministic dynamics. This is a direct consequence of the fact that p_{eq} is an equilibrium measure. As a result, the time taken by the instanton to climb from A to M is equal to the time taken by the deterministic dynamics to relax from M to A (up to order $\mathcal{O}(\epsilon)$): $\tau_I(A \to M) = \tau_{\text{Relax}}(M \to A)$. Therefore, it is possible to assume a priori equilibrium and estimate the duration of the instanton from A to B as:

$$\tau_I(A \to B) = \tau_I(A \to M) + \tau_{\text{Relax}}(M \to B) \overset{\text{Equilibrium}}{=} \tau_{\text{Relax}}(M \to A) + \tau_{\text{Relax}}(M \to B)$$

As the pseudo-potential is unknown, the estimate will be exact up to order $\mathcal{O}(\epsilon)$ if a posteriori equilibrium holds, otherwise it will be a good initial guess for the bisection method 1 described above.

In the generic case of $\tau_I(r_0 \to r_f)$ with no fixed points in $[r_0, r_f]$, the time taken by the relaxation path is easy to estimate because the deterministic dynamics arrives in r_f in finite time:

$$\tau_I(r_0 \to r_f) = \tau_{\text{Relax}}(r_0 \to r_f) = \{T : r(T) = r_f \text{ where } r(t) \text{ solving } \dot{r}(t) = f(r(t)),\, r(0) = r_0\} \quad (22)$$

If otherwise a fixed point is in $[r_0, r_f]$, some tolerances should be used because the time taken by the relaxation path to reach (or deviate from) a fixed point in the deterministic dynamics is infinity. Consider the worst case scenario[27] $r_0 = M$ and $r_f = A$. In practice, one evaluates the time taken by the relaxation path starting in $M - \sigma_M$ to reach $A + \sigma_A$, with $0 < \sigma_M, \sigma_A \ll 1$ tolerances that can be estimated as follows[28]. For σ_A, the idea is that there are numerous typical paths going to A from a point sufficiently close-by, driven by typical fluctuations. The instanton is reliable only in atypical regions and there it is not a good prediction of the trajectory. It is then reasonable to set the tolerance of the order[29] of the typical fluctuations around the minimum of the potential: $\sigma_A = \sqrt{\epsilon/V''(A)}$. For the unstable fixed point M the same reasoning is not applicable because it is an atypical position. The idea is that the deterministic force driving the position towards the minima dominates over the noise as soon as the position $r(t)$ slightly deviates from the maximum M, so let $[M - \sigma_M, M + \sigma_M]$ the interval where the noise is dominating the dynamics. $\sigma_M > 0$ because the noise dominates the Langevin equation 5 at least in $r = M$ because $f(M) = 0$. In the vicinity of M, $V(r) \approx V(M) + \frac{1}{2}V''(M)(r-M)^2$ so the deterministic typical escaping time from M is $\tau_M = |1/V''(M)|$. Then define σ_M as the deviation from M reached in τ_M thanks to the noise. As near M the deterministic forcing is very weak, the motion is Brownian around the maximum and the typical deviation from the starting point M in time τ_M will be $\sigma_M = \sqrt{2\epsilon\tau_M}$ by diffusion.

Wrapping up for the case $A \to B$, the "equilibrium assumption" assumes equilibrium and estimates the duration of the instanton as:

$$\tau_I(A \to B) \gtrsim \tau_I(A + \sigma_A \to B - \sigma_B) \approx \tau_{\text{Relax}}(M - \sigma_M \to A + \sigma_A) + \frac{2}{|V''(M)|} + \tau_{\text{Relax}}(M + \sigma_M \to B - \sigma_B) \quad (23)$$

with $\sigma_A = \sqrt{\epsilon/V''(A)}$, $\sigma_B = \sqrt{\epsilon/V''(B)}$ and $\sigma_M = \sqrt{2\epsilon/|V''(M)|}$. Notice that the time spent spontaneously by an atypical path near the maximum (so in $[M - \sigma_M, M + \sigma_M]$) is estimated as $2 \cdot \tau_M = \frac{2}{|V''(M)|}$ (recall the definition of τ_M as the time to deviate of σ_M from M).

3 Results: Macroscopic Fluctuation Theory for turbulence

After having seen the simple but insightful application of MFT to the Langevin equation case (section 2), this chapter will treat the case of a stochastic Navier-Stokes equation, which represents a quite general model of turbulence as motivated in section 1.1.

[27] If the interval (r_0, r_f) contains a fixed point, split it on its left and right.

[28] For the path from M to B signs are inverted.

[29] The order of the typical fluctuations around a minimum in a system described by the Langevin equation 5 can be found by a parabolic approximation of the potential near the minimum A. Using the results for the Ornstein–Uhlenbeck process [23], the standard deviation from the minimum is $\sqrt{\epsilon/V''(A)}$.

3.1 From the hypothesis to a final hydrodynamics amenable for MFT

3.1.1 Hypothesis

Referring to section 1.1.4 for the details, the stochastic hydrodynamics to be considered is:

$$\begin{cases} \partial_t \overline{u}^l(x,t) + \overline{u}^l(x,t)\cdot\nabla\overline{u}^l(x,t) = -\nabla\overline{p}^l(x,t) + \nu\Delta\overline{u}^l(x,t) + \nabla\cdot\xi(\overline{u}^l,x,t) \\ \nabla\cdot\overline{u}^l(x,t) = 0, \qquad \overline{u}^l(x,0) = u_{\text{in}}(x), \qquad t\in[0,t_f]. \end{cases} \tag{24}$$

where the vectorial notation was and will be omitted except when necessary.

The distribution of the noise tensor ξ is hard to give a priori, so the following 4 assumptions are made.

Approximation 1: zero average noise. Following [20], define the coarse grained scales $u-\overline{u}^l := \tilde{u}^l$ and rewrite the sub-grid tensor (section 1.1.4) using $\overline{\overline{u}^l}^l = \overline{u}^l$ and $\overline{\tilde{u}^l}^l = 0$ as:

$$\tau_l := \overline{u\otimes u} - \overline{u}\otimes\overline{u} = \overline{(\overline{u}+\tilde{u})\otimes(\overline{u}+\tilde{u})} - \overline{(\overline{u}+\tilde{u})}\otimes\overline{(\overline{u}+\tilde{u})} = -(\overline{\overline{u}\otimes\overline{u}} - \overline{u}\otimes\overline{u}) - (\overline{\overline{u}\otimes\tilde{u}} + \overline{\tilde{u}\otimes\overline{u}} + \overline{\tilde{u}\otimes\tilde{u}}) \tag{25}$$

where the superscript l was dropped for readability. The first parenthesis $(\overline{\overline{u}\otimes\overline{u}} - \overline{u}\otimes\overline{u})$ is a systematic contribution, whereas the second is assumed to be the stochastic one. In this report, since l is considered to be small, the systematic contribution is a sub-leading deterministic term in equation 24 and it will be disregarded in the following. This means that the sub-grid tensor is approximated to be a zero-mean noise: $\tau_l = \xi_l$. In [20] this average term is not neglected as it is useful for the application of techniques from statistical field theory.

Assumption 2: vanishing amplitude noise. As already mentioned, $\tau_l \to 0$ if $l \to 0$ and this means that the noise should be assumed to have a vanishing amplitude $n_0(l)$: $\xi_l = \sqrt{n_0(l)}\hat{\xi}$. It is possible to define n_0 such that $\hat{\xi}$ adimensional (so[30] $[n_0] = (\text{m/s})^4$). The definition of the sub grid tensor does not give any information about the scaling of n_0 with l, so it is assumed to be competing with the non-linear convective term of Navier-Stokes equation for small scales. Imposing $\mathcal{O}[\overline{u}^l\cdot\nabla\overline{u}^l] = \mathcal{O}[\nabla\cdot\sqrt{n_0(l)}\hat{\xi}]$ entails $n_0(l)\sim l^{4h}$ as $l\to 0$ (see section 3.1.2 for more details).

Assumption 3: multiplicative Gaussian noise. As the stochastic contribution of τ_l depends explicitly on $\overline{u}^l$ (see equation 25), a priori $\langle\hat{\xi}^i_a(\overline{u}^l,x,t)\hat{\xi}^j_b(\overline{u}^l,y,s)\rangle = \chi^{ij}_{ab}(\overline{u}^l,x-y,t-s)$ depends on the full velocity field realization. For simplicity, a local dependence of the correlation on the field will be assumed: $\chi([\overline{u}^l],x-y,t-s) = \chi(\overline{u}^l(x,t),\overline{u}^l(y,s),x-y,t-s)$. Furthermore, it will be supposed that the multiplicative part of the correlation matrix is fully specified by a positive scalar function σ such that $\chi(\overline{u}^l(x,t),\overline{u}^l(y,s),x-y,t-s) = \sqrt{\sigma(\overline{u}^l(x,t)/u_0)}C(x-y,t-s)\sqrt{\sigma(\overline{u}^l(y,s)/u_0)}$. The macroscopic velocity scale u_0 (see section 3.1.2) is included so that σ is correctly adimensional.

Wrapping up, the only really unmotivated assumptions up to now are the Gaussian distribution for the noise and the fact that σ is scalar. A future perspective will be to use direct numerical simulations of Navier-Stokes equation (DNS) to infer the shape of the distribution and the correlations of the sub-grid tensor. For the latter, a theoretical alternative is to impose a self-consistency loop such that the correlations of observable $A = \overline{\tau}_l(x,t_f) \overset{l\to 0}{\approx} (\overline{\overline{u}\otimes\overline{u}} - \overline{u}\otimes\overline{u})|_{(x,t_f)}$ match with the assumed noise correlation. The use of a Gaussian noise can be justified by the hydrodynamical construction that underlies this coarse grained Navier-Stokes equation 24, as discussed in section 1.2.

So far the model for the sub-grid tensor reads:

$$\tau_l(x,t) = \sqrt{n_0(l)\sigma(u(x,t)/u_0)}\eta(x,t) \quad \text{with} \quad \langle\eta^{ai}(x,t)\eta^{bj}(y,s)\rangle = C^{ij}_{ab}(x-y,t-s) \tag{26}$$

and with $\eta(x,t)$ Gaussian zero-mean noise. As all correlation operators, $C^{ij}_{ab}(x-y,t-s)$ is symmetric under the transformation $(i,a,x,t)\leftrightarrow(j,b,y,s)$ and is positive defined.

Assumption 4: a solenoidal noise. A further assumption is that the noisy term $\nabla\cdot\sqrt{n_0\sigma}\eta$ is solenoidal (it respects $\nabla\cdot u = 0$), which mathematically means $\sum_{ij}\partial_i\partial_j\sqrt{n_0\sigma}\eta^{ij} := \text{tr}[\nabla\otimes\nabla\sqrt{\sigma(v)}\eta] = 0$. This is never evidently used in the following discussion up to section 3.3, but it is reasonable because any non solenoidal component of this term would be compensated by the pressure field.

[30] The physical dimensions of a quantity Q is labelled as $[Q]$ and given in the SI units of measure.

Approximation on the discretization effect. The choice of the discretization for the stochastic PDE 24 will be considered irrelevant in the vanishing noise limit coherently with what discussed in section 2.1.1. The literature [7, 20] mentions further reasons that are specific to the NSE and will be briefly discussed in section 3.2.1. Within this approximation, the Itô discretization is chosen for simplicity.

The spatial boundary conditions. They would be very system dependent and here the discussion is kept general to see how Macroscopic Fluctuation Theory can be applied to the stochastic Navier-Stokes equation. For simplicity, they are chosen to have vanishing boundary terms arising from the spatial integrations by parts that will be performed in section 3.2.2. A possibility is to use PBCs on an hyper-cubic box $[0, L]^d$, which is useful for numerical simulations [35], or take $x \in \mathbb{R}^d$ with $u(x,t) \to 0$ as $|x| \to \infty$. Otherwise, if the system is confined in a domain of boundary $\partial\Lambda$, the boundary terms of the spatial integration by parts are still set to zero by the local equilibrium assumption that underlies any hydrodynamical description (see section 1.2.1). The reason is that this hypothesis forces the system to be constantly in local equilibrium with the boundaries, so atypical deviations are not allowed [5], meaning $\pi(x,t) = 0\ \forall x \in \partial\Lambda\ \forall t$ (which makes the boundary terms vanish).

Assumption: the limit Re $\to \infty$ faster than $l \to 0$. The interesting limit of equation 24 is the one such that spontaneous stochasticity is relevant: the noise should vanish sufficiently slowly in the perfect coarse-graining limit ($l \ll$ size of the system) with respect to the inverse of the Reynolds number[31] $\mathrm{Re}^{-1} \to 0$. From a mathematical perspective, this limit is necessary to see the effect on NSE of the singularities that are present in the 3d Euler equations. The reasons why this double limit should be taken carefully are mainly two.

From a coarse-graining perspective, the effect of the noise on the macro-scale is relevant only if the disregarded scales can foster the real butterfly effect (section 1.1.3): a sufficient condition is $l > \eta_K$ with $\eta_K \sim (\nu^3/E_D)^{1/4}$ being the Kolmogorov scale[32]. If the $l \to 0$ is taken too fast, the coarse-grained scales are highly dissipative and the singularities amplifying the noise cannot develop, resulting in a completely negligible effect of the noise.

From a stochastic regularization perspective, the vanishing noise is a way to probe spontaneous stochasticity in the Re $\to \infty$ limit when Navier-Stokes equation approaches the Euler equation which has been proven to have singular and non unique solutions [9]. If the noise is sent to zero before the singularities start playing a role, the solutions of the Navier-Stokes equation will be unique and regular, and spontaneous stochasticity will not be observed.

One may object that this is a rather narrow limit and, consequently, spontaneous stochasticity is not physically relevant, but this is not the case. As an example, in [1] the Landau-Lifshitz Navier-Stokes equation is considered, namely equation 24 with only thermal noise ($\sigma = 1$), which is the weakest form of noise one could consider in the system. The authors find that an amplitude of the noise vanishing as $\mathrm{Re}^{-15/4}$ is a sufficient condition for the thermal noise to trigger spontaneous stochasticity. This scaling with the Reynolds number is not unphysical because it originates from imposing a generalization of the Einstein fluctuation-dissipation relation.

For the purposes of this report, it suffices that the vanishing noise limit $l \to 0$ is taken after (or sufficiently slower than) the high Reynolds number limit in order to have a non-vanishing effect of the noise. A more precise condition will be stated in section 3.1.2.

The aim of applying Macroscopic Fluctuation Theory to the stochastic Navier-Stokes equation 24 is to recover the large deviations function of any observable $A[u]$ in the small noise limit that is implied by $n_0 \to 0$ as $l \to 0$. The observable $A[u]$ can generally depend on the full trajectory over time, but the applications will be to observables depending only on the final time, e.g. a component of the vorticity or the dissipation $\frac{1}{2}||u(x,t)||^2$. In the next section 3.2 the application of Macroscopic Fluctuation Theory to this setting will be discussed with all the mathematical technicalities. Before (sections 3.1.2 and 3.1.3), the aim will be to rewrite equation 24 in a more amenable way for the

[31] Re $\to \infty$ is physically relevant because a fully developed turbulent flow has Re > 5000 and in the atmospheric boundary layer Re $\sim 10^8$.

[32] This is a classical result of Kolmogorov theory of turbulence and is based essentially on applying dimensional analysis under the assumption that at small scales in turbulent flows only the viscosity and the rate of energy dissipation per unit mass E_D are relevant.

application of MFT. Finally, in section 3.3 the correlation matrix will be specialized to a simpler case in order to get some more intuition. From now on, for brevity let $u := \overline{u}^l$, but the fact that the coarse-graining has been done should never be forgotten.

3.1.2 Making the stochastic Navier-Stokes equation adimensional

Plugging equation 26 into equation 24 the stochastic Navier-Stokes equation reads:

$$\partial_t u + u \cdot \nabla u = -\nabla P + \nu \Delta u + \nabla \cdot \sqrt{n_0(l)\sigma(u/u_0)}\eta$$

Dimensions are: $[\eta] = [\sigma] = 1, [P] = \mathrm{m}^2/\mathrm{s}^2$, $[n_0] = \mathrm{m}^4/\mathrm{s}^4$, $[\nu] = \mathrm{m}^2/\mathrm{s}$. The following rescaling is introduced: $x' = x/x_0$, $t' = t/t_0$, $u' = u/u_0$, $P' = P/u_0^2$, $\mathrm{Re} = x_0 u_0/\nu$ with $x_0, t_0, u_0 := x_0/t_0$ being respectively the macroscopic characteristic length, time and velocity scales. The rescaled equation is:

$$\partial_{t'} u' + (u' \cdot \nabla_{x'})u' = -\nabla_{x'} P' + (\mathrm{Re})^{-1}\Delta_{x'} u' + \sqrt{\frac{n_0}{u_0^4}}\nabla_{x'} \cdot \sqrt{\sigma(u')}\eta$$

The idea is to choose x_0, u_0 (and so t_0) such that the amplitude of the rescaled noisy term $\epsilon := n_0(l)/u_0^4$ vanishes as $l \to 0$. In this way, the saddle point approximation that underlies MFT (discussed for Langevin equations in section 2 and used again for NSE in section 3.2.2) is sensible. As equation 24 results from a coarse-graining where fluctuations are considered important, both the macroscopic and the microscopic scales play a role. At the macro-scale, the system has a characteristic length-scale x_0 that can be fixed to the integral length, which is the one the external forcing sustaining turbulence is applied to. The macroscopic velocity field u_0 can be chosen as the characteristic value of the mean velocity field, obtained by averaging out the noise. At the micro-scale, the length-scale is the coarse-graining cut-off l. For the microscopic velocity, the small scale fluctuations $\delta u(x;\ell) := u(x+\ell) - u(x)$ have a characteristic value δu which is known from multi-fractal theory of turbulence [22] to scale as:

$$\delta u(x;\ell) := u(x+\ell) - u(x) \sim \begin{cases} U \cdot (|\ell|/x_0)^h & \text{if } |\ell| > \eta_h \\ U' \cdot (|\ell|/x_0) & \text{if } |\ell| \leq \eta_h \end{cases}$$

where $h \in (0, 1]$ is the Hölder exponent, $\eta_h \sim \nu^{\frac{1}{1+h}}$ is the Hölder regularization length[33], U and U' are further typical velocities that do not scale with ℓ and are introduced only for dimensional reasons. A natural choice is $|\ell| = l$ and, generalizing what discussed in section 3.1.1, the limits $l, \mathrm{Re}^{-1} \to 0$ are assumed[34] to be taken such that at least $l > \eta_{h_{\min}}$ where $h_{\min} \leq 1/3$ is the minimal Hölder exponent the system shows. In this way, the effect of singularities will emerge and spontaneous stochasticity will be triggered.
Wrapping up, $\delta u \sim U \cdot (l/x_0)^h$ with $h \leq 1/3$.

Aiming to show that $\epsilon \to 0$ in $l \to 0$, the scaling of the amplitude of the noise $n_0(l)$ is needed. The noisy term must be competing with the non linear convective one in order to produce a non negligible effect, so they must be of the same order at the small scale l: $\mathcal{O}[u \cdot \nabla u] = \mathcal{O}[\nabla \cdot \sqrt{n_0(l)\sigma(u/u_0)}\eta]$. Reasoning in orders of magnitude, since $\mathcal{O}[\nabla] = l^{-1}$ (the gradient-scale is the inverse of the smallest length-scale in the system because ∇ is large for large variations at small scales) and at small scales $\mathcal{O}[u] = \delta u$, the balance is obtained if n_0 is such that $\delta u^2 l^{-1} \sim l^{-1}\sqrt{n_0}$ so $n_0 \sim \delta u^4 \sim U^4 \cdot (l/x_0)^{4h}$. Putting all together, as the macroscopic values u_0 and x_0 are not vanishing, $\epsilon = n_0/u_0^4 \sim (\frac{U}{u_0})^4 (\frac{l}{x_0})^{4h}$ correctly vanishes in the coarse-graining limit $l \to 0$.

Notice that during this scaling analysis the parameter l/x_0 emerged. It essentially corresponds to "quality" of coarse-graining as it is the analogous of $l_{\mathrm{meso}}/l_{\mathrm{macro}}$ that was defined in section 1.2.1, coherently with the MFT literature [5].

In conclusion, it was shown that it is always possible to find a rescaling such that $\epsilon \to 0$ in $l \to 0$ and the equation for the rescaled coarse-grained velocity $u := \overline{u}^l$ is (omitting the $'$ notation):

$$\partial_t u + (u \cdot \nabla)u = -\nabla P + \mathrm{Re}^{-1}\Delta u + \nabla \cdot \sqrt{\epsilon\sigma(u)}\eta, \quad \nabla \cdot u = 0, \quad u(\cdot, 0) = u_0 \tag{27}$$

[33] By definition, η_h is the largest length-scale below which the velocity field is once-differentiable and the scaling can be found by imposing $\mathcal{O}(u \cdot \nabla u) = \mathcal{O}[\nu \Delta u]$ with $\mathcal{O}(u) = \delta u \sim \ell^h$; the Kolmogorov length η_K corresponds to the case $h = 1/3$.

[34] In this scaling argument there is no distinction between $\mathrm{Re}^{-1} \to 0$ and $\nu \to 0$ as $\mathrm{Re} = u_0 x_0/\nu$ with x_0, u_0 finite.

3.1.3 The Leray projector in place of the pressure

The aim from now on will be to build the large deviation function of any observable up to leading order in $\epsilon \to 0$ and the strategy will be analogous to what discussed in section 2. As it will be necessary to take functional derivatives $\frac{\delta}{\delta u(x,t)}$ of the terms appearing in equation 27 and it would be inconvenient to evaluate the functional derivative of the pressure term. It is instead smarter [35] to impose the pressure such that the incompressibility equation $\nabla \cdot u$ is satisfied, thus formally removing the pressure from the stochastic Navier-Stokes equation. This is possible through the Leray projector, which projects onto the space of divergence-free vector fields and reads:

$$\mathcal{P} = \mathbb{1} - \nabla\Delta^{-1}\nabla\cdot \qquad \text{or, in components: } \mathcal{P}^{ij} = \delta^{ij} - \partial_i\Delta^{-1}\partial_j$$

where Δ^{-1} inverse Laplacian operator is defined as $\varphi := \Delta^{-1}\psi$ if $\psi = \Delta\varphi$.

Applying $\mathcal{P}$ to both sides of equation 27 eliminates the pressure term ($\mathcal{P}\nabla p = \nabla p - \nabla\Delta^{-1}\nabla^\top\nabla p = \nabla p - \nabla\Delta^{-1}\Delta p = 0$), yielding:

$$\partial_t u + \mathcal{P}(u\cdot\nabla u) = \mathrm{Re}^{-1}\Delta u + \mathcal{P}[\nabla\cdot\sqrt{\epsilon\sigma}\eta] \quad \text{with } \nabla\cdot u = 0 \tag{28}$$

where it was used that $\mathcal{P}u = u$ because of incompressibility. Now the equation is closed for the velocity.

In order to deal with the presence of the projector, following [35], the calculations are performed on an unconstrained velocity v such that $u = \mathcal{P}v$. Commuting $\mathcal{P}$ with all the differential operators, equation 28 maps to $\mathcal{P}[\partial_t v + (\mathcal{P}(v)\cdot\nabla\mathcal{P}(v)) - \mathrm{Re}^{-1}\Delta v - \nabla\cdot\sqrt{\epsilon\sigma}\eta] = 0$ and so the unrestricted velocity field v follows:

$$\partial_t v + \mathcal{P}(v)\cdot\nabla\mathcal{P}(v) - \mathrm{Re}^{-1}\Delta v - \nabla\cdot\sqrt{\epsilon\sigma}\eta = 0 \tag{29}$$

Why is it necessary to introduce v? An alternative might be to apply Macroscopic Fluctuation Theory to equation 28 directly. However, this would lead to inconsistent instanton equations where the divergence-less condition is not satisfied, making the instanton an unphysical path that lives outside of the Hilber space of the incompressible velocity fields $\mathbb{H}_I := \{\psi : \forall x \in \mathbb{R}^d, \forall t \in [0,t_f]\,\psi(x,t) \in \mathbb{R}^d,\, \nabla\cdot\psi(x,t) = 0\}$. On the other hand, the change of variable $u = \mathcal{P}(v)$ makes equation 29 defined on the unconstrained velocity field space $\mathbb{H}_U := \{\psi : \forall x \in \mathbb{R}^d, \forall t \in [0,t_f]\,\psi(x,t) \in \mathbb{R}^d\}$, which is closed to the operations that will be performed in section 3.2. The physical result for u will be obtained through the projection thus avoiding all incoherences.

3.2 The MFT action and the instanton equations

In this section the MFT action and the instanton equations will be derived for the unrestricted field v governed by equation 29 and then they will be projected back into the divergence-less velocity field space so to recover the results for u. The calculations are performed in the general case of the initial condition drawn from a distribution $v_0 \sim \exp[-F]$ (take $\exp[-F]$ to be a delta function if the initial condition is fixed). Let $A[u]$ the observable of interest, that might depend on the full realization of the velocity. The notation $f[v] := f[u = \mathcal{P}(v)]$ will be used for brevity for any $f[u]$ functional of the divergence-less u. Finally, recall that $v \in \mathbb{H}_U := \{\psi : \forall x \in \mathbb{R}^d, \forall t \in [0,t_f]\,\psi(x,t) \in \mathbb{R}^d\}$ where scalar product is defined as:

$$\langle\psi|\phi\rangle := \sum_{i=1}^{d}\int_0^{t_f} dt \int_{\mathbb{R}^d} d^dx\,\langle\psi|x,t,i\rangle\,\langle x,t,i|\phi\rangle = \sum_{i=1}^{d}\int_0^{t_f} dt \int_{\mathbb{R}^d} d^dx\psi^i(x,t)\phi^i(x,t)$$

Furthermore, if an operator $O^{ij}(x,x',t,t')$ is applied:

$$\langle\psi|O|\phi\rangle := \sum_{i,j=1}^{d}\int_0^{t_f} dt \int_0^{t_f} dt' \int_{\mathbb{R}^d} d^dx \int_{\mathbb{R}^d} d^dx'\,\psi^i(x,t)O^{ij}(x,x',t,t')\phi^j(x',t')$$

For brevity, the following compact notation for the sums and integrals is adopted:
$\sum_{i=1}^{d}\int_0^{t_f} dt \int_{\mathbb{R}^d} d^dx := \int_{i,t,x}$.

3.2.1 MFT action for the unrestricted velocity field v

Let $v[v(0),\eta]$ the solution of equation 29 given the initial condition and the realization of the noise. Under ensemble equivalence explained in section 2.1.1, consider the following canonical partition function and use the "Hamiltonian" approach introduced in section 2.1.2:

$$\begin{aligned}
Z(\lambda) &= \mathbb{E}_{\eta,v(0)}[e^{\epsilon^{-1}\lambda A[v[v(0),\eta]]}] = \int \mathcal{D}v(0) e^{-F[v(0)]} \int \mathcal{D}\eta \mathbb{P}[\eta] e^{\epsilon^{-1}\lambda A[v[v(0),\eta]]} = \\
&= \int \mathcal{D}v(0) e^{-F[v(0)]} \int \mathcal{D}\eta \mathbb{P}[\eta] e^{\epsilon^{-1}\lambda A[v[v(0),\eta]]} \int \mathcal{D}v \delta(v - v[v(0),\eta]]) = \\
&= \int \mathcal{D}v(0) \int \mathcal{D}v\, e^{-F[v(0)]} e^{\epsilon^{-1}\lambda A[v]} \mathbb{E}_\eta \Big[\delta\Big(\partial_t v + \mathcal{P}(v)\cdot\nabla\mathcal{P}(v) - \mathrm{Re}^{-1}\Delta v - \nabla\cdot\sqrt{\epsilon\sigma(v)}\eta \Big) |\det \mathcal{J}[\eta]| \Big]
\end{aligned}$$

where the Jacobian $\mathcal{J}[\eta]$ is arising from the change of variable inside the delta and is analogous to the one of equation 12. There are many reasons for neglecting this Jacobian. First, its effect in the MFT action is subleading in $\epsilon \to 0$, as shown for the Langevin equation case in section 2.1.1; moreover, for situations where the nonlinear part of the dynamics satisfies a Liouville theorem, the Jacobian is in fact only a field-independent factor [20] and it was proven in [31] that the equation obtained from Navier-Stokes equation by discretizing and cutting-off the small length-scales[35] falls in this category. This further motivates the irrelevance of the discretization scheme, which for simplicity was chosen to be Itô.

The Hamiltonian momentum π is introduced by the Fourier representation of Dirac's delta: $\delta(\psi) = \int \mathcal{D}K \exp\{i\langle K|\psi\rangle\} \overset{\pi=-i\epsilon K}{=} \int \mathcal{D}\pi \exp\{-\epsilon^{-1}\langle\pi|\psi\rangle\}$. Incorporating $\mathcal{D}v(0)$ in $\mathcal{D}v$:

$$\begin{aligned}
Z(\lambda) &= \int \mathcal{D}[v,\pi] \exp\Big[-F[v(0)] + \frac{1}{\epsilon}\big\{ \lambda A[v] - \langle \pi | \partial_t v + \mathcal{P}(v)\cdot\nabla\mathcal{P}(v) - \mathrm{Re}^{-1}\Delta v\rangle \big\} \Big] \mathbb{E}_\eta \Big[e^{\sum_a \langle \pi | \partial_a \sqrt{\frac{1}{\epsilon}\sigma(v)}\eta^a\rangle} \Big] = \\
&\quad \text{By parts on } \partial_a \text{ and then using Gaussian integrals on the last term} \\
&= \int \mathcal{D}[v,\pi] \exp\Big[-F[v(0)] + \frac{1}{\epsilon}\Big(\lambda A[v] - \langle \pi | \partial_t v + \mathcal{P}(v)\cdot\nabla\mathcal{P}(v) - \mathrm{Re}^{-1}\Delta v\rangle \Big) + \frac{1}{2\epsilon}\sum_{a,b} \langle \partial_a\pi | \sqrt{\sigma(v)} C_{ab} \sqrt{\sigma(v)} | \partial_b \pi\rangle \Big] \\
&= \int \mathcal{D}[v,\pi]\, e^{-\epsilon^{-1} S_\lambda[v,\pi]}
\end{aligned}$$

where S_λ is the MFT Hamiltonian action and reads:

$$S_\lambda[v,\pi] = \underbrace{\epsilon F[v(0)]}_{\text{I.C.}} \underbrace{-\lambda A[v]}_{\text{Constraint}} + \langle \pi | \underbrace{\partial_t v + \mathcal{P}(v)\cdot\nabla\mathcal{P}(v) - \mathrm{Re}^{-1}\Delta v}_{\text{Deterministic part of 29 (NSE)}} \rangle \underbrace{- \sum_{a,b=1}^{d} \frac{1}{2} \langle \partial_a\pi | \sqrt{\sigma(v)} C_{ab} \sqrt{\sigma(v)} | \partial_b \pi\rangle}_{\text{Effective contribution of the noise}} \tag{30}$$

or, in explicit notation:

$$\begin{aligned}
S_\lambda[v,\pi] &= -\lambda A[v] + \epsilon F[v(0)] + \\
&\quad + \int_{j,y,s} \Big[\pi^j(y,s) \frac{\partial}{\partial s} v^j(y,s) + \sum_k \pi^j(y,s) \mathcal{P}^k(v(y,s)) \frac{\partial}{\partial y_k} \mathcal{P}^j(v(y,s)) - \mathrm{Re}^{-1} \pi^j(y,s) \Delta v^j(y,s) \Big] + \\
&\quad - \frac{1}{2} \int_{a,b,j,y,s,k,z,s'} \Big(\frac{\partial \pi^j(y,s)}{\partial y^a} \Big) \sqrt{\sigma(v(y,s))} C^{jk}_{ab}(y-z, s-s') \sqrt{\sigma(v(z,s'))} \Big(\frac{\partial \pi^k(z,s')}{\partial z^b} \Big)
\end{aligned}$$

The physical interpretation is the same as in the Langevin equation case (section 2.2.1). Notice that in the small ϵ limit the initial condition is subleading, provided it is non singular (e.g. a delta function).

3.2.2 Instanton equations of the stochastic Navier-Stokes equation

As it was argued in section 2.2.1, the action introduces a probability measure over the space of paths. As $\epsilon \to 0$, the measure is dominated by the instanton, namely the path minimizing the action, which is interpreted as the most likely path (v,π) to realize in time t_f the rare event $A = a$ starting from a

[35] Also called Fourier-Galerkin truncation of the Navier- Stokes dynamics.

state drawn randomly form $\exp[-F]$. The cumulant generating function follows from a saddle point approximation on the MFT action (consider the analogue of equation 13). Therefore, the aim of this section is to find the stationary equations of this action. The calculation of the functional derivatives is particularly technical and mechanical, so it will be omitted here and reported in the appendix 5.1 (see the final equations 43, 44, 47, 48, 49). The instanton equations are written below with the contribution of the deterministic part of equation 29 on the left hand side and the effect of the noise and the constraint on the observable on the right hand side:

$$\partial_t v^i(x,t) + \mathcal{P}(v(x,t))\cdot\nabla\mathcal{P}^i(v(x,t)) - \mathrm{Re}^{-1}\Delta v^i(x,t) = \\ -\sum_{a,b}\frac{\partial}{\partial x^a}\Big[\sqrt{\sigma(v(x,t))}\int_{j,y,s} C^{ij}_{ab}(x-y,t-s)\sqrt{\sigma(v(y,s))}\frac{\partial\pi^j(y,s)}{\partial y^b}\Big] \tag{31a}$$

$$\partial_t\pi^i(x,t) + \mathcal{P}^i\Big[\sum_j \mathcal{P}^j(v(x,t))\vec{\nabla}\pi^j(x,t) + \mathcal{P}^j(v(x,t))\partial_j\vec{\pi}(x,t)\Big] + \mathrm{Re}^{-1}\Delta\pi^i(x,t) = -\lambda\frac{\delta A[v]}{\delta v^i(x,t)} + \\ -\mathcal{P}^i\Big[\frac{(\vec{\nabla}_u\sigma(u(x,t)))}{2\sqrt{\sigma(u(x,t))}}\sum_{j,k,a,b}\int_{s,y}\frac{\partial\pi^j(x,t)}{\partial x^a}C^{jk}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^k(y,s)}{\partial y^b}\Big]\Big|_{u=\mathcal{P}(v)} \tag{31b}$$

$$t=0\colon\ \pi(x,0) = -\lambda\frac{\delta A[v]}{\delta v(x,0)} + \epsilon\frac{\delta F[v]}{\delta v(x,0)} \qquad t=t_f\colon\ \pi(x,t_f) = \lambda\frac{\delta A[v]}{\delta v(x,t_f)} \tag{31c}$$

where the notation $\vec{\ }$ is only to indicate which is the vector the projector is applying to when it is not obvious. Projecting the equation for $\partial_t v$, using equation 49, the substitution $u=\mathcal{P}(v)$, recalling the short-hand notation $\sigma(v)=\sigma(\mathcal{P}(v))$, the instanton equations for u,π read:

$$\partial_t u^i(x,t) + \mathcal{P}^i[u(x,t)\cdot\nabla\vec{u}(x,t)] - \mathrm{Re}^{-1}\Delta u^i(x,t) = \\ -\sum_l \mathcal{P}^{il}\Big\{\sum_{a,b}\frac{\partial}{\partial x^a}\Big[\sqrt{\sigma(u(x,t))}\int_{j,y,s} C^{lj}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^j(y,s)}{\partial y^b}\Big]\Big\} \tag{32a}$$

$$\partial_t\pi^i(x,t) + \mathcal{P}^i\Big[u(x,t)\cdot\nabla\vec{\pi}(x,t) + \sum_j u^j(x,t)\vec{\nabla}\pi^j(x,t)\Big] + \mathrm{Re}^{-1}\Delta\pi^i(x,t) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,t)}\Big] + \\ -\mathcal{P}^i\left[\frac{(\vec{\nabla}_u\sigma(u(x,t)))}{2\sqrt{\sigma(u(x,t))}}\sum_{j,k,a,b}\int_{s,y}\frac{\partial\pi^j(x,t)}{\partial x^a}C^{jk}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^k(y,s)}{\partial y^b}\right] \tag{32b}$$

$$t=0:\ \pi^i(x,0) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,0)}\Big] + \mathcal{P}^i\Big[\epsilon\frac{\delta F[u(0)]}{\delta u(x,0)}\Big] \qquad t=t_f:\ \pi^i(x,t_f) = \lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,t_f)}\Big] \tag{32c}$$

This is the main original result of this report. Notice that:

- As a coherence check, in the zero noise case ($\sigma=0$) the incompressible Navier-Stokes equation is recovered. Notice also that applying the divergence to equation 32a gives 0, so the instanton velocity is divergence-less and defines a physically admissible path;
- As it will be even more evident in section 3.3, equation 32a is in a momentum conserving form (notice the divergence $\sum_a \frac{\partial}{\partial x^a}$ appearing in the noise contribution). This is coherent with the physical interpretation that π is an effective stochastic force exerted by the noise to reach the large deviation $A=a$ (section 2.2.1), so it is expected to maintain the properties of the noise;
- The instanton momentum π was already invariant under $\mathcal{P}$ in equation 31b. This invariance holds also for the initial and final conditions (equation 31c) by the use of equation 49;
- The term linked to the initial condition F is subdominant if non singular (if $v(0)$ is fixed, the initial constraint on $\pi(0)$ would be replaced by $v(0)=v_{\mathrm{in}}$, as in [29]);
- Comparing equation 32 with the instanton equations for Langevin equation (equation 15), notice that here the momentum acts on $\partial_t u$ only via the derivative of π ($\partial_b\pi$, see equation 32a). This is another effect of assuming the conservative form of the noise term ($\nabla\cdot\xi$);

- Notice that the the deterministic contribution (left hand side) in equation 32b resembles the Navier Stokes operator $(\partial_t + \mathcal{P}[u\cdot\nabla] - \mathrm{Re}^{-1}\Delta)$ applied to π with inverse time arrow (see section 3.3.2) and with different convective term being[36] $\mathcal{P}\{u\cdot[(\nabla\otimes\pi)+(\nabla\otimes\pi)^\top]\}$ instead of $\mathcal{P}\{u\cdot(\nabla\otimes\pi)\}$. This means that both equations 32a and 32b are as "non-linear" and difficult to deal with as Navier-Stokes equation. As a side comment, notice that the resemblance between the deterministic contributions of equations 32a and 32b is non trivial. For the generic hydrodynamics $\partial_t u(x,t)+\nabla\cdot J(u(x,t)) = \nabla\cdot[\sqrt{\epsilon\sigma(u(x,t))}\eta(x,t)]$ a similar calculation gives the following instanton equation for π:

$$\partial_t\pi(x,t) - [\nabla_u\otimes(\nabla\cdot J(u(x,t))]\pi(x,t) = -\lambda\frac{\delta A[u]}{\delta u(x,t)} + \\ - \frac{\nabla_u\sigma(u(x,t))}{2\sqrt{\sigma(u(x,t))}}\int_{y,s}\mathrm{tr}\Big[(\nabla\otimes\pi(x,t))C(x-y,t-s)(\nabla\otimes\pi(y,s))\Big]$$

  so it is not guaranteed that $[\nabla_u\otimes(\nabla\cdot J(u(x,t))]\pi(x,t)$ has a similar shape to $J(u(x,t))$ (which is the deterministic contribution to the instanton equation of u);

3.2.3 Quasi-potential and Lagrangian phase transitions

The instanton equations 31 can plugged back into the action S_λ (equation 30) to get the cumulant generating function of the observable A:

$$w_A(\lambda) := \epsilon\ln\Big[\langle e^{\frac{\lambda}{\epsilon}A}\rangle\Big] \approx \epsilon\ln\int\mathcal{D}[v,\pi]e^{-\frac{1}{\epsilon}S_\lambda[v,\pi]} \overset{\text{saddle}}{\approx} -S_\lambda[v_I,\pi_I]$$

where v_I,π_I solutions of the instanton equations 31. By plugging equation 32a into the definition of the action[37], the cumulant generating function reads:

$$w_A(\lambda) = -\epsilon F[u_I(0)] + \lambda A[u_I] + \sum_{a,b}\langle\pi_I|\partial_a\sqrt{\sigma(u_I)}C_{ab}\sqrt{\sigma(u_I)}|\partial_b\pi_I\rangle + \frac{1}{2}\sum_{a,b}\langle\partial_a\pi_I|\sqrt{\sigma(u_I)}C_{ab}\sqrt{\sigma(u_I)}|\partial_b\pi_I\rangle \overset{\text{Parts}}{=} \\ = -\epsilon F[u_I(0)] + \lambda A[u_I] - \frac{1}{2}\sum_{a,b}\langle\partial_a\pi_I|\sqrt{\sigma(u_I)}C_{ab}\sqrt{\sigma(u_I)}|\partial_b\pi_I\rangle \tag{33}$$

with u_I,π_I solutions of the instanton equations 32. Formally the large deviation function of the observable $\phi_A(a)$ is $\phi_A(a) = \sup_\lambda\Big[\lambda a - w_A(\lambda)\Big]$ where notice that the dependence of $w_A(\lambda)$ on λ is also hidden in u_I,π_I (see equations 32).

All the observations made in section 2.2.3 about the choice of the initial and final conditions in the instanton equations straightforwardly generalize here. Moreover, the results of the previous sections generalize to an observable $\boldsymbol{A}[u](x)$ that is a vectorial field by considering a vectorial conjugated field and sending $\lambda A \mapsto \int dx\boldsymbol{\lambda}(x)\cdot\boldsymbol{A}[u](x)$. For instance, in order to build the quasi-potential for the velocity field $\phi[u] := -\epsilon\ln\mathrm{P}_{\mathrm{ss}}\big[\{u(x)\}_{x\in\mathbb{R}^d}\big]$, one can fix the initial velocity profile to $u_{\mathrm{steady}}(x)$, which is a steady state of the deterministic dynamics (so that $(u_{\mathrm{steady}}(t),\pi(t)=0)$ is a solution of the instanton equations), choose the observable $A = u(t_f)$ and vary its final profile $u_{\mathrm{f}}(x)$.

No theorem ensures the uniqueness of solutions of the shooting problem for the instanton equations 32; in fact, in general this is not the case because these equations are as non linear as the Navier-Stokes equation and the shooting problem typically has many solutions. This might have far-reaching consequences. Assume for simplicity that the initial value problem of equation 32 admits unique solutions, so that for each (u_0,π_0) initial condition only one instanton path is defined by the instanton dynamics. Nothing prevents that for some u_{f} there exists a set of different paths $\alpha = 1,...,n$ (i.e. different $\pi_I^{(\alpha)}(0)$ initial momenta) solving the instanton equations with initial velocity u_{steady} and final velocity u_{f}. Usually only one of this path will correspond to the global optimum of the action 30, but

[36] The transpose $\top$ is to be interpreted in the matrix sense, i.e. $[(\nabla\otimes\pi(x,t))^\top]_{ij} := \partial_j\pi^i(x,t)$

[37] More precisely, use $\pi_I = \mathcal{P}(\pi_I)$ (see the comments below equation 32) in equation 30, use $\langle\mathcal{P}(\pi_I)|\psi\rangle = \langle\pi_I|\mathcal{P}(\psi)\rangle$ and then substitute $u_I = \mathcal{P}(v_I)$.

it might happen the global optimum jumps from one path to another when crossing a critical manifold of u_f where both these paths are global minimizers manifold. This would result in a discontinuity of the first derivative of the action $S_0[u_I, \pi_I]$ with respect to u_f final condition, which will translate in a discontinuity in the quasi-potential and correspond to a first order phase transition. This naive picture built in the Hamiltonian perspective is formalized in [5], where these transitions are called "Lagrangian phase transitions" and can only happen in out of equilibrium systems.
This might have a relevance in real systems because, changing the point of view, the same final condition u_f might be reached following macroscopically different instantons if the initial condition is changed of a tiny amount around a critical initial condition (the one such that two globally optimal paths exist). This might be seen as another explanation of the spontaneous stochasticity discussed in section 1.1.2 as a microscopic perturbation changes the macroscopic behaviour (in finite time) and, if the system is started at the critical initial condition, it has to choose randomly which of the two globally optimal paths to follow. A future aim would be to check if spontaneous stochasticity can be explained by this instead of the loss of uniqueness of solutions discussed in section 1.1.1.

3.3 The case of an isotropic local correlation matrix

3.3.1 Instanton equations for an isotropic local correlation matrix

The general case treated in section 3.2 is too abstract to be intuitively understandable. In the following sections the noise will be assumed to be delta correlated and isotropic, resulting in the following correlation matrix [1]:

$$\langle \eta^{aj}(x,t)\eta^{bk}(y,s)\rangle = \delta(x-y)\delta(t-s)\left[\delta^{ab}\delta^{jk} + \delta^{ak}\delta^{bj} - \frac{2}{3}\delta^{aj}\delta^{bk}\right] \tag{34}$$

This corresponds to make the easiest choice for the spatial and time dependence and assume that the noise tensor is isotropic, symmetric and traceless for the vectorial part. First, isotropicity is a physical assumption if there are no preferred directions in the system; the most general space-dependent 4-indexes isotropic tensor is [36]:

$$C^{ij}_{ab}(x,y) = C_1(|x-y|)\delta^{ij}\delta^{ab} + C_2(|x-y|)\delta^{ja}\delta^{ib} + C_3(|x-y|)\delta^{ia}\delta^{jb}$$

so under the assumption of delta correlation in time the correlation matrix will be assumed to have this shape. The assumption of symmetry $\eta^{ai} = \eta^{ia}$ is naturally descending from the definition of the sub-grid tensor represented by the noise ($\tau_l^{ij} = \overline{u^i}\,\overline{u^j} - \overline{u^i u^j}$, sections 1.1.4 and 3.1.1) and implies $C^{ij}_{ab} = C^{ib}_{aj} = C^{aj}_{ib} \implies C_1 = C_2$. Finally, the traceless noise tensor ($\sum_i \eta^{ii} = 0$) comes from the fact that any trace can be included in the pressure field by redefining $\tilde{p}(x,t) = p(x,t) - \sum_a \eta^{aa}(x,t)$ and $\tilde{\eta}^{bi} = \eta^{bi} - \frac{1}{d}\sum_a \eta^{aa}$, resulting in $\sum_{ij} C^{ij}_{ij} = 0 \implies C_3 = -\frac{2C_1}{d}$.
For what concerns the spatial dependence, the hypothesis of solenoidality[38] can be proven to entail $C_1''(r := |x-y|) = 0$ for $r > 0$, which, for instance, is satisfied by a delta function.

With this choice for the correlation tensor, equation 27 is a delta correlated stochastic PDE, which is ill-defined from a mathematical point of view. The idea is to always consider a discretization in space and time which regularizes the unphysical delta correlations appearing above, for instance by introducing a wave number cut-off Λ and defining $\delta_\Lambda(\mathbf{x}-\mathbf{y}) = \sum_{\mathbf{k}:|\mathbf{k}|<\Lambda}(2\pi)^{-d}\exp[i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})]$ with Λ^{-1} that should be chosen smaller than the mean free path and the average intermolecular distance. Referring to [1] (supplementary materials section 1) for the technical details, any stochastic PDEs contained in this report should be interpreted in a discretized way. Recall that effect of discretization

[38] A solenoidal noise is defined such that the noisy term is divergence-less: $\sum_{ai}\partial_a\partial_i\sqrt{\sigma(u(x,t))}\eta^{ai}(x,t) := \mathrm{tr}[\nabla\otimes\nabla\sqrt{\sigma(u(x,t))}\eta(x,t)] = 0\ \forall x,t$. This straightforwardly implies $\langle \mathrm{tr}[\nabla\otimes\nabla\sqrt{\sigma(v(x,t))}\eta(x,t)]\,\mathrm{tr}[\nabla\otimes\nabla\sqrt{\sigma(v(y,s))}\eta(y,s)]\rangle = 0$, which gives: $\sum_{abhk}\mathcal{P}^{ih}_x\mathcal{P}^{jk}_y\frac{\partial}{\partial x^a}\frac{\partial}{\partial y^b}\sqrt{\sigma(u(x,t))}C^{hk}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))} = \sum_{ab}\frac{\partial}{\partial x^a}\frac{\partial}{\partial y^b}\sqrt{\sigma(u(x,t))}C^{ij}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}$, where $\mathcal{P}$ is the Leray projector that was introduced in section 3.1.3. In the case of the correlation matrix of equation 34, this specialize to $\sum_{ai}\partial_a\partial_i\eta^{ai}(x,t) = 0\ \forall x,t \implies \sum_{ij}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\langle\eta^{ij}(x,t)\eta^{kl}(y,s)\rangle = 0\ \forall t,s,k,l, x\neq y$, from which a tedious calculation shows $C_1''(r := |x-y|) = 0$.

can be neglected in Macroscopic Fluctuation Theory at first approximation and for simplicity the Itô discretization was chosen (see section 3.1.1).

The calculation to obtain the instanton equations in this case are reported in appendix 5.2. Using equations 51, 53 and 54, the instanton equations 32 rewrite as:

$$\partial_t u^i(x,t) + \mathcal{P}^i[u(x,t)\cdot\nabla\vec{u}(x,t)] - \mathrm{Re}^{-1}\Delta u^i(x,t) = \\ -\mathcal{P}^i\Big\{\nabla\cdot\Big[\sigma(u(x,t))\big[(\nabla\otimes\pi(x,t)) + (\nabla\otimes\pi(x,t))^\top\big]\Big]\Big\} \tag{35a}$$

$$\partial_t \pi^i(x,t) + \mathrm{Re}^{-1}\Delta\pi^i(x,t) + \mathcal{P}^i\Big[u(x,t)^\top\big[(\nabla\otimes\pi(x,t)) + (\nabla\otimes\pi(x,t))^\top\big]\Big] = -\lambda\mathcal{P}^i\Big(\frac{\delta A[u]}{\delta u(x,t)}\Big) + \\ -\mathcal{P}^i\Big\{\frac{1}{2}(\vec{\nabla}_u\sigma(u(x,t)))\Big[\mathrm{tr}\Big\{(\nabla\otimes\pi(x,t))\big[(\nabla\otimes\pi(x,t)) + (\nabla\otimes\pi(x,t))^\top\big]\Big\}\Big]\Big\} \tag{35b}$$

$$t=0:\ \pi^i(x,0) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,0)}\Big] + \mathcal{P}^i\Big[\epsilon\frac{\delta F[u(0)]}{\delta u(x,0)}\Big] \qquad t=t_f:\ \pi^i(x,t_f) = \lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,t_f)}\Big] \tag{35c}$$

where the matrix $(\nabla\otimes\pi(x,t))_{ij} := \partial_i\pi^j(x,t)$ naturally emerged and the transpose $\top$ is to be interpreted in the matrix sense[39], i.e. $[(\nabla\otimes\pi(x,t))^\top]_{ij} := \partial_j\pi^i(x,t)$. Furthermore, observe the emergence of its symmetric counterpart $\big[(\nabla\otimes\pi(x,t)) + (\nabla\otimes\pi(x,t))^\top\big]$ appearing in both equations, so introduce:

$$S_\pi(x,t) := (\nabla\otimes\pi(x,t)) + (\nabla\otimes\pi(x,t))^\top, \quad S_\pi^{ij} = \partial_i\pi^j + \partial_j\pi^i, \quad S_\pi^\top = S_\pi \tag{36}$$

$$\Omega_\pi(x,t) := (\nabla\otimes\pi(x,t)) - (\nabla\otimes\pi(x,t))^\top, \quad \Omega_\pi^{ij} = \partial_i\pi^j - \partial_j\pi^i, \quad \Omega_\pi^\top = -\Omega_\pi \tag{37}$$

$$B_\pi(x,t) := (\nabla\otimes\pi(x,t)) = \frac{1}{2}(S_\pi + \Omega_\pi), \quad B_\pi^{ij} = \partial_i\pi^j \tag{38}$$

Notice that the last term of equation 35b can be rewritten using:

$$\mathrm{tr}\Big\{(\nabla\otimes\pi(x,t))\big[(\nabla\otimes\pi(x,t)) + (\nabla\otimes\pi(x,t))^\top\big]\Big\} = \frac{1}{2}\mathrm{tr}\Big[(S_\pi+\Omega_\pi)S_\pi\Big] = \frac{1}{2}\mathrm{tr}(S_\pi^2) + \frac{1}{2}\mathrm{tr}(S_\pi\Omega_\pi) = \frac{1}{2}\mathrm{tr}(S_\pi^2)$$

where it was used that $\mathrm{tr}(S_\pi\Omega_\pi) = \mathrm{tr}[(S_\pi\Omega_\pi)^\top] = -\mathrm{tr}[\Omega_\pi S_\pi] = -\mathrm{tr}[S_\pi\Omega_\pi] \implies \mathrm{tr}[S_\pi\Omega_\pi] = 0$. Then it is easy to see that the instanton equations 35 rewrite as:

$$\partial_t u^i(x,t) + \mathcal{P}^i[u(x,t)\cdot\nabla\vec{u}(x,t)] - \mathrm{Re}^{-1}\Delta u^i(x,t) = -\mathcal{P}^i\Big\{\nabla\cdot\Big[\sigma(u(x,t))S_\pi(x,t)\Big]\Big\} \tag{39a}$$

$$\partial_t \pi^i(x,t) + \mathcal{P}^i\Big[S_\pi(x,t)u(x,t)\Big] + \frac{\Delta\pi^i(x,t)}{\mathrm{Re}} = -\lambda\mathcal{P}^i\Big(\frac{\delta A[u]}{\delta u(x,t)}\Big) - \mathcal{P}^i\Big\{\frac{1}{4}(\vec{\nabla}_u\sigma(u(x,t)))\mathrm{tr}\big(S_\pi(x,t)^2\big)\Big\} \tag{39b}$$

$$t=0: \quad \pi^i(x,0) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,0)}\Big] + \mathcal{P}^i\Big[\epsilon\frac{\delta F[u(0)]}{\delta u(x,0)}\Big] \qquad t=t_f: \quad \pi^i(x,t_f) = \lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,t_f)}\Big] \tag{39c}$$

First, notice that equation 39a is still momentum conserving and it can be recast in a continuity equation, while equation 39b has not this property. This asymmetry between u and π is coherent with the fact that π is interpretable as stochastic force driving the system to the large deviation and there is no physical reason why its integral over space should be constant over time.
Moreover, the physical meaning of the symmetric tensor S_π is of a strain-rate tensor of the momentum π: it represents the local rate of deformation of fluid elements due to the stochastic force π. The fact that only the symmetric part of the gradient of π appears is linked to isotropicity.

3.3.2 An equilibrium solution to the instanton equations in the case of additive noise

One of the main results of Macroscopic Fluctuation Theory is that, in a stationary state, the spontaneous emergence of a macroscopic fluctuation (large deviation) takes place most likely following a

[39] For instance, in component notation this term reads:
$\Big(\nabla\cdot\Big[\sigma(u(x,t))\big[(\nabla\otimes\pi(x,t)) + (\nabla\otimes\pi(x,t))^\top\big]\Big]\Big)^j := \sum_a\partial_a\Big[\sigma(u(x,t))\big[(\partial_a\pi^j(x,t) + \partial_j\pi^a(x,t))^\top\big]\Big]$.

trajectory (the instanton) which is the time reversal of the relaxation path according to the adjoint hydrodynamics [5, 27]. This was briefly discussed in sections 2.2.2 and 2.3.3. Now if the stationary state is also an equilibrium state, the adjoint hydrodynamics coincides with the hydrodynamics, so at equilibrium the instanton path leading to the rare state will be time reversal of the relaxation path from the rare state to equilibrium following the same hydrodynamical equations.

It is completely non-trivial to know under which condition equation 27 admits an equilibrium state. Therefore, in this section it will be shown that a solution of the instanton equations 39 with constant $\sigma \geq 0$ and $A = u(t_f)$ can be found by assuming that the system admits an equilibrium state. The inspiration comes from [8], where the calculation that it will be performed below is done for the case of a Langevin equation, which admits an equilibrium state when the force is potential and the noise is Gaussian, white and additive.

Assume σ constant, $A = u(t_f)$, fix the initial condition $u(0) = u_0$ and choose the final state $u(t_f) = u_f$. Then the instanton equations 39 rewrite as:

$$\partial_t u^i(x,t) + \mathcal{P}^i[u(x,t)\cdot\nabla\vec{u}(x,t)] - \mathrm{Re}^{-1}\Delta u^i(x,t) = -\sigma\Delta\pi^i(x,t) \tag{40a}$$

$$\partial_t \pi^i(x,t) + \mathrm{Re}^{-1}\Delta\pi^i(x,t) + \mathcal{P}^i\Big[S_\pi(x,t)u(x,t)\Big] = 0 \tag{40b}$$

$$t = 0: \quad u^i(x,0) = u_0^i(x) \qquad t = t_f: \ u^i(x,t_f) = u_f^i(x) \tag{40c}$$

where for equation 40a it was used that $\sigma\mathcal{P}^i\Big\{\nabla\cdot\Big[S_\pi(x,t)\Big]\Big\} = \sigma\mathcal{P}^i\Big\{\Delta\pi(x,t)\Big\} = \sigma\Delta\pi(x,t)$ as the instanton momentum π is naturally such that $\mathcal{P}\pi = \pi$ (see the comments under the instanton equations 32). The x dependence will be omitted and the vectorial notation will be used from now on.

Since the relaxation path u_R, π_R does not need to be activated by the noise (as it is a typical path, [8]), $\pi_R(t) = 0\ \forall t$ and u_R is the solution of the noiseless hydrodynamics (NSE):

$$\partial_t u_R(t) + \mathcal{P}[u_R(t)\cdot\nabla\vec{u}_R(t)] - \mathrm{Re}^{-1}\Delta u_R(t) = 0, \tag{41a}$$

$$u_R(0) = -u_f, \qquad u_R(t_f) = -u_0 \tag{41b}$$

where the initial and final conditions get a further minus as the velocity is odd under time reversal. Considering the result from Macroscopic Fluctuation Theory mentioned above, if the system is at equilibrium the instanton velocity is $u(t) = \Theta[u_R(t)] = -u_R(t_f - t)$, where u_R follows equation 41 and Θ is the time reversal operator. The instanton momentum $\pi(t)$ must be chosen such that the instanton equations 40 are satisfied with the ansatz introduced[40]. By substituting the ansatz in equation 40a and using equation 41a, it is easy to find that π should satisfy:

$$\sigma\Delta\pi(t) = -\frac{2}{\mathrm{Re}}\Delta u_R(t_f - t) = \frac{2}{\mathrm{Re}}\Delta u(t) \tag{42}$$

Case $\sigma > 0$: a simple solution to equation 42 is $\pi(t) = -\frac{2}{\sigma\mathrm{Re}}u_R(t_f - t) = \frac{2}{\sigma\mathrm{Re}}u(t)$ and actually $(u(t), \pi(t)) = (-u_R(t_f - t), -\frac{2}{\sigma\mathrm{Re}}u_R(t_f - t))$ turns out to be a solution of the instanton equations 40. Indeed, the initial and final conditions 40c are matched by equation 41b and equation 40a is trivially satisfied by the choice of π. It remains to check the validity of equation 40b:

$$\partial_t\pi(t) + \mathcal{P}[\sum_j u^j(t)\nabla\pi^j(t) + u(t)\cdot\nabla\pi(t)] + \mathrm{Re}^{-1}\Delta\pi(t) \stackrel{?}{=} 0$$

$$\iff -\frac{2}{\sigma\mathrm{Re}}\partial_t u_R(t_f - t) + \frac{2}{\sigma\mathrm{Re}}\mathcal{P}[\sum_j u_R^j(t_f - t)\nabla u_R^j(t_f - t) + u_R(t_f - t)\cdot\nabla u_R(t_f - t)] - \frac{2}{\sigma\mathrm{Re}^2}\Delta u_R(t_f - t) \stackrel{?}{=} 0$$

$$\stackrel{s=t_f-t}{\iff} \partial_s u_R(s) + \mathcal{P}[u_R(s)\cdot\nabla u_R(s)] - \mathrm{Re}^{-1}\Delta u_R(s) + \mathcal{P}[\sum_j u_R^j(s)\nabla u_R^j(s)] \stackrel{?}{=} 0$$

$$\stackrel{eq.41a}{\iff} \mathcal{P}[\sum_j u_R^j(s)\nabla u_R^j(s)] \stackrel{?}{=} 0$$

[40] π cannot be obtained by the time reversal of the momentum of the relaxation path ($\pi = \Theta[\pi_R]$) because it is not known how the time reversal operator is acting on π; instead, in a Lagrangian perspective $\pi(t)$ can be obtained once $u(t)\ \forall t$ is known by Legendre duality.

which is true in distributional sense: equation 46 with $\pi = u$ implies that $\mathcal{P}[\sum_j u_R^j(s)\nabla u_R^j(s)] = -\mathcal{P}[\sum_j u_R^j(s)\nabla u_R^j(s)]$ and so it is 0.

Case 2 $\sigma = 0$: The case $\sigma = 0$ is the deterministic limit and has not a physical meaning, but it is still interesting from a mathematical point of view. The only way to satisfy equation 42 is $\mathrm{Re}^{-1} = 0$. This is reasonable because, if there is viscosity but there is not a noise giving back the dissipated energy, an arrow of time can clearly be identified (e.g. looking at the derivative of energy in time), so the equilibrium assumption cannot hold. Notice that taking $\sigma, \mathrm{Re}^{-1} = 0$ in the original model (equation 27) leads to inviscid Euler equation, which is clearly time reversible. As a consequence, the ansatz $u(t) = -u_R(t_f - t)$ satisfies equation 40a for all π. So π should be just chosen such that the instanton equation 40b is satisfied. A simple solution is $\pi(t) = \alpha u(t) = -\alpha u_R(t_f - t)$ with $\alpha \in \mathbb{R}$ (again because of the equivalence of 46). Notice that this solution can be recovered from the solution $\pi(t) = \frac{2}{\sigma \mathrm{Re}} u(t)$ found in *case* $\sigma > 0$ by performing the double limit $\sigma \to 0$ and $\mathrm{Re} \to \infty$ with $\alpha = \frac{2}{\sigma \mathrm{Re}}$ constant.

In conclusion, it was proven that a solution of the instanton equations 40 is this "equilibrium instanton" $(u(t), \pi(t)) = (-u_R(t_f - t), -\frac{2}{\sigma \mathrm{Re}} u_R(t_f - t))$ without any restriction of σ, Re. This means that this path is stationary for the MFT action $S_\lambda[u, \pi]$ and it is then a candidate[41] for being the global optimal path. It would be unexpected that the existence of an equilibrium state holds for any choice of the parameters σ, Re. Moreover, a Lagrangian phase transition might be present (this concept was defined in section 3.2.3): there might be a region in σ, Re space in which equilibrium holds and the "equilibrium instanton" is the global optimum of the action, and a region in which a NESS outperforms the "equilibrium instanton" leading to a regime of out of equilibrium turbulence. Analogous considerations on the Lagrangian phase transitions discussed in section 3.2.3 apply here too.

[41] Uniqueness of the solution to the instanton equations is not guaranteed.

4 Conclusions

To summarize the content of this report, section 1.1 proved that turbulent flows do not behave deterministically due to spontaneous stochasticity arising in Navier-Stokes equation at high Reynolds numbers. Phenomenologically, the reason is an amplification of perturbation at small scales. This suggested to keep a vanishing noise affecting Navier-Stokes equation after a coarse-graining that is performed in order to get an effective macroscopic description of the flow in the turbulent regime. The details of the adopted model were discussed in section 3.1.1.
Section 1.2 motivated that turbulence, being out of equilibrium and described by a stochastic hydrodynamics, can be studied by Macroscopic Fluctuation Theory (MFT). This theory, its results and its interpretations were discussed in section 2 by applying it to a Langevin equation. MFT was shown to provide a way to approximate the large deviation function of any observable to leading order in the vanishing noise. It required the integration of the instanton equations, namely the ones for the most likely path leading to the large deviation. The practical example of a multiplicative Langevin equation with an asymmetric double-well potential was discussed in section 2.3.
The application of MFT to the coarse-grained Navier-Stokes equation defined in section 3.1.1 was detailed in a very general case in section 3.2 and in a simpler but still physical case in section 3.3. The possibility of an out of equilibrium phase transition between an equilibrium regime and a non-equilibrium one was discussed in sections 2.2.2 and 3.3.2.

Besides re-deriving known results of MFT applied to Langevin equations from a Hamiltonian perspective in section 2.2.2, the original part of this work is contained in section 3. It was shown explicitly that Macroscopic Fluctuation Theory can formally be applied to a stochastic version of Navier-Stokes equation with momentum-conserving multiplicative noise and, for the first time, the instanton equations 32 were derived. This work is far from answering the question whether Macroscopic Fluctuation Theory can produce an effective statistical model of turbulence and, indeed, many paths remain to be explored.

First of all, a way to get predictions on the statistics of some observables by the use of the instanton equations 32 is needed. There is probably not much to be done analytically on these equations as they are. Some analytical results might be obtained by perturbation theory in $\lambda \to 0$ of the instanton equations or from the study of the Restricted Euler equations that are discussed in appendix 5.3.
From a numerical point of view, the challenge is also quite ambitious as the instanton equations have the same computational difficulty of Navier-Stokes equation, plus it is a shooting problem. Just as the Navier-Stokes equation at high Reynolds numbers, it is currently unfeasible to integrate them directly [17]. A possible numerical alternative is to project them onto logarithmic lattices [10], which makes the integration exponentially less expensive.

A further step would be to infer the shape of the correlation function of the noise considered in section 3.1.1 from experiments and direct numerical simulations of the Navier-Stokes equation. Theoretically, it could be possible to obtain this self-consistently by imposing that the correlations of the sub-grid tensor τ_l evaluated with MFT match with the assumed noise ones (this was briefly discussed in section 3.1.1).

Once predictions of this theory will be available, they can be checked by experiments where rare events are observed [17] in turbulent flows. Would this furnish an approximated but effective model of turbulence, it could be used for weather and climate forecast in place of the empirical and parametric model used nowadays, which could have a positive impact in building resilience against climate change.

A possible field of application of MFT beyond turbulence is climate, as Langevin equations with small noise are not new among conceptual models for this system [4]. Another possibility is to apply this theory to weather, where having early warnings of extreme events would be extremely useful and they can in principle be inferred by the instanton equations; the idea is that, experimentally, the only way a rare atypical event can arise is following the instanton path (up to small fluctuations around it), as it is overwhelmingly more probable than any other.

Bibliography

[1] Dmytro Bandak et al. "Spontaneous Stochasticity Amplifies Even Thermal Noise to the Largest Scales of Turbulence in a Few Eddy Turnover Times". In: *Physical Review Letters* 132.10 (2024), p. 104002. DOI: 10.1103/PhysRevLett.132.104002. URL: https://doi.org/10.1103/PhysRevLett.132.104002.

[2] A.A. Belavin et al. "Pseudoparticle solutions of the Yang-Mills equations". In: *Physics Letters B* 59.1 (1975), pp. 85–87. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(75)90163-X. URL: https://www.sciencedirect.com/science/article/pii/037026937590163X.

[3] Santiago J. Benavides and Dwight Barkley. "Model for transitional turbulence in a planar shear flow". In: *arXiv preprint arXiv:2309.12879* (2023). URL: https://arxiv.org/abs/2309.12879.

[4] Roberto Benzi et al. "A Theory of Stochastic Resonance in Climatic Change". In: *SIAM Journal on Applied Mathematics* 43.3 (1983), pp. 565–578. DOI: 10.1137/0143037.

[5] Lorenzo Bertini et al. "Macroscopic fluctuation theory". In: *Reviews of Modern Physics* 87.2 (2015), pp. 593–636. DOI: 10.1103/RevModPhys.87.593.

[6] Rehab Bitane, Holger Homann, and Jérémie Bec. "Geometry and violent events in turbulent pair dispersion". In: *Journal of Turbulence* 14 (2013), pp. 23–45. DOI: 10.1080/14685248.2013.766747. URL: https://arxiv.org/abs/1209.2296.

[7] Freddy Bouchet, Jason Laurie, and Oleg Zaboronski. "Langevin dynamics, large deviations and instantons for the quasi-geostrophic model and two-dimensional Euler equations". In: *Journal of Statistical Physics* 156 (2014), pp. 1066–1092. DOI: 10.1007/s10955-014-1015-1.

[8] Serena Bradde and Giulio Biroli. "The Generalized Arrhenius Law in Out of Equilibrium Systems". In: *arXiv preprint arXiv:1204.6027* (2012). URL: https://arxiv.org/abs/1204.6027.

[9] Ciro S. Campolina and Alexei A. Mailybaev. "Chaotic Blowup in the 3D Incompressible Euler Equations on a Logarithmic Lattice". In: *Physical Review Letters* 121.6 (2018), p. 064501. DOI: 10.1103/PhysRevLett.121.064501. URL: https://doi.org/10.1103/PhysRevLett.121.064501.

[10] Ciro S. Campolina and Alexei A. Mailybaev. "Fluid dynamics on logarithmic lattices". In: *Nonlinearity* 34.7 (2021), pp. 4684–4715. DOI: 10.1088/1361-6544/abef73. URL: https://doi.org/10.1088/1361-6544/abef73.

[11] Leticia F. Cugliandolo and Vivien Lecomte. "Rules of calculus in the path integral representation of white noise Langevin equations: the Onsager–Machlup approach". In: *Journal of Physics A: Mathematical and Theoretical* 50.34 (2017), p. 345001. DOI: 10.1088/1751-8121/aa7dd6. URL: https://doi.org/10.1088/1751-8121/aa7dd6.

[12] Bernard Derrida. "Non-equilibrium steady states: Fluctuations and large deviations of the density and of the current". In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.07 (2007), P07023. DOI: 10.1088/1742-5468/2007/07/P07023.

[13] Bernard Derrida and Tridib Sadhu. "Large deviations conditioned on large deviations I: Markov chain and Langevin equation". In: *Journal of Statistical Physics* 176.4 (2019), pp. 773–805. DOI: 10.1007/s10955-019-02321-4. URL: https://doi.org/10.1007/s10955-019-02321-4.

[14] Bernard Derrida and Tridib Sadhu. "Large deviations conditioned on large deviations II: Fluctuating hydrodynamics". In: *arXiv preprint arXiv:1904.01170* (2019). URL: https://arxiv.org/abs/1904.01170.

[15] J. R. Dormand and P. J. Prince. "A family of embedded Runge-Kutta formulae". In: *Journal of Computational and Applied Mathematics* 6.1 (1980), pp. 19–26.

[16] Theodore D. Drivas, Alexei A. Mailybaev, and Artem Raibekas. "Statistical determinism in non-Lipschitz dynamical systems". In: *Ergodic Theory and Dynamical Systems* 44.7 (2023), pp. 1856–1884. DOI: 10.1017/etds.2023.74. URL: https://doi.org/10.1017/etds.2023.74.

[17] Bérengère Dubrulle et al. "How many modes are needed to predict climate bifurcations? Lessons from an experiment". In: *Nonlinear Processes in Geophysics* 29.1 (2022), pp. 17–35. DOI: 10.5194/npg-29-17-2022.

[18] Bérengère Dubrulle et al. "Can we predict the weather? New tools for an old problem". In: *Phys. Rev. Fluids* (June 2025). DOI: 10.1103/25dz-vjmt. URL: https://link.aps.org/doi/10.1103/25dz-vjmt.

[19] Gregory Eyink. "Onsager's 'Ideal Turbulence' Theory". In: *Journal of Fluid Mechanics* 988 (2024), P1–P64. DOI: 10.1017/jfm.2024.415. URL: https://doi.org/10.1017/jfm.2024.415.

[20] Gregory L. Eyink. "Turbulence Noise". In: *Journal of Statistical Physics* 83.5-6 (1996), pp. 955–1019. DOI: 10.1007/BF02179551. URL: https://doi.org/10.1007/BF02179551.

[21] Etienne Forest and Ronald D. Ruth. "Fourth-order symplectic integration". In: *Physica D Nonlinear Phenomena* 43.1 (May 1990), pp. 105–117. DOI: 10.1016/0167-2789(90)90019-L.

[22] Uriel Frisch. *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge: Cambridge University Press, 1995. ISBN: 9780521457132.

[23] C. W. Gardiner. *Handbook of stochastic methods for physics, chemistry and the natural sciences*. 2nd ed. Springer Series in Synergetics. Berlin: Springer Berlin, Heidelberg, 1996, pp. xix, 442. ISBN: 978-3540616344.

[24] Pedro L. Garrido. *Notes about the Macroscopic Fluctuating Theory*. https://arxiv.org/abs/2006.08261. Lecture notes. 2020.

[25] Benjamin Gess, Daniel Heydecker, and Zhengyan Wu. "Landau-Lifshitz-Navier-Stokes equations: Large deviations and relationship to the energy equality". In: *arXiv preprint arXiv:2310.02094* (2023).

[26] S. S. Girimaji and C. G. Speziale. "A modified restricted Euler equation for turbulent flows with mean velocity gradients". In: *Physics of Fluids* 7.6 (1995), pp. 1438–1446. DOI: 10.1063/1.868530.

[27] G. Jona-Lasinio. "Review article: Large fluctuations in non-equilibrium physics". In: *Nonlinear Processes in Geophysics* 30.2 (2023), pp. 253–262. DOI: 10.5194/npg-30-253-2023. URL: https://npg.copernicus.org/articles/30/253/2023/.

[28] P. L. Krapivsky and Baruch Meerson. "Fluctuations of current in nonstationary diffusive lattice gases". In: *Phys. Rev. E* 86 (3 Sept. 2012), p. 031106. DOI: 10.1103/PhysRevE.86.031106. URL: https://doi.org/10.1103/PhysRevE.86.031106.

[29] Paul L Krapivsky, Kirone Mallick, and Tridib Sadhu. "Tagged particle in single-file diffusion". In: *arXiv preprint arXiv:1506.00865* (2015). URL: https://arxiv.org/abs/1506.00865.

[30] L. D. Landau and E. M. Lifshitz. *Statistical Physics, Part 2*. 1st English edition. Vol. 9. Course of Theoretical Physics. Sections 88–89. Pergamon Press, 1980. ISBN: 9780080230386.

[31] T. D. Lee. "On some statistical properties of hydrodynamical and magneto-hydrodynamical fields". In: *Quarterly of Applied Mathematics* 10.1 (1952), pp. 69–74. DOI: 10.1090/qam/51081. URL: https://doi.org/10.1090/qam/51081.

[32] Kirone Mallick, Hiroki Moriya, and Tomohiro Sasamoto. "Exact solution of the macroscopic fluctuation theory for the symmetric exclusion process". In: *Physical Review Letters* 129.4 (2022), p. 040601. DOI: 10.1103/PhysRevLett.129.040601. URL: https://doi.org/10.1103/PhysRevLett.129.040601.

[33] Matteo Marsili. *Lecture Notes on Probability and Information Theory*. Available online at https://www.dropbox.com/scl/fi/tt8f71dogql4z6bxevwlt/book.pdf?rlkey=dmzkn7muoe8g1wouvap5flz5s&e=1&dl=0, the full course being at https://sites.google.com/view/courseprobability2020/home. 2023.

[34] T. N. Palmer, A. Döring, and G. Seregin. "The Real Butterfly Effect". In: *Nonlinearity* 27.9 (2014), R123–R141. DOI: 10.1088/0951-7715/27/9/R123.

[35] Timo Schorlepp et al. "Spontaneous Symmetry Breaking for Extreme Vorticity and Strain in the 3D Navier-Stokes Equations". In: *Philosophical Transactions of the Royal Society A* 380.2218 (2022), p. 20210051. DOI: 10.1098/rsta.2021.0051.

[36] William Smyth. "C.5 Fourth-Order Isotropic Tensors". In: *All Things Flow: Fluid Mechanics for the Natural Sciences*. CC BY-NC 4.0 licensed. Oregon State University (LibreTexts), 2019. Chap. Appendix C.5. URL: https://eng.libretexts.org/Bookshelves/Civil_Engineering/Book:_All_Things_Flow_-_Fluid_Mechanics_for_the_Natural_Sciences_(Smyth)/14:_Appendix_C-_Isotropic_Tensors/14.05:_C.5_4th-order_isotropic_tensors.

[37] Andre A. Souza. *Instantons as a Means to Probe Chaotic Attractors*. https://gfd.whoi.edu/wp-content/uploads/sites/18/2018/03/Andre_Souza_x_243644.pdf. GFD Summer Program Report, Woods Hole Oceanographic Institution. 2015.

[38] Simon Thalabard, Jérémie Bec, and Alexei A. Mailybaev. "From the butterfly effect to spontaneous stochasticity in singular shear flows". In: *Communications Physics* 3.1 (2020), p. 122. DOI: 10.1038/s42005-020-0391-6. URL: https://doi.org/10.1038/s42005-020-0391-6.

[39] Hugo Touchette. "The large deviation approach to statistical mechanics". In: *Physics Reports* 478.1–3 (2009), pp. 1–69. DOI: 10.1016/j.physrep.2009.05.002. arXiv: arXiv:0804.0327 [cond-mat.stat-mech].

5 Appendix

5.1 Appendix A: functional derivatives for the instanton equations

In this appendix the functional derivatives of the trickiest terms in the MFT action (equation 30) are evaluated. These are used in section 3.2.2 to get the instanton equations 31. Notice that in all space integrations by parts the boundary terms is zero because of what was discussed in section 3.1.1. Essentially, if the system is confined in a domain of boundary $\partial\Lambda$, the local equilibrium assumption that underlies any hydrodynamical description (see section 1.2.1) implies $\pi(x,t) = 0\ \forall x \in \partial\Lambda\ \forall t$ (the system at the boundaries is constantly in local equilibrium with them and atypical deviations are not allowed [5]).

1. The Legendre term:

This term is the origin of the initial and final condition on π.

$$\frac{\delta \langle \pi | \partial_t v \rangle}{\delta v^i(x,t)} = \frac{\delta}{\delta v^i(x,t)} \Big[\Big(\sum_j \int d^d y\, \pi^j(y,s) v^j(y,s) \Big) \Big|_{s=0}^{s=t_f} - \langle \partial_t \pi | v \rangle \Big] =$$
$$= \pi^i(x,t_f)\mathbb{1}(t=t_f) - \pi^i(x,0)\mathbb{1}(t=0) - \partial_t \pi^i(x,t)\mathbb{1}(t \in (0,t_f)) \tag{43}$$

where the indicator function of a clause $\mathbb{1}(\text{clause}) = 1$ if the clause is satisfied and zero otherwise.

2. Projector term:

Here it will be proven that:

$$\frac{\delta}{\delta v^i(x,t)} \langle \pi | \mathcal{P}(v) \cdot \nabla \mathcal{P}(v) \rangle = -\mathcal{P}^i \Big[(\mathcal{P}(v(x,t)) \cdot \nabla) \vec{\pi}(x,t) + \sum_j \mathcal{P}^j(v(x,t)) \vec{\nabla} \pi^j(x,t) \Big] \tag{44}$$

where the notation $\vec{}$ is introduced to underline the vector which the projector is applied to.

Recall that the Leray projector ($\mathcal{P}^{ij} = \delta^{ij} - \partial_i \Delta^{-1} \partial_j$ and $(\mathcal{P}^{ij})^\dagger = \mathcal{P}^{ji}$) is such that $\nabla \cdot \mathcal{P}(v) = 0\ \forall v$:

$$\frac{\delta}{\delta v^i(x,t)} \langle \pi | \mathcal{P}(v) \cdot \nabla \mathcal{P}(v) \rangle = \frac{\delta}{\delta v^i(x,t)} \langle \pi | \nabla \cdot \big(\mathcal{P}(v) \otimes \mathcal{P}(v) \big) \rangle =$$
$$= \frac{\delta}{\delta v^i(x,t)} \int_{y,s} \sum_{j,k,h,l} \pi^j(y,s) \partial_k \Big[\Big(\mathcal{P}^{kh} v^h(y,s) \Big) \Big(\mathcal{P}^{jl} v^l(y,s) \Big) \Big] \overset{\text{Parts}}{=}$$
$$= -\frac{\delta}{\delta v^i(x,t)} \int_{y,s} \sum_{j,k,h,l} \Big[\Big(\mathcal{P}^{kh} v^h(y,s) \Big) \Big(\mathcal{P}^{jl} v^l(y,s) \Big) \Big] \partial_k \pi^j(y,s) =$$
$$= -\int_{y,s} \sum_{j,k,h,l} \Big(\mathcal{P}^{kh} \frac{\delta v^h(y,s)}{\delta v^i(x,t)} \Big) \Big(\mathcal{P}^{jl} v^l(y,s) \Big) \partial_k \pi^j(y,s) - \int_{y,s} \sum_{j,k,h,l} \Big(\mathcal{P}^{kh} v^h(y,s) \Big) \Big(\mathcal{P}^{jl} \frac{\delta v^l(y,s)}{\delta v^i(x,t)} \Big) \partial_k \pi^j(y,s) \overset{\text{Adjoint}}{=}$$
$$= -\int_{y,s} \sum_{j,k,h,l} \Big(\frac{\delta v^h(y,s)}{\delta v^i(x,t)} \Big) \mathcal{P}^{hk} \Big[\Big(\mathcal{P}^{jl} v^l(y,s) \Big) \partial_k \pi^j(y,s) \Big] - \int_{y,s} \sum_{j,k,h,l} \Big(\frac{\delta v^l(y,s)}{\delta v^i(x,t)} \Big) \mathcal{P}^{lj} \Big[\Big(\mathcal{P}^{kh} v^h(y,s) \Big) \partial_k \pi^j(y,s) \Big]$$

and now use $\frac{\delta v^h(y,s)}{\delta v^i(x,t)} = \delta^{ih} \delta(x-y) \delta(t-s)$:

$$= -\sum_k \Big[\mathcal{P}^{ik} \Big(\sum_{jl} \mathcal{P}^{jl}(v^l(x,t)) \partial_k \pi^j(x,t) \Big) \Big] - \sum_j \Big[\mathcal{P}^{ij} \Big(\sum_{kh} \mathcal{P}^{kh}(v^h(x,t)) \partial_k \pi^j(x,t) \Big) \Big] =$$

passing back to vectorial notation $\sum_b \mathcal{P}^{ab}(v^b) = \mathcal{P}^a(\vec{v})$

$$= -\mathcal{P}^i \Big[\sum_j \mathcal{P}^j(v(x,t)) \vec{\nabla} \pi^j(x,t) + \sum_j \mathcal{P}^j(v(x,t)) \partial_j \vec{\pi}(x,t) \Big]$$

which proves equation 44.

The same term can also be written as:

$$\frac{\delta}{\delta v^i(x,t)} \langle \pi | \mathcal{P}(v) \cdot \nabla \mathcal{P}(v) \rangle = \mathcal{P}^i \big[-(\mathcal{P}(v(x,t)) \cdot \nabla) \vec{\pi}(x,t) + \sum_j \pi^j(x,t) \vec{\nabla} \mathcal{P}^j(v(x,t)) \big] \tag{45}$$

This can be proven either by repeating the calculation above but skipping the first passage, or as a consequence of the following equivalence in distributional sense:

$$\langle\psi|\mathcal{P}\Big[\sum_j u^j\nabla\pi^j\Big]\rangle = -\langle\psi|\mathcal{P}\Big[\sum_j \pi^j\nabla u^j\Big]\rangle \tag{46}$$

where $u=\mathcal{P}(v)\in\mathbb{H}_I:=\{\psi:\psi(x,t)\in\mathbb{R}^d,\nabla\cdot\psi=0\}$ and all functions are evaluated at (x,t). The proof of the equivalence is:

$$\begin{aligned}\Big\langle\varphi\Big|\mathcal{P}\Big[-\sum_j u^j\vec{\nabla}\pi^j\Big]\Big\rangle &= -\sum_j\Big\langle\mathcal{P}(\varphi)\Big|u^j\vec{\nabla}\pi^j\Big\rangle = -\sum_{i,j}\int_{x,t}\mathcal{P}^i(\varphi)u^j\partial_i\pi^j \overset{\text{Parts}}{=}\\ &= \sum_{i,j}\int_{x,t}\pi^j\partial_i\left[\mathcal{P}^i(\varphi)u^j\right] = \sum_{i,j}\int_{x,t}\mathcal{P}^i(\varphi)\pi^j\partial_i u^j = \Big\langle\varphi\Big|\mathcal{P}\Big[\sum_j\pi^j\vec{\nabla}u^j\Big]\Big\rangle\end{aligned}$$

where it was used that the projector is hermitian and that $\mathcal{P}(\varphi)$ is divergence-less by definition of the Leray operator.
In this report the first formulation (equation 44) was always chosen, coherently with [35].

3. π term:
Using the symmetry of the correlation tensor under the transformation $(k,a,y,s)\leftrightarrow(j,b,z,s')$, it is easy to see that:

$$\begin{aligned}\frac{\delta}{\delta\pi^i(x,t)}\frac{1}{2}\sum_{a,b}\langle\partial_a\pi|\sqrt{\sigma(v)}C_{ab}\sqrt{\sigma(v)}|\partial_b\pi\rangle &= \text{Use symmetry } (k,a,y,s)\leftrightarrow(j,b,z,s')\\ &= \sum_{a,b,k,j}\int_{z,s',y,s}\Big[\frac{\delta}{\delta\pi^i(x,t)}\frac{\partial\pi^k(z,s')}{\partial z^a}\Big]\sqrt{\sigma(v(z,s'))}C^{kj}_{ab}(z-y,s'-s)\sqrt{\sigma(v(y,s))}\frac{\partial\pi^j(y,s)}{\partial y^b}\\ &= -\sum_{a,b}\frac{\partial}{\partial x^a}\sqrt{\sigma(v(x,t))}\int_{j,y,s}C^{ij}_{ab}(x-y,t-s)\sqrt{\sigma(v(y,s))}\frac{\partial}{\partial y^b}\pi^j(y,s)\end{aligned} \tag{47}$$

4. σ term:

$$\frac{\delta}{\delta v^i(x,t)}\frac{1}{2}\sum_{a,b}\langle\partial_a\pi|\sqrt{\sigma(v)}C_{ab}\sqrt{\sigma(v)}|\partial_b\pi\rangle = \langle\partial_a\pi|\frac{\delta\sqrt{\sigma(v)}}{\delta v^i(x,t)}C_{ab}\sqrt{\sigma(v)}|\partial_b\pi\rangle$$

where Leibniz rule and the symmetry of the expression[42] were used, in particular $C^{jk}_{ab}(x-y,t-s)=C^{kj}_{ba}(y-x,s-t)$ (symmetry of a well defined correlation matrix, see section 3.1.1).

Let's discuss from a distributional point of view the functional derivative; recall that $\sigma(v):=\sigma(u=\mathcal{P}(v))$, use the matrix representation of the Leray projector $\mathcal{P}^j[v(y,s)]=\sum_l\mathcal{P}^{jl}v^l(y,s)$ with $\mathcal{P}^{jl}=\delta_{jl}-\partial_j\Delta^{-1}\partial_l$, and take a test function $f(y,s)$ belonging to the Hilbert space $\{\psi:\psi(x,t)\in\mathbb{R}\}$

[42] Essentially it is a symmetry due to the hermitian correlation operator and the fact that on both sides of the scalar product there is $\nabla\otimes\pi$. Explicitly:
$\frac{\delta}{\delta v^i(x,t)}\sum_{a,b}\langle\partial_a\pi|\sqrt{\sigma(v)}C_{ab}\sqrt{\sigma(v)}|\partial_b\pi\rangle=\sum_{ab}\langle\partial_a\pi|\frac{\delta\sqrt{\sigma(v)}}{\delta v^i(x,t)}C_{ab}\sqrt{\sigma(v)}|\partial_b\pi\rangle+\langle\partial_a\pi|\sqrt{\sigma(v)}C_{ab}\frac{\delta\sqrt{\sigma(v)}}{\delta v^i(x,t)}|\partial_b\pi\rangle$
Then use $\langle\psi|A|\phi\rangle=\langle\phi|A^\dagger|\psi\rangle$ and $(C^{jk}_{ab}(y-z,s-s'))^\dagger=C^{kj}_{ba}(z-y,s'-s)$. Finally, write explicitly the integrals, rename $(j,a,y,s)\leftrightarrow(k,b,z,s')$ in the second bra-ket and use the symmetry $C^{kj}_{ba}(z-y,s'-s)=C^{jk}_{ab}(y-z,s-s')$ to prove that the second bra-ket is the same as the first. A faster way to see this would be to define the scalar product between matrices $\langle\nabla\otimes\pi|O|\nabla\otimes\pi\rangle:=\sum_{a,b}\langle\partial_a\pi|O_{ab}|\partial_b\pi\rangle$ and use $O=O^\dagger$ with $O=\sqrt{\sigma}C\sqrt{\sigma}$.

(not $\mathbb{R}^d$):

$$\int_{y,s} f(y,s)\frac{\delta\sqrt{\sigma(\mathcal{P}(v(y,s)))}}{\delta v^i(x,t)} = \int_{y,s}\frac{f(y,s)}{2\sqrt{\sigma(\mathcal{P}(v(y,s)))}}\frac{\delta\sigma(\mathcal{P}(v(y,s)))}{\delta v^i(x,t)} \overset{\text{Chain}}{=}$$

$$= \int_{y,s}\frac{f(y,s)}{2\sqrt{\sigma(\mathcal{P}(v(y,s)))}}\sum_h \frac{\partial\sigma(u)}{\partial u^h}\Big|_{u=\mathcal{P}(v(y,s))}\frac{\delta\mathcal{P}^h(v(y,s))}{\delta v^i(x,t)} =$$

Introduce the matrix representation of $\mathcal{P}$

$$= \int_{y,s,h,l}\frac{f(y,s)}{2\sqrt{\sigma(\mathcal{P}(v(y,s)))}}\frac{\partial\sigma(u)}{\partial u^h}\Big|_{u=\mathcal{P}(v(y,s))}\mathcal{P}^{hl}\frac{\delta v^l(y,s)}{\delta v^i(x,t)} = [\text{use } (\mathcal{P}^{hl})^\dagger = \mathcal{P}^{lh}]$$

$$= \int_{y,s,h,l}\mathcal{P}^{lh}\Big[\frac{f(y,s)}{2\sqrt{\sigma(\mathcal{P}(v(y,s)))}}\frac{\partial\sigma(u)}{\partial u^h}\Big|_{u=\mathcal{P}(v(y,s))}\Big]\delta^{li}\delta(x-y)\delta(t-s) =$$

$$= \sum_h \mathcal{P}^{ih}\Big[\frac{f(x,t)}{2\sqrt{\sigma(\mathcal{P}(v(x,t)))}}\frac{\partial\sigma(u)}{\partial u^h}\Big|_{u=\mathcal{P}(v(x,t))}\Big] = \mathcal{P}^i\Big[\frac{f(x,t)}{2\sqrt{\sigma(\mathcal{P}(v(x,t)))}}\vec{\nabla}_u\sigma(u)\Big|_{u=\mathcal{P}(v(x,t)))}\Big]$$

This result applies straightforwardly with $f(y,s) = \int_{a,b,j,k,s',z}(\partial_a\pi^j(y,s))C^{jk}_{ab}(y-z,s-s')\sqrt{\sigma(v(z,s'))}(\partial_b\pi^k(z,s'))$ and so:

$$\frac{\delta}{\delta v^i(x,t)}\frac{1}{2}\sum_{a,b}\langle\partial_a\pi|\sqrt{\sigma(v)}C_{ab}\sqrt{\sigma(v)}|\partial_b\pi\rangle = \sum_{a,b}\langle\partial_a\pi|\frac{\delta\sqrt{\sigma(v)}}{\delta v^i(x,t)}C_{ab}\sqrt{\sigma(v)}|\partial_b\pi\rangle =$$

$$= \mathcal{P}^i\Big[\frac{(\vec{\nabla}_u\sigma(u(x,t))}{2\sqrt{\sigma(u(x,t))}}\sum_{j,k,a,b}\int_{s,y}\frac{\partial\pi^j(x,t)}{\partial x^a}C^{jk}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^k(y,s)}{\partial y^b}\Big]\Big|_{u=\mathcal{P}(v)} \tag{48}$$

5. Observable term:
Recalling that $A[v] := A[u = \mathcal{P}(v)]$:

$$\frac{\delta A[\mathcal{P}(v)]}{\delta v^i(x,t)} = \int_{j,y,s}\frac{\delta A[u]}{\delta u^j(y,s)}\frac{\delta u^j(y,s)}{\delta v^i(x,t)}\Big|_{u=\mathcal{P}(v)} = \int_{j,y,s}\frac{\delta A[u]}{\delta u^j(y,s)}\Big|_{u=\mathcal{P}(v)}\sum_h\frac{\delta\mathcal{P}^{jh}v^h(y,s)}{\delta v^i(x,t)} =$$

$$= \int_{j,y,s,h}\frac{\delta A[u]}{\delta u^j(y,s)}\Big|_{u=\mathcal{P}(v)}\mathcal{P}^{jh}\Big[\frac{\delta v^h(y,s)}{\delta v^i(x,t)}\Big] \overset{(\mathcal{P}^{jh})^\dagger=\mathcal{P}^{hj}}{=} \int_{j,y,s,h}\mathcal{P}^{hj}\Big[\frac{\delta A[u]}{\delta u^j(y,s)}\Big|_{u=\mathcal{P}(v)}\Big]\delta(x-y)\delta(t-s)\delta^{hi} =$$

$$= \sum_j\mathcal{P}^{ij}\Big[\frac{\delta A[u]}{\delta u^j(x,t)}\Big]\Big|_{u=\mathcal{P}(v)} = \mathcal{P}^i\Big[\frac{\delta A[u]}{\delta\vec{u}(x,t)}\Big]\Big|_{u=\mathcal{P}(v)}$$

which in summary means:

$$\frac{\delta A[\mathcal{P}(v)]}{\delta v(x,t)} = \mathcal{P}\Big[\frac{\delta A[u]}{\delta\vec{u}(x,t)}\Big]\Big|_{u=\mathcal{P}(v)} \tag{49}$$

The same applies for $\frac{\delta F}{\delta\vec{v}(x,0)}$.

5.2 Appendix B: specializing to the local isotropic correlation matrix

The calculations reported here are the ones performed to obtain equation 35 from equation 32 by specializing to the correlation matrix of equation 34.

Aiming at the instanton equation for π, let's evaluate:

$$\sum_{j,k,a,b}\int_{s,y}\frac{\partial\pi^j(x,t)}{\partial x^a}C^{jk}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^k(y,s)}{\partial y^b}\overset{(34)}{=}$$

$$=\sum_{j,k,a,b}\Big[\delta^{ab}\delta^{jk}+\delta^{ak}\delta^{bj}-\frac{2}{3}\delta^{aj}\delta^{bk}\Big]\frac{\partial\pi^j(x,t)}{\partial x^a}\sqrt{\sigma(u(x,t))}\frac{\partial\pi^k(x,t)}{\partial x^b}$$

$$=\sqrt{\sigma(u(x,t))}\Bigg[\sum_{a,j}\Big[\frac{\partial\pi^j(x,t)}{\partial x^a}\frac{\partial\pi^j(x,t)}{\partial x^a}\Big]+\sum_{a,j}\Big[\frac{\partial\pi^j(x,t)}{\partial x^a}\frac{\partial\pi^a(x,t)}{\partial x_j}\Big]-\frac{2}{3}\sum_{k,j}\Big[\frac{\partial\pi^j(x,t)}{\partial x_j}\frac{\partial\pi^k(x,t)}{\partial x_k}\Big]\Bigg]=$$

rename k=a in the last term

$$\boxed{=\sqrt{\sigma(u(x,t))}\sum_{a,j}\left[\frac{\partial\pi^j(x,t)}{\partial x^a}\frac{\partial\pi^j(x,t)}{\partial x^a}+\frac{\partial\pi^j(x,t)}{\partial x^a}\frac{\partial\pi^a(x,t)}{\partial x_j}-\frac{2}{3}\frac{\partial\pi^j(x,t)}{\partial x_j}\frac{\partial\pi^a(x,t)}{\partial x^a}\right]}\tag{50}$$

or, in matrix notation

$$=\sqrt{\sigma(u(x,t))}\Big[\mathrm{tr}\big[(\nabla\otimes\pi(x,t))(\nabla\otimes\pi(x,t))^\top\big]+\mathrm{tr}\big[(\nabla\otimes\pi(x,t))(\nabla\otimes\pi(x,t))\big]-\frac{2}{3}(\nabla\cdot\pi(x,t))^2\Big]$$

$$\boxed{=\sqrt{\sigma(u(x,t))}\Big[\mathrm{tr}\Big\{\big(\nabla\otimes\pi(x,t)\big)\big[(\nabla\otimes\pi(x,t))+(\nabla\otimes\pi(x,t))^\top\big]\Big\}-\frac{2}{3}\underbrace{(\nabla\cdot\pi(x,t))^2}_{=0}\Big]}\tag{51}$$

where $(\nabla\otimes\pi(x,t))_{ij}:=\partial_i\pi^j(x,t)$ and the transpose $\top$ is to be interpreted in the matrix sense, i.e. $[(\nabla\otimes\pi(x,t))^\top]_{ij}:=\partial_j\pi^i(x,t)$. Notice that π is divergence-less so the last term vanishes (see comments below equation 32).

Aiming at the instanton equation for u, let's evaluate:

$$\sum_l\mathcal{P}^{il}\Big\{\sum_{a,b}\frac{\partial}{\partial x^a}\Big[\sqrt{\sigma(u(x,t))}\int_{j,y,s}C^{lj}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^j(y,s)}{\partial y^b}\Big]\Big\}=$$

rename indexes:$j\mapsto k\,,\ l\mapsto j$

$$=\sum_j\mathcal{P}^{ij}\Big\{\sum_{a,b}\frac{\partial}{\partial x^a}\Big[\sqrt{\sigma(u(x,t))}\int_{k,y,s}C^{jk}_{ab}(x-y,t-s)\sqrt{\sigma(u(y,s))}\frac{\partial\pi^k(y,s)}{\partial y^b}\Big]\Big\}=$$

$$=\sum_{j,k,a,b}[\delta^{ab}\delta^{jk}+\delta^{ak}\delta^{bj}-\frac{2}{3}\delta^{aj}\delta^{bk}]\mathcal{P}^{ij}\Big\{\frac{\partial}{\partial x^a}\Big[\sigma(u(x,t))\frac{\partial\pi^k(x,t)}{\partial x^b}\Big]\Big\}$$

$$=\sum_{a,j}\mathcal{P}^{ij}\partial_a[\sigma(u(x,t))\partial_a\pi^j(x,t)]+\sum_{a,j}\mathcal{P}^{ij}\partial_a[\sigma(u(x,t))\partial_j\pi^a(x,t)]-\frac{2}{3}\sum_{b,j}\mathcal{P}^{ij}\partial_j[\sigma(u(x,t))\partial_b\pi^b(x,t)]$$

$$\boxed{=\sum_{a,j}\mathcal{P}^{ij}\Big\{\partial_a\big[\sigma(u(x,t))\big(\partial_a\pi^j(x,t)+\partial_j\pi^a(x,t)\big)\big]-\frac{2}{3}\partial_j[\sigma(u(x,t))\partial_a\pi^a(x,t)]\Big\}}\tag{52}$$

or, in matrix form:

$$\boxed{=\mathcal{P}^i\Bigg\{\nabla\cdot\Big[\sigma(u(x,t))\big[(\nabla\otimes\pi(x,t))+(\nabla\otimes\pi(x,t))^\top\big]\Big]-\frac{2}{3}\Big[\nabla[\sigma(u(x,t))\underbrace{\nabla\cdot\pi(x,t)}_{=0}]\Big]\Bigg\}}\tag{53}$$

Again π is divergence-less so the last term vanishes (or also because $\mathcal{P}[\vec{\nabla}f(x)]=0$ for all f scalar).

Notice finally that the following term appearing in the instanton equation for π can also be written with respect to the emerging matrix $\nabla\otimes\pi$:

$$(\sum_j u^j(x,t)\vec{\nabla}\pi^j(x,t)+u^j(x,t)\partial_j\vec{\pi}(x,t))^k=\sum_j u^j(x,t)\partial_k\pi^j(x,t)+u^j(x,t)\partial_j\pi^k(x,t)=$$

$$=\Big\{u^\top(x,t)\big[(\nabla\otimes\pi(x,t))+(\nabla\otimes\pi(x,t))^\top\big]\Big\}^k\tag{54}$$

5.3 Appendix C: the restricted Euler equations in the case of the local isotropic correlation matrix

5.3.1 Reintroducing the pressures in place of the Leray projector

Notice that the Leray projector is present to enforce the incompressibility constraint and it was very useful to avoid taking care of the pressure; however, aiming to get some physics out of these equations, the pressure fields P, Q respectively for u, π can introduce back such that, by definition, they enforce the incompressibility constraint. From equation 39:

$$\partial_t u(x,t) + (u(x,t)\cdot\nabla)u(x,t) - \mathrm{Re}^{-1}\Delta u(x,t) = -\nabla P(x,t) - \nabla\cdot\Big[\sigma(u(x,t))S_\pi(x,t)\Big]$$
$$\partial_t \pi(x,t) + \mathrm{Re}^{-1}\Delta\pi(x,t) = -\nabla Q(x,t) - \lambda\frac{\delta A[u]}{\delta u(x,t)} - S_\pi(x,t)u(x,t) - \frac{1}{4}\mathrm{tr}\big(S_\pi(x,t)^2\big)\big(\nabla_u\sigma(u(x,t))\big)$$

with the pressures defined such as the divergences of the two equations vanish (provided that π, u divergence-less), which means (in Einstein notation for indexes):

$$\partial_i[u^j\partial_j u^i] = -\partial_i\partial_i P - \partial_i\partial_j[\sigma S_\pi^{ij}] \iff \Delta P = -\partial_i[u^j\partial_j u^i] - \partial_i\partial_j[\sigma S_\pi^{ij}]$$
$$0 = -\partial_i\partial_i Q - \lambda\partial_i\frac{\delta A}{\delta u^i} - \partial_i[S_\pi^{ij}u^j] - \frac{1}{4}\partial_i[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma}{\partial u^i}] \iff \Delta Q = -\lambda\partial_i\frac{\delta A}{\delta u^i} - \partial_i[S_\pi^{ij}u^j] - \frac{1}{4}\partial_i[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma}{\partial u^i}]$$

But now let's notice some properties of the tensor S_π due to the incompressibility condition ($\nabla\cdot\pi = 0$):

$$\sum_i \partial_i S_\pi^{ij} = \sum_i \partial_i(\partial_j\pi^i + \partial_i\pi^j) = \sum_i \partial_i\partial_i\pi^j = \Delta\pi^j \tag{55}$$

which also implies $\sum_{ij}\partial_i\partial_j S_\pi^{ij} = 0$. Then, starting from the equation for P:

$$\begin{aligned}\Delta P &= -\partial_i[u^j\partial_j u^i] - \partial_i\partial_j[\sigma S_\pi^{ij}] \overset{\text{Leib.}}{=} \\ &= -(\partial_i u^j)(\partial_j u^i) - u^j\partial_j\partial_i u^i - \Big[S_\pi^{ij}\partial_i\partial_j\sigma + (\partial_i S_\pi^{ij})(\partial_j\sigma) + (\partial_j S_\pi^{ij})(\partial_i\sigma) + \sigma\partial_i\partial_j S_\pi^{ij}\Big] = \\ &\quad \text{use equation 55 and incompressibility, rename indexes when needed} \\ &= -(\partial_i u^j)(\partial_j u^i) - S_\pi^{ij}\partial_i\partial_j\sigma - 2(\Delta\pi^j)(\partial_j\sigma) = -\mathrm{tr}[(\nabla\otimes u)^2] - S_\pi^{ij}\partial_i\partial_j\sigma - 2(\Delta\pi^j)(\partial_j\sigma)\end{aligned}$$

Whereas for Q, noticing that $\partial_i[S_\pi^{ij}u^j] = S_\pi^{ij}(\partial_i u^j) + u^j\partial_i S_\pi^{ij} = \mathrm{tr}[S_\pi(\nabla\otimes u)] + u^j\Delta\pi^j$:

$$\Delta Q = -\lambda\partial_i\frac{\delta A}{\delta u^i} - \partial_i[S_\pi^{ij}u^j] - \frac{1}{4}\partial_i[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma}{\partial u^i}] = -\lambda\partial_i\frac{\delta A}{\delta u^i} - \mathrm{tr}[S_\pi(\nabla\otimes u)] - u^j\Delta\pi^j - \frac{1}{4}\partial_i\Big[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma}{\partial u^i}\Big]$$

The strain tensor for u emerges, as in equation 35 for π, which suggests to define:

$$S_u(x,t) := (\nabla\otimes u(x,t)) + (\nabla\otimes u(x,t))^\top, \quad S_u^{ij} = \partial_i u^j + \partial_j u^i = S_u^{ji}, \tag{56}$$
$$\Omega_u(x,t) := (\nabla\otimes u(x,t)) - (\nabla\otimes u(x,t))^\top, \quad \Omega_u^{ij} = \partial_i u^j - \partial_j u^i = -\Omega_u^{ji}. \tag{57}$$
$$B_u(x,t) := (\nabla\otimes u(x,t)) = \frac{1}{2}(S_u + \Omega_u), \quad B_u^{ij} = \partial_i u^j \tag{58}$$

and notice that $\partial_i\sigma := \frac{\partial\sigma(u(x,t))}{\partial x_i} = \sum_k B_u^{ik}(x,t)\frac{\partial\sigma(u)}{\partial u^k}|_{(x,t)}$ i.e. $\nabla\sigma = B_u\nabla_u\sigma$. Use these definitions to rewrite:

$$\begin{aligned}\Delta P &= -\mathrm{tr}[(\nabla\otimes u)^2] - S_\pi^{ij}\partial_i\partial_j\sigma - 2(\Delta\pi^j)(\partial_j\sigma) = -\mathrm{tr}[B_u^2] - S_\pi^{ij}B_u^{ik}B_u^{jh}\frac{\partial^2\sigma(u)}{\partial u^k\partial u^h} \\ &= -\frac{1}{4}\mathrm{tr}[S_u^2 + \Omega_u^2] - \mathrm{tr}[S_\pi B_u\mathrm{Hess}_u(\sigma)B_u^\top] - 2\Delta\pi\cdot B_u\nabla_u\sigma = \\ &\quad \text{Use that } \mathrm{tr}(SA)\text{=0 for all matrices } S, A \text{ symmetric and antisymmetric} \\ &= -\frac{1}{4}\mathrm{tr}\Big[S_u^2 + \Omega_u^2 + S_\pi(S_u\mathrm{Hess}_u(\sigma)S_u - \Omega_u\mathrm{Hess}_u(\sigma)\Omega_u)\Big] - 2\Delta\pi\cdot B_u\nabla_u\sigma\end{aligned}$$

By plugging B_u definition into the equation for ΔQ above, on gets:

$$\Delta Q = -\lambda\partial_i\frac{\delta A}{\delta u^i} - \mathrm{tr}[S_\pi(\nabla\otimes u)] - u^j\Delta\pi^j - \frac{1}{4}\partial_i\Big[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma}{\partial u^i}\Big] =$$
$$= -\lambda\nabla\cdot\frac{\delta A[u]}{\delta u(x,t)} - \frac{1}{2}\mathrm{tr}[S_\pi(x,t)S_u(x,t)] - u(x,t)\cdot\Delta\pi(x,t) - \frac{1}{4}\nabla\cdot\Big[\mathrm{tr}(S_\pi^2)\nabla_u\sigma(u(x,t))\Big]$$

As a side note, observe that the properties $\nabla\cdot S_\pi = \Delta\pi$ and $S_\pi = S_\pi^\top$ imply that:

$$\nabla\cdot\Big[\sigma(u(x,t))S_\pi(x,t)\Big] = S_\pi\nabla\sigma + \sigma\Delta\pi$$

This property was never used because it is more meaningful to keep evident that the effect of the noise on u is momentum-conservative.

The conclusion is that the instanton equations rewrite as:

$$\partial_t u(x,t) + (u(x,t)\cdot\nabla)u(x,t) - \mathrm{Re}^{-1}\Delta u(x,t) + \nabla P(x,t) = -\nabla\cdot\Big[\sigma(u(x,t))S_\pi(x,t)\Big] \tag{59a}$$

$$\partial_t\pi(x,t) + \mathrm{Re}^{-1}\Delta\pi(x,t) + S_\pi(x,t)u(x,t) + \nabla Q(x,t) = -\lambda\frac{\delta A[u]}{\delta u(x,t)} - \frac{1}{4}\mathrm{tr}\big(S_\pi(x,t)^2\big)\big(\nabla_u\sigma(u(x,t))\big) \tag{59b}$$

$$\Delta P(x,t) = \Big\{-\frac{1}{4}\mathrm{tr}\Big[S_u^2 + \Omega_u^2 + S_\pi\big[S_u\mathrm{Hess}_u(\sigma(u))S_u - \Omega_u\mathrm{Hess}_u(\sigma(u))\Omega_u\big]\Big] - 2\Delta\pi\cdot B_u\nabla_u\sigma(u)\Big\}\Big|_{(x,t)} \tag{59c}$$

$$\Delta Q(x,t) = -\lambda\nabla\cdot\frac{\delta A[u]}{\delta u(x,t)} - \frac{1}{2}\mathrm{tr}[S_\pi(x,t)S_u(x,t)] - u(x,t)\cdot\Delta\pi(x,t) - \frac{1}{4}\nabla\cdot\Big[\mathrm{tr}(S_\pi^2)\nabla_u\sigma(u(x,t))\Big] \tag{59d}$$

$$\text{t=0:}\quad \pi^i(x,0) = -\lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,0)}\Big] + \mathcal{P}^i\Big[\epsilon\frac{\delta F[u(0)]}{\delta u(x,0)}\Big] \tag{59e}$$

$$\text{t=T:}\quad \pi^i(x,T) = \lambda\mathcal{P}^i\Big[\frac{\delta A[u]}{\delta u(x,T)}\Big] \tag{59f}$$

The initial and final conditions are exactly the same as before and, in order to remove the Leray projector, just consider a divergence-less initial condition.

5.3.2 Restricted Euler equation for the stress tensors $\partial_i u^j$ and $\partial_i\pi^j$

In fluid dynamics, the restricted Euler equation is an approximated equation for the velocity gradient $B_u = \nabla\otimes u$. It is obtained from Navier-Stokes equation assuming the inviscid limit and that the pressure Hessian (the second derivative of the pressure) is isotropic and proportional to the trace of the velocity gradient $(\nabla\otimes u)$ squared. This is a strong simplification that connects the pressure field to the strain (via the velocity gradient tensor) [26]. The resulting equation is:

$$\frac{dB_u^{ij}}{dt} + \sum_k B_u^{ik}B_u^{kj} = \frac{1}{3}\,\mathrm{tr}(B_u^2)\,\delta^{ij}$$

which is very simple and prone to analytical approaches, but still preserves key features of the inviscid 3d Euler equation like finite time blows up [26]. The aim of this section is to apply this simplification to equation 59 to see if it provides with analytical results, but it will be clear soon that this is not the case.

Let $B_u = \nabla\otimes u$ and $B_\pi = \nabla\otimes\pi$, let's try to obtain the analogous of restricted Euler equation for the instanton equations 59. Take the ∂_i of the j-component of the equations 59a and 59b:

$$\partial_t B_u^{ij} + \sum_k\partial_i(u^k\partial_k u^j) - \mathrm{Re}^{-1}\Delta B_u^{ij} + \partial_i\partial_j P = -\sum_k\partial_i\partial_k(\sigma(u)S_\pi^{jk})$$
$$\partial_t B_\pi^{ij} + \mathrm{Re}^{-1}\Delta B_\pi^{ij} + \sum_k\partial_i\big[S_\pi^{jk}u^k\big] + \partial_i\partial_j Q = -\lambda\partial_i\frac{\delta A}{\delta u^j} - \frac{1}{4}\partial_i\Big[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma(u)}{\partial u^j}\Big]$$

Use Leibniz rule on the terms: $\sum_k\partial_i(u^k\partial_k u^j) = (B_u^2)^{ij} + u\cdot\nabla B_u^{ij}$ and $\sum_k\partial_i\big[S_\pi^{jk}u^k\big] = (B_uS_\pi)^{ij} + \sum_k u_k\partial_i(B_\pi^{jk} + B_\pi^{kj})$. The latter can be rewritten using $\partial_i B_\pi^{kj} = \partial_i\partial_k\pi^j = \partial_k B_\pi^{ij}$. Eventually the

equations for B_u and B_π can be written as:

$$\begin{cases} \partial_t B_u^{ij} + u\cdot\nabla B_u^{ij} + ((B_u)^2)^{ij} - \mathrm{Re}^{-1}\Delta B_u^{ij} + \partial_i\partial_j P = -\sum_k \partial_i\partial_k(\sigma(u)S_\pi^{jk}) \\ \partial_t B_\pi^{ij} + u\cdot\nabla B_\pi^{ij} + \mathrm{Re}^{-1}\Delta B_\pi^{ij} + \partial_i\partial_j Q + (B_u S_\pi)^{ij} + \sum_k u^k\partial_i B_\pi^{jk} = \\ \qquad -\lambda\,\partial_i\frac{\delta A}{\delta u^j} - \frac{1}{4}\partial_i\left[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma(u)}{\partial u^j}\right] \end{cases} \tag{60}$$

or equivalently, acting on the coloured terms:

$$\begin{cases} \partial_t B_u^{ij} + u\cdot\nabla B_u^{ij} + ((B_u)^2)^{ij} - \mathrm{Re}^{-1}\Delta B_u^{ij} + \partial_i\partial_j P = \\ \qquad -\sum_{k,h}\partial_i\left[S_\pi^{jk}B_u^{kh}\frac{\partial\sigma(u)}{\partial u^h}\right] - \sigma\Delta B_\pi^{ij} - (\sum_k B_u^{ik}\frac{\partial\sigma(u)}{\partial u^k})(\Delta\pi^j) \\ \partial_t B_\pi^{ij} + \mathrm{Re}^{-1}\Delta B_\pi^{ij} + \partial_i\partial_j Q + (B_u S_\pi)^{ij} + \sum_k u^k\partial_i S_\pi^{jk} = -\lambda\,\partial_i\frac{\delta A}{\delta u^j} - \frac{1}{4}\partial_i\left[\mathrm{tr}(S_\pi^2)\frac{\partial\sigma(u)}{\partial u^j}\right] \end{cases} \tag{61}$$

In order to get equations that can provide analytical results, it useful to take the trace of equation 60 because $\mathrm{tr}[B_u] = \mathrm{tr}[B_\pi] = 0$ by incompressibility ($\nabla\cdot u = \nabla\cdot\pi = 0$). Notice that by Leibniz rule $\sum_{ik}\partial_i\partial_k(\sigma S_\pi^{ik}) = \nabla\cdot(S_\pi\nabla\sigma) + (\nabla\cdot S_\pi)\cdot(\nabla\sigma)$ and the chain rule $\nabla\sigma = B_u\nabla_u\sigma$. Observe also that:

$$\sum_{i,k} u^k\partial_i B_\pi^{ik} \overset{\nabla\cdot\pi=0}{=} \sum_{i,k} u^k\partial_i S_\pi^{ik} \overset{55}{=} u\cdot\Delta\pi$$

Eventually, the equations for the traces read:

$$\boxed{\begin{cases} \mathrm{tr}[B_u^2] + \mathrm{tr}[\mathrm{Hess}(P)] = -\nabla\cdot\left[S_\pi B_u\nabla_u\sigma(u)\right] - (B_u\nabla_u\sigma(u))\cdot(\nabla\cdot S_\pi) \\ \frac{1}{2}\mathrm{tr}[S_u S_\pi] + \mathrm{tr}[\mathrm{Hess}(Q)] + \sum_{i,k} u^k\partial_i S_\pi^{ik} = -\lambda\nabla\cdot\frac{\delta A[u]}{\delta u} - \frac{1}{4}\nabla\cdot[\mathrm{tr}(S_\pi^2)\nabla_u\sigma] \end{cases}} \tag{62}$$

These equations show the following features:

- equation 62 only depend on S_π i.e. the symmetric part of B_π, whereas the dependence on Ω_u (the antisymmetric part of B_u) is unavoidable ($\mathrm{tr}[B_u^2] = \frac{1}{4}\mathrm{tr}[S_u^2 + \Omega_u^2]$)
- The
- Only derivatives of the noise come into place; thus, if $\sigma\in\mathbb{R}$ constant, $\forall\sigma$ the equations are the same (this is to be expected because a constant σ can be reabsorbed in ϵ).
- If $\sigma\in\mathbb{R}$ constant, the first equation is the same that is obtained in the standard NSE

Now the idea is to apply the Restricted Euler ansatz and hope to be able to achieve analytical results, at least in the simple case of σ constant and $A = A[u(t_f)]$. One would send $\mathrm{Re}\to 0$ in equation 60 and, inspired by equation 62 would impose the ansatz:

$$\mathrm{Hess}(P) = -\frac{1}{d}\mathrm{tr}(B_u^2)\mathbb{1}_d \qquad\qquad \mathrm{Hess}(Q) = -\frac{1}{d}\left[\frac{1}{2}\mathrm{tr}[S_u S_\pi] + \sum_{ik} u^k\partial_i S_\pi^{ik}\right]\mathbb{1}_d$$

inside equation 60 and obtain:

$$\begin{cases} \partial_t B_u^{ij} + u\cdot\nabla B_u^{ij} - \mathrm{Re}^{-1}\Delta B_u^{ij} + \left[(B_u^2)^{ij} - \frac{1}{d}\mathrm{tr}(B_u^2)\delta^{ij}\right] = -\sigma\Delta B_\pi^{ij} \\ \partial_t B_\pi^{ij} + u\cdot\nabla B_\pi^{ij} + \mathrm{Re}^{-1}\Delta B_\pi^{ij} + \left[(B_u S_\pi)^{ij} - \frac{1}{d}\mathrm{tr}(B_u S_\pi)\delta^{ij}\right] + \left[\sum_k u^k\partial_i B_\pi^{jk} - \frac{1}{d}\sum_{k,l} u^k\partial_l B_\pi^{lk}\right] = 0 \end{cases} \tag{63}$$

However, even in the $\mathrm{Re}^{-1}\to 0$ limit, the term $\sum_{ik} u^k\partial_i S_\pi^{ik}$ makes difficult to go on with analytical derivation for the moment. Nevertheless, equations 63 are simpler than the instanton equations 39 and it remains a promising way to achieve some analytical results.