Statistical optimization for passive scalar transport: maximum entropy production vs maximum Kolmogorov-Sinay entropy

M. Mihelich, D. Faranda, B. Dubrulle and D. Paillard*
Laboratoire SPHYNX, CEA/IRAMIS/SPEC, CNRS URA 2464,
F-91191 Gif-sur-Yvette, France; E-Mail: berengere.dubrulle@cea.fr
(Dated: september 16, 2014)

We derive rigorous results on the link between the principle of maximum entropy production and the principle of maximum Kolmogorov- Sinai entropy for a Markov model of the passive scalar diffusion called the Zero Range Process. We show analytically that both the entropy production and the Kolmogorov-Sinai entropy, seen as functions of a parameter f connected to the jump probability, admit a unique maximum denoted $f_{max_{EP}}$ and $f_{max_{KS}}$. The behavior of these two maxima is explored as a function of the system disequilibrium and the system resolution N. The main result of this article is that $f_{max_{EP}}$ and $f_{max_{KS}}$ have the same Taylor expansion at first order in the deviation from equilibrium. We find that $f_{max_{EP}}$ hardly depends on N whereas $f_{max_{KS}}$ depends strongly on N. In particular, for a fixed difference of potential between the reservoirs, $f_{max_{EP}}(N)$ tends towards a non-zero value, while $f_{max_{KS}}(N)$ tends to 0 when N goes to infinity. For values of N typical of those adopted by Paltridge and climatologists working on MEP ($N \approx 10 \sim 100$), we show that $f_{max_{EP}}$ and $f_{max_{KS}}$ coincide even far from equilibrium. Finally, we show that one can find an optimal resolution N_* such that $f_{max_{EP}}$ and $f_{max_{KS}}$ coincide, at least up to a second order parameter proportional to the non-equilibrium fluxes imposed to the boundaries. We find that the optimal resolution N^* depends on the non equilibrium fluxes, so that deeper convection should be represented on finer grids. This result points to the inadequacy of using a single grid for representing convection in climate and weather models. Moreover, the application of this principle to passive scalar transport parametrization is therefore expected to provide both the value of the optimal flux, and of the optimal number of degrees of freedom (resolution) to describe the system.

---

* REVTeX Support: revtex@aps.org

I. INTRODUCTION

A major difficulty in the modeling of nonlinear geophysical or astrophysical processes is the taking into account of all the relevant degrees of freedom. For example, fluid motions obeying Navier-Stokes equations usually require of the order of $N = Re^{9/4}$ modes to faithfully describe all scales between the injection scale and the dissipative scale (Frisch 1995). In atmosphere, or ocean, where the Reynolds number exceeds 10^9, this amount to $N = 10^{20}$, a number too large to be handled by any existing computers (Wallace and Hobbs 2006). The problem is even more vivid in complex systems such as planetary climate, where the coupling of lito-bio-cryo-sphere with ocean and atmosphere increases the number of degrees of freedom beyond any practical figure. This justifies the long historical tradition of parametrization and statistical model reduction, to map the exact equations describing the system onto a set of simpler equations involving few degrees of freedom. The price to pay is the introduction of *free parameters*, describing the action of discarded degrees of freedom, that needs to be prescribed.

When the number of free parameters is small, their prescription can be successfully done empirically through calibrating experiments or by a posteriori tuning (Rotstayn 2000). When the number of parameters is large, such as in climate models where it reaches several hundreds (Murphy et al. 2004), such empirical procedure is inapplicable, because it is impossible to explore the whole parameter space. In that respect, it is of great interest to explore an alternative road to parametrization via application of a statistical optimization principle, such as minimizing or maximizing of a suitable cost functional. As discussed by (Turkington 2013) and (Pascale et al. 2012), this strategy usually leads to closed reduced equations with adjustable parameters in the closure appearing as weights in the cost functional and can be computed explicitly. A famous example in climate is given by a principle of maximum entropy production (MEP) that allowed (Paltridge 1975) to derive the distribution of heat and clouds at the Earth surface with reasonable accuracy, without any parameters and with a model of a dozen of degrees of freedom (boxes). Since then, refinements of Paltrige model have been suggested to increase its generality and range of prediction (Herbert et al. 2011). MEP states that a stationary nonequilibrium system chooses its final state in order to maximize the entropy production as is explain in (Martyushev and Seleznev 2006). Rigorous justifications of its application have been searched using e.g. information theory (Dewar and Maritan 2014) without convincing success. More recently, we have used the analogy of the climate box model of Paltridge with the asymmetric exclusion Markov process (ASEP) to establish numerically a link between the MEP and the principle of maximum Kolmogorov- Sinai entropy (MKS)(Mihelich et al. 2014). The MKS principle is a relatively new concept which extends the classical results of equilibrium physics (Monthus 2011). This principle applied to Markov Chains provides an approximation of the optimal diffusion coefficient in transport phenomena (Gómez-Gardeñes and Latora 2008) or simulates random walk on irregular lattices (Burda et al. 2009). It is therefore a good candidate for a physically relevant cost functional in passive scalar modeling.

The goal of the present paper is to derive rigorous results on the link between MEP and MKS using a Markov model of the passive scalar diffusion called the Zero Range Process (Andjel 1982). We find that there exists an optimal resolution N_* such that both maxima coincide to second order in the distance from equilibrium. The application of this principle to passive scalar transport parametrization is therefore expected to provide both the value of the optimal flux, and of the optimal number of degrees of freedom (resolution) to describe the system. This suggests that the MEP and MKS principle may be unified when the Kolmogorov- Sinai entropy is defined on opportunely coarse grained partitions.

II. FROM PASSIVE SCALAR EQUATION TO ZRP MODEL

The equation describing the transport of a passive scalar like temperature in a given velocity field $u(x, t)$ reads:

$$\partial_t T + u\partial_x T = \kappa \partial_x^2 T, \tag{1}$$

with appropriate boundary conditions. Here κ is the diffusivity. To solve this equation, one must know both the velocity field and the boundary conditions, and use as many number of modes as necessary to describe all range of scales up to the scales at which molecular diffusivity takes place i.e. roughly $(RePr)^{3/2}$ modes, where Re is the Reynolds number of the convective flow, and Pr is its Prandtl number. In geophysical flows, this number is too large to be handled even numerically (Troen and Mahrt 1986). Moreover, in typical climate studies, the velocity flow is basically unknown as it must obey a complicated equation involving the influence of all the relevant climate components. In order to solve the equation, one must necessarily prescribe the heat flux $f = -uT + \kappa\nabla T$. The idea of Paltridge was then to discretize the passive scalar equation in boxes and prescribethe heat flux $f_{i(i+1)}$ between boxes i and $i+1$ by maximizing the associated thermodynamic entropy production $\dot{S} = \sum_i f_{i(i+1)}(\frac{1}{T_{i+1}} - \frac{1}{T_i})$. Here, we slightly modify the Paltridge discretization approximation to make it amenable to rigorous mathematical results on Markov Chains. For simplicity, we stick to a one dimensional case (corresponding to boxes varying only in latitude) and impose the boundary conditions through two reservoirs located at each end of the chain (mimicking the solar heat flux at pole and equator). We consider a set of N boxes that can contain an arbitrary number $n \in \mathbb{N}$ of particles. We then allow transfer of particles in between two adjacent boxes via decorrelated jumps (to the right or to the left) following a 1D Markov dynamics governed by a coupling with the two reservoirs imposing a difference of chemical potential at the ends. The resulting process is called the Zero Range Process (Andjel 1982). The different jumps are described as follow. At each time step a particle can jump right with probability pw_n or jump left with probability qw_n where w_n is a parameter depending of the number of particles inside the box. Physically it represents the interactions between particles. At the edges of the lattice the probability rules are different: At the left edge a particule can enter with probability α and exit with probability γw_n whereas at the right edge a particle can exit with probability βw_n and enter with probability δ. Choices of different w_n give radically different behaviors. For example $w_n = 1 + b/n$ where $b \geq 0$ described condensation phenomena (Großkinsky et al. 2003) whereas $w_1 = w$ et $w_n = 1$ if $n \geq 2$ has been used to modeled road traffic. We will consider in this article the particular case where $w = 1$ by convenience of calculation. Moreover without loss of generality we will take $p \geq q$ which corresponds to a particle flow from the left to the right and note $f = p - q$. After a sufficiently long time the system reaches a non-equilibrium steady state. The interest of this toy model is that it is simple enough so that exact computations are analytically tractable.

Taking the continuous limit of this process, it may be checked that the fugacity z,which is a quantity related to the average particle density (see 8 below), of stationary solutions of a system consisting of boxes of size $\frac{1}{N}$ follows the continuous equation (Levine et al. 2005) :

$$f\frac{\partial z}{\partial x} - \frac{1}{2N}\frac{\partial^2 z}{\partial x^2} = 0, \tag{2}$$

corresponding to stationary solution of a passive scalar equation with velocity f and diffusivity $\frac{1}{2N}$. Therefore, the fugacity of the Zero Range Process is a passive scalar obeying a convective-diffusion equation. We thus see that $f = 0$ corresponds to a purely conductive regime whereas the larger f the more convective the regime. In the sequel, we calculate the entropy production and the Kolmogorov-Sinai entropy function of f. These two quantities reach a maximum noted respectively $f_{max_{EP}}$ and $f_{max_{KS}}$. The MEP principle (resp. the MKS principle) states that the system will choose $f = f_{max_{EP}}$ (resp $f = f_{max_{KS}}$).
We will show first of all in this article that numerically $f_{max_{EP}} \approx f_{max_{KS}}$ even far from equilibrium for a number of boxes N roughly corresponding to the resolution taken by Paltridge (1975) in his climate model. This result is similar to what we found for the ASEP model (Mihelich et al. 2014) and thus gives another example of a system in which the two principles are equivalent. Moreover we will see analytically that $f_{max_{EP}}$ and $f_{max_{KS}}$ have the same behavior in first order in the difference of the chemical potentials between the two reservoirs for N large enough. These results provide a better understanding of the relationship between the MEP and the MKS principles.

III. NOTATIONS AND USEFUL PRELIMINARY RESULTS

This Markovian Process is a stochastic process with a infinite number of states in bijection with $\mathbb{N}^N$. In fact, each state can be written $n = (n_1, n_2, ..., n_N)$ where n_i is the number of particule lying in site i. We call P_n the stationary probability to be in state n. In order to calculate this probability it is easier to use a quantum formalism than the Markovian formalism as explained in the following articles (Domb 2000, Levine et al. 2005).

The probability to find m particles in the site k is equal to: $p_k(n_k = m) = \frac{z_k^m}{Z_k}$ where Z_k is the analogue of the grand canonical repartition function and z_k is the fugacity between 0 and 1. Moreover $Z_k = \sum_{i=0}^{\infty} z_k^i = \frac{1}{1-z_k}$. So, finally

$$p_k(n_k = m) = (1 - z_k)z_k^m, \tag{3}$$

We can show that the probability P over the states is the tensorial product of the probability p_k over the boxes:

$$P = p_1 \otimes p_2 \otimes .... \otimes p_N,$$

Thus events $(n_k = m)$ and $(n'_k = m')$ for $k \neq k'$ are independent and so:

$$P(m_1, m_2, ..., m_N) = p_1(n_1 = m_1) * ... * p_N(n_N = m_N), \tag{4}$$

So finally

$$P(m_1, m_2, ..., m_N) = \prod_{k=1}^{N} (1 - z_k) z_k^{m_k}. \tag{5}$$

Moreover, with the Hamiltonian equation found from the quantum formalism we can find the exact values of z_k function of the system parameters:

$$z_k = \frac{(\frac{p}{q})^{k-1}[(\alpha+\delta)(p-q) - \alpha\beta + \gamma\delta] - \gamma\delta + \alpha\beta(\frac{p}{q})^{N-1}}{\gamma(p-q-\beta) + \beta(p-q+\gamma)(\frac{p}{q})^{N-1}}, \tag{6}$$

and the flux of particles c:

$$c = (p-q)\frac{-\gamma\delta + \alpha\beta(\frac{p}{q})^{N-1}}{\gamma(p-q-\beta) + \beta(p-q+\gamma)(\frac{p}{q})^{N-1}}. \tag{7}$$

Finally, the stationary density is related to the fugacity by the relation:

$$\rho_k = z_k \frac{\partial \log Z_k}{\partial z_k} = \frac{z_k}{1 - z_k}. \tag{8}$$

A. Entropy Production

For a system subject to internal forces X_i and associated fluxes J_i the macroscopic entropy production is well known (Onsager 1931)and takes the form:

$$\sigma = \sum_i J_i * X_i.$$

The Physical meaning of this quantity is a measure of irreversibility: the larger σ the more irreversible the system. In the case of the zero range process irreversibility is created by the fact that $p \neq q$. We will parametrize this irreversibility by the parameter $f = p - q$ and we will take $p + q = 1$. In the remaining of the paper, we take, without loss of generality, $p \leq q$ which corresponds to a flow from left to right. Moreover, the only flux to be considered is here the flux of particules c and the associated force is due to the gradient of the density of particules $\rho : X = \nabla \log \rho$ (Balian 1992).
Thus, when the stationary state is reached ie when c is constant:

$$\sigma = \sum_{i=1}^{N-1} c.(\log(\rho_i) - \log(\rho_{i+1})) = c.(\log(\rho_1) - log(\rho_N)). \tag{9}$$

Thus, according to Eqs. (6), (7), (8) and (9) when N tends to $+\infty$ we obtain:

$$\sigma(f) = \frac{\alpha f}{f+\gamma}(\log(\frac{\alpha}{f+\gamma-\alpha}) - \log(\frac{(\alpha+\delta)f + \gamma\delta}{f(\beta-\alpha-\delta) + \beta\gamma - \gamma\delta})). \tag{10}$$

Because $f \geq 0$ the entropy production is positive if and only if $\rho_1 \geq \rho_N$ iff $z_1 \geq z_N$. This is physically coherent because fluxes are in the opposite direction of the gradient. We remark that if $f = 0$ then $\sigma(f) = 0$. Moreover, when f increases $\rho_1(f)$ decreases and $\rho_2(f)$ increases till they take the same value. Thus it exists f, large enough, for which $\sigma(f) = 0$. Between these two values of f the entropy production has at least one maximum.

B. Kolmogorov-Sinai Entropy

There are several ways to introduce the Kolmogorov-Sinai entropy which is a mathematical quantity introduced by Kolmogorov and developed by famous mathematician as Sinai and Billingsley (Billingsley 1965). Nevertheless, for a Markov process we can give it a simple physical interpretation: the Kolmogorov-Sinai entropy is the time derivative of the Jaynes entropy (entropy over the path).

$$S_{Jaynes}(t) = -\sum_{\Gamma_{[0,t]}} p_{\Gamma_{[0,t]}}.\log(p_{\Gamma_{[0,t]}}), \tag{11}$$

For a Markov Chain we have thus:

$$S_{Jaynes}(t) - S_{Jaynes}(t-1) = -\sum_{(i,j)} \mu_{i_{stat}} p_{ij} \log(p_{ij}), \tag{12}$$

where $\mu_{stat} = \mu_{i_{stat}} i = 1...N$ is the stationary measure and where the p_{ij} are the transition probabilities. Thus the Kolmogorov-Sinai entropy takes the following form:

$$h_{KS} = -\sum_{(i,j)} \mu_{i_{stat}} p_{ij} \log(p_{ij}), \tag{13}$$

For the Zero Range Process ,we show in appendix that it can be written as:

$$\begin{aligned} h_{KS} = &-(\alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + \beta\log\beta + (N-1)(p\log(p) + q\log(q))) + (p\log(p) + q\log(q))\sum_{i=1}^{N}(1-z_i) \\ &+ (\gamma\log(\gamma) + p\log(p))(1-z_1) + (\beta\log(\beta) + q\log(q))(1-z_N). \end{aligned} \tag{14}$$

IV. RESULTS

We will start first by pointing to some interesting properties of $f_{max_{EP}}$ and $f_{max_{KS}}$, then by presenting numerical experiments on the ZRP model and finally concluding with some analytical computations.
Let us first note that for N,α,β,γ,δ fixed the entropy production as well as the Kolmogorov-Sinai entropy seen as functions of f admit both a unique maximum. When N tends to infinity and $f = 0$, using Eq.(6) (i.e. the symmetric case), we find that $z_1 = \frac{\alpha}{\gamma}$ and $z_N = \frac{\delta}{\beta}$. Thus, the system is coupled with two reservoirs with respective chemical potential $\frac{\alpha}{\gamma}$(left) and $\frac{\delta}{\beta}$ (right). For $\frac{\alpha}{\gamma} \neq \frac{\delta}{\beta}$ the system is out of equilibrium. We assume, without loss of generality, $z_1 \geq z_N$ which corresponds to a flow from left to right. As a measure of deviation from equilibrium we take $s = z_1 - z_N$: the larger s, the more density fluxes we expect into the system.
First we remark that $f_{max_{EP}}$ hardly depends on N whereas $f_{max_{KS}}$ depends strongly on N. This is easily understood because σ depends only on z_1 and z_N whereas h_{KS} depends on all the z_i. Moreover, the profile of the z_i depends strongly on N. In particular, for a fixed difference of potential between the reservoirs , $f_{max_{EP}}(N)$ tends towards a non-zero value, while $f_{max_{KS}}(N)$ tends to 0 when N goes to infinity.
Moreover, $f_{max_{EP}}$ and $f_{max_{KS}}$ coincide even far from equilibrium for N corresponding to the choice of Paltridge (1975) $N \approx 10 \sim 100$. For N fixed, as large as one wants, and for all ϵ, as small as one wants, it exists ν such that for all $s \in [0;\nu]$ $|f_{max_{EP}} - f_{max_{KS}}| \leq \epsilon$.
These observations are confirmed by the results presented in Figures 1 and 3 where EP and KS are calculated using Eq. (6) and (14) for $s = 0.13$ and three different partitions: $N = 20$ $N = 100$ et $N = 1000$. The figure shows that $f_{max_{EP}}$ and $f_{max_{KS}}$ coincide with good approximation for $N = 20$ and $N = 100$. But then when N increases $f_{max_{KS}}(N)$ tends to 0 whereas $f_{max_{EP}}(N)$ tends to a non-zero value.
In Figure 2 we represent the Entropy Production (top) and KS Entropy (bottom) function of f for $N = 1000$ and for three value of s: $s = 0.13$; $s = 0.2$; $s = 0.04$. This supports the claim that for N fixed, we could tried different values of s such that $s \in [0;\nu]$ $|f_{max_{EP}} - f_{max_{KS}}| \leq \epsilon$. We recover this result in Figure 3.
Such numerical investigations suggest to understand why $f_{max_{KS}}(N)$ and $f_{max_{EP}}(N)$ have different behaviors function of N, and why for N large enough $f_{max_{KS}}$ and $f_{max_{EP}}$ have the same behavior of first order in the deviation from equilibrium measured by the parameter s. We will see that we can get a precise answer to such questions by doing calculations and introducing a sort of Hydrodynamics approximation.

A. Taylor expansion

From Eq. (14) it is apparent that $f_{max_{KS}}$ depends on N whereas from Eq. (9) we get that $f_{max_{EP}}$ hardly depends on N. Indeed there is a difference between $f_{max_{EP}}$ and $f_{max_{KS}}$, i.e. a difference between the two principles for the Zero Range Process. Nevertheless, we have seen numerically that there is a range of N, namely $N \approx 10 \sim 100$ for which the maxima fairly coincide.

Using Eqs. (14) (6) (10) we compute analytically the Taylor expansion of $f_{max_{EP}}$ and $f_{max_{KS}}$ in s. We will show the main result: $f_{max_{EP}}$ and $f_{max_{KS}}$ have the same Taylor expansion in first order in s for N large enough. Their Taylor expansions are different up to the second order in s but it exists an N, *i.e.* a resolution, such that $f_{max_{EP}}$ and $f_{max_{KS}}$ coincident up to the second order.

Let us start by computing $f_{max_{KS}}$. It does not depend of the constant terms of h_{KS} in Eq.(14) and therefore we need only concern ourselves with :

$$-(p\log(p)+q\log(q))(\sum_{i=1}^{N}(z_i)-1)+(\gamma\log(\gamma)+p\log(p))(1-z_1)+(\beta\log(\beta)+q\log(q))(1-z_N)=N.H(f,N,\alpha,\gamma,\beta,\delta). \tag{15}$$

Using Eq.(6), the expression of $H(f,N,\alpha,\gamma,\beta,\delta)$ takes an easy form. To simplify the calculations, we restrict the space of parameter by assuming $\alpha+\gamma=1$ and $\beta+\delta=1$ and we parametrize the deviation from equilibrium by the parameter $\bar{s}=\alpha-\delta$. Moreover let's note $a=\frac{1}{N}$. Thus, we have $H(f,N,\alpha,\gamma,\beta,\delta)=H(f,a,\alpha,\bar{s})$. In order to know the Taylor expansion to the first order in $\bar{s}$ of $f_{max_{KS}}$ we develop $H(f,a,\alpha,\bar{s})$ up to the second order in f; *i.e.* we have $H(f,a,\alpha,\bar{s})=C+Bf+Af^2+o(f^2)$ then we find $f_{max_{KS}}=-B/2A$ that we will develop in power of $\bar{s}$. This is consistent if we assume $f \ll a$.

After some tedious but straightforward calculations, we get at the first order in $\bar{s}$

$$f_{max_{KS}}(\bar{s})=\frac{1}{4}\frac{(1-\alpha)-a(\alpha+2)}{\alpha(1-\alpha)+2a\alpha(\alpha-1)}\bar{s}+o(\bar{s}). \tag{16}$$

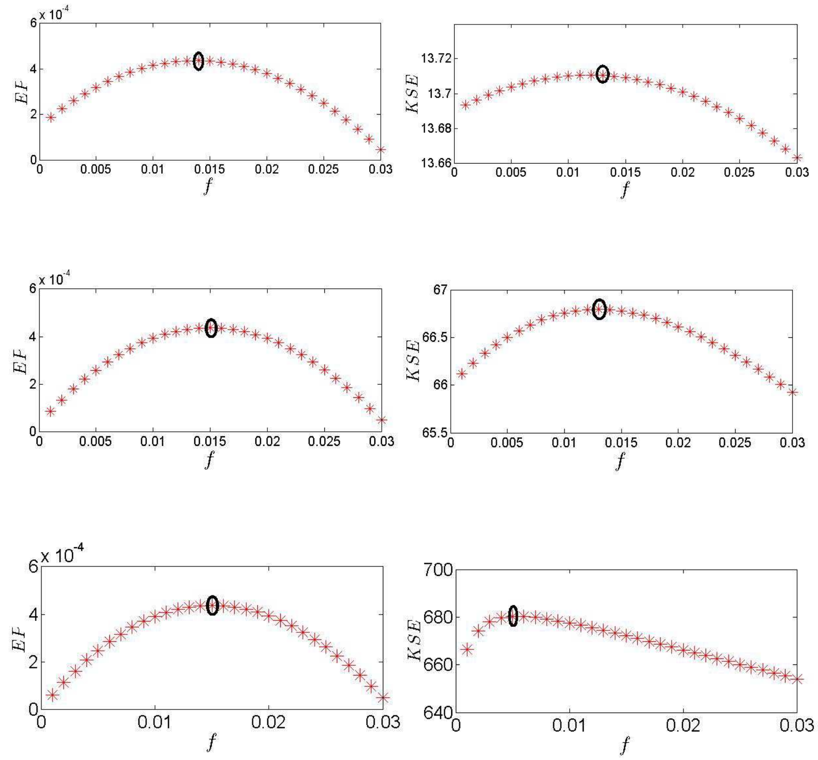

FIG. 1. Entropy Production calculate using 10 (left) and KS Entropy calculate using 6 and 14 (right) function of f for $s=0.13$ and respectively $N=20$ $N=100$ et $N=1000$

and so,

$$f_{max_{KS}}(\bar{s}) = \frac{1}{4\alpha}\bar{s} + \frac{3a}{4(\alpha-1)}\bar{s} + o(\bar{s}) + o(a\bar{s}). \tag{17}$$

We repeat the same procedure starting from Eq.(10) and we obtain:

$$f_{max_{EP}}(\bar{s}) = \frac{\bar{s}}{4\alpha} + o(\bar{s}) + o(a). \tag{18}$$

Thus, since $a = \frac{1}{N} \ll 1$ the behaviour of $f_{max_{KS}}(\bar{s})$ and $f_{max_{EP}}(\bar{s})$ is the same for $\bar{s}$ small enough. We remark that we can strictly find the same result by solving the hydrodynamics continuous approximation given by Eq. (2). This equation is a classical convection-diffusion equation. We remark that, by varying f, we change the

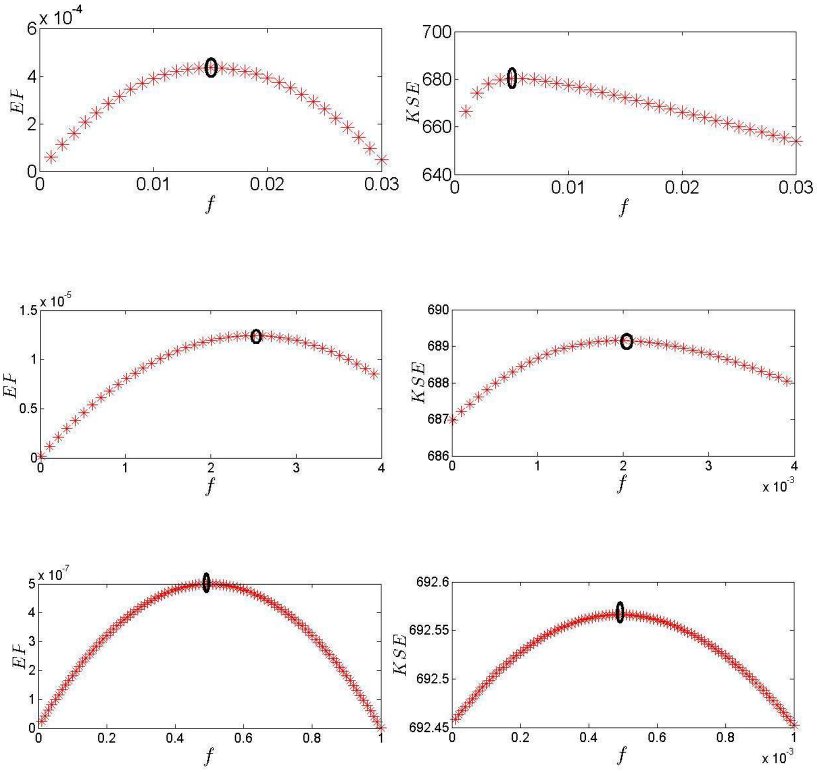

FIG. 2. Entropy Production (left) and KS Entropy (right) function of f for $N = 1000$ and respectively $s = 0.13$; $s = 0.2$; $s = 0.04$

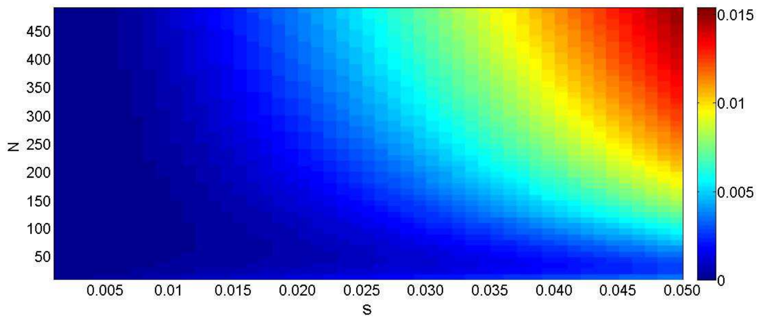

FIG. 3. New figure: 2D plot representing $\Delta f_{max} = f_{max_{EP}} - f_{max_{KS}}$ in the (N, s) space.

convective behavior: $f = 0$ corresponds to a purely diffusive regime whereas by increasing f we enhance the role of convection. If the system is near equilibrium then $f_{max_{EP}} \approx f_{max_{KS}} \approx 0$ and the system is purely diffusive. When the system is out of equilibrium $f_{max_{EP}}$ and $f_{max_{KS}}$ are different from 0 and corresponds to an (optimal) trade-off between purely diffusive and convective behavior.
One can verify this numerically: We first calculate the exact values of the Entropy Production function of f using Eq. (6) and the Kolmogorov-Sinai Entropy function of f using Eqs. (6) (14). Then we approximate these two curves with a cubic spline approximation in order to find $f_{max_{EP}}$ and $f_{max_{KS}}$.

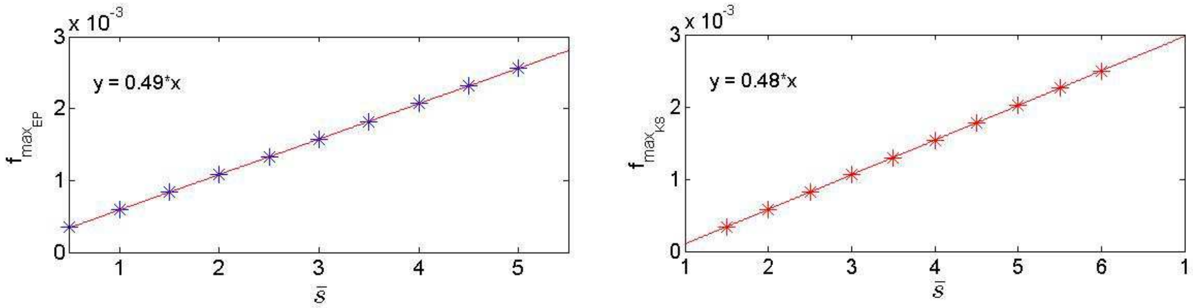

FIG. 4. $f_{max_{EP}}$ (left) and $f_{max_{KS}}$ (right) function of $\bar{s}$ for $\alpha = 0.5$ and $N = 100$. We remark than $f_{max_{KS}}$ and $f_{max_{EP}}$ have both a linear behaviour with slope respectively 0.48 and 0.49 which is really close to $\frac{1}{4\alpha} = 0.5$

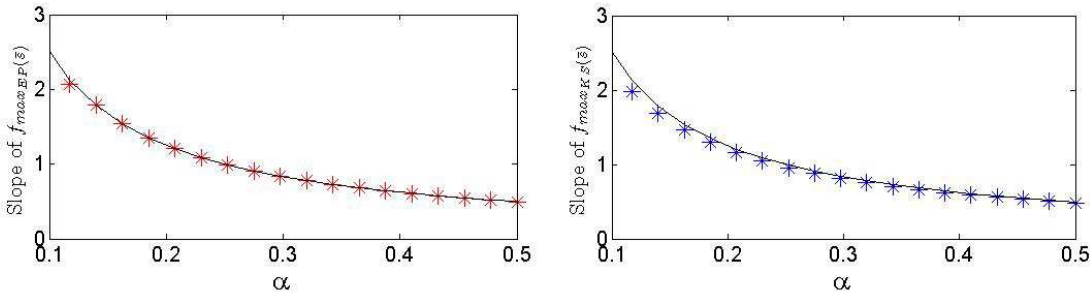

FIG. 5. We plot the slope of $f_{max_{KS}(\bar{s})}$ (left) and $f_{max_{EP}(\bar{s})}$ (right) function of α and in black the curve $f(\bar{s}) = \frac{1}{4\alpha}\bar{s}$. We remark than the approximation $f_{max_{KS}}(\bar{s}) \approx f_{max_{EP}}(\bar{s}) \approx= \frac{1}{4\alpha}\bar{s}$ is good

In order to find the optimal resolution N_* we can go one step further by expanding $f_{max_{EP}}$ and $f_{max_{KS}}$ up to the second order in $\bar{s}$:

$$f_{max_{EP}}(\bar{s}) = \frac{\bar{s}}{4\alpha} + \frac{\bar{s}^2(\alpha+1)}{8\alpha^2(\alpha-1)} + o(\bar{s}^2) + o(a). \tag{19}$$

$$f_{max_{KS}}(\bar{s}) = \frac{1}{4}\frac{(1-\alpha)-a(\alpha+2)}{\alpha(1-\alpha)+2a\alpha(\alpha-1)}\bar{s} + \frac{(1-\alpha)^2+a(\alpha^2-2\alpha+1)}{8\alpha^2(\alpha-1)^2(1-2a)}\bar{s}^2 + o(\bar{s}^2). \tag{20}$$

Thus, $f_{max_{EP}}$ and $f_{max_{KS}}$ coincide in second order in $\bar{s}$ iff a satisfies the quadratic equation:

$$(4\alpha - 6\alpha^2 + 6\alpha^3 - 4\bar{s} + 3\alpha^2\bar{s})a^2 - \frac{1}{2}(8\alpha - 8\bar{s} + 3\alpha^2\bar{s} - 6\alpha^2 + 6\alpha^3)a - (1-\alpha) = 0. \tag{21}$$

This equation has a unique positive solution because theleading coefficient is positive for s small enough $(4\alpha - 6\alpha^2 + 6\alpha^3 - 4\bar{s} + 3\alpha^2\bar{s}) \geq 0$ and the constant term is negative $-(1-\alpha) \leq 0$. We remark that the optimal resolution $N_* = \frac{1}{a*}$ depends on the parameters of the system namely on the degree of non-equilibrium. This fact can be the explanation for two well known issues in climate/weather modeling. First, it explains that, when downgrading or upgrading the resolution of convection models, the relevant parameters must be changed as they depend on the grid size. Second, it suggests that if the resolution is well tuned to represent a particular range of convective phenomena, it might fail in capturing the dynamics out of this range: since finer grids are needed to better represent deep convection phenomena, the deviations between model and observations observed in the distribution of extreme convective precipitation may be due to an inadequacy of the grid used.

V. CONCLUSION

We have shown how a simple 1D Markov Process, the Zero Range Process, can be used to obtain rigorous results on the problem of parametrization of the passive scalar transport problem, relevant to many geophysical applications including temperature distribution in climate modeling. Using this model, we have derived rigorous results on the link between a principle of maximum entropy production and the principle of maximum Kolmogorov- Sinai entropy using a Markov model of the passive scalar diffusion called the Zero Range Process. The Kolmogorov-Sinai entropy seen as function of the convective velocity admit a unique maximum. We show analytically that both have the same Taylor expansion at the first order in the deviation from equilibrium. The behavior of these two maxima is explored as a function of the resolution N (equivalent to the number of boxes, in the box approximation). We found that for a fixed difference of potential between the reservoirs , the maximal convective velocity predicted by the maximum entropy production principle tends towards a non-zero value, while the maximum predicted using Kolmogorov-Sinai entropy tends to 0 when N goes to infinity. For values of N typical of those adopted by climatologists ($N \approx 10 \sim 100$), we show that the two maxima nevertheless coincide even far from equilibrium. Finally, we show that there is an optimal resolution N_* such that the two maxima coincide to second order in $\bar{s}$, a parameter proportional to the non-equilibrium fluxes imposed to the boundaries. The fact that the optimal resolution depends on the intensity of the convective phenomena to be represented, points to new interesting research avenues, e.g. the introduction of convective models with adaptive grids optimized with maximum entropy principles on the basis of the convective phenomena to be represented.

On another hand, the application of this principle to passive scalar transport parametrization is therefore expected to provide both the value of the optimal flux, and of the optimal number of degrees of freedom (resolution) to describe the system. It would be interesting to apply it to more realistic passive scalar transport problem, to see if it yield to model that can be numerically handled (i.e. corresponding to a number of bow that is small enough to be handled by present computers). Moreover, on a theoretical side, it will be interesting to study whether for general dynamical systems, there exists a smart way to coarse grain the Kolmogorov- Sinai entropy such that its properties coincide with the thermodynamic entropy production. This will eventually justify the use of the MEP principle and explain the deviations as well as the different representations of it due to the dependence of the dynamic (Kolmogorov Smirnov, Tsallis, Jaynes) entropies on the kind of partition adopted.

Andjel, E. D.: Invariant measures for the zero range process, The Annals of Probability, pp. 525–547, 1982.

Balian, R.: Physique statistique et themodynamique hors équilibre, Ecole Polytechnique, 1992.

Billingsley, P.: Ergodic theory and information, Wiley, 1965.

Burda, Z., Duda, J., Luck, J. M., and Waclaw, B.: Localization of the maximal entropy random walk, Phys. Rev. Lett., 102, 160 602, 2009.

Dewar, R. C. and Maritan, A.: A theoretical basis for maximum entropy production, in: Beyond the Second Law, pp. 49–71, Springer, 2014.

Domb, C.: Phase transitions and critical phenomena, vol. 19, Academic Press, 2000.

Frisch, U.: Turbulence: the legacy of AN Kolmogorov, Cambridge university press, 1995.

Gómez-Gardeñes, J. and Latora, V.: Entropy rate of diffusion processes on complex networks, Physical Review E, 78, 065 102, 2008.

Großkinsky, S., Schütz, G. M., and Spohn, H.: Condensation in the zero range process: stationary and dynamical properties, Journal of statistical physics, 113, 389–410, 2003.

Herbert, C., Paillard, D., Kageyama, M., and Dubrulle, B.: Present and Last Glacial Maximum climates as states of maximum entropy production, Q. J. R. Meteorol. Soc., 137, 1059–1069, 2011.

Levine, E., Mukamel, D., and Schütz, G.: Zero-range process with open boundaries, Journal of statistical physics, 120, 759–778, 2005.

Martyushev, L. M. and Seleznev, V. D.: Maximum entropy production principle in physics, chemistry and biology, Phys. Rep., 426, 1–45, 2006.

Mihelich, M., Dubrulle, B., Paillard, D., and Herbert, C.: Maximum Entropy Production vs. Kolmogorov-Sinai Entropy in a Constrained ASEP Model, Entropy, 16, 1037–1046, 2014.

Monthus, C.: Non-equilibrium steady states: maximization of the Shannon entropy associated with the distribution of dynamical trajectories in the presence of constraints, J. Stat. Mech., p. P03008, 2011.

Murphy, J. M., Sexton, D. M., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004.

Onsager, L.: Reciprocal relations in irreversible processes. I., Physical Review, 37, 405, 1931.

Paltridge, G. W.: Global dynamics and climate-a system of minimum entropy exchange, Q. J. R. Meteorol. Soc., 101, 475–484, 1975.

Pascale, S., Gregory, J. M., Ambaum, M. H., and Tailleux, R.: A parametric sensitivity study of entropy production and kinetic energy dissipation using the FAMOUS AOGCM, Climate dynamics, 38, 1211–1227, 2012.

Rotstayn, L. D.: On the tuning of autoconversion parameterizations in climate models, Journal of Geophysical Research: Atmospheres (1984–2012), 105, 15 495–15 507, 2000.

Troen, I. and Mahrt, L.: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation, Boundary-Layer Meteorology, 37, 129–148, 1986.

Turkington, B.: An optimization principle for deriving nonequilibrium statistical models of hamiltonian dynamics, Journal of Statistical Physics, 152, 569–597, 2013.

Wallace, J. M. and Hobbs, P. V.: Atmospheric science: an introductory survey, vol. 92, Academic press, 2006.

VI. APPENDIX: COMPUTATION OF THE K-S ENTROPY

In this appendix, we compute the Kolmogorov-Sinai entropy for the Zero Range Process, starting from its definition Eq. (13). In the frame of our Zero Range Process , we use Eqs. (13) and (5) to write it as:

$$\begin{aligned} h_{KS} = -\sum_i \mu_{i_{stat}} \sum_j p_{ij}\log(p_{ij}) &= -\sum_{m_1=0}^{+\infty} ... \sum_{m_N=0}^{+\infty} P(m_1, m_2, ..., m_N) \sum_j p_{(m_1,...,m_N)\to j} \log(p_{(m_1,...,m_N)\to j}) \\ &= -\sum_{m_1=0}^{+\infty} P(m_1) ... \sum_{m_N=0}^{+\infty} P(m_N) \sum_j p_{(m_1,...,m_N)\to j} \log(p_{(m_1,...,m_N)\to j}) \qquad (22) \end{aligned}$$

We thus have to calculate $\sum_j p_{(m_1,...,m_N)\to j}\log(p_{(m_1,...,m_N)\to j})$ that we will refer to as $(*)$. We will take $p+q=\alpha+\delta=\beta+\gamma=1$ and $dt=\frac{1}{N}$ in order to neglect the probabilities to stay in the same state compare to the probabilities of changing state. There are five different cases to consider:

1. if $\forall i\ m_i \geq 1$ so the possible transitions are:
   $(m_1, m_2, ..., m_N) \to (m_1 \pm 1, m_2, ..., m_N)$ with respective probabilities α and δ
   $(m_1, m_2, ..., m_N) \to (m_1, m_2, ..., m_N \pm 1)$ with respective probabilities γ and β
   and $(m_1, ..., m_k, ..., m_N) \to (m_1, ..., m_k \pm 1, ..., m_N)$ with respective probabilities p and q

   Thus,

$$(*) = \alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + \beta\log\beta + (N-1)(p\log(p) + q\log(q)) \qquad (23)$$

2. if $m_1 \geq 1$ and $m_N \geq 1$ and let i be the number of m_i between 2 and $N-1$ equal to 0. With the same argument as previously we have:

$$(*) = \alpha\log\alpha + \delta\log\delta + \gamma\log\gamma + \beta\log\beta + (N-1-i)(p\log(p) + q\log(q)) \qquad (24)$$

3. if $m_1 = 0$ and $m_N \geq 1$ and let i the number of m_i between 2 and $N-1$ equal to 0 we have:

$$(*) = \alpha\log\alpha + \delta\log\delta + \beta\log\beta + (N-2-i)p\log(p) + (N-1-i)q\log(q) \qquad (25)$$

4. The same applies if $m_1 \geq 1$ and $m_N = 0$ and let i the number of m_i between 2 and $N-1$ equal to 0 we have:

$$(*) = \alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + (N-1-i) p \log(p) + (N-2-i) q \log(q) \tag{26}$$

5. finally, if $m_1 = 0$ and $m_N = 0$ and let i the number of m_i between 2 and $N-1$ equal to 0 we have:

$$(*) = \alpha \log \alpha + \delta \log \delta + (N-2-i)(p \log(p) + q \log(q) \tag{27}$$

Using equation 3 we find that $P(m_k = 0) = 1 - z_k$ and $\sum_{i=1}^{+\infty} P(m_k = i) = z_k$

We thus obtain than h_{KS} writes:

$$\begin{aligned} h_{KS} = &-(\alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + \beta \log \beta + (N-1)(p \log(p) + q \log(q)) \\ &+ (p \log(p) + q \log(q))(\sum_{r=0}^{N} r \sum_{i_1 \dots i_N} \prod_{i=i_1, \dots i_r} (1 - z_i) \prod_{i \neq i_1 \dots i_r} z_i) \\ &+ (\gamma \log(\gamma) + p \log(p)) z_N (1 - z_1)(\sum_{i_2 \dots i_{N-1}} \prod_{i=i_2, \dots i_r} (1 - z_i) \prod_{i \neq i_2 \dots i_r} z_i) \\ &+ (\beta \log(\beta) + q \log q) z_1 (1 - z_N)(\sum_{i_2 \dots i_{N-1}} \prod_{i=i_2, \dots i_r} (1 - z_i) \prod_{i \neq i_2 \dots i_r} z_i) \\ &+ (\beta \log(\beta) + \gamma \log \gamma + p \log p + q \log q)(\sum_{i_2 \dots i_{N-1}} \prod_{i=i_2, \dots i_r} (1 - z_i) \prod_{i \neq i_2 \dots i_r} z_i) \end{aligned} \tag{28}$$

This expression, though complicated at first sight, can be simplified. Indeed interested in the function $F(a) = \prod_1^N (z_k + a(1 - z_k))$ and by deriving subject to a we show that:

$$\sum_{r=0}^{N} r \sum_{i_1 \dots i_N} \prod_{i=i_1, \dots i_r} (1 - z_i) \prod_{i \neq i_1 \dots i_r} z_i = \sum_{i=1}^{N} (1 - z_i) \tag{29}$$

Thus we can simplify the last equation and we obtain:

$$\begin{aligned} h_{KS} = &-(\alpha \log \alpha + \delta \log \delta + \gamma \log \gamma + \beta \log \beta + (N-1)(p \log(p) + q \log(q))) + (p \log(p) + q \log(q)) \sum_{i=1}^{N} (1 - z_i) \\ &+ (\gamma \log(\gamma) + p \log(p))(1 - z_1) + (\beta \log(\beta) + q \log(q))(1 - z_N) \end{aligned} \tag{30}$$

Response to the referee 1:

Interactive comment on "Statistical optimization for passive scalar transport: maximum entropy production vs. maximum Kolmogorov–Sinay entropy" by M. Mihelich et al.

Anonymous Referee #1

Received and published: 29 November 2014

The Maximum Entropy Production (MEP) conjecture is a much debated scientific issue and has attracted lots of interest and criticism over the last thirty years. The main problem with it is that, in spite of some empirical evidence built up mostly in climate science, a rigorous, general demonstration does not exist yet. This piece of work by Mihelich et al. adds a useful contribution to this debate as it shows that, for a simple statistical model of diffusion (a Markov model of the passive scalar diffusion) between two reservoirs, the entropy production is linked to the Kolmogorov-Sinai entropy, a well know quantity in information theory. Results by Mihelich et al. are not general as they hold only for this specific model – and for another similar case (Mihelich et al. (2013))– but may open new avenues of research.

The paper is, overall, interesting and gives a useful contribution to the scientific discussion on the Maximum Entropy Production conjecture. However some more work is required to have the manuscript in its final, publishable form. Therefore no recommendation for publication can be made until the comments and suggestions, listed below, are addressed.

Major points and general remarks

1) Results could be displayed in a more convincing and complete way. The authors should make a more systematic exploration of resolutions and far-fromequilibrium setups. I suggest the author to plot the difference (or percentual difference) between f_{MPE} and f_{MKS} as a function of N and s, that is a 2D contour plot, with s going from 0 to a value typical of far-from-equilibrium conditions and N from $O(1)$ to , e.g., $O(1000)$. This would summarize very effectively the main findings of this study and show clear patterns in the $(N; s)$ space in a wide range of N and s;

This remark is entirely justified. Thus I add a new 2D contour plot representing the difference between f_{max_{ep}} and f_{max_{ks}} in the (N,s) space.

2) page 1695. Here the definition of _S is not correct and the notation used for the fluxes between contiguous boxes confusing. Paltridge used the divergence of the meridional heat flux in a certain latitudinal box divided by the box temperature, not the flux itself divided by the temperature, i.e.

R

(r _ F)=T , not
R
R F=T. Then
(r _ F)=T =
R
F _ r(1=T) because F = 0 at the boundaries (poles). Moreover, the notation f_{ij} is confusing because it looks like there can be a heat exchange between any i and j, so also noncontiguous boxes, which is not the case.

Indeed, I have corrected the error by writing: $\dot S=\sum_i f_{i(i+1)}(\frac{1}{T_{i+1}}-\frac{1}{T_{i}})$

3) English. There are several typos and minor English mistakes. I'll list a few ones in the following, but this is not an exhaustive list. Therefore the manuscript should be carefully edited to correct minor grammatical errors;

4) References. The scientific literature cited in this study is very, very limited indeed.

In some cases, references cited by the authors are old and more updated studies

could be instead cited. For example, when introducing the "macroscopic" entropy production _ the authors cite Balian (1992) at the beginning of Section 3.1. Now, even having all the respect for Balian, it is odd that they do not mention previous authors such as Onsager (1931), or De Groot and Mazur (1962), or Glansdorff and Prigogine (1971).
At page 1693 they cite Yang et al. (2012) which deals with a convective scheme, but there is nothing about parameter tuning in a General Circulation Model (e.g. Murphy et al. (2004)). At the same page and line 27 the cite Dewar (2003) – which is an outdated study about demonstrating MEP – but they omit more recent studies such as Dewar and Maritan (2014) and references therein.
Also, the authors mention an alternative method for parameter tuning (page 1693, line 17) in the case of complex models based on maximizing (minimizing) a suitable functional (e.g. entropy production), but totally ignore previous studies (Kunz et al. (2008); Pascale et al. (2012)) in which such an idea has been tested for GCMs of various complexity. Concerning efforts made to extend MEP generality, the authors might want also consider the work by Gjermundsen et al. (2014).
In the revised version the authors should therefore pay more attention to this aspect, which is important to put their work in the right wider scientific context.

I naturally added the reference of Onsager (1931) when I introduce the entropy production. I also add the recent work of Dewar and Maritan and the work of Pascale.

5) Diffusion. At page 1703, line 6, the authors say that "If the system is at equilibrium

then $f_{maxEP} = f_{maxKS} = 0$ and the system is purely diffusive". This is not true, (molecular) diffusion is also an irreversible process which leads to entropy production

R

(_jrTj2)=T 2dV _ 0. This also points out to me that such an entropy production in not taken into account when the simple ZRP is considered. Perhaps the authors concentrate on f because this, in a real atmosphere, is associated with the nonlinear quasi-turbulent atmospheric flow (midlatitude baroclinic eddies), but this has to be clarified in the revised manuscript.

In this article I did not say that if the system is diffusive there is no entropy production. I show that for the ZRP, if the system is close to equilibrium, the state chosen by MEP corresponds to a diffusive state. Nevertheless, in order to make this clearer I change " if the system is at equilibrium" by "if the system is close to equilibrium"

Minor points and suggestions

1) Some typos and minor mistakes: (p. 1692, l 9) deviation of/from equilibrium; (p. 1693, l 3) to/too large; (p. 1693, l 17) the/an alternate/alternative road..; (p. 1694, l 3) the citation should be within brackets; (p. 1694, l 11) distance to/from equilibrium; (p. 1695, l 17) do not start a new statement with a mathematical symbol (furthermore in lower case); (p. 1696, l 25) particule ???; (p. 1696, l 25) stationnary/stationary; (p. 1697, l 2) explain/explained; (p. 1698, l 9) The P/physical interpretation; (p. 1698, l 17) reach ie/ reached; (p. 1699, l 11) picked up???; (p. 1700, l 23) for/For N fixed; (p. 1701, l 8) fixe/fixed; (p. 1702, l 1) by compute/computing; (p. 1702, l 1) It is not depends of ?????; (p. 1702, l 8) let's/let us; (p. 1703, l 1) We remark than/that; (p. 1703, l 7) different than/from 0; (p. 1703, l 17) equation of second degrees???/second order equation; (p 1704, l 4) it might fails/fail; (p 1704, l 15) seen as functions/function; (p 1704, l 23) typical of that/those adopted; (p 1705, l 1) research patterns/research avenues.

I thank a lot the referee for all these corrections that I have all changed.

2) Often in the text, new variable or mathematical symbols are suddenly introduced without a previous definition. This is quite annoying and very confusing. For example, immediately in the abstract the symbol f is thrown (line 6). But how can the reader know what f stands for and thus understand that sentence? At page 1696 line 1 the fugacity z is mentioned without being previously defined; at the same page, line 19 the "chemical potential"; at page 1697 an Hamiltonian is mentioned (line 15), and this is completely out of the blue; at page 1698 the flux of mass c unexpectedly appears in an involved relationship (eq. 7). I really suggest the authors to introduce/define these quantities when they first discuss the model.

In fact, the symbol f line 6 was badly introduced. Thus, I add that " f is parameter connected to the jump probability".
Concerning the fugacity, I add that this is a quantity related to the average particle density.
I also precise that the Hamiltonian equation is found from the quantum formalism.

3) page 1695, line 4: shouldn't it be f = □uT + _rT?

I have corrected the error

4) page 1698, line 16: Why is the thermodynamic force X equal to rlog _ and not

r_?

For the thermodynamics force I took the definition of Balian where X is proportional to the gradient of $\log\rho$.

5) page 1692, line 1-3: The way it's written, this sentence seems to mean that, through a Markov model, the authors demonstrate the link between MEP and MKS in general, which is not the case. I would therefore say: "We derive rigorous results on the link between the principle of maximum entropy production and the principle of maximum Kolmogorov-Sinai entropy for a Markov model of the passive scalar diffusion called the Zero Range Process"

I change the sentence as suggested by the referee.

6) page 1692, line 13: Climatologist also use GCMs, actually nowadays climatologists hardly use box-models as those in Paltridge (1975), except people studying MEP. Same applies for page 1704, line 23. Please make this sentence more precise.

In order to make this sentence more precise I add "climatologist working on MEP".

7) page 1703, eq. 21: Actually I can't see any "=" in the equation;

I corrected the error.

8) page 1703, line 20: what's a "dominant" coefficient?

I have changed dominant by "leading "coefficient.

9) Eq. (1): given that, spatially, u is a function only of x, wouldn't it be more precise to write @$_x$ and @^2_x in place of r and r^2?

This remark is well justified and I corrected the equation.

10) page 1695, line 1-3: Said like that, it seems that such an untold equation is something mysterious and esoteric; but this is just the conservation of momentum (NS equation) and, for baroclinic fluids, also energy and mass conservation;

11) page 1695, line 25; page 1698, line 14: right to left, or left to right?

I corrected the mistake page 1698

I really want to thank the referee for all these constructive comments.

Response to the referee 2

Interactive comment on "Statistical optimization for passive scalar transport: maximum entropy production vs. maximum Kolmogorov–Sinay entropy" by M. Mihelich et al.

Anonymous Referee #2

Received and published: 28 December 2014

This paper is not easy to understand. There is a mixture of turbulence (passive scalar transport), of maximum entropy production, of Kolmogorov Sinai entropy and of zero range process. I suggest to reject this paper since the content is too narrow and far from geosciences, thus not adapted to NPG.

Major points:

1) The title is not adapted to the content: the title mentions passive scalar transport in turbulence, but in fact the manuscript is dealing only with a 1D toy model of passive scalar, called ASEP (asymmetric exclusion Markov process).

The title does not mention turbulence. Moreover this article is not about the ASEP model, which is cited only twice (p1693 l29 and p1696 l17) in reference to previous works. This article addresses problem of passive scalar transport for which the Zero Range Process (ZRP) is a simple but very insightful model. The expression "Statistical Optimization" naturally implies the discussion of MEP and other optimization principles like the Maximum Kolmogorov Sinai Entropy Principle (MKSEP). Therefore we believe that the title is appropriate to the content of the paper since i) MEP and MKSEP are the physical principles discussed. ii) Passive scalar transport via the ZRP model is the object of the study.
Since both the passive scalar and the MEP are widely discussed and published in the geoscience literature, we believe our paper is appropriate to NPG. Moreover, as recognized below by the referee, we establish a clear link between our results and geosciences in the discussion section.

With such restriction, the topic of the present paper seems rather far from geosciences. The link with ASEP and numerical models used in the geosciences is not obvious, and only justified in the perspectives and conclusion of this manuscript.

As previously said, the paper is not at all about ASEP model. It is about the ZRP, a toy model which both exactly mimics the general approach to MEP in geophysics and enables exact analytical calculation that allow to explore the validity of such principle in a simple case. The link between the ZRP and, e.g., Paltridge's work is explained not only in the introduction and in the perspectives but also in the main body of the paper, where the meaning of the terms introduced in the ZRP are linked with general thermodynamic quantities used in geophysics. See specifically Page 1695, lines 4-8.

While the mathematical content of the paper seems correct [compute analytically the heat flux f for maximum entropy production and for Kolmogorov Sinai entropy, equations (19) and (20), consider for which cases the maximum coincides in both analytical expressions], its scope seems very narrow [to show that a toy model has two ways to estimate the heat flux corresponding to a maximum entropy situation] to be useful for geosciences applications.

The passive scalar transport Eq. 1 modelled via the ZRP is one of the fundamental equation of any geophysical models as transport in ocean, soils, atmospheres is understood in terms of Eq. 1. Therefore, saying that the results obtained both analytically and numerically here do not have implications on geosciences application is neglecting the role of Eq. 1 in the dynamics of geophysical systems. The result of the paper is to provide a theoretical explication for MEP (which has been successfully used in several applications in geophysics), via a simple exact model where all the calculations can be performed analytically. This provides firmer ground to the use of this maximization tools to more complicated systems, and open new perspectives as to how to perform it in the more efficient way. We do not understand why this should not be relevant to geosciences.

2) The paper is not self-contained and it is very difficult to understand the point without reading other papers. The model ASEP cannot be understood by reading this manuscript.

As we have already pointed out, this paper is not about the ASEP model but the Zero Range Process. This model itself is very simple, and is described in the first section and via the Figure 1. The paper is self-contained, in the sense that the calculations presented do not require any further knowledge or material. So there is no need to read other papers to understand the calculations, nor the model.

Equation (2) uses z, the fugacity, which is not precisely defined. One is lost at this point. The maximum entropy production concept is used in the title and in many places in the manuscript, but its meaning is not recalled.
Other points: The review paper Martyushev and Selesnev (Maximum entropy production principle in physics, chemistry and biology, Physics Report 426 (2006) 1-45) should be cited since it nicely and clearly introduces the MEP.

We agree with the referee and we provide further explications on the fugacity: ”which is a quantity related to the average particle density (see \ref{eq:3} below)”

the concept of MEP and further references including Martyushev and Selesnev’s work: “MEP states that a stationary nonequilibrium system chooses its final state in order to maximize the entropy production as is explain in \citep{martyushev2006maximum}”

Typos: line 4 page 1695 -> passive scalar; line 11 same page: decor related jumps ->

We fixed the typos