# Link between stochastic grid perturbation and location uncertainty framework

Simon Clement, Eric Blayo, Laurent Debreu, Jean-Michel Brankart, Pierre Brasseur, Long Li, Etienne Mémin

# Link between stochastic grid perturbation and location uncertainty framework

**S. Clément[1], E. Blayo[1], L. Debreu[1], J-M. Brankart[2], P. Brasseur[2], L. Li[3], E. Mémin[3]**

[1]Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, Grenoble, France
[2]Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, Grenoble, France
[3]Univ. Rennes, Inria, IRMAR, Rennes, France

**Key Points:**

- A rapidly moving grid can be a valid implementation of the Location Uncertainty (LU) framework in the absence of stochastic forcing
- Correlated noise in time and a compensating advection term are introduced to preserve both LU properties and structure of the original grid
- A statistical analysis highlights the importance of the compensation term in the implementation of the method

Corresponding author: Eric Blayo, `eric.blayo@univ-grenoble-alpes.fr`

**Abstract**

This paper investigates the relationship between a Stochastic Grid Perturbation (SGP) and Location Uncertainty (LU) in the context of ocean modeling. The LU formulation, which introduces random velocity fluctuations, has shown efficacy in organizing large-scale flow and replicating long-term statistical characteristics. SGP was created as a simpler approach which perturbs the computational grid for ensemble members, aiming to simulate small uncertainties in high-resolution predictability studies. We aim to clarify the link between SGP and LU. After introducing the LU formalism, we derive the SGP method and discuss its connection to LU. Correlated noise in time is introduced in the SGP method to preserve the structure of the original grid. A compensating advection term is shown to preserve LU properties despite the latter correlated noise. Numerical experiments on a 3-layer Quasi-Geostrophic model compare various SGP implementations with an explicit LU implementation, highlighting the importance of the compensating advection term to achieve strict equivalence.

## Plain Language Summary

This paper examines two subgrid stochastic parameterizations to improve ocean and climate predictions: Stochastic Grid Perturbation (SGP) and Location Uncertainty (LU).

LU is a general framework for deriving stochastic representations of fluid dynamics. It decomposes the Lagrangian velocity into a smooth-in-time resolved component and an unresolved random component that is smooth in space but white in time. The LU dynamics enables the representation of large-scale patterns while preserving important physical properties, such as energy balance. It also incorporates the statistical effects associated with the inhomogeneity of unresolved small-scale dynamics.

In a more empirical way, SGP slightly shifts the positions of grid points in numerical models to study the impact of small uncertainties in ocean forecasts, originally used for high-resolution studies of specific regions.

We explore the relationship between SGP and LU, finding that SGP can be adjusted to work similarly to LU. Through computer simulations, we validate our findings and examine some consequences of the possible discrepancies between the theoretical LU method and the practical SGP method. Understanding the link between SGP and LU can lead to better description of uncertainties in ocean models, and thus to improved forecasts. Our goal is to make advanced techniques like LU more accessible for broader scientific use.

## 1 Introduction

The need for uncertainty quantification in designing probabilistic simulations, constructing plausible scenarios in climate modeling, and developing efficient ensemble methods for data assimilation supports the adoption of stochastic parameterization in geophysical flow models. Starting with classical subgrid modeling, parameterizations with energy backscattering ability and stochastic perturbation have long been proposed (Leith, 1990; Mason & Thomson, 1992). These techniques, notably based on Markovian quasi-Gaussian closures (Kraichnan, 1959) and expressed in the spectral domain, have been successfully explored for the theoretical study of turbulence (Lesieur & Métais, 1996). These ideas of stochastic perturbations have been further developed in climate modeling (see the reviews Berner et al. (2017); Gottwald et al. (2017); Kitsios et al. (2023); O'Kane et al. (2023) and references therein, as well as the seminal work of Hasselmann (1976)). Most of the models studied in these works, however, are built on empirical grounds and lack generality. Unlike deterministic models, they are not derived through a general methodology that respects classical physical invariants. They are conceived by considering additional stochastic forcing, perturbation of the parameters, or semi empirical fast-slow scale decomposition. As a

result, they may face issues such as uncontrolled variance growth if not compensated for by additional dissipation terms. Additionally, for very low noise levels, they may deviate strongly from the associated underlying deterministic system (Chapron et al., 2018).

To avoid such inconsistency, ensure good physical and mathematical properties, and guarantee numerical stability, it is crucial to derive these stochastic representations using a rigorous framework. Along this path, two companion methodologies, based on stochastic transport and time decorrelation assumption, have been introduced by Mémin (2014) and Holm (2015), providing rigorously justified approaches to defining stochastic large-scale flow representations (Crisan et al., 2019; Debussche et al., 2023), conserving energy and circulation, respectively. The first method, which is energy-preserving and referred to as modeling under Location Uncertainty (LU), is a Newtonian framework, while the second is a variational framework. As in the deterministic setting, these two approaches are complementary. In this work, we will focus exclusively on the first methodology.

LU has proven to be versatile in developing various models, ranging from reduced-order models (Chapron et al., 2018) and quasi-geostrophic models (Li et al., 2023a) to wave models (Mémin et al., 2024) and stochastic primitive hydrostatic models (Tucciarone et al., 2023). These models have been shown to be efficient in organizing large-scale flow (Bauer et al., 2020a) and replicating long-term statistical characteristics (Bauer et al., 2020b; Li et al., 2023a) within barotropic and baroclinic quasi-geostrophic models. Due to its construction, based on a stochastic representation of the Reynolds transport theorem, LU models exhibit the same theoretical energy conservation properties as the corresponding deterministic models (e.g., energy or tracer moments) (Brecht et al., 2021; Li et al., 2023a; Resseguier et al., 2017a). They also maintain key theoretical convergence properties (Debussche et al., 2023; Lang et al., 2023), when expressed and understood in a stochastic sense. These desirable theoretical results for LU models stem from an intrinsic balance between variance production and dissipation. The formulation naturally incorporates in its core (through stochastic transport) a direct expression of the fluctuation-dissipation theorem, maintaining theoretically a strict equilibrium between the energy introduced by the noise and the energy dissipated by the system (Bauer et al., 2020a; Resseguier et al., 2017a). Moreover, the convergence to deterministic equations as noise vanishes (Chapron et al., 2018; Debussche et al., 2023) ensures physical consistency in large-scale representation. It is worth noting that such convergence is not trivial, especially for non-Gaussian multiplicative noise; ad hoc noise, at the vanishing limit (i.e., for very small, but non zero, noise), may exhibit different attractors or velocity increment probability distributions (see for instance Chapron et al. (2018) for an example in the Lorenz 63 system). It is desirable for the system to behave consistently with the underlying deterministic dynamics when noise is minimal.

However, even though stochastic equations are increasingly recommended for subgrid parameterizations and ensemble simulations, the perceived difficulty of understanding and implementing them remains a barrier to widespread adoption.

Using another point of view, Leroux et al. (2022) propose a method called here "Stochastic Grid Perturbation" (SGP) to control the covariance of their ensemble by rapidly perturbing the computational grid for each ensemble member. The aim is to simulate a small uncertainty in the numerical model in order to study its impact as part of a very high-resolution predictability study of the Western Mediterranean. This method resembles the Location Uncertainty formalism, where the position of the points fluctuate randomly.

In a data assimilation context, Zhen et al. (2023) replaced the inflation step by the application of an uncorrelated-in-time transport map (i.e. the state variables are modified along a displacement in the underlying coordinates). They formally find links to the LU framework and to the stochastic advection by Lie transport proposed by Holm (2015). Since remapping the physical space is analogous to perturbing the grid points, this may indicate a connection between "Stochastic Grid Perturbation" and the LU approach.

In this paper, we present a first attempt to clarify this link between SGP and LU and discuss the benefits of such Lagrangian parameterization. The LU formalism is first introduced in Section 2. The derivation of the SGP method is then discussed in Section 3 with a link to the LU formalism. Numerical experiments are carried out to compare several SGP implementations with an explicit LU implementation in Section 4. Finally, the link between LU and SGP is summarized in Section 5 together with a discussion on the application scenarios of SGP.

## 2 Equations of fluids under Location Uncertainty

This section presents the main points of the derivation of the Location Uncertainty framework. A more explicit derivation can be found in Bauer et al. (2020a) and Resseguier et al. (2017a). The notations and properties are presented in a concise way in §2.1, and §2.2 details the derivation which will be linked to the Stochastic Grid Perturbation method.

### 2.1 Location Uncertainty formalism

The central hypothesis of the LU framework is the identification of two scales of motion: similarly to a Reynolds averaging, the evolution of the position of a particle $\mathbf{X}_t$ has a large-scale contribution which is smooth and explicitly resolved and a small-scale unresolved component that is assumed uncorrelated in time at the characteristic time scale of the resolved component, and referred to as "noise":

$$d\mathbf{X}_t = \underbrace{\mathbf{v}(\mathbf{X}_t,t)dt}_{\text{Large scale motion}} + \underbrace{\sigma(\mathbf{X}_t,t)d\mathbf{B}_t}_{\text{Small scale uncertainty}} \tag{1}$$

The cylindrical Brownian motion $\mathbf{B}_t(\mathbf{x})$ is uncorrelated in time but correlated in space. Notably, the idealized assumption of temporal decorrelation makes it possible, by means of stochastic calculus, to explicitly take into account the correlation terms between the resolved and unresolved components, which are usually added afterward (e.g. large-scale diffusion, bolus velocity, modified advection...). In particular, this allows to explain in a different way certain models proposed in a fully deterministic framework (see for instance the relations with the Craik-Leibovich system for Langmuir turbulence (Bauer et al., 2020a)). Note that the resolved velocity component is also a stochastic process – i.e. unlike the Reynolds average decomposition it is not an expectation.

The absence of time correlation can be seen as the assumption that the characteristic time scale of turbulence is much smaller than the time step of the numerical model. This simplification of the true continuum of spatial and temporal scales of turbulence assumes that turbulent fluctuations rapidly decorrelate within each time step, allowing them to be treated as independent from one step to the next. In this sense LU can be seen as an expression of the dynamics in the temporal decorrelation limit of the small scales (Debussche & Mémin, 2024). The spatial correlation of the noise is enforced through the operator $\sigma(\mathbf{X}_t,t)$. The latter is a Hilbert-Schmidt integral operator and can thus be defined through a kernel $\breve{\sigma}$:

$$\sigma(\mathbf{X}_t,t)d\mathbf{B}_t \triangleq \int_\Omega \breve{\sigma}(\mathbf{X}_t,\mathbf{y},t)d\mathbf{B}_t(\mathbf{y})d\mathbf{y}.$$

At this point, we need to make several comments on our notations:

- bold letters indicate that the values are in $\mathbb{R}^2$. $\mathbf{B}_t$ has hence two components and the kernel $\breve{\sigma}$ is a $2\times 2$ matrix.
- The computational domain $\Omega$ is a bounded domain of $\mathbb{R}^2$.
- The differential notation used in (1) actually represents the equality between the integrated terms $\mathbf{X}_t-\mathbf{X}_0=\int_0^t \mathbf{v}dt+\int_0^t \sigma d\mathbf{B}_t$ where the last integral is a Itô stochastic integral. The Itô integral is a martingale of null ensemble mean, which is why it is preferred to the Stratonovich integral in most settings based on LU. Despite the

martingale property, the noise features a non-zero mean component through the introduction of a Girsanov drift (see Li et al., 2023b): it consists in a change of probability measure and resembles a smooth transport term.

Finally, if $q$ is a quantity such that $dq = 0$ (i.e. $q(\mathbf{X}_{t+\delta t}, t+\delta t) = q(\mathbf{X}_t, t)$) one can use the stochastic calculus tools to obtain the definition of the *stochastic transport operator*:

$$\mathbb{D}_t q \triangleq d_t q + (\mathbf{v}^\star dt + \sigma d\mathbf{B}_t) \cdot \nabla q - \frac{1}{2}\nabla \cdot (\mathbf{a}\nabla q)dt = 0. \tag{2}$$

The variance $\mathbf{a}$ of the noise (or one-point-one-time correlation matrix) is defined by $\mathbf{a}(\mathbf{x}, t) \triangleq \int_\Omega \breve{\sigma}(\mathbf{x}, \mathbf{y}, t)\breve{\sigma}^T(\mathbf{x}, \mathbf{y}, t)d\mathbf{y}$. The effective velocity $\mathbf{v}^\star$ is defined by $\mathbf{v}^\star \triangleq \mathbf{v} - \frac{1}{2}\nabla \cdot \mathbf{a} + \sigma^T(\nabla \cdot \sigma)$. When expressed in Stratonovich notation, the diffusion term is implicit and the transport of a passive scalar quantity corresponds to the model proposed by Kraichnan (1994), which has been extensively used for theoretical studies (Falkovich et al., 2001; Gawedzky & Kupiainen, 1995; Majda & Kramer, 1999). Although based on a strong idealized decorrelation property, this model exhibits in particular anomalous dissipation.

The differences between the total (Lagrangian) derivative $dq \triangleq q(\mathbf{X}_{t+\delta t}, t+\delta t) - q(\mathbf{X}_t, t)$, the temporal Eulerian increment $d_t q \triangleq q(\mathbf{X}_t, t+\delta t) - q(\mathbf{X}_t, t)$ and the stochastic transport operator $\mathbb{D}_t q$ are essential. Indeed, as we will see in §3, the Stochastic Grid Perturbation method consists in connecting the first two operators and aims to simulate the last one. Let us mention that the transport operator is formally defined. As in the deterministic case, the stochastic partial differential equations (SPDEs) derived from it are only defined in a weak sense, meaning that the functions are not necessarily differentiable in a strong sense. The noise is assumed to be smooth in space (in the weak sense) while the function considered are not smoother than in the deterministic case. Interested readers may find precise mathematical assumptions for the noise in the LU Navier-Stokes equations in (Debussche et al., 2023). It is also worth pointing out that for a linear transport equation, the type of noise we are dealing with (called transport noise) allows for relaxing the Lipschitz condition on the resolved drift (Fedrizzi & Flandoli, 2013)

### 2.2 Detailed derivation

This section contains the technical derivation of the stochastic transport operator $\mathbb{D}_t$ which is crucial for the upcoming proof of the Stochastic Grid Perturbation method. Readers focused on the final results can skip this Section without losing the main thread.

The starting point of the LU derivation is the Ito-Wentzell formula (Kunita, 1990) which is a chain rule for random (semimartingale) functions with random (semimartingale) arguments:

**Ito-Wentzell formula.** *Let $f(\mathbf{x}, t)$ be a $C^2$-process over $\mathbf{x} \in \mathbb{R}^d$ and $C^1$-semimartingale over $t \in \mathbb{R}^+$; let $\mathbf{X}$ be a semimartingale over $\Omega \subset \mathbb{R}^d$. $f(\mathbf{X}, t)$ is a semimartingale satisfying*

$$df = d_t f + \sum_{i=1}^{d} \frac{\partial f}{\partial x_i} dX^i + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i \partial x_j} d\langle X^i, X^j\rangle_t + \sum_{i=1}^{d} d\langle \frac{\partial f}{\partial x_i}(\mathbf{X}, \cdot), X^i\rangle_t \tag{3}$$

Applying the Ito-Wentzell formula (3) to $dq(\mathbf{X}_t, t)$ decomposes the total derivative $dq$ into a time increment $d_t q$ (which plays the role of a partial derivative), a transport term $(v + \sigma d\mathbf{B}_t) \cdot \nabla q$ and additional quadratic co-variation terms:

$$dq = d_t q + \sum_{i=1}^{d} \frac{\partial q}{\partial x_i}(\sigma d\mathbf{B}_t)^i + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 q}{\partial x_i \partial x_j} d\langle X^i, X^j\rangle_t + \sum_{i=1}^{d} d\langle \frac{\partial q}{\partial x_i}(\mathbf{X}, \cdot), X^i\rangle_t + v \cdot \nabla q dt \tag{4}$$

Following the derivation of Li et al. (2023a), the quadratic co-variation $d\langle X^i, X^j\rangle_t$ is equal to $a_{ij}dt$. The time increment $d_t q$ is then decomposed into its smooth component and its

time uncorrelated component:

$$d_t q = f(\mathbf{X}_t, t)dt + \int_\Omega \sum_{k=1}^{d} g_k(\mathbf{X}_t, y, t) d\mathbf{B}_t^k(y) dy. \tag{5}$$

The quadratic co-variation depends only on the uncorrelated parts. The gradient of the tracer can hence be replaced by its uncorrelated component in the quadratic co-variation with the stochastic flow:

$$\langle \frac{\partial q}{\partial x_i}, X^i \rangle_t = \int_\Omega \sum_{j=1}^{d} \breve{\sigma}_{ij}(\mathbf{X}_t, \mathbf{y}, t) \frac{\partial g_j}{\partial x_i}(\mathbf{X}_t, \mathbf{y}, t) d\mathbf{y}, \quad \forall i = 1, ..., d \tag{6}$$

This leads to use the canonical decomposition of semi-martingales (i.e. the decomposition (5) is unique) to identify $f$ and $g$ in the equation $dq = 0$:

$$\begin{aligned} f &= -\frac{1}{2} \sum_{i,j=1}^{d} a_{ij} \frac{\partial^2 q}{\partial x_i \partial x_j} - \int_\Omega \sum_{j=1}^{d} \breve{\sigma}_{ij}(\mathbf{X}_t, \mathbf{y}, t) \frac{\partial g_j}{\partial x_i}(\mathbf{X}_t, \mathbf{y}, t) d\mathbf{y} - v \cdot \nabla q \\ g_j &= -\sum_{k=1}^{d} \breve{\sigma}_{jk} \frac{\partial q}{\partial x_k}. \end{aligned} \tag{7}$$

This identification specifies $d_t q$ for a tracer conserved by the flow. It depends notably on the quadratic co-variation (6) which can be further expanded with the substitution of $g_j$. After some calculus, one obtains

$$\int_\Omega \breve{\sigma}_{ij} \frac{\partial g_j}{\partial x_i} = \underbrace{-\frac{1}{2} a_{ik} \frac{\partial^2 q}{\partial x_i \partial x_k}}_{\text{cancelled (in } f)} - \frac{1}{2} \underbrace{\frac{\partial}{\partial x_i} \left( a_{ik} \frac{\partial q}{\partial x_k} \right)}_{\nabla \cdot (a \nabla q)} - \underbrace{\frac{1}{2} \frac{\partial a_{ik}}{\partial x_i} \frac{\partial q}{\partial x_k} + \sigma_{jk} \frac{\sigma_{ij}}{\partial x_i} \frac{\partial q}{\partial x_k}}_{(v^\star - v) \cdot \nabla q} \tag{8}$$

where the summation are dropped for sake of simplicity. By replacing $f$ and $g$ in (5), one obtains the definition of the stochastic transport operator (2):

$$\mathbb{D}_t q \triangleq d_t q + (\mathbf{v}^\star dt + \sigma d\mathbf{B_t}) \cdot \nabla q - \frac{1}{2} \nabla \cdot (\mathbf{a} \nabla q) dt = 0.$$

This derivation was based on the assumption $dq = 0$ but it can be extended to other equations involving $dq$ (e.g. Resseguier et al., 2017b).

## 3 Stochastic Grid Perturbation

Let us consider a Finite Difference space discretization where the grid points are noted $\mathbf{x}_i$, $i$ being a space index. The Stochastic Grid Perturbation method consists in moving grid points in a Lagrangian manner following the small scale uncertainty (see Figure 1). The continuous formulation in time reads:

$$\mathbf{x}_i^{\text{Brownian}}(t) - \mathbf{x}_i(0) = \int_0^t (\sigma d\mathbf{B}_t)_i \tag{9}$$

This method is illustrated in §3.1 with the simple example of a transport equation and is compared with the Location Uncertainty (LU) framework. Then §3.2 presents how time correlation in the grid motion (to keep the structure of the grid) can be compensated by an additional advection term. The exact link between the Stochastic Grid Perturbation method and Location Uncertainty is then discussed in §3.3.

### 3.1 Transport equation on a perturbed grid

We first show the effect of using the SGP method on the transport equation $\partial_t q \; dt = -\mathbf{v} \nabla q \; dt$. The left-hand side loses its time differentiability and becomes

$$q\left(\mathbf{x}_i^{\text{Brownian}}(t + \delta t), t + \delta t\right) - q\left(\mathbf{x}_i^{\text{Brownian}}(t), t\right).$$

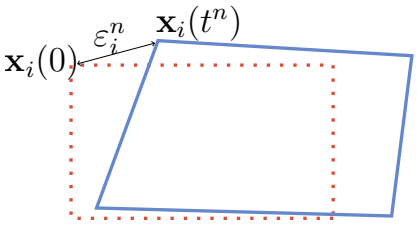


**Figure 1.** Illustration of the grid perturbation: dotted lines represent the initial grid and solid lines represent the perturbed grid at time $t^n$. All points are drifted to the right in order to illustrate the fact that the noise is spatially correlated.

- In the particular case where $\mathbf{v} = 0$, one can recognize a Lagrangian discretization of a cancelling total derivative $dq = 0$. The total derivative and the presence of noise is at the heart of the LU framework. Following step by step the derivation of the stochastic transport operator in §2.2 with $\mathbf{v} = 0$, the Ito-Wentzell formula applied to $q\left(\mathbf{x}_i^{\text{Brownian}}(t+\delta t), t+\delta t\right) - q\left(\mathbf{x}_i^{\text{Brownian}}(t), t\right)$ yields the right-hand side of (4) which contains a time increment $d_t q$, the perturbation term $\sigma d\mathbf{B}_t \cdot \nabla q$ and additional co-variation terms. Using the canonical martingale decomposition (5) of the time increment and exploiting its uniqueness to identify smooth and uncorrelated terms leads to (7), using (6) to simplify the smooth parts. With the slight simplification that $v = 0$, the expansion of the quadratic covariation (8) finally gives

$$\begin{cases} d\mathbf{x}_i^{\text{Brownian}} & = (\sigma d\mathbf{B}_t)_i \\ dq(\mathbf{x}_i, t) & = 0 \end{cases} \implies \mathbb{D}_t q = 0, \quad (\mathbf{v} = 0) \tag{10}$$

- In the general case with any $\mathbf{v}$, the left-hand side can be noted $d_t(q \circ \mathbf{x}_i^{\text{Brownian}})$ (where $\circ$ is the composition) and represents a material derivative taking only the uncorrelated part of the velocity into account (see §3.3). The derivation of §2.2 can still be followed by applying the Ito-Wentzell formula (3) on $d_t(q \circ \mathbf{x}_i^{\text{Brownian}}, t) + v \cdot \nabla q dt$ which gives the right-hand side of (4). The canonical decomposition of semimartingales is then used on the equation $d_t(q \circ \mathbf{x}_i^{\text{Brownian}}, t) = -v \cdot \nabla q dt$, which yields system (7). Assuming a perfect discretization of $\nabla q$, one finally obtains

$$\begin{cases} d\mathbf{x}_i^{\text{Brownian}} & = (\sigma d\mathbf{B}_t)_i \\ d_t(q \circ \mathbf{x}_i^{\text{Brownian}}) & = -\mathbf{v}(\mathbf{x}_i, t) \cdot \nabla q dt \end{cases} \implies \mathbb{D}_t q = 0 \tag{11}$$

For the transport equation, the perturbed grid $(\mathbf{x}_i^{\text{Brownian}})$ hence implements the stochastic transport operator $\mathbb{D}_t$ defined in (2).

However, $\int_0^t \sigma d\mathbf{B}_t$ is not bounded and the distance between a grid point $\mathbf{x}_i^{\text{Brownian}}(t)$ and its original location $\mathbf{x}_i^{\text{Brownian}}(0)$ possibly increases infinitely. As consequences, the grid loses its structure with time and the evaluation of $\nabla q$ potentially reduces its accuracy. Two solutions can be applied to keep the grid close to its original disposition:

- Restore the grid to its original state by interpolating all the variables. However, this global interpolation has a significant cost. It is interesting to note that in the particular case of some semi-Lagrangian models, an interpolation is already implemented to realign the grid points to their original positions.
- Introduce a time-correlated noise which keeps $\mathbf{x}_i(t) - \mathbf{x}_i(0)$ close to zero. This is the subject of §3.2.

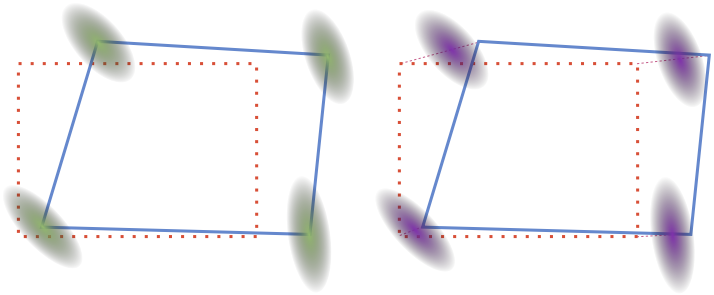

**Figure 2.** Representation of Brownian (left) and Auto-Regressive (right) noises for the grid perturbation. Dotted lines are the initial grid, solid lines are the current perturbed grid, and the ellipses represent the positions of the future grid points. Auto-Regressive processes perform a translation of the probability density function from the current grid towards the initial grid.

### 3.2 Ornstein-Uhlenbeck processes in Stochastic Grid Perturbation

In their grid perturbation, Leroux et al. (2022) did not use a Brownian motion but an auto-regressive process. The crucial property of auto-regressive processes is that it is possible to bound their variance, which removes the problem of grid divergence mentioned earlier. Let us note the current perturbation of a grid point $\varepsilon_i^n = \mathbf{x}_i(t_n) - \mathbf{x}_i(0)$ as in Figure 1. The first order Auto-Regressive process AR(1) is

$$\varepsilon_i^n = e^{-\alpha\delta t}\varepsilon_i^{n-1} + \sigma(\mathbf{B}(t_n) - \mathbf{B}(t_{n-1})) \tag{12}$$

where $e^{-\alpha\delta t}$ is slightly smaller than 1 to ensure that the variance of AR(1) is bounded for $n \to \infty$. The continuous counterparts of AR(1) are Ornstein-Uhlenbeck processes $R_t$, which satisfy the stochastic differential equation

$$d\mathbf{R}_t = -\alpha\mathbf{R}_t dt + \sigma d\mathbf{B}_t, \qquad R_0 = 0. \tag{13}$$

Injecting equation (13) in (11) yields

$$\begin{cases} d\mathbf{x}_i^{\text{Brownian}} & = \alpha\mathbf{R}_t dt + d\mathbf{R}_t \\ d_t(q \circ \mathbf{x}_i^{\text{Brownian}}) & = -\mathbf{v}(\mathbf{x}_i, t) \cdot \nabla q dt \end{cases} \implies \mathbb{D}_t q = 0 \tag{14}$$

Similarly to §3.1, the nudging term $\alpha\mathbf{R}_t dt$ can be "transferred" from the first line to the second line of the system. The effect is to change the noise of the perturbation and to compensate this change by an additional advection term which depends on the current displacement of the grid:

$$\begin{cases} d\mathbf{x}_i^{\text{O-U}} & = d\mathbf{R}_t \\ d_t(q \circ \mathbf{x}_i^{\text{O-U}}) & = (-\mathbf{v}(\mathbf{x}_i, t) + \alpha\mathbf{R}_t) \cdot \nabla q dt \end{cases} \implies \mathbb{D}_t q = 0 \tag{15}$$

Figure 2 represents the difference between Brownian and Auto-Regressive processes in the Stochastic Grid Perturbation method. It should be noted that the probability density function is simply translated towards the initial grid in the second case.

**Remark.** *In a discrete-in-time setting, the AR(1) process can replace Ornstein-Uhlenbeck process in the previous operations and the compensation can be derived at the discrete level. One then obtains*

$$\underbrace{\sigma(\mathbf{B}(t_n) - \mathbf{B}(t_{n-1}))}_{\text{“}\sigma d\mathbf{B}_t\text{”}} = \underbrace{(\varepsilon_i^n - \varepsilon_i^{n-1})}_{\text{“}d\mathbf{R}_t\text{”}} + \underbrace{(1 - e^{-\alpha\delta t})\varepsilon_i^{n-1}}_{\text{“}\alpha\mathbf{R}_t dt\text{”}}$$

*Following the continuous case, the compensation term $\alpha\mathbf{R}_t \cdot \nabla q\ dt$ can be discretized by $((1-e^{-\alpha\delta t})\varepsilon_i^{n-1}) \cdot \nabla q$. It is interesting to note that the compensation uses the grid disposition at the former time step and not at the current one. The grid at the former time step must hence be stored to implement this discrete compensation term.*

### 3.3 General link between SGP and LU

Let us define the operator $\partial_t^{\mathrm{SGP}}$ which represents (at the continuous level in time) a partial derivative in time under the Stochastic Grid Perturbation method: for a quantity $\theta$, its definition is $\partial_t^{\mathrm{SGP}}\theta \triangleq d_t(\theta \circ \mathbf{x}_i^{\mathrm{Brownian}})$. If a Ornstein-Uhlenbeck process is used, as proposed in §3.2, in order to keep the grid close to its original state, then one can use an alternative (but equivalent) definition: $\partial_t^{\mathrm{SGP}}\theta \triangleq d_t(\theta \circ \mathbf{x}_i^{\mathrm{O-U}}) - \alpha\mathbf{R}_t \cdot \nabla\theta dt$. This section clarifies the links and differences between $\partial_t^{\mathrm{SGP}}\theta$, the total variation $d\theta$ and the stochastic transport operator $\mathbb{D}_t\theta$.

Applying the Ito-Wentzell formula (3) to $\partial_t^{\mathrm{SGP}}\theta$, one obtains:

$$\partial_t^{\mathrm{SGP}}\theta = d_t\theta + \sum_{i=1}^{d} \frac{\partial\theta}{\partial x_i}(\sigma d\mathbf{B}_t)^i + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2\theta}{\partial x_i \partial x_j} d\langle X^i, X^j\rangle_t + \sum_{i=1}^{d} d\langle \frac{\partial\theta}{\partial x_i}(\mathbf{X},\cdot), X^i\rangle_t. \quad (16)$$

This operator $\partial_t^{\mathrm{SGP}}$ only takes into account the uncorrelated component of the velocity. The variations $d\theta$ and $\partial_t^{\mathrm{SGP}}\theta$ (given by equations (4) and (16), respectively) differ only in the presence of a transport term by the smooth component of the velocity:

$$\partial_t^{\mathrm{SGP}}\theta = d\theta - \mathbf{v}\cdot\nabla\theta dt. \quad (17)$$

The SGP operator is hence equivalent to the total derivative without the smooth velocity.

Note that, in (16), the last term $d\langle \frac{\partial\theta}{\partial x_i}(\mathbf{X},\cdot), X^i\rangle_t$ depends on the equation under consideration. For instance for $\partial_t^{\mathrm{SGP}}\theta + \mathbf{v}\cdot\nabla\theta dt = 0$, this term leads (as explained in §2.2) to $\partial_t^{\mathrm{SGP}}\theta + \mathbf{v}\cdot\nabla\theta = \mathbb{D}_t\theta$. Conversely, the presence of a non-zero right-hand side can instigate a modification in the very definition of $\partial_t^{\mathrm{SGP}}$, causing it to diverge from $\mathbb{D}_t$ (which is defined independently of the equation). The rest of this section focuses on explaining this difference between the Stochastic Grid Perturbation and Eulerian implementations of the Location Uncertainty parameterization.

Appendix B of Resseguier et al. (2017a) derives the difference between the material derivative and the stochastic transport operator. This difference is non-zero only in the presence of a stochastic forcing. They decompose $\mathbb{D}_t\theta$ into a smooth part $f dt$ and an uncorrelated component $\mathbf{h}^T d\mathbf{B}_t$:

$$\mathbb{D}_t\theta = f dt + \mathbf{h}^T d\mathbf{B}_t. \quad (18)$$

This decomposition leads them to $d\theta = \mathbb{D}_t\theta + tr(\sigma^T\nabla\mathbf{h}^T)dt$, which gives a relation between the stochastic transport operator and the partial derivative in time under Stochastic Grid Perturbation:

$$\partial_t^{\mathrm{SGP}}\theta + \mathbf{v}\cdot\nabla\theta dt = d\theta = \mathbb{D}_t\theta + tr(\sigma^T\nabla\mathbf{h}^T)dt. \quad (19)$$

To illustrate the importance of the term $tr(\sigma^T\nabla\mathbf{h}^T)$, let us consider the example of the Quasi-Geostrophic (QG) equations. In the numerical experiments of Section 4, the QG equations will be based on the transport of potential vorticity and feature no stochastic forcing (see §4.1). This formulation ensures that $tr(\sigma^T\nabla\mathbf{h}^T) = 0$. Conversely, in Li et al. (2023a) one can see that, when considering QG equations based on the momentum $\mathbf{u} = (u, v)$, one has $\mathbf{h}^T d\mathbf{B}_t = \beta y\mathbf{k}\times\sigma_h d\mathbf{B}_t$. More precisely,

$$\begin{pmatrix}\mathbf{h}_u^T d\mathbf{B}_t\\ \mathbf{h}_v^T d\mathbf{B}_t\end{pmatrix} = \beta y\mathbf{k}\times\sigma_h d\mathbf{B}_t = \beta y\begin{pmatrix}-(\sigma_h d\mathbf{B}_t)_y\\ (\sigma_h d\mathbf{B}_t)_x\end{pmatrix} = \beta y\begin{pmatrix}-\sigma_{21} & -\sigma_{22}\\ \sigma_{11} & \sigma_{12}\end{pmatrix} d\mathbf{B}_t.$$

This noise is additive and a simplified form of $tr(\sigma^T\nabla\mathbf{h}_u^T), tr(\sigma^T\nabla\mathbf{h}_v^T)$ can be derived. The difference between $(\partial_t^{\mathrm{SGP}}\mathbf{u} + \mathbf{v}\cdot\nabla\mathbf{u}\ dt)$ and $\mathbb{D}_t\mathbf{u}$ is finally

$$\begin{pmatrix}tr(\sigma^T\nabla\mathbf{h}_u^T)\\ tr(\sigma^T\nabla\mathbf{h}_v^T)\end{pmatrix} dt = \beta y\begin{pmatrix}-\sum_i \sigma_{i\bullet}\partial_i\cdot\sigma_{2\bullet}\\ \sum_i \sigma_{i\bullet}\partial_i\cdot\sigma_{1\bullet}\end{pmatrix} dt + \beta\begin{pmatrix}\sigma_{21}^2+\sigma_{22}^2\\ \sigma_{1\bullet}\cdot\sigma_{2\bullet}\end{pmatrix} dt, \quad (20)$$

where $\bullet$ stands for any index value. Injecting the values of $\beta, y, \sigma$ used in the numerical experiments of Section 4, we numerically obtain that $\frac{||tr(\sigma^T \nabla \mathbf{h}^T)dt||_\infty}{||d\mathbf{u}||_\infty} < 10^{-4}$ and $\frac{||tr(\sigma^T \nabla \mathbf{h}^T)dt||_2}{||d\mathbf{u}||_2} < 10^{-3}$. This suggests that the difference between SGP and LU is minimal in practice.

In the opposite case of systems which do not contain any stochastic forcing, the SGP is strictly equivalent to the stochastic transport operator:

$$\partial_t^{\mathrm{SGP}} q + \mathbf{v} \cdot \nabla q dt = \mathbb{D}_t q \tag{21}$$

This strict equivalence notably occurs in the equations numerically studied in the following section.

## 4 Numerical experiments

In this section we examine the Stochastic Grid Perturbation method applied on a 3-layer Quasi-Geostrophic model described in §4.1. Several simulations using different approaches are described in §4.2. Those simulations are compared qualitatively in §4.3 then quantitatively in §4.4.

### 4.1 Numerical model description

The numerical model we used was created by Thiry et al. (2023) and slightly modified to make the Stochastic Grid Perturbation possible. This model is described here again for completeness. In all the numerical experiments the prognostic variables are the potential vorticity $\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$ (in $\mathrm{s}^{-1}$) and the dynamic pressure $\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$ (in $\mathrm{m}^2.\mathrm{s}^{-2}$), stacked in three isopycnal layers. The 3-layer QG equations read:

$$\partial_t \mathbf{q} = \frac{1}{f_0} J(\mathbf{q}, \mathbf{p}) + f_0 B\mathbf{e} - \frac{a_4}{f_0} \Delta^2(\Delta \mathbf{p}), \tag{22}$$

$$\big(\Delta - f_0^2 A\big)\mathbf{p} = f_0 \mathbf{q} - f_0 \beta (y - y_0), \tag{23}$$

where $\Delta$ is the horizontal Laplacian, $\Delta^2$ the biharmonic operator, $J(a,b) = \partial_x a \partial_y b - \partial_x b \partial_y a$ stands for the Jacobi operator, $f_0 + \beta(y - y_0)$ is the Coriolis parameter under beta-plane approximation with the meridional axis centered on $y_0$, and $a_4$ is the biharmonic viscosity coefficient. In the case of grid perturbation, the time derivative $\partial_t$ in (22) is replaced by its perturbed counterpart $\partial_t^{\mathrm{SGP}}$, with spatial derivatives also influenced by the grid perturbation. For the LU framework, we specifically consider a stochastic transport of $\mathbf{q}$ (with forcing and dissipation), namely

$$\mathbb{D}_t q = \left( f_0 B\mathbf{e} - \frac{a_4}{f_0} \Delta^2(\Delta \mathbf{p}) \right) dt, \tag{24}$$

where the same kinematic relationship as in (23) applies.

The three-layer QG equations are integrated in a mid-latitude-centered rectangular ocean box model configuration, similar to the implementation used by Hogg et al. (2005). The ocean circulation is driven by a stationary, symmetrical wind stress at the surface and a linear Ekman stress at the bottom, generating an eastward turbulent jet featuring an idealized Gulf Stream in which spontaneous eddy production takes place. Despite its simplicity, such an idealized model is capable of producing meaningful mesoscale dynamics provided horizontal resolution is sufficiently high and viscosity is relatively low. In this paper, simulations were carried out with a horizontal resolution of 40, 10 and 5 km. The numerical model setup was further modified by Thiry et al. (2023) to make the Stochastic Grid Perturbation method possible.

| | Double-gyre 3-layer QG |
|---|---|
| $L_x, L_y$ | $(3840 \times 4800)$km |
| $H_k$ | $(350, 750, 2900)$m |
| $g'_{k+0.5}$ | $(0.025, 0.0125)$ ms$^{-2}$ |
| $\delta_{ek}$ | 2m |
| $\tau_0$ | $2 \times 10^{-5}$m$^2$s$^{-1}$ |
| Simulation time | 1 year |
| Time step | 2 hours |
| Temporal scheme | SSRK3 |
| $a_4$ | $5 \times 10^{11}$ m$^4$s$^{-1}$ |
| $f_0$ | $9.37 \times 10^{-5}$ s$^{-1}$ |
| $\beta$ | $1.75 \times 10^{-11}$ m$^{-1}$s$^{-1}$ |
| $\alpha$ | $5 \times 10^{-5}$ s$^{-1}$ |

**Table 1.** Model parameters shared by all simulations.

Table 1 gives the values of the parameters shared by all simulations. The values of $B$ and $\mathbf{e}$, given by Thiry et al. (2023), are:

$$B = \begin{bmatrix} \frac{1}{H_1} & \frac{-1}{H_1} & 0 & 0 \\ 0 & \frac{1}{H_2} & \frac{-1}{H_2} & 0 \\ 0 & 0 & \frac{1}{H_3} & \frac{-1}{H_3} \end{bmatrix}, \ \mathbf{e} = \begin{bmatrix} \partial_x \tau^y - \partial_y \tau^x \\ 0 \\ 0 \\ \frac{\delta_{\mathrm{ek}}}{2|f_0|} \Delta p_3 \end{bmatrix}, \ \vec{\tau} = \tau_0 \begin{bmatrix} -\cos(2\pi y / L_y) \\ 0 \end{bmatrix}, \quad (25)$$

where $\tau_0$ represents the surface wind magnitude, $H_k$ is the background thickness of the $k$-th layer and $\delta_{ek}$ is the thickness of the lower Ekman layer. The vertical stratification of this model is described by the term $-f_0^2 A \mathbf{p}$ where

$$A = \begin{bmatrix} \frac{1}{H_1 g'_{1.5}} & \frac{-1}{H_1 g'_{1.5}} & 0 \\ \frac{-1}{H_2 g'_{1.5}} & \frac{1}{H_2}\left(\frac{1}{g'_{1.5}} + \frac{1}{g'_{2.5}}\right) & \frac{-1}{H_2 g'_{2.5}} \\ 0 & \frac{-1}{H_3 g'_{2.5}} & \frac{1}{H_3 g'_{2.5}} \end{bmatrix}, \quad (26)$$

with $g'_{k+0.5}$ the reduced gravity defined between layers $k$ and $k+1$.

The noise $\sigma d\mathbf{B}_t$ is computed through offline analysis of high-resolution simulations based on Empirical Orthogonal Functions (see Li et al., 2023b). Specifically, high-pass filtering is applied to high-resolution simulation data to extract streamfunction fluctuations, from which velocity fluctuations are derived and aggregated into $N = 50$ orthogonal modes via the EOF procedure. The semi-martingale associated with the noise is then represented as a linear combination of these modes, weighted by standard Brownian motions:

$$\sigma(\mathbf{x}) d\mathbf{B}_t = \sum_{i<N} \sqrt{\lambda_i} \mathbf{\Phi}_i(\mathbf{x}) d\xi_t^i + \mathbf{\Gamma}(\mathbf{x}) dt, \quad (27)$$

where $\lambda_i$ is the eigenvalue associated with the spatial mode $\mathbf{\Phi}_i$ and $(d\xi_t^i)_{i<N}$ denote independent and identically distributed standard Brownian motions. The drift term $\mathbf{\Gamma}$ is introduced through Girsanov transformation (change of probability measure), with the aim of replicating the mid-latitude jet observed in the high-resolution experiment. This field is concentrated in mid-latitudes and can be found in the input data of the provided code (Clement & Li, 2024).

This choice of $\sigma$ is tailored for the LU model to resemble a high-resolution simulation. Consequently, if a method differs from the LU model, using the same noise parameterization is likely to result in a solution that does not resemble the high-resolution simulation. This discrepancy indicates a divergence from the LU model rather than an absolute measure of the method's performance. In fact, it would be possible to parameterize $\sigma$ specifically for this method to achieve better results.

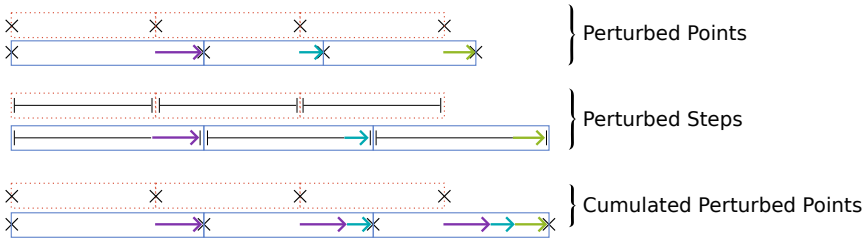


**Figure 3.** Illustration of the grid perturbation methods "Perturbed Points" (top) where the noise is applied to the grid points and "Perturbed Steps" (center) where the noise is applied to the space steps. The dotted orange rectangles represent the initial grid and blue rectangles represent the perturbed grid. On the bottom is an illustration of the equivalence between "Perturbed Points" with a cumulative sum of the noise and "Perturbed Steps".

### 4.2 Stochastic Grid Perturbation and Metric Perturbation

The perturbation described in Leroux et al. (2022) does not consist in perturbing the grid points but instead the grid metric. The associated experiment is hence called "Perturbed Steps" and will be compared (despite its lack of direct theoretical relation with LU and its possible artefacts on the eastern boundary) to the Stochastic Grid Perturbation presented in this paper, named "Perturbed Points". Figure 3 illustrates the difference between the two approaches. In a rectangular grid, the perturbation of the space steps amounts to modify $\sigma d\mathbf{B}_t$ with a cumulative sum from the lower space indices to the greater ones. This is illustrated on the bottom of Figure 3 where the grid points with a cumulative noise are aligned with the grid points of the "Perturbed steps" method. Inversely, moving the grid points would amounts to modify the "Perturbed steps" noise by taking its difference between consecutive steps. There is hence no fundamental difference behind the distinction between "Perturbed Steps" and "Perturbed Points". Nevertheless, we will see that the choice between the two methods significantly changes the behavior of the model if $\sigma$ is kept the same.

In both cases, an Auto-Regressive process is used to preserve the structure of the grid. As it is explained in §3.2, a compensation (i.e. the additional advective term) is introduced in the "Perturbed points" method to recover the effects of an uncorrelated-in-time noise.

**Remark.** *The choice of the nudging intensity parameter $\alpha$ depends on the structure and intensity of the noise. In our numerical simulations, $\alpha = 5 \times 10^{-5}$ s$^{-1}$ which gives $e^{-\alpha\delta t} \approx 0.7$ for our choice of $\delta t$. The goal was to minimise $\alpha$ while keeping the perturbation small compared to the space step: $\epsilon_i \ll \delta x$. The value of $\alpha$ was chosen by hand through numerical experiments to prevent negative space steps and avoid reliance on numerical thresholds.*

To focus on the effect of the grid perturbation, the same advective term is applied to the "Perturbed steps" method. As it is illustrated on the bottom of Figure 3, the displacement of the grid points can be large in the latter method, especially with space-correlated noise. The control of the parameters of such a noise is also a difficult task because of a possible strong accumulation effect. Using this displacement to compute the compensation term leads to high advection speed which would require to drastically decrease the time step. Consequently, at the price of a possible theoretical inconsistency, the displacement of the "Perturbed points" method was used to compute the compensation, even in the "Perturbed steps" method. We consider 6 ensemble simulations (30 members each):

- Deterministic : high (10km) and low (40km) resolution
- Explicit LU

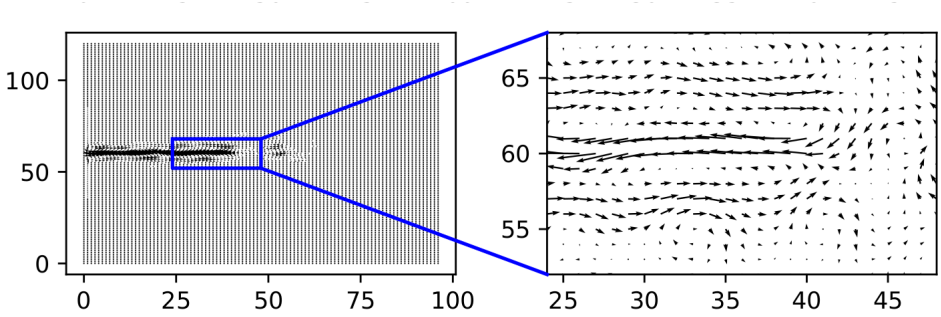


**Figure 4.** Visualization of the perturbation of the grid generated by the SGP method for one of the ensemble members at the end of a one-year simulation. Arrows linking the initial grid to the perturbed grid are plotted: the length of the arrows has been increased tenfold to better visualise the direction of the perturbation. On the left panel, the whole grid is represented and the right panel displays a zoom on a smaller portion.

- Perturbed Steps
- Perturbed Points
- Perturbed Points without the compensation term

Note that, since the role of (hyper)viscosity is to damp scales that are not properly resolved by the grid, the parameter $a_4$ was reduced to $10^{10}$ $\mathrm{m}^4.\mathrm{s}^{-1}$ in the high resolution simulation. Keeping this parameter constant would not significantly change the results compared to the low resolution simulation (but would require a significantly smaller time step).

Figure 4 shows the perturbation applied in the "Perturbed Points" case, and also serves as a representation of the noise (note that the amplitude of the perturbation is small, and that the length of the arrows has been multiplied by a factor of 10 to better visualize the displacement of the grid points). The noise features a so-called Girsanov drift to obtain a jet that is similar to the one in the deterministic High Resolution (HR) simulation (Li et al., 2023a).

*Creation of the initial ensemble* A very high resolution (5km, with the hyper-viscosity coefficient reduced to $a_4 = 2.10^9$ $\mathrm{m}^4\mathrm{s}^{-1}$) deterministic simulation at statistical equilibrium is used to create the initial ensemble. Assuming that the turbulence is sufficiently decorrelated in time, a snapshot of the very high resolution simulation is taken every 6 months in a 15-year simulation. To adjust the level of energy to the low resolution experiments, the (low resolution) explicit LU code then integrates in time all those snapshots for one year. For the 10km-resolution deterministic experiments, the snapshots of the 5km-resolution are used directly.

### 4.3 Qualitative study

In this section we focus on the solution profile at the end of the one-year simulation. Recall that the LU method and "perturbed points" are theoretically the same (up to possibly higher-order error terms), while "perturbed steps" and "without compensation" differ at leading order. The deterministic solutions serve as references to understand how significant the differences are. For consistency, the 10 km deterministic simulation data were downsampled to the 40 km coarse grid using a low-pass filtering procedure. A Gaussian kernel with a standard deviation of $\sigma = 2$ (half of the resolution ratio) and a width of $6\sigma$ (capturing 99.7% of the Gaussian distribution) was applied to filter the inner points of the high-resolution pressure fields. The boundary value of each downsampled pressure was

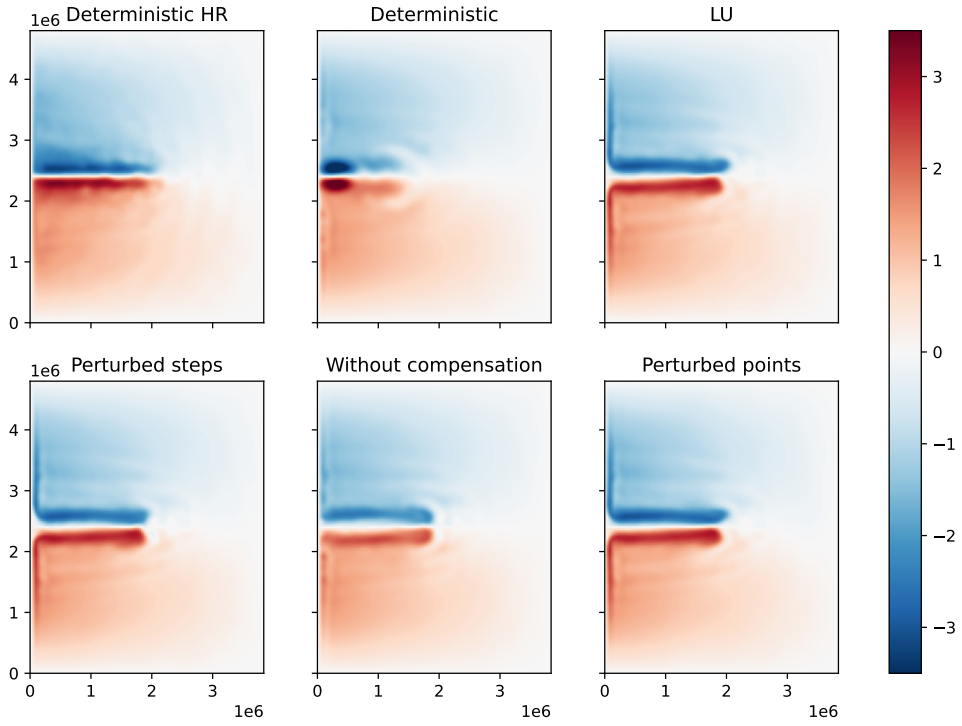


**Figure 5.** Surface pressure (in $m^2.s^{-2}$) averaged over the ensemble and over the last week of a one-year simulation.

determined by enforcing mass conservation constraints. Subsequently, other fields, such as potential vorticity, were derived on the coarse grid.

Figures 5 and 6 respectively present the averaged pressure and potential vorticity. The average is computed both over the ensemble and over the last week of a one-year simulation. Firstly, it can be observed that the LU formalism allows for the generation of a jet that is much more pronounced than without noise injection. The solutions of the "Perturbed Points" and "Perturbed Steps" methods do exhibit such an energetic jet: the behaviour of SGP is hence similar to LU in this regard. However, in the absence of the compensation term (experiment "Without compensation"), the jet is less pronounced.

Figure 7 provides the standard deviation of the surface pressure of the ensemble simulations. The standard deviation is also computed both over the ensemble and over the last week of the one-year simulation. It shows that, without the compensation term ("Without compensation"), the standard deviation is more concentrated at the center of the domain: it is stronger at the jet level and globally weaker in the rest of the basin. The compensation term seems hence to play a crucial role in transferring ensemble variability from grid points to the state variables. Interestingly, the "Perturbed steps" simulation also seems to exhibit a higher ensemble variance at the jet level compared to the LU approach. The LU method with EOF-based noise (27) produces lower variability compared to the high-resolution deterministic run, primarily due to the stationary assumption of the EOF approach. This limitation has been addressed in our previous works, where more advanced noise parameterization techniques and Girsanov drift corrections, such as state-dependent projections (Li et al., 2023a) and dynamic mode decomposition (Li et al., 2023b), were shown to improve the results. However, a detailed exploration of these methods is beyond the scope of this

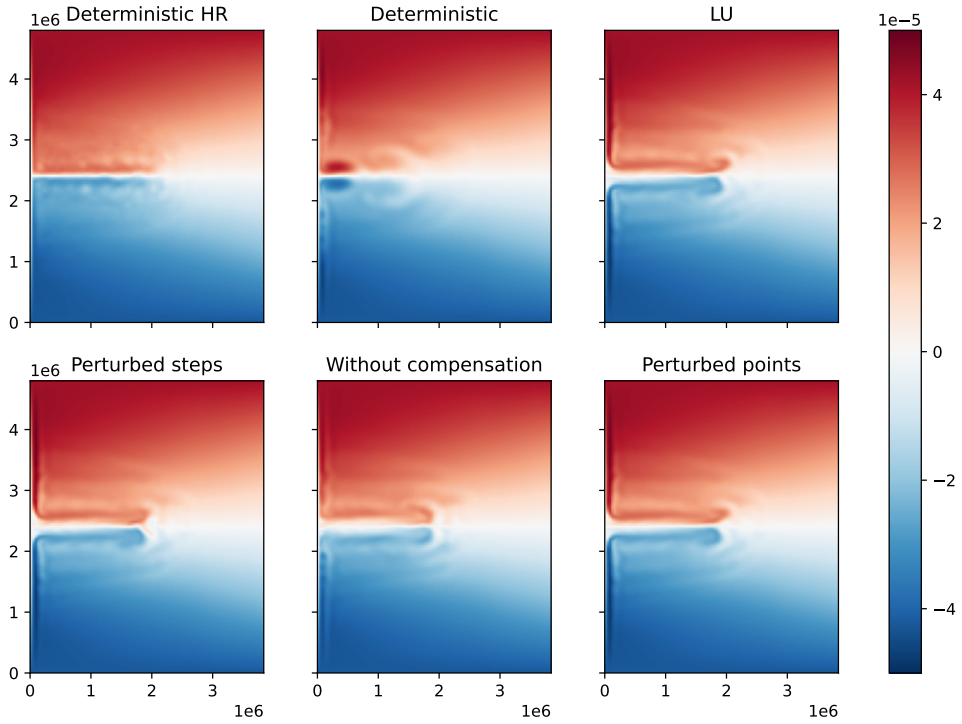


**Figure 6.** Surface potential vorticity (in $s^{-1}$) averaged over the ensemble and over the last week of a one-year simulation.

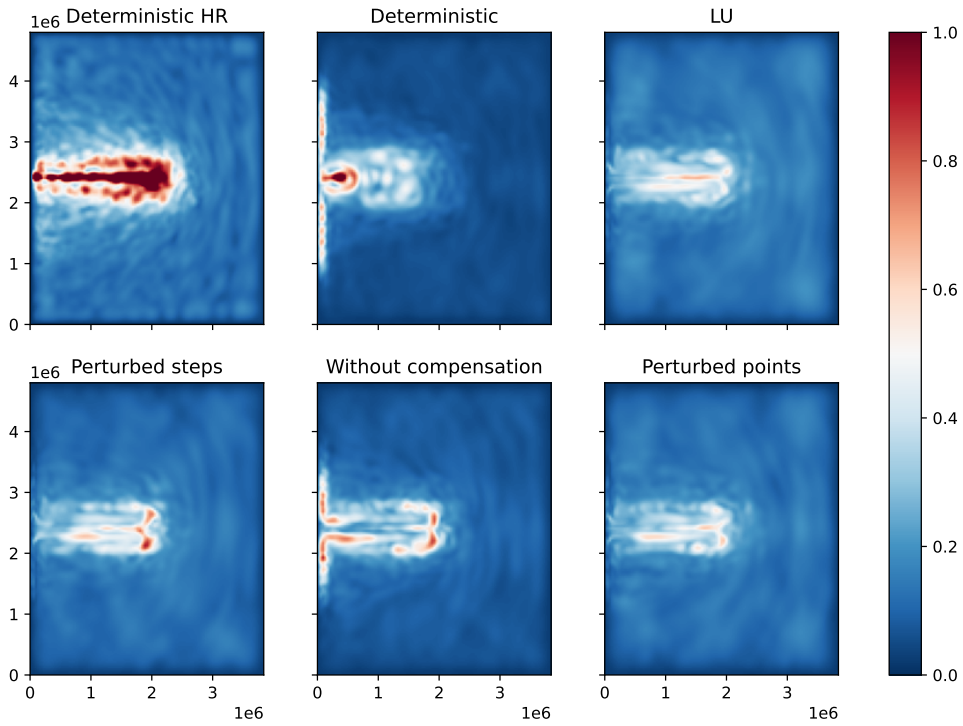


**Figure 7.** Point-wise standard deviation of surface pressure computed over the ensemble and over the last week of a one-year simulation.

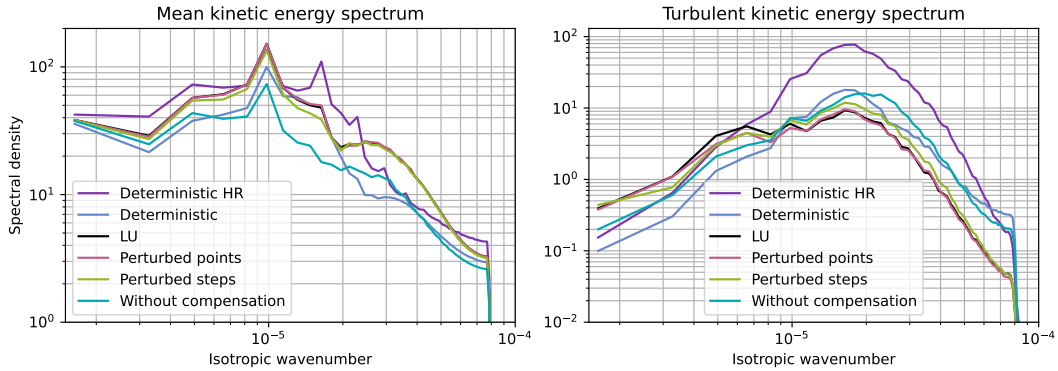


**Figure 8.** Mean (left) and turbulent (right) kinetic energy spectrum computed over the ensemble at the end of simulation.

work. The objective here is to establish the connection between grid perturbation methods and the LU framework, for which a simple yet efficient noise parameterization is sufficient to draw meaningful conclusions.

Figure 8 compares the mean kinetic energy (MKE) and turbulent kinetic energy (TKE) spectra of different models at the end of the simulation. The MKE spectrum reflects the energy of the ensemble mean, while the TKE spectrum shows the energy of deviations within the ensemble, indicating variance across scales. The MKE spectrum reveals that both the low-resolution deterministic run and the Perturbed Points method without compensation perform poorly at large and small scales. In contrast, the LU, Perturbed Steps, and Perturbed Points methods effectively transfer energy from small to intermediate scales, showing good backscattering. The method without compensation shows only weak backscattering and performs worse at large scales, highlighting the value of the compensation term. The TKE spectrum shows that the LU, Perturbed Points, and Perturbed Steps methods produce similar results, with the correct slope and no energy buildup at the cutoff scale. The Perturbed Steps method slightly increases TKE, while the method without compensation produces higher TKE but suffers from energy accumulation at the cutoff scale. This issue highlights its weaker ability to transfer energy from fluctuations to the mean, again showing the importance of compensation.

### 4.4 Quantitative study

Figures 9 and 10 display the first four spatially computed standardized moments of $u$ and $q$, respectively. We note $m_i(u,t) = \frac{1}{V}\int_V \left(u(\mathbf{x},t) - \overline{u}(\mathbf{x},t)\right)^i d\mathbf{x}$ the $i$-th order central moment ($i \geq 2$), where $V$ is the basin volume and $\overline{u}(\mathbf{x},t)$ is the ensemble average of $u$. The standard deviation, skewness and kurtosis are respectively defined by $\sqrt{m_2}$, $\frac{m_3}{(m_2)^{1.5}}$ and $\frac{m_4}{(m_2)^2}$. One can see on those figures that the statistics of the "Perturbed points" ensemble are close to those of the "LU" ensemble. The other two methods presented approximate some statistics at the expense of others:

- The mean vorticity with the "Perturbed steps" method significantly differs from other methods. Cumulative noise from low to high indices results in stronger noise in one gyre, creating average vorticity.
  It was also tested to remove the compensation term in this "Perturbed steps" experiment: doing so also halves the creation of vorticity. The displacement of the "Perturbed points" method used to compute the compensation (see §4.1) may be a cause of creation of vorticity. However there is still an unidentified source of vorticity,

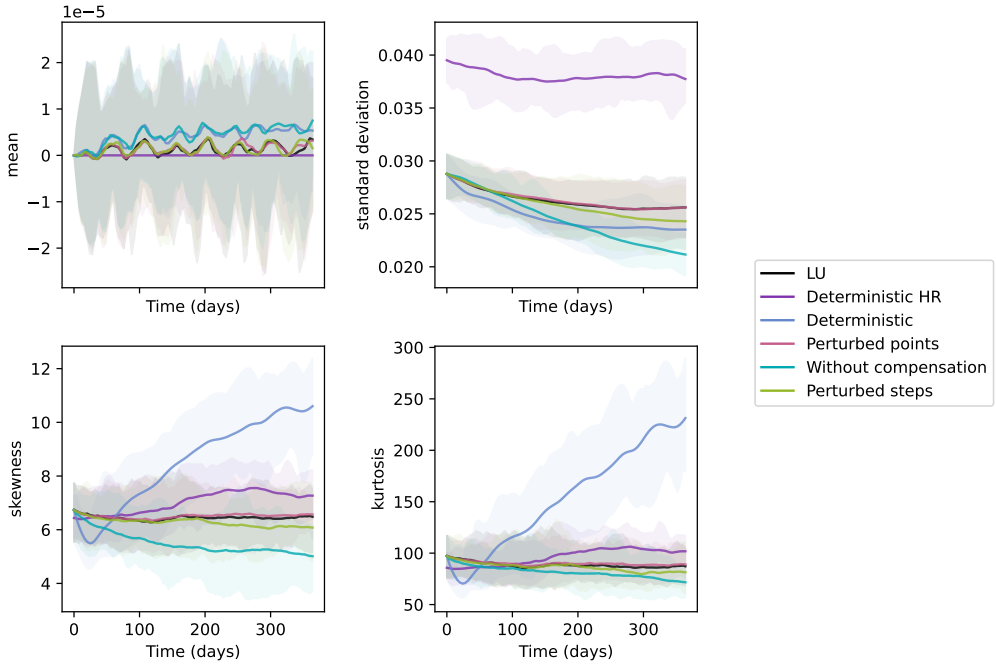


**Figure 9.** First four standardized moments of the variable $u$ (reconstructed from $p$). Each solid line represents the average of a moment over the ensemble members, and the envelopes contain all the moment values.

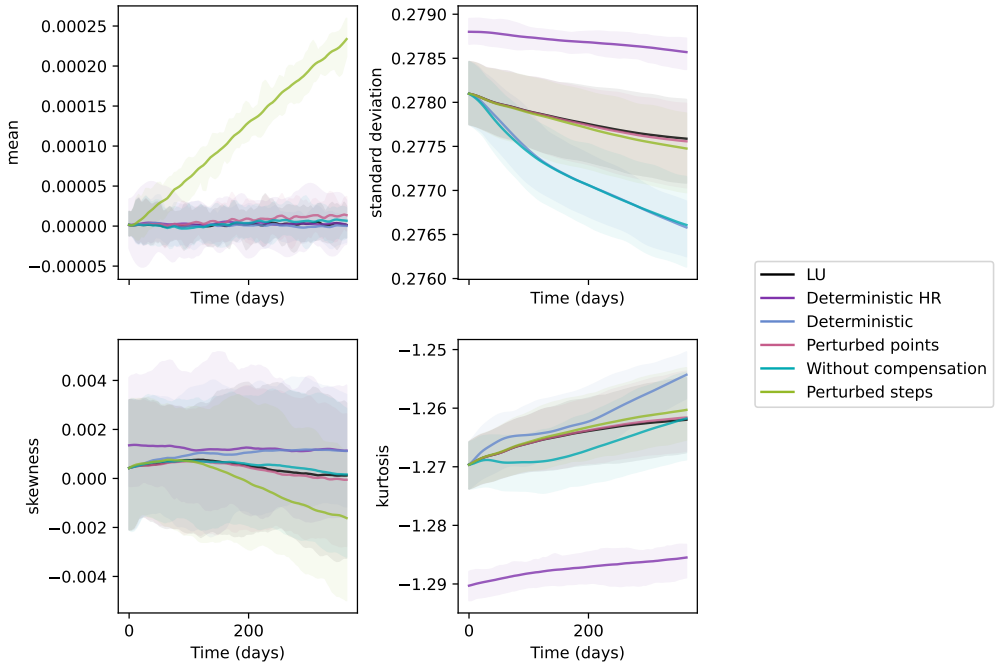


**Figure 10.** Same as Figure 9 but for the vorticity (variable $q$).

as removing the compensation term while using the cumulative sum in each gyre does not completely cancel the creation of vorticity.

- The standard deviation for both variables is underestimated with the method "Without compensation". The compensation term of the "Perturbed points" method is indeed one of the keys of the transfer of variability from the grid points to the variables $u$ and $q$.

Figure 9 also features interesting characteristics of skewness and kurtosis. First, the positive skewness of the variable $u$ suggests that extreme values are predominantly positive. The latter observation is expected, as this double-gyre configuration is characterized by the east-directed flow concentrating the extreme values. Second, the kurtosis of $u$ in the deterministic case is much larger than in the stochastic cases, indicating more extreme values. It can be noted that the deterministic simulation appears to lead to a slightly higher variance than the simulation without compensation. This suggests that the high skewness and kurtosis of the low resolution deterministic simulation does not come from small values of the variance. Figure 6 hints that in the absence of a clear jet, the deterministic (LR) experiment may create extreme values located on the west side between the two gyres. Note that the spatial statistics of the high-resolution deterministic experiments are displayed for completeness but are difficult to compare to the low-resolution experiments.

An additional analysis of the moments (not shown) centered on the jet area yields the same conclusions, except that the variance, skewness and kurtosis of the low-resolution deterministic simulation are much smaller, indicating the absence of a clear jet. The creation of vorticity of the "Perturbed steps" method is also largely reduced, indicating that it does not originate from the jet.

Finally, these figures (for the sake of readability) only display 3 out of the 30 ensemble members. Comparing the statistics of two ensembles can be done using the Wasserstein distance. This involves finding the permutation $\pi : [\![1, 30]\!] \to [\![1, 30]\!]$ that minimizes the term-wise distance: $\inf_\pi \left( \frac{1}{30} \sum_{i=1}^{30} ||X_i - Y_{\pi(i)}||^2 \right)^{\frac{1}{2}}$. This allows for the comparison of ensemble simulations for a given statistic. Tables 2 and 3 (summarized in Figure 11, (a)) provide the *relative distance to LU*. A reference is calculated by taking the Wasserstein distance between the first 15 and the last 15 members of the same ensemble from the 'LU' experiment (for a given turbulent statistic). The *relative distance to LU* is the Wasserstein distance between two ensembles divided by the latter reference. It can be understood as the member-to-member distance divided by the distance between members of the LU ensemble.

Figure 11 presents the relative distance to LU for each standardized moments of $u$ and $q$, gathered on radar charts. The left panel (a) is a summary of Tables 2 and 3 where the moments take into account the whole computational domain. Conversely, the moments are computed only on an area around the jet to create the right panel (b). Those radar charts are in line with the previous conclusions: the statistics of the "Perturbed points" ensemble are indeed close to those of the "LU" ensemble. All methods introducing a stochastic noise are closer to LU than the deterministic simulation, with the exception of the "Perturbed steps" method, due to its systematic creation of vorticity. The latter remark does not apply as strongly when the moments are computed only in the vicinity of the jet. Finally, we also note that the relative difference to LU is often less than 1. This means that the variability of the LU ensemble for those moments is larger than the difference between the studied approaches.

## 5 Conclusion

### 5.1 Summary

This paper has presented a Stochastic Grid Perturbation method based on autoregressive processes that is equivalent to an explicit implementation of the LU method. To achieve

| | Deterministic | Perturbed points | Without compensation | Perturbed steps |
|---|---|---|---|---|
| MEAN | 1.502 | 0.048 | 1.730 | 0.074 |
| STD | 2.272 | 0.155 | 4.436 | 0.416 |
| SKEWNESS | 33.139 | 0.471 | 5.772 | 0.738 |
| KURTOSIS | 92.389 | 0.420 | 1.435 | 0.567 |

**Table 2.** Relative distance to LU for the velocity $u$.

| | Deterministic | Perturbed points | Without compensation | Perturbed steps |
|---|---|---|---|---|
| MEAN | 0.168 | 1.466 | 0.293 | 643.695 |
| STD | 7.442 | 0.005 | 7.423 | 0.056 |
| SKEWNESS | 0.746 | 0.040 | 0.172 | 0.907 |
| KURTOSIS | 1.581 | 0.012 | 0.771 | 0.098 |

**Table 3.** Relative distance to LU for the vorticity $q$.

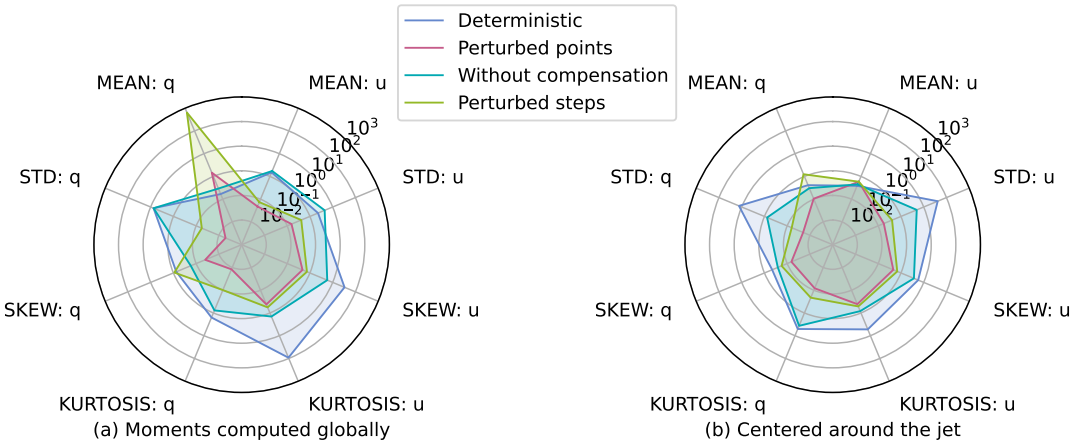


**Figure 11.** Log scaled radar chart showing the relative distance to LU for the first four standardized moments of the variables $u$ and $q$. Left panel (a): summary of Tables 2 and 3: the standardized moments are computed on the whole computational domain. Right panel (b): additional analysis where the standardized moments are computed on a smaller portion of the domain, including the jet area.

strict equivalence, a compensation term has been added in the form of advection. The importance of this compensation has been numerically studied, and the validity of the grid perturbation method has been confirmed. Perturbing grid steps differs from perturbing grid points: transitioning between the two is akin to modifying the operator $\sigma$ fundamentally, altering the statistical characteristics of the solutions.

In the model employed here, the stochastic transport operator applied to $q$ was a significant choice. As it was mentioned in Section 3.3, considering velocity $u$ as a prognostic variable under LU assumptions results in a system with distinct properties. It has been shown that, with the presence of a stochastic right-hand side, the effect of grid perturbation can differ from the stochastic transport operator although a numerical estimation hints that this difference may be small for the Quasi-Geostrophic equations.

This paper addresses the link between LU and SGP by modifying the SGP method to make it similar to LU. This approach is quite natural because LU offers more theoretical guarantees such as energy conservation along trajectories. However, one could alternatively design a noise term $\sigma d\mathbf{B}_t$ to express the SGP method as an instance of LU and study its properties. This noise would feature a Girsanov drift based on its past trajectory. Investigating such a noise within the LU framework would be interesting but challenging.

### 5.2 Application scenarios and limitations of SGP

The starting point of this study is the implementation of SGP in NEMO (Leroux et al., 2022). This implementation was motivated by the apparent simplicity of perturbing the grid to control the spread of an ensemble. However, their goal was not to implement exactly a LU method and the noise introduced was homogeneous. Indeed, the main difficulty of most LU implementations is the parameterization of the operator $\sigma$. This can be achieved for instance through offline analysis of high-resolution simulations based on Empirical Orthogonal Functions (Tucciarone et al., 2023) or more elaborate approach (Tucciarone et al., 2024).

When the SGP described in this paper is integrated into an Eulerian code with the compensation term, it combines the stochastic Eulerian transport of the compensation term with grid perturbation. In contrast, explicitly implementing LU would solely necessitate a stochastic Eulerian transport.

On another hand, there is presently an emerging interest in the use of Arbitrary Lagrangian Eulerian (ALE) methods (e.g. (Donea et al., 2004)) in the ocean modelling community. In such a perspective, it is interesting to note that within a code built upon an ALE framework, grid points can undergo direct perturbation using $\sigma d\mathbf{B}_t$ without the need for Ornstein-Uhlenbeck processes and a compensation term.

Applying the SGP method amounts to treat the smooth component of the motion in an Eulerian manner while using a Lagrangian method for the uncorrelated motion. This is somewhat similar to the so-called fractional-step methods (e.g. Huerta & Casadei, 1994), widely used in ALE frameworks. Therefore, this study presents the option of a semi-Lagrangian implementation of the LU equations, which holds more significance for semi-Lagrangian frameworks. Future work could focus on adapting this semi-Lagrangian setting to other discretization methods, such as Finite Volume methods.

## Open Research Section

The numerical model, the code for statistical analyses and the code generating the Figures and Tables can be found in the repository "Stochastic Grid Perturbation" (Clement & Li, 2024) at address `https://gricad-gitlab.univ-grenoble-alpes.fr/clemensi/qgm_sgp`. The latter Python3.10 code is under MIT license. Moreover, the datasets containing

the outputs of the simulations used in the Figures are available in Zenodo (Clement, 2024) at `https://doi.org/10.5281/zenodo.11650897`.

**Acknowledgments**

Parts of the computations presented in this paper were performed using the GRICAD infrastructure (https://gricad.univ-grenoble-alpes.fr), which is supported by Grenoble research communities. This work was supported by funding from the French government through the National Research Agency as part of France 2030, under references ANR-22-POCE-0003 and ANR-19-CE46-0011. L.L. and E.M. acknowledge the support of the ERC EU project 856408-STUOD.

## References

Bauer, W., Chandramouli, P., Chapron, B., Li, L., & Mémin, E. (2020a). Deciphering the role of small-scale inhomogeneity on geophysical flow structuration: a stochastic approach. *Journal of Physical Oceanography*, *50*(4), 983–1003.

Bauer, W., Chandramouli, P., Li, L., & Mémin, E. (2020b). Stochastic representation of mesoscale eddy effects in coarse-resolution barotropic models. *Ocean Modelling*, *151*, 101646.

Berner, J., Achatz, U., Batté, L., Bengtsson, L., Cámara, A., Christensen, H., ... others (2017). Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, *98*(3), 565–588.

Brecht, R., Li, L., Bauer, W., & Mémin, E. (2021). Rotating shallow water flow under location uncertainty with a structure-preserving discretization. *Journal of Advances in Modeling Earth Systems*, *13*(12), e2021MS002492.

Chapron, B., Dérian, P., Mémin, E., & Resseguier, V. (2018). Large-scale flows under location uncertainty: a consistent stochastic framework. *QJRMS*, *144*(710), 251-260.

Clement, S. (2024). *Data of the Stochastic Grid Perturbation comparison with Location Uncertainty framework.* Zenodo. doi: 10.5281/zenodo.11654106

Clement, S., & Li, L. (2024). *Stochastic Grid Perturbation on Quasi-Geostrophic model.* Retrieved from `https://hal.science/hal-04611878`

Crisan, D., Flandoli, F., & Holm, D. (2019). Solution properties of a 3D stochastic Euler fluid equation. *Journal of Nonlinear Science*, *29*(3), 813–870.

Debussche, A., Hug, B., & Mémin, E. (2023). A consistent stochastic large-scale representation of the navier–stokes equations. *Journal of Mathematical Fluid Mechanics*, *25*(1), 19. doi: 10.1007/s00021-023-00764-0

Debussche, A., & Mémin, E. (2024). Variational principles for fully coupled stochastic fluid dynamics across scales. *ArXiv*(2409.12654).

Donea, J., Huerta, A., Ponthot, J.-P., & Rodriguez-Ferran, A. (2004). Arbitrary lagrangian eulerian methods. In *Encyclopedia of computational mechanics* (chap. 14). John Wiley & Sons, Ltd. doi: 10.1002/0470091355.ecm009

Falkovich, G., Gawedzki, K., & Vergassola, M. (2001). Particles and fields in fluid turbulence. *Rev. Mod. Phys.*, *73*, 913–975. doi: 10.1103/RevModPhys.73.913

Fedrizzi, E., & Flandoli, F. (2013). Noise prevents singularities in linear transport equations. *Journal of Functional Analysis*, *264*(6), 1329-1354. doi: https://doi.org/10.1016/j.jfa.2013.01.003

Gawedzky, K., & Kupiainen, A. (1995). Anomalous scaling of the passive scalar. *Physical review letters*, *75*, 3834–3837.

Gottwald, G., Crommelin, D., & Franzke, C. (2017). Stochastic climate theory. In *Nonlinear and stochastic climate dynamics* (p. 209-240). Cambridge University Press.

Hasselmann, K. (1976, 2024/06/25). Stochastic climate models Part I. Theory. *Tellus*, *28*(6), 473–485.

Hogg, A. M. C., Killworth, P. D., Blundell, J. R., & Dewar, W. K. (2005). Mechanisms of decadal variability of the wind-driven ocean circulation. *Journal of Physical Oceanog-*

*raphy*, *35*(4), 512–531.

Holm, D. D. (2015). Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *471*(2176), 20140963.

Huerta, A., & Casadei, F. (1994). New ALE applications in non-linear fast-transient solid dynamics. *Engineering Computations*, *11*(4), 317–345.

Kitsios, V., Frederiksen, J. S., & O'Kane, T. J. (2023). Subgrid parameterization of eddy, meanfield and topographic interactions in simulations of an idealized antarctic circumpolar current. *Journal of Advances in Modeling Earth Systems*, *15*(5), e2022MS003412. doi: 10.1029/2022MS003412

Kraichnan, R. (1959). The structure of isotropic turbulence at very high reynolds numbers. *J. of Fluids Mech.*, *5*, 477–543.

Kraichnan, R. (1994). Anomalous scaling of a randomly advected passive scalar. *Physical Review Letters*, *72*(7), 1016.

Kunita, H. (1990). *Stochastic flows and stochastic differential equations* (Vol. 24). Cambridge university press.

Lang, O., Crisan, D., & Mémin, E. (2023). Analytical properties for a stochastic rotating shallow water model under location uncertainty. *Journal of Mathematical Fluid Mechanics*, *25*(2), 29. doi: 10.1007/s00021-023-00769-9

Leith, C. (1990). Stochastic backscatter in a subgrid-scale model: plane shear mixing layer. *Phys. of Fluids*, *2*(3), 1521–1530.

Leroux, S., Brankart, J.-M., Albert, A., Brodeau, L., Molines, J.-M., Jamet, Q., ... Brasseur, P. (2022). Ensemble quantification of short-term predictability of the ocean dynamics at a kilometric-scale resolution: a Western Mediterranean test case. *Ocean Science*, *18*(6), 1619–1644.

Lesieur, M., & Métais, O. (1996). New trends in large-eddies simulation of turbulence. *Annu. Rev. Fluid. Mech.*, *28*, 45–82.

Li, L. (2021). *Stochastic modeling and numerical simulation of ocean dynamics* (Unpublished doctoral dissertation). Université de Rennes 1.

Li, L., Deremble, B., Lahaye, N., & Mémin, E. (2023a). Stochastic Data-Driven Parameterization of Unresolved Eddy Effects in a Baroclinic Quasi-Geostrophic Model. *Journal of Advances in Modeling Earth Systems*, *15*(2), e2022MS003297.

Li, L., Mémin, E., & Tissot, G. (2023b). Stochastic parameterization with dynamic mode decomposition. In B. Chapron, D. Crisan, D. Holm, E. Mémin, & A. Radomska (Eds.), *Stochastic transport in upper ocean dynamics* (pp. 179–193). Cham: Springer International Publishing.

Majda, A., & Kramer, P. (1999). Simplified models for turbulent diffusion:theory, numerical modelling, and physical phenomena. *Physics report*, *314*, 237–574.

Mason, P., & Thomson, D. (1992). Stochastic backscatter in large-eddy simulations of boundary layers. *J. of Fluid Mech.*, *242*, 51–78.

Mémin, E., Li, L., Lahaye, N., Tissot, G., & Chapron, B. (2024). Linear wave solutions of a stochastic shallow water model. In *Stochastic transport in upper ocean dynamics ii* (pp. 223–245). Springer International Publishing.

Mémin, E. (2014). Fluid flow dynamics under location uncertainty. *Geophysical & Astrophysical Fluid Dynamics*, *108*(2), 119–146.

O'Kane, T. J., Fiedler, R., Collier, M. A., & Kitsios, V. (2023). Ocean model response to stochastically perturbed momentum fluxes. *Journal of Climate*, *36*(6), 1895–1922.

Resseguier, V., Mémin, E., & Chapron, B. (2017a). Geophysical flows under location uncertainty, Part I Random transport and general models. *Geophysical & Astrophysical Fluid Dynamics*, *111*(3), 149–176.

Resseguier, V., Mémin, E., & Chapron, B. (2017b). Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading. *Geophysical & Astrophysical Fluid Dynamics*, *111*(3), 177–208.

Thiry, L., Li, L., & Mémin, E. (2023). Modified (hyper-)viscosity for coarse-resolution ocean models. In B. Chapron, D. Crisan, D. Holm, E. Mémin, & A. Radomska (Eds.),

*Stochastic transport in upper ocean dynamics* (pp. 273–285). Cham: Springer International Publishing.

Tucciarone, F. L., Mémin, E., & Li, L. (2023). Primitive equations under location uncertainty: Analytical description and model development. In B. Chapron, D. Crisan, D. Holm, E. Mémin, & A. Radomska (Eds.), *Stochastic transport in upper ocean dynamics* (pp. 287–300). Cham: Springer International Publishing.

Tucciarone, F. L., Mémin, E., & Li, L. (2024). Data driven stochastic primitive equations with dynamic modes decomposition. In B. Chapron, D. Crisan, D. Holm, E. Mémin, & A. Radomska (Eds.), *Stochastic transport in upper ocean dynamics ii* (pp. 321–336). Cham: Springer Nature Switzerland.

Zhen, Y., Resseguier, V., & Chapron, B. (2023). Physically constrained covariance inflation from location uncertainty. *Nonlinear Processes in Geophysics*, *30*(2), 237–251.